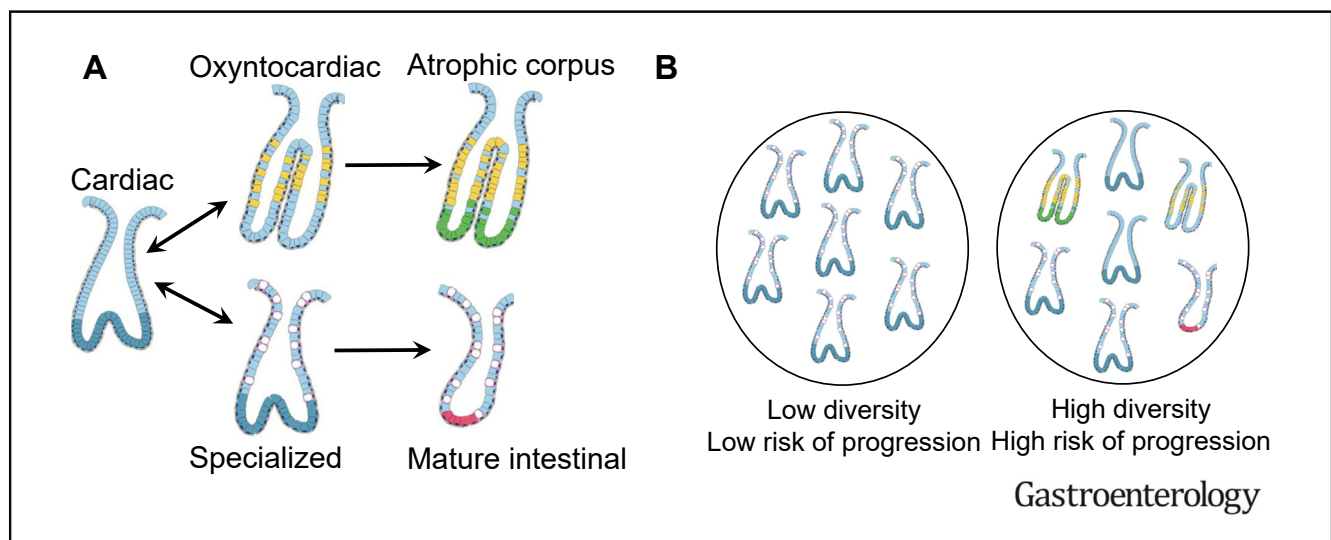


Clonal Transitions and Phenotypic Evolution in Barrett's Esophagus



James A. Evans,^{1,*} Emanuela Carlotti,^{1,*} Meng-Lay Lin,¹ Richard J. Hackett,¹ Magnus J. Haughey,² Adam M. Passman,¹ Lorna Dunn,³ George Elia,¹ Ross J. Porter,⁴ Mairi H. McLean,⁴ Frances Hughes,⁵ Joanne ChinAleong,⁶ Philip Woodland,⁷ Sean L. Preston,⁷ S. Michael Griffin,^{2,8} Laurence Lovat,^{9,10} Manuel Rodriguez-Justo,¹¹ Weini Huang,² Nicholas A. Wright,¹² Marnix Jansen,^{11,13} and Stuart A. C. McDonald¹

¹Clonal Dynamics in Epithelia Laboratory, Queen Mary University of London, London, United Kingdom; ²School of Mathematical Sciences, Queen Mary University of London, London, United Kingdom; ³Northern Institute for Cancer Research, Newcastle University, Newcastle, United Kingdom; ⁴Department of Gastroenterology, University of Aberdeen, Aberdeen, United Kingdom; ⁵Department of Surgery, Barts Health NHS Trust, Royal London Hospital, London, United Kingdom; ⁶Department of Histopathology, Barts Health NHS Trust, Royal London Hospital, London, United Kingdom; ⁷Endoscopy Unit, Barts Health NHS Trust, Royal London Hospital, London, United Kingdom; ⁸Royal College of Surgeons of Edinburgh, Edinburgh, United Kingdom; ⁹Oesophagogastric Disorders Centre, Department of Gastroenterology, University College London Hospitals, London, United Kingdom; ¹⁰Research Department of Tissue and Energy, University College London Division of Surgical and Interventional Science, University College London, London, United Kingdom; ¹¹Department of Cellular Pathology, University College London Hospitals, London, United Kingdom; ¹²Epithelial Stem Cell Laboratory, Centre for Cancer Genomics and Computational Biology, Barts Cancer Institute, Queen Mary University of London, London, United Kingdom; and ¹³UCL Cancer Institute, University College London, London, United Kingdom



See Covering the Cover synopsis on page 999.

BACKGROUND & AIMS: Barrett's esophagus (BE) is a risk factor for esophageal adenocarcinoma but our understanding of how it evolves is poorly understood. We investigated BE gland phenotype distribution, the clonal nature of phenotypic change, and how phenotypic diversity plays a role in progression. **METHODS:** Using immunohistochemistry and histology, we analyzed the distribution and the diversity of gland phenotype between and within biopsy specimens from patients with nondysplastic BE and those who had progressed to dysplasia or had developed postesophagectomy BE. Clonal relationships were determined by the presence of shared mutations between distinct gland types using laser capture microdissection

sequencing of the mitochondrial genome. **RESULTS:** We identified 5 different gland phenotypes in a cohort of 51 nondysplastic patients where biopsy specimens were taken at the same anatomic site (1.0–2.0 cm superior to the gastroesophageal junction). Here, we observed the same number of glands with 1 and 2 phenotypes, but 3 phenotypes were rare. We showed a common ancestor between parietal cell-containing, mature gastric (oxyntocardiac) and goblet cell-containing, intestinal (specialized) gland phenotypes. Similarly, we have shown a clonal relationship between cardiac-type glands and specialized and mature intestinal glands. Using the Shannon diversity index as a marker of gland diversity, we observed significantly increased phenotypic diversity in patients with BE adjacent to dysplasia and predysplasia compared to nondysplastic BE and postesophagectomy BE, suggesting that diversity develops over

time. **CONCLUSIONS:** We showed that the range of BE phenotypes represents an evolutionary process and that changes in gland diversity may play a role in progression. Furthermore, we showed a common ancestry between gastric and intestinal-type glands in BE.

Keywords: Barrett's Esophagus (BE); Clonal; Esophageal Adenocarcinoma (EA); Diversity; Evolution.

Barrett's esophagus (BE) is the only known precursor condition of esophageal adenocarcinoma (EA) and is characterized by the metaplastic replacement of the normal squamous epithelium of the distal esophagus with a columnar epithelial phenotype that frequently contains intestinal metaplasia (IM).¹ In some countries, the diagnosis of BE is based on the histopathologic presence of IM in esophageal biopsy specimens,² but in others, only the endoscopic presence of columnar epithelium in the distal esophagus is required.³ It is widely assumed that the presence of IM assists in stratifying patients' cancer risks, but there is evidence to suggest that this is not always the case.⁴⁻⁶ We have previously shown that glands that do not contain goblet cells can clonally expand, accumulate oncogenic *TP53* mutations, and be the source of EA.⁷ This finding highlights an important lack of understanding of the evolution of the BE epithelial phenotype.^{8,9}

BE displays a rich diversity of morphologically distinct glands¹⁰⁻¹³ that contain an admixture of both gastric and intestinal epithelial cell lineages.¹⁴ At present, we do not fully understand the scope of epithelial lineage diversity nor whether these lineages represent an evolutionary pathway that may be altered in patients who progress to dysplasia. Previous studies have documented that genotypic diversity predicts the risk of BE progressing to cancer,¹⁵⁻¹⁷ but it is currently unknown whether this is reflected in phenotypic diversity across the segment. There is controversy as to the mechanism by which genetic diversity evolves, with some data showing it is acquired over time,¹⁸ and others showing that diversity is inherent to BE¹⁷ and is always increased relative to nonprogressors.¹⁵ To date, the significance of the evolution of gland phenotype in BE has not been fully appreciated. This is an important omission when we consider that all current diagnoses are entirely based on histopathologic analysis³ and that natural selection acts fundamentally on phenotype, not genotype.¹⁹

Here, we address this unresolved issue by showing the frequency distribution of gland phenotypes in a cross-sectional BE patient cohort, both at a fixed point within the metaplastic segment (1.0–2.0 cm proximal of the gastroesophageal junction) and throughout the BE segment. Using mitochondrial DNA mutations as clonal marks,^{20,21} we then demonstrate ancestral clonal relationships between different gland phenotypes within nondysplastic BE, indicating phenotypic evolution. Finally, we measure gland phenotype diversity in nondysplastic BE from patients who have no history of dysplasia and compare this to BE biopsy samples taken before a diagnosis of dysplasia, nondysplastic BE

WHAT YOU NEED TO KNOW

BACKGROUND AND CONTEXT

We do not fundamentally understand the phenotypic evolution of Barrett's esophagus, nor do we fully understand the distribution of gland phenotype diversity and if this is associated with progression.

NEW FINDINGS

There is heterogeneity of Barrett's esophagus gland phenotype close to the gastroesophageal junction that results from the evolution between gastric and intestinal gland types. Diversity of gland phenotypes is associated with progression.

LIMITATIONS

More longitudinal studies are needed to determine specific evolutionary steps to dysplasia.

IMPACT

This study shows for the first time how local diversity affects the Barrett's esophagus lesion and that gastric and intestinal gland types show a common ancestry. Understanding how diversity affects patients who progress to dysplasia may prove an important predictive biomarker of cancer risk.

adjacent to dysplasia, and postesophagectomy nondysplastic BE. Together, these data show, for the first time to our knowledge, the phenotypic evolutionary pathway within BE.

Methods


Patients

Patients were recruited from the surveillance BE endoscopic clinic at Barts Health National Health Service Trust and from the archives of both the Royal London Hospital and University College London Hospital approved under multicenter ethical approval from London research ethics committee (11/LO/1613 and 15/LO/2127). Postesophagectomy BE esophagus specimens were accessed under County Durham and Tees Valley 2 research ethics committee approval (08/H0908/25) and from University College London hospital (as described earlier). Snap-frozen biopsy specimens and formalin-fixed paraffin-embedded (FFPE) archival specimens were used in this study.

Cohort 1a. A series of 64 biopsy specimens from 51 BE patients were collected from 1.0–2.0 cm proximal of the

* Authors share co-first authorship.

Abbreviations used in this paper: AP, alkaline phosphatase; BE, Barrett's esophagus; DAB, diaminobenzidine; EA, esophageal adenocarcinoma; EGA, European Genome-Phenome Archive; EMR, endoscopic mucosal resection; FFPE, formalin-fixed, paraffin-embedded; HRP, horseradish peroxidase; IHC, immunohistochemistry; IM, intestinal metaplasia; mtDNA, mitochondrial DNA; NGS, next-generation sequencing; PCR, polymerase chain reaction.

 Most current article

© 2022 The Author(s). Published by Elsevier Inc. on behalf of the AGA Institute. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

0016-5085

<https://doi.org/10.1053/j.gastro.2021.12.271>

gastroesophageal junction and were FFPE preserved. All biopsy specimens met the following inclusion criteria: (1) they were taken at the same anatomic height within the esophagus, regardless of BE maximum length; (2) they were taken from the BE lesion identified during endoscopy; and (3) no dysplasia or cancer was observed at the time of endoscopy or any history of dysplasia. The mean age of the patients within cohort 1 was 62.2 years (range, 27–89 years), the female-to-male ratio was 1:4.9, and the mean maximum BE segment length was 4.5 cm (range, 1.5–14 cm; median, 4.0 cm). For 25 of these patients, we obtained further archival FFPE H&E sections from all biopsy specimens taken at the same surveillance endoscopy. Further anonymized clinical information is detailed in [Supplementary Table 1](#).

Cohort 1b. The cohort included fresh-frozen adjacent biopsy specimens taken from BE, the anatomic gastric cardia, and squamous esophagus of 20 patients. These were used for either lineage tracing or mitochondria next-generation sequencing (NGS). Further anonymized clinical information is detailed in [Supplementary Table 2](#).

Cohort 2. The cohort included 99 FFPE-preserved BE biopsy specimens (19 patients) showing no dysplasia or history of dysplasia, 21 endoscopic mucosal resection (EMR) specimens (18 patients) for high-grade dysplasia with adjacent nondysplastic BE (age range, 43–65 years), and 47 nondysplastic biopsy specimens (12 patients) from patients who eventually progressed to dysplasia (BE predysplasia).

Cohort 3. This cohort included 31 biopsy specimens from 19 patients with postesophagectomy BE (neo-BE). All patient biopsy specimens were taken a minimum of 2 years after esophagectomy for adenocarcinoma. No data were collected for patient age or BE length. All samples were FFPE.

Gland Phenotype Identification by Histology and Immunohistochemistry

Two independent experienced pathologists determined gland phenotype in FFPE and frozen BE tissue sections (M.J. and N.A.W.) by identifying the presence of either chief cells, parietal cells, goblet cells, foveolar cells, or Paneth cells by H&E and by immunohistochemistry (IHC) analysis ([Supplementary Figures 1 and 2](#)).

Serial 5- μ m FFPE tissue sections were dewaxed in xylene and hydrated through a graded ethanol series to water. Antigen retrieval was performed in boiling Tris-EDTA pH 8.0 (Sigma) or sodium citrate pH 6.0 (FisherChemicals) for 10 minutes depending on each primary antibody ([Supplementary Table 3](#)). Endogenous peroxidase activity was blocked with 3% hydrogen peroxide (FisherChemical) for 10 minutes followed by Protein Block, Serum-Free solution (Agilent Technologies Ltd) for 30 minutes. Typically, no endogenous alkaline phosphatase (AP) was detected.

Double IHC was performed in a specific sequence on serial sections of primary antibodies to H⁺K⁺ATPase (horseradish peroxidase [HRP]-3,3'-diaminobenzidine tetra hydrochloride [DAB]) and pepsinogen (AP-blue) (set 1) and for MUC5AC (HRP-DAB) and MUC2 (AP-blue) (set 2) and defensin 5 α (HD5, frozen only) or 6 α (HD6, FFPE only) (AP-blue) (set 3) ([Supplementary Figure 1A–C](#), respectively).

Dilutions were performed in Ready-to-Use diluent (Agilent), and primary antibodies were incubated for 1 hour at room temperature, followed by incubation with either a goat anti-

mouse IgG (Agilent) or swine anti-rabbit (Sigma) at a 1:200 dilution for 45 minutes at room temperature depending on the primary antibody ([Supplementary Table 3](#)). Streptavidin HRP (Agilent) was then added (1:100) and incubated for 30 minutes, and 3,3'-diaminobenzidine tetra hydrochloride peroxidase substrate was added until a brown color developed (Vectro Labs Ltd). This was followed by incubation with a second round of primary antibodies, a secondary-biotinylated antibody, and then a tertiary streptavidin conjugated to AP. Vector blue substrate (Vector Labs Ltd) was then added until a blue color developed. The same protocol was followed for frozen sections with the exception of not performing antigen retrieval, and section thickness was 10 μ m.

