scientific reports

OPEN

Wood cookstove use is associated with gastric cancer in Central America and mediated by host genetics

Samara B. Rifkin[1,12], Anna K. Miller[2,12], Eleazar E. Montalvan-Sanchez[3,4], Dalton A. Norwood[3,5], Enrique Martinez[6], Tim Waterboer[7], T. Mark Beasley[8,9], Ricardo L. Dominguez[3], Scott M. Williams[2,10,13] & Douglas R. Morgan[11,13]✉

Biomass cookstove food preparation is linked to aero-digestive cancers, mediated by ingested and inhaled carcinogens (e.g., heterocyclic amines, and polycyclic aromatic hydrocarbons). We investigated the association between gastric adenocarcinoma, wood cookstove use, *H. pylori* CagA infection and risk modification by variants in genes that metabolize and affect the internal dose of carcinogens. We conducted a population-based, case–control study (814 incident cases, 1049 controls) in rural Honduras, a high-incidence region with a homogeneous diet and endemic *H. pylori* infection, primarily with the high-risk CagA genotype. We investigated factors including wood cookstove use, *H. pylori* CagA serostatus, and 15 variants from 7 metabolizing genes, and the interactions between wood stove use and the genetic variants. Male sex (OR 2.0, 1.6–2.6), age (OR 1.04, 1.03–1.05), wood cookstove use (OR 2.3, 1.6–3.3), and CagA serostatus (OR 3.5, 2.4–5.1) and two SNPs in *CYP1B1* (rs1800440 and rs1056836) were independently associated with gastric cancer in multivariate analysis. In the final multivariate model, a highly significant interaction (OR 3.1, 1.2–7.8) was noted between wood cookstove use and the rs1800440 metabolizing genotype, highlighting an important gene-environment interaction. Lifetime wood cookstove use associates with gastric cancer risk in the high-incidence regions of Central America, and the association is dependent on the rs1800440 genotype in *CYP1B1*. *H. pylori* CagA infection, wood cookstove use and the rs1800440 genotype, all of which are highly prevalent, informs who is at greatest risk from biomass cookstove use.

Gastric adenocarcinoma (GC) is the leading cause of infection-related cancer mortality globally, and overall is the third leading cause of cancer death[1-5]. One million incident cases occur annually, and high-incidence regions include Latin America, eastern Asia, and eastern Europe. In Latin America, a large burden of disease is concentrated in the mountainous regions along the Pacific littoral[6]. This pattern has been described as the gastric cancer "altitude enigma", and suggests a clustering of risk exposures in the rural mountain villages[6,7]. *Helicobacter pylori* (*H. pylori*) infection is the primary risk factor for gastric cancer, acting in concert with host genetics and responses, dietary factors, and environmental exposures[8]. *H. pylori* genetic variation affects risk, and CagA is the principal *H. pylori* virulence genotype for gastric adenocarcinoma[8].

[1]Department of Gastroenterology and Hepatology, University of Michigan, Ann Arbor, MI, USA. [2]Department of Genetics and Genome Sciences, Case Western Reserve University, Cleveland, OH, USA. [3]Hospital de Occidente, Ministry of Health, Santa Rosa de Copan, Copan, Honduras. [4]Department of Medicine, Indiana University School of Medicine, Indianapolis, IN, USA. [5]Division of Preventive Medicine, Department of Medicine, The University of Alabama at Birmingham, Birmingham, AL, USA. [6]Hospital Evangelico, Siguatepeque, Honduras. [7]Infections and Cancer Epidemiology, German Cancer Research Center (DKFZ), Heidelberg, Germany. [8]Department of Biostatistics, School of Public Health, The University of Alabama at Birmingham, Birmingham, AL, USA. [9]Birmingham/Atlanta Geriatric Research, Education, and Clinical Center, Birmingham Veterans Affairs Medical Center, Birmingham, AL, USA. [10]Department of Population and Quantitative Health Sciences and Cleveland Institute for Computational Biology, Case Western Reserve University, Cleveland, OH, USA. [11]Medicine and Epidemiology, UAB Division of Gastroenterology and Hepatology, The University of Alabama at Birmingham (UAB), Birmingham, AL 35294, USA. [12]These authors contributed equally: Samara B. Rifkin and Anna K. Miller. [13]These authors jointly supervised this work: Scott M. Williams and Douglas R. Morgan. ✉email: drmorgan@uabmc.edu

Cooking and heating with biomass fuel is an environmental risk for a variety of cancers. One-third of the world's population burns organic matter (wood, charcoal, dung) for domestic energy, leading to multiple health and environmental consequences[9]. The mechanisms by which biomass cooking and heating increases cancer risk is hypothesized to be exposures to organic carcinogens, such as heterocyclic amines (HCAs), and inhaled airborne pollutants, including polycyclic aromatic hydrocarbons (PAHs)[10]. The effects of this exposure is exacerbated by poor ventilation in the mountainous regions[11]. Metal carcinogens from the cooking surface may also contribute to risk[12,13]. Red meat cooked on biomass stoves contribute to cancer risk by the ingestion of carcinogens, including HCAs and PAHs, created during cooking at high temperatures[14]. HCAs are formed when amino acids, sugars, and creatine in meat react at high cooking temperatures, while PAHs are formed when fat vaporizes and then adheres to the surface of the meat[15]. High temperature cooking of other foods may also introduce risk[16]. The International Agency for Research on Cancer (IARC) has therefore classified red and processed meat as a Class I carcinogen, linked to the increased relative risk of colon cancer, gastric cancer, and other cancers[14].

The relationship between exposure to HCAs and PAHs and gastric cancer may be modified by genetic variation in genes encoding enzymes that activate or deactivate HCAs and PAHs[17–22]. These enzymes include phase I enzymes, such as CYP1A2 and CYPB1, and phase II enzymes, including N-acetyltransferases (NAT) and glutathione S transferases (GSTs). Phase I enzymes tend to activate HCAs and PAHs into carcinogenic metabolites that create DNA adducts and lead to mutations, whereas some phase II enzymes can detoxify carcinogenic compounds and prevent further damage[17]. Finally, the Aryl hydrocarbon receptor (AHR) can mediate expression of both classes of enzymes. These activating and detoxifying enzymes in combination may determine the internal exposure ("internal dose") to environmental carcinogens, and hence modify cancer risk due to environmental exposures[17]. Polymorphisms of enzymes encoding genes that increase the activity or level of phase I enzymes and decrease the activity or level of phase II enzymes can lead to a higher risk of cancer[17].

