Carcinogenesis vol.32 no.8 pp.1216–1222, 2011
doi:10.1093/carcin/bgr095
Advance Access publication May 23, 2011

Genetic associations with sporadic neuroendocrine tumor risk

Monica Ter-Minassian[1,2,*], Zhaoxi Wang[1], Kofi Asomaning[1], Michael C.Wu[3], Chen-Yu Liu[1], Jessica K.Paulus[4], Geoffrey Liu[5], Penelope A.Bradbury[5], Rihong Zhai[1], Li Su[1], Christine S.Frauenhoffer[2], Susanne M.Hooshmand[2], Immaculata De Vivo[4], Xihong Lin[3], David C.Christiani[1,4,6] and Matthew H.Kulke[2]

[1]Environmental and Occupational Medicine and Epidemiology Program, Department of Environmental Health, Harvard School of Public Health, 677 Huntington Avenue, Boston, MA 02115, USA, [2]Department of Medical Oncology, Dana-Farber Cancer Institute, Boston, MA 02115, USA, [3]Department of Biostatistics, Harvard School of Public Health, 677 Huntington Avenue, Boston, MA 02115, USA, [4]Department of Epidemiology, Harvard School of Public Health, 677 Huntington Avenue, Boston, MA 02115, USA, [5]Princess Margaret Hospital, Ontario Cancer Institute, Toronto, Ontario, Canada M5G 2C4 and [6]Department of Medicine, Massachusetts General Hospital, Harvard Medical School, Boston, MA 02115, USA

[*]To whom correspondence should be addressed. Tel: 617 432 6458; Fax: 617 432 0219;
Email: mtermina@hsph.harvard.edu

Genetic risk factors for sporadic neuroendocrine tumors (NET) are poorly understood. We tested risk associations in patients with sporadic NET and non-cancer controls, using a custom array containing 1536 single-nucleotide polymorphisms (SNPs) in 355 candidate genes. We identified 18 SNPs associated with NET risk at a *P*-value <0.01 in a discovery set of 261 cases and 319 controls. Two of these SNPs were found to be significantly associated with NET risk in an independent replication set of 235 cases and 113 controls, at a *P* value ≤0.05. An SNP in interleukin 12A (*IL12A* rs2243123), a gene implicated in inflammatory response, replicated with an adjusted odds ratio (95% confidence interval) (aOR) = 1.47 (1.03, 2.11) *P*-trend = 0.04. A second SNP in defender against cell death, (*DAD1* rs8005354), a gene that modulates apoptosis, replicated at aOR = 1.43 (1.02, 2.02) *P*-trend = 0.04. Consistent with our observations, a pathway analysis, performed in the discovery set, suggested that genetic variation in inflammatory pathways or apoptosis pathways is associated with NET risk. Our findings support further investigation of the potential role of *IL12A* and *DAD1* in the etiology of NET.

Introduction

The diagnosed incidence of neuroendocrine tumors (NET) is increasing; however, their etiology remains poorly understood. In a recent analysis of the Surveillance, Epidemiology and End Results database, the annual incidence of NET in the USA had increased from a baseline of 1–2/100 000 individuals to 5.25/100 000 individuals in the past few decades; furthermore, the estimated prevalence of living individuals diagnosed with NETs in the USA exceeded 100 000 (1).

Classic cancer risk factors have not been implicated in the development of most NET. Although smoking has been associated with aggressive neuroendocrine cancer histologies such as small cell lung cancer (2), no clear association between smoking and well differentiated NETs, which comprise the majority of NET cases, has been identified (3,4) . Similarly, no association has been reported between NET and alcohol use.

NETs are often subclassifed as either pancreatic NET or carcinoid tumors. Pancreatic NET, in particular, may develop in the context of rare autosomal-dominant inherited genetic syndromes particularly in multiple endocrine neoplasia 1 or 2 (MEN1 or 2) and Von Hippel–Lindau syndrome, but less frequently in tuberous sclerosis (TSC) or neurofibromatosis type 1 (5,6). Several studies have also reported a familial aggregation of carcinoid tumors (7,8) and a higher prevalence of NETs in patients whose first-degree relatives develop any cancer (9). These observations suggest a possible heritable basis for sporadic NETs.

Large-scale candidate gene association studies have previously identified genetic variants associated with more common malignancies, but have not, to our knowledge, been performed in NET (10–13). We sought to establish whether common genetic variants in established cancer-related genes contribute to the risk of sporadic NET. We evaluated 1536 single-nucleotide polymorphisms (SNPs) for potential risk associations in an initial discovery case–control set and then confirmed potential associations in an independent replication set of NET cases and healthy controls.

Abbreviations: AML, admixture maximum likelihood; DFCI, Dana-Farber Cancer Institute; FDR, false discovery rate; MAF, minor allele frequency; NET, neuroendocrine tumor; SNP, single-nucleotide polymorphism; TSC, tuberous sclerosis.

Materials and methods

Study population

This study was approved by the Human Subjects Committees of Massachusetts General Hospital, Dana-Farber Cancer Institute (DFCI) and the Harvard School of Public Health in Boston, MA. Patients with a confirmed diagnosis of NET were recruited at the DFCI. Patients completed a questionnaire prior to time of first visit that included demographic, socioeconomic and lifestyle data as well as past medical and family cancer history. Unreported clinical information, including age, gender, smoking history, tumor type and tumor stage was extracted from the medical record. Patients that reported not currently smoking and had no other smoking history in the medical record were coded as never-smokers. Over 95% of all NET patients seen at DFCI consented to participate in the study, and of the consenting cases, 89% permitted blood sample collection. Cases in the discovery set were recruited from 1 July 2003 to 23 March 2007. Dates of initial diagnosis ranged from 15 June 1967 to 22 March 2007. Cases in the replication set were recruited from 8 August 2003 to 24 February 2009 (79% after 23 March 2007) with initial dates of diagnosis ranging from 1 July 1969 to 24 February 2009.