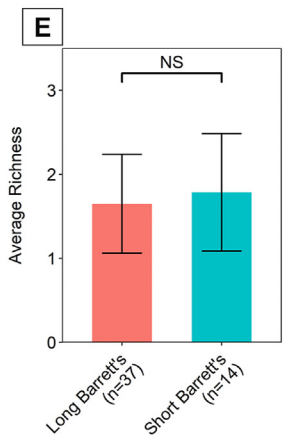
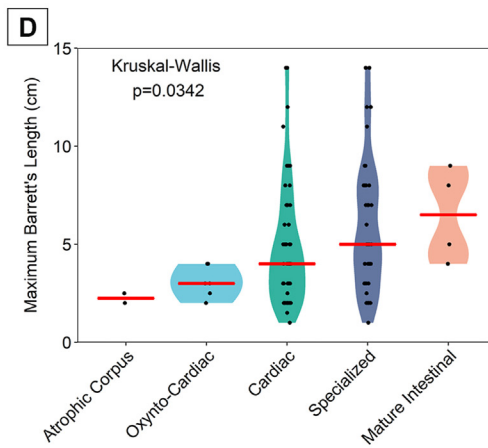
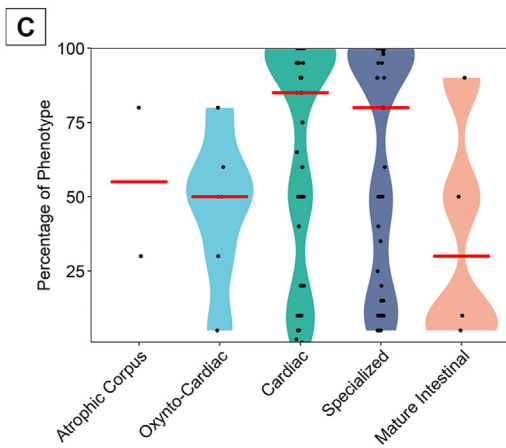
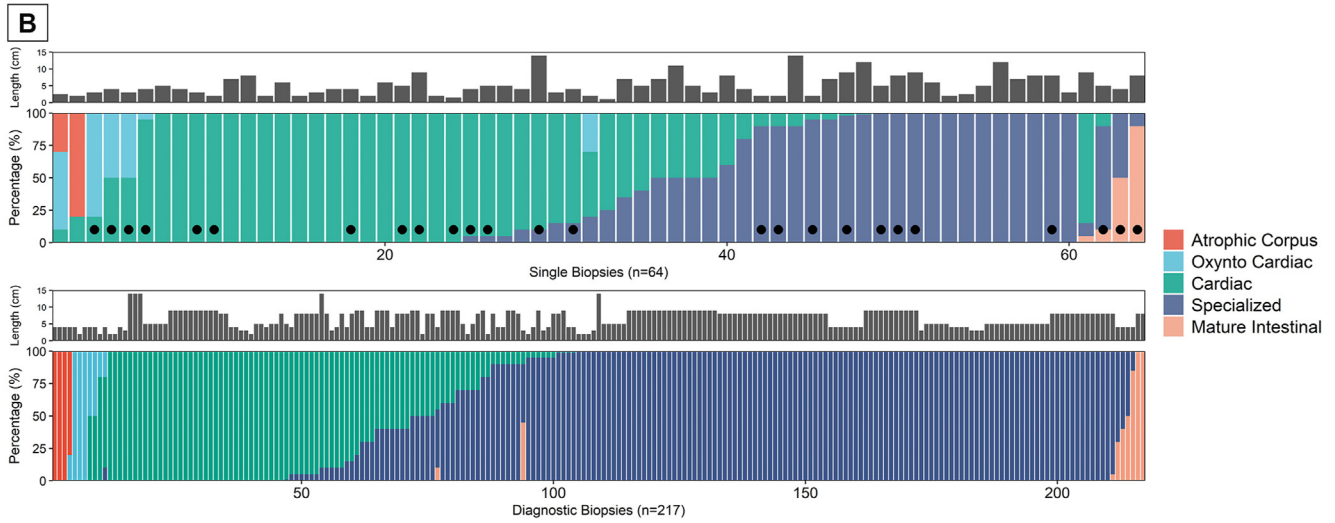
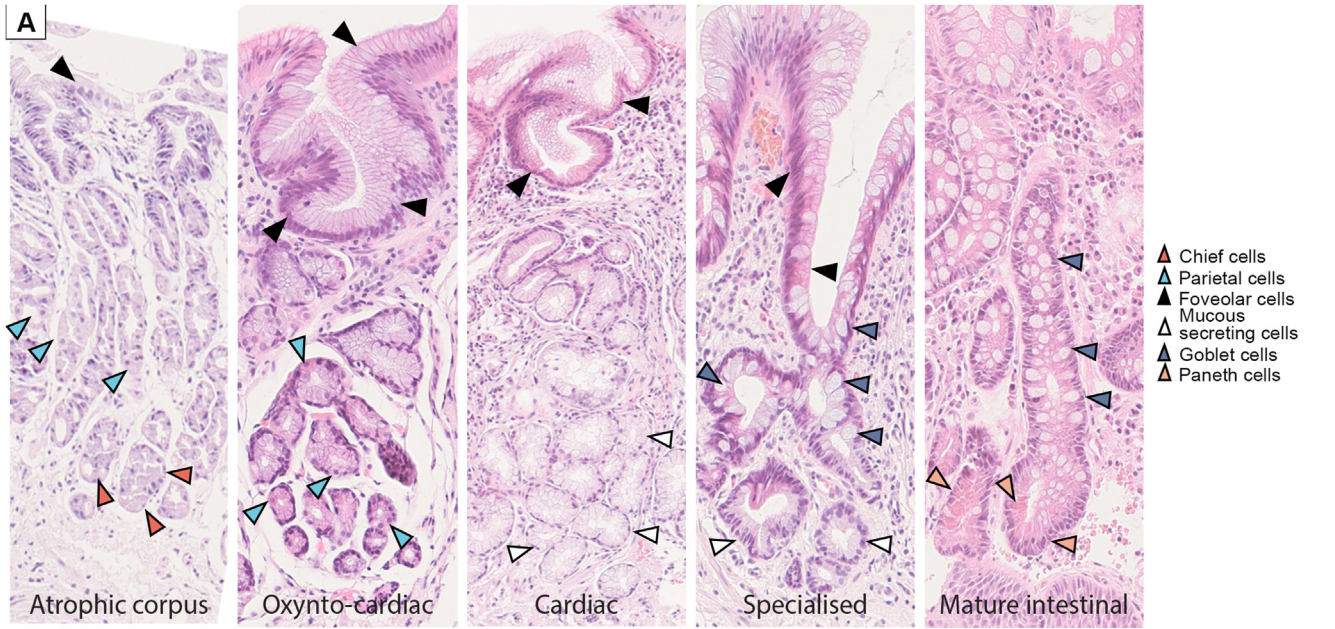
A workflow of gland-specific antibodies is shown in [Supplementary Method Figure 1](#). IHC was used to complement H&E analysis, and we show in [Supplementary Figure 1D and E](#) the efficacy of identifying phenotypically distinct glands in a patient whose biopsy contained both MUC2⁺ MUC5AC⁺ (specialized) and MUC2⁻MUC5AC⁺ (cardiac) glands ([Supplementary Figure 1Di–iii and Ei–iii](#), respectively). Additionally, we show a similar level of distinction by IHC in mature intestinal MUC2⁺MUC5AC⁻ and specialized glands ([Supplementary Figure 2A–Ci](#)) and the distinction between atrophic corpus and oxyntocardiac glands ([Supplementary Figure 2D–F](#)). The addition of MUC6 to HD6 staining in [Supplementary Figure 2C](#) was included purely to highlight gland bases.

Laser-Capture Microdissection

Serial 10- μ m frozen sections were cut onto P.A.L.M. membrane slides (Zeiss) previously treated with UV exposure for 30–40 minutes. To delineate gland outline in frozen material, sections were subjected to dual enzyme histochemistry for cytochrome *c* oxidase and succinate dehydrogenase as per previously published protocols.^{20,22} In all cases, sections were left to dry, and then microdissection was performed using a P.A.L.M laser dissection microscope (Zeiss). Microdissected cells were digested in 14 μ L Picopure digestion buffer (Life Technologies) at 65°C for 3 hours, followed by proteinase K inactivation at 95°C for 5 minutes.

Mitochondrial Polymerase Chain Reaction Sequencing

A nested PCR protocol was used as previously published.²⁰ Briefly, the mitochondrial genome from each microdissected area was amplified into 9 2-kilobase fragments, which were subsequently reamplified into 500-base pair fragments. Primer sequences and PCR conditions were used as previously described ([Supplementary Table 4](#)).²⁰ The second-round PCR primers contained an M13 sequence to facilitate sanger sequencing. PCR products were ExoSAP-treated according to manufacturer's protocol (GE Healthcare) and Sanger sequenced by Eurofins Genomics. Obtained sequences were viewed using 4Peaks software (<https://nucléobite.com>) and compared to the revised Cambridge reference sequence using online tools provided at www.mitomap.org/MITOMAP. Polymorphisms and nonepithelial mutations were eliminated from analysis by comparison with sequences from a microdissected area of stroma. Each mutation was confirmed using the same PCR sequencing protocol repeated from the original DNA sample.



BASIC AND TRANSITIONAL AT

The mitochondrial DNA (mtDNA) NGS methodology is described in the [Supplementary Methods](#). NGS sequencing data have been deposited at the European Genome-Phenome Archive (EGA), which is hosted by the European Bioinformatics Institute Centre for Genomic Regulation, under accession number EGAS00001005729. Further information about the EGA can be found at <https://ega-archive.org>.

Statistics

Statistical analysis was performed using a 1-way analysis of variance (Kruskal-Wallis) test for assessing phenotype distribution. An unpaired Student *t* test or a Mann-Whitney *U* test was used when comparing BE length and diversity, changes in diversity, and changes in phenotype where data were either normally distributed or not, respectively.

Diversity was measured using both a richness score (a record of the number of gland types visible) and the Shannon diversity index¹⁶ that takes into account the number of gland types and their relative frequency within a tissue specimen. The Shannon diversity index was calculated as

$$\text{Shannon diversity index} = - \sum_i^s p_i \ln(p_i),$$

where *s* is the number of species, *p_i* is the frequency of each gland phenotype (*i*) within a tissue specimen. The Shannon diversity index was calculated as the sum of the natural log (ln) of every *p_i* [ln(*p_i*)]. Significance was determined using a 2-sample Mann-Whitney *U* test.

Results

Identification and Distribution of Different Gland Phenotypes in Barrett's Esophagus

Here, we provide a detailed analysis of gland phenotype from cohort 1 of BE patients (described in the Methods section). [Figure 1A](#) shows representative H&E FFPE sections of 5 histologically confirmed phenotypes detected in our BE cohort. We additionally developed a lineage-specific expression profile using IHC to assist in identifying gland phenotype ([Supplementary Methods Figure 1](#)). Gland species were identified as atrophic corpus ($\text{H}^+\text{K}^+\text{ATPase}^+/\text{pepsinogen}^+$), oxyntocardiac ($\text{H}^+\text{K}^+\text{ATPase}^+/\text{pepsinogen}^-$), simple cardiac type (MUC5AC^+ , MUC2^-), specialized BE (MUC5AC^+ , MUC2^+), and mature intestinal (MUC5AC^- , MUC2^+ , HD6^+) ([Supplementary Figures 1 and 2](#)).

Single biopsy specimens (*n* = 64) were taken at 1.0–2.0 cm above the gastroesophageal junction in a cohort of 51 patients, 10 of whom had follow-up biopsies. The

proportion of each gland phenotype within these is shown in [Figure 1B \(top\)](#), and the corresponding diagnostic biopsy specimens taken during the same endoscopy where available (marked with ●) are shown in [Figure 1B \(bottom\)](#) (*n* = 25 patients, 217 biopsy specimens). In the single biopsy cohort, 30 of 64 (46.98%) biopsy specimens displayed 1 phenotype, 30 of 64 (46.98%) displayed 2 phenotypes, and 4 of 64 (6.24%) displayed 3. In the small number of cases where 3 phenotypes were observed within single biopsy specimens, these displayed either atrophic corpus or mature intestinal glands. Within the diagnostic biopsy cohort, 149 of 217 (68.67%) displayed one phenotype, 65 of 217 (29.95%) contained 2, and 3 of 217 (1.38%) had 3. Notably, more single phenotypes were observed than in the single biopsy cohort. The individual patient phenotype proportions based on the location of these biopsies within the BE lesion are presented in [Supplementary Figure 3](#). From this, the previously reported distribution of gland phenotype from proximal to distal ends of the lesion¹⁰ appears to hold, but the distribution of the number of gland phenotypes along this axis was not apparent in the majority of cases. This may be due to cohort differences or overall Barrett segment length. There is a clear dominance of cardiac and specialized gland types from the single biopsy cohort ([Figure 1B \[top\] and C](#)) as well as from the diagnostic biopsy cohort ([Figure 1B \[bottom\]](#)). There was a positive relationship between the overall size of the lesion and the individual gland types detected at 1.0–2.0 cm, where specialized and mature intestinal glands occur more frequently in the longer rather than the shorter BE segments ([Figure 1D](#)). This finding is in line with a previous report.²³ Interestingly, the presence and diversity of specific gland phenotypes from single biopsy specimens at 1.0–2.0 cm was not dependent on BE lesion maximum length ([Figure 1E](#)), nor was it dependent on patient age ([Supplementary Figures 4 and 5](#)). Phenotypic richness was not associated with patient segment length or the number of biopsy samples taken at any single endoscopy as per bootstrapping analysis ([Supplementary Figure 5](#)). Additionally, 10 patients had follow-up over time data available, 5 of which (50%) showed no change in diversity ([Supplementary Figure 6](#)).

Mitochondrial DNA Mutations in Barrett's Esophagus, Anatomic Gastric Cardia, and Squamous Tissue

To determine the concentration and distribution of somatic mtDNA mutations, we performed NGS on laser-capture-microdissected epithelium from frozen sections of

Figure 1. (A) The histologic gland phenotypes observed in BE, from left to right: Atrophic corpus glands containing parietal (▲), chief (▲), and foveolar cells (▲); oxyntocardiac glands containing parietal (▲) and foveolar (▲) but not chief cells; cardiac-type glands containing foveolar (▲) and mucous-secreting cells (▲) only; specialized glands containing goblet (▲), foveolar (▲), and mucous-secreting cells (▲); and mature intestinal glands containing goblet (▲) and Paneth cells (▲). (B, top) Proportion of gland types within single biopsy specimens, taken from 1.0–2.0 cm proximal of the gastroesophageal junction in observable salmon-pink mucosa. (B, bottom) Proportion of gland types in biopsy specimens taken throughout the BE lesion from patients marked with ● in B (top). (C) A summary of the frequency of gland types through all of the single-biopsy cohort. (D) Relationship between specific gland types and the maximum length of the BE lesion. (E) The average phenotypic richness within long vs short BE lesions.

biopsy specimens taken from 10 patients with matched BE, anatomic gastric cardia, and squamous samples (Supplementary Materials and Methods). Single-gland sequencing was not possible because of the concentration of mtDNA being below the sensitivity of the NGS assay. Mutations were found in all samples except the gastric cardia of 1 patient and showed no significant hotspots within the mitochondrial genome of any patient (Supplementary Figure 7A–C). A similar concentration of BE mtDNA mutations was detected compared to other reports.²⁴ A variable mutation count (Supplementary Figure 7D) and proportion of total mutations of each tissue type in each patient (Supplementary Figure 7E) was observed. There were significantly more mutations in the gastric cardia compared to matched BE (Supplementary Figure 7F). No shared somatic mtDNA mutations were observed between any patient samples. Furthermore, there were no significant differences between any BE phenotype and mutation numbers; however, there was a trend with younger patients displaying the fewest mtDNA mutations (Pearson correlation $R^2 = 0.72$, $P = .02$) (Supplementary Figure 7G). This fits with previous studies showing that somatic mtDNA mutations are age dependent.²⁵ These data show the presence of frequent mtDNA mutations in BE independent of gland phenotype.

Evolution of Gland Phenotypes in Barrett's Esophagus

These analyses underscore significant phenotypic heterogeneity, even in biopsy specimens from uncomplicated Barrett segments. Multiple phenotypes in a single biopsy may be a consequence of independent parallel evolution or, alternatively, shared branching evolution. To investigate phenotypic gland evolution in BE, we determined if divergent gland phenotypes within biopsy specimens share a common ancestor. In total, 10 snap-frozen biopsy specimens from cohort 1b that showed at least 2 gland types, as verified by IHC, were subjected to laser capture microdissection mtDNA Sanger sequencing. Glands of each individual phenotype were microdissected. Of these cases, 6 did not show any mtDNA mutations, and 1 demonstrated a mutation that was shared by nonepithelial cells (and was excluded). Of the remaining 3 cases, 1 case (H&E [Figure 2A] and MUC5AC [Figure 2B]) showed the presence of both $H^+K^+-ATPase^+$ parietal cell-containing oxyntocardiac glands (Figure 2C and Ci at high power), MUC5AC⁺MUC2⁺ goblet cell-containing specialized glands, and MUC5AC⁺MUC2⁻ cardiac-type glands (Figure 2B, D, and Di at high power). Cytochrome *c* oxidase staining on an adjacent slide was shown as pre-LCM (Figure 2Cii and Dii), and post-LCM micrographs are shown (Figure 2Ciii and Diii). A shared homoplasmic somatic *m.303-311 Cins* mtDNA mutation in the H-strand replication origin region (*MT-OHR*) (Figure 2E and F) was observed between the entire parietal cell-rich area (glands 1 [surface] and 2 [base]), 2 MUC5AC⁺MUC2⁺ specialized glands (glands 3 and 4), and a single cardiac-type MUC5AC⁺MUC2⁻ gland (gland 5). This mutation was not present in a nearby MUC5AC⁺MUC2⁻

gland (gland 6) and stroma. This indicates the presence of clonal lineages within multiple gland phenotypes and shows, for the first time to our knowledge, a clonal relationship between gastric-like glands (oxyntocardiac and cardiac) and specialized glands in BE.

Furthermore, we investigated an additional frozen biopsy specimen (H&E) (Figure 3A) wherein all glands bar a single gland were MUC2⁺ (Figure 3B). MUC5AC (Figure 3C) staining showed a mixture of positive and negative glands, indicating the presence of cardiac-type, specialized, and mature intestinal glands. LCM of the single MUC2⁻MUC5AC⁺ cardiac-type gland and a neighboring MUC2⁺MUC5AC⁻ mature intestinal gland (cytochrome *c* oxidase-stained before and after LCM) (Figure 3D and E) showed a shared homoplasmic *m.10492 T>C* mutation in the *MT-ND4L* region (Figure 3F) that was not present elsewhere in the biopsy specimen (Figure 3G).

In a third patient, we observed the presence of specialized (HD5⁻MUC5AC⁺ or ¹⁰MUC2⁺) and mature intestinal (HD5⁺MUC5AC⁻MUC2⁺) glands (Figure 4A [H&E], B [HD5], C [MUC5AC], and D [MUC2]) and with high-power images (Figure 4Ei–x). We microdissected 7 glands in total (3 HD5⁺, 4 HD5⁻). We detected 2 homoplasmic somatic mtDNA mutations (*m.2283 T>C* and *m.2217 T>C*), both located in the *MT-RNR2* region (Figure 4Fi–iv) in 2 mature intestinal HD5⁺MUC5AC⁻MUC2⁺ glands (glands 2 and 3) (Figure 4Eiii–v). Additionally, 1 HD5⁻MUC5AC¹⁰MUC2⁺ and 1 HD5⁻MUC5AC⁺MUC2⁺ gland (both considered specialized) also contained both variants (Figure 4Eiii–v). However, a neighboring mature intestinal HD5⁺MUC5AC⁻MUC2⁺ and a distant specialized HD5⁻MUC5AC⁺MUC2⁺ gland (glands 1 and 6, respectively) were both wild type for both variants (Figure Eiii–v and viii–x). These data show that mature intestinal and specialized glands can share a common ancestor. Furthermore, this provides additional evidence of a multiclonal landscape in BE. Taking all 3 patients together, these data show shared common ancestry between phenotypically distinct glands in BE, suggesting an evolutionary process that accounts for the presence of phenotypic diversity.

Phenotypic Diversity in Nondysplastic Barrett's Esophagus Adjacent to Dysplasia

Overall, these data show the diversity and clonal relationship of gland phenotype evolution in nondysplastic BE. Our analysis suggests that phenotypic transitions are bottleneck events. Successive bottlenecks increase the risk of progression. Therefore, we hypothesized that BE associated with dysplasia would show greater evidence of phenotypic diversity. To determine if gland phenotype is altered in BE progression to dysplasia or cancer, we investigated an additional group of patients (cohort 2; see Methods section) that included 19 patients who had no history of dysplasia or cancer (99 biopsy specimens), 12 patients (47 biopsy specimens) with nondysplastic BE taken before a diagnosis of dysplasia (predysplastic BE), 18 patients (33 EMR specimens) with confirmed dysplasia or cancer who also displayed surrounding nondysplastic BE,

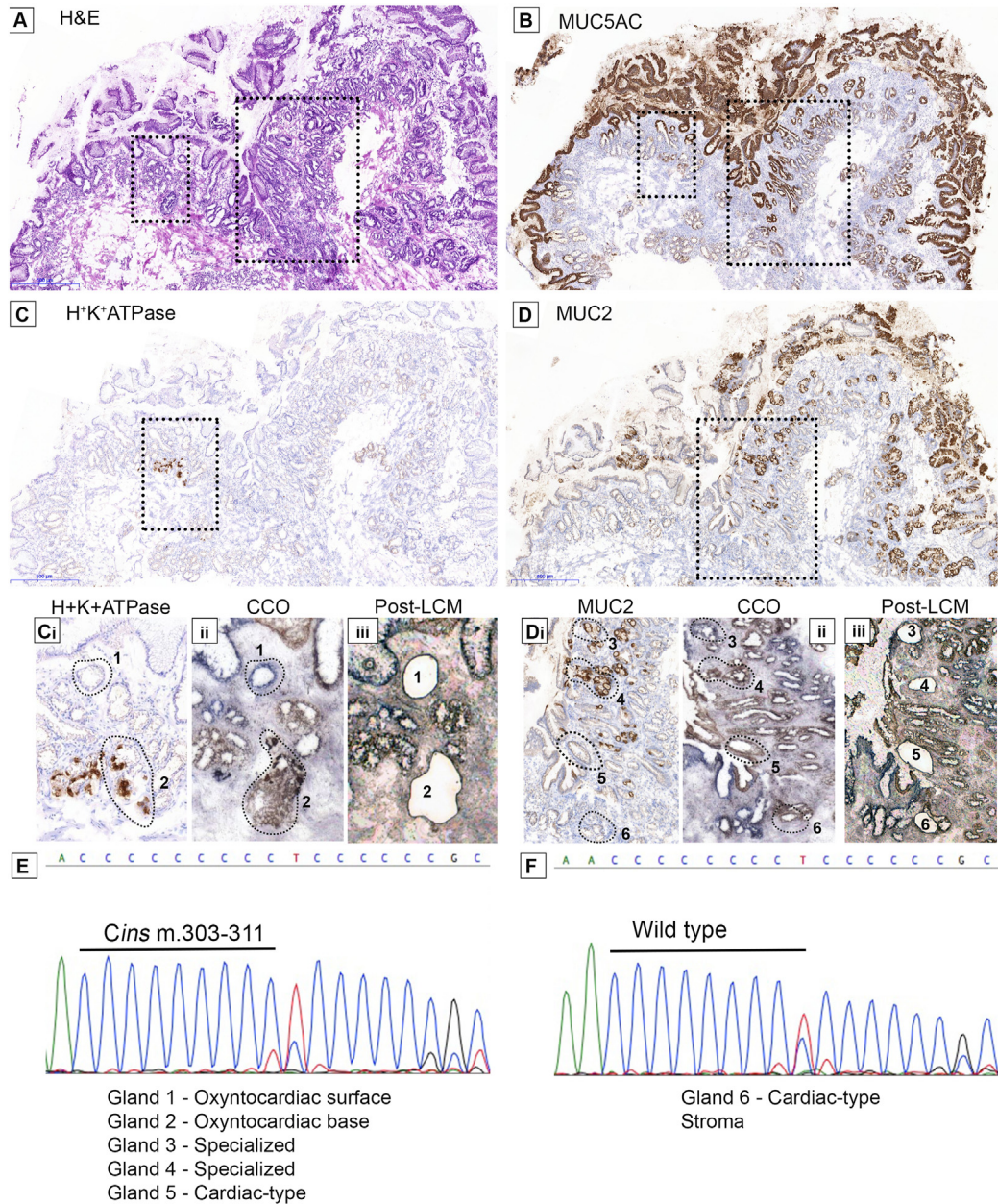


Figure 2. A common ancestry between oxyntocardiac, cardiac-type, and specialized epithelium in BE. (A) H&E with areas of interest outlined. (B) MUC5AC is extensively expressed on all surface foveolar cells. (C) H⁺K⁺ATPase⁺ parietal cell-containing glands and (D) MUC2⁺ goblet cell glands on the same section. (Ci-iii) and (Di-iii) show high-power magnification images of C and D, respectively, with adjacent cytochrome c oxidase staining (Cii, Dii) and postmicrodissection images (Ciii, Diii). (E) PCR sequencing showed a common, complex *m.303-311* *Cins* mutation in the *MT-OHR* region of the mitochondrial genome present in oxyntocardiac MUC5⁺MUC2⁻ surface (1) and H⁺K⁺ATPase⁺ base (2) glands as well as MUC5AC⁺MUC2⁺ specialized glands (3 and 4) and a MUC5AC⁺MUC2⁻ cardiac-type gland (5). (F) The stroma and 1 other cardiac-type MUC5AC⁺MUC2⁻ gland (6) sequenced were wild type. All microdissected glands are traced by dotted outlines.

and 19 patients (31 biopsy specimens) who developed nondysplastic postesophagectomy BE (cohort 3, neo-BE) within 2 years of removal of the lower esophagus.

Figure 5A shows a representative H&E section of an endoscopic resection specimen showing nondysplastic BE (white dashed square) adjacent to BE neoplasia (intramucosal adenocarcinoma, green dashed circle). It is clear from this section that there are multiple gland phenotypes in the nondysplastic BE adjacent to BE neoplasia and at high

power magnification, it is clear that these are distinct glands (Figure 5A, inset). Representative H&E images from patients with neo-BE are shown in Figure 5B and C. We then compared the phenotypic proportions of BE, BE pre-dysplasia, BE adjacent to dysplasia, and neo-BE using a biopsy-by-biopsy approach (Figure 5Di). We calculated the Shannon diversity index (see Methods section) based on the percentage of glands of a given phenotype within each specimen and within each of the 4 groups (Figure 5Dii).

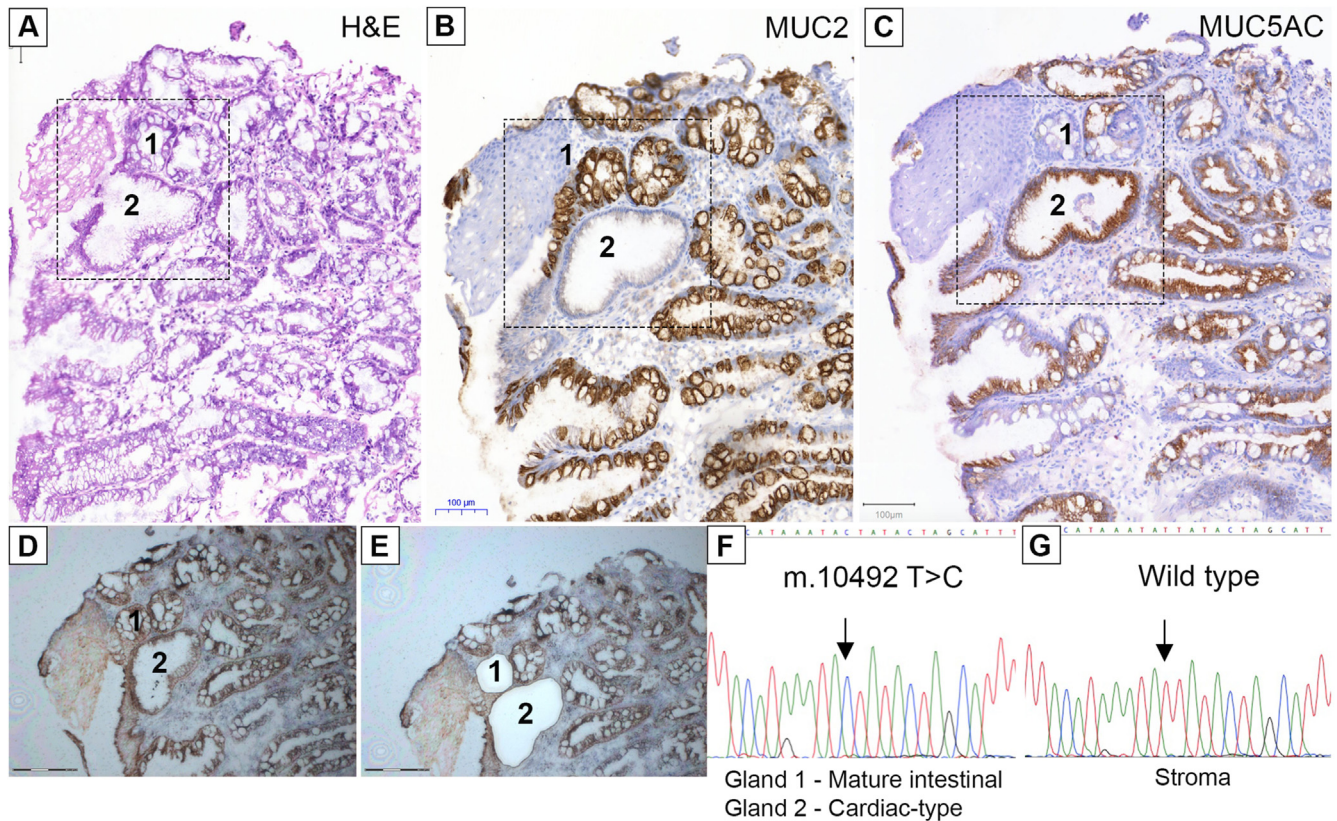


Figure 3. A common ancestry between cardiac-type and mature intestinal epithelium. (A) H&E of a BE biopsy specimen containing both cardiac and mature intestinal epithelium (outlined region). (B) MUC2⁺ and MUC2⁻ and (C) MUC5AC⁺, MUC5AC⁰, and MUC5AC⁻ glands are present. (D, E) Pre- and post-LCM images, respectively. Gland 1 is a MUC2⁺MUC5AC⁻ mature intestinal gland. Gland 2 is a MUC2⁻MUC5AC⁺ cardiac-type gland. (F, G) Glands 1 and 2 showed a common *m.10492T>C* mutation in the *MT-ND4L* region of the mitochondrial genome not present in stroma and other glands.

Interestingly, we find that BE predysplasia and BE adjacent to dysplasia is significantly more phenotypically diverse compared to BE biopsy specimens from patients without dysplasia or those with neo-BE (Figure 5Dii). Individual biopsy specimens capture a smaller mucosal surface area compared to shoulder regions of EMR specimens, which may underestimate diversity in the biopsy group. To address this, bootstrapping analysis was performed on all the biopsy specimens in our cohort (Supplementary Materials and Methods), which showed that randomly sampling a larger number of glands from these biopsy specimens would not have altered the outcome of the Shannon diversity index (Supplementary Figure 8).

Neo-BE biopsy specimens showed no per-biopsy diversity, and the dominant phenotype observed was cardiac ($n = 22/31$ biopsy specimens in 13/19 patients), as has been observed before.²⁶ Specialized and atrophic corpus-type glands were identified in a minority of patients ($n = 5/31$ biopsy specimens in 4/19 patients and $n = 4/31$ biopsy specimens in 4/19 patients, respectively). Two neo-BE patients showed more than 1 phenotype (Supplementary Figure 9) but in distinct biopsy specimens only. A patient-per-patient biopsy phenotype distribution for all 4 groups is shown in Supplementary Figure 10. These data show that areas of nondysplastic BE adjacent to dysplasia and BE predysplasia are significantly more diverse than BE or neo-

BE, indicating that diversity develops before the onset of dysplasia. This suggests that phenotypic diversity could act as a potential predictive biomarker for progression in surveillance biopsies.

Discussion

The study of gland phenotype in BE has been largely overlooked.¹⁰⁻¹² Considering that all BE diagnoses are based on phenotypic observations, there is an unmet need to determine the range, evolution, mechanism, and diversity of these phenotypes. Our cohorts appear to reflect those included in previous studies by identifying similar gland types,¹⁰ but with a much broader interpretation of phenotypes and their diversity. The ancestral relationships we uncover between gland phenotypes provide important information about the evolution of BE to cancer. We show that evolution of BE can involve a change of phenotype at the stem cell level (niche succession^{27,28}). Importantly, we show that a mature gastric gland phenotype expressing parietal cells shares a common ancestor with specialized BE, indicating a phenotypic adaptation of BE along a gastric differentiation, presumably as a result of natural selection acting on gastric phenotypes. We show that multiple phenotypic changes represent clonal phenotypic evolution within BE and that an increase in their diversity is associated with progression. However, de novo postesophagectomy BE does

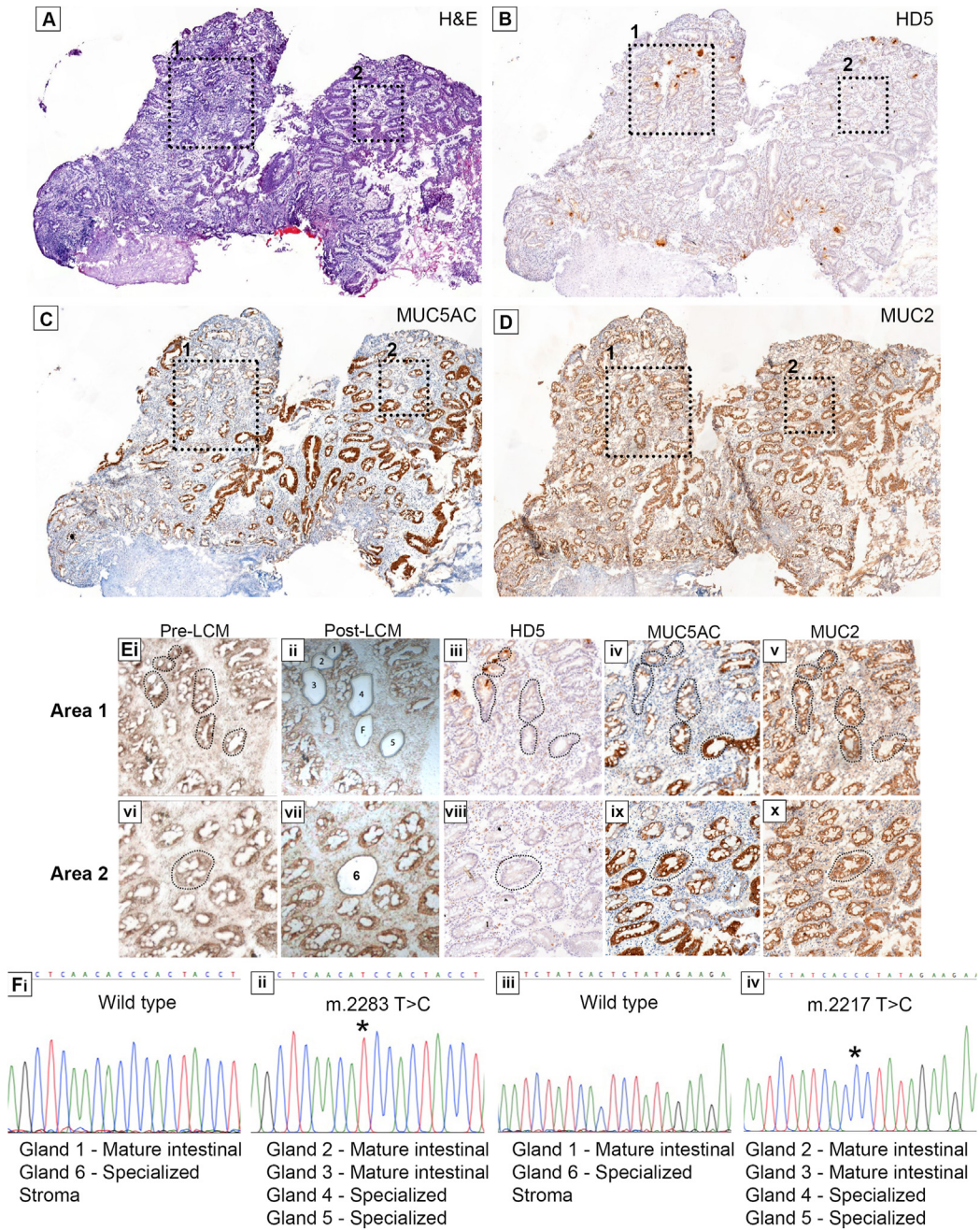


Figure 4. A common ancestry of specialized and mature intestinal glands in BE. (A) H&E of a biopsy containing both gland phenotypes confirmed by IHC for (B) HD5⁺ and HD5⁻ glands, (C) MUC5AC^{+/-} and MUC5AC⁻ glands, and (D) MUC2⁺ glands. Dash-lined boxes indicate 2 areas of interest. (Ei-x) Show the 2 areas of interest at higher power, with pre- and post-LCM images (dotted outlines) as well as mucin and defensin staining of each area. Gland marked F (Eii) failed amplification. (Fi-iv) Two somatic mtDNA mutations (m.2283 T>C and m.2217 T>C) in HD5⁺MUC5AC⁻MUC2⁺ mature intestinal glands (2 and 3) and in 2 specialized glands; 1 being HD5⁺MUC5AC^{lo}MUC2⁺ and the other HD5⁻MUC5AC⁺MUC2⁺ (glands 4 and 5, respectively). The variants were not present in a neighboring HD5⁺MUC5AC⁻MUC2⁺ mature intestinal gland (1) and a distant HD5⁻MUC5AC⁺MUC2⁺ specialized gland (6) and the surrounding stroma. Mutations are marked as *.

not show any within-biopsy diversity, and this suggests that gland phenotype diversity is acquired over time and may not be present during the early development of BE.

We observe 5 well-defined gland phenotypes in our cohort of biopsy specimens taken at the same anatomic location within the esophagus (1.0–2.0 cm proximal of the gastric folds). Of note, we did not observe any pancreatic metaplasia in any of our cohorts.²⁹ Previous reports have

suggested that there is a distribution of gland type within the length of the BE segment.¹⁰ Although we observe a relationship between the proportions of individual gland types present at 1.0–2.0 cm and the overall size of the BE lesion, diversity of gland type is not dependent on lesion size, and cases where more than 2 phenotypes are present in the same biopsy specimen are rare. This is confirmed when we investigate all biopsy specimens taken at each

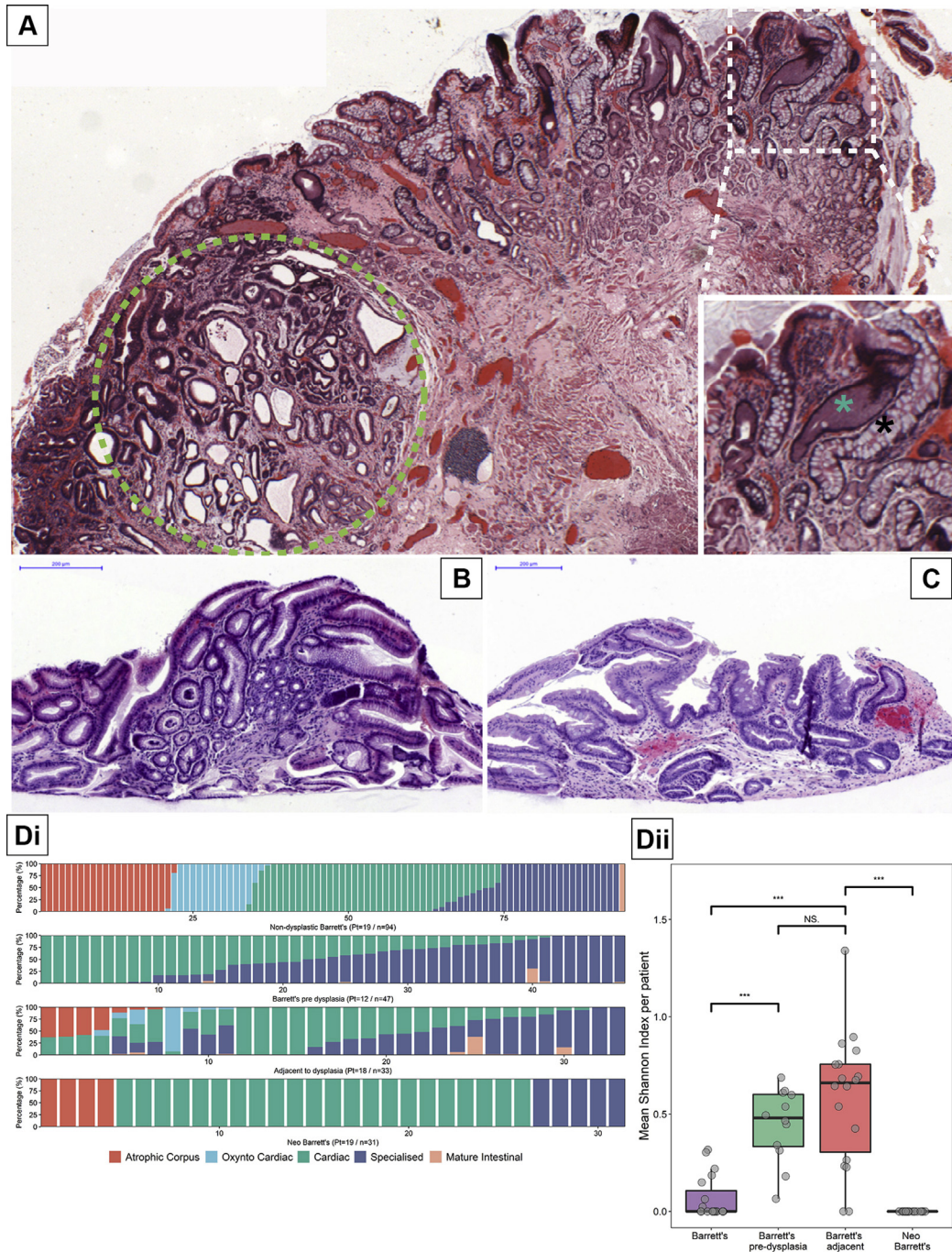


Figure 5. BE gland phenotype diversity is associated with dysplasia. (A) An endoscopic mucosal resection H&E showing an area of high-grade dysplasia (green dashed circle) with surrounding nondysplastic BE showing the presence of multiple gland phenotypes (white dashed box). The inset shows a distinct cardiac-type gland (*) and a mature intestinal gland (*) adjacent to each other. (B) and (C) display postesophagectomy BE (neo-BE) biopsy specimens showing cardiac-type epithelium and specialized epithelium, respectively. (Di) The percentage of each phenotype observed in 94 biopsy specimens (from 19 patients) with BE and no history of dysplasia, 47 biopsy specimens (from 12 patients) with nondysplastic BE before developing dysplasia, 33 specimens of EMRs (from 18 patients) with BE adjacent to dysplasia, and 31 neo-BE biopsy specimens (from 19 patients). (Dii) The mean Shannon diversity index per patient for each patient set from Di (***) = $P < 0.001$, NS = not significant).

particular endoscopy and observe no change in gland phenotype diversity between the single biopsy set (Figure 1B, top) compared with the diagnostic biopsy set (Figure 1B, bottom). The distribution of gland phenotype is

likely to be determined by the surrounding microenvironment.

Here, we report that gland phenotype can evolve and show that distinct gland types, each with a specific

combination of differentiated epithelial cells, can share common ancestry. In particular, we show that cardiac-type and mature intestinal glands in direct proximity to each other have a common mtDNA mutation that can only be explained by the mutation arising in an ancestral gland and being passed to its daughter glands as it divides by gland fission.³⁰ The odds of 2 glands with distinct epithelial phenotypes possessing the same homoplasmic or highly heteroplasmic mtDNA mutation independently is vanishingly small.²⁰ It is, therefore, highly likely that gland fission is the mechanism of clonal expansion within the metaplastic esophagus as it is in the normal, unperturbed gastrointestinal tract.²⁰ Importantly, we observed a shared somatic mtDNA mutation between gastric-type glands (oxyntocardiac and cardiac) and specialized glands, showing that intestinal and mature gastric lineages can share ancestry. Although glandular differentiation within the BE segment is traditionally depicted as one of increasing intestinalization per se, these data show that metaplastic glands can also follow a gastric line of differentiation (Figure 2). Individual patient characteristics such as smoking, obesity, or age may influence gland phenotype selection. Although we do not have sufficient cases to infer a relationship between smoking or obesity, we have shown that age is independent of phenotype (Supplementary Figures 4 and 7G).

Somatic mtDNA mutations are not the sole means to experimentally identify clonal relationships. We and others have previously used somatic genomic mutations to demonstrate clonality in BE and intestinal metaplasia of the stomach.³¹⁻³³ However, the level of interpatient heterogeneity of specific genomic mutations and the broad range of genes mutated³⁴ made detecting such mutations on a gland-by-gland approach for our study difficult. MtDNA mutations are common (Supplementary Figure 7), particularly considering the size of their genome; however, they do not provide any information on the likelihood of progression to cancer because of their neutral impact on cell behavior.²⁸ This excludes any inference that the mtDNA mutation is playing a role in the observed phenotypic evolution in our cases. Determination of clonality will be dependent on the clonal markers used and may result in discordance between mitochondrial and genomic analyses of such. Regardless, this does not compromise the phenotypic clonal relationships presented here, but it may result in an underestimation of clone size. Next-generation whole-genome sequencing could address any potential discordance; however, mtDNA Sanger sequencing was deemed a more practical method of clonal analysis when coupled with small amounts of LCM material.

Our clonal analysis data suggest that the cardiac-type gland is the fulcrum on which all other gland types observed in BE are based. Cardiac-type glands are characterized as simple glands, containing only foveolar cells at their surface and mucous-secreting cells at their base. It has been suggested that the presence of non-goblet cell-containing columnar-lined esophagus is associated with shorter lengths of BE and a lower cancer risk.³⁵ However, we have previously shown that columnar-lined esophagus can evolve to cancer and can contain oncogenic driver mutations in genes such as

TP53.⁷ We therefore propose that evolution of cardiac-type glands is the initial basis of progression within BE.

The presence of multiple gland phenotypes across the BE segment may yield important information on the risk of progression to dysplasia or cancer. In terms of physical size, BE changes very little, if at all, over time.³⁶ This belies the rate of clonal evolution within the BE segment itself, in particular those patients who are at risk of developing cancer.¹⁶ Somatic genetic alterations and their diversity in particular are increased in BE before the onset of cancer.¹⁵⁻¹⁷ Controversy surrounds the evolution of genetic diversity, with some studies showing that this occurs only 2-4 years before the onset of cancer,¹⁷ whereas others indicate that BE from patients who progress always has increased genetic diversity.¹⁵ This applies to multiple cancer types.³⁷ Here, phenotypic diversity is shown to be increased in the nondysplastic areas of BE surrounding dysplasia. We considered that analyzing biopsy specimens may underestimate SIs. Bootstrapping our sampling data, however, showed that a stable Shannon diversity index was achieved with the number of samples analyzed. Although our data support the hypothesis that increased diversity of metaplastic phenotype adjacent to dysplasia reflects overall lesion diversity (as supported by our extensive data sampling of nondysplastic biopsy specimens in Figure 1), we were not able to collect sufficient specimen material to represent the entire length of the Barrett lesion in those patients who had undergone EMR. Although our data cannot completely exclude the possibility that the presence of a dysplastic lesion drove the increased gland diversity in its neighboring mucosa, we consider this less likely because our data show that biopsy specimens taken before the onset of dysplasia also showed a diverse epithelial landscape, suggesting that phenotypic diversity arose before the onset of dysplasia. Indeed, data from the patients with neo-BE suggest that no phenotypic diversity is present in recently developed BE and that this is therefore acquired rather than inherent. Furthermore, in cases that progress to dysplasia, maximal genetic diversity has been shown to occur toward the gastroesophageal junction.^{38,39} Our samples taken in this same region also show phenotypic diversity regardless of segment length. Phenotypic diversity in BE and other conditions⁴⁰ in the progression to cancer is understudied, and when we consider that tissue diagnoses rely on phenotypic analysis, it is surprising that diversity is not investigated more often.

Overall, this suggests that BE gland phenotypes are diverse; the presence of individual phenotypes are not related to the size of the lesion and can be observed simultaneously in the same location within the esophagus. Our data show definitively that BE phenotypes represent an evolutionary process and suggest that diversity of gland phenotype may play a role in progression.

Supplementary Material

Note: To access the supplementary material accompanying this article, visit the online version of *Gastroenterology* at www.gastrojournal.org, and at <https://doi.org/10.1053/j.gastro.2021.12.271>.

References

1. Pophali P, Halland M. Barrett's oesophagus: diagnosis and management. *BMJ* 2016;353:i2373.
2. Shaheen NJ, Falk GW, Iyer PG, et al. ACG clinical guideline: diagnosis and management of Barrett's esophagus. *Am J Gastroenterol* 2016;111:30–50.
3. Fitzgerald RC, di Pietro M, Ragnath K, et al. British Society of Gastroenterology guidelines on the diagnosis and management of Barrett's oesophagus. *Gut* 2014;63:7–42.
4. Bhat S, Coleman HG, Yousef F, et al. Risk of malignant progression in Barrett's esophagus patients: results from a large population-based study. *J Natl Cancer Inst* 2011;103:1049–1057.
5. Kelty CJ, Gough MD, Van Wyk Q, et al. Barrett's oesophagus: intestinal metaplasia is not essential for cancer risk. *Scand J Gastroenterol* 2009;42:1271–1274.
6. Srivastava A, Golden KL, Sanchez CA, et al. High goblet cell count is inversely associated with ploidy abnormalities and risk of adenocarcinoma in Barrett's esophagus. *PLoS One* 2015;10(7):e0133403.
7. Lavery DL, Martinez P, Gay LJ, et al. Evolution of oesophageal adenocarcinoma from metaplastic columnar epithelium without goblet cells in Barrett's oesophagus. *Gut* 2016;65:907–913.
8. McDonald SAC, Lavery D, Wright NA, et al. Barrett oesophagus: lessons on its origins from the lesion itself. *Nat Rev Gastroenterol Hepatol* 2015;12:50–60.
9. Rhee H, Wang DH. Cellular origins of Barrett's esophagus: the search continues. *Curr Gastroenterol Rep* 2018;20(11):51.
10. Going JJ, Fletcher-Monaghan AJ, Neilson L, et al. Zoning of mucosal phenotype, dysplasia, and telomerase activity measured by telomerase repeat assay protocol in Barrett's esophagus. *Neoplasia* 2004;6:85–92.
11. Paull A, Trier JS, Dalton MD, et al. The histologic spectrum of Barrett's esophagus. *N Engl J Med* 1976;295:476–480.
12. Chandrasoma PT, Der R, Dalton P, et al. Distribution and significance of epithelial types in columnar-lined esophagus. *Am J Surg Pathol* 2001;25:1188–1193.
13. Quante M, Graham TA, Jansen M. Insights into the pathophysiology of esophageal adenocarcinoma. *Gastroenterology* 2018;154:406–420.
14. Lavery DL, Nicholson AM, Poulosom R, et al. The stem cell organisation, and the proliferative and gene expression profile of Barrett's epithelium, replicates pyloric-type gastric glands. *Gut* 2014;63:1854–1863.
15. Martinez P, Timmer MR, Lau CT, et al. Dynamic clonal equilibrium and predetermined cancer risk in Barrett's oesophagus. *Nat Commun* 2016;7:12158.
16. Maley CC, Galipeau PC, Finley JC, et al. Genetic clonal diversity predicts progression to esophageal adenocarcinoma. *Nat Genet* 2006;38:468–473.
17. Li X, Galipeau PC, Paulson TG, et al. Temporal and spatial evolution of somatic chromosomal alterations: a case-cohort study of Barrett's esophagus. *Cancer Prev Res (Phila.)* 2014;7:114–127.
18. Reid BJ, Paulson TG, Li X. Genetic insights in Barrett's esophagus and esophageal adenocarcinoma. *Gastroenterology* 2015;149:1142–1152.
19. McDonald SAC, Graham TA, Lavery DL, et al. The Barrett's gland in phenotype space. *Cell Mol Gastroenterol Hepatol* 2014;1:41–54.
20. Greaves LC, Preston SL, Tadrous PJ, et al. Mitochondrial DNA mutations are established in human colonic stem cells, and mutated clones expand by crypt fission. *Proc Natl Acad Sci U S A* 2006;103:714–719.
21. Gutierrez-Gonzalez L, Deheragoda M, Elia G, et al. Analysis of the clonal architecture of the human small intestinal epithelium establishes a common stem cell for all lineages and reveals a mechanism for the fixation and spread of mutations. *J Pathol* 2009;217:489–496.
22. McDonald SAC, Greaves LC, Gutierrez-Gonzalez L, et al. Mechanisms of field cancerization in the human stomach: the expansion and spread of mutated gastric stem cells. *Gastroenterology* 2008;134:500–510.
23. Harrison R, Perry I, Haddadin W, et al. Detection of intestinal metaplasia in Barrett's esophagus: an observational comparator study suggests the need for a minimum of eight biopsies. *Am J Gastroenterol* 2007;102:1154–1161.
24. Lee S, Lee S, Han M-J, et al. Frequent occurrence of mitochondrial DNA mutations of mitochondrial DNA mutations in Barrett's metaplasia without the presence of dysplasia. *PLoS One* 2012;7(5):e37571.
25. Taylor RW, Barron MJ, Borthwick GM, et al. Mitochondrial DNA mutations in human colonic crypt stem cells. *J Clin Invest* 2003;112:1351–1360.
26. Dunn LJ, Shenfine J, Griffin SM. Columnar metaplasia in the esophageal remnant after esophagectomy: a systematic review. *Dis Esophagus* 2015;28:32–41.
27. Shibata D. Inferring human stem cell behaviour from epigenetic drift. *J Pathol* 2009;217:199–205.
28. Baker A-M, Cereser B, Melton S, et al. Quantification of crypt and stem cell evolution in the normal and neoplastic human colon. *Cell Rep* 2014;8:940–947.
29. Krishnamurthy S, Dayal Y. Pancreatic metaplasia in Barrett's esophagus. An immunohistochemical study. *Am J Surg Pathol* 1995;19:1172–1180.
30. Nicholson AM, Graham TA, Simpson A, et al. Barrett's metaplasia glands are clonal, contain multiple stem cells and share a common squamous progenitor. *Gut* 2012;61:1380–1389.
31. Leedham SJ, Preston SL, McDonald SAC, et al. Individual crypt genetic heterogeneity and the origin of metaplastic glandular epithelium in human Barrett's oesophagus. *Gut* 2008;57:1041–1048.
32. Ross-Innes CS, Becq J, Warren A, et al. Whole-genome sequencing provides new insights into the clonal architecture of Barrett's esophagus and esophageal adenocarcinoma. *Nat Genet* 2015;47:1038–1046.
33. Gutierrez-Gonzalez L, Graham TA, Rodriguez-Justo M, et al. The clonal origins of dysplasia from intestinal metaplasia in the human stomach. *Gastroenterology* 2011;140:1251–1260.

34. Weaver JM, Ross-Innes CS, Shannon N, et al. Ordering of mutations in preinvasive disease stages of esophageal carcinogenesis. *Nat Genet* 2014;46:837–843.
35. Desai TK, Krishnan K, Samala N, et al. The incidence of oesophageal adenocarcinoma in non-dysplastic Barrett's oesophagus: a meta-analysis. *Gut* 2012;61:970–976.
36. Cameron AJ, Lomboy CT. Barrett's esophagus: age, prevalence, and extent of columnar epithelium. *Gastroenterology* 1992;103:1241–1245.
37. Andor N, Graham TA, Jansen M, et al. Pan-cancer analysis of the extent and consequences of intratumor heterogeneity. *Nat Med* 2016;22:105–113.
38. Martinez P, Mallo D, Paulson TG, et al. Evolution of Barrett's esophagus through space and time at single-crypt and whole-biopsy levels. *Nat Commun* 2018;9(1):794.
39. Timmer MR, Martinez P, Lau CT, et al. Derivation of genetic biomarkers for cancer risk stratification in Barrett's oesophagus: a prospective cohort study. *Gut* 2016;65:1602–1610.
40. Almendro V, Kim HJ, Cheng Y-K, et al. Genetic and phenotypic diversity in breast tumor metastases. *Cancer Res* 2014;74:1338–1348.

Received April 20, 2021. Accepted December 22, 2021.

Correspondence

Address correspondence to: Stuart A.C. McDonald, PhD, Centre for Cancer Genomics and Computational Biology, John Vane Science Centre, Barts Cancer Institute, Barts and the London School of Medicine and Dentistry, Queen Mary University of London, Charterhouse Square, London, EC1M 6BQ, United Kingdom. e-mail: s.a.mcdonald@qmul.ac.uk.

Acknowledgments

We would like to thank all the patients that contributed their tissue samples to this study. We would like to thank the Pathology Core Service at the Barts Cancer Institute, Queen Mary University of London, for providing all sectioning of tissues.

The current institution for Lorna Dunn Current is Northumbria Healthcare NHS Foundation Trust, North Tyneside Hospital, North Shields, Tyne and Wear, United Kingdom. The current institution for Ross J. Porter is the Centre for Inflammation Research, Queens Medical Research Institute, University of Edinburgh, United Kingdom. The current institution for Mairi H. McLean is the Division of Molecular and Clinical Medicine, School of Medicine, University of Dundee, United Kingdom.

Conflicts of interest

The authors disclose no conflicts.

Funding

This study was funded through a Guts UK Derek Butler clinical fellowship (to James A. Evans) and a Cancer Research UK Programme Foundation award (to Stuart A.C. McDonald) (A12446). Institute support was provided through a Cancer Research UK Centre grant to the Barts Cancer Centre (C16420/A18066).

Supplementary Methods

Mitochondrial Next-Generation Sequencing

A total of 30 samples collected from 3 different anatomic sites, across 10 patients (Supplementary Figure 7) were included in the study. For each biopsy, we microdissected $2 \times 10^6 \mu\text{m}^2$ of tissue in 0.5 mL UV-irradiated tubes adhesive caps (Zeiss Technologies). DNA was then extracted using the QIAamp DNA extraction kit (Qiagen) according to the manufacturer's instructions with the addition of 1 μL of carrier RNA to increase the yield of DNA. Each sample was eluted in 60 μL of nuclease-free water. Extracted DNA was quantified using the Qubit double-stranded DNA High Sensitivity Assay kit (Life Technologies).

The whole mtDNA genome was analyzed using a PCR strategy by generating 2 overlapping amplicons: MTL-1, 9.1 kilobases, and FRAG1, 11.2 kilobases (primers listed in Supplementary Table 4). To minimize PCR errors, the entire mitochondrial genome was sequenced in duplicate independently from each other. The PCR was performed using the TaKaRa LA TaqDNA polymerase kit (Clontech), and the reaction was performed in a volume of 25 μL containing 1 \times PCR buffer, 0.4 $\mu\text{mol/L}$ primers, 0.4 mmol/L dNTPs, 1.25 units of TaKaRa LA Taq, and 8 μL of template DNA, and the following PCR protocol was used: 94°C for 5 minutes, 33 cycles of 98°C for 15 seconds, 68°C for 10 seconds (slow ramp from 68°C to 60°C at 0.2°C/second) followed by primer annealing at 60°C for 15 seconds and then extension at 68°C for 11 minutes, and a final elongation at 72°C for 10 minutes.

A qualitative and quantitative analysis of the PCR products was performed using a 2200 TapeStation instrument (Life Technologies). All samples were normalized to 0.2 ng/ μL and 5 μL (1 ng) of DNA was used for generating libraries with the NexteraXT DNA Sample Preparation Kit (Illumina) according to the protocol. Libraries were then sequenced via the Illumina MiSeq platform v2, 300 cycles (150 nucleotide paired end). Sequencing data have been deposited at the EGA, which is hosted by the EBI and the CRG, under accession number EGAS00001005729. Further information about EGA can be found at <https://ega-archive.org>.

Mutation Calling for Next-Generation Sequencing

Data were processed using FastQC, and the average read depth was obtained for each independent BE, cardia, or squamous sample. Samples with low mean read depths (approximately <3500) were resequenced. All variant calls not shared between technical replicates were discarded, and single-nucleotide variants (SNVs) with a variant frequency of <1% were also discarded. Reads were aligned to the revised Cambridge Reference Genome (GenBank accession number NC_012920.1) using the Burrow's-Wheeler Aligner - Maximal Exact Match algorithm¹ with a default gap open penalty and gap extension penalty of 6 and 1, respectively. For compatibility with Picard's MarkDuplicates tool,² shorter split hits were marked as secondary during read alignment. Aligned reads were then sorted by coordinate using Picard's SortSam tool, and duplicates were marked

and removed with Picard's MarkDuplicates tool. SNVs were identified using the deepSNV R package³. Bases with a Phred quality score of less than 20 (corresponding to an error probability of $P = 1\%$ in the original base call) were not considered in SNV calling. A Benjamini-Hochberg correction was applied for multiple testing, and SNVs with a P value of less than .05 were considered statistically significant. Rare cases were identified in which sequences (up to approximately 10 base pairs in length) of contiguous mtDNA deletions were detected by deepSNV. These were removed manually from the somatic SNV data on the basis that they are likely to be the result of 1 single deletion, rather than multiple independent SNV events.

Distribution of mitochondrial SNVs was presented as individual Circos plots for each tissue site. Statistical comparison between anatomic sites was performed using a 2-sided paired Student t test.

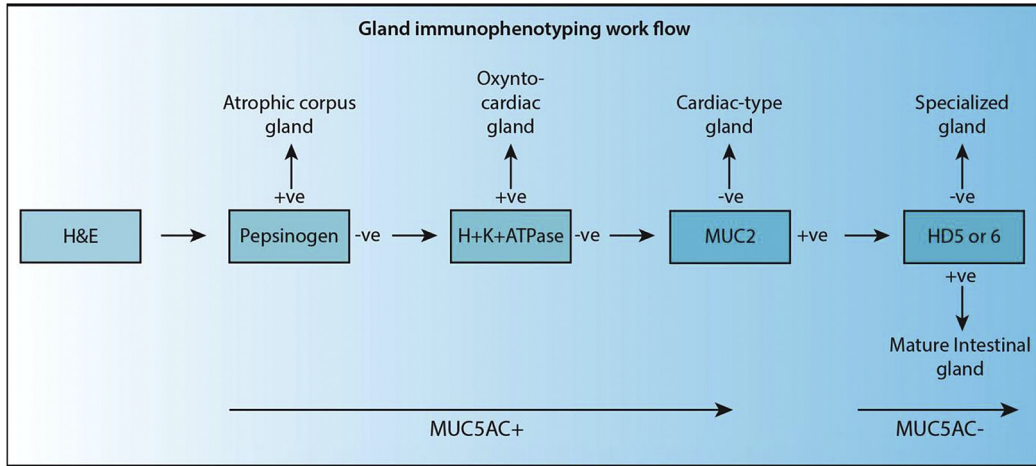
Spearman Correlation of Richness

The Spearman correlation between richness of phenotype data and the number of biopsy specimens was measured using the Hmisc package (version 4.3.0) in R, implementing a bootstrapping strategy. Each patient (index i) has n_i biopsy specimens. A richness score was calculated by the presence of unique phenotypes over the set of biopsy specimens, that is, $R(n_i)$. The richness R was calculated ($R(n_j)$) for patient i , and a random subset of biopsy specimens, ranging from $j = 1, 2, \dots, n_i$. Each random subset obtained is a single bootstrap replicate and was repeated 1000 times. $R(n_j)$ was calculated for each patient on their subsetted biopsy specimens. The Spearman rank correlation coefficient of $R(n_i)$ was calculated between the number of biopsy specimens sampled and the richness score $R(n_j)$ for each of the 1000 replicates for each patient. The final coefficient values across 22 patients (excluding patients with unchanging richness across all biopsy specimens) were ranked from low to high before visualization using the ggplot2 package (version 3.2.1) in R.

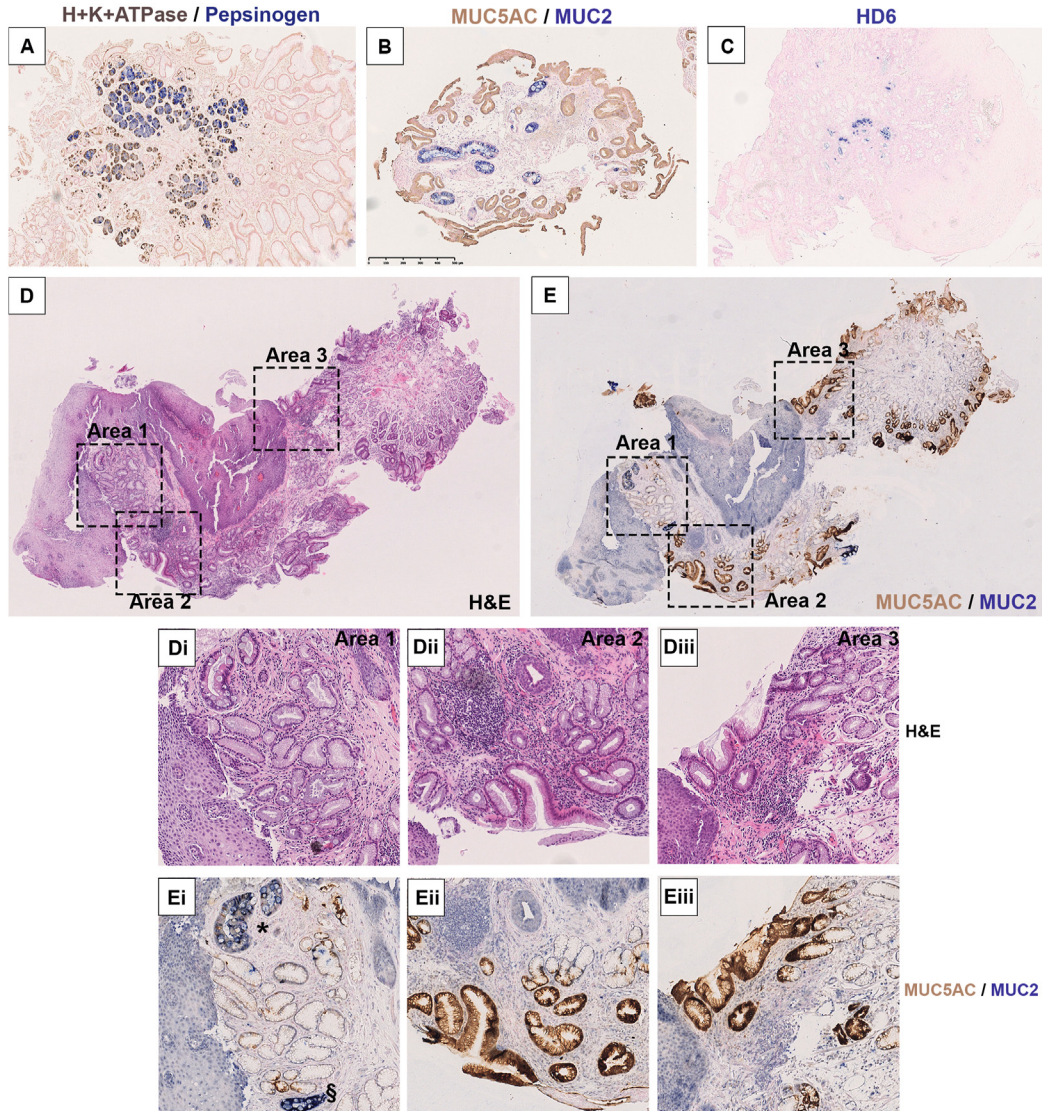
To ensure that sufficient sampling was achieved for diversity per biopsy between nondysplastic BE and BE adjacent to dysplasia, biopsy specimens from each disease state were independently subsampled 1000 times with a random number of glands in each subset (each subset being a bootstrap). Each individual patient's data were plotted separately.

Supplementary References

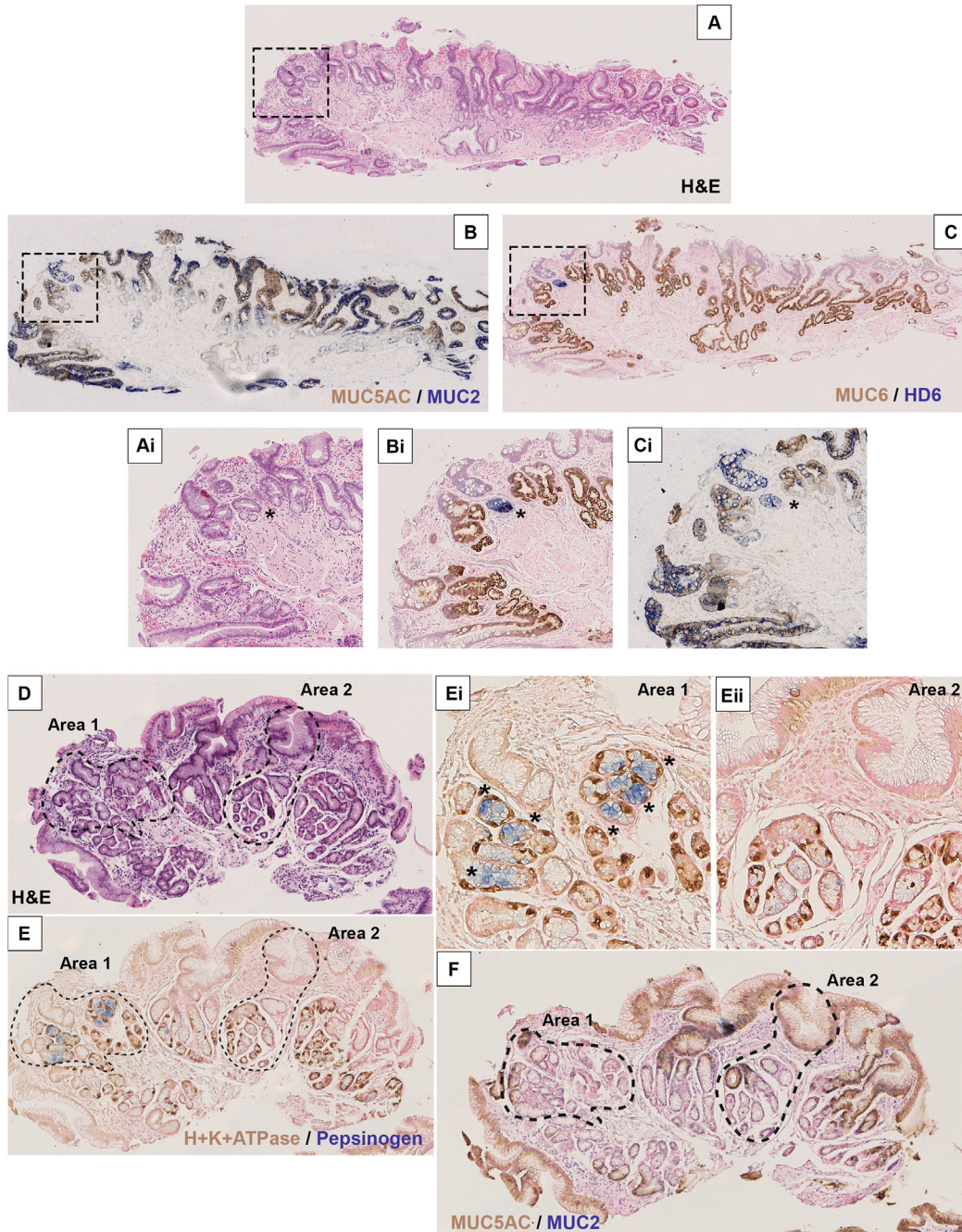
- Li H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. arXiv:1303.3997. Preprint posted online March 16, 2013 (v1); May 26, 2013 (v2).
- Broad Institute. Picard tools 2018 2.6. Available at: <http://broadinstitute.github.io/picard/>. Accessed February 4, 2022.
- Gerstung M, Papaemmanuil E, Campbell PJ. Subclonal variant calling with multiple samples and prior knowledge. *Bioinformatics* 2014;30:1198–1204.



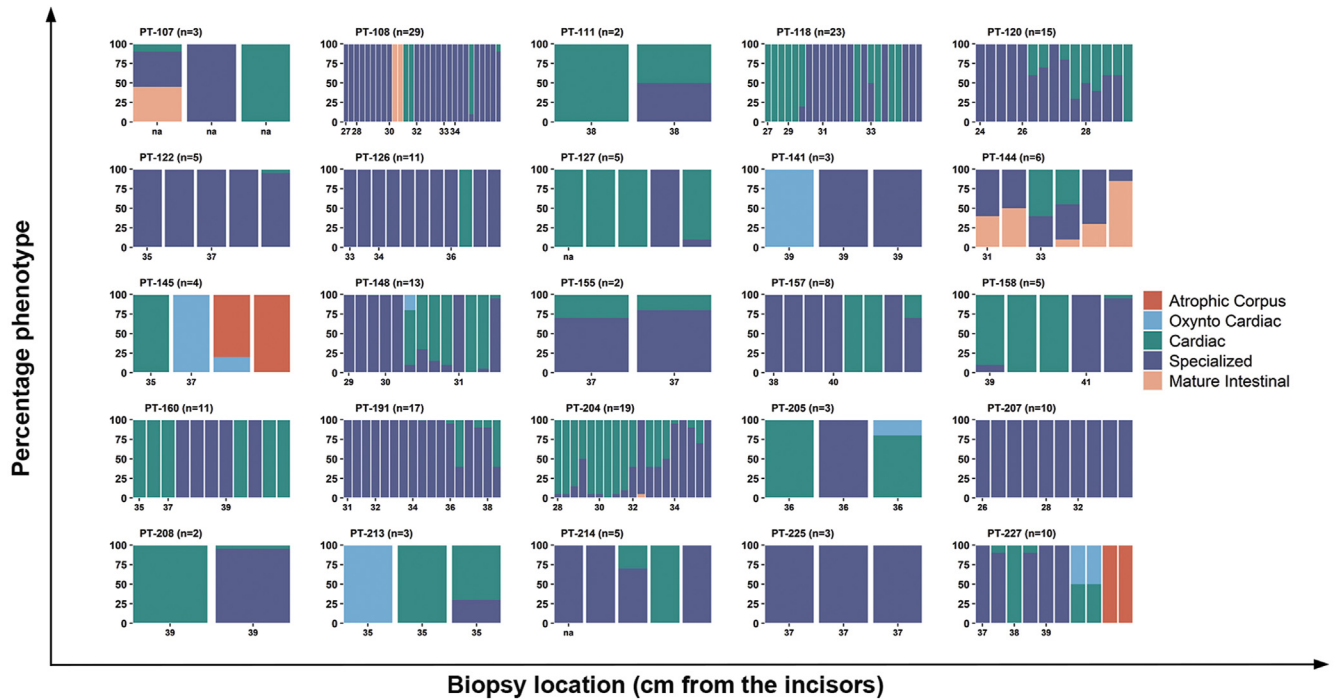
Supplementary Methods Figure 1. H&E and IHC workflow using specific antibodies to identify gland phenotypes. H&E is the primary histologic guide. Only intestinal glands do not express MUC5AC; however, glands transitioning from specialized to a mature intestinal phenotype may have weak MUC5AC (MUC5AC^{lo}) staining and are considered specialized. -ve, negative; +ve, positive.



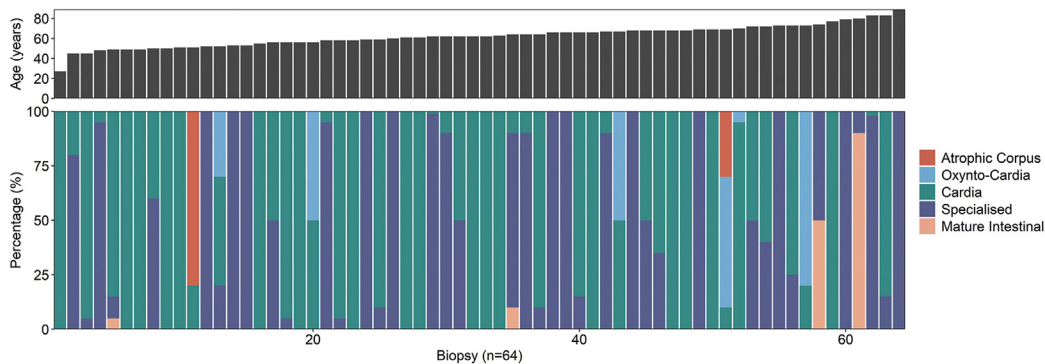
Supplementary Figure 1. Representative IHC for lineage-identifying antibodies. (A) Set 1: H⁺K⁺ATPase (DAB-brown) and pepsinogen (AP-blue). (B) Set 2: MUC5AC (DAB-brown) and MUC2 (AP-blue). (C) Set 3: HD6 (AP-blue). Additionally, we show that distinct gland types are evident through double IHC. (D) An FFPE biopsy specimen that shows both squamous and Barrett's epithelium stained with H&E and (E) a serial section stained with MUC2 and MUC5AC. Three areas are highlighted in higher power (*Di-iii*, *Ei-iii*). Area 1 shows MUC2⁺ with light MUC5AC⁺ cells within the same specialized gland (*) and a MUC2⁺MUC5AC⁻ (§) mature intestinal gland. In areas 2 and 3 (*Dii* and *iii*, *Eii* and *iii*) MUC2⁺ cells are completely absent. The IHC clearly defines specialized and cardiac-type epithelium as well as a likely mature intestinal gland. (HD6 staining is not present to confirm the presence of Paneth cells.)



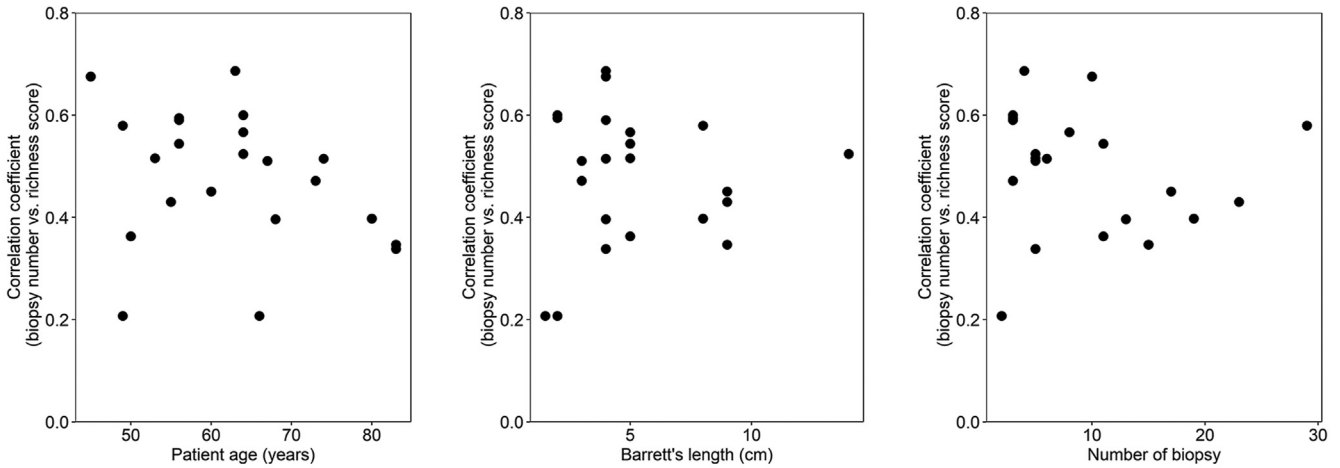
Supplementary Figure 2. Distinct gland types within biopsy specimens. (A) An H&E of a biopsy specimen displaying glandular architecture. (B, C) Double IHC for MUC2 (AP-blue) + MUC5AC (DAB-brown) and HD6 (AP-blue) + MUC6 (brown), respectively. (Ai, Bi, Ci) Higher-power views of a single mature intestinal gland that is HD6⁺ (*), where all other glands are specialized, and HD5⁻ with gland bases is evidenced by MUC6 staining. (D) A second biopsy specimen shows glandular architecture with a gastric phenotype (H&E). (E) Double IHC for H⁺K⁺ATPase (DAB-brown) and pepsinogen (AP-blue) shows 2 areas: area 1 is positive for both markers and is therefore an atrophic corpus phenotype with both parietal and chief cells present (E and Ei), whereas area 2 only shows positivity for H⁺K⁺ATPase parietal cells (DAB-brown) and is therefore oxyntocardiac (E and Eii). (F) MUC5AC (DAB-brown) + MUC2 (AP-blue) IHC did not show the presence of goblet cells in this biopsy.



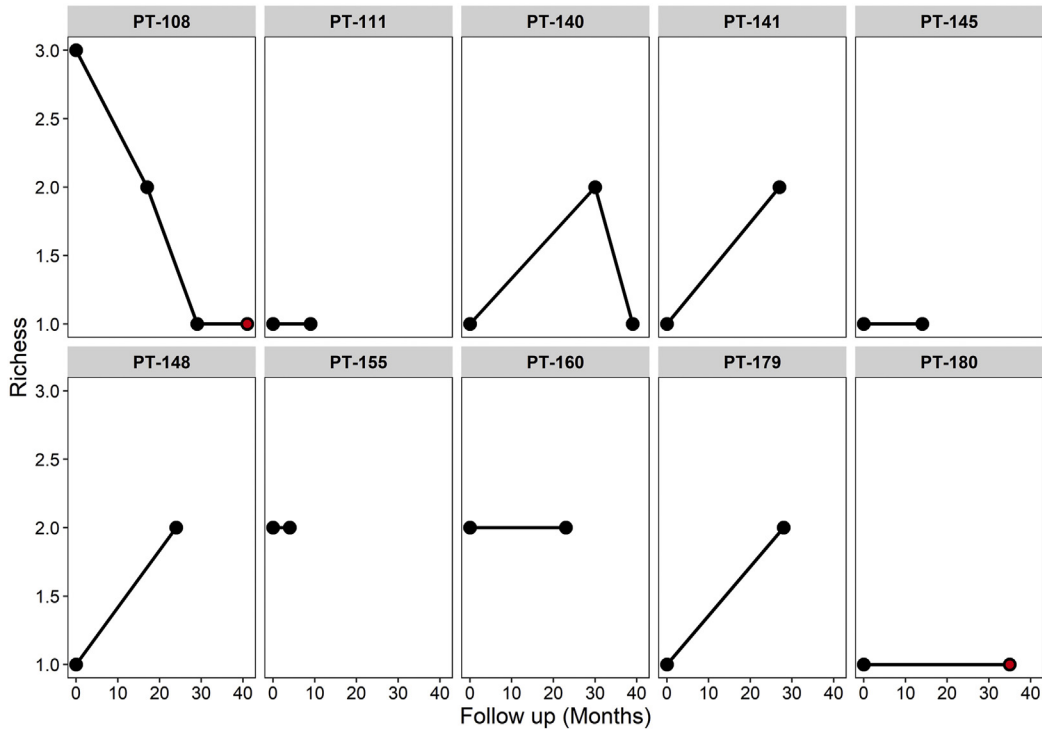
Supplementary Figure 3. Individual patient distributions of gland phenotype based on available diagnostic tissue blocks. All biopsy specimens from a total of 25 patients, from the cohort described in Figure 1B (top), taken at the same endoscopy during which the single biopsy specimen was taken, were phenotyped. H&E sections were reviewed by expert pathologists (M.J. and N.A.W.) and 2 experienced researchers (E.C. and S.A.C.M.). Data are presented as percentages of each phenotype within each biopsy specimen at the specific depth in the esophagus at which it was taken.



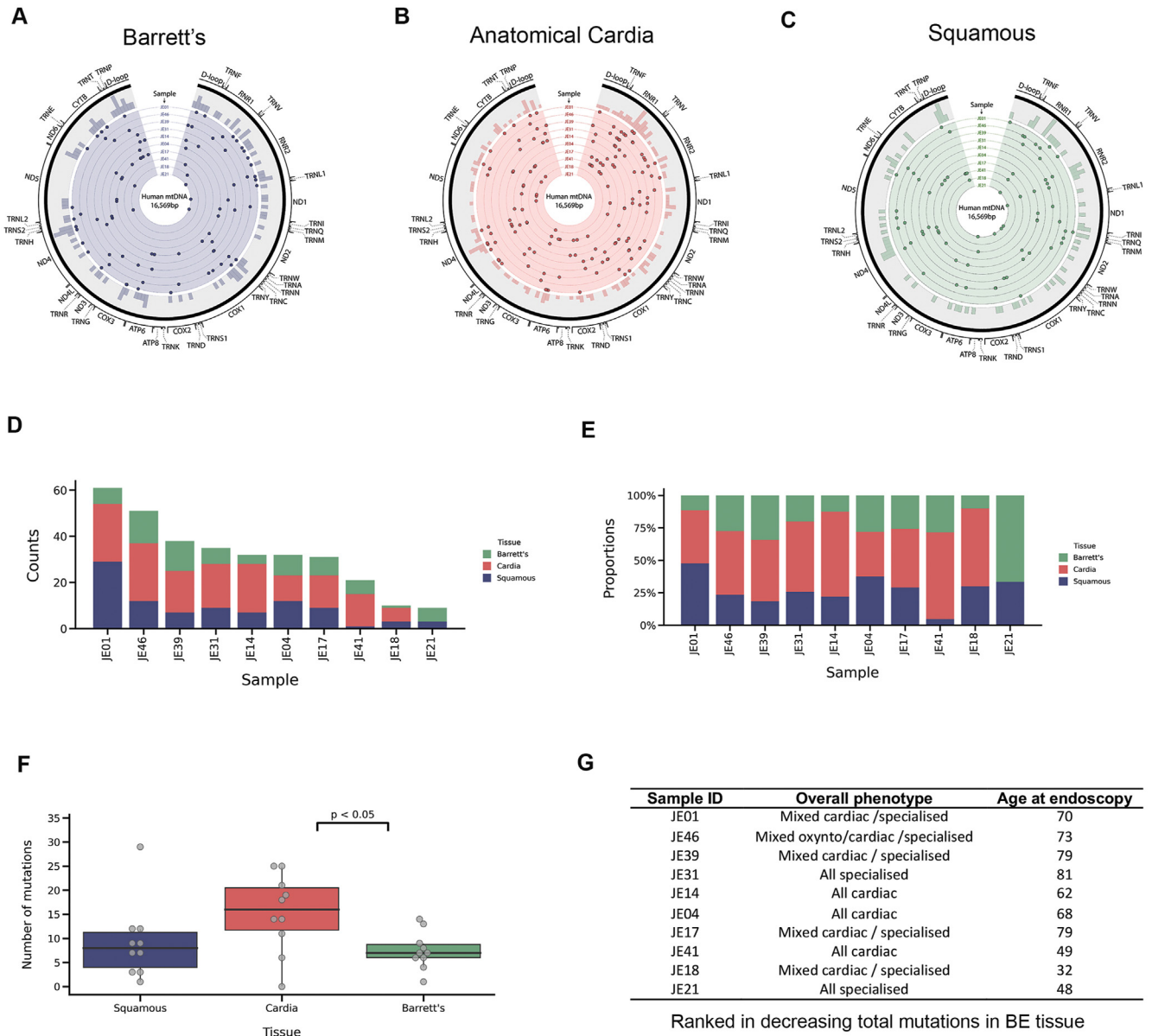
Supplementary Figure 4. Diversity of BE gland phenotype is not associated with patient age. Each biopsy specimen taken from Figure 1B (top) was redistributed according to patient age at the time of endoscopy. There is no correlation between patient age and phenotypic diversity.



Supplementary Figure 5. Phenotypic richness in BE is not related to (A) patient age, (B) lesion size or and (C) the number of biopsy specimens taken at endoscopy. The Spearman correlation coefficient was calculated between the number of biopsy specimens sampled and the richness score for each of 1000 bootstrapping replicates for every patient (n = 22). Three patients were removed from this analysis because of unchanging richness across all biopsy specimens.

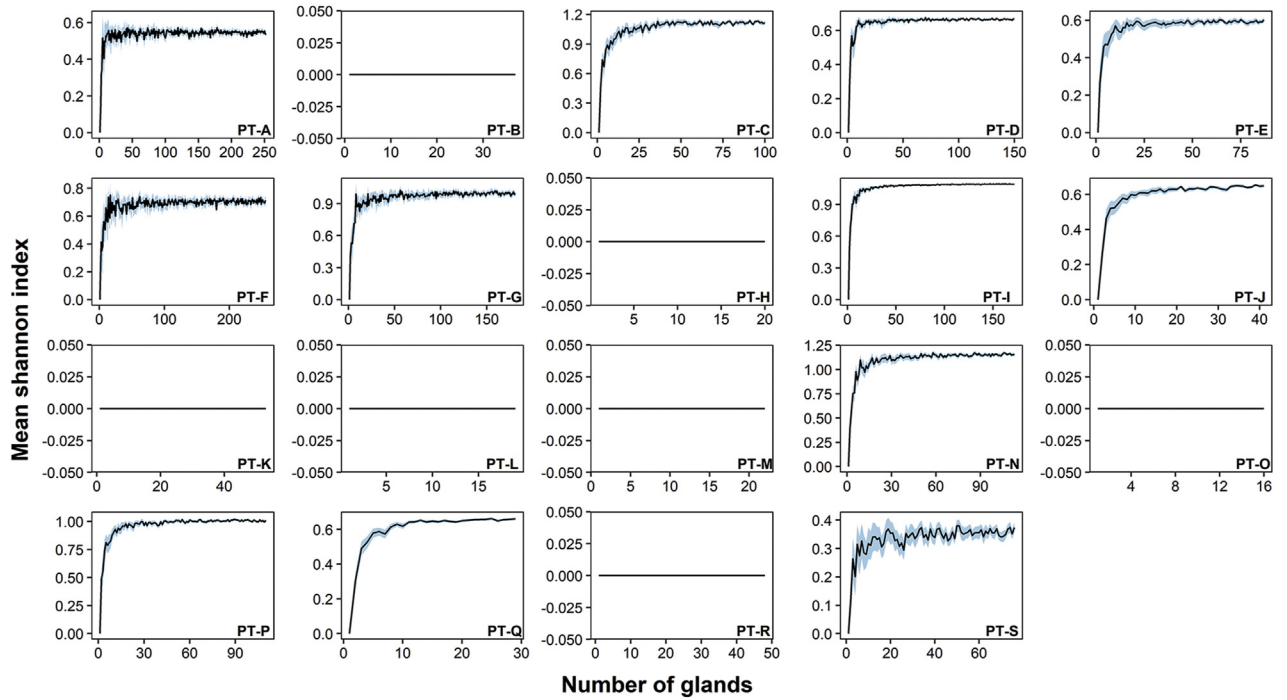


Supplementary Figure 6. Diversity on follow-up is highly variable. Ten patients with nondysplastic BE showed variable gland richness (number of phenotypes) over time. Five patients showed no change in diversity, 3 showed an increase, 1 showed a decrease, and 1 was variable. A red dot represents a change in phenotype with no change in richness. PT, patient.

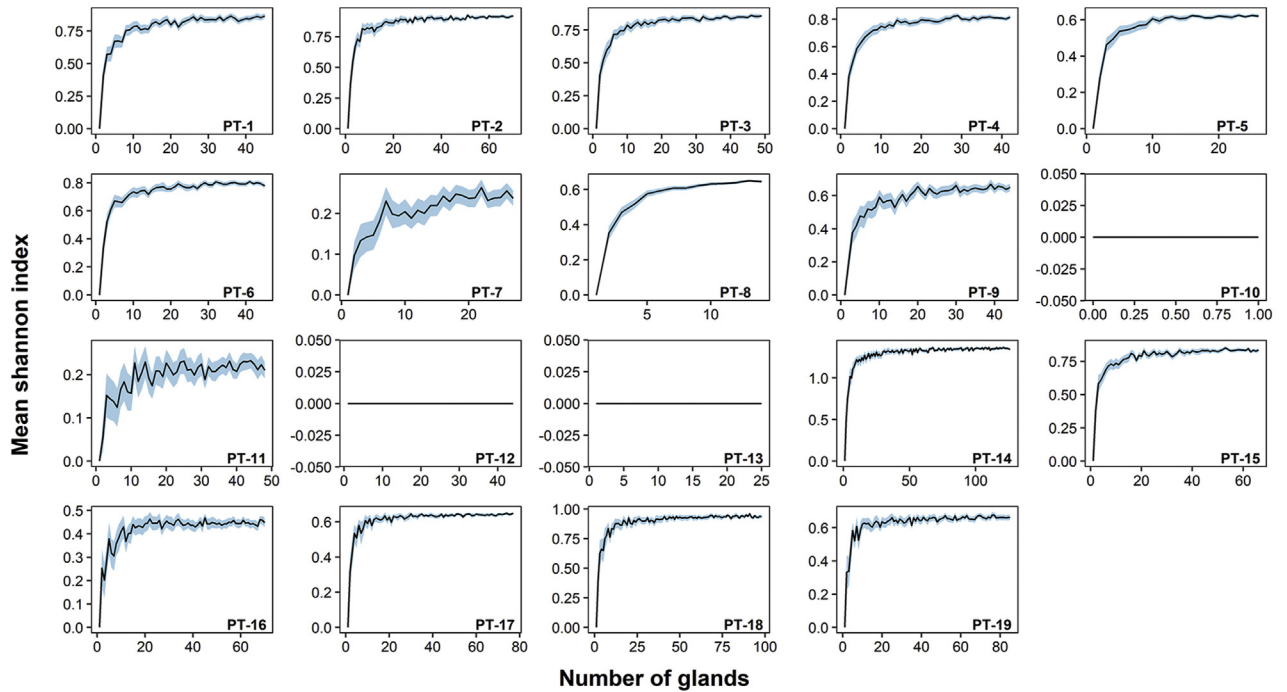


Supplementary Figure 7. NGS of BE, gastric cardia, and adjacent squamous mtDNA from matched patients. (A–C) Circos plots showing the distribution of all somatic variants in (A) Barrett’s epithelium, (B) anatomic gastric cardia epithelium, and (C) squamous. $n = 10$ for each. (D, E) Mutation burden as per (D) normalized mutation count for each tissue type and (E) the proportion of mtDNA mutations in each tissue for each patient. (F) A comparison of the number of variants observed in each tissue. A significantly greater number of mutations was observed in the cardia compared with BE ($P < .05$). (G) Mutation burden was not associated with gland phenotype. Patients are ordered by decreasing number of mutations, and a trend was observed in mutation frequency with increasing age (Pearson’s correlation $R^2 = 0.72$, $P = .02$).

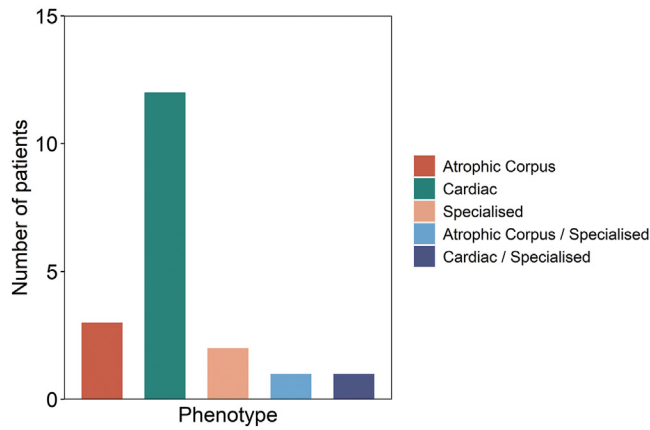
A



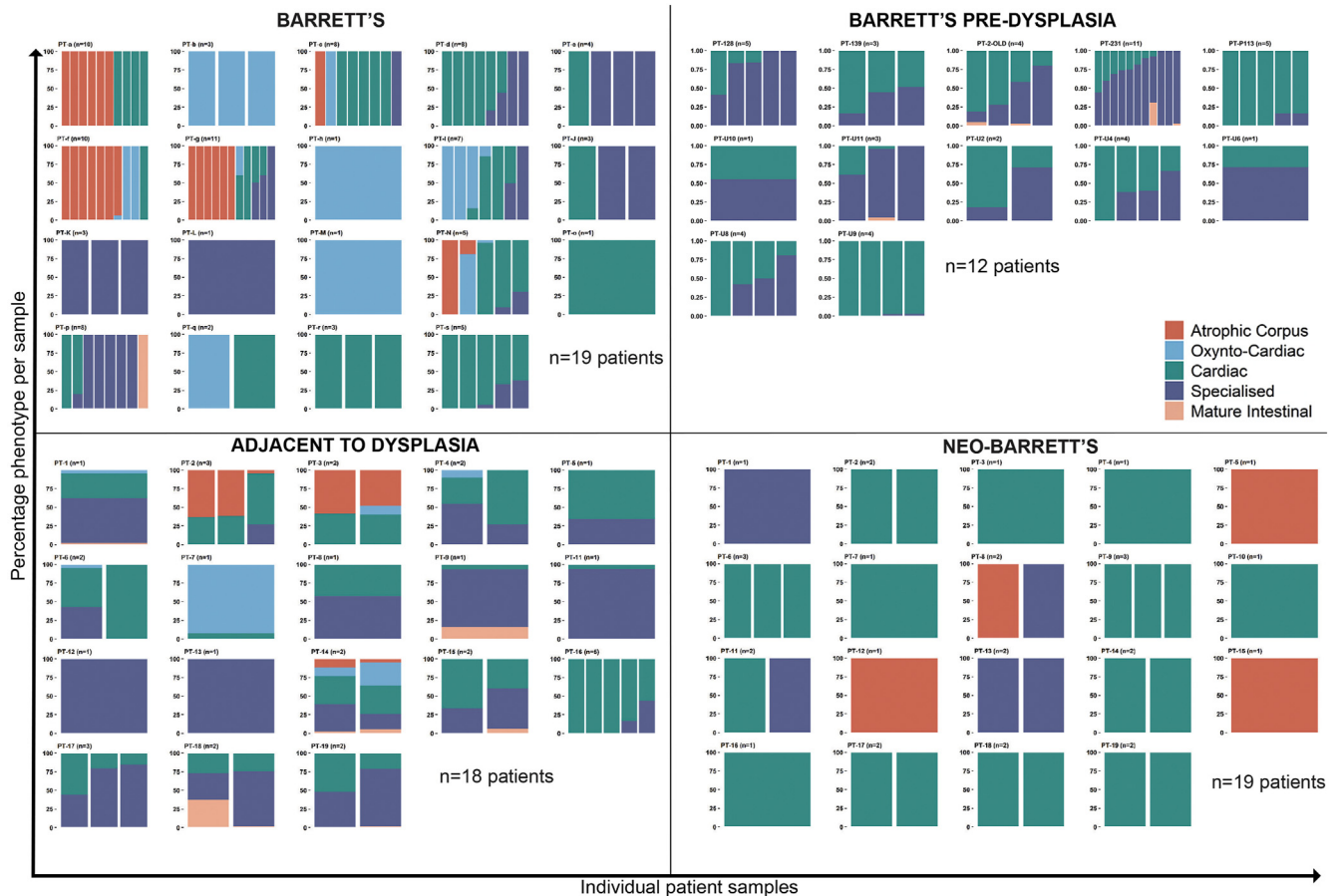
B



Supplementary Figure 8. The number of glands sampled in the BE and BE adjacent to dysplasia cohorts was sufficient to provide a reliable Shannon diversity index. Each graph represents an individual patient. The blue shaded area represents the 95% confidence interval. The Shannon diversity index is derived from a bootstrapping subset ($n = 5000$) from either the BE or the BE adjacent to dysplasia cohort. In the vast majority of cases, a stable Shannon diversity index was achieved in the simulation below or at the number of glands used in the experimental data set. Neo-BE was not included because no diversity was observed. PT, patient.



Supplementary Figure 9. Gland phenotype in neo-BE cases. Thirty-one biopsy specimens from 19 neo-B patient biopsies were phenotyped. The majority (n = 22) of patient biopsy specimens showed a pure cardiac-type phenotype.



Supplementary Figure 10. A per-patient analysis of observed gland phenotypes. Phenotypic diversity within nondysplastic BE (top left), predysplastic BE (top right) adjacent to dysplasia, BE (bottom left), and neo-BE (bottom right). Each column represents a single biopsy taken at a known site within the esophagus.

Supplementary Table 1. Clinical Details of Cohort 1a

Patient code	Biopsy code	Prague	Sex	Age at scope, y	Distance from GEJ, cm	Hiatus hernia, ^a cm	Smoking status	Body mass index, kg/m ²
107	RH19	C0M3	F	73	0.5	3	N/A	N/A
108	JE41	C7M8	M	49	1	5	N/A	N/A
	JE119	C6M8		50	1	6		N/A
	RH41	C7M9		52	1	6		N/A
	RH82	C8M8		53	1	7		N/A
110	JE48	C0M2.5	M	69	1	3	N/A	28.8
111	JE105	C0M1.5	M	49	1	3	Yes	N/A
	RH05	C2M2		49	2	4		35.6
118	RH26	C9M9	M	55	1	9	N/A	29.5
120	JE120	C8M9	F	83	1	5	N/A	23.4
122	RH38	C0M4	M	83	1	4	N/A	23.4
126	JE125	C2M5	M	50	1	2	N/A	N/A
127	RH60	C14M14	M	64	1	4	Yes	37.5
133	JE29 B6	C5M7	M	89	1	Yes	No	27.8
136	JE32P	C0M0.5-1	M	73	1	Yes	Ex	36.2
137	JE33 B5	C2M6	M	61	1	Yes	Yes	24.2
140	JE36 B5	C12M12	M	59	1	8 to 9	No	30.7
	RH45	C14M14		62	1	9		N/A
	RH72	C12M12		62	1	10		N/A
141	JE37 B5	C0M2	M	56	1	N/A	Yes	N/A
	RH28	C0M2		58	1	2		23.7
143	JE39	C4M5	M	79	1	N/A	Ex	25.6
144	RH31	C0M4	F	74	1	5	Yes	31.0
145	JE42	C2M4	F	62	1	4	No	22.0
	JE111	C1M4		63	1	N/A		N/A
147	JE44	C0M2	M	51	1	N/A	No	31.4
148	JE45	C1M2.5	M	68	1	4	N/A	N/A
	RH20	C0M4		70	1	10		N/A
152	JE50	C2M3	M	27	1	4	N/A	N/A
154	JE52	C0M2	F	52	1	Yes	N/A	N/A
155	JE53	C1M3	M	66	1	6	N/A	N/A
	JE71	C0M2		67	1	4		31.8
157	JE55	C1M5	M	64	1	5	No	35.0
158	RH27	C0M3	M	67	1	N/A	Yes	29.8
160	JE58	C1M5	M	56	1	N/A	Yes	35.0
	RH30	C0M5		58	1	2		N/A
161	JE59	C0M2	F	77	1	Yes	Yes	28.3
165	JE64	C5M6	M	61	1	N/A	Yes	N/A
167	RH32	C0M2	M	51	1	3	No	30.3

Supplementary Table 1. Continued

Patient code	Biopsy code	Prague	Sex	Age at scope, y	Distance from GEJ, cm	Hiatus hernia, ^a cm	Smoking status	Body mass index, kg/m ²
168	JE67	C3M7	M	48	1	3	N/A	N/A
172	JE72	C7M8	M	68	1	4	Yes	30.5
174	JE74	C0M4	M	59	1	2	Yes	22.2
178	JE78	C0M5	M	72	1	5	Ex	36.0
179	JE80	C1M2	M	66	1	3	No	19.6
	RH71	C1M3		68	1	3		N/A
180	JE81	C1M6	M	66	1	5	N/A	35.0
	RH91	C0M5		69	1	3		N/A
191	RH12	C8M9	M	60	2	5	N/A	N/A
197	RH14	C7M7	F	68	1	8	N/A	N/A
204	JE113b	C7M8	M	80	1	5	Yes	27.3
205	JE115	C0M2	F	64	1	2	Ex	26.1
207	JE117	C7M8	F	73	1	4 to 5	Ex	N/A
208	JE118	C1M2	M	66	1	N/A	N/A	23.3
213	RH03	C3M4	F	56	1	3	No	26.7
214	RH06	C0M5	M	53	1	N/A	No	31.6
224	RH29	C2M5	M	62	1	4	No	33.1
225	RH33	C0M3	M	62	1	5	No	24.1
227	RH39	C1M4	M	45	1	3	N/A	29.1
230	RH59	C2M4	M	45	2	7	No	29.6
240	RH73	C0M4	M	58	1	2	Ex	26.5
243	RH76	C3M3	F	69	2	7	No	31.2
244	RH77	C7M7	M	68	1	6	Yes	27.2
246	RH85	C4M6	M	72	2	5	Ex	N/A
247	RH88	C5M7	M	56	1	2	Ex	22.3

ex, ex-smoker; F, female; GEJ, gastroesophageal junction; M, male; N/A, not available; no, never smoked; yes, smoker.

^aYes indicates hiatus hernia present but length unknown.

Supplementary Table 2.Clinical Information of Cohort 1b

Patient code	Biopsy code	Prague	Sex	Age at scope, y	Study	Smoking status	Hiatus hernia, ^a cm	Body mass index, kg/m ²
133	JE29 B4	C5M7	M	89	Sanger	No	Yes	27.8
180	JE81 B2	C1M6	M	66	Sanger	N/A	5	35.0
146	JE43 B2	C5M6	M	67	Sanger	N/A	N/A	N/A
107	JE01 squamous JE01 BE JE01 cardia	C2M3	F	70	NGS	No	No	N/A
110	JE04 squamous JE04 BE JE04 cardia	C2M2	M	69	NGS	N/A	Yes	28.83
119	JE14 squamous JE14 BE JE14 cardia	C3M7	M	62	NGS	N/A	2	29.23
122	JE17 squamous JE17 BE JE17 cardia	C1M4	M	79	NGS	N/A	4	23.4
123	JE18 squamous JE18 BE JE18 cardia	C2M4	M	32	NGS	N/A	4–5	29.39
126	JE21 squamous JE21 BE JE21 cardia	C1M5	M	48	NGS	Ex	2	N/A
135	JE31 squamous JE31 BE JE31 cardia	C3M4	F	81	NGS	N/A	N/A	N/A
143	JE39 squamous JE39 BE JE39 cardia	C4M5	F	79	NGS	Ex	N/A	25.7
108	JE41 squamous JE41 BE JE41 cardia	C7M9	M	49	NGS	No	5	29.09
149	JE46 squamous JE46 BE JE46 cardia	C0M5	M	73	NGS	N/A	3	N/A

ex, ex-smoker; F, female; M, male; N/A, not available; no, never smoked; yes, smoker.

^aYes indicates hiatus hernia present but length unknown.

Supplementary Table 3. Characteristics of the Primary Antibodies Used for IHC

Antibody	Source	Host/isotype	Dilution FFPE	Dilution FF	Secondary	Tertiary
Pepsinogen	Abcam	Mouse/IgG1	1:700	N/A	Goat anti-mouse	Streptavidin AP/blue
H ⁺ K ⁺ ATPase	DAKO	Mouse/IgG1	1:1000	1:4000	Goat anti-mouse	Streptavidin HRP/DAB
MUC5AC	Abcam	Mouse/IgG1	1:500	1:50/1:100	Goat anti-mouse	Streptavidin HRP/DAB
MUC2	Thermo Fisher Scientific	Mouse/IgG1	1:500	1:200	Goat anti-mouse	Streptavidin AP/blue
Defensin6 α	Sigma	Rabbit/polyclonal	1:5000	N/A	Swine anti-rabbit	Streptavidin AP/blue
Defensin5 α	Abcam	Mouse/IgG1	N/A	1:100	Rabbit anti-mouse	Streptavidin HRP/DAB
MUC6	Abcam	Mouse/IgG1	1:500	N/A	Goat anti-mouse	Streptavidin HRP/DAB

NOTE. Antibody sources, dilutions, and processing used in this article. FF, fresh frozen; N/A = not applicable (antibody not used for this purpose).

Supplementary Table 4. Primer Sequences for Whole mtDNA Amplification

Primer	Sequence 5' → 3'	Position of amplicon in mitochondrial genome
MTL1-F	AAAGCACATACCAAGGCCAC	9397–1892
MTL1-R	TTGGCTCTCCTTGCAAAGTT	
FRAG1-F	TATCCGCCATCCCATACATT	15195–9796
FRAG1-R	AATGTTGAGCCGTAGATGCC	