Wood cookstove ("fogon") use is common in the mountainous regions of Pacific Latin America, where gastric cancer incidence is amongst the highest in the world, and cookstove use may contribute to the excess burden of gastric cancer. In the Central America Four (CA-4, Honduras, Guatemala, El Salvador, Nicaragua) region, wood stove use is the predominant method of cooking and *H. pylori* infection is endemic (70–90% of the adult population). The CA-4 countries comprise the core low/middle income country (LMIC) region in the western hemisphere, with a population of over 40 million, and is linked to an immigrant population in the U.S. of 6 million[6,23].

We hypothesized that wood cookstove use increases the risk of gastric adenocarcinoma and that the risk is modified by HCA- and PAH- metabolizing enzyme genotypes. The effects of wood stove use exposures may be most discernable in populations with homogeneous diets and external HCA-exposures, and endemic high-risk *H. pylori* infection (e.g., CagA genotype), as in the rural populations of Central America[24,25].

Materials and methods

Study design

We conducted a population-based, case–control study based in the mountainous regions of Honduras. This region has among the highest incidence rates in the western hemisphere, with a high prevalence of chronic *H. pylori* infection (80–90%)[25–29]. Incident GC cases were enrolled prospectively from the two district hospitals (Santa Rosa de Copán and Siguatepeque) that serve the mountainous rural areas of west-central Honduras. The diagnosis of GC was based on endoscopic appearance and confirmatory histopathology. Household interviews were conducted for randomly selected healthy controls from a wide distribution of villages in the region and the catchment area, as described previously[25,28]. Once consented, patients and control subjects would undergo the study interview with review of exclusionary health criteria, as well as the demographic, health assessment, and epidemiology questionnaires. Lifetime woodstove use was dichotomized (yes, no) since we observed minimal lifetime variation in usage nor transition to other types of cook stoves in the Honduras rural populations.

H. pylori infection and CagA assessment

A validated multiplex serology was used to determine *H. pylori* and CagA serostatus, the dominant bacterial risk genotype for GC[30–32]. We focused on CagA, the principal oncoprotein, as the region has an extremely high, *H. pylori* prevalence. The multiplex serology panel was developed by the Germany Cancer Research Center ("DKFZ"). In brief, the *H. pylori* proteins were recombinantly expressed as Glutathione-S-transferase (GST)-tag fusion proteins in *Escherichia coli* BL21 and affinity-purified on glutathione-coated fluorescently labeled polystyrene beads (Luminex Corp.). A mixture of the differently labeled and antigen-loaded beads was incubated with serum to allow binding of serum antibodies to the *H. pylori* proteins. Bound serum antibodies were detected by a biotin-labeled anti-human IgM/IgA/IgG secondary antibody and Streptavidin-R-phycoerythrin. The Luminex 200 analyzer (Luminex Corp.) distinguished between the bead type and the bound antigen and quantified the amount of bound serum antibody as median fluorescence intensity (MFI) of 100 beads per type measured. The CagA antigen-specific cutoff was used[30].

Candidate genes

We selected 15 variants from 7 HCA metabolizing genes previously associated with colorectal polyp risk[17] (Table S1). The genes and single nucleotide polymorphisms (SNPs) of interest include: Epoxide Hydrolase 1 (*EPHX1*) (rs1051740), Cytochrome P450 Family 1 Subfamily B Member 1 (*CYP1B1*) (rs1800440 and rs1056836), Aryl hydrocarbon receptors (*AHR*) (rs2066853), N-Acetyltransferase 1 (*NAT1*) (rs1799931, rs15561, rs1208, 1799930, rs1041983, rs1799929, rs1801279, and rs1801280), Cytochrome P450 Family 2 Subfamily E Member 1 (*CYP2E1*) (rs2031920), UDP Glucuronosyltransferase Family 1 Member A7 (*UGT1A7*) (rs61261057) and Cytochrome P450 Family 1 Subfamily A Member 2 (*CYPIA2*) (rs762551). Primers for each of these SNPs are presented in Table S2.

Genotype analysis

Human DNA was isolated from whole blood samples with the Qiagen Puregene® kit and genotyped on the MassARRAY® Sequenom platform at Vanderbilt University Medical Center (VANTAGE Core facility). We tested for Hardy–Weinberg equilibrium and removed SNPs that deviated in cases and controls at $p < 1.0 \times 10^{-6}$. We also removed SNPs that had a minor allele frequency (MAF) < 0.05 using PLINK (version 1.9)[33–35]. We tested for Linkage Disequilibrium in PLINK and Haploview (version 4.2) regardless of case–control designation to determine the number of independent tests for the False Discovery Rate (FDR) threshold and limit multiple testing[36].

Statistical analysis

We compared differences between case and control subjects using t-tests for the continuous variable age or the chi-square test for the categorical variables sex, wood stove use, and bacterial CagA serostatus. Age was tested for normality using the Shapiro–Wilk's method with the R "stats" package (R version 4.0.4). Age was negatively skewed as GC is the final step of progressive gastric disease, and the population was predominately older[37]. Therefore, age was log transformed, and the analyses were run with the transformed variable, but the deviation from a normal distribution did not impact the results substantively. The untransformed age variable is presented in all the results.

Univariate logistic regression models were used to estimate risk of GC associated with individual variables, including all SNPs. Logistic regression was used to estimate odds ratios (OR) for adjusted models that included combinations of the variables age, sex, CagA serostatus, wood stove use, and each SNP individually. Daily tobacco use was very low (20 pack-year history, 4.8%), and was not included in the model. For each SNP, we also assessed multivariable models for individuals without missing data that included an interaction term with wood stove use. A likelihood ratio test using R package "lmtest" was used to assess the goodness of fit of competing statistical models, i.e., those without the interaction and those with the interaction (R version 4.0.4). The R command "p.adjust" was used to calculate the q-values using the Benjamini–Hochberg test with a False Discovery Rate (FDR) level of 0.1 for all tests (R version 4.0.4). We confirm that all research was performed in accordance with relevant guidelines/regulations and informed consent was obtained from all participants and/or their legal guardians. Research involving was performed in accordance with the Declaration of Helsinki.

Ethics committee approvals

The study was approved by the institutional review boards of The University of Alabama at Birmingham, Vanderbilt University and the Ministry of Health in western Honduras.