Two different control groups were used for discovery and replication. To be included in the study, controls were required to have complete information on age, gender and smoking history as well as an adequate quantity of available DNA. For the discovery set, controls were selected from healthy friends and non-blood related family members (usually spouses), who had been recruited and volunteered for the Harvard Lung Cancer Susceptibility Study conducted from 1992–2007 at the Massachusetts General Hospital/Harvard Cancer Center. Interviewer-administered questionnaires adapted from American Thoracic Society questionnaire (14) were used to obtain information on demographic and detailed smoking histories from each subject (15,16). For the replication set, healthy controls were recruited from friends and non-blood related family members who accompanied the patients to the DFCI clinic from 20 January 2006 to 24 February 2009 and volunteered to participate in the study. Controls in the replication set completed the same questionnaire as the cases.

Sample collection and preparation

Peripheral whole blood samples for cases and controls were stored at −80°C. DNA was extracted from peripheral blood samples using the Puregene DNA Isolation Kit (Qiagen, Valencia, CA) and stored at −20°C. Concentrations of genomic DNA were measured by PicoGreen dsDNA quantitation reagent Molecular Probes (Invitrogen, Carlsbad, CA) at a dilution of 50 ng/μl.

Gene and SNP selection and genotyping

We designed an Illumina GoldenGate custom 1536 loci panel on a BeadArray chip (San Diego, CA) to include both tagging and functional SNPs in a broad range of genes implicated in the development of malignancy (supplementary Table I is available at *Carcinogenesis* Online). The selected genes included those known to be involved in key cell-signaling pathways, including mammalian target of rapamycin signaling, DNA repair, cell cycle control, apoptosis, inflammation and xenobiotic and endogenous metabolism. Predicted functional and tagging status was the primary selection criterion. Predicted genotyping performance was considered a secondary selection criterion and SNPs with both excellent and moderate performance characteristics were included (17).

Genotyping the Illumina Goldengate BeadArray for the discovery set was performed at the Broad Institute, Cambridge, MA, on a BeadStation system by laboratory personnel blinded to case–control status, with 48 duplicate experimental control samples. We identified eighteen SNPs with the discovery set associated with risk at $P \leq 0.01$, excluding SNPs in strong linkage disequilibrium [$D' = 1$ with HapMap CEU (CEPH, Centre d'Etude du Polymorphisme Humain; (Utah residents with ancestry from northern and western Europe)] release 24 data (18). We then genotyped these SNPs with the replication case–control set. These SNPs were genotyped by Partners Healthcare Center for Personalized Genomic Medicine, Cambridge, MA with Sequenom iPLEX with 40 duplicate experimental control samples.

The concordance rate of the 48 duplicate samples in the discovery set was 99.6%. Two authors reviewed all genotyping results independently. From 1536 SNPs genotyped, 1334 SNPs in 354 known cancer genes remained that had <10% missing genotypes, Hardy–Weinberg equilibrium χ^2-test *P*-value of >0.001 and minor allele frequency (MAF) >0.01 in Caucasian controls. Of these, 1311 SNPs had <5% missing genotypes. SNP MAFs did not differ from HapMap CEU or other Caucasian frequencies in dbSNP by >15%. The concordance rate of the 40 duplicate samples in the replication set was 100%. Three of 18 SNPs failed genotyping due to poor extension of the primer. The remaining 15 SNPs had <15% missing genotypes, with an average of 6% missing. All successfully genotyped SNPs passed the Hardy–Weinberg equilibrium χ^2-test at *P*-value of ≥0.05 and had a MAF <1% in Caucasian controls.

Statistical analysis

Individuals of all races were recruited for this study. To reduce population stratification, we restricted our analyses to self-reported Caucasians only. We retained subjects with complete information on age, gender and each SNP individually.

We tested each SNP for association with NET susceptibility using multiple logistic regression under dominant and additive (test for linear trend) genetic models, adjusting for age, gender and smoking status in the discovery set and adjusting for age and gender in the replication set. We used SAS/Genetics software (version 9.1.3; SAS Institute, Cary, NC) to perform the analyses. Cases and controls were classified as either current smokers, former smokers (those who quit smoking for >1 year prior to diagnosis or enrollment) or never-smokers (<100 cigarettes in their lifetime). Age was modeled as a continuous covariate. *P*-values were reported along with a Benjamini–Hochberg false discovery rate (FDR) multiple testing adjustment reported as *q* values (19). In the discovery set, we tested the association of each SNP in subgroups of tumor type: pancreatic islet cell tumor only and small bowel carcinoids only. For replicated SNPs, we tested the difference in the SNP associations with the risk of NET between the two datasets statistically, with a cross-product interaction term between the SNP (additive model) and an indicator term for discovery versus replication datasets in a multivariate regression model. If there was no significant difference, we combined the datasets for an overall aOR.

For the discovery set, at $\alpha = 0.01$ under the additive genetic model, we estimated >80% power to detect genetic odds ratios 1.7 for SNPs with MAFs equal to 15% and 1.53 for SNPs with MAF equal to 30%. For the replication set at $\alpha = 0.05$, under the additive genetic model, we estimated >80% power to detect genetic odds ratios of 1.8 for SNPs with MAF equal to 15% and 1.63 model for SNPs with MAF equal to 30%.

To identify pathways potentially important in NET risk, we tested SNPs in 13 cancer related pathways or functional gene groups for association with NET using the admixture maximum likelihood (AML) test in our discovery set (11,20). We used 1000 permutations and fixed the maximum proportion of associated SNPs at 20%.

Results

Population characteristics

We identified 315 cases from DFCI and 350 controls from the Harvard Lung Cancer Susceptibility study at Massachusetts General Hospital for a discovery set, and 271 cases and 155 friend/non-blood related family member controls from DFCI for a replication set. For subsequent analysis, we excluded cases and controls with non-Caucasian self-reported race or known genetic syndromes (e.g. MEN1, MEN2, Von Hippel–Lindau or TSC) or incomplete information on age or sex. We also excluded discovery cases and controls with >10% missing genotype information and missing smoking status or replication cases and controls with >30% missing genotype information. A total of 261 cases and 319 controls were used for the discovery analysis; and a total of 235 cases and 113 controls were used for replication.

