Gastroenterology 2017;153:657–673

REVIEWS IN BASIC AND CLINICAL GASTROENTEROLOGY AND HEPATOLOGY

Ernst J. Kuipers and Vincent W. Yang, Section Editors

The Evolving Genomic Landscape of Barrett's Esophagus and Esophageal Adenocarcinoma

Gianmarco Contino,[1] Thomas L. Vaughan,[2,3] David Whiteman,[4] and Rebecca C. Fitzgerald[1]

[1]*Medical Research Council Cancer Unit, Hutchison/Medical Research Council Research Centre, University of Cambridge, Cambridge, UK;* [2]*Cancer Epidemiology Program, Fred Hutchinson Cancer Research Center, Seattle, Washington;* [3]*Department of Epidemiology, School of Public Health, University of Washington, Seattle, Washington; and* [4]*Cancer Control, QIMR Berghofer Medical Research Institute, Brisbane, Australia*

We have recently gained unprecedented insight into genetic factors that determine risk for Barrett's esophagus (BE) and progression to esophageal adenocarcinoma (EA). Next-generation sequencing technologies have allowed us to identify somatic mutations that initiate BE and track genetic changes during development of tumors and invasive cancer. These technologies led to identification of mechanisms of tumorigenesis that challenge the current multistep model of progression to EA. Newer, cost-effective technologies create opportunities to rapidly translate the analysis of DNA into tools that can identify patients with BE at high risk for cancer, detect dysplastic lesions more reliably, and uncover mechanisms of carcinogenesis.

Keywords: Esophagus; Genome-wide Association Study; Mutational Signature; Chromothripsis; Cytosponge.

Over the past 40 years, the incidence of esophageal adenocarcinoma (EA) has increased more than sixfold in Western countries.[1–3] The overall age-adjusted incidence in the United States is 2.7 cases per 100,000, a figure that reaches 6.0 and 9.4 per 100,000 among American and British white men, respectively.[2,3] Overall 5-year survival is approximately 20% and approximately half of patients die within a year of diagnosis.[1] However, fewer than 50% of patients diagnosed early enough for curative treatment (surgery and neoadjuvant chemotherapy or chemoradiotherapy) survive for 5 years.[4,5]

Barrett's esophagus (BE) is a precursor to development of EA. In the absence of dysplasia (NDBE), the risk for transformation of BE to invasive cancer is 0.3% per year (reviewed in Anaparthy and Sharma[6]). BE, which is found in approximately 1% to 2% of the general population, is a squamocolumnar metaplasia that develops in response to gastroesophageal reflux (GER). BE pathogenesis involves a combination of anatomic (hiatus hernia), genetic, and lifestyle risk factors.[6–10] Neoplastic transformation of NDBE usually occurs through progressive grades of dysplasia. Endoscopic treatment is recommended on identification of dysplasia, which is associated with a risk of progression to cancer of 10% per year or higher.[11–15] Unfortunately, clinical strategies for BE, which focus on endoscopic surveillance and endoscopic therapy, have not reduced the incidence or mortality of EA in the general population.[16–18] This is because most cases of EA present without a previous diagnosis of BE. It has been estimated that 40% of EA cases have no history of GER symptoms and an additional 52% of cases have a history of GER but did not receive a diagnosis or undergo endoscopic surveillance.[7]

Even when diagnosed, there are no systems to stratify patients with BE, based on cancer risk, for surveillance and endoscopic therapy. Limited sensitivity of current endoscopic imaging technologies and sampling bias causes many dysplastic lesions to be missed. There is also low interobserver reproducibility among pathologists in grading dysplasia, leading to overdiagnosis or underdiagnosis. When patients with invasive EA are identified, there are few therapeutic options.

Some of these issues can be improved by increasing our understanding of molecular factors associated with development of EA, including inherited (germline, Supplementary Figure 1) and acquired (somatic, Figure 1) genetic alterations (Figure 2). Development of massively parallel and less costly sequencing techniques (next-generation sequencing) has led to a number of genome-wide datasets, which can be used to study the genomic features of EA (Supplementary Table 1). We review the germline and somatic variants identified in different stages of the NDBE to EA spectrum, and discuss the challenges to translating findings from genomic analyses into screening, diagnostic, and therapeutic strategies.

Abbreviations used in this paper: **APOBEC, apolipoprotein B mRNA editing enzyme catalytic polypeptide-like; BE, Barrett's esophagus; BEACON, Barrett's and Esophageal Adenocarcinoma Consortium; BMI, body mass index; EA, esophageal adenocarcinoma; GER, gastroesophageal reflux; GWAS, genome-wide association study; NDBE, nondysplastic Barrett's esophagus; SE, standard error; SNP, single nucleotide polymorphism; WGS, whole-genome sequencing.**