Consent to participate

Informed consent was obtained from all individual participants included in the study.

Results

A total of 814 gastric cancer patients and 1049 population controls were enrolled, of whom genetic and *H. pylori* data were available for 1,425 participants (Table 1). The median ages of the cases and controls were 65 and 54, respectively. Approximately one-fifth of cases were under the age of 55 (n = 170 of 814). Overall, 87% and 85% of subjects were positive for *H. pylori* infection and CagA, respectively. Wood stove use was observed in 92% of the gastric cancer cases and in 78% of the population-based controls. Gastric cancer cases were more likely

Characteristics	Control population[a] n (%)	Gastric cancer cases[a] n (%)	P-value
Age			
Mean (SD)	54.3	64.1	< 0.0001
Sex			
Female	550 (49.7)	260 (31.9)	< 0.0001
Male	557 (50.3)	554 (68.1)	
Wood stove use			
Yes	784 (77.9)	713 (91.9)	< 0.0001
No	222 (22.1)	61 (7.9)	
H. pylori serostatus			
Yes	677 (87.4)	568 (87.5)	0.93
No	98 (12.6)	81(12.5)	
CagA serostatus			
Yes	605 (78.1)	600 (92.5)	< 0.0001
No	170 (21.9)	48 (7.5)	

Table 1. The characteristics of the gastric cancer cases and population controls. [a]There were 814 gastric cancer patients and 1,049 population controls enrolled; Complete genotyping and *H. pylori* serology data were available for 1,425 participants.

to be older, male, have lifetime wood stove use, and be positive for CagA serostatus. In univariate analyses, age ($\beta = 1.04$, $p < 2.00 \times 10^{-16}$), sex (OR = 2.51, $p = 2.53 \times 10^{-16}$), wood stove use (OR = 2.57, $p = 5.52 \times 10^{-8}$), and bacterial CagA serostatus (OR = 3.51, $p = 4.20 \times 10^{-13}$) all associated with an increased risk of gastric cancer (Table 2). In a multivariate model including age, sex, wood stove use and bacterial CagA serostatus, these variables all remained significantly associated with gastric cancer.

Of the 15 SNPs genotyped, none were out of Hardy–Weinberg equilibrium ($p < 1.0 \times 10^{-6}$ for both cases and controls), but 2 SNPs had a minor allele frequency of less than 0.05 and were excluded from further analyses (Table S1). Allele frequencies of the SNPs were within the expected ranges, based on the Latin American populations in the 1000 Genomes databases (Version 3). Linkage Disequilibrium (L.D.) was tested for the 6 SNPs in the NAT1 gene that met inclusion criteria. Three pairs showed evidence of significant correlation ($r^2 > 0.4$), and we randomly removed one SNP from each pair to reduce the multiple testing burden (Fig. S1). The data cleaning resulted in 10 SNPs being tested in the final analysis. Genotype counts are presented in Table S3.

Initial unadjusted association analyses identified two SNPs that associated with GC under a dominant model where the minor allele was used as the referent. The two SNPs, rs1800440 and rs1056836, in cytochrome P450 family 1 subfamily B member 1 (*CYP1B1*), were both significantly associated with GC after FDR adjustment for multiple testing (adjusted q = 0.07 and q = 0.003, respectively) and were not in L.D. ($R^2 = 0.03$) (Table 3). In models adjusted for age, sex, bacterial CagA serostatus, wood stove use, and each SNP singly, both the *CYP1B1* SNP rs1800440 (adjusted q = 0.023) and the *CYP1B1* SNP rs1056836 (adjusted q = 0.023) remained significant (Table 3). Additive and recessive genetic models were also run but no SNPs were significant after FDR correction in the univariate analyses (Table S4).

We used three statistical models of increasing complexity to assess patterns of association (Table S5). The first multivariable model included the core variables age, sex, wood stove use, bacterial CagA serostatus, and the two *CYP1B1* SNPs. All variables were significant (Table 4). Multivariable model 2 included the core variables, rs1800440, and the interaction of wood stove use and rs1800440. All variables remained significant except for wood stove use and rs1800440 (Table 4). Multivariable model 3 included the same variables as multivariable model 2 and also included SNP rs1056836 with all variables remaining significant except for wood stove use and

Characteristics	Univariate OR (95% CI)	Multivariate OR (95% CI)[a]
Age	1.04 (1.03, 1.05)	1.04 (1.03, 1.05)
Sex		
Male	2.51 (2.02, 3.13)	1.99 (1.56, 2.55)
Female	Referent	Referent
Wood stove use		
Yes	2.57 (1.84, 3.64)	2.33 (1.62, 3.39)
No	Referent	Referent
Bacterial CagA serostatus		
Yes	3.51 (2.52, 4.98)	3.49 (2.41, 5.13)
No	Referent	Referent

Table 2. Univariate and multivariate odds ratios for gastric cancer outcomes. *95% CI* 95% confidence interval. [a]Multivariate model includes age, sex, wood stove use, and CagA positive serostatus.

Chr	SNP	Gene	Major allele	Minor allele	MAF	Univariate OR (95% CI)	Univariate q-value	Multivariate* OR (95% CI)	Multivariate q-value*
1	rs1051740	EPHX1	T	C	0.385	0.99 (080, 1.23)	0.937		
2	rs1800440	CYP1B1	A	G	0.09	1.42 (1.08, 1.89)	**0.070**	1.43 (1.04, 1.97)	**0.023**
2	rs1056836	CYP1B1	C	G	0.243	1.48 (1.20, 1.84)	**0.003**	1.33 (1.04, 1.69)	**0.023**
7	rs2066853	AHR	G	A	0.152	0.90 (0.71, 1.14)	0.645		
8	rs15561	NAT1	A	C	0.433	1.10 (0.81, 1.49)	0.803		
8	rs1801280	NAT1	T	C	0.28	0.96 (0.77, 1.19)	0.864		
8	rs1799930	NAT1	G	A	0.19	0.99 (0.79, 1.23)	0.937		
8	rs1799931	NAT1	G	A	0.128	1.17 (0.92, 1.50)	0.493		
10	rs2031920	CYP2E1	C	T	0.14	1.19 (0.93, 1.52)	0.493		
15	rs762551	CYPIA2	A	C	0.239	1.13 (0.92, 1.40)	0.493		

Table 3. Associations of single nucleotide polymorphisms (SNPs) with gastric cancer risk under the univariate and multivariate models. Univariate analysis N = 1440, multivariable N = 1284. The referent allele is minor allele. *Chr* chromosome, *MAF* minor allele frequency. FDR significant q-values are bolded. *Adjusted for age, sex, CagA serostatus and wood cookstove use.

