IAL AT

TRANSLATION

Polymorphisms Near TBX5 and GDF7 Are Associated With Increased Risk for Barrett's Esophagus

Claire Palles,^{1,*} Laura Chegwidden,^{2,*} Xinzhong Li,³ John M. Findlay,¹ Garry Farnham,² Q17 Francesc Castro Giner,¹ Maikel P. Peppelenbosch,⁴ Michal Kovac,¹ Claire L. Adams,¹ Hans Prenen,⁵ Sarah Briggs,¹ Rebecca Harrison,⁶ Scott Sanders,⁷ David MacDonald,⁸ Chris Haigh,⁹ Art Tucker,¹⁰ Sharon Love,¹¹ Manoj Nanji,¹² John deCaestecker,¹³ David Ferry,¹⁴ Barrie Rathbone,¹⁵ Julie Hapeshi,¹⁶ Hugh Barr,¹⁷ Barbara Zietek,¹² Neera Maroo,¹² Laura Gay,¹² Tim Underwood,¹⁸ Lisa Boulter,¹⁸ Hugh McMurtry,¹⁹ David Monk,²⁰ Praful Patel,²¹ Krish Ragunath,²² David Al Dulaimi,²³ Iain Murray,²⁴ Konrad Koss,²⁵ Andrew Veitch,¹⁴ Nigel Trudgill,²⁶ Chuka Nwokolo,²⁷ Bjorn Rembacken,²⁸ Paul Atherfold,²⁹ Elaine Green,³⁰ Yeng Ang,^{31,32} Ernst J. Kuipers,³³ Wu Chow,³⁴ Stuart Paterson,³⁵ Sudarshan Kadri,¹⁵ Ian Beales,³⁶ Charles Grimley,³⁷ Paul Mullins,³⁸ Conrad Beckett,³⁹ Mark Farrant,⁴⁰ Andrew Dixon,⁴¹ Sean Kelly,⁴² Matthew Johnson,⁴³ Shahjehan Wajed,⁴⁴ Anjan Dhar,⁴⁵ Elinor Sawyer,⁴⁶ Rebecca Roylance,⁴⁷ Lynn Onstad,⁴⁸ Marilie D. Gammon,⁴⁹ Douglas A. Corley,⁵⁰ Nicholas J. Shaheen,⁵¹ Nigel C. Bird,⁵² Laura J. Hardie,⁵³ Brian J. Reid,^{48,54} Weimin Ye,⁵⁵ Geoffrey Liu,⁵⁶ Yvonne Romero,^{57,58} Leslie Bernstein,⁵⁹ Anna H. Wu,⁶⁰ Alan G. Casson,⁶¹ Rebecca Fitzgerald,⁶² David C. Whiteman,⁶³ Harvey A. Risch,⁶⁴ David M. Levine,⁶⁵ Tom L. Vaughan,⁴⁸ Auke P. Verhaar,⁴ Jan van den Brande,⁶⁶ Eelke L. Toxopeus,⁶⁷ Manon C. Spaander,⁴ Bas P. L. Wijnhoven,⁶⁷ Luc J. W. van der Laan,⁶⁷ Kausilia Krishnadath,⁶⁸ Cisca Wijmenga,⁶⁹ Gosia Trynka,⁶⁹ Ross McManus,⁷⁰ John V. Reynolds,⁷¹ Jacintha O'Sullivan,⁷¹ Padraic MacMathuna,⁷² Sarah A. McGarrigle,⁷¹ Dermot Kelleher,⁷³ Severine Vermeire,⁵ Isabelle Cleynen,⁵ Raf Bisschops,⁵ Ian Tomlinson,¹ and Janusz Jankowski^{2,12} Q2

¹Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, UK; ²Plymouth University Peninsula School of Medicine and Dentistry, Plymouth, Devon, UK; ³Centre of Biostatistics, Bioinformatics and Biomarkers, Plymouth University Peninsula Schools of Medicine and Dentistry, Plymouth, Devon, UK; ⁴Department of Gastroenterology and Hepatology, Erasmus MC, University Medical Centre Rotterdam, Rotterdam, The Netherlands; ⁵Department of Digestive Oncology, University Hospital Gasthuisberg, Leuven, Belgium; ⁶Department of Pathology, Leicester Royal Infirmary, Leicester, UK; ⁷Department of Cellular Pathology, Warwick Hospital, Warwick, UK; ⁸Department of Oral Biological and Medical Sciences, University of British Columbia, Vancouver, British Columbia, Canada; ⁹Department of Gastroenterology, Wansbeck General Hospital, Ashington, Northumberland, UK; ¹⁰William Harvey Research Institute, The Ernest Cooke Vascular & Microvascular Unit, Centre for Clinical Pharmacology, St Bartholomew's Hospital, London, EC1A 7BE; ¹¹Centre for Statistics in Medicine and Oxford Clinical Trials Research Unit, Oxford, UK; ¹²Centre for Digestive Diseases, Queen Mary University of London, London, UK; ¹³Department of Gastroenterology, Leicester General Hospital, Leicester, UK; ¹⁴Department of Oncology, New Cross Hospital, Royal Wolverhampton Hospitals NHS Trust, Wolverhampton, UK;¹⁵Department for Gastroenterology, Leicester Royal Infirmary, Leicester, UK; ¹⁶Gloucestershire Royal Hospital, Great Western Road, Gloucester, UK; ¹⁷Department of Upper Gl Surgery, Gloucestershire, Royal Hospital, Gloucester, UK; ¹⁸University of Southampton, Southampton General Hospital, Surgery, Gloucestershire, Royal Hospital, Gloucester, UK; ¹⁵University of Soutnampton, Soutnampton General Hospital, Tremona Road, Southampton, UK; ¹⁹Lancashire Teaching Hospitals NHS Foundation Trust, Royal Preston Hospital, Lancashire, UK; ²⁰General Surgery, Countess of Chester Hospital, Chester, UK; ²¹Southampton University Hospitals NHS Trust, Southampton, UK; ²²Wolfson Digestive Diseases Centre, Queens Medical Centre, Nottingham, UK; ²³Worcestershire Acute Hospitals NHS Trust, Alexandra Hospital, Redditch, UK; ²⁴Department of Gastroenterology, Royal Cornwall Hospital, Truro, Cornwall, UK; ²⁵Macclesfield General Hospital, Macclefield, Cheshire, UK; ²⁶Department of Gastroenterology, Sandwell General Hospital, Lyndon, West Bromwich, UK; ²⁷Department of Gastroenterology, University Hospital of Coventry, Coventry, UK; ²⁸Department of Gastroenterology, Leeds General Infirmary, Leeds, UK; ²⁹Department of Clinical Pharmacology University of Oxford, UK; ³⁰School of Biomedical & Healthcare Sciences, Plumouth University Peningula Schools of Medicine and of Oxford, Oxford, UK; ³⁰School of Biomedical & Healthcare Sciences, Plymouth University Peninsula Schools of Medicine and Dentistry, Plymouth, UK; ³¹Gastroenterology, Royal Albert Edward Infirmary NHS Trust, Wigan, UK; ³²GI Science Centre, Salford Royal NHS Foundation Trust, University of Manchester, Salford, UK; ³³Department of Gastroenterology and Hepatology, Department of Internal Medicine, Erasmus MC, University Medical Centre Rotterdam, Rotterdam, The Netherlands; ³⁴ Forth Valley Royal Hospital, Larbert, Scotland, UK; ³⁵Norfolk and Norwich University Hospitals NHS Foundation Trust, Norfolk and Norwich University Hospital, Labert, Scotland, OK, "Norfolk and Norwich Oniversity Hospitals NTS Foundation Gastroenterology, University Hospital of Northern BC, Prince George, British Columbia, Canada; ³⁸Bradford Teaching Hospitals NHS Foundation Trust, Bradford Royal Infirmary, Bradford, UK; ³⁹Royal United Hospital Bath NHS Trust, Royal United Hospital, Avon, Bath, Somerset, UK; ⁴⁰Kettering General Hospital NHS Foundation Trust, Kettering General Hospital, Rothwell Road,