The demographics of the two study populations are summarized in Table I. In the discovery set, 54 (21%) patients had pancreatic NET (PET) and 207 (79%) patients had carcinoid; among the carcinoid patients, 91 had primary small bowel carcinoid. The distribution of

Table I. Descriptive characteristics of the NET cases and controls[a]

Characteristics	Discovery set		Replication set	
	Cases (n = 261)	Controls (n = 319)	Cases (n = 235)	Controls (n = 113)
Age[b]	52 (15–86)	56 (30–83)	55 (18–83)	53 (26–85)
Gender				
Females	140 (53.6%)	138 (43.3%)	112 (47.7%)	72 (63.7%)
Males	121 (46.4%)	181 (56.7%)	123 (53.3%)	41 (36.3%)
Smoking status[c]				
Never-smoker	120 (45.9%)	112 (35.1%)	117 (51.5%)	56 (50%)
Ex-smoker	124 (48.3%)	143 (44.8%)	96 (42.3%)	44 (39.3%)
Current smoker	17 (5.8%)	64 (20.1%)	14 (7.0%)	12 (10.7%)
Site of origin				
Pancreatic islet cell	54 (20.7%)		47 (20%)	
Small bowel	91 (34.9%)		92 (39.2%)	
Lung	24 (9.1%)		25 (10.6%)	
Appendix	19 (7.3%)		11 (4.7%)	
Stomach	8 (3.1%)		5 (2.1%)	
Other[d]	23 (8.8%)		23 (9.8%)	
Unknown primary	42 (16.1%)		32 (13.6%)	
Stage[e]				
M0	118 (45.2%)		106 (45.1%)	
M1	143 (54.7%)		129 (54.9%)	

[a]Discovery cases were recruited for the study at the DFCI (2003–2007) and discovery controls at Massachusetts General Hospital (1992–2007); replication cases and controls were recruited at the DFCI (2003–2009), 79% after 2007.
[b]Median (range).
[c]In the replication set, eight cases and one control had missing data on smoking status.
[d]Other sites include colon (5, 4), rectum (7, 8), anus (1, 0), thorax (2, 0), larynx (2, 0), heart (1, 0), thyroid (1, 0) and rare (4, 11) in the discovery and replication sets, respectively.
[e]M0, no metastasis after resection at initial diagnosis; M1, metastasis at initial diagnosis.

tumor subgroups in our cases differs from population estimates (21) and reflects a high proportion of gastrointestinal NETs, probably due to accrual of cases in the gastrointestinal cancer clinic at DFCI. When restricted to gastrointestinal carcinoids alone, the frequencies are similar to those reported in population-based series (21).

Tumor stage was evenly distributed between patients who had undergone complete tumor resection (M0) [117 (45%) cases] and those who were diagnosed with metastatic disease (M1) [144 (55%)]. Primary tumor site and stage distributions in the replication set were similar to those in the discovery set. We did note imbalances among the groups in both smoking characteristics and gender distribution. Although smoking characteristics were similar among all the cases and in the controls recruited from DFCI, we observed a lower proportion of never-smokers in the control group from the Harvard Lung Cancer Susceptibility study. We additionally observed a higher proportion of females among the friend/non-blood related family member controls recruited from the DFCI compared with the other groups. Allele frequencies between the two control sets, however, were similar; the average MAF difference was 4% and the average genotype frequency difference was 6% among 15 SNPs subsequently genotyped in the replication cohort. The median difference for both was 3%.

Association of SNPs and overall NET risk

We used multiple logistic regression analysis to evaluate risk associations among 1334 selected SNPs, representing 354 genes in 14 pathways (Table II) in the discovery case and control set. We adjusted for age, gender and smoking status in the analysis, given potential associations of these variables with NET and observed differences in our control groups. Our discovery analysis revealed that 18 SNPs were associated with NET risk at a *P*-value $\leq$0.01, under either the additive or dominant genetic model, meeting our criterion for replication.

Of the 18 SNPs identified in the discovery analysis, we successfully genotyped 15 SNPs in our replication case and control sets. After adjusting for the significant confounders of age and sex, we found that two SNPs, *IL12A* rs2243123 (an intronic tag SNP in interleukin 12A) and *DAD1* rs8005354 (an intronic tag SNP in defender against cell death one), were associated with NET risk at a *P*-value of $\leq$0.05 (Table III).

In the discovery analysis, the adjusted odds ratio (aOR) under the additive model for *IL12A* rs2243123 (95% CI) was 1.43 (1.09, 1.89), $P = 0.01$, FDR $q = 0.61$, (*P*-dominant = 0.004); in the replication set, the aOR was 1.47 (1.03, 2.12), $P = 0.04$, FDR $q = 0.24$ (Table IV). To evaluate potential differences in SNP associations in the discovery and replication datasets, we tested for the interaction between *IL12A* rs2243123 and the two datasets; we found no evidence for an interaction. In the combined discovery and replication samples, the aOR for *IL12A* rs2243123 under the additive model remained significant, 1.37 (1.12, 1.69), $P = 0.002$.

For *DAD1* rs8005354, the aOR under the additive model was 1.31 (1.03, 1.67), $P = 0.028$, FDR $q = 0.62$, (*P*-dominant = 0.006) in the discovery set. In the replication set, the aOR was 1.43 (1.02, 2.02), $P = 0.04$, FDR $q = 0.24$. The MAF of *DAD1* rs8005354 were similar in the discovery and replication cases, and in the discovery and replication controls, again supporting a potential association (Table IV). We found no evidence of a difference between *DAD1* rs8005354 risk associations in the two datasets. In the combined discovery and replication samples, the aOR for *DAD1* rs8005354 under the additive model was 1.32 (1.09, 1.59), $P = 0.004$.

The SNP most strongly associated with NET risk in our discovery analysis was *TSC2* rs13337626, a synonymous SNP in a gene encoding a subunit of the TSC complex. This SNP failed genotyping in the replication set using the initial Sequenom iPLEX genotyping platform. Given known associations between TSC and NET, we subsequently evaluated this SNP using Taqman as an alternate genotyping platform and also evaluated five additional SNPs in strong linkage disequilibrium with *TSC2* rs13337626 for potential associations with NET. We failed to observe any significant risk associations with these approaches, suggesting that the initial risk association observed with rs13337626 was spurious.

Associations of SNPs with specific tumor subtypes

A limited number of cases precluded our ability to comprehensively identify and replicate risk associations for different tumor subtypes in the discovery and replication sets. However, in the discovery analysis, different sets of genes emerged as significant in the pancreatic NET and small bowel carcinoid subgroups (Table V). Eighteen SNPs were associated with pancreatic NETs at the $P \leq 0.01$ level, although none passed multiple testing adjustment. These included SNPs in the genes *LIG3* (ligase 3), *CDKN2A* (cyclin D kinase2A), *BCL2* (b-cell CLL/lymphoma 2), *VEGFR* (vascular endothelial growth factor receptor) and *PTEN* (phosphatase and tensin homolog). Twenty-two SNPs (excluding *TSC2* rs13337626 for reasons described above) were associated with small bowel carcinoid tumors at the $P \leq 0.01$ level. *IL1RN* rs380092 and *CYP1B1* (cytochrome P450 1B1) rs162562 passed FDR multiple testing adjustment at the 20% level. Neither *IL12A* rs2243123 nor *DAD1* rs8005354 reached significance in either tumor subgroup at $P \leq 0.01$ in the discovery set alone. However, *IL12A* rs2243123 was the only SNP that reached significance in either tumor subgroup at $P \leq 0.01$ in the replication set: additive aOR = 2.14 (1.27, 3.60), $P = 0.004$ associated with pancreatic NET.

Associations of NET with pathways

To identify which pathways may be important in NET risk, we tested NET risk associations within gene functional groups using the AML in our discovery set. AML tests for the presence of more positive associations than expected by chance, estimates the proportion of associated SNPs and their typical effect size, and evaluates statistical significance via Monte Carlo simulation (11,20). Of the 13 pathways evaluated (see Table II), only the apoptosis (*P*-heterogeneity = 0.015, *P*-trend = 0.009) and inflammation pathways (*P*-heterogeneity = 0.013, *P*-trend = 0.004) were associated with NET risk using AML. After excluding SNPs from the most highly associated gene in each pathway, only the apoptosis pathway remained associated with NET risk (*P*-heterogeneity = 0.04, *P*-trend = 0.04).

Table II. Genes and SNPs evaluated by pathway

	Number of genes ($n = 354$)[a]	Number of SNPs ($n = 1334$)[a]
Metabolism	71	202
DNA repair	50	136
Cell cycle	46	165
Inflammation	38	228
Apoptosis	36	233
Cell growth/IGF	33	62
Mammalian target of rapamycin	27	214
Metastasis	15	18
Angiogenesis	11	126
Hormone	8	45
Immunity	8	14
Transporter	8	16
DNA methylation	3	15
Other	18	33

[a]Total is not a sum due to overlap in gene and SNP count.

Discussion

Few prior studies have evaluated the effect of heritable genetic variation on NET risk (22,23) In one small case–control study, variants of the tumor necrosis factor alpha gene and interleukin 2 gene, *TNFA*-1031 and *IL2*–330 T > G, were associated with gastroenteropancreatic NET risk (22,24). Our findings also suggest a possible role for inflammation in the etiology of NET. In this study, we performed a large-scale analysis of genetic variation in candidate genes and NET

Table III. SNP associations with NET in the discovery and replication sets (selected SNPs are those associated with NET risk in the discovery set with adjusted[a] *P*-value under 0.01)

Gene	Variable	Discovery set (n = 261 cases, 319 controls)					Replication set (n = 235 cases, 113 controls)				
		Dominant aOR* (95% CI)	Dominant *P*-value	Additive aOR* (95% CI)	Additive *P*-value	HWE-*P*	Dominant aOR* (95% CI)	Dominant *P*-value	Additive aOR[a] (95% CI)	Additive *P*-value	HWE-*P*
TSC2	rs13337626	2.82 (1.89, 4.19)	3.18E-07	2.46 (1.69, 3.57)	2.69E-06	0.50	Failed genotyping				
IL1RN	rs380092	1.87 (1.32, 2.65)	0.0004	1.65 (1.27, 2.15)	0.0002	0.61	0.61 (0.38, 0.98)	0.04	0.69 (0.49, 0.97)	0.03[b]	0.88
CYP1B1	rs162562	1.75 (1.23, 2.48)	0.002	1.51 (1.11, 2.04)	0.008	0.69	1.21 (0.75, 1.95)	0.44	1.20 (0.79, 1.82)	0.41	0.54
BIRC5	rs1508147	1.76 (1.22, 2.53)	0.002	1.37 (1.07, 1.76)	0.014	0.19	0.93 (0.57, 1.50)	0.75	0.92 (0.65, 1.29)	0.61	0.45
AKAP9	rs6964587	0.58 (0.41, 0.83)	0.003	0.70 (0.55, 0.91)	0.006	0.47	1.46 (0.89, 2.37)	0.13	1.27 (0.91, 1.77)	0.17	0.65
IL12A	**rs2243123**	1.68 (1.18, 2.38)	**0.004**	1.43 (1.09, 1.89)	**0.011**	0.40	1.53 (0.97, 2.43)	**0.07**	1.47 (1.03, 2.12)	**0.04**	**0.62**
BCL2	rs7234941	0.58 (0.40, 0.84)	0.004	0.59 (0.42, 0.84)	0.013	0.04	1.21 (0.72, 2.02)	0.48	1.18 (0.74, 1.88)	0.48	0.37
APAF1	rs1007573	0.53 (0.34, 0.82)	0.004	0.60 (0.40, 0.90)	0.012	0.56	0.89 (0.48, 1.64)	0.70	0.99 (0.55, 1.64)	0.85	0.96
BCL2	rs1982673	0.57 (0.38, 0.85)	0.005	0.59 (0.42, 0.84)	0.003	0.50	Failed genotyping				
DAD1	**rs8005354**	1.65 (1.15, 2.35)	**0.006**	1.31 (1.03, 1.67)	**0.028**	0.04	1.52 (0.96, 2.41)	**0.07**	1.43 (1.02, 2.02)	**0.04**	**0.62**
APAF1	rs2288713	0.55 (0.36, 0.85)	0.007	0.62 (0.42, 0.92)	0.018	0.60	0.90 (0.49, 1.68)	0.75	0.98 (0.55, 1.75)	0.94	0.29
CYP1B1	rs10916	1.62 (1.13, 2.30)	0.008	1.44 (1.06, 1.95)	0.019	0.70	Failed genotyping				
MS4A6A	rs1019670	0.67 (0.47, 0.95)	0.026	0.67 (0.52, 0.87)	0.002	0.72	1.00 (0.62, 1.62)	0.99	0.94 (0.68, 1.31)	0.73	0.47
FRAP1	rs12124983	1.52 (1.07, 2.14)	0.019	1.51 (1.16, 1.97)	0.002	0.10	0.83 (0.53, 1.32)	0.44	0.90 (0.63, 1.27)	0.53	0.67
CASP7	rs4342983	1.80 (1.14, 2.86)	0.013	1.85 (1.19, 2.89)	0.007	0.21	0.95 (0.48, 1.86)	0.87	0.92 (0.49, 1.73)	0.79	0.51
FRAP1	rs1064261	1.42 (1.01, 2.01)	0.05	1.45 (1.11, 1.90)	0.007	0.21	0.84 (0.53, 1.34)	0.46	0.99 (0.71, 1.37)	0.94	0.52
TERT	rs2075786	1.38 (0.97, 1.97)	0.07	1.41 (1.10, 1.80)	0.007	0.73	1.47 (0.93, 2.33)	0.10	1.29 (0.87, 1.74)	0.25	0.68
ADH1C	rs698	1.46 (1.02, 2.08)	0.04	1.38 (1.08, 1.77)	0.01	0.94	1.11 (0.68, 1.79)	0.68	1.14 (0.82, 1.59)	0.45	0.46

CI, confidence interval. SNPs considered to replicate in bold.
[a]Adjusted for age, sex and smoking (discovery), adjusted for age and sex (replication).
[b]Risk association in discovery and replication are in opposite directions, so not reported as replicating.

Table IV. Genotype and allele frequencies of the replicating SNPs in Caucasian neuroendocrine cases and controls

Gene	SNP	Discovery				Replication			
		Case count	Case frequency	Control count	Control frequency	Case count	Case frequency	Control count	Control frequency
IL12A	rs2243123								
	TT	115	46.37%	180	56.78%	113	48.29%	65	58.04%
	CT	113	45.56%	114	35.96%	94	40.17%	40	35.71%
	CC	20	8.06%	23	7.26%	27	11.54%	7	6.25%
	C allele		0.31		0.25		0.32		0.24
DAD1	rs8005354								
	TT	84	43.89%	140	32.56%	93	39.74%	57	50.44%
	CT	129	40.75%	130	50.00%	107	45.73%	46	40.71%
	CC	45	15.36%	49	17.44%	34	14.53%	10	8.85%
	C allele		0.42		0.36		0.37		0.29

risk. We identified potential associations in a discovery case–control set and then sought to replicate these associations in an independent set of cases and controls. Our study evaluated 1334 SNPs derived from 354 genes implicated in cancer-related pathways. We identified two SNPs, in two different genes, associated with overall NET risk in both the discovery and replication sets: *IL12A* rs2243123 and *DAD1* rs8005354. Consistent with these observations, a pathway analysis, performed on the discovery set, implicated apoptosis and inflammation pathways in NET risk.

IL12A rs2243123, located in intron 1, was selected as a tagging SNP for this study and has no documented or predicted function (25,26). However, several other SNPs in *IL12A*, some strongly correlated with rs2243123 (by D′, HapMap CEU, release 27), have been associated with cancer risk such as cervical cancer (27), lung cancer (28) and gastric cancer (29). Studies have reported associations of *IL12A* SNPs [rs755004, rs485497 (D′ = 0.54 and D′ = 0.48, respectively)] in Caucasians for non-Hodgkins lymphomas as well (30,31). A SNP strongly associated with childhood ALL in Caucasians, rs583911, is adjacent to our SNP with D′ = 1 (32). The SNP rs2243123 has been directly associated with an increased risk of primary biliary cirrhosis (33); other SNPs within or immediately 5′ to *IL12A* have been associated with celiac disease [rs17810546 (D′ = 0.87)] (34,35) and with multiple sclerosis (36), identified with corroborating genome wide association studies.

IL12A encodes the p35 protein, which, together with the p40 subunit, forms the IL12 heterodimer (27). Produced by antigen-presenting cells, dendritic cells and macrophages, IL12 activates the T helper 1 (T_h1) response, signaling CD8+ cytotoxic T-cell differentiation and stimulating natural killer cells and T_h1 cells to produce interferon-γ (37). The increased production of interferon gamma also has an anti-angiogenic effect, mediated by the increased production of inducible protein-10 (IP-10 or CXCL10) chemokine, which inhibit endothelial cell chemotaxis and differentiation into tube-like structures (38). IL12 may also inhibit angiogenesis by downregulating vascular endothelial growth factor and basic fibroblast growth factor (38). Therefore, a genetic variant that impairs IL12 function in T-cell development and anti-angiogenesis could be important to the carcinogenesis and progression of NET.

Like *IL12A* rs2243123, *DAD1* rs8005354 was selected for our study as a tagging SNP. It is intronic and has no known or predicted function. However, it is located in a region of high linkage disequilibrium by D′ (HapMap release 27) that extends ~40 kb, including the 27 kb

Table V. Risk associations in NET subgroups (identified in the discovery set with a $P \leq 0.01$)

Pancreatic NET						Small bowel NET					
Gene	SNP	Dominant aOR (95% CI)	Dominant adjusted *P*-value	Additive aOR (95% CI)	Additive adjusted *P*-value	Gene	SNP	Dominant aOR (95% CI)	Dominant adjusted *P*-value	Addititve aOR (95% CI)	Additive adjusted *P*-value
LIG3	rs1052536	0.32 (0.17, 0.60)	0.0003	0.42 (0.26, 0.69)	0.0005	*CYP1B1*	rs162562	2.50 (1.54, 4.06)	0.0002	2.03 (1.37, 3.02)	0.0004
CDKN2A	rs3731198	2.80 (1.50, 5.22)	0.001	2.50 (1.44, 4.35)	0.001	*CFLAR*	rs7573529	0.39 (0.22, 0.68)	0.0008	0.47 (0.29, 0.76)	0.002
CDKN2A	rs3731217	2.79 (1.49, 5.21)	0.001	2.51 (1.44, 4.38)	0.001	*CYP1B1*	rs162557	2.11 (1.30, 3.43)	0.002	1.82 (1.23, 2.68)	0.003
CDKN2A	rs2518719	2.75 (1.47, 5.11)	0.001	2.48 (1.43, 4.30)	0.001	*CYP1B1*	rs10916	2.14 (1.31, 3.52)	0.003	1.90 (1.27, 2.83)	0.002
BCL2	rs7234941	0.32 (0.15, 0.68)	0.003	0.37 (0.19, 0.75)	0.005	*IL1RN*	rs380092	2.13 (1.30, 3.48)	0.003	1.97 (1.39, 2.80)	0.0001
BCL2	rs12957119	0.34 (0.16, 0.70)	0.004	0.44 (0.23, 0.83)	0.011	*ALOX5*	rs3824612	0.48 (0.30, 0.78)	0.003	0.64 (0.44, 0.94)	0.022
ADPRT	rs1136410	2.56 (1.34, 4.87)	0.004	2.17 (1.20, 3.94)	0.010	*DAD1*	rs5742747	0.44 (0.25, 0.77)	0.004	0.54 (0.35, 0.95)	0.030
VEGFR1	rs2387632	0.43 (0.23, 0.79)	0.007	0.45 (0.26, 0.76)	0.003	*PIK3CA*	rs3729692	2.49 (1.33, 4.68)	0.005	2.49 (1.33, 4.68)	0.005
SLC10A2	rs3803258	0.33 (0.15, 0.75)	0.008	0.40 (0.19, 0.81)	0.011	*TNFRSF6*	rs2296600	2.14 (1.27, 3.62)	0.005	1.70 (1.20, 2.41)	0.003
TSC1	rs4962081	2.60 (1.27, 5.32)	0.009	2.79 (1.44, 5.41)	0.002	*LRMP*	rs7969931	2.00 (1.23, 3.25)	0.005	1.52 (1.01, 2.27)	0.044
PERP	rs3734299	2.39 (1.24, 4.58)	0.009	1.63 (1.08, 2.46)	0.018	*BCL2*	rs1982673	0.42 (0.23, 0.78)	0.006	0.45 (0.26, 0.79)	0.006
BCL2	rs7243091	0.41 (0.21, 0.80)	0.009	0.52 (0.30, 0.91)	0.023	*PTGIS*	rs508757	1.96 (1.21, 3.19)	0.007	1.65 (1.09, 2.49)	0.018
ERCC1	rs3212961	2.33 (1.23, 4.42)	0.009	2.10 (1.18, 3.74)	0.012	*TNFRSF10A*	rs11780345	2.02 (1.22, 3.36)	0.007	1.44 (1.02, 2.04)	0.040
CDKN2A	rs3731211	1.96 (1.02, 3.76)	0.042	2.11 (1.28, 3.47)	0.003	*IL1RN*	rs2071459	2.07 (1.22, 3.52)	0.007	2.07 (1.22, 3.52)	0.007
MMP14	rs2236302	2.29 (1.18, 4.44)	0.014	2.59 (1.38, 4.86)	0.003	*ABCC1*	rs246221	1.94 (1.19, 3.14)	0.008	1.46 (1.00, 2.12)	0.049
IFNGR2	rs1059293	2.17 (1.01, 4.51)	0.047	1.84 (1.19, 2.85)	0.006	*BCL2*	rs9807663	2.12 (1.21, 3.72)	0.009	2.06 (1.19, 3.58)	0.010
TNFA	rs1800629	2.26 (1.20, 4.24)	0.011	2.16 (1.23, 3.78)	0.007	*IL1RN*	rs454078	0.53 (0.32, 0.85)	0.009	0.61 (0.41, 0.90)	0.013
CYP24A1	rs3787555	1.73 (0.94, 3.17)	0.077	1.63 (0.98, 2.69)	0.010	*MS4A6A*	rs1019670	0.61 (0.38, 0.98)	0.043	0.59 (0.41, 0.85)	0.005
						IGFBP1	rs1995051	1.73 (1.07, 2.82)	0.027	1.73 (1.07, 2.82)	0.005
						IL17RB	rs1043261	1.72 (1.17, 2.52)	0.015	2.10 (1.15, 3.83)	0.006
						PGR	rs1042839	0.49 (0.27, 0.88)	0.016	0.48 (0.28, 0.83)	0.009
						CHEK2	rs2267130	0.69 (0.42, 1.14)	0.149	0.63 (0.45, 0.90)	0.010

CI, confidence interval.

length of *DAD1* and the adjacent locus control region on chromosome 14q11 (39). Therefore, it is possible that rs8005354 is correlated with a functional variant in this region that is associated with enhanced function or increased expression of DAD1.

The DAD1 protein is a core component of the multisubunit oligosaccharyltransferase that catalyzes *N*-glycosylation of nascent polypeptide chains within the lumen of the endoplasmic reticulum (40). DAD1 has been implicated as a regulator of apoptosis: absence of DAD1 is associated with increased apoptosis in mouse embryos (41).

Overexpression of DAD1 has been associated with poor prognosis in Hodgkins lymphoma and in prostate, small cell lung and hepatocellular carcinomas (41). A potential role for DAD1 in carcinoid tumorigenesis has been suggested by studies of genomic alterations in small bowel carcinoid tumors. These studies have revealed amplification of 14q11, which encompasses the locus control region for *DAD1*, as well as evidence of DAD1 overexpression in carcinoid tumors (41,42). Studies correlating *DAD1* rs8005354 with DAD1 protein expression are warranted and could shed further light on the potential role of this SNP in neuroendocrine tumorigenesis.

Our study has several limitations. Although both cases and controls in the discovery and replication sets were identified from medical centers in the greater Boston area, we observed imbalances in demographic features that could potentially have influenced our risk analysis. In particular, we observed a higher proportion of current or former smokers among the controls in the discovery set and a higher proportion of females among the controls in the replication set. Although an association between smoking and NET risk has not been clearly demonstrated, we adjusted for smoking as well as other potential risk factors (e.g. gender and age) in our regression analysis.

Our study utilized a large collection of cases considering the overall rarity of NETs but was also limited to some extent by small sample size. Our sample size limited our ability to detect specific risk associations in tumor subgroups. The smaller sample size of the replication set (compared with the discovery set) may also have limited our ability to replicate associations overall, particularly after adjusting for multiple testing.

Finally, because our cases included both incident and prevalent cases, a genetic variant associated with a better prognosis could be observed to be more prevalent in cases than in controls, resulting in an apparent but artificial risk association. Although we did not observe obvious associations between *DAD1* rs8005354 or *IL12A* rs2243123 and survival, a low number of death events in our cases at the time of our analysis precluded more definitive evaluation of this possibility.

In summary, we performed a large-scale analysis of genetic risk factors in sporadic NETs. We found that genetic variation *IL12A* and *DAD1* is potentially associated with NET risk, and more broadly, that inflammatory and apoptosis pathways may play a role in neuroendocrine tumorigenesis. Larger studies confirming these associations and more specifically evaluating NET subtypes are warranted.

Supplementary material

Supplementary Table I can be found at http://carcin.oxfordjournals.org/

Funding

Centers for Disease Control and Prevention, The National Institute for Occupational Safety and Health [(T42-OH008416) to D.C.C., M.T.], National Research Service Award [(T32 ES 007069) to M.T.]; John F. and Virginia B.Taplin Scholarship Fund to M.T. and National Institutes of Health [(CA074386, CA119650) to D.C.C.]; this work was also supported by a gift from Saul and Gitta Kurlat and the Dana-Farber/Harvard Cancer Center SPORE in Gastrointestinal Cancer [(P50 CA127003) to M.H.K. and a grant from Novartis Pharmaceuticals to M.H.K.]. The Broad Institute Center for Genotyping and Analysis is supported by grant U54 RR020278 from the National Center for Research Resources.

Acknowledgements

We thank Drs Thomas J.Lynch and John C.Wain and the staff of the MGH Thoracic Oncology Center for participant recruitment and Andrea Shafer and Salvatore Mucci for data collection and management. We thank the team at the Broad Institute Center for Genotyping and Analysis headed by Dr Daniel B.Mirel for their assistance in genotyping. We also thank Alison Brown and Patrice Vamivakas at Partners Healthcare Center for Personalized Medicine for their assistance in genotyping.

Conflict of Interest Statement: None declared.

References

1. Yao,J.C. *et al*. (2008) One hundred years after ''carcinoid'': epidemiology of and prognostic factors for neuroendocrine tumors in 35,825 cases in the United States. *J. Clin. Oncol.*, **26**, 3063–3072.
2. Ferolla,P. *et al*. (2007) Epidemiology of non-gastroenteropancreatic (neuro)endocrine tumours. *Clin. Endocrinol. (Oxf)*, **66**, 1–6.
3. Kaerlev,L. *et al*. (2002) The importance of smoking and medical history for development of small bowel carcinoid tumor: a European population-based case-control study. *Cancer Causes Control*, **13**, 27–34.
4. Hassan,M.M. *et al*. (2008) Risk factors associated with neuroendocrine tumors: a U.S.-based case-control study. *Int. J. Cancer*, **123**, 867–873.
5. Anlauf,M. *et al*. (2007) Hereditary neuroendocrine tumors of the gastroenteropancreatic system. *Virchows Arch.*, **451** (suppl. 1), S29–S38.
6. Toumpanakis,C.G. *et al*. (2008) Molecular genetics of gastroenteropancreatic neuroendocrine tumors. *Am. J. Gastroenterol.*, **103**, 729–732.
7. Hiripi,E. *et al*. (2009) Familial gastrointestinal carcinoid tumours and associated cancers. *Ann. Oncol.*, **20**, 950–954.
8. Jarhult,J. *et al*. (2010) First report on metastasizing small bowel carcinoids in first-degree relatives in three generations. *Neuroendocrinology*, **91**, 318–323.
9. Hassan,M.M. *et al*. (2008) Family history of cancer and associated risk of developing neuroendocrine tumors: a case-control study. *Cancer Epidemiol. Biomarkers Prev.*, **17**, 959–965.
10. Garcia-Closas,M. *et al*. (2007) Large-scale evaluation of candidate genes identifies associations between VEGF polymorphisms and bladder cancer risk. *PLoS Genet.*, **3**, e29.
11. Pharoah,P.D. *et al*. (2007) Association between common variation in 120 candidate genes and breast cancer risk. *PLoS Genet.*, **3**, e42.
12. Hazra,A. *et al*. (2008) Large-scale evaluation of genetic variants in candidate genes for colorectal cancer risk in the Nurses' Health Study and the Health Professionals' Follow-up Study. *Cancer Epidemiol. Biomarkers Prev.*, **17**, 311–319.
13. Hosgood,H.D.III *et al*. (2008) Pathway-based evaluation of 380 candidate genes and lung cancer susceptibility suggests the importance of the cell cycle pathway. *Carcinogenesis*, **29**, 1938–1943.
14. Ferris,B.G. (1978) Epidemiology Standardization Project (American Thoracic Society). *Am. Rev. Respir. Dis.*, **118**, 1–120.
15. Zhou,W. *et al*. (2005) Gene-smoking interaction associations for the ERCC1 polymorphisms in the risk of lung cancer. *Cancer Epidemiol. Biomarkers Prev.*, **14**, 491–496.
16. Miller,D.P. *et al*. (2002) Combinations of the variant genotypes of GSTP1, GSTM1, and p53 are associated with an increased lung cancer risk. *Cancer Res.*, **62**, 2819–2823.
17. Illumina Inc. (2005) *GoldenGate Genotyping Assay Design Tool*. Illumina Inc., San Diego, CA http://www.illumina.com/documents/products/technotes/technote_goldengate_design.pdf. (September 2008, date last accessed).
18. The International HapMap Consortium. (2005) A haplotype map of the human genome. *Nature*, **437**, 1299–1320.
19. Benjamini,Y. *et al*. (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. Ser. B*, **57**, 289–300.
20. Tyrer,J. *et al*. (2006) The admixture maximum likelihood test: a novel experiment-wise test of association between disease and multiple SNPs. *Genet. Epidemiol.*, **30**, 636–643.
21. Maggard,M.A. *et al*. (2004) Updated population-based review of carcinoid tumors. *Ann. Surg.*, **240**, 117–122.
22. Berkovic,M. *et al*. (2006) TNF-alpha promoter single nucleotide polymorphisms in gastroenteropancreatic neuroendocrine tumors. *Neuroendocrinology*, **84**, 346–352.
23. Berkovic,M.C. *et al*. (2007) IL-6-174 C/G polymorphism in the gastroenteropancreatic neuroendocrine tumors (GEP-NETs). *Exp. Mol. Pathol.*, **83**, 474–479.

24. Berkovic,M.C. *et al*. (2010) IL-2-330 T/G SNP and serum values-potential new tumor markers in neuroendocrine tumors of the gastrointestinal tract and pancreas (GEP-NETs). *J. Mol. Med.*, **88**, 423–429.
25. Yuan,H.Y. *et al*. (2006) FASTSNP: an always up-to-date and extendable service for SNP function analysis and prioritization. *Nucleic Acids Res.*, **34**, W635–W641.
26. FastSNP Database. (2006) Institute of Biomedical Sciences and Institute of Information Science. Academia Sinica: Taipei, Taiwan http://fastsnp.ibms.sinica.edu.tw/pages/input_CandidateGeneSearch.jsp. (September 2008, date last accessed).
27. Chen,X. *et al*. (2009) Interactions of IL-12A and IL-12B polymorphisms on the risk of cervical cancer in Chinese women. *Clin. Cancer Res.*, **15**, 400–405.
28. Lee,K.M. *et al*. (2007) Polymorphisms in immunoregulatory genes, smoky coal exposure and lung cancer risk in Xuan Wei, China. *Carcinogenesis*, **28**, 1437–1441.
29. Navaglia,F. *et al*. (2005) Interleukin 12 gene polymorphisms enhance gastric cancer risk in H pylori infected individuals. *J. Med. Genet.*, **42**, 503–510.
30. Butterbach,K. *et al*. (2011) Association of JAK-STAT pathway related genes with lymphoma risk: results of a European case-control study (EpiLymph). *Br. J. Haematol*, **153**, 318–333.
31. Lan,Q. *et al*. (2011) Genetic variation in Th1/Th2 pathway genes and risk of non-Hodgkin lymphoma: a pooled analysis of three population-based case-control studies. *Br. J. Haematol*, **153**, 341–350.
32. Chang,J.S. *et al*. (2010) Genetic polymorphisms in adaptive immunity genes and childhood acute lymphoblastic leukemia. *Cancer Epidemiol. Biomarkers Prev.*, **19**, 2152–2163.
33. Liu,X. *et al*. (2010) Genome-wide meta-analyses identify three loci associated with primary biliary cirrhosis. *Nat. Genet.*, **42**, 658–660.
34. Hunt,K.A. *et al*. (2008) Newly identified genetic risk variants for celiac disease related to the immune response. *Nat. Genet.*, **40**, 395–402.
35. Zhernakova,A. *et al*. (2010) Evolutionary and functional analysis of celiac risk loci reveals SH2B3 as a protective factor against bacterial infection. *Am. J. Hum. Genet.*, **86**, 970–977.
36. International Multiple Sclerosis Genetics Consortium (IMSGC). (2010) IL12A, MPHOSPH9/CDK2AP1 and RGS1 are novel multiple sclerosis susceptibility loci. *Genes Immun.*, **11**, 397–405.
37. Del Vecchio,M. *et al*. (2007) Interleukin-12: biological properties and clinical application. *Clin. Cancer Res.*, **13**, 4677–4685.
38. Duda,D.G. *et al*. (2000) Direct *in vitro* evidence and *in vivo* analysis of the antiangiogenesis effects of interleukin 12. *Cancer Res.*, **60**, 1111–1116.
39. Ortiz,B.D. *et al*. (2001) Function and factor interactions of a locus control region element in the mouse T cell receptor-alpha/Dad1 gene locus. *J. Immunol.*, **167**, 3836–3845.
40. Sanjay,A. *et al*. (1998) DAD1 is required for the function and the structural integrity of the oligosaccharyltransferase complex. *J. Biol. Chem.*, **273**, 26094–26099.
41. Kulke,M.H. *et al*. (2008) High-resolution analysis of genetic alterations in small bowel carcinoid tumors reveals areas of recurrent amplification and loss. *Genes Chromosomes Cancer*, **47**, 591–603.
42. Andersson,E. *et al*. (2009) High-resolution genomic profiling reveals gain of chromosome 14 as a predictor of poor outcome in ileal carcinoids. *Endocr. Relat. Cancer*, **16**, 953–966.

Received January 18, 2011; revised May 4, 2011; accepted May 17, 2011