Characteristics	Multivariable model 1 OR (95% CI)[a]	Multivariable model 1 p-value	Multivariable model 2 OR (95% CI)[a]	Multivariable model 2 p-value	Final multivariable model (3) OR (95% CI)[a]	Final multivariable model (3) p-value
Age	1.04 (1.03, 1.05)	$< 2.00 \times 10^{-16}$	1.04 (1.03, 1.05)	$< 2.00 \times 10^{-16}$	1.04 (1.03, 1.05)	$< 2.00 \times 10^{-16}$
Sex	1.99 (1.55, 2.55)	5.41×10^{-8}	1.99 (1.56, 2.55)	4.89×10^{-8}	1.98 (1.55, 2.55)	5.70×10^{-8}
Wood stove use	2.26 (1.57, 3.29)	1.70×10^{-5}	0.95 (0.41, 2.29)	0.91	0.89 (0.39, 2.55)	0.80
Bacterial CagA serostatus	3.47 (2.38, 5.14)	2.32×10^{-10}	3.45 (2.37, 5.12)	2.44×10^{-10}	3.44 (2.36, 5.11)	3.14×10^{-10}
rs1800440	1.57 (1.14, 2.19)	0.007	0.57 (0.24, 1.41)	0.21	0.59 (0.25, 1.47)	0.25
rs1056836	1.41 (1.10, 1.80)	0.007	NA	NA	1.42 (1.11, 1.82)	0.006
Wood stove use:rs1800440	NA	NA	2.93 (1.11, 7.52)	0.027	3.05 (1.16, 7.81)	0.021

Table 4. Associations of multivariable models including interaction terms with gastric cancer. Multivariable model 1 variables: age, sex, wood stove use, bacterial CagA serostatus, rs1800440, and rs1056836. Multivariable model 2 variables: age, sex, wood stove use, bacterial CagA serostatus, rs1800440, and the interaction of wood stove use with rs1800440. Multivariable model 3 variables: age, sex, wood stove use, bacterial CagA serostatus, rs1800440, the interaction of wood stove use with rs1800440, and rs1056836. *95% CI* 95% confidence interval. p-values ≤ 0.05 are bolded. [a]Study N = 1284.

rs1800440 (Table 4). In both models 2 and 3, the interaction term was statistically significant. The likelihood ratio test identified model 3 with the interaction of wood stove use and rs1800440 as the best model (p = 0.025), and it was used for interpretation. Remarkably, the odds ratio of the interaction term in this model was of similar magnitude as the highly associated bacterial CagA serostatus.

Our results showed that genotype did not associate with GC risk for individuals who did not use a wood stove (p = 0.25, OR = 0.59). Nor did wood stove use associate with risk for individuals with GG/GA genotypes (p = 0.80, OR = 0.90). However, wood stove use was associated with GC risk in AA individuals compared to AA individuals who did not use a wood stove (p < 0.0001, OR = 2.73). Finally, for individuals who used a wood stove, AA genotype individuals were at increased risk (p = 0.0007, OR = 1.83) whereas in the absence of wood stove use the same genotype did not associate with GC (Fig. 1; Table 5). Similarly, wood stove use was not associated with GC for people who were GG/GA (Table S6). Multivariable models for rs105836, including the interaction of rs1056836 and wood stove showed no significant interaction between this SNP and wood stove use (Table S7). None of the remaining eight SNPs showed any interactions with woodstove use (Table S8).

Discussion

In the rural Central America Four region, we found a significant association between wood cookstove use and gastric adenocarcinoma. Wood stove use increased the risk of GC by more than 2.3 times after adjusting for demographic and risk factors (age, sex, CagA serostatus). This association was independent of the CagA genotype, the principal *H. pylori* virulence and GC risk factor. Importantly, the wood stove use association was significantly modified by the rs1800440 variant of *CYP1B1*, a gene involved in HCA and PAH metabolism, a determinant of the "internal dose". This SNP was only significant in the context of wood cookstove use, and neither wood stove use nor the SNP were significant when the interaction of this SNP and wood stove use was included in the model, implying strong biological mediation by genotype. Specifically, the rs1800440 AA genotype was high risk when wood stoves were used, but not otherwise, and wood stoves were only significant in the presence of this genotype. These two risk factors are highly prevalent in Central American populations.

The gene-environment interaction between wood stove use and genetics greatly informs our understanding with respect to who is most susceptible to adverse effects of cooking with a wood stove. Individuals who use a wood stove are at an increased risk of developing gastric cancer when they are homozygous for the major rs1800440 allele, whereas they do not appear to be at increased risk in the presence of a minor allele. Similarly, those who do not use a wood stove do not have a significant association based on genotype. This finding supports the concept that elevated risk associated with wood stove use may be mediated by HCA and PAH exposure generated in food preparation with wood stove use and is consistent with the recognized link between meat consumption and gastric and colon cancers. Latin American populations have a high prevalence of the rs1800440 risk allele frequency (86%-88%), as do African and Asian populations with 100% and 99% prevalence, respectively[38].

Studies examining the rs1800440 *CYP1B1* polymorphism and cancer risk have been inconclusive. The interaction of exposures, such as wood stove use and the variant, may explain the discordant literature related to *CYP1B1* genotypes[39–50]. Our study found that the common variant of the phase I *CYP1B1*, rs1800440, conferred approximately 40% increased risk of gastric cancer even after adjusting for age, sex, CagA status and wood stove use in analyses without a SNP-wood stove interaction term. However, we also noted that the rs1800440 SNP is only significant in the presence of wood stove exposure, a potential explanation of why some prior studies did not find any association between *CYP1B1* rs1800440 and gastric cancer[39]. In colorectal cancer studies, a meta-analysis found no association between colorectal cancer and rs1800440; however, there was considerable heterogeneity among the studies[36]. Carriers of the minor allele have been shown to confer either decreased risk of colorectal cancer, increased risk, or no risk in different studies[41,42]. Interestingly, in a Chinese population, the rs1800440 genotype AG associated with decreased hepatocellular carcinoma risk compared to the AA (homozygous major allele) genotype (adjusted OR = 0.33) when adjusting for similar variables to our study as well as smoking[51].
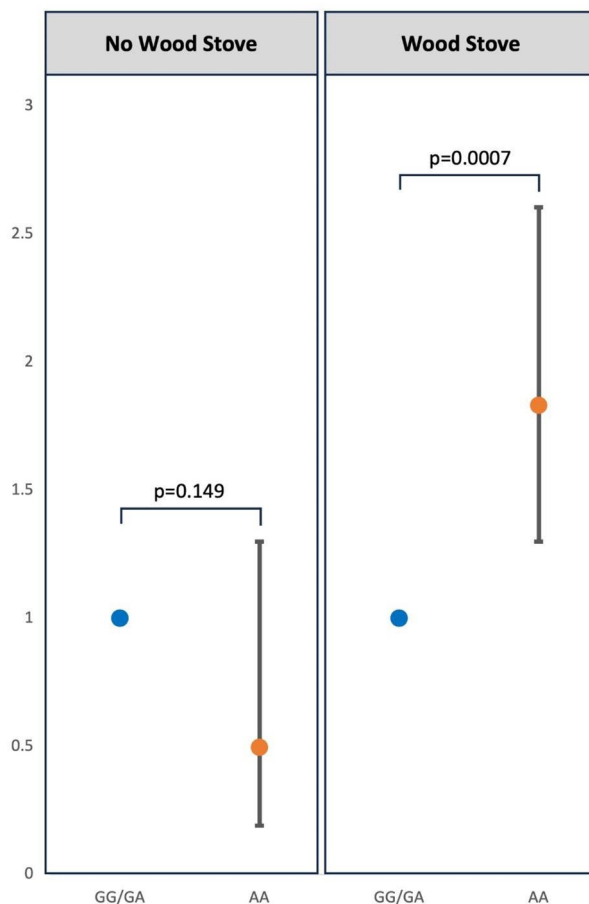
Figure 1. Interaction between wood stove use and the rs1800440 *CYP1B1* SNP. The effect of the rs1800440 AA genotype is dependent on wood stove use. Those that have the rs1800440 AA genotype and use a wood stove have an increased risk of gastric cancer compared to those with the rs1800440 GG or GA genotype and use a wood stove (OR = 1.83; 1.29, 2.60); p = 0.0007). The results derived from Model 3 (Table 5).

Characteristics	No wood stove (N = 184)		Wood stove (N = 1100)	
	OR (95% CI)ᵃ	p-value	OR (95% CI)ᵃ	p-value
Age	1.054 (1.031, 1.080)	< 0.0001	1.036 (1.027, 1.045)	< 0.0001
Sex	2.568 (1.223, 5.609)	0.0147	1.944 (1.494, 2.534)	< 0.0001
CagA serostatus	3.759 (1.407, 12.067)	0.0140	3.350 (2.231, 5.121)	< 0.0001
rs1800440	0.498 (0.191, 1.300)	0.1495	1.833 (1.296, 2.605)	0.0007
rs1056836	0.853 (0.410, 1.769)	0.6694	1.528 (1.173, 1.993)	0.0017

Table 5. Associations of multivariable models with gastric cancer stratified by wood stove use.

Importantly, the rs1800440 *CYP1B1* minor allele (G) *e*ncodes an amino acid substitution Asn453Ser with Ser being likely deleterious with respect to protein function according to Polyphen-2 (score = 0.906) (Fig. S2)[52]. This is consistent with a functional allele increasing internal doses of carcinogens, as in the case of the AA genotypes associating with risk in wood stove users. Our analyses showing a significant interaction between genotype and wood stove use may explain the inconsistencies as prior evidence also indicated that this phase I enzyme minor allele polymorphism may degrade more rapidly than the wild type enzyme and have a drastically shorter half-life, and therefore reduced metabolic activation of estrogens, PAHs, and HCAs[48].

Limited studies have examined the relationship between wood cookstove use and cancer, and specifically gastric cancer. Compounds generated by wood stoves are modified by metabolizing enzymes, and therefore an interaction between genes for these enzymes is not unexpected. The limited literature indicates that exposure to biomass burning and wood stove cooking has been linked to a variety of cancers of the upper aero-digestive tract, including lung, gastric and esophageal cancer[24,53–62]. The use of wood stoves, diets low in fruits and vegetables,

and higher elevation households have all been noted as significant risk factors in an epidemiologic profile of GC in Peru[24] and the association between cancers of the upper aero-digestive tract was confirmed in southern Brazil after adjusting for tobacco, alcohol, or dietary factors[53]. In addition, the use of biomass fuels was linked to gastrointestinal cancers, including esophageal cancer and gastric cancer, in a large Iranian cohort of men and women[62]. However, how these factors interact with enzymes that activate and deactivate carcinogens has not to our knowledge been studied previously.

Regular wood stove use also contribute to gastric cancer risk through the generation of PAHs. An association has been detected between biomass smoke exposure and gastric cancer, as well as biomass smoke exposure and urinary 1-hydroxypyrene (1-OHP), a PAH metabolite, with a significant trend for dose in a study from Zambia. However, no association was detected between the metabolite and gastric cancer status[61]. Notably, the exposures in sub-Saharan Africa may be less than Latin America as the cook stoves are typically outdoors and often involve the boiling of meats and foods. A recent study of Shanghai women found that women diagnosed with gastric cancer had higher urinary levels of 1-hydroxypyrene glucuronide (1-OHPG), a closely related PAH metabolite[63].

There is considerable potential for mitigation with improved cookstove design. The Global Alliance for Clean Cook stoves has proposed deployment of 100 million improved stoves[64]. Various designs and interventions have been implemented for the past two decades in order to decrease smoke exposure and deforestation, yet exposure to organic carcinogens in well-done meat and foods cooked at high temperature remain a challenge. Populations with a high prevalence of the risk rs1800440 genotype may benefit most from reduced wood stove use or a changes in food preparation.

The strengths of our study include the regional approach in a rural LMIC population with a high gastric cancer incidence, endemic *H. pylori* cagA infection, and a homogeneous diet. Not having precise quantification of HCA and PAH exposures was one of the study limitations. We did not examine dietary information and meat intake to further quantify HCA and PAH exposure, but this potential influence is dampened by the generally uniform diet in this rural region over time. Lastly, our asymptomatic population-based controls did not undergo endoscopy, and a limited number of control subjects may have had precancerous conditions or early gastric cancer, but this would only decrease the power to detect associations.