Q1

2 Palles et al

Gastroenterology Vol. ■, No. ■

181

182

183

184

185

186

187

188

189

190

191

192

193

194

195

196

197

198

199

200

201

202

203

204

205

206

207

208

209

210

211

212

213

214

215

216

217

218

219

220

221

222

223

224

225

226

227

228

229

230

231

232

233

234

235

236

240

Kettering, Northants, UK; ⁴¹York Teaching Hospital NHS Foundation Trust, York, UK; ⁴²Luton and Dunstable University 121 Hospital NHS Foundation Trust, Luton, Bedfordshire, UK; ⁴³Department of Thoracic and Upper Gastrointestinal Surgery, Royal Devon and Exeter NHS Foundation Trust, Exeter, UK; ⁴⁴County and Durham and Darlington NHS Foundation Trust, Bishop 122 123 Auckland, County Durham, UK; ⁴⁵Guy's and St Thomas' NHS Foundation Trust, London, UK; ⁴⁶Barts Cancer Institute, Barts and The London School of Medicine and Dentistry, Queen Mary University of London, Charterhouse Square, London, UK; 124 ⁴⁷Division of Public Health Sciences, Fred Hutchinson Cancer Research Centre, Seattle, Washington; ⁴⁸Department of 125 Epidemiology, University of North Carolina School of Public Health, Chapel Hill, North Carolina;⁴⁹Division of Research and San 126 Francisco Medical Center, Kaiser Permanente Northern California, California; ⁵⁰Division of Gastroenterology and Hepatology, 127 UNC School of Medicine, University of North Carolina, Chapel Hill, North Carolina; ⁵¹Department of Oncology, The Medical School, University of Sheffield, UK; ⁵²Division of Epidemiology, University of Leeds, Leeds, UK; ⁵³Division of Human Biology, Fred Hutchinson Cancer Research Center, Seattle, Washington; ⁵⁴Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden; ⁵⁵Princess Margaret Cancer Centre, Ontario Cancer Institute, Toronto, ON, Canada; 128 129 130 ⁵⁶Division of Gastroenterology and Hepatology, Mayo Clinic, Rochester, Minnesota;
 ⁵⁷Registry, On behalf of the Romero;
 ⁵⁸Department of Population Sciences, Beckman Research Institute and City of Hope Comprehensive Cancer Center, Duarte, 131 132 California; ⁵⁹Department of Preventive Medicine, University of Southern California/Norris Comprehensive Cancer Center, Los 133 Angeles, California; 60 Department of Surgery, University of Saskatchewan, Saskatoon, Canada; 61 MRC Cancer Cell Unit, 134 Hutchison-MRC Research Centre and University of Cambridge, Cambridge, UK; 62 Cancer Control, QIMR Berghofer Medical 135 Research Institute, Queensland, Australia; ⁶³Department of Chronic Disease Epidemiology, Yale School of Public Health, New 136 Haven, Connecticut; ⁶⁴Department of Biostatistics, University of Washington School of Public Health, Seattle, Washington; 137 ⁶⁵Department of Gastroenterology, Tergooi Hospital, Hilversum, The Netherlands; ⁶⁶Department of Surgery, Erasmus MC, University Medical Centre Rotterdam, Rotterdam, The Netherlands; ⁶⁷Department of Gastroenterology and Hepatology, 138 Academic Medical Centre, Amsterdam, The Netherlands; ⁶⁸Department of Genetics, University Medical Centre Groningen and University of Groningen, The Netherlands; ⁶⁹Department of Clinical Medicine & Institute of Molecular Medicine, Trinity Centre for Health Sciences, Trinity College Dublin, St James's Hospital, Dublin, Ireland; ⁷⁰Department of Surgery, Trinity Centre for 139 140 141 Health Sciences, Trinity College Dublin, St. James' Hospital, Dublin, Ireland; ⁷¹Gastrointestinal Unit, Mater Misericordiae 142 University Hospital, University College Dublin, Dublin, Ireland; ⁷²Faculty of Medicine, Imperial College London, South 143 Kensington Campus, London, UK 144

BACKGROUND & AIMS: Barrett's esophagus (BE) increases risk for esophageal adenocarcinoma (EAC). Increased risk for BE has been associated with single nucleotide polymorphisms (SNPs) on chromosome 6p21 (within the HLA region) and on basic and Translational at 16q23, where the closest protein-coding gene is *FOXF1*. The Barrett's and Esophageal Adenocarcinoma Consortium (BEA-CON) identified risk loci for BE and esophageal adenocarcinoma in CRTC1 and BARX1, and within 100 kb FOXP1. We aimed to identify SNPs that increased risk for BE in a genome-wide association study (GWAS) and to validate previously reported associations. METHODS: We performed a GWAS to identify variants associated with BE and further analyzed promising variants identified by the BEACON. We performed genotype analysis of 10,158 patients with BE and 21,062 controls. **RESULTS:** We identified 2 SNPs not previously associated with BE: rs3072 (2p24.1; odds ratio [OR] = 1.14; 95% CI: 1.09–1.18; $P = 1.8 \times 10^{-11}$) and rs2701108 (12q24.21; OR = 0.90; 95%) CI: 0.86–0.93; $P = 7.5 \times 10^{-9}$). The closest protein-coding genes were GDF7 (rs3072), which encodes a ligand in the bone morphogenetic protein pathway, and *TBX5* (rs2701108), which encodes a transcription factor that regulates esophageal and cardiac development. We also identified 3 SNPS already identified by the BEACON (rs2687201, rs11789015, and rs10423674). Meta-analysis of all data identified another SNP associated with BE and esophageal adenocarcinoma: rs3784262, near ALDH1A2 (OR = 0.90; 95% CI: 0.87-0.93; $P = 3.72 \times 10^{-9}$). **CONCLUSIONS:** We identified 2 loci associated with risk for BE and provide data to support a locus previously associated with risk in the BEACON. The genes we found to be associated with risk for BE encode transcription factors involved in thoracic, diaphragmatic, and esophageal development or proteins involved in the inflammatory response.

Darrett's esophagus (BE) is a common premalignant **B** condition that affects up to 2% of the adult population in the Western world.¹ BE comprises the second stage in the esophagitis-metaplasia-dysplasia-adenocarcinoma sequence. BE confers a 2%-4% lifetime risk of esophageal adenocarcinoma (EAC).¹ Chronic gastric acid reflux is the predominant etiologic factor for BE. In addition, BE cooccurs with conditions such as intestinal metaplasia, hiatal hernia, obesity, and hypercholesterolemia.²⁻⁵ Several factors, including the degree of acid reflux, hiatal hernia size, and the percentage of intestinal metaplasia-positive glands, can affect the progression to cancer. A role for genetics in the pathogenesis of gastroesophageal reflux disease, including BE and EAC, has been implicated on the basis of 3 observations: concordance in monozygous and dizygous twins $^{6-8}$; the increased risk of disease in those with a positive family history^{9,10}; and, recently, the identification of single nucleotide polymorphisms (SNPs) associated with BE in Genome-Wide Association Studies (GWAS).^{11,12} The proportion of variation in BE risk explained by common variants has been estimated to be 35%.¹³

Our GWAS previously identified 2 SNPs, on chromosomes 6p21 (rs9257809; $P = 4.1 \times 10^{-9}$) and 16q24

*Authors share co-first authorship.

Keywords: EAC; Intestinal Metaplasia; Susceptibility; Cancer.

© 2015 by the AGA Institute 237 0016-5085/\$36.00 238 http://dx.doi.org/10.1053/j.gastro.2014.10.041 239

145

146

156

157

158

159

160

161

162

163

164

165

166

167

168

169

170

171

172

173

174

175

176

177

Abbreviations used in this paper: ASE, allele-specific expression; BE, Barrett's esophagus; BEACON, Barrett's and Esophageal Adenocarcinoma Consortium; CI, confidence interval; EAC, esophageal carcinoma; eQTL, expression quantitative trait locus; GWAS, genome-wide association study; LD, linkage disequilibrium; OR, odds ratio; PC, principal component; SNP, single nucleotide polymorphism; TCGA, The Cancer Genome Atlas.

Q8

(rs9936833; $P = 2.7 \times 10^{-10}$), that are associated with BE.¹¹ One of these loci lies within the HLA region and the other close to *FOXF1*, which is involved in esophageal structure and development. Both SNPs have also been shown to be associated with risk for EAC.¹⁴ More recently, in a combined analysis of BE and EAC cases, the Barrett's and Esophageal Adenocarcinoma Consortium (BEACON) identified susceptibility SNPs near *CRTC1*, *BARX1*, and *FOXP1*.¹² The last 2 of these genes are known to be involved in esophageal development.^{15,16}

We aimed to identify further BE predisposition SNPs from our GWAS by performing wider and deeper independent replication of existing SNPs that had promising disease associations.

Methods

Patient and Sample Collection Criteria

Figure 1 outlines this study and the numbers of samples that contributed to each phase. As described previously,¹¹ Discovery

Phase cases were diagnosed with histologically confirmed BE and ascertained through the UK-based Aspirin Esomeprazole Che-moprevention Trial (AspECT) in Barrett's Metaplasia, a clinical trial of proton-pump inhibitor (esomeprazole) and aspirin as preventive agents for progression of BE to EAC.¹⁷ Replication Phase UK, Irish, Dutch, and Belgian patient samples were obtained from the Chemoprevention Of Premalignant Intestinal Neoplasia (ChOPIN) genetic study and the Esophageal Adenocarcinoma GenEtics (EAGLE) consortium.¹ Replication Phase patients were diagnosed with BE with lengths of >1 cm (C1M1) circumferential disease or ≥ 2 cm tongue patterns (COM2), according to the Pra-gue criteria.¹⁸ Patient collection was in accordance with British Society of Gastroenterology criteria for BE¹⁹ and followed verification of endoscopic findings and proven BE on histopathologic records. Presence of EAC at presentation or subsequently was recorded, but was not an inclusion criterion.

Sample Sets

Discovery Phase. BE patients (n = 1852) were UK Q^9 participants in the AspECT study (Clinical Investigator:



 and Levine et al SNP: rs3784262. ^Δ7 SNPs: Our SNPs: rs3072, rs6751791, rs2731672, rs2701108, rs189247, rs2043633.