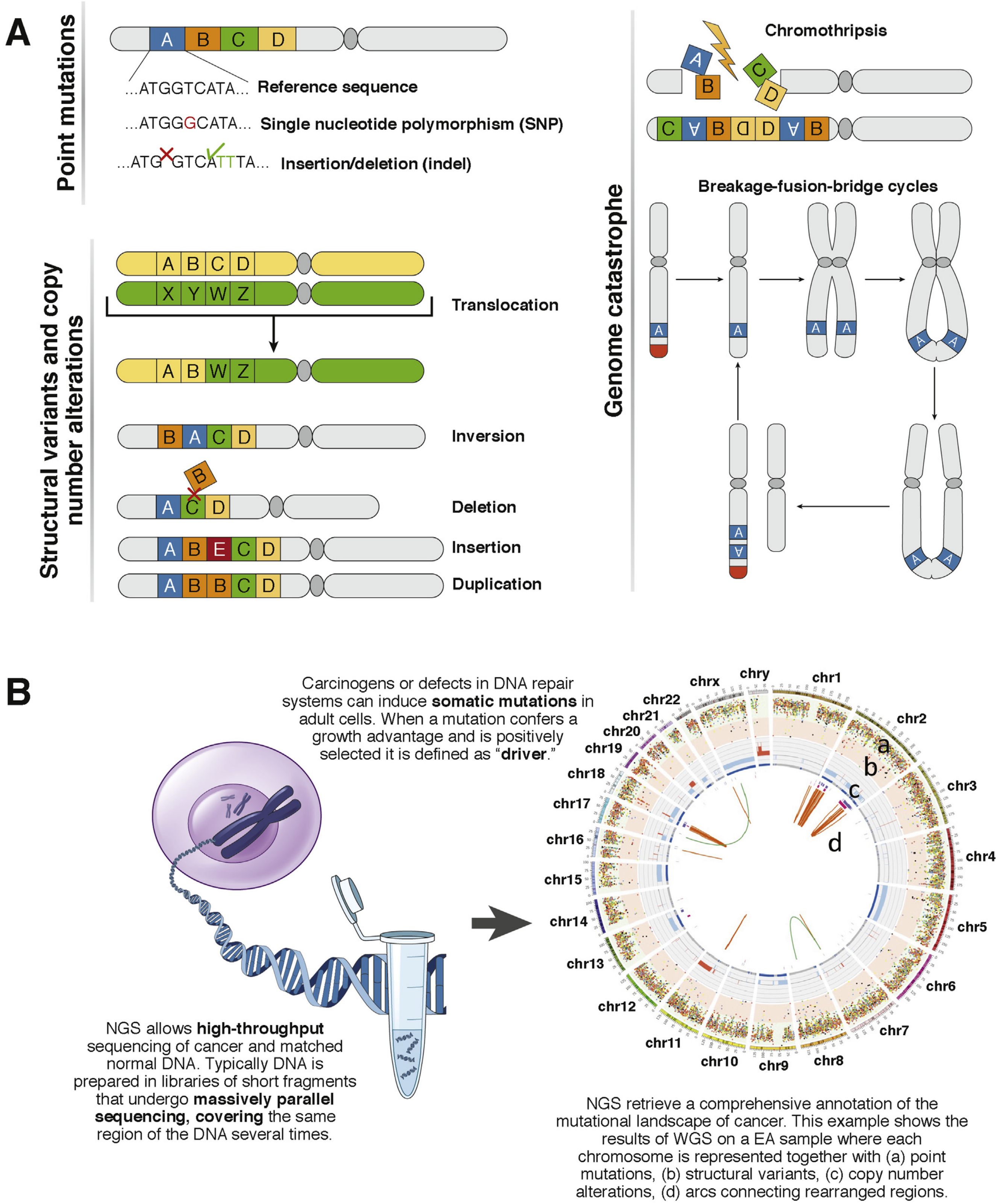
Most current article

© 2017 by the AGA Institute
0016-5085/$36.00
http://dx.doi.org/10.1053/j.gastro.2017.07.007

Germline Variations and Susceptibility

Family Studies

Evidence that germline mutations contribute to development of EA originated from reports of familial aggregation of this cancer[19] and BE.[20–23] Orloff et al[24] performed linkage analyses comparing 21 concordant affected sibling pairs (42 siblings with BE and/or EA) and 11 discordant sibling pairs using a 100K single nucleotide polymorphism (SNP) set. Subsequent fine-mapping of


A
Point mutations
A
B
C
D
...ATGGTCATA... Reference sequence
...ATGGGCATA... Single nucleotide polymorphism (SNP)
...ATGGTCATTTA... Insertion/deletion (indel)
Structural variants and copy number alterations
A B C D
X Y W Z
Translocation
A B W Z
B A C D
Inversion
B
A C D
Deletion
A B E C D
Insertion
A B B C D
Duplication
Genome catastrophe
Chromothripsis
A
B
C
D
C A B D D A B
Breakage-fusion-bridge cycles
A
B
chrx
chry
chr1
chr2
chr3
chr4
chr5
chr6
chr7
chr8
chr9
chr10
chr11
chr12
chr13
chr14
chr15
chr16
chr17
chr18
chr19
chr20
chr21
chr22
a
b
c
d

Carcinogens or defects in DNA repair systems can induce **somatic mutations** in adult cells. When a mutation confers a growth advantage and is positively selected it is defined as "**driver**."

NGS allows **high-throughput** sequencing of cancer and matched normal DNA. Typically DNA is prepared in libraries of short fragments that undergo **massively parallel sequencing**, **covering** the same region of the DNA several times.

NGS retrieve a comprehensive annotation of the mutational landscape of cancer. This example shows the results of WGS on a EA sample where each chromosome is represented together with (a) point mutations, (b) structural variants, (c) copy number alterations, (d) arcs connecting rearranged regions.

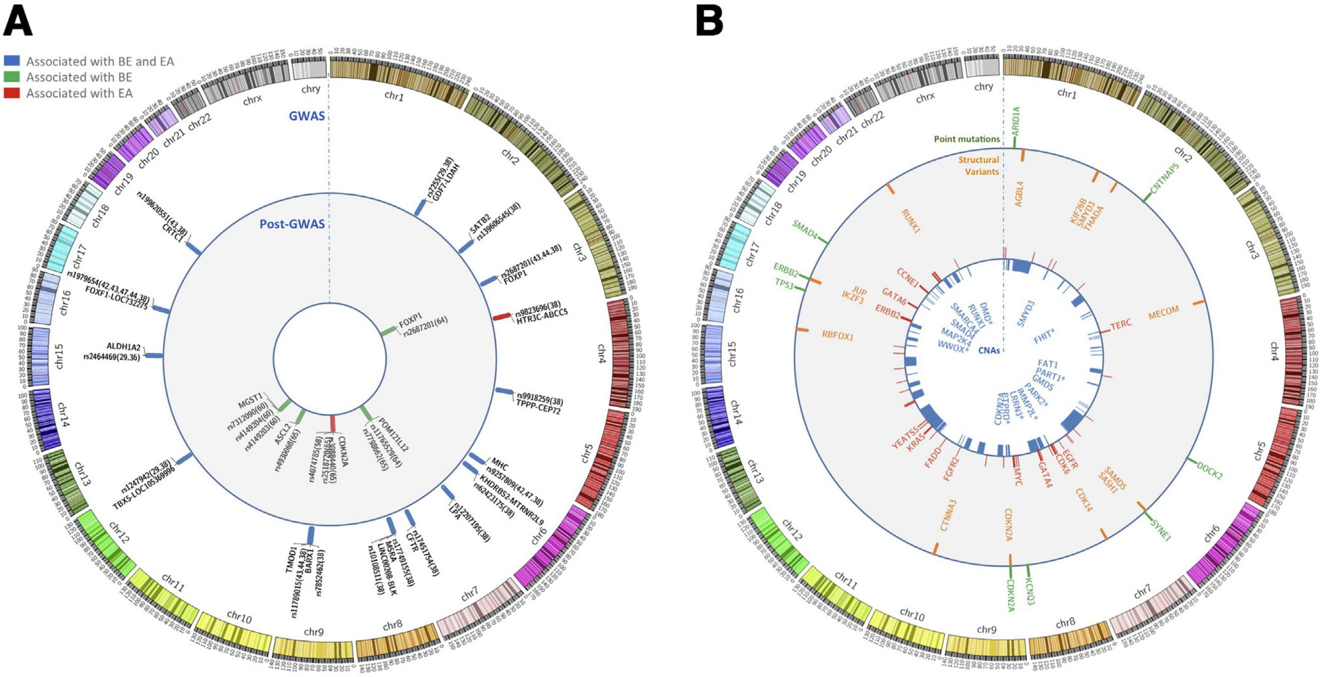


Figure 2. Variants that increase risk for BE and EA and genomic alterations frequently detected in EAs. (*A*) Circos plot of the loci associated with BE or EA risk in GWAS and in post-GWAS, reference to the first report followed by reference to confirmatory reports is shown in *brackets*. (*B*) Circos plot of genomic alterations frequently detected in EAs. From the center of the circos to the outer ring: (1) significant regions of copy number losses (*blue*) according to the Gistic analysis (a tool to identify somatic copy number alterations; Broad Institute, US) reported by Secrier et al,[78] Nones et al,[85] and Kim et al[91] on their respective cohorts; (2) copy number gains (*red*) according to the previously described criteria; (3) most frequent recurrent gene hits by SVs reported by Secrier et al,[78] fragile sites were excluded; (4) recurrent point mutations in driver genes according to MutSig and MutSigCV (bioinformatic tools to identify driver mutations; Broad Institute, US) in $\geq$ 10% of cases by Dulak et al,[54] Secrier et al,[78] and Kim et al.[91] *Common Fragile Site Genes. For an extended annotation of the data shown, see Supplementary Table 2.

regions of interest in an independent set of persons with BE or EA and controls, integration with publicly available gene expression data, and mutational analyses revealed 3 candidate genes for validation, performed in an independent set of 58 persons with BE or EA. Variants in *MSR1*, on chromosome 8p22, were significantly associated with BE or EA in the validation sample and in the pooled sample.[24] More recently, analyses of 42 multiplex pedigrees linked BE and EA with 3 chromosome regions (2q31, 4p14, and 12q23), and an additional region (15q26), in 18 female, affected pedigrees.[25] The specific variants that mediate these associations have not been identified.

The extent to which BE or EA (including adenocarcinoma of the gastroesophageal junction) in siblings determines risk of BE or EA was examined using a training data set of 879 BE pedigrees and a validation set of data from 643 pedigrees, obtained from the Barrett's Esophagus Translational Research Network.[26] In male and female individuals, having a sibling with BE or EA is associated with increased risk. For example, a 50-year-old man with 1 unaffected brother was estimated to have a 3.2% baseline risk for BE or associated cancers. With 1 or 2 affected brothers, his risk increases 2.8-fold (to 9.1%) and 8.3-fold (to 26.6%), respectively. Similar increases in relative risk were estimated for a 50-year-old woman, but applied to a much lower baseline risk (0.5%.) However, when the discrimination accuracy (determined from area under the curve) of a risk prediction model containing only demographic and clinical risk factors was compared with a model that contained family history, there was only minimal improvement (from 0.803 to 0.806). This likely reflects the relative rarity of a positive history in siblings in the general population,

Figure 1. Somatic mutations and next-generation sequencing of cancer. (*A*) Tumor tissues can have point mutations, structural variations, copy number alterations, and genome catastrophes. Possible mechanisms of mutation are shown in a chromosome (2 arms linked by a dark gray centromere); these can involve a large segment of genome (*lettered rectangles*) or single DNA base pairs. Structural variations can cause loss or gain of genetic material and result in copy number changes. Complex structural variations occur in regions of genome catastrophes, such as chromothripsis and breakage–fusion–bridge cycles.[100,101,103,115] In cycles of breakage–fusion–bridge, an unprotected DNA end is generated following the loss of the telomeres (*red*) or a double-strand break.[115] During anaphase, the broken chromatids can fuse (anaphase bridge) and then tear unevenly when the 2 chromatids are pulled apart. This event can be repeated through multiple cycles, leading to amplification of oncogenes. (*B*) Next-generation sequencing of DNA extracted from cancer cells can identify somatic mutations that arise during carcinogenesis.

and the strength and higher prevalence of the other established risk factors that were included in the models.

Heritability

An estimate of heritability (genetic variance explained) of EA and BE among unrelated individuals was calculated using pooled genome-wide association study (GWAS, Supplementary Figure 1) data from 1509 patients with EA, 2383 patients with BE, and 2170 control participants, contributed by 14 epidemiologic studies in the Barrett's and Esophageal Adenocarcinoma Consortium (BEACON). Using autosomal markers and genome-wide complex trait analysis, Ek et al[27] estimated that 25% (standard error [SE], 5%; 1-sided $P = .0000002$) of EA cases and 35% of BE cases (SE, 6%; 1-sided $P = .000000001$) were determined by the composite effect of many common mutations of small individual relative risk.[28] Furthermore, they demonstrated substantial polygenic overlap between EA and BE, indicating that shared genes influence the development of the 2 disorders. No other studies have reported on the EA genetic variance explained, nor on the overlap between EA and BE. However, Palles et al[29] reported a lower figure for BE (10.0%; SE, 1.2%) for genetic variance explained. However this was based on the combined contributions of fewer SNPs (521,744 compared with 797,518 from the BEACON study).

A portion of the heritability of EA and BE may be explained by germline variants that affect development and severity of risk factors for these conditions, including symptomatic GER and obesity.[30–33] For example, a study based on self-administered questionnaires found that GER symptoms were substantially more prevalent among first-degree relatives of persons with BE or EA than among first-degree relatives of their spouses.[34] Twin studies of symptomatic GER support the concept of an important susceptibility component, with heritability estimates ranging from 13% to 41%.[35–37] Gharahkhani et al[38] estimated heritability based on genotype arrays and reported that 7% of the variance in GER symptoms could be explained by genetic factors. Furthermore, they found evidence for substantial genetic overlap between symptomatic GER and BE and EA. The heritability of obesity, measured by body mass index (BMI), waist circumference, and waist-hip ratio, appears to be even higher than for symptomatic GER, with estimates ranging from 40% to 70% from twin and family studies.[39] GWAS have identified close to 100 loci at the genome-wide level of significance ($P < .00000005$), and estimated that more than 20% of variation in BMI can be accounted for by common variants.[40] Using Mendelian randomization methods, researchers associated a risk score based on 29 BMI-associated variants was with a 12% to 16% increase in risk of BE and EA, respectively, per 1 kg/m^2 increase in BMI.[41]

GWAS

The first GWAS of BE was based on a discovery dataset of 1852 case and 5172 control participants from the Wellcome Trust Case Control Consortium.[42] After replication, researchers confirmed that 2 SNPs were associated with BE risk. One was located on chromosome 6p21, within the major histocompatibility complex, and the other on chromosome 16q24. A multiphase extension of this study identified 3 additional loci, on chromosomes 2p24, 12q24, and 15q21, respectively, that were significantly associated with risk of BE.[29]

A larger GWAS, which was the first to include EA cases (n = 2390) in addition to BE cases (n = 3175), was conducted by the BEACON consortium.[43] This study took advantage of the previous finding of extensive genetic overlap between EA and BE,[27] pooling BE and EA cases in the main analyses to increase statistical power. The researchers found 3 additional novel loci, on chromosomes 3p14, 9q22, and 19p13. They also observed that the previously reported association between BE and a locus on 16q24 also extended to risk of EA. Confirmatory evidence for an association between risk of EA and 3 of the 4 BEACON-reported SNPs (3p14, 9q22, and 16q24) was reported in a study from Germany using targeted genotyping.[44]

A meta-analysis of data from 4 GWAS, performed in 6 countries, included 4112 cases of EA, 6167 cases of BE, and 17,159 control participants of European ancestry.[45] The analysis confirmed associations among BE, EA, and the combined case group, with 7 of the 8 previously reported loci at the traditional level of statistical significance ($P = .00000008$). The eighth, on chromosome 9q22, narrowly missed this threshold ($P = .00000062$). This analysis also identified 9 additional loci, 8 of which were associated with BE and EA, and 1, on chromosome 3q27, which was associated with only EA.

In summary, a total of 17 independent loci associated with the development of BE and/or EA have been identified by traditional GWAS (Figure 2, Supplementary Table 2). One striking finding is that many of the identified SNPs are located in or near genes that regulate development and differentiation of the esophagus, stomach, and intestine (such as *FOXP1*, *FOXF1*, *BARX1*, *GDDF1*, and *ABCC5*).[29,42,43,45–47] Given the importance of GER in development of BE and EA, and the fact that hiatal hernia substantially predisposes to GER, the findings identify mechanisms by which these variants might affect development of BE and EA. Support for this concept was provided by pathway analyses, which identified processes related to muscle cell differentiation, as well as mesenchyme development and differentiation, associated with these conditions.[45]

The large meta-analysis identified an intriguing association between an SNP on chromosome 7q31, located within the *CFTR* gene, and risk of BE and EA.[45] This gene is mutated in patients with cystic fibrosis, a condition characterized by severe dysfunction of the respiratory and gastrointestinal tract beginning in childhood, including a high prevalence of GER (in 35%–81% of patients).[48,49] The incidence of cystic fibrosis is approximately sixfold higher in persons of European ancestry vs African ancestry, as are incidences of BE and EA.[50] It was highlighted that *CFTR* and *ABCC5* each encode proteins belonging in the same class of transmembrane ion transporters (ATP-binding cassette), indicating an interesting area for research into pathogenic mechanisms of these disorders.[51]

After GWAS

Moving beyond GWAS, investigators have used a variety of analytic approaches to explore the influence of genetic factors on EA pathogenesis, including integrating knowledge of somatic mutations with germline data, performing pathways-based analyses, and using epidemiologic data to examine genetic associations with risk factors. Somatic mutations occur at high frequencies in the *CDKN2A* and *TP53* tumor suppressor genes in EAs (and other malignancies)[52–54]; loss of heterozygosity at these loci has been associated with progression from BE to cancer (see section on somatic mutation analyses).[55–57] Reasoning that these loci may be implicated in susceptibility to cancer, investigators from the BEACON consortium tested 13 SNPs at the *TP53* locus and 24 SNPs in *CDKN2A*, which were within 2-kb flanking regions and satisfied quality control constraints. Although none of the SNPs in *TP53* were associated with EA risk, 3 polymorphisms in *CDKN2A* were associated with a 10% to 16% reduction in risk for EA ($P < .05$) (Figure 2, Supplementary Table 2).[58] The investigators then tested whether any of the variants predicted neoplastic progression in a separate prospective cohort of 408 patients with BE, and found that 2 of the variants (rs2518720; hazard ratio, 0.57 and rs3088440; hazard ratio, 0.34) were independently significantly associated with reduced risks of progression. Expression of one of the variants (rs3088440) in cell lines indicated that it reduces microRNA-mediated repression of the *CDKN2A* mRNA.

Systemic and local (esophageal) inflammation, caused by factors such as abdominal obesity, GER, and cigarette smoking, may be a common pathway in the development of BE and EA.[59] The role of genetic variation in inflammatory responses was investigated using a principal components-based approach in the BEACON GWAS. Variants in the cyclooxygenase (COX) pathway were significantly associated with risk of BE. Gene-level analyses identified an association with *MGST1* (on chromosome 12p12), and a meta-analysis, which added BE and control participants from the Wellcome Trust GWAS, confirmed associations between 4 SNPs and risk of BE (Figure 2, Supplementary Table 2).[60] Analyses of GWAS data examining the role of germline variation in other pathways, including the biogenesis and activity of microRNAs,[61] androgens,[62] and the estrogen and oxytocin pathways,[63] also indicated associations, but these have not been replicated.

Gene–Environment Interactions

Some genetic factors affect susceptibility to BE or EA depending on other factors. Another approach to identifying so-called risk-modifying genes is therefore to test for differences in statistical associations across strata of exposure to those factors (eg, BMI, sex). Using the well-annotated BEACON GWAS, Dai et al[64] examined the first 7 SNPs identified as associated with BE or EA at the genome-wide level of significance for interactions with BMI, GER symptoms, and smoking status. They found that the previously identified variant near *FOXP1* (rs2687201) significantly modified the association between GER symptoms occurring at least weekly and risk of BE, such that the association was stronger (odds ratio, 6.2) among persons with 0 minor alleles, compared with those with 1 or 2 (odds ratios, 3.6 and 4.0, respectively,) ($P_{\text{interaction}} = .0005$; false discovery rate $= 0.042$.)

Dai et al[65] developed a set of constrained testing methods to increase statistical power for tests of gene–environment interactions in settings in which several risk factors may act through a common pathway. Inflammation frequently accompanies cigarette smoking, abdominal obesity, and GER. When the constrained score statistics were applied to the BEACON dataset, 3 loci were identified that simultaneously interacted with smoking, obesity, and GER (Supplementary Table 2). Further explorations in this area will likely require much larger datasets that also include accurate annotation of key environmental risk factors.

Pleiotropic Analysis of Risk Loci

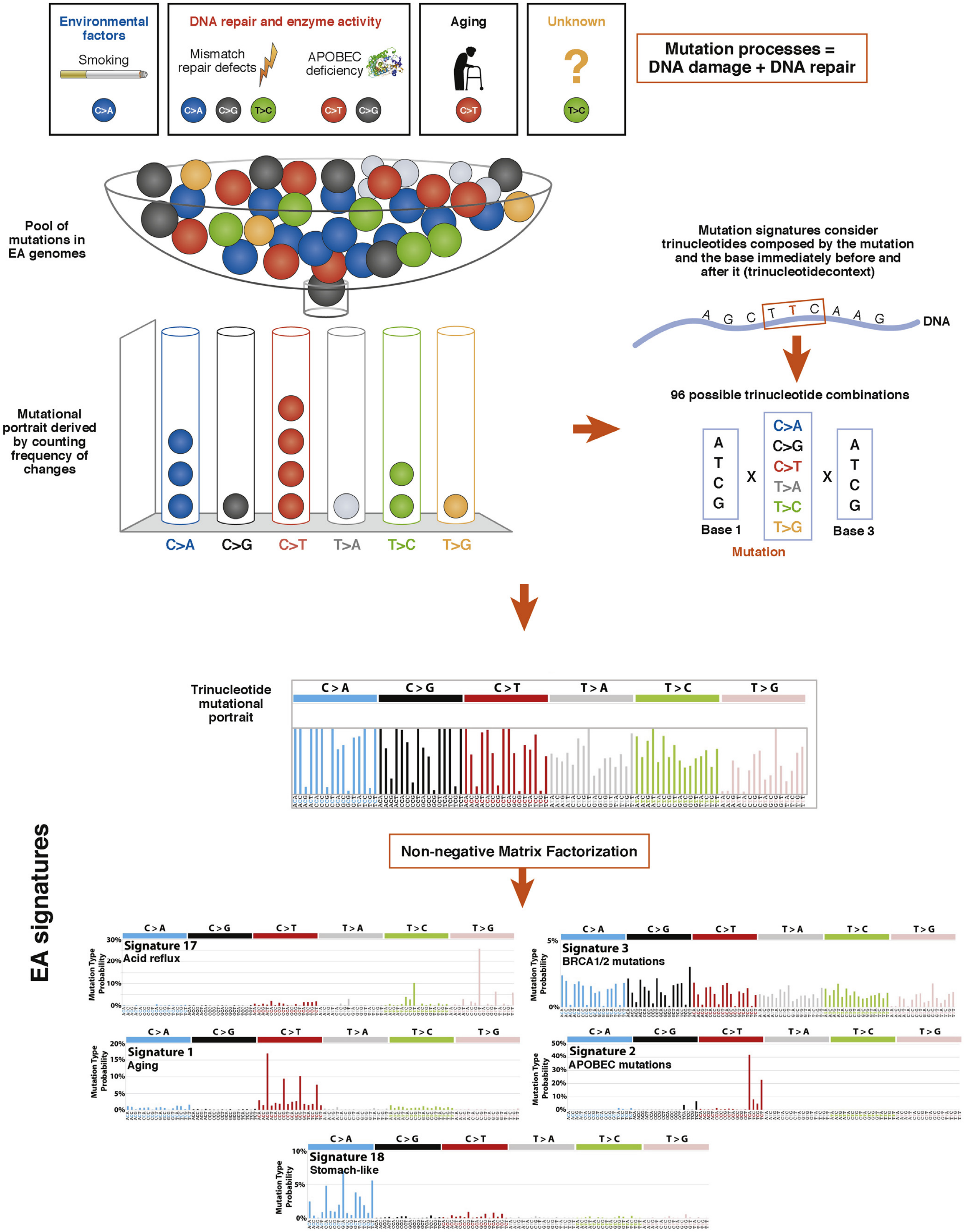
To investigate whether risk-associated loci from GWAS of other cancer sites might also modify risk of BE or EA, Lee et al[66] tested 387 candidate SNPs. None were found to be associated with risk of BE or EA, and there was no evidence for interactions with smoking, obesity, or GER symptoms.

Somatic Mutations That Affect BE Progression

With the advent of next-generation sequencing, mutations have been reported from hundreds of cases in studies of coding regions (whole-exome sequencing) and the entire genome (whole-genome sequencing [WGS]). These data can be obtained from 2 large pan-cancer consortia: the Cancer Genome Atlas (https://tcga-data.nci.nih.gov) and the International Cancer Genome Consortium (http://icgc.org). New data are being added every day.

Progression from Premalignant BE to EA

There was a reasonable expectation that sequencing the genomes of BE or EA tissues would identify somatic alterations required for progression from BE to EA. This was expected to lead to biomarkers that could assist clinicians in identifying preneoplastic lesions at highest risk for progression to invasive cancer. In our current model, dysplasia progresses to invasive EA via early loss of *CDKN2A*, emergence of dysplastic clones with mutations in *TP53* and/or additional somatic alterations, and increases in copy number.[55,57,67–73] Although the basics of this model, largely characterized before the advent of next-generation sequencing techniques, appear to hold true, sequencing studies have shown the BE genome to be highly complex, even when nondysplastic for many years, and that progression can be nonlinear.[74,75]


Environmental factors
Smoking
C>A
DNA repair and enzyme activity
Mismatch repair defects
C>A
C>G
T>C
APOBEC deficiency
C>T
C>G
Aging
C>T
Unknown
?
T>C
Mutation processes = DNA damage + DNA repair
Pool of mutations in EA genomes
Mutational portrait derived by counting frequency of changes
C>A
C>G
C>T
T>A
T>C
T>G
Mutation signatures consider trinucleotides composed by the mutation and the base immediately before and after it (trinucleotidecontext)
A G C T T C A A G
DNA
96 possible trinucleotide combinations
A T C G
Base 1
X
C>A C>G C>T T>A T>C T>G
Mutation
X
A T C G
Base 3
Trinucleotide mutational portrait
C > A
C > G
C > T
T > A
T > C
T > G
Non-negative Matrix Factorization
EA signatures
Mutation Type Probability
Signature 17
Acid reflux
30%
20%
10%
0%
Signature 3
BRCA1/2 mutations
5%
0%
Signature 1
Aging
20%
15%
10%
5%
0%
Signature 2
APOBEC mutations
50%
40%
30%
20%
10%
0%
Signature 18
Stomach-like
10%
5%
0%

It is now apparent that point mutations accumulate during early stages of disease and BE lesions often have a higher rate of mutation rate than many common, invasive cancers.[52,70,75] At the time BE becomes dysplastic, the tissue has a mutation rate comparable to that of EA.[52,75] Mutations are found in a number of tumor suppressor genes important in chromatin remodeling, such as *ARID1A* and *SMARCA4*[75] (Supplementary Table 3). Mutations in *TP53* and *SMAD4* are usually found only in patients with high-grade dysplasia and EA, respectively. In contrast to patients with NDBE with no history of disease progression, mutations in *TP53* are found in NDBE tissues adjacent to EA.[67,70] This observation is consistent with the high allele fraction of *TP53* mutations in many different cancer types, indicating that either this mutation appears early during tumorigenesis or it is able to promote expansion of a dominating clone.[76]

Mutations in *PIK3CA* and *CTNNB1* have also been found in BE, although accumulation of activating mutations and amplifications in oncogenes is a marker of invasive EA.[70] Similarities in mutation patterns provide evidence for the common origin of BE and EA (see Figure 3).[53,70] However, fewer than 20% of specific variants overlap between adjacent BE and EA, so either the cancer clone diverged at an early stage or originated separately.[53,70] Analysis of patients with BE suggested that the genetic diversity of different clones did not change significantly over time, but the extent of divergence of clones at baseline was the strongest predictor of progression.[77]

The mutational landscape found in BE and EA differs more dramatically at a chromosomal scale. For example, compared with BE epithelium, EAs have marked differences in genomic copy number profiles. Genomes of BE tissues are relatively stable compared with those of invasive tumors, in which almost 40% of the genome is nondiploid (median, range, 2%–97%). The only common copy number alteration found in BE is 9p loss of heterozygosity (*CDKN2A*).[53,70,71] Invasive tumors have increased copy numbers of several oncogenes (*GATA4*, *KLF5*, *MYB*, *PRKCI*, *CCND1*, *FGF3*, *FGF4*, *FGF19*, and *VEGFA*) and loss of common fragile sites (*FHIT*, *WWOX*, *PDE4D*, *PTPRD*, and *PARK2*).[53,70,78–80]

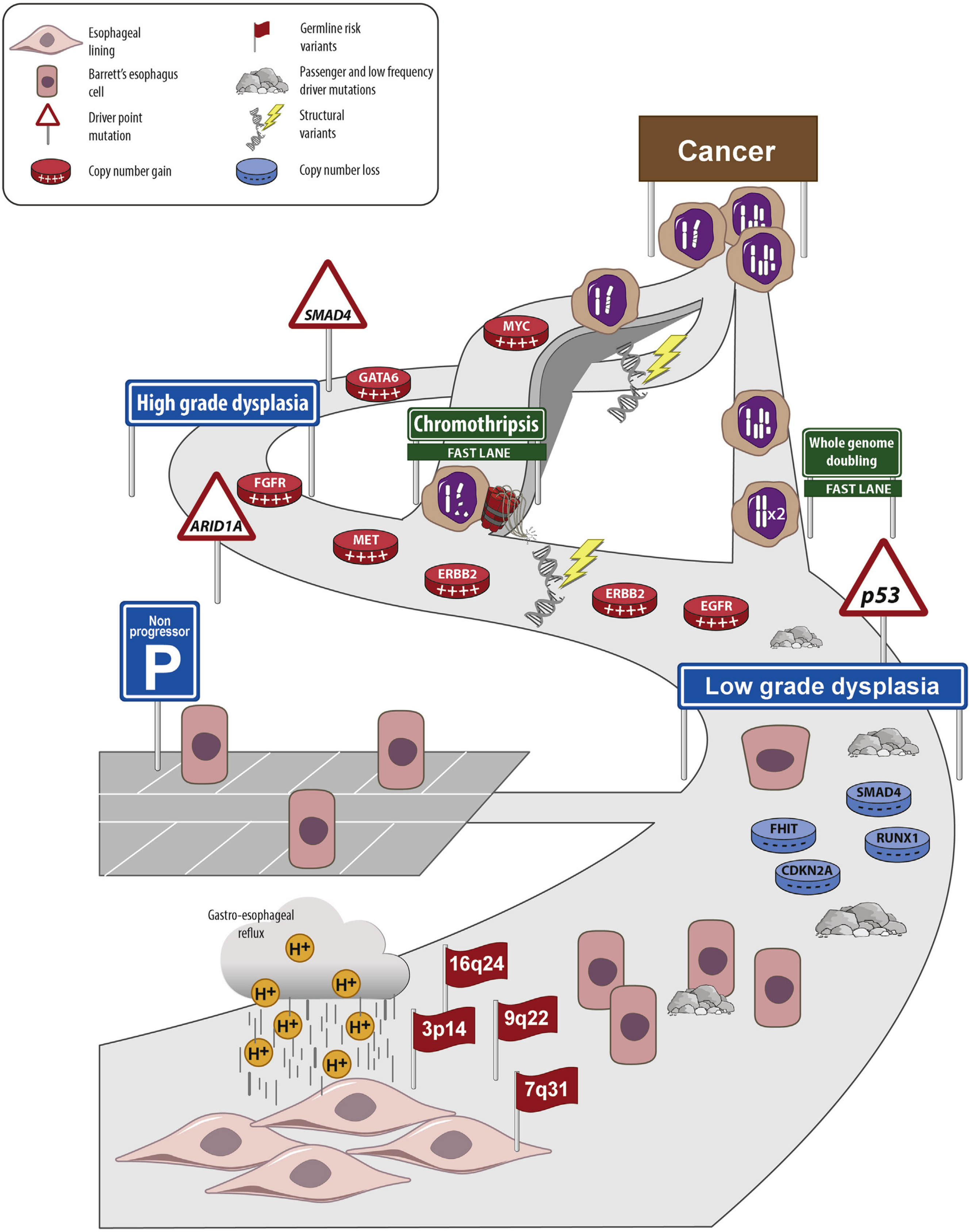
The stochastic and gradual accrual of copy number alterations fits into the linear multistep process of BE progression, but does not entirely account for the frequent whole-genome doubling observed by Stachler at al,[70] particularly in EA tissues with *TP53* mutations. The authors propose that following *TP53* loss, whole-genome doubling occurs, which accelerates tumor progression and requires few other mutations. It is also observed that BE can progress to cancer via multiple different pathways, and suddenly accelerate, due to crises involving large regions of the genome (genomic catastrophes). Tumors with unstable genomes are more likely to progress rapidly,[55,59,81] so the frequency of copy number changes is a good biomarker for development of EA. In the 24 months before a patient is diagnosed with esophageal cancer, biopsies from BE tissues show a marked increase in DNA content.[69] These findings indicate that the time course and pathways to tumor development vary to a greater extent than previously appreciated (Figure 4).

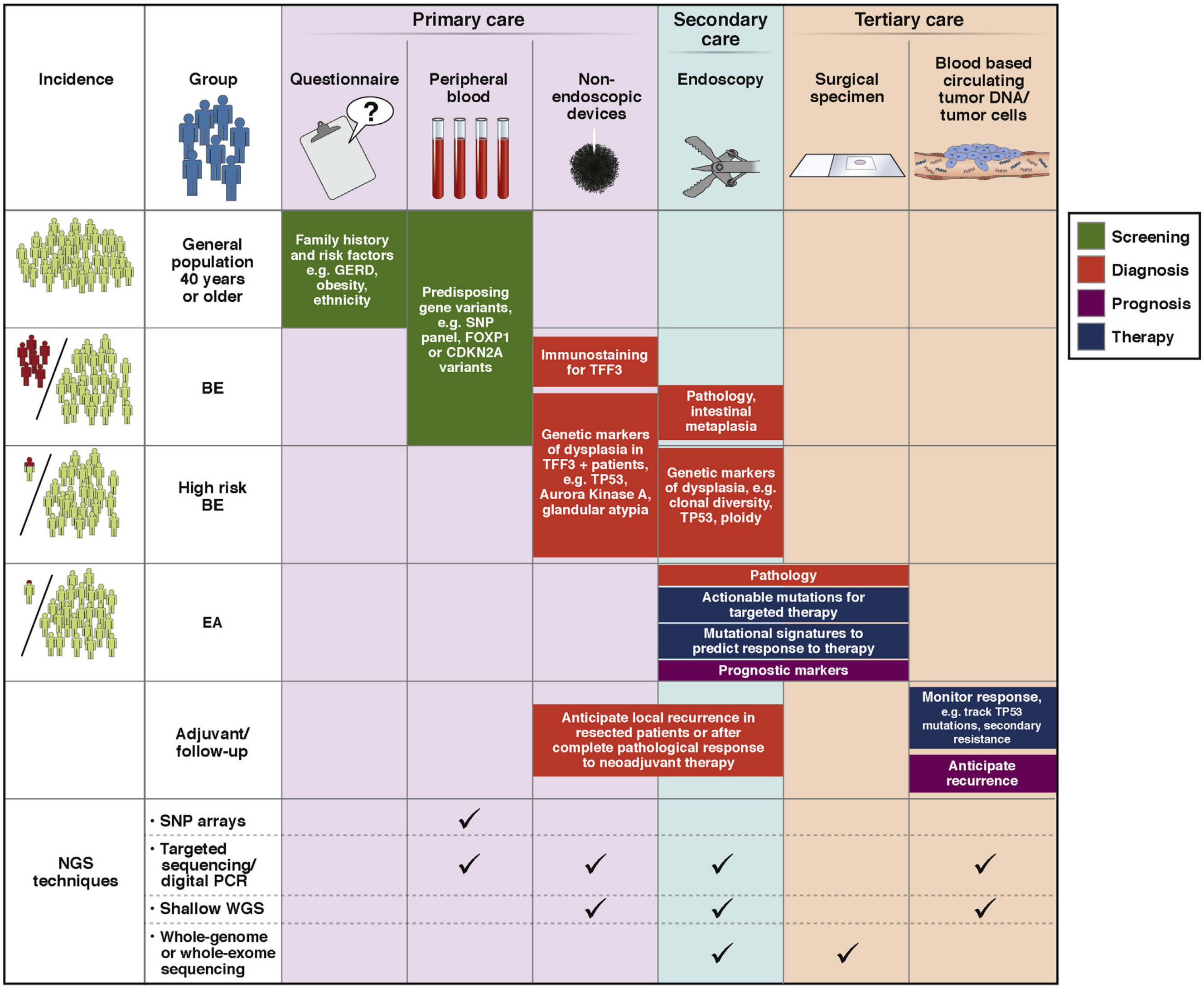
On a practical note, it is a challenge to predict the lifetime course of a patient's BE progression. In the past, when esophagectomy was the only therapy available, patients were followed until it was clear they had invasive cancer. Now, intervention is appropriate earlier in the disease course,[11,12] due to the availability of outpatient-based endoscopic techniques, such as endoscopic mucosal resection and radiofrequency ablation. The agenda has therefore shifted toward identifying early genomic events that distinctly mark the presence of dysplasia, awaiting more refined risk models for NDBE. The modality of tissue sampling is critical, because BE is a polyclonal disease and endoscopic biopsies have inherent sampling bias. Fortunately, several new modes of sample collection have been developed, which could overcome some of these limitations.

One of these approaches, the Cytosponge sampling device, collects cells from the entire length of the esophagus; it is simple to perform and inexpensive, allowing for repeated sample collection in a primary care setting.[82,83] The diagnostic yield of the Cytosponge for new cases of BE in individuals with a history of reflux is being compared with standard of care in a cluster randomized clinical trial of 9000 patients in primary care (registration number: REC 16-EE-0546). As well as diagnosing BE, as noted previously, risk stratification is essential. Analysis of a single Cytosponge sample was able to recapitulate the same sequencing results as samples collected from polyclonal lesions in multiple biopsies.[53,75] Furthermore, a panel of biomarkers can be applied to BE cells from the same sample

◀

Figure 3. Mutational signatures of tumors. (*A*) Mutational processes are biological activities (eg, aging, smoking, UV light exposure, unknown carcinogens) that generate patterns of mutations (mutational signatures) through a damage of the DNA sequence and its attempt to repair it by DNA repair mechanisms. (*B*) The mutational portrait is the total pattern of genetic changes in cancer cell that derive from the sum of all the mutational signatures occurring in a lifetime.[86] (*C*) Mathematical approaches, such as non-negative matrix factorization, can be used to extract mutational signatures from the mutational portraits of groups of patients' cancer genomes. The pattern includes all base substitutions and flanking nucleotides (96 possible combinations shown in *bar charts*). Non-negative matrix factorization estimates the relative contribution of each signature to the mutational portrait and can highlight cancers that are predominantly driven by some mutational signatures. A comprehensive catalogue of the signatures identified by Alexandrov et al[87] is available in the catalogue of somatic mutations in cancer (COSMIC, www.cancer.sanger.ac.uk). Mutation signatures associated with EA include (1) S17, also called an acid signature, there are 2 forms, S17A and B; (2) S3, associated with defects in the BRCA1/2-led homologous recombination pathway; (3) S1, associated with aging; (4) S2, caused by APOBEC mutations; and (5) S18, detected in gastric cancer and neuroblastoma, arises via an unknown mechanism.[78,85]


Esophageal lining
Barrett's esophagus cell
Driver point mutation
++++
Copy number gain
Germline risk variants
Passenger and low frequency driver mutations
Structural variants
Copy number loss
Cancer
SMAD4
MYC ++++
GATA6 ++++
High grade dysplasia
Chromothripsis
FAST LANE
FGFR ++++
ARID1A
MET ++++
ERBB2 ++++
ERBB2 ++++
EGFR ++++
x2
Whole genome doubling
FAST LANE
p53
Non progressor
P
Low grade dysplasia
SMAD4
FHIT
RUNX1
CDKN2A
Gastro-esophageal reflux
H+
16q24
3p14
9q22
7q31



Incidence	Group	Primary care			Secondary care	Tertiary care	
		Questionnaire	Peripheral blood	Non-endoscopic devices	Endoscopy	Surgical specimen	Blood based circulating tumor DNA/ tumor cells
	General population 40 years or older	Family history and risk factors e.g. GERD, obesity, ethnicity	Predisposing gene variants, e.g. SNP panel, FOXP1 or CDKN2A variants				
	BE			Immunostaining for TFF3 Genetic markers of dysplasia in TFF3 + patients, e.g. TP53, Aurora Kinase A, glandular atypia	Pathology, intestinal metaplasia		
	High risk BE				Genetic markers of dysplasia, e.g. clonal diversity, TP53, ploidy		
	EA				Pathology Actionable mutations for targeted therapy Mutational signatures to predict response to therapy Prognostic markers		
	Adjuvant/ follow-up			Anticipate local recurrence in resected patients or after complete pathological response to neoadjuvant therapy			Monitor response, e.g. track TP53 mutations, secondary resistance Anticipate recurrence
NGS techniques	• SNP arrays		✓				
	• Targeted sequencing/ digital PCR		✓	✓	✓		✓
	• Shallow WGS			✓	✓		✓
	• Whole-genome or whole-exome sequencing				✓	✓	

🟩 Screening
🟥 Diagnosis
🟪 Prognosis
🟦 Therapy

Figure 5. Translating findings from genetic studies into clinical practice. Genetic data can be used to determine an individual's risk for developing BE or EA, and to manage patients at different stages of disease progression. Tests are available for use in primary (*pink*), secondary (*light blue*), and tertiary (*orange*) care settings. For each group (*left*), we provide examples of clinical applications. The most suitable technique for each test is presented in the *bottom row*. Further details about available NGS techniques are available in Supplementary Table 1. The *left column* indicates the group size relative to the general population.

(identified as Trefoil Factor 3–positive cells at immunostaining) to stratify patients into 3 risk groups according to the following criteria: presence of glandular atypia, p53 abnormality and a ploidy measure (Aurora kinase A positivity), along with joint effects of major risk factors, such as age, obesity, and length of the Barrett's segment (if known). Using this algorithm, 35% of patients fell into the low-risk category, and were eligible for a less-intense surveillance regimen, and this was reliable in a validation cohort[84] (Figure 5).

In summary, it seems that regardless of the sampling method, more informative assays are required to identify genomic instability and increasing copy number in patients requiring endoscopic therapy; this would avoid reliance on the subjective diagnosis of dysplasia as the basis for clinical decision making.[55,69,77]

Figure 4. Paths of BE Progression to EA. Findings from next-generation sequencing studies indicate BE progression can accelerate via genome doubling, genome catastrophes, and other unknown mechanisms, even at early stages of tumor progression. The main path represents the multistep progression of BE to EA through dysplasia. BE and EA pathogenesis includes genetic risk factors (each *flag* indicates GWAS-identified regions), exposure to environmental risk factors (eg, acid reflux), and the accumulation of different types of driver and passenger mutations. Genomic catastrophes, such as chromothripsis and whole-genome doubling, can occur at any stage and dramatically accelerate progression of BE.

Whole-Exome and Whole-Genome Analyses of EA

Point Mutations and Indels

The main finding of next generation sequencing studies are summarized in Table 1. EAs have a high degree of intersample genomic heterogeneity and a high mutation burden (Table 1). Each tumor genome has a median of 8 mutations/Mb (range, 1.5–35 mutations/Mb), one of the highest mutation rates observed in tumors, along with bladder, colorectal, lung tumors and melanoma.[85] Other tumor types, such as breast and ovarian tumors, have fewer than 2 mutations/Mb, respectively.[76] EAs might have a high mutation rate depending on the exposure to environmental mutagens, the efficiency in DNA repair, the rate of proliferation, and the inflammatory response. Although no mutagen has been convincingly proven to cause EA, carcinogenesis is believed to involve acid and bile reflux. Little is known about the mechanisms by which these luminal constituents might cause DNA damage; inherited mismatch repair gene deficiencies are not commonly observed.[78]

One method to identify and classify mutational processes is through the statistical analysis of the frequency of base-changes (eg, A>C, T>G) throughout the entire genome (mutational portrait) (Figure 3). This can be carried out by analysis of 1 base at a time or in the context of the base either side (so-called trinucleotide context). Analyses of base changes occurring in a large number of normal and cancer tissue genomes have identified mutation signatures. These have, in some cases, been associated with mutagens, such as ultraviolet radiation, cigarette smoke, or aging[86] (Figure 3). Alexandrov et al[87] created a catalogue of these signatures, using a non-negative matrix factorization algorithm. Tumors can therefore be characterized according to the most commonly occurring signatures (S, number) (Figure 3).

One interesting aspect of EA is the frequency of T>G substitutions in a CTT context, called the S17 signature. This mutation signature has been associated with gastric acid reflux and often referred to as an acid signature.[78,85] Other signatures include one associated with aging (S1), a complex pattern caused by defects in the BRCA1/2-regulated homologous recombination pathway (S3); C>T mutations in a TCA/TCT context, due to apolipoprotein B mRNA editing enzyme catalytic polypeptide-like (APOBEC) mutations (S2); and C>A/T dominant in a GCA/TCT context (S18), also found in gastric carcinoma and neuroblastoma.[78,85] The APOBEC signature has been associated with characteristic clusters of localized hypermutations named kataegis, in which a single strand accumulates a high burden of C>T and C>G mutations.[88] Further analysis of these signatures may help to elucidate mechanisms of carcinogenesis and to aid in classification and treatment[78] (Figure 3).

For a cancer to occur, it is estimated that at least 3 driver gene mutations are required.[89] Despite the large number of mutations found in EA tissues, they contain an average of 1.7 driver mutations per case. Bioinformatic tools can be used to identify driver mutations, such as MutSig and more recently MutSigCV. These have identified only 8 genes that are consistently mutated (in more than 10% of cases)[54,75,78,90,91] (Figure 2, Supplementary Tables 2 and 3). *TP53* is by far the most frequently mutated gene: more than 70% of samples contain loss of function mutations in *TP53*. Studies are needed to determine the combination of mutations required for EA tumorigenesis.

Copy Number Alterations and Structural Variants

Two WGS studies have highlighted that EA genomes are predominantly characterized by large-scale genomic rearrangements (ie, structural variants) and gains or losses of genomic regions (copy number alterations)[78,85] (Figure 1 and Table 1). Chromosome instability stands out as a hallmark of EA when compared with squamous esophageal cancer and gastric adenocarcinoma.[10,91] Copy number alterations of genes encoding EGFR, ERBB2, MET, and FGFR2 and other receptor tyrosine kinases are also common in EA and show a high degree of redundancy with downstream targets[79,92,93] (Figures 2 and 4, Supplementary Table 2).

Rearrangements are variably distributed in the genomes of EA samples. Nones et al[85] proposed a classification of EA genomes: unstable (with 450 or more structural variations), scattered (fewer than 450 structural variations, evenly distributed across the genome), and complex localized (with a concentration of clustered structural variations in a single or few chromosomes), based on the pattern of structural variations distribution.

Highly recurrent rearrangements have been mainly reported in common fragile sites, but their biological significance is unclear. For instance, the fragile histidine triad gene (*FHIT* or *FRA3B*) and WW domain containing oxidoreductase gene (*WWOX* or *FRA16D*) contain rearrangements in up to 95% of cases. Despite evidence that these are tumor suppressor genes,[94,95] their loci are frequently rearranged following perturbation of DNA replication and replication stress.[96,97] Besides common fragile sites, structural variations could be a common mechanism of recurrent mutation in EA. *RUNX1*, a gene translocated in acute myeloid leukemia, and *SMYD3*, are rearranged in 39% and 27% of cases of EA.[78] Although functional studies are needed to confirm a driver role in EA, these alterations are possibly the most common after *TP53* mutations.

In addition, a peculiar class of structural variations is represented by mobile element insertions that occur as a consequence of the excision and reinsertion of repeated L1 and Alu sequences that are transposed as DNA or through the reverse transcription of an mRNA intermediate. In EA, L1 insertions have been reported in the coding sequence of several genes (*ERBB4, CTNNA3, CTNNA2, CDH18,* and *SOX5*). Mobile element activity represents the most relevant contributor to the total structural variant SV burden in several EA genomes but further work is required to clarify their functional consequences.[78,98,99]

WGS has revealed that many EA samples have evidence of genomic catastrophes, which result in the accumulation of

Table 1. Whole Exome and Whole Genome Sequencing Studies on EA

Consortium	Sequencing technique	Cohort (*no. of cases*)	Sample Type	Main findings	Study
N/A	WES + Sanger sequencing	11 chemonaive EAs, 12 chemonaive esophageal squamous cell carcinoma, 2 matched BE	Frozen biopsies	- When compared, EA have more A:T>C:G transversions, squamous carcinomas have more C:G>G:C - NOTCH1 mutations are frequent in esophageal squamous carcinoma (21%) but not in EAs - Most mutations in EAs are already present in matched BE	Agrawal et al, 2012[52]
The Cancer Genome Atlas (TCGA)	WES + WGS on selected samples	Chemonaive EAs (149)	Frozen biopsies	- 26 significantly mutated genes (5 genes in more than 10% of cases) - 3 bp mutational signatures reveal prevalence of A>C transversions at AA dinucleotides - ELMO1 and DOCK2 mutations suggest potential activation of RAC1 pathway, involved in cell invasion	Dulak et al, 2013[54]
International Cancer Genome Consortium (ICGC)	WGS + targeted sequencing	Chemonaive EAs (112), BE (84), HGDs (61)	Frozen biopsies, Cytosponge	- Similar mutation burden between BE and EAs, mostly shared mutations - Only TP53 and SMAD4 mutations occur at stage-specific manner (EA and HGD, respectively) - Mutations can be identified in a Cytosponge samples	Weaver et al, 2014[75]
ICGC	WGS + targeted sequencing	Paired BE and chemonaive EA samples (23), longitudinal sampling of BE (1)	Frozen biopsies, paraffin, Cytosponge	- BE is polyclonal and highly mutated even in the absence of dysplasia - EA development is mainly driven by copy number increases - The mutational context suggests a common causative insult for BE and EAs	Ross-Innes et al, 2015[53]
TCGA	WGS + targeted sequencing	Paired BEs and chemonaive EA samples (25), multiple sampling of BEs and EAs (5)	Frozen biopsies, paraffin	- 62.5% of EA emerges following genome doubling - Genome doubling follows TP53 mutations and leads to the acquisition of oncogenic amplification, providing an alternative model to gradual accumulation of tumor suppressor alterations	Stachler et al, 2015[70]
N/A	WGS	Chemonaive EAs (22)	Frozen biopsies	- High frequency of genomic catastrophes (32% of cases with chromothriptic and 27% with Bridge Fusion Breakages leading to oncogene amplification (*MYC*, *MDM2*, *KRAS,* and *RRC3*) - Extreme genomic instability may be driven by somatic BRCA2 mutations	Nones et al, 2015[85]
N/A	WES	Multiregion, paired pre + post chemotherapy (8)	Frozen biopsies	- Heterogeneity of driver mutation, parallel evolution, and early genome doubling - Poor response to platinum containing neoadjuvant chemotherapy correlates with tumor heterogeneity - Platinum containing neoadjuvant chemotherapy is associated to increase of C>A mutations in CpC context	Murugaesu et al, 2015[117]
ICGC	WGS	Chemonaive EAs (129)	Frozen biopsies	- High heterogeneity with few recurrent point mutations and many large-scale events (copy number alterations and rearrangements) - Co-amplification of receptor tyrosine kinases - Three distinct mutational signature combinations that define molecular subtypes with potential therapeutic relevance	Secrier et al, 2016[78]
TCGA	WES + SNP arrays integrated with DNA methylation and mRNA sequencing	Gastroesophageal samples (592) of which 171 chemonaive EAs and 90 chemonaive esophageal squamous cell carcinoma	Frozen biopsies	- EA is similar to chromosomally instable gastric adenocarcinoma - Hypermethylation is present in 70% of EAs - EA is molecularly distinct from esophageal squamous cell carcinoma	Kim et al, 2017[91]

Table 1. Continued

Consortium	Sequencing technique	Cohort (*no. of cases*)	Sample Type	Main findings	Study
ICGC	WGS	Chemonaive EAs (62), and chemotherapy treated EAs (58) and comparison of matched pre and post chemo EAs (10)	Frozen biopsies	- No significant differences in the overall mutation rate, mutation signatures, specific recurrent point mutations, or copy-number events in respect to chemotherapy status - Whole-genome sequencing of samples obtained following neoadