Germline variation in the insulin-like growth factor pathway and risk of Barrett’s esophagus and esophageal adenocarcinoma


1Department of Cancer Prevention and Control, Roswell Park Comprehensive Cancer Center, Buffalo, New York 14263 USA
2Department of Biostatistics and Bioinformatics, Roswell Park Comprehensive Cancer Center, Buffalo, New York 14263 USA
3Division of Public Health Sciences, Fred Hutchinson Cancer Research Center, Seattle, WA, 98109 USA
4QIMR Berghofer Medical Research Institute, Brisbane, Queensland, Australia
5Department of Biostatistics, University of Washington, School of Public Health, Seattle, Washington, USA.
6Institute of Cancer and Genomic Sciences, University of Birmingham, Birmingham, UK.
7Department of Molecular medicine and Surgery, Karolinska Institutet, Stockholm, Sweden
8Department of Epidemiology, University of North Carolina, Chapel Hill, North Carolina, USA
9Division of Gastroenterology and Hepatology, Mayo Clinic, Rochester, Minnesota, USA
10Department of Epidemiology and Public Health, Queen's University of Belfast, Royal Group of Hospitals, Northern Ireland.
11Pharmacogenomic Epidemiology, Ontario Cancer Institute, Toronto, Ontario, Canada
12Department of Preventive Medicine, University of Southern California/ Norris Comprehensive Cancer Center, Los Angeles, California, USA
13Department of Epidemiology, MD Anderson Cancer Center, Houston, Texas, USA
14Department of Chronic Disease Epidemiology, Yale School of Public Health, New Haven, Connecticut, USA
15School of Cancer and Pharmaceutical Sciences, King’s College London.

© The Author(s) 2020. Published by Oxford University Press. All rights reserved. For Permissions, please email: journals.permissions@oup.com.
Division of Gastroenterology and Hepatology, University of North Carolina School of Medicine, University of North Carolina, Chapel Hill, North Carolina, USA

Department of Population Sciences, Beckman Research Institute and City of Hope Comprehensive Cancer Center, Duarte, California, USA

Division of Research, Kaiser Permanente Northern California, Oakland, California, USA

San Francisco Medical Center, Kaiser Permanente Northern California, San Francisco, California, USA

Oncology Department, University Hospital Antwerp, Edegem, Belgium

Digestive Diseases Centre, University Hospitals of Leicester, Leicester, UK

Department of Oral Biological and Medical Sciences, University of British Columbia, Vancouver, British Columbia, Canada

Farncombe Family Digestive Health Research Institute, Department of Medicine, McMaster University, Hamilton, Ontario, Canada

Department of Upper GI Surgery, Gloucestershire Royal Hospital, Gloucester, UK

Centre for Statistics in Medicine and Oxford Clinical Trials Research Unit, Oxford, UK

University of Cambridge Metabolic Research Laboratories, Wellcome Trust-MRC Institute of Metabolic Science, Addenbrooke’s Hospital, Cambridge, UK

Department of General Surgery, North Tyneside General Hospital, North Shields, UK

Department of Medicine, Institute of Clinical Science, Royal Victoria Hospital, Belfast, UK

Department of Pathology, Leicester Royal Infirmary, Leicester, UK

Department of General, Visceral and Transplantation Surgery, University of Heidelberg, Heidelberg, Germany

Department of General, Visceral and Thorax Surgery, RoMed Klinikum Rosenheim, Rosenheim, Germany

Institute for Urban Public Health, University Hospitals, University of Duisburg-Essen, Essen, Germany.

Department of Gastroenterology, Hepatology and Infectious Diseases, Otto-von-Guericke University Hospital, Magdeburg, Germany

Department of General, Visceral and Transplant Surgery, University Medical Center, University of Mainz, Mainz, Germany

Gastroenterologie am Burgweiher, Bonn, Germany

Institute for Medical Biometry, Informatics, and Epidemiology, University of Bonn, Bonn, Germany

Institute of Pathology, Klinikum Bayreuth, Bayreuth, Germany

Department of Internal Medicine, Evangelisches Krankenhaus, Düsseldorf, Germany

Gastroenterologische Gemeinschaftspraxis, Koblenz, Germany

Department of Visceral, Transplant, Thoracic and Vascular Surgery, University Hospital of Leipzig, Leipzig, Germany

Department of Surgery, Asklepios Harzklinik Goslar, Goslar, Germany
Institute of Human Genetics, Medical Faculty, University of Bonn, Bonn, Germany

General, Visceral and Thoracic Surgery Department and Clinic. University Medical Center Hamburg-Eppendorf. Hamburg. Germany

Department of Internal Medicine II, Frankfurt Hoechst Hospital, Frankfurt, Germany

Department of Interdisciplinary Endoscopy, University Hospital Hamburg-Eppendorf, Hamburg, Germany

Clinic for General, Visceral and Trauma Surgery, Contilia Center for Esophageal Diseases. Elisabeth Hospital Essen, Germany.

Department of Interdisciplinary Endoscopy, University Hospital Hamburg-Eppendorf, Hamburg, Germany

Department of Gastroenterology and Interdisciplinary Endoscopy, Vivantes Wenekebach-Klinikum, Berlin, Germany

Department of Gastroenterology and Interventional Endoscopy, St. John of God Hospital, Regensburg, Germany

Department of Internal Medicine and Gastroenterology, Elisabeth Hospital, Essen, Germany

Department of General, Visceral and Cancer Surgery, University of Cologne, Cologne, Germany

Department of General and Visceral Surgery, Sana Klinikum, Offenbach, Germany

Institute of Pathology, Klinikum Bayreuth, Bayreuth, Germany

Department of Gastroenterology, Oncology and Pneumology, Asklepios Paulinen Klinik, Wiesbaden, Germany

Center for Human Genetics, University Hospital of Marburg, Marburg, Germany

Department of Medicine II, Sana Klinikum, Offenbach, Germany

Cancer Control, QIMR Berghofer Medical Research Institute, Brisbane, Australia.

Division of Medicine Kings Mill Hospital, Sherwood Hospitals NHS Trust, Nottinghamshire

Comprehensive Clinical Trials Unit, University College London.

Center for Human Genetics, University Hospital of Marburg, Marburg, Germany.

Department of Epidemiology, University of Washington, School of Public Health, Seattle, Washington, USA.

Division of Epidemiology, University of Leeds, Leeds, UK.

*Correspondence: matthew.buas@roswellpark.org (M.F.B) and L.J.Hardie@leeds.ac.uk (L.J.H.)
ABSTRACT

Genome-wide association studies (GWAS) of esophageal adenocarcinoma (EAC) and its precursor, Barrett’s esophagus (BE), have uncovered significant genetic components of risk, but most heritability remains unexplained. Targeted assessment of genetic variation in biologically relevant pathways using novel analytical approaches may identify missed susceptibility signals. Central obesity, a key BE/EAC risk factor, is linked to systemic inflammation, altered hormonal signaling, and insulin-like growth factor (IGF) axis dysfunction. Here, we assessed IGF-related genetic variation and risk of BE and EAC. Principal components analysis (PCA) was employed to evaluate pathway-level and gene-level associations with BE/EAC, using genotypes for 270 SNPs in or near 12 IGF-related genes, ascertained from 3295 BE cases, 2515 EAC cases, and 3207 controls in the Barrett’s and Esophageal Adenocarcinoma Consortium (BEACON) GWAS. Gene-level signals were assessed using Multi-marker Analysis of GenoMic Annotation (MAGMA) and SNP summary statistics from BEACON and an expanded GWAS meta-analysis (6167 BE cases, 4112 EAC cases, 17,159 controls). Global variation in the IGF pathway was associated with risk of BE (P=0.0015). Gene-level associations with BE were observed for GHR (growth hormone receptor; p=0.00046, FDR q=0.0056) and IGF1R (IGF1 receptor; p=0.0090, q=0.0542). These gene-level signals remained significant at q<0.1 when assessed using data from the largest available BE/EAC GWAS meta-analysis. No significant associations were observed for EAC. This study represents the most comprehensive evaluation to date of inherited genetic variation in the IGF pathway and BE/EAC risk, providing novel evidence that variation in two genes encoding cell-surface receptors, GHR and IGF1R, may influence risk of BE.
INTRODUCTION

Despite diagnostic and therapeutic advances over recent decades, esophageal adenocarcinoma (EAC) remains one of the most lethal of all cancers, with a median survival of less than one year [1-4]. EAC arises from an epithelial precursor lesion, Barrett’s esophagus (BE). Gastroesophageal reflux, central obesity, tobacco smoking, male sex, and European ancestry are well-established BE/EAC risk factors [5, 6]. Inherited genetics represents another important contributor to disease etiology. In the last eight years, genome-wide association studies (GWAS) have identified nearly 20 susceptibility loci for BE/EAC and revealed substantial heritable components ($h^2$~25-35%) of risk [7-12]. Nevertheless, the majority of heritability remains unexplained, underscoring the need for further discovery efforts. To maximize the value of existing GWAS data, we and others have adopted pathway-level and gene-level approaches to aggregate distributed genetic signals, reduce dimensionality, and boost statistical power to detect further associations. Post-GWAS assessments have identified additional candidate susceptibility genes, such as $MGST1$ and $CDKN2A$, in biologically plausible pathways (e.g., inflammation and tumor suppression) related to BE/EAC pathogenesis [13, 14].

Strong epidemiologic associations between central obesity and risk of BE/EAC have suggested a potential role for metabolic signaling disturbances, such as in the insulin-like growth factor (IGF) axis, in the pathophysiology of BE/EAC [15-19]. Visceral fat is known to affect glucose and lipid metabolism, and alter levels of bioactive molecules and hormones such as insulin, insulin-like growth factors, and pro-inflammatory cytokines [20]. Such hormonal and pro-inflammatory alterations may lead to a dysfunctional IGF system, which has been associated with risk of multiple cancers, including breast, colorectal, prostate, lung and ovary [21-25]. The core IGF pathway comprises growth hormone (GH), the growth hormone receptor (GHR), ligand proteins, insulin-like growth factor 1 and 2 (IGF1, IGF2), insulin-like growth factor 1 receptor (IGF1R), and 6 IGF binding proteins (IGFBP1-IGFBP6) (Figure S1) [26]. The binding of circulating GH to GHR, which is expressed on the cell surface in multiple tissues,
triggers the intracellular synthesis of IGF1, the primary downstream effector. IGF1 in circulation is bound to IGFBPs, which are cleaved at target organ sites, allowing free IGF1 to bind IGF1R and promote tissue growth. Pathway hyper-activation supports risk of transformation and tumorigenesis through promotion of cellular proliferation and angiogenesis, and reduction of apoptosis [26].

Additional lines of evidence have implicated the IGF pathway in esophageal carcinogenesis. Immunohistochemical assessment of patient specimens revealed elevated IGF1R protein expression during BE to EAC progression [27]. Studies from The Cancer Genome Atlas (TCGA) found IGF1R gene amplifications in 10% of EAC tumors analyzed [28]. Upregulation of IGFBP3 has been reported in BE and EAC tissues, relative to normal esophageal epithelium, and IGFBP3 and IGFBP4 overexpression was seen in BE tissues of patients with concurrent EAC, compared to those without cancer [29]. In-vitro studies have demonstrated increased EAC cell-line proliferation in response to IGF1 [30]. While serum IGFBP3 and IGF1 levels were not associated with EAC risk in one BE cohort [31], high serum IGF1 levels have been noted in viscerally obese EAC cases [15, 30], along with increased IGF1R gene and protein expression [15].

A previous study described associations between polymorphisms in GHR and IGF1 and reduced risk of EAC and BE, respectively [32]. However, this study included only a small number of cases (n<500) from a single country (Ireland) and evaluated only ~100 SNPs in seven IGF-related genes. Thus, comprehensive assessment of germline genetic variation in the IGF pathway is needed to understand the extent to which such variation may influence BE/EAC risk and interact with known risk factors. To address this gap, we used complementary statistical approaches to evaluate associations between IGF-related inherited variation and BE/EAC risk, with data from the Barrett’s and Esophageal Adenocarcinoma Consortium (BEACON) GWAS, and further assessed gene-level results in an expanded GWAS meta-analysis dataset.
METHODS

Study populations and SNP genotyping

Detailed descriptions of study participants and GWAS datasets have been published previously [9, 10]. The first phase of our analysis included 2515 EAC, 3295 BE cases and 3207 controls from the BEACON GWAS [10]. All participants were of European ancestry. DNA was isolated from buffy coat or whole-blood and genotyped using the Illumina Omni1M Quad platform. The second phase of our study used data from a larger GWAS meta-analysis (Table S1), which included the BEACON GWAS participants and additional cases and controls from GWAS conducted in Germany and the United Kingdom [9]. All participants were of European ancestry. DNA samples were obtained from blood or saliva and genotyped on high-density Illumina arrays [9]. Informed consent was obtained from all participants in individual studies and every participating institution received ethics approval from their respective Institutional Review Boards (IRBs).

Selection of genes and SNPs

Twelve core IGF-related genes, GH1, GHR, IGF1, IGF2, IGF1R, IGFBP1, IGFBP2, IGFBP3, IGFBP4, IGFBP5, IGFBP6 and IGFALS were selected a priori for analysis. Omni1M SNPs that passed Illumina quality metrics, satisfied additional quality control criteria and had minor allele frequencies (MAFs) ≥1% were eligible for inclusion [10]. Variants selected for analysis are located within hg19 consensus gene boundaries, or within 2.0 kb flanking sequences proximal to the transcriptional start site and distal to the 3’ untranslated region (Table S2) [10]. SHAPEIT was used to impute missing values of genotyped SNPs in the BEACON dataset [33].

Statistical analysis

A principal component analysis (PCA) framework, developed and described previously [13], was applied to examine the association between global genetic variation in the IGF pathway and the risks of BE and EAC, separately, using the BEACON GWAS genotype data. Briefly, a genotype
matrix was constructed using all eligible SNPs assigned to the 12 genes under study. Each SNP genotype was coded 0, 1 or 2, based on the number of designated minor alleles, and standardized across participants to obtain a mean of 0 and SD of 1. The first N principal components (PCs) that captured ≥50% of the genetic variance were selected. A likelihood ratio test statistic was used to assess the association between genetic variation and disease risk. A full model, containing the N pathway-level PCs, age, sex and the first four PCs derived from ancestry-informative markers (PC1_{AIM}-PC4_{AIM}), was compared to a reduced model, containing age, sex and PC1_{AIM}-PC4_{AIM} only. A two-sided pathway-level P value <0.05 was considered statistically significant.

Gene-level analysis was conducted for all genes (n=10) with three or more eligible genotyped SNPs, based on a similar PCA approach [13]. We also employed Multi-marker Analysis of GenoMic Annotation (MAGMA v1.08) as a complementary method for gene-level evaluation, using SNP summary statistics. MAGMA is a fast, flexible and robust tool for analyzing the joint associations of multiple genetic markers simultaneously while accounting for linkage disequilibrium (LD) [34, 35]. At the gene level, correction for multiple comparisons was conducted via the Benjamini-Hochberg false discovery rate (FDR) method [36]. MAGMA was similarly applied to SNP summary statistics derived from the larger GWAS meta-analysis dataset (Tables S1-S2) [9]. The Phase 3 1000 Genomes (EUR) reference dataset (hg19) was used to calculate LD [37]. Meta-analysis SNP-level summary data for the top two genes identified in BEACON were visualized graphically using LocusZoom plots [38] and characterized for functional potential using HaploReg and the Genotype-Tissue Expression (GTEx) project [39, 40].

In exploratory studies, genetic variants in the BEACON dataset were evaluated for gene-environment interactions with BE risk factors, including sex, tobacco smoking, body mass index (BMI), and gastro-esophageal reflux symptoms. We also examined interactions by waist circumference (WC: controls=593, BE cases=1113) and waist-hip ratio (WHR: controls=376, BE
cases=874), in a limited subset of participants with relevant data. Based on guidelines from the World Health Organization and Centers for Disease Control, normal WHR was defined as ≤ 0.85 for women and ≤ 0.90 for men; and normal WC was defined as ≤ 35 inches (women) and ≤ 40 inches (men) [41, 42]. All analyses were conducted using STATA/SE V.15 (College Station, Texas, USA).

RESULTS

Characteristics of study participants
Our primary analysis included data from 3295 BE cases, 2515 EAC cases, and 3207 controls from the BEACON GWAS [10]. Demographic and lifestyle characteristics of participants are presented in Table 1. The majority of case patients were white and older than controls; smoking, reflux symptoms, and use of non-steroidal anti-inflammatory drugs (NSAIDs) were more prevalent among BE and EAC cases versus controls.

Pathway-level associations using BEACON GWAS data
Components of the IGF pathway function biologically together as a unit, exhibiting signal amplification and feedback loops [44]. To assess overall genetic variation in the IGF axis in relation to disease risk, we used a PCA-based method, and identified a significant association with risk of BE (p= 0.0015), but not EAC (p=0.36) (Table 2). This association remained significant after application of the stringent Bonferroni correction accounting for five pathways examined in our previous publication [13] and the IGF pathway in this report (0.05/6=0.0083).

Gene-level associations using BEACON GWAS data
To further localize the pathway-level signal, we applied the PCA method at the gene level for individual IGF axis genes with three or more SNPs. Among 10 eligible genes with ≥3 SNPs, we found a significant association between variation in GHR and risk of BE (p=0.0022, FDR
q=0.022) (Table 3A). As a complementary approach, utilizing SNP summary statistics from the BEACON GWAS, we repeated gene-level analyses using MAGMA [34, 35]. As in the PCA-based assessment, GHR was significantly associated with risk of BE (p=0.00046, q=0.0056). At FDR<0.1, a second significant signal was observed for IGF1R (p=0.0090, q=0.054) (Table 3B), which also ranked second in our PCA analysis (P=0.078). Exploratory examination of gene-level associations with risk of EAC revealed nominally significant signals for IGF1 (p=0.030) and IGFBP6 (0.048) using PCA, but these were lost after correcting for multiple comparisons (FDR=0.24) (Table S3A). Significant (p<0.05) associations with risk of EAC were not observed using MAGMA (Table S3B).

Gene-level associations using GWAS meta-analysis data
To extend our analysis to a larger dataset, we acquired SNP summary statistics from a GWAS meta-analysis (described earlier) [9]. Using MAGMA, we conducted gene-level analyses with summary statistics for all available variants assigned to each of the 12 IGF genes. Consistent with our previous results, at FDR<0.1 we found significant gene-level associations for GHR (p=0.0097, q=0.071) and IGF1R (p=0.0176, q=0.071) in relation to risk of BE; a third signal was identified for IGFBP3 (p=0.0135, q=0.071) (Table 4). Significant (p<0.05) gene-level associations were not observed for EAC (Table S4).

SNP-level associations and in-silico functional characterization
We examined odds ratios (ORs) and 95% confidence intervals (CIs) for the top 10 SNPs in GHR and IGF1R, ranked by GWAS meta-analysis P value (Table S5). The strongest associations had ORs ranging between 0.42-2.69 for GHR and 0.87-1.15 for IGF1R, using a per-allele additive model. LocusZoom plots were assembled to visualize associations of all analyzed meta-analysis SNPs in these genes (Figure 1). At the GHR locus, 257 SNPs were nominally associated with risk of BE (p<0.05). These variants were geographically scattered
across the region, with no visible peaks in signal strength, suggesting that cumulative
associations of multiple weak signals may account for the observed gene-level association
(Figure 1A). The top GHR variants (p<0.01) modify predicted sequence motifs for several
transcription factors, including FOXP1, SOX, STAT, and MEF2 (Table S6A). We also found
evidence of enhancer histone marks in the immediate vicinity of risk-associated SNPs, in the
liver, which is crucial to IGF axis function, particularly IGF1 synthesis. None of the GHR SNPs
remained significant at FDR q<0.05. Of interest, 155 of the 257 SNPs satisfying p<0.05 are
expression quantitative trait loci (eQTLs) for GHR in tissues such as adipose, skeletal muscle,
heart, lung, pancreas, and thyroid, based on data from the Genotype-Tissue Expression Project
(GTEx) [45]. Of the 155, 47 are GHR eQTLs in esophageal muscularis; at each of these eQTLs,
the allele associated with reduced GHR expression is also associated with reduced risk of BE,
in our GWAS data (data not shown). At the IGF1R locus, 144 SNPs in the meta-analysis were
nominally associated with risk of BE (p<0.05), with 11 satisfying FDR q<0.05. The most
significant signals were clustered at the 5’ end of the locus (Figure 1B). Several of these
intrinsic variants modify predicted sequence motifs for several transcription factors, including
FOX, CTCF, Nf-kappaB, STAT, and EWSR1-FL1 (Table S6B). One SNP (rs4305005) exhibits
particularly strong regulatory potential, and is located in a conserved region marked by DNAsel-
hypersensitivity and enhancer histone modifications in multiple tissues including stomach, with
FOXA2/p300 protein occupancy reported in a liver cancer cell line (HepG2). Five variants are
GTEx eQTLs for IGF1R in non-esophageal tissues.

Assessment of gene-environment interactions and mediation relationships
We assessed associations of GHR SNPs and IGF1R SNPs with BE risk in subgroups stratified
by individual BE/EAC risk factors, using BEACON GWAS data. No statistically significant
interactions were noted for sex, smoking history (ever/never or categories of pack-years), BMI
categories (normal, overweight, obese), or history of reflux, after FDR correction across all
comparisons (data not shown). An interaction between one of the top IGF1R risk SNPs (rs1319869) and BMI, however, did satisfy FDR q<0.1 when accounting for 131 IGF1R SNPs tested (P=0.00067, FDR q=0.088). We also evaluated effect modification by two measures of central obesity, WC and WHR, in subsets of participants with the relevant data. While nominally significant (p<0.05) interactions were observed for certain GHR and IGF1R variants with WC/WHR, none remained significant after correction for multiple comparisons (Tables S7-S8).

To investigate obesity as a potential mediator in the association between IGF-related inherited variation and BE risk, we first examined IGF variants (n=270) in relation to obesity status (BMI≥30) among healthy controls in our study sample. Using logistic regression, we identified associations (p<0.05) for 16 variants (1 GHR, 13 IGF1R, 1 IGFBP3 and 1 IGFBP6) (Table S9A). Next, among the subset of participants with available BMI measurements, we assessed these 16 SNPs in relation to BE risk, with or without adjustment for BMI. Six of the 16 variants were associated with BE (p<0.05) in the unadjusted analysis, and resulting odds ratios were highly similar after inclusion of BMI in logistic regression models (Table S9B). At the pathway level, we further evaluated the extent to which obesity-associated variants might contribute to the observed risk association with BE, via a sensitivity analysis. After excluding the 16 obesity-associated SNPs from the pool of 270, we re-ran the pathway-level PCA test and found that the resulting LRT p value (p=0.0020) was only modestly attenuated compared to that reported in Table 2 (p=0.0015). Together, these findings do not exclude a possible mediating role for obesity, but indicate that potential genetic influences on obesity alone are unlikely to account for the relationship between inherited variation in the core IGF axis and BE.
DISCUSSION

Metabolic disturbances such as IGF axis dysfunction are hypothesized to be involved in obesity-related cancers [15, 16, 18]. This study represents the first large consortium-based assessment of germline genetic variation in the IGF pathway and in its component genes with risks of BE and EAC. Through application of complementary statistical approaches, we identified both pathway-level and individual gene-level associations with risk of BE. First, using BEACON GWAS data, we observed a significant association between global variation in the IGF pathway and BE risk. This association remained significant after correcting for the previously examined five inflammatory pathways in our earlier analysis [13]. Subsequent gene-level examination further resolved this pathway-level signal into gene-level associations for GHR and IGF1R. These gene-based associations with BE risk remained significant in an expanded GWAS meta-analysis dataset, with an additional signal for IGFBP3. At each locus, we further identified several non-coding risk-associated variants with regulatory potential.

The GHR gene, located on the short arm of chromosome 5, encodes the growth hormone receptor, a cell-surface receptor found abundantly in liver, muscle, and adipose tissues, among others, in early embryonic life and after birth. The IGF1R gene, located on the long arm of chromosome 15, encodes the insulin-like growth factor 1 receptor, a cell-surface receptor highly expressed across many tissues (including the esophagus) [46]. GHR is located upstream in the IGF pathway, while IGF1R is a downstream effector. GH, synthesized in the pituitary and released into circulation, binds to the extracellular region of GHR and triggers IGF1 production. Secreted IGF1 binds to binding proteins such as IGFBP3, which are cleaved off at target organ sites, eventually allowing free IGF1 to attach to IGF1R to stimulate cell proliferation via RAS-RAF-MAP kinase, and survival and cell cycle progression via PI3K/AKT and mTOR [47]. Visceral adipose tissue releases pro-inflammatory and procoagulant adipocytokines, which can lead to IGF pathway dysfunction and insulin resistance, particularly where the volume of adipose tissue is increased, as in obesity [30]. These changes may increase insulin secretion.
and reduce IGFBP levels, ultimately elevating free IGF1, which has been suggested to foster cell proliferation and contribute to carcinogenesis [18, 30].

Since BE is an established precursor for EAC, factors that alter BE risk may change EAC predisposition. In this report, we found evidence that pathway-level and gene-level IGF-related genetic variation was significantly associated with risk of BE, but not with risk of EAC. There are multiple possible reasons for these findings. First, the sample size available for BE was ~30% larger than that for EAC, resulting in increased statistical power in the analysis of BE versus EAC. Of interest, we did find that EAC odds ratios (ORs) obtained for multiple SNPs mapping to GHR and IGF1R were directionally concordant with (but smaller in magnitude than) ORs obtained in the BE analysis (Table S5B). With a larger sample of EAC cases, more of these weak associations may have reached statistical significance. Second, BE is a heterogeneous condition with respect to its propensity to progress to EAC, likely reflecting a complex interplay between inherited genetics, somatic alterations, and environmental exposures. If a BE susceptibility variant were associated with altered risk of indolent (but not aggressive) forms of BE, the overall risk estimate for EAC could be attenuated in magnitude relative to the risk estimate for BE. Third, past studies have also estimated that common variant heritability for BE ($h^2_g=35\%$) is larger than for EAC ($h^2_g=25\%$) [8], suggesting a more substantial overall role for inherited variation in the etiology of the precursor (BE) versus the cancer (EAC).

Visceral adiposity is a well-established risk factor for BE and EAC. It is noteworthy that IGF axis abnormalities are commonly observed in obese individuals, and those with metabolic syndrome [44, 48, 49], a condition recently linked to an elevated risk of BE (but not EAC) [50]. If IGF-related genetic variation associated with BE risk results in subtle downstream functional alterations to this signaling axis, e.g. through changes in gene expression of cell-surface receptors, such risk could be modified by exposures which themselves interface biologically with the IGF system. While our present analysis did not reveal convincing evidence of gene-environment (GxE) interactions between adiposity measures and IGF-related variants, more
complete ascertainment of WC/WHR is required in larger sample sizes to reach firmer conclusions. While our study does not exclude possible genetic influences on obesity by variants in the IGF pathway, such effects appear unlikely alone to account for the observed relationship between inherited variation in the IGF axis and BE risk. Identification and characterization of GxE interactions, coupled with assessment of the relationships between BE risk SNPs and environmental factors, can provide further insight into potential causal pathways which integrate genetic and non-genetic inputs. It may also be of interest to investigate interactions between IGF-related genetic variation and additional relevant exposures such as nutritional intake and physical activity, as well as studies of the degree to which GERD conveys the increased obesity risk for BE.

Few studies have examined genetic variation in the IGF axis in relation to BE/EAC susceptibility. A case-control study of 102 SNPs in and around seven IGF-related genes—IGF1, IGF2, IGF1R, IGFBP3, GH1, GH2, GHR—using a sample of ~500 BE/EAC cases and ~250 controls [32], found one GHR variant (rs6898743) associated with reduced risk of EAC, and one IGF1 variant (rs6214) associated with reduced risk of BE [32]. Gene-level and pathway-level analyses were not performed. In our larger GWAS meta-analysis dataset, neither rs6898743 nor rs6214 was associated with BE or EAC risk (p>0.2). In another case-control study of 431 participants, obese individuals with the polymorphic A-variant at G1013A in IGF1R were found to have increased risk of BE and EAC [51]. In our GWAS meta-analysis data, this variant (rs2229765) was not associated (p>0.2) with either BE or EAC.

IGF axis gene variants have been associated with risks of several other cancers including lung, breast, colorectal, prostate, and ovary [21-24]. Most notably, a lung cancer GWAS reported 11 risk-associated SNPs in genes encoding components of the GH-IGF pathway, including GHR (rs6183, MAF=0.001, OR = 12.98, p = 0.0019) [52]. Candidate gene studies in the past have also described associations between polymorphisms in IGF component
genes such as *IGF1R, IGF1*, and *IGFBP-3* and risks of breast, prostate and colorectal cancers [53-55]. Of interest, mutations in *GHR* observed in Laron syndrome, leading to GH insensitivity and IGF1 deficiency, have been linked to reduced risks of various cancers [56].

The present study has several strengths. First, our analysis made use of the largest BE/EAC GWAS dataset currently available, providing a sample size ~20-fold higher than that used in previous studies, and thus greater power to detect genetic associations. Second, we employed two complementary statistical methods, PCA and MAGMA, to aggregate signals from multiple variants at the gene and pathway levels and determine joint relationships with BE/EAC risk. In reducing data dimensionality and the number of required statistical comparisons, these methods boost power and help focus association evaluation on blocks of the genome with suspected biological relevance to the disease of interest. Both approaches yielded *GHR* as the top gene-level signal. MAGMA, which can be applied to SNP summary statistics, enabled us to leverage summary data from a recent GWAS meta-analysis and extend our gene-level evaluation to a considerably larger study sample, yielding similar findings.

This study also has certain limitations. First, given the overlap in study participants between the GWAS meta-analysis and the BEACON GWAS, gene-level associations reported here will require formal validation using additional independent datasets. Second, the assessment of gene-environment interactions was limited by a substantial amount of missing or unavailable data for pertinent risk factors such as history of reflux, NSAID use, and obesity, which reduced statistical power and interpretability. Since visceral obesity, in particular, appears linked to risk of BE and EAC, WC/WHR represent the optimal measures for evaluating interactions, rather than BMI, but were only available for small subsets of BEACON participants. Third, in focusing on common genetic polymorphisms in and around 12 core genes in the IGF pathway, our analysis did not account for rare variants or variation within distal regulatory elements relevant to IGF axis function. Fourth, while there is some correlative evidence that *GHR* and *IGF1R* variants may influence mRNA expression levels, experimental studies will be
needed in the future to establish projected causal linkages between such variants and altered gene expression.

In conclusion, using complementary approaches and large well-annotated GWAS datasets, we report novel associations between genetic variation in \( GHR \) and \( IGF1R \) and risk of BE, while no associations with EAC were revealed. Additional studies are required to validate these findings in even larger, more diverse cohorts; comprehensively investigate gene-environment interactions; and determine whether, and to what extent, genetic modulation of IGF pathway component gene expression may contribute to the etiology of BE.
The authors declare no potential conflicts of interest.

Acknowledgments

We thank Drs. Rebecca Fitzgerald, Paul D. Pharoah and Carlos Caldas from the Oesophageal Cancer Clinical and Molecular Stratification (OCCAMS) Consortium for provision of data towards this project. We also thank Terri Watson, Tricia Christopherson, and Paul Hansen for their contributions in project management and organization of biospecimens/data. Genotyping data generated in the BEACON GWAS has been deposited into dbGaP under accession number phs000869.v1.p1. The MD Anderson controls were drawn from dbGaP (study accession: phs000187.v1.p1). Genotyping of these controls was performed through the University of Texas MD Anderson Cancer Center (UTMDACC) and the Johns Hopkins University Center for Inherited Disease Research (CIDR). We acknowledge the principal investigators of this study: Christopher Amos, Qingyi Wei, and Jeffrey E Lee. Controls from the Genome-Wide Association Study of Parkinson Disease were obtained from dbGaP (study accession: phs000196.v2.p1). This work, in part, used data from the National Institute of Neurological Disorders and Stroke (NINDS) dbGaP database from the CIDR:NeuroGenetics Research Consortium Parkinson's disease study. We acknowledge the principal investigators and coinvestigators of this study: Haydeh Payami, John Nutt, Cyrus Zabetian, Stewart Factor, Eric Molho, and Donald Higgins. Controls from the Chronic Renal Insufficiency Cohort (CRIC) were drawn from dbGaP (study accession: phs000524.v1.p1). The CRIC study was performed by the CRIC investigators and supported by the National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK). Data and samples from CRIC reported here were supplied by NIDDK Central Repositories. This report was not prepared in collaboration with investigators of the CRIC study and does not necessarily reflect the opinions or views of the CRIC study, the NIDDK Central Repositories, or the NIDDK. We acknowledge the principal investigators and the
project officer of this study: Harold I Feldman, Raymond R Townsend, Lawrence J Appel, Mahboob Rahman, Akinlolu Ojo, James P Lash, Jiang He, Alan S Go, and John W Kusek.


Financial support: This work was principally supported by the National Institutes of Health [R21DK099804 to M.M.M.; R01CA136725 to T.L.V. and D.C.W.; P30CA016056 to the Roswell Park Comprehensive Cancer Center]. Support for studies related to the Bonn dataset was granted by the Else Kröner Fresenius Stiftung (EKFS) (grant number 2013_A118 awarded to I.G. and J.S.). M.M.N. received support from the Alfried Krupp von Bohlen und Halbach-Stiftung and is a member of the DFG funded Excellence Cluster ImmunoSensation. The Heinz Nixdorf Recall cohort was established with the generous support of the Heinz Nixdorf Foundation, Germany. Additional funding sources for individual studies included in the BEACON GWAS, and for BEACON investigators, have been acknowledged previously. Study sponsors played no role in the study design, collection, analysis, or interpretation of the data, or in the writing of the report.
REFERENCES


Table 1. Characteristics of BEACON GWAS study participants.

<table>
<thead>
<tr>
<th></th>
<th>Controls (n=3207)</th>
<th>BE cases (n=3295)</th>
<th>EAC cases (n=2515)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>%</td>
<td>N</td>
</tr>
<tr>
<td><strong>Age (years)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;50</td>
<td>726</td>
<td>22.6</td>
<td>449</td>
</tr>
<tr>
<td>50-59</td>
<td>885</td>
<td>27.6</td>
<td>780</td>
</tr>
<tr>
<td>60-69</td>
<td>963</td>
<td>30.0</td>
<td>1011</td>
</tr>
<tr>
<td>70+</td>
<td>633</td>
<td>19.8</td>
<td>1048</td>
</tr>
<tr>
<td><strong>Sex</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>880</td>
<td>27.4</td>
<td>806</td>
</tr>
<tr>
<td>Male</td>
<td>2327</td>
<td>72.6</td>
<td>2489</td>
</tr>
<tr>
<td><strong>BMI</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;25</td>
<td>786</td>
<td>36.3</td>
<td>608</td>
</tr>
<tr>
<td>25-29.99</td>
<td>944</td>
<td>43.6</td>
<td>1191</td>
</tr>
<tr>
<td>30-34.99</td>
<td>307</td>
<td>14.2</td>
<td>657</td>
</tr>
<tr>
<td>35+</td>
<td>130</td>
<td>6.0</td>
<td>278</td>
</tr>
<tr>
<td><strong>Smoking status</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>889</td>
<td>40.9</td>
<td>1081</td>
</tr>
<tr>
<td>Yes</td>
<td>1284</td>
<td>59.1</td>
<td>1994</td>
</tr>
<tr>
<td><strong>Smoking pack years</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>None</td>
<td>889</td>
<td>41.3</td>
<td>1081</td>
</tr>
<tr>
<td>&lt;15</td>
<td>889</td>
<td>41.3</td>
<td>1081</td>
</tr>
<tr>
<td>15-29</td>
<td>358</td>
<td>16.6</td>
<td>465</td>
</tr>
<tr>
<td>30-44</td>
<td>326</td>
<td>15.1</td>
<td>348</td>
</tr>
<tr>
<td>45+</td>
<td>273</td>
<td>12.7</td>
<td>279</td>
</tr>
<tr>
<td>309</td>
<td>14.3</td>
<td>306</td>
<td>12.3</td>
</tr>
<tr>
<td><strong>Regular NSAID use</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Never</td>
<td>814</td>
<td>44.0</td>
<td>1050</td>
</tr>
<tr>
<td>Ever</td>
<td>1038</td>
<td>56.0</td>
<td>770</td>
</tr>
<tr>
<td><strong>Weekly reflux/heartburn</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>1448</td>
<td>80.6</td>
<td>1058</td>
</tr>
<tr>
<td>Yes</td>
<td>349</td>
<td>19.4</td>
<td>1186</td>
</tr>
</tbody>
</table>

Numbers may not add to total due to missing data.

BE, Barrett’s esophagus; EAC, esophageal adenocarcinoma; BMI, body mass index; NSAID, non-steroidal anti-inflammatory drugs
Table 2. Pathway level associations with risk of BE and EAC.

<table>
<thead>
<tr>
<th>Pathway</th>
<th>Genes</th>
<th>SNPs&lt;sup&gt;a&lt;/sup&gt;</th>
<th>BE</th>
<th>EAC</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>PCs&lt;sup&gt;b&lt;/sup&gt;</td>
<td>P&lt;sup&gt;c&lt;/sup&gt;</td>
</tr>
<tr>
<td>IGF</td>
<td>12</td>
<td>270</td>
<td>17</td>
<td>0.0015</td>
</tr>
</tbody>
</table>

* a: Number of single-nucleotide polymorphisms (SNPs) selected for analysis
  b: Number of pathway-level principal components included in logistic model
  c: Likelihood ratio p value
Table 3. Assessments of insulin-like growth factor (IGF) gene-level associations with risk of BE using BEACON GWAS data via application of (A) PCA and (B) MAGMA.

A.

<table>
<thead>
<tr>
<th>Gene</th>
<th>Gene name</th>
<th>Chr</th>
<th>SNPs&lt;sup&gt;a&lt;/sup&gt;</th>
<th>PCs&lt;sup&gt;b&lt;/sup&gt;</th>
<th>P&lt;sup&gt;c&lt;/sup&gt;</th>
<th>FDR q&lt;sup&gt;d&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>GHR</td>
<td>5</td>
<td>39</td>
<td>3</td>
<td>0.0022</td>
<td>0.022</td>
</tr>
<tr>
<td>2</td>
<td>IGF1R</td>
<td>15</td>
<td>131</td>
<td>8</td>
<td>0.078</td>
<td>0.39</td>
</tr>
<tr>
<td>3</td>
<td>IGF2</td>
<td>11</td>
<td>25</td>
<td>3</td>
<td>0.16</td>
<td>0.46</td>
</tr>
<tr>
<td>4</td>
<td>IGFBP6</td>
<td>12</td>
<td>6</td>
<td>3</td>
<td>0.18</td>
<td>0.46</td>
</tr>
<tr>
<td>5</td>
<td>IGFBP3</td>
<td>7</td>
<td>5</td>
<td>3</td>
<td>0.24</td>
<td>0.48</td>
</tr>
<tr>
<td>6</td>
<td>IGFBP5</td>
<td>2</td>
<td>16</td>
<td>3</td>
<td>0.32</td>
<td>0.50</td>
</tr>
<tr>
<td>7</td>
<td>IGFBP1</td>
<td>7</td>
<td>6</td>
<td>3</td>
<td>0.35</td>
<td>0.50</td>
</tr>
<tr>
<td>8</td>
<td>IGFBP2</td>
<td>2</td>
<td>8</td>
<td>3</td>
<td>0.47</td>
<td>0.59</td>
</tr>
<tr>
<td>9</td>
<td>IGF1</td>
<td>12</td>
<td>25</td>
<td>3</td>
<td>0.62</td>
<td>0.62</td>
</tr>
<tr>
<td>10</td>
<td>IGFBP acid labile subunit</td>
<td>16</td>
<td>4</td>
<td>3</td>
<td>0.62</td>
<td>0.62</td>
</tr>
</tbody>
</table>

a: Number of SNPs per gene  

b: Number of gene-level PCs included in the logistic regression model  
c: Likelihood ratio p value, d: False discovery rate q value

B.

<table>
<thead>
<tr>
<th>Gene</th>
<th>Chr</th>
<th>SNPs&lt;sup&gt;a&lt;/sup&gt;</th>
<th>P&lt;sup&gt;b&lt;/sup&gt;</th>
<th>FDR q&lt;sup&gt;c&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>5</td>
<td>39</td>
<td><strong>0.00046</strong></td>
<td><strong>0.0056</strong></td>
</tr>
<tr>
<td>2</td>
<td>15</td>
<td>130</td>
<td><strong>0.0090</strong></td>
<td><strong>0.054</strong></td>
</tr>
<tr>
<td>3</td>
<td>2</td>
<td>16</td>
<td>0.12</td>
<td>0.44</td>
</tr>
<tr>
<td>4</td>
<td>11</td>
<td>25</td>
<td>0.21</td>
<td>0.44</td>
</tr>
<tr>
<td>5</td>
<td>7</td>
<td>5</td>
<td>0.21</td>
<td>0.44</td>
</tr>
<tr>
<td>6</td>
<td>17</td>
<td>3</td>
<td>0.23</td>
<td>0.44</td>
</tr>
<tr>
<td>7</td>
<td>12</td>
<td>6</td>
<td>0.26</td>
<td>0.44</td>
</tr>
<tr>
<td>8</td>
<td>17</td>
<td>2</td>
<td>0.40</td>
<td>0.59</td>
</tr>
<tr>
<td>9</td>
<td>16</td>
<td>4</td>
<td>0.51</td>
<td>0.65</td>
</tr>
<tr>
<td>10</td>
<td>7</td>
<td>6</td>
<td>0.54</td>
<td>0.65</td>
</tr>
<tr>
<td>11</td>
<td>2</td>
<td>8</td>
<td>0.69</td>
<td>0.75</td>
</tr>
<tr>
<td>12</td>
<td>12</td>
<td>25</td>
<td>0.77</td>
<td>0.77</td>
</tr>
</tbody>
</table>

a: Number of SNPs per gene, ga012685 not included in 1000G Phase 3 EUR reference data  
b: Gene-level p value, c: False discovery rate q value
Table 4. Assessment of gene-level associations with risk of BE using GWAS meta-analysis data via application of MAGMA.

<table>
<thead>
<tr>
<th>Gene</th>
<th>Chr</th>
<th>SNPs&lt;sup&gt;a&lt;/sup&gt;</th>
<th>P value&lt;sup&gt;b&lt;/sup&gt;</th>
<th>FDR q&lt;sup&gt;c&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>GHR</td>
<td>5</td>
<td>882</td>
<td>0.0097</td>
</tr>
<tr>
<td>2</td>
<td>IGFBP3</td>
<td>7</td>
<td>37</td>
<td>0.0135</td>
</tr>
<tr>
<td>3</td>
<td>IGF1R</td>
<td>15</td>
<td>1226</td>
<td>0.0176</td>
</tr>
<tr>
<td>4</td>
<td>GH1</td>
<td>17</td>
<td>27</td>
<td>0.09</td>
</tr>
<tr>
<td>5</td>
<td>IGFBP1</td>
<td>7</td>
<td>29</td>
<td>0.11</td>
</tr>
<tr>
<td>6</td>
<td>IGFBP4</td>
<td>17</td>
<td>51</td>
<td>0.22</td>
</tr>
<tr>
<td>7</td>
<td>IGF2</td>
<td>11</td>
<td>84</td>
<td>0.48</td>
</tr>
<tr>
<td>8</td>
<td>IGFBP5</td>
<td>2</td>
<td>87</td>
<td>0.48</td>
</tr>
<tr>
<td>9</td>
<td>IGF1</td>
<td>12</td>
<td>223</td>
<td>0.70</td>
</tr>
<tr>
<td>10</td>
<td>IGFBP6</td>
<td>12</td>
<td>18</td>
<td>0.83</td>
</tr>
<tr>
<td>11</td>
<td>IGFALS</td>
<td>16</td>
<td>58</td>
<td>0.87</td>
</tr>
<tr>
<td>12</td>
<td>IGFBP2</td>
<td>2</td>
<td>99</td>
<td>0.88</td>
</tr>
</tbody>
</table>

<sup>a</sup>: Number of SNPs per gene  
<sup>b</sup>: Gene-level p value, <sup>c</sup>: False discovery rate q value
Figure Legends

Figure 1. LocusZoom plots for (A) GHR and (B) IGF1R SNPs in relation to BE risk (meta-analysis).
Figure 1

A.

B.