Conclusions

Wood cookstove use contributes to gastric cancer risk in the high-incidence regions of mountainous Central America, where *H. pylori* cagA infection is endemic. We found that the effect is mediated by host genetics, specifically the *CYP1B1* genotype, wherein wood cookstove usage increases risk only in individuals genetically predisposed, to thereby increase the conversion of pro-carcinogenic compounds to carcinogens. Further studies are indicated, as our results underscore the broad health and environmental impact of biomass cookstove use.

Data availability

The datasets generated and/or analyzed during the current study are available in the Mendeley Data repository, [Morgan, Douglas; Norwood, Dalton (2023), "Wood cookstove use is associated with gastric cancer in Central America, modified by the CYP1B1 genotype rs1800440, and independent of H. pylori cagA serostatus", Mendeley Data, V1, https://doi.org/10.17632/mrrjjp4j2w.1].

References

1. Torre, L. A. *et al.* Global cancer statistics, 2012. *CA Cancer J. Clin.* **65**, 87–108. https://doi.org/10.3322/caac.21262 (2015).
2. Bray, F. *et al.* Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J. Clin.* **68**, 394–424. https://doi.org/10.3322/caac.21492 (2018).
3. de Martel, C., Georges, D., Bray, F., Ferlay, J. & Clifford, G. M. Global burden of cancer attributable to infections in 2018: A worldwide incidence analysis. *Lancet Glob. Health* **8**, e180–e190. https://doi.org/10.1016/s2214-109x(19)30488-7 (2020).
4. Bray, F., Jemal, A., Grey, N., Ferlay, J. & Forman, D. Global cancer transitions according to the Human Development Index (2008–2030): A population-based study. *Lancet Oncol.* **13**, 790–801. https://doi.org/10.1016/s1470-2045(12)70211-5 (2012).
5. Norwood, D. A., Sanchez, E. M., Dominguez, R. L. & Morgan, D. R. Gastric cancer: Emerging trends in prevention, diagnosis, and treatment. *Gastroenterol. Clin.* **51**, 501–518. https://doi.org/10.1016/j.gtc.2022.05.001 (2022).
6. Torres, J. *et al.* Gastric cancer incidence and mortality is associated with altitude in the mountainous regions of Pacific Latin America. *Cancer Causes Control* **24**, 249–256. https://doi.org/10.1007/s10552-012-0114-8 (2013).
7. Montalvan-Sanchez, E. E. *et al.* 1140: The altitude enigma and gastric cancer incidence and mortality: A systematic review and meta-analysis. *Gastroenterology* **162**, S267–S268. https://doi.org/10.1016/S0016-5085(22)60634-6 (2022).
8. Chmiela, M., Karwowska, Z., Gonciarz, W., Allushi, B. & Staczek, P. Host pathogen interactions in Helicobacter pylori related gastric cancer. *World J. Gastroenterol.* **23**, 1521–1540. https://doi.org/10.3748/wjg.v23.i9.1521 (2017).
9. Fullerton, D. G., Bruce, N. & Gordon, S. B. Indoor air pollution from biomass fuel smoke is a major health concern in the developing world. *Trans. R Soc. Trop. Med. Hyg.* **102**, 843–851. https://doi.org/10.1016/j.trstmh.2008.05.028 (2008).
10. Turesky, R. J. & Le Marchand, L. Metabolism and biomarkers of heterocyclic aromatic amines in molecular epidemiology studies: Lessons learned from aromatic amines. *Chem. Res. Toxicol.* **24**, 1169–1214. https://doi.org/10.1021/tx200135s (2011).
11. Clark, M. L. *et al.* Indoor air pollution, cookstove quality, and housing characteristics in two Honduran communities. *Environ. Res.* **110**, 12–18. https://doi.org/10.1016/j.envres.2009.10.008 (2010).
12. Agudo, A. *et al.* Aromatic DNA adducts and risk of gastrointestinal cancers: A case-cohort study within the EPIC-Spain. *Cancer Epidemiol. Biomarkers Prev.* **21**, 685–692. https://doi.org/10.1158/1055-9965.Epi-11-1205 (2012).
13. Yaman, M. Comprehensive comparison of trace metal concentrations in cancerous and non-cancerous human tissues. *Curr. Med. Chem.* **13**, 2513–2525. https://doi.org/10.2174/092986706778201620 (2006).
14. Bouvard, V. *et al.* Carcinogenicity of consumption of red and processed meat. *Lancet Oncol.* **16**, 1599–1600. https://doi.org/10.1016/S1470-2045(15)00444-1 (2015).

15. Cross, A. J. & Sinha, R. Meat-related mutagens/carcinogens in the etiology of colorectal cancer. *Environ. Mol. Mutagen* **44**, 44–55. https://doi.org/10.1002/em.20030 (2004).
16. le Ngoan, T. *et al.* Cooking temperature, heat-generated carcinogens, and the risk of stomach and colorectal cancers. *Asian Pac. J Cancer Prev.* **10**, 83–86 (2009).
17. Fu, Z. *et al.* Using gene-environment interaction analyses to clarify the role of well-done meat and heterocyclic amine exposure in the etiology of colorectal polyps. *Am. J. Clin. Nutr.* **96**, 1119–1128. https://doi.org/10.3945/ajcn.112.040345 (2012).
18. Zhuo, W. *et al.* CYP2E1 RsaI/PstI polymorphism and gastric cancer susceptibility: Meta-analyses based on 24 case-control studies. *PLoS ONE* **7**, e48265. https://doi.org/10.1371/journal.pone.0048265 (2012).
19. Boccia, S. *et al.* Polymorphisms in metabolic genes, their combination and interaction with tobacco smoke and alcohol consumption and risk of gastric cancer: A case-control study in an Italian population. *BMC Cancer* **7**, 206. https://doi.org/10.1186/1471-2407-7-206 (2007).
20. Kato, S. *et al.* Cytochrome P4502E1 (CYP2E1) genetic polymorphism in a case-control study of gastric cancer and liver disease. *Pharmacogenetics* **5**, S141-144. https://doi.org/10.1097/00008571-199512001-00016 (1995).
21. Nishimoto, I. N. *et al.* Cytochrome P450 2E1 polymorphism in gastric cancer in Brazil: Case-control studies of Japanese Brazilians and non-Japanese Brazilians. *Cancer Epidemiol. Biomarkers Prev.* **9**, 675–680 (2000).
22. Wu, M. S. *et al.* Genetic polymorphisms of cytochrome p450 2E1, glutathione S-transferase M1 and T1, and susceptibility to gastric carcinoma in Taiwan. *Int. J. Colorectal Dis.* **17**, 338–343. https://doi.org/10.1007/s00384-001-0383-2 (2002).
23. Corral, J. E. *et al.* The descriptive epidemiology of gastric cancer in Central America and comparison with United States Hispanic populations. *J. Gastrointest. Cancer* **46**, 21–28. https://doi.org/10.1007/s12029-014-9672-1 (2015).
24. Chirinos, J. L., Carbajal, L. A., Segura, M. D., Combe, J. & Akiba, S. Gastric cancer: Epidemiologic profile 2001–2007 in Lima, Peru. *Rev. Gastroenterol. Peru* **32**, 58–64 (2012).
25. Dominguez, R. L. *et al.* Gastric cancer incidence estimation in a resource-limited nation: Use of endoscopy registry methodology. *Cancer Causes Control* **24**, 233–239. https://doi.org/10.1007/s10552-012-0109-5 (2013).
26. Estevez-Ordonez, D. *et al.* Health barriers and patterns of gastric cancer care in rural central american resource-limited settings. *JAMA Oncol.* **4**, 1131–1133. https://doi.org/10.1001/jamaoncol.2018.2570 (2018).
27. Norwood, D. A. *et al.* Western honduras copan population-based cancer registry: Initial estimates and a model for rural central America. *JCO Glob. Oncol.* **7**, 1694–1702. https://doi.org/10.1200/GO.21.00273 (2021).
28. Miller, A. K. *et al.* Ornithine decarboxylase (ODC1) gene variant (rs2302615) is associated with gastric cancer independently of *Helicobacter pylori* CagA serostatus. *Oncogene* **40**(40), 5963–5969. https://doi.org/10.1038/s41388-021-01981-5 (2021).
29. Salvador, I., Mercado, A., Bravo, G. L., Baldeon, M. & Fornasini, M. Risk and protective factors for gastric metaplasia and cancer: A hospital-based case-control study in ecuador. *Nutr. Hosp.* **32**, 1193–1199. https://doi.org/10.3305/nh.2015.32.3.9257 (2015).
30. Michel, A., Waterboer, T., Kist, M. & Pawlita, M. *Helicobacter pylori* multiplex serology. *Helicobacter* **14**, 525–535. https://doi.org/10.1111/j.1523-5378.2009.00723.x (2009).
31. Butt, J. *et al.* Differences in antibody levels to H. pylori virulence factors VacA and CagA among African Americans and whites in the Southeast USA. *Cancer Causes Control* **31**, 601–606. https://doi.org/10.1007/s10552-020-01295-z (2020).
32. Varga, M. G. *et al.* Immunostimulatory membrane proteins potentiate H. pylori-induced carcinogenesis by enabling CagA translocation. *Gut Microbes* **13**, 1–13. https://doi.org/10.1080/19490976.2020.1862613 (2021).
33. Purcell, S. *et al.* PLINK: A tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* **81**, 559–575. https://doi.org/10.1086/519795 (2007).
34. Fujii-Kuriyama, Y. & Mimura, J. Molecular mechanisms of AhR functions in the regulation of cytochrome P450 genes. *Biochem. Biophys. Res. Commun.* **338**, 311–317. https://doi.org/10.1016/j.bbrc.2005.08.162 (2005).
35. Turesky, R. J. Formation and biochemistry of carcinogenic heterocyclic aromatic amines in cooked meats. *Toxicol. Lett.* **168**, 219–227. https://doi.org/10.1016/j.toxlet.2006.10.018 (2007).
36. Barrett, J. C., Fry, B., Maller, J. & Daly, M. J. Haploview: Analysis and visualization of LD and haplotype maps. *Bioinformatics* **21**, 263–265. https://doi.org/10.1093/bioinformatics/bth457 (2005).
37. Pilotto, A. & Franceschi, M. *Helicobacter pylori* infection in older people. *World J. Gastroenterol.* **20**, 6364–6373. https://doi.org/10.3748/wjg.v20.i21.6364 (2014).
38. Phan, L., Zhang, H., Qiang, W., Shekhtman, E., Shao, D., Revoe, D., Villamarin, R., Ivanchenko, E., Kimura, M., Wang, Z. Y., Hao, L., Sharopova, N., Bihan, M., Sturcke, A., Lee, M., Popova, N., Wu, W., Bastiani, C., Ward, M., Holmes, J. B., Lyoshin, V., Kaur, K., Moyer, E., Feolo, M., & Kattman, B. L. *ALFA: Allele Frequency Aggregator*, www.ncbi.nlm.nih.gov/snp/docs/gsr/alfa/ (2020).
39. Freedman, N. D. *et al.* Polymorphisms in estrogen- and androgen-metabolizing genes and the risk of gastric cancer. *Carcinogenesis* **30**, 71–77. https://doi.org/10.1093/carcin/bgn258 (2009).
40. Mei, Q., Zhou, D., Han, J., Lu, H. & Tang, B. CYP1B1 Asn453Ser polymorphism and colorectal cancer risk: A meta-analysis. *Metabolism* **61**, 1321–1329. https://doi.org/10.1016/j.metabol.2012.02.010 (2012).
41. Hlavata, I. *et al.* Association between exposure-relevant polymorphisms in CYP1B1, EPHX1, NQO1, GSTM1, GSTP1 and GSTT1 and risk of colorectal cancer in a Czech population. *Oncol. Rep.* **24**, 1347–1353. https://doi.org/10.3892/or_00000992 (2010).
42. Cotterchio, M. *et al.* Red meat intake, doneness, polymorphisms in genes that encode carcinogen-metabolizing enzymes, and colorectal cancer risk. *Cancer Epidemiol. Biomarkers Prev.* **17**, 3098–3107. https://doi.org/10.1158/1055-9965.Epi-08-0341 (2008).
43. Liu, J. Y. *et al.* Association between the CYP1B1 polymorphisms and risk of cancer: A meta-analysis. *Mol. Genet. Genomics* **290**, 739–765. https://doi.org/10.1007/s00438-014-0946-x (2015).
44. Shen, M. *et al.* Quantitative assessment of the influence of CYP1B1 polymorphisms and head and neck squamous cell carcinoma risk. *Tumour Biol.* **35**, 3891–3897. https://doi.org/10.1007/s13277-013-1516-2 (2014).
45. Gajjar, K., Owens, G., Sperrin, M., Martin-Hirsch, P. L. & Martin, F. L. Cytochrome P1B1 (CYP1B1) polymorphisms and ovarian cancer risk: A meta-analysis. *Toxicology* **302**, 157–162. https://doi.org/10.1016/j.tox.2012.09.009 (2012).
46. Wang, F., Zou, Y. F., Sun, G. P., Su, H. & Huang, F. Association of CYP1B1 gene polymorphisms with susceptibility to endometrial cancer: A meta-analysis. *Eur. J. Cancer Prev.* **20**, 112–120. https://doi.org/10.1097/CEJ.0b013e3283410193 (2011).
47. Economopoulos, K. P. & Sergentanis, T. N. Three polymorphisms in cytochrome P450 1B1 (CYP1B1) gene and breast cancer risk: A meta-analysis. *Breast Cancer Res. Treat.* **122**, 545–551. https://doi.org/10.1007/s10549-009-0728-z (2010).
48. Abnet, C. C. *et al.* The influence of genetic polymorphisms in Ahr, CYP1A1, CYP1A2, CYP1B1, GST M1, GST T1 and UGT1A1 on urine 1-hydroxypyrene glucuronide concentrations in healthy subjects from Rio Grande do Sul Brazil. *Carcinogenesis* **28**, 112–117. https://doi.org/10.1093/carcin/bgl131 (2007).
49. Schults, M. A. *et al.* Genetic polymorphisms in catalase and CYP1B1 determine DNA adduct formation by benzo(a)pyrene ex vivo. *Mutagenesis* **28**, 181–185. https://doi.org/10.1093/mutage/ges070 (2013).
50. Church, T. R. *et al.* Interaction of CYP1B1, cigarette-smoke carcinogen metabolism, and lung cancer risk. *Int. J. Mol. Epidemiol. Genet.* **1**, 295–309 (2010).
51. Liu, F. *et al.* Polymorphisms of the CYP1B1 gene and hepatocellular carcinoma risk in a Chinese population. *Gene* **564**, 14–20. https://doi.org/10.1016/j.gene.2015.03.035 (2015).
52. Adzhubei, I. A. *et al.* A method and server for predicting damaging missense mutations. *Nat. Methods* **7**, 248–249. https://doi.org/10.1038/nmeth0410-248 (2010).
53. Pintos, J., Franco, E. L., Kowalski, L. P., Oliveira, B. V. & Curado, M. P. Use of wood stoves and risk of cancers of the upper aerodigestive tract: A case-control study. *Int. J. Epidemiol.* **27**, 936–940. https://doi.org/10.1093/ije/27.6.936 (1998).

54. Velema, J. P. *et al.* Burning wood in the kitchen increases the risk of cervical neoplasia in HPV-infected women in Honduras. *Int. J. Cancer* **97**, 536–541. https://doi.org/10.1002/ijc.1622 (2002).
55. Sierra-Torres, C. H., Arboleda-Moreno, Y. Y. & Orejuela-Aristizabal, L. Exposure to wood smoke, HPV infection, and genetic susceptibility for cervical neoplasia among women in Colombia. *Environ. Mol. Mutagen* **47**, 553–561. https://doi.org/10.1002/em.20228 (2006).
56. Kayamba, V. *et al.* HIV infection and domestic smoke exposure, but not human papillomavirus, are risk factors for esophageal squamous cell carcinoma in Zambia: A case-control study. *Cancer Med.* **4**, 588–595. https://doi.org/10.1002/cam4.434 (2015).
57. Wang, Y., Vogelsang, M., Schäfer, G., Matejcic, M. & Parker, M. I. MicroRNA polymorphisms and environmental smoke exposure as risk factors for oesophageal squamous cell carcinoma. *PloS One* **8**, e78520. https://doi.org/10.1371/journal.pone.0078520 (2013).
58. Mota, O. M., Curado, M. P., Oliveira, J. C., Martins, E. & Cardoso, D. M. Risk factors for esophageal cancer in a low-incidence area of Brazil. *Sao Paulo Med. J.* **131**, 27–34. https://doi.org/10.1590/s1516-31802013000100005 (2013).
59. Dandara, C., Li, D. P., Walther, G. & Parker, M. I. Gene-environment interaction: The role of SULT1A1 and CYP3A5 polymorphisms as risk modifiers for squamous cell carcinoma of the oesophagus. *Carcinogenesis* **27**, 791–797. https://doi.org/10.1093/carcin/bgi257 (2006).
60. Li, D., Dandara, C. & Parker, M. I. The 341C/T polymorphism in the GSTP1 gene is associated with increased risk of oesophageal cancer. *BMC Genet.* **11**, 47. https://doi.org/10.1186/1471-2156-11-47 (2010).
61. Kayamba, V. *et al.* Biomass smoke exposure is associated with gastric cancer and probably mediated via oxidative stress and DNA damage: A case-control study. *JCO Glob. Oncol.* **6**, 532–541. https://doi.org/10.1200/go.20.00002 (2020).
62. Sheikh, M. *et al.* Household fuel use and the risk of gastrointestinal cancers: The Golestan cohort study. *Environ. Health Perspect.* **128**, 67002. https://doi.org/10.1289/EHP5907 (2020).
63. Liao, L. M. *et al.* Polycyclic aromatic hydrocarbons and risk of gastric cancer in the Shanghai Women's Health Study. *Int. J. Mol. Epidemiol. Genet.* **5**, 140–144 (2014).
64. Subramanian, M. Global health: Deadly dinners. *Nature* **509**, 548–551. https://doi.org/10.1038/509548a (2014).

Acknowledgements

Author contributions

D.R.M. and R.L.D. designed and implemented the study. R.L.D., E.M., E.M.S., D.N., and D.R.M. executed the study. T.W. designed, performed, and analyzed the *H. pylori* multiplex assay. A.K.M., S.M.W., S.B.R., M.B., and D.R.M. provided the overall statistical analysis and oversight. All authors provided the critical review of the manuscript and approved the final manuscript.

Funding

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-023-42973-7.

Correspondence and requests for materials should be addressed to D.R.M.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.