4 Palles et al

361

362

363

364

365

366

367

368

369

370

371

372

373

374

375

376

377

378

379

380

381

382

383

384

385

386

BASIC AND

Gastroenterology Vol. ■, No. ■

Jankowski), HANDEL study (Clinical Investigator: Jankowski), ChOPIN study (Clinical Investigator: Jankowski), and population controls of white Caucasian origin (n = 5172) were from the common Wellcome Trust Case Control Consortium 2 (WTCCC2) set.¹¹

Replication Phase 1. UK Replication 1 totaled 1105 BE patients from ChOPIN and 6819 controls. The controls comprised People of the British Isles (Clinical Investigator: W. Bodmer) (n = 2578) and WTCCC2 (Clinical Investigator: P. Donnelly) samples (n = 4241) that were not genotyped in the Discovery Phase. The Dutch replication samples consisted of 473 BE patients and 1780 controls from the University Medical Centre, Groningen. An additional 64 Dutch cases and 206 controls, provided since 2012 from Nijmegen and Rotterdam as part of EAGLE, were genotyped for the 7 SNPs taken into Replication Phase 3.

Replication Phase 2. UK Replication 2 comprised 1765 BE patients from the ChOPIN study. Controls (n = 1586) were from the Colorectal Tumour Gene Identification (CoRGI) Consortium²⁰ (Clinical Investigator: Tomlinson), comprising spouses or partners unaffected by cancer and without a family history (to 2nd-degree relative level) of colorectal neoplasia. All were of white UK ethnic origin. The Irish replication samples were 245 BE patients and 473 controls of white Caucasian origin from St James's Hospital and Mater Misericordiae University Hospital, Dublin. Healthy donor controls were provided by Trinity Biobank.

Replication Phase 3. UK Replication 3 comprised 997 BE patients from the ChOPIN study and 974 female controls from the Genetics of Lobular Carcinoma In Situ in Europe (GLACIER) study (Clinical Investigators: Sawyer, Roylance) with no personal or family history of breast cancer and of white Caucasian origin.²¹ The Belgian replication samples consisted of 362 cases and 848 controls from Leuven. Finally, 3295 BE patients and 3204 controls predominantly of northern European descent from the BEACON consortium GWAS (Clinical Investigators: Vaughan, Whiteman, Levine) were included.¹² All studies received ethical board approval (details in Supplementary Material). Two SNPs described in Su et al¹¹ had been genotyped previously in Replication Phase 2 and BEA-CON/BEAGESS samples. All other Replication Phase 3 samples were new to this study.

Genotyping

For all samples, genomic DNA was extracted from peripheral blood. Various genotyping methods were used, depending on the phase of the study and on pre-existing data from some sample sets. In brief, Discovery Phase genotyping was performed using the Illumina 660W-Quad array for cases and a custom Human 1.2M-Duo array for controls at the Wellcome Trust Sanger Institute.¹¹ Replication Phase 1 genotyping was performed using the Illumina Immunochip at the Wellcome Trust Sanger Institute¹¹ or as described in Trynka et al.²²

In Replication Phase 2, samples underwent custom genotyping for SNPs that met one of the following criteria: $P_{association} < 10^{-4}$ in combined Discovery and Replication Phase 1 analysis (n = 63); $P_{association} < 10^{-4}$ in Discovery Phase, but not included in Replication Phase 1 (n = 12); and $P_{association} < 10^{-4}$ in a sex-stratified analysis of the Discovery phase (n = 5); and candidate polymorphisms previously

reported as associated with BE and not well tagged by the Discovery Phase or Immunochip arrays, specifically, MSR1 p.Arg293Gly,²³ and variants in IGF1R and GHR²⁴ (Supplementary Table 1). Sequenom iPLEX assays were successfully designed for 65 of these SNPs and genotyping was performed at the Wellcome Trust Sanger Institute. Genotypes were assigned using MassArray TyperAnalyzer 4.0 (Sequenom). Samples with sex discrepancies between manifests or with overall call rates <95% were excluded, as were SNPs with call rates of <95%. Where a SNP was in the top 40 of the prioritized SNPs (by P value) and had failed at the design stage of the iPLEX (n = 3), genotyping was performed by KASPar in the full Replication Phase 2 sample set. Seventy-seven of two hundred and forty-five Irish cases passing quality control were genotyped using iPLEX assays, and the other cases were genotyped using the Immunochip. Eighteen iPLEX SNPs were not analyzed in the Irish cohort, as the SNPs were not present on the Immunochip. The samples were also genotyped for 4 SNPs after publication of Levine et al.¹²

Replication Phase 3 samples (Supplementary Table 2) were genotyped using KASPar for 7 SNPs prioritized after analysis of the previous phases. The samples were also genotyped for 4 SNPs after publication of Levine et al.¹² All SNPs had call rates >95%. Sample exclusions were as for Replication Phase 2.

We had previously demonstrated >99% concordance between genome-wide array, Immunochip, and KASPar assays for other SNPs.¹¹ Sequenom call rate was >96%. For samples analyzed only by KASPar, genotyping QC was tested using duplicate DNA samples within studies and SNP assays, together with direct sequencing of subsets of samples to confirm genotyping accuracy. For all SNPs, >98% concordant results were obtained.

Association Analysis

Case-control analysis was performed using frequentist tests under a missing data logistic regression model, as implemented in SNPTEST (version 2.4.1). Principal component analysis was performed for all samples typed on GWAS arrays (Discovery Phase) and has been described in Su et al.¹¹ As described previously, principal component 1 (PC1) was included as a covariate in all analyses of the Discovery Phase. Each SNP was tested as a quantitative explanatory variable, coded as 0, 1, 2. We used GWAMA (version 2.1) to implement fixed inverse variance-based methods for meta-analysis.²⁵ The software tests for heterogeneity of effects between studies²⁶ and enables sexspecific meta-analysis.

Replication and Validation of Single Nucleotide Polymorphisms From BEACON/BEAGESS Meta-analysis

In order to examine the 4 genome-wide significant BE+EAC SNPs and 83 other SNPs with $P_{assoc} < 10^{-4}$ in the BEACON/ BEAGESS data, we performed association testing using AspECT Discovery Phase cases. Because our Discovery Phase controls overlapped with those used by Levine et al,¹² we used 1898 white European controls from colorectal cancer GWAS studies CoRGI and Colon Cancer Family Registry.²⁷ Genotypes were imputed where necessary, with strict cut-offs for imputation quality.^{28,29}

480

421

422

423

424

425

426

427

428

429

430

431

432

433

434

435

436

437

438

439

440

441

396

397

404

405

406

407

408

409

410

411

412

413

414

415

416

Other Analyses

Details of imputation, fine mapping, pathway analyses, estimation of heritability, and URLs are provided in the Supplementary Material.

Results

Identification of Two New Barrett's Esophagus Predisposition Single Nucleotide Polymorphisms

In order to identify further loci associated with BE, we 491 prioritized 65 SNPs (Supplementary Table 1) with the best 492 evidence of association with BE from our previous GWAS 493 Discovery Phase and Replication Phase 1 (details in 494 Methods).¹¹ These SNPs were genotyped in an additional 495 1765 cases and 1586 controls from the UK and in an Irish 496 cohort of 245 cases and 473 controls (Replication Phase 2, 497 previously used to genotype rs9257809 and rs9936833 498 described in Su et al¹¹). After meta-analysis of these new 499 data together with Discovery Phase and Replication Phase 1, 500 seven SNPs showing evidence of associations with BE risk at 501 $P_{\rm meta} < 5 \times 10^{-6}$ were identified and genotyped in Repli-502 cation Phase 3 samples (Table 1). After Replication Phase 3, 503 two SNPs-rs3072 and rs2701108 on chromosome 2p24 504 and 12q24, respectively-reached the level of significance 505 conventionally used for GWAS ($P = 5 \times 10^{-8}$) (Table 1). 506 Combined P_{meta} values were 1.8×10^{-11} for rs3072 (OR = 507 1.14; 95% CI: 1.09–1.18) and 7.5 \times 10⁻⁹ for rs2701108 508 (OR = 0.90; 95% CI: 0.86-0.93), derived from a total sample 509 of 10,158 BE cases and 21,062 controls (Supplementary 510 Table 3). The associations remained at or near genome-511 wide significance upon restricting the analysis to the 8521 512 cases with histologically proven intestinal metaplasia 513 (rs3072: $P = 1.3 \times 10^{-9}$; OR = 1.13; 95% CI: 1.09-1.17; 514 rs2701108: $P = 6.2 \times 10^{-8}$; OR = 0.90; 95% CI: 0.86-0.94). 515 There was no evidence of sex heterogeneity for either SNP 516 (Supplementary Table 4). 517

In Silico Fine Mapping and Annotation of the 518 Chromosome 2p24 and 12q24 Loci. rs3072 lies be-519 tween 2 genes, mapping 7.5 kb downstream of GDF7 (also 520 known as BMP12) and 6.5 kb downstream of C2orf43 521 (Figure 2). rs2701108 is 117 kb downstream of TBX5 and 522 270 kb upstream of RBM19. We imputed in our Discovery 523 Phase all SNPs in 1-Mb regions flanking each of the lead 524 SNPs. At chromosome 2p24, rs3072 remained the most 525 strongly associated SNP, but at chromosome 12q24, 526 rs1920562 was more strongly associated with disease risk 527 ($P_{Discovery} = 1.4 \times 10^{-5}$; OR = 0.84) than the lead genotyped 528 SNP ($P_{Discovery} = 1.4 \times 10^{-3}$; OR = 0.88). rs1920562 529 (linkage disequilibrium [LD] with rs2701108; $r^2 = 0.6$) lies 530 131 kb downstream of TBX5 and 256 kb upstream of 531 RBM19. Nonsynonymous SNPs in the genes flanking the 532 signals on chromosomes 2 and 12 were not in strong LD (r^2 533 < 0.4; D' < 0.8) with the lead genotyped or imputed SNPs, 534 suggesting that the functional variants may have effects on 535 gene expression and regulation rather than protein 536 sequence. Haploregv2³⁰ and Annovar³¹ were used to 537 annotate SNPs in strong LD ($r^2 > 0.4$) with the 2 lead 538 tagging SNPs. 539

488

489

490

481

482

483

484

which mark enhancers (data from lymphoblastoid cell line GM12878). Three other SNPs in LD with rs3072 map to the enhancer region detected in GM12878. One of these, rs7255, maps to a site of high evolutionary conservation/constraint; another SNP, rs9306894, whilst not at a conserved site (Supplementary Table 5), is predicted as "likely to affect

5

541

542

543

544

545

546

547

548

549

550

551

552

553

554

555

556

557

558

559

576

596

597

598

599

600

protein binding and linked to expression of a gene target"

Barrett's Esophagus SNPs Near TBX5 and GDF7

rs3072, which may alter a GATA binding motif, lies

within a region of histone modifications, such as H3K4Me1,

according to RegulomeDB.³² We examined associations between SNPs in this region and gene expression in The Cancer Genome Atlas (TCGA) EAC data.³³ Genotypes were only available for rs9306894 in the chromosome 2 locus and gene expression data had been obtained using RNASeq. After correcting for copy number, we determined associations between rs9306894 genotype and total RNA levels for expression quantitative trait locus (eQTL) analysis and bias in allelic expression of coding SNPs (allele-specific expression [ASE]/ASE analysis). There was no significant association with expression of the closest genes GDF7, HS1BP3 and C2orf43 ($P_{eOTL} > .20$; $P_{ASE} > .38$; n = 62) and no genomewide association with expression of any other gene was present (q > .05, details not shown). In public data sets based on monocytes³⁴ and on lymphoblastoid cell lines and adipose tissue,³⁵ C2orf43 is the suggested target of rs9306894 following eQTL studies (GENevar; $P = 7 \times$ 10^{-4}). rs9306894 genotype was not associated with *GDF7* expression in these cell types.

rs2701108 itself is not likely to be a functionally regulatory SNP, but rs1920562, which showed the strongest regional association after imputation, is a more promising candidate (Supplementary Table 6). This SNP maps to a highly conserved base and a region containing enhancer marks in human embryonic stem cells (h1-ESC) and lung fibroblasts (NHLF). rs1920562 and an additional SNP (rs1247938) in moderate LD ($r^2 = 0.52$) with rs2701108, are highlighted by Regulome DB as being the most likely SNPs in this region to affect protein binding. CTCF and RAD21 binding are predicted to be affected by rs1247938 and the ability of IKZF1 to bind is predicted to be altered by rs1920562. Expression analyses were performed for the rs2701108 region, in the same way as for rs9306894. However, none of the three rs2701108 region SNPs was associated with TBX5, TBX3 or RBM19 expression in the TCGA data ($P_{eOTL} > .39$; $P_{ASE} > .43$; n = 62), was an eQTL in whole-transcriptome analysis, or was an eQTL in the public databases (details not shown).

Pathway/Geneset Enrichment Analysis

Improved Gene Set Enrichment Analysis for Genome Wide Association Study (iGSEA4GWAS) and SNP ratio test respectively found 26 and 34 pathways significantly enriched in cases at False Discovery Rate–corrected P < .05. Genetic Genomics Analysis of Complex Data (Gengen) did not identify any pathways with corrected P < .05, but 10 pathways had P < .25 (Supplementary Table 7). Three pathways (type 1 diabetes mellitus, KEGG antigen processing and presentation, and KEGG autoimmune thyroid

RTIC

 Table 1. Meta-analysis of Discovery and Replication Phase Sample Sets for SNPs Taken Into Replication Phase 3

SNP	Chr	Position (build 37)	Minor/ major	Discovery MAF (cases /controls)	Discovery	Replication phase 1 (rep 1)	Meta 1 (discovery + rep 1)	Replication phase 2 (rep 2)	Meta 2 (meta 1+ rep 2)	Replication phase 3 (rep 3)	Final meta (meta 2 + rep 3)	Final meta I ²	n
rs3072	2	20878406	G/A	0.41/0.36	1.23 (1.14–1.33) 2.64 × 10 ⁻⁷	1.10 (1.02–1.19) 1.33 × 10 ⁻²	1.16 (1.10–1.23) 8.13 × 10 ⁻⁸	1.13 (1.03–1.24) 7.00 × 10 ⁻³	1.16 (1.10–1.21) 2.27 × 10 ⁻⁹	1.11 (1.04–1.17) 1.02 × 10 ⁻³	1.14 (1.09–1.18) 1.75 × 10 ^{–11}	0.42	8 Q14
rs6751791	2	35581997	A/G	0.51/0.48	1.15 (1.06–1.23) 5.03 × 10 ⁻⁴	1.16 (1.07 - 1.25) 1.45 \times 10 ⁻⁴	1.15 (1.09-1.21) 2.68 \times 10 ⁻⁷	1.07 (0.97 - 1.16) 1.64 × 10 ⁻¹	1.13 (1.08 - 1.18) 2.91 × 10^{-7}	0.99 (0.93-1.05) 7.99 × 10 ⁻¹	1.08 (1.04 - 1.12) 7.65 $\times 10^{-5}$	0.60	8
rs2731672	5	176842474	A/G	0.27/0.24	1.18 (1.09 - 1.28) 1.64 \times 10 ⁻⁴	1.14 (1.04 - 1.24) 3.87 \times 10 ⁻³	1.16 (1.09 - 1.23) 2.54×10^{-6}	1.09 (0.99 - 1.21) 8.20 \times 10 ⁻²	1.14 (1.08 - 1.20) 8.33 \times 10 ⁻⁷	0.95 (0.89-1.02) 1.81 × 10 ⁻¹	1.07 (1.03 - 1.12) 1.66 $\times 10^{-3}$	0.63	8
rs2701108	12	114674261	G/A	0.38/0.41	0.88 (0.81–0.95) 1.00 × 10 ⁻³	0.87 (0.81–0.94) 4.40×10^{-4}	0.87 (0.83–0.92) 1.51×10^{-6}	0.89 (0.81-0.97) 1.10 × 10 ⁻²	0.88 (0.84–0.92) 5.68×10^{-8}	0.93 (0.87–0.99) 1.42×10^{-2}	0.90 (0.86–0.93) 7.48 \times 10 ⁻⁹	0.14	8
rs189247	15	97586630	A/G	0.41/0.37	1.18 (1.09 - 1.27) 5.67 \times 10 ⁻⁵	1.14 (1.05 - 1.23) 1.25 \times 10 ⁻³	1.15 (1.09 - 1.22) 2.91 \times 10 ⁻⁷	1.10 (1.00 - 1.20) 4.90×10^{-2}	1.14 (1.09 - 1.19) 6.12×10^{-8}	0.96 (0.90 - 1.02) 1.73 × 10 ⁻¹	1.10 (1.06 - 1.14) 3.55 $\times 10^{-7}$	0.20	8
rs2043633	16	5819274	C/A	0.37/0.41	0.85 (0.79-0.92) 6.04×10^{-5}	0.88 (0.82-0.95) 9.80 × 10 ⁻⁴	0.87 (0.82-0.92) 2.49 × 10 ⁻⁷	0.88 (0.80-0.96) 5.00 × 10 ⁻³	0.87 (0.83-0.91) 4.74×10^{-9}	0.99 (0.94 - 1.05) 8.39×10^{-1}	0.92 (0.88-0.95) 2.25 × 10 ⁻⁶	0.58	8
rs12985909	19	18439383	G/A	0.48/0.45	1.12 (1.04 – 1.21) 2.94 × 10 ⁻³	1.12 (1.04–1.21) 2.73 × 10 ⁻³	1.12 (1.06-1.18) 2.45 \times 10 ⁻⁵	1.11 (1.01 – 1.21) 2.70 × 10 ⁻²	1.12 (1.07 – 1.17) 1.99 × 10 ⁻⁶	$1.07(1.01 - 1.13)2.63 \times 10^{-2}$	1.10 (1.06 – 1.14) 3.28 × 10 ⁻⁷	0.00	8

gna jisab Ta janoitajenart

NOTE. For each phase, association data show (top to bottom) OR, (95% Cl), and P_{assoc} . Results are presented with respect to the minor allele. rs3072 and rs2701108 reached genome-wide significance. In BEACON, rs7598399 was used as a proxy for rs6751791 ($r^2 = 1$) and rs189247 was imputed from 4 genotyped SNPs (rs991757, rs2670927, rs2670930, and rs234540) with accuracy approximately 98%. The Dutch extension samples were analyzed with the previously described Dutch replication samples as part of Rep 1. The *P* value threshold for including a SNP in Phase 2 was 1 × 10⁻⁴ and that for inclusion in Phase 3 was 5 × 10⁻⁶. Chr, chromosome; I², I² heterogeneity index; MAF, minor allele frequency; N, number of studies.





Figure 2. Regional plots of association (*left y-axis*) and recombination rates (*right y-axis*) for the chromosomes 2p24 and 12q24 loci after imputation. The lead genotyped SNP is marked with a *purple square*. Imputed SNPs are plotted as *circles* and genotyped SNPs as *squares*.

disease) were identified by all methods and the SNPs mapping to each pathway were subjected to set-based association tests using PLINK, producing empirical P values of .0021, .025, and .0317, respectively. The SNPs within these pathways that showed replication at P < .05 in Replication Phase 1 all mapped to chromosome 6p21, either close to or within HLA genes. Upon removal of the HLA genes from all the pathways, 20/26 pathways originally with False Discovery Rate P < .05 according to the iGSEA4GWAS approach remained significant, but only 1/34 pathways identified by SNP ratio test remained significant. No HLAdepleted pathways were even suggestive of enrichment (all P > .25) by Gengen. The top networks identified by Ingenuity Pathway Analysis were Cardiovascular System Development and Function, Embryonic Development and Organ Development (Supplementary Figure 1). The 5 genes implicated by the BE susceptibility SNPs here are all involved in development at a cellular, embryonic, organ, and organism level. Bone morphogenetic protein 4 was the most significant upstream regulator ($P_{overlap}$ = 1.99×10^{-6}).

Barrett's Esophagus Heritability

Genome-wide haplotype-tagging SNP data on the 1852 cases and 5172 controls in the Discovery Phase of this study were used in Genome-wide Complex Trait Analysis to estimate the proportion of variation in risk of BE that can be explained by common genetic variants. In line with our previous, disease score test analysis,¹¹ we found a statistically significant component of BE risk to be polygenic (9.99% [SE 1.2%]). This was a lower estimate than that recently derived by Ek et al.¹³

Replication Testing of Previously Reported Barrett's Esophagus Susceptibility Single Nucleotide Polymorphisms at Candidate Loci

Using a systematic review, we identified 26 polymorphisms reported in the literature to be associated with BE (Supplementary Table 8). In our Discovery Phase samples, 20 of 26 SNPs were directly genotyped or were in strong LD $(r^2 > 0.7)$ with a directly genotyped SNP. Only one of these SNPs showed a nominally significant association (P < .05) in our data (rs909253, proxy for rs1041981, $r^2 = 0.93$; OR = 1.12; P = .005). This SNP was also present on the Immunochip and showed additional evidence of replication $(P_{\text{meta}} = 3.1 \times 10^{-4}, \text{ OR} = 1.07)$. rs909253 maps to a highly conserved base (based on SiPhy score) in an intron of LTA (tumor necrosis factor- β) where histone marks associated with both promoters and enhancers are present in LCLs.³⁰Q11 PBX3, PU1, POL2, YY1, and nuclear factor- κ B have all been found to bind here in ENCODE ChIP-seq experiments. No data were available for this SNP in RegulomeDB. We were able to genotype 3 other candidate SNPs previously reported for BE susceptibility (rs41341748, rs2715425, rs6898743) on the Sequenom iPLEX panel used to genotype UK Replication 2. None of these SNPs showed associations with BE risk (Supplementary Table 1). Because rs41341748 (MSR1 p.Arg293Gly) has a low minor allele frequency (<5%) and consequently our power to detect an association was relatively low, we additionally genotyped it in UK Replication 3. After meta-analysis of Replication 2 and 3, we remained unable to replicate the previously reported association²³ between this SNP and BE risk (OR = 1.07; 95% CI: 0.70-1.43; P = .79).

Assessment of Previously Reported Barrett's Esophagus + Esophageal Adenocarcinoma Single Nucleotide Polymorphisms and Meta-analysis With BEAGESS Data

Three new genome-wide significant BE + EAC loci (4 SNPs) were recently identified by Levine et al¹² in a combined analysis of EAC and BE: rs10419226 and rs10423674 in *CRTC1*, rs11789015 in *BARX1* and rs2687201 within 100 kb of *FOXP1*. None of these associations was genome-wide significant at $P < 5 \times 10^{-8}$ when Levine et al restricted their analysis to BE cases alone, although one SNP, rs10419226, near *CRTC1*, reached $P = 5.5 \times 10^{-8}$. In our datasets, only rs10423674 had been directly genotyped, but the remaining SNPs were all reliably

781

782

783

784

816

806

838

839

8 Palles et al

841

842

843

844

845

846

847

848

849

850

851

852

853

854

855

856

857

858

859

860

861

862

863

864

865

866

basic and Translational at

876

877

878

879

880

881

882

883

884

885

886

887

888

889

890

891

892

893

894

895

896

897

898

899

900

Q12

Gastroenterology Vol. ■, No. ■

imputed in our Discovery Phase samples (Info scores >0.95). However, the controls used in the replication phase of the Levine et al¹² study overlapped entirely with the controls used in our Discovery phase (WTCCC2 controls) and we therefore used alternative UK controls from the CoRGI study (see Methods).

Of the 4 Levine SNPs, rs10423674, 1 of 2 SNPs in *CRTC1* and rs2687201 near *FOXP1*, were supported in this study (P = 0.02; OR = 1.14; 95% CI: 1.03–1.27 and P = 0.05; OR = 0.94; 95% CI: 0.88–1.00, respectively). There was also some support for rs11789015, near *BARX1* (P = 0.07; OR = 0.90; 95% CI: 0.81–1.01). However, the association at rs10419226, near *CRTC1*, was not replicated in our data (P = 0.87; OR = 1.01; 95% CI: 0.91–1.11). All 4 SNPs still reached genome-wide significance ($P < 5 \times 10^{-8}$) upon meta-analysis of our BE and Levine's BE + EAC datasets. In a BE-only meta-analysis, the associations improved with the inclusion of our data for 3 out of the 4 SNPs, with 1 (rs2687201, *FOXP1*) reaching genome-wide significance for BE (Table 2).

We then addressed the other 87 other SNPs with $P_{assoc} < 10^{-4}$ in the Levine data (Supplementary Table 3 of Levine et al¹²). Of these, 73 were directly genotyped in our samples or could be imputed with an IMPUTE2 info score of >0.95. Of the 10 SNPs that could not be imputed with high quality, only one had $P_{assoc} < 10^{-5}$ in the original Levine data; we therefore genotyped this SNP (rs11771429) using KASPar in our cases and controls. We did not obtain genotypes for the remaining 9 SNPs. On performing a metaanalysis of the Levine BE + EAC cases with our UK Discovery Phase, 4 SNPs (rs1497205, rs254348, rs3784262, and rs4523255) showed P_{meta} < 10^{-5} and were not strongly correlated with 1 of the 4 BE + EAC SNPs reported previously. We therefore genotyped these 4 SNPs in our Replication Phase samples; rs3784262 (near ALDH1A2) was associated with BE + EAC (OR = 0.90; 95% CI: 0.87-0.93; $P = 3.72 \times 10^{-9}$). No SNP was formally associated with BE alone (Table 2). eQTL and ASE analysis (see XXX) did not show associations for the rs3784262 proxy rs7165247 in TCGA data or other public data sets (details not shown).

Discussion

We have added 2 new BE predisposition SNPs, rs3072 on chromosome 2p24 and rs2701108 on chromosome 12q24, to the 2 BE SNPs on chromosome 6p21 (HLA region) and chromosome 16q23 (near FOXF1) that we reported previously.¹¹ Both of the new SNPs remained at or very near genome-wide significance when analysis was restricted to cases with intestinal metaplasia. In silico fine mapping provided evidence that rs3072 and/or 1 of 3 nearby SNPs might be functional because they map to putative enhancer regions. The nearby gene, GDF7, is the best functional candidate, because this encodes the BMP12 protein and the BMP pathway has previously been implicated in the development of BE.³⁶ The importance of this pathway in BE is also suggested in Ingenuity Pathway Analysis, where bone morphogenetic protein 4 acts upstream of proteins encoded by genes close to the BE predisposition SNPs. GDF7 plays a

role in the neural system and tendon/ligament development and repair,^{37,38} and also regulates Hedgehog and Wnt signaling pathways that impact on esophageal development through *FOXF1* and *TBX5*. In the chromosome 12q24 region, rs1920562 (the top imputed SNP) provided the strongest association signal and maps to a possible enhancer. Gene expression analysis did not suggest the target of the chromosome 12q24 variation, although *TBX5* is a very strong functional candidate. It is involved in cardiac development and its deficiency causes thoracic malformations and abnormalities of the diaphragmatic musculature,^{39,40} which could predispose patients to hiatus hernia and acid reflux, 2 subphenotypes of BE.

Messenger RNA expression analysis using TCGA EAC data and public data from leukocytes and adipocytes provided little evidence that rs3072 or rs2701108 (or other SNPs in strong LD) were eQTLs or influenced ASE. The absence of these associations is typical for GWAS or cancer or precancerous traits. Even for the "prototypic" multicancer SNP rs6983267, convincingly demonstrating the effects of SNP alleles on gene expression has required considerable additional work in a variety of systems, and even now, consistent eQTL and ASE associations have not been shown.⁴¹⁻⁴⁴ The likely major reason for the lack of eQTLs at GWAS SNPs is that the SNPs have their effects in a restricted set of cells or at a particular time. There is, for example, evidence that the forkhead box (FOX) proteins are most strongly expressed during embryogenesis, and that the levels of these transcription factors are critical for proper development.⁴⁵⁻⁴⁷ Given this, our first choice in searching for eQTLs would be cells in the developing human thorax. Unfortunately, such sample collections do not currently exist.

We showed rs2687201 (*FOXP1*) to be associated with disease in a BE-only analysis. Our data generally support the report by Levine et al¹² of associations between BE + EAC and SNPs on chromosome 3 (*FOXP1*), chromosome 9 (*BARX1*), and one of the SNPs on chromosome 19 (*CRTC1*). However, we were not able to replicate the association observed for another SNP (rs10419226) in *CRTC1*. For this last SNP, the meta-analysis showed evidence of significant heterogeneity between the BEACON/BEAGESS data and our data (Table 2), and in the absence of clear reasons for this difference, we caution against drawing firm conclusions here.

We found another SNP, rs3784262 (*ALDH1A2*), to be formally associated with BE + EAC upon meta-analysis of our data with the Levine BE + EAC dataset. *ALDH1A2* encodes retinaldehyde dehydrogenase 2, which catalyzes the synthesis of retinoic acid and may also be involved in alcohol metabolism.⁴⁸ Of the candidate SNPs we assessed (Supplementary Table 8), we found supporting evidence, albeit short of genome-wide significance, for rs909253 ($P = 3.1 \times 10^{-4}$), mapping to an intronic region of *LTA* within the HLA region, but not in LD with rs9257809, the other HLA BE SNP.

We previously reported that our original GWAS provided evidence that multiple common variants, each with small effects contribute to BE susceptibility.¹¹ Ek et al¹³ recently 901

902

903

904

905

906

907

	000000000000000000000000000000000000	
	× × × × × × 0 0 0 0 0 0 0 0 0 0 0 0 0 0	8 8 8 8 8 1 1 1 1 1 1 1 1 1 6 6 6 6 6 6
21111111000000	0 9 8 7 6 5 4 3 2 1 0 9 8 7 6 5	4 3 2 1 0 9 8 7 6 5 4 3 2 1 0 9 8 7 6 5 4 3 2 1
$\vec{O} \cdot \vec{U} \cdot \vec{A} \cdot \vec{V} \cdot \vec{D} \cdot \vec{V} \cdot \vec{D} \cdot \vec{V} \cdot \vec{O} \cdot \vec{V} \cdot \vec{V} \cdot \vec{O} \cdot \vec{V} \cdot $		

SNP	Chr	Position	Nearby genes	Minor/ major allele	BE+/ -EAC	Levine et al meta OR (95% Cl)	P _{Levine} et al	This study meta OR (95% Cl)	P _{This Study}	This study + Levine et al ¹² meta OR (95% Cl)	P _{meta}	Meta I ²	No. of studies
rs2687201	3	70928930	FOXP1	T/G	BE	1.18 (1.10–1.26)	$2.00 imes 10^{-6}$	1.14 (1.03–1.27)	1.18×10^{-2}	1.16 (1.10–1.23)	4.61×10^{-8}	0.00	3
					BE/EAC	1.18 (1.12-1.25)	$5.47 imes10^{-9}$			1.17 (1.11–1.23)	6.70×10^{-10}	0.00	3
rs11789015	9	96716028	BARX1	G/A	BE	0.85 (0.79-0.91)	5.08×10^{-6}	0.90 (0.81-1.01)	6.63×10^{-2}	0.86 (0.81-0.92)	$1.38 imes10^{-6}$	0.00	3
					BE/EAC	0.83 (0.79-0.88)	1.02×10^{-9}			0.85 (0.81-0.89)	1.14×10^{-10}	0.00	3
rs10419226	19	18803172	CRTC1	A/C	BE	1.19 (1.12–1.26)	5.54×10^{-8}	1.01 (0.91–1.11)	8.65×10^{-1}	1.13 (1.08–1.20)	$2.14 imes 10^{-6}$	0.82	3
					BE/EAC	1.18 (1.12–1.24)	3.55×10^{-10}			1.14 (1.09–1.19)	$1.17 imes 10^{-8}$	0.82	3
rs10423674	19	18817903	CRTC1	T/G	BE	0.85 (0.80–0.91)	1.92×10^{-6}	0.94 (0.88–1.00)	4.88×10^{-2}	0.89 (0.85-0.93)	2.99×10^{-7}	0.40	5
					BE/EAC	0.84 (0.80-0.89)	1.75×10^{-9}			0.88 (0.84–0.91)	4.87×10^{-11}	0.49	5
rs1497205	4	76169067	PARM1,	C/T	BE	0.86 (0.80-0.92)	2.86×10^{-5}	0.92 (0.87-0.98)	7.59×10^{-1}	0.90 (0.86-0.94)	$2.57 imes 10^{-6}$	0.00	6
			RCHY1		BE/EAC	0.87 (0.82-0.93)	1.28×10^{-5}			0.90 (0.86–0.94)	3.68×10^{-7}	0.00	6
rs254348	16	65980789		T/C	BE	0.88 (0.83-0.94)	$1.15 imes10^{-4}$	0.95 (0.91-1.01)	8.88×10^{-2}	0.93 (0.89-0.97)	5.49×10^{-4}	0.53	6
					BE/EAC	0.89 (0.84-0.94)	$1.40 imes 10^{-5}$			0.92 (0.89–0.96)	2.81×10^{-5}	0.53	6
rs3784262	15	58253106	ALDH1A2	G/A	BE	0.85 (0.80-0.90)	3.62×10^{-7}	0.93 (0.89-0.98)	5.13×10^{-3}	0.91 (0.87-0.94)	$1.37 imes 10^{-6}$	0.12	9
					BE/EAC	0.88 (0.83-0.92)	6.72×10^{-7}			0.90 (0.87-0.93)	3.72×10^{-9}	0.16	9
rs4523255	8	8713038	MFHAS1	A/G	BE	1.13 (1.06–1.21)	2.46×10^{-4}	1.07 (1.01–1.12)	2.11×10^{-2}	1.09 (1.05-1.14)	$2.48 imes 10^{-5}$	0.36	6
					BE/FAC	1 13 (1 07-1 20)	4.15×10^{-5}			1.09(1.05 - 1.14)	9.24×10^{-6}	0 46	6

Table 2. Meta-analysis With Our Data for 4 BE/EAC SNPs and 4 Other Selected SNPs with $P < 1 \times 10^{-4}$ from Levine et al¹²

NOTE. The minimum meta-analysis comprised the Levine et al¹² Discovery and Replication Phases and our Discovery Phase (with amended controls, as described in Methods). rs10423674 was additionally genotyped in our UK and Dutch Replication Phase 1. rs1497205, rs254348, rs3784262, and rs4523255 were genotyped in our UK and Dutch Replication Phase 1. rs1497205, rs254348, rs3784262, and rs4523255 were genotyped in our UK and Dutch Replication Phase 1. rs1497205, rs254348, rs3784262, and rs4523255 were genotyped in our UK and Dutch Replication Phase 1 and UK Replication Phase 2. rs3784262 was also genotyped in Irish Replication Phase 2 samples and UK and Belgium Replication Phase 3 samples. For rs10419226, which shows evidence of inter-study heterogeneity, random effects model *P* values for BE and BE/EAC are .10 and .04, respectively. Chr, chromosome.

BASIC AND TRANSLATIONAL AT

FLA 5.2.0 DTD \blacksquare YGAST59419_proof \blacksquare 12 December 2014 \blacksquare 1:48 am \blacksquare

8

g

Barrett's Esophagus SNPs

Near

10 Palles et al

estimated that the heritability of BE is 35% (SE 6%). We 1081 also found that the heritability of BE is highly significant, but 1082 explains only 9.99% of BE risk (SE 1.2%). Our GWAS con-1083 sisted of UK cases and controls, and the BEACON/BEAGESS 1084 samples used by Ek et al originated from 3 continents 1085 (Europe, North America, and Australia). Cryptic population 1086 stratification could perhaps explain the larger estimate of 1087 heritability obtained using the BEACON/BEAGESS GWAS. In 1088 addition, we used software to calculate LD-adjusted kin-1089 ships, such that the SNPs used in the heritability analysis 1090 were weighted according to local LD structure. It has been 1091 found that heritability estimation from genome-wide SNPs 1092 is highly sensitive to uneven LD; causal SNPs in regions of 1093 high LD can lead to overestimation of heritability and 1094 conversely causal SNPs in regions of low LD can result in an 1095 underestimation of heritability.49 1096

Although the BE GWAS have not yet identified the 1097 functional SNPs in each region or their gene targets, the 1098 information generated already permits the generation of 1099 hypotheses regarding processes that may be involved in BE. 1100 First, transcription factors involved in development and 1101 structure of the thorax, diaphragm, and esophagus may be 1102 important: the SNPs near FOXF1, FOXP1, BARX1, and TBX5 1103 might act in this way and the genes appear to be function-1104 ally related (Supplementary Figure 1). Second, the inflam-1105 matory response may be important: the SNPs within the 1106 HLA region (rs9257809 and, perhaps, rs909253) might act in this way and pathway analysis provided suggestive evidence of a role for type 1 diabetes genes in BE etiology. A basic and translational at plausible, testable hypothesis is that these 2 groups of SNPs respectively influence the tendency to gastroesophageal reflux disease, perhaps through thoracic and diaphragmatic structure (hiatal hernia defect), and the inflammatory response to the refluxed gastric acid. Given the limited scope for clinical intervention in the former processes, we await with interest the outcome of trials such as AspECT 1116 that target the inflammatory response to gastric reflux.^{1,1} 1117

Supplementary Material

Note: To access the supplementary material accompanying this article, visit the online version of *Gastroenterology* at www.gastrojournal.org, and at http://dx.doi.org/10.1053/j.gastro.2014.10.041.

References

- 1. **Jankowski J, Barr H**, Wang K, et al. Diagnosis and management of Barrett's oesophagus. BMJ 2010; 341:c4551.
- Jankowski J, Odze R. Biomarkers in gastroenterology; between hype and hope comes histopathology. Am J Gastroenterol 2009;104:1093–1096.
- Duggan C, Onstad L, Hardikar S, et al. Association between markers of obesity and progression from Barrett's esophagus to esophageal adenocarcinoma. Clin Gastroenterol Hepatol 2013;11:934–943.
- 4. Iyer PG, Borah BJ, Heien HC, et al. Association of Barrett's esophagus with type II diabetes mellitus: results
- 1139 1140

1118

1119

1120

1121

1122

1123

1124

1125

1126

1127

1128

1129

1130

1131

1132

1133

1134

1135

1136

1137

1138

Q13

Gastroenterology Vol. ■, No. ■

1141

1142

1143

1144

1145

1146

1147

1148

1149

1150

1151

1152

1153

1154

1155

1156

1157

1158

1159

1160

1161

1162

1163

1164

1165

1166

1167

1168

1169

1170

1171

1172

1173

1174

1175

1176

1177

1178

1179

1180

1181

1182

1183

1184

1185

1186

1187

1188

1189

1190

1191

1192

1193

1194

1195

1196

1197

from a large population-based case-control study. Clin Gastroenterol Hepatol 2013;11:1108–1114.e5.

- DeVault K, McMahon BP, Celebi A, et al. Defining esophageal landmarks, gastroesophageal reflux disease, and Barrett's esophagus. Ann N Y Acad Sci 2013; 1300:278–295.
- Cameron AJ, Lagergren J, Henriksson C, et al. Gastroesophageal reflux disease in monozygotic and dizygotic twins. Gastroenterology 2002;122:55–59.
- Mohammed I, Cherkas LF, Riley SA, et al. Genetic influences in gastro-oesophageal reflux disease: a twin study. Gut 2003;52:1085–1089.
- 8. Lembo A, Zaman M, Jones M, et al. Influence of genetics on irritable bowel syndrome, gastro-oesophageal reflux and dyspepsia: a twin study. Aliment Pharmacol Ther 2007;25:1343–1350.
- 9. Groves C, Jankowski J, Barker F, et al. A family history of Barrett's oesophagus: another risk factor? Scand J Gastroenterol 2005;40:1127–1128.
- Chak A, Ochs-Balcom H, Falk G, et al. Familiality in Barrett's esophagus, adenocarcinoma of the esophagus, and adenocarcinoma of the gastroesophageal junction. Cancer Epidemiol Biomarkers Prev 2006;15:1668–1673.
- 11. **Su Z, Gay LJ**, Strange A, et al. Common variants at the MHC locus and at chromosome 16q24.1 predispose to Barrett's esophagus. Nat Genet 2012;44:1131–1136.
- 12. Levine DM, Ek WE, Zhang R, et al. A genome-wide association study identifies new susceptibility loci for esophageal adenocarcinoma and Barrett's esophagus. Nat Genet 2013;45:1487–1493.
- Ek WE, Levine DM, D'Amato M, et al. Germline genetic contributions to risk for esophageal adenocarcinoma, Barrett's esophagus, and gastroesophageal reflux. J Natl Cancer Inst 2013;105:1711–1718.
- 14. Dura P, van Veen EM, Salomon J, et al. Barrett associated MHC and FOXF1 variants also increase esophageal carcinoma risk. Int J Cancer 2013;133:1751–1755.
- 15. Woo J, Miletich I, Kim BM, et al. Barx1-mediated inhibition of Wnt signaling in the mouse thoracic foregut controls tracheo-esophageal septation and epithelial differentiation. PLoS One 2011;6:e22493.
- 16. Shu W, Lu MM, Zhang Y, et al. Foxp2 and Foxp1 cooperatively regulate lung and esophagus development. Development 2007;134:1991–2000.
- 17. Jankowski J, Barr H. Improving surveillance for Barrett's oesophagus: AspECT and BOSS trials provide an evidence base. BMJ 2006;332:1512.
- Sharma P, Dent J, Armstrong D, et al. The development and validation of an endoscopic grading system for Barrett's esophagus: the Prague C & M criteria. Gastroenterology 2006;131:1392–1399.
- Bennett C, Vakil N, Bergman J, et al. Consensus statements for management of Barrett's dysplasia and early-stage esophageal adenocarcinoma, based on a Delphi process (BADCAT). Gastroenterology 2012;143: 336–346.
- 20. Houlston RS, Cheadle J, Dobbins SE, et al. Meta-analysis of three genome-wide association studies identifies susceptibility loci for colorectal cancer at 1q41, 3q26.2, 12q13.13 and 20q13.33. Nat Genet 2010;42:973–977.

2015

Barrett's Esophagus SNPs Near TBX5 and GDF7 11

- 1201
 1202
 1203
 21. Petridis C, Shinomiya I, Kohut K, et al. Germline CDH1 mutations in bilateral lobular carcinoma in situ. Br J Cancer 2013;100:1053–1057.
- 1205
 1204
 1205
 1206
 1206
 1206
 1207
 1207
 1208
 1209
 1209
 1209
 1209
 1209
 1209
 1209
 1209
 1209
 1209
 1209
 1209
 1209
 1209
 1209
 1209
 1209
 1209
 1209
 1209
 1209
 1209
 1209
 1209
 1209
 1209
 1209
 1209
 1209
 1209
 1209
 1209
 1209
 1209
 1209
 1209
 1209
 1209
 1209
 1209
 1209
 1209
 1209
 1209
 1209
 1209
 1209
 1209
 1209
 1209
 1209
 1209
 1209
 1209
 1209
 1209
 1209
 1209
 1209
 1209
 1209
 1209
 1209
 1209
 1209
 1209
 1209
 1209
 1209
 1209
 1209
 1209
 1209
 1209
 1209
 1209
 1209
 1209
 1209
 1209
 1209
 1209
 1209
 1209
 1209
 1209
 1209
 1209
 1209
 1209
 1209
 1209
 1209
 1209
 1209
 1209
 1209
 1209
 1209
 1209
 1209
 1209
 1209
 1209
 1209
 1209
 1209
 1209
 1209
 1209
 1209
 1209
 1209
 1209
 1209
 1209
 1209
 1209
 1209
 1209
 1209
 1209
 1209
 1209
 1209
 1209
 1209
 1209
 1209
 1209
 1209
 1209
 1209
 1209
 1209
 1209
 1209</l
- 1207
 1208
 1209
 1210
 1210
 1210
 1210
 1211
 1212
 1212
 1212
 1213
 1213
 1214
 1214
 1214
 1215
 1214
 1216
 1217
 1218
 1218
 1218
 1219
 1219
 1210
 1211
 1211
 1211
 1211
 1211
 1211
 1211
 1211
 1211
 1211
 1211
 1211
 1211
 1211
 1211
 1211
 1211
 1211
 1211
 1211
 1211
 1211
 1211
 1211
 1211
 1211
 1211
 1211
 1211
 1211
 1211
 1211
 1211
 1211
 1211
 1211
 1211
 1211
 1211
 1211
 1211
 1211
 1211
 1211
 1211
 1211
 1211
 1211
 1211
 1211
 1211
 1211
 1211
 1211
 1211
 1211
 1211
 1211
 1211
 1211
 1211
 1211
 1211
 1211
 1211
 1211
 1211
 1211
 1211
 1211
 1211
 1211
 1211
 1211
 1211
 1211
 1211
 1211
 1211
 1211
 1211
 1211
 1211
 1211
 1211
 1211
 1211
 1211
 1211
 1211
 1211
 1211
 1211
 1211
 1211
 1211
 1211
 1211
 1211
 1211
 1211
 1211
 1211
 1211
 1211
 1211
 1211
 1211
 1211
 1211
 1211
 1211
 1211
 1211
 1211
 1211
 1211
 1211
 1211
 1211</l
- 1211
 1212
 1213
 1213
 1214
 1215
 24. McElholm AR, McKnight AJ, Patterson CC, et al. A population-based study of IGF axis polymorphisms and the esophageal inflammation, metaplasia, adenocarcinoma sequence. Gastroenterology 2010;139: 204–212.
- 121625. Magi R, Morris AP. GWAMA: software for genome-wide
association meta-analysis. BMC Bioinform 2010;11:288.
- Higgins JP, Thompson SG, Deeks JJ, et al. Measuring
 inconsistency in meta-analyses. BMJ 2003;327:557–560.
- 1220
 1221
 1221
 1222
 1222
 1223
 27. Dunlop MG, Dobbins SE, Farrington SM, et al. Common variation near CDKN1A, POLD3 and SHROOM2 influences colorectal cancer risk. Nat Genet 2012;44: 770–776.
- 1224
 1225
 1226
 28. Delaneau O, Zagury JF, Marchini J. Improved wholechromosome phasing for disease and population genetic studies. Nat Methods 2013;10:5–6.
- 1227
 1228
 1228
 1229
 1230
 29. Howie BN, Donnelly P, Marchini J. A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. PLoS Genet 2009; 5:e1000529.
- 30. Ward LD, Kellis M. HaploReg: a resource for exploring
 chromatin states, conservation, and regulatory motif alterations within sets of genetically linked variants.
 Nucleic Acids Res 2012;40:D930–D934.
- 1235 31. Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. Nucleic Acids Res 2010;38:e164.
- 1238 32. Boyle AP, Hong EL, Hariharan M, et al. Annotation of
 1239 functional variation in personal genomes using Reg1240 ulomeDB. Genome Res 2012;22:1790–1797.
- 124133. Dulak AM, Stojanov P, Peng S, et al. Exome and whole-
genome sequencing of esophageal adenocarcinoma
identifies recurrent driver events and mutational
complexity. Nat Genet 2013;45:478–486.
- 124534. Zeller T, Wild P, Szymczak S, et al. Genetics and
beyond—the transcriptome of human monocytes and
disease susceptibility. PLoS One 2010;5:e10693.
- 1248
 1249
 1250
 35. Nica AC, Parts L, Glass D, et al. The architecture of gene regulatory variation across multiple human tissues: the MuTHER study. PLoS Genet 2011;7:e1002003.
- 1251
 1252
 1253
 1254
 1255
 36. Castillo D, Puig S, Iglesias M, et al. Activation of the BMP4 pathway and early expression of CDX2 characterize non-specialized columnar metaplasia in a human model of Barrett's esophagus. J Gastrointest Surg 2012; 16:227–237; discussion 237.
- 37. Berasi SP, Varadarajan U, Archambault J, et al. Divergent activities of osteogenic BMP2, and tenogenic BMP12 and BMP13 independent of receptor binding affinities. Growth Factors 2011;29:128–139.
- 1259 1260

- Currle DS, Cheng X, Hsu CM, et al. Direct and indirect roles of CNS dorsal midline cells in choroid plexus epithelia formation. Development 2005;132:3549–3559.
- 39. Arora R, Metzger RJ, Papaioannou VE. Multiple roles and interactions of Tbx4 and Tbx5 in development of the respiratory system. PLoS Genet 2012;8:e1002866.
- 40. Mori AD, Bruneau BG. TBX5 mutations and congenital heart disease: Holt-Oram syndrome revealed. Curr Opin Cardiol 2004;19:211–215.
- 41. Pomerantz MM, Ahmadiyeh N, Jia L, et al. The 8q24 cancer risk variant rs6983267 shows long-range interaction with MYC in colorectal cancer. Nat Genet 2009; 41:882–884.
- Tuupanen S, Turunen M, Lehtonen R, et al. The common colorectal cancer predisposition SNP rs6983267 at chromosome 8q24 confers potential to enhanced Wnt signaling. Nat Genet 2009;41:885–890.
- Ling H, Spizzo R, Atlasi Y, et al. CCAT2, a novel noncoding RNA mapping to 8q24, underlies metastatic progression and chromosomal instability in colon cancer. Genome Res 2013;23:1446–1461.
- 44. Takatsuno Y, Mimori K, Yamamoto K, et al. The rs6983267 SNP is associated with MYC transcription efficiency, which promotes progression and worsens prognosis of colorectal cancer. Ann Surg Oncol 2013; 20:1395–1402.
- Mahlapuu M, Enerback S, Carlsson P. Haploinsufficiency of the forkhead gene Foxf1, a target for sonic hedgehog signaling, causes lung and foregut malformations. Development 2001;128:2397–2406.
- 46. Shaw-Smith C. Genetic factors in esophageal atresia, tracheo-esophageal fistula and the VACTERL association: roles for FOXF1 and the 16q24.1 FOX transcription factor gene cluster, and review of the literature. Eur J Med Genet 2010;53:6–13.
- Wang DH, Tiwari A, Kim ME, et al. Hedgehog signaling regulates FOXA2 in esophageal embryogenesis and Barrett's metaplasia. J Clin Invest 2014;10:1172–1178.
- Gyamfi MA, Kocsis MG, He L, et al. The role of retinoid X receptor alpha in regulating alcohol metabolism. J Pharmacol Exp Ther 2006;319:360–368.
- 49. Speed D, Hemani G, Johnson MR, et al. Improved heritability estimation from genome-wide SNPs. Am J Hum Genet 2012;91:1011–1021.

Author names in bold designate shared co-first authorship.

Received May 6, 2014. Accepted October 21, 2014.

Reprint requests

Address requests for reprints to: Janusz Jankowski, XX, Associate Dean's Office, Faculty of Health and Life Sciences, Coventry University, Priory Street, Coventry, CV1 5FB. e-mail: XXX; fax: XXX; Claire Palles, XX, Molecular and Population Genetics Laboratory, Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, England, OX3 7BN, United Kingdom. e-mail: XXX; fax: XXX; or Ian Tomlinson, XX, Molecular and Population Genetics, Laboratory, Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, England, OX3 7BN, United Kingdom. e-mail: XXX; fax: XXX; or Ian Tomlinson, XX, Molecular and Population Genetics, Laboratory, Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, England, OX3 7BN, United Kingdom. e-mail: XXX; fax: XXX;

Acknowledgments

FLA 5.2.0 DTD ■ YGAST59419 proof ■ 12 December 2014 ■ 1:48 am ■ ce

The authors would like to thank the WTCCC2 consortium, the BEACON consortium and the AspECT, BOSS, and ChOPIN trial teams (comprising the EAGLE consortium) and participants in all these studies. The Discovery Phase and Immunochip replication were funded by the Wellcome Trust IPOD

1315

1316

1317

1318

1319

1320

1261

1262

1263

1264

1265

1266

1267

1268

1269

1270

1271

1272

1273

1274

1275

1276

1277

1278

1279

1280

1281

1282

1283

1284

1285

1286

AT

TRANSLATIONAL

12 Palles et al

TRANSLATIO

Gastroenterology Vol. ■, No. ■

grant (084722/Z/08/Z). In addition, sample collection for all phases was core supported by the Cancer Research UK funded AspECT, ChOPIN and Handel trials. CoRGI and GLACIER were also funded by Cancer Research UK. In addition, the authors thank AstraZeneca for an educational grant for the tissue and blood collection. Core funding to the Wellcome Trust Centre for Human Genetics was provided by the Wellcome Trust (090532/Z/09/Z). We also thank Liam J. Murray and Wong-Ho Chow, who provided samples as part of the BEACON consortium. The authors would also like to thank the Experimental Cancer Medicine Centre-supported University of Southampton Faculty of Medicine Tissue Bank. The Cancer Genome Atlas (TCGA) esophageal carcinoma data were also gratefully used for expression

analyses. Finally, the authors want to thank Dr Vincent Plagnol of University College London for advice on statistical analyses and presentation.

Conflicts of interest

This author discloses the following: Janusz Jankowski was an AstraZeneca consultant from 2002 to 2012. The remaining authors disclose no conflicts. Q6

Funding

FLA 5.2.0 DTD ■ YGAST59419 proof ■ 12 December 2014 ■ 1:48 am ■ ce

This work was supported by Cancer Research UK, Peninsula Schools of ^{Q15} Medicine and Dentistry School, Wellcome Trust (084722/Z/08/Z and 090532/ ^{Q16} Z/09/Z), and an AstraZeneca UK educational grant. Q7

bhas DPIN Rese al g Trus ded ike of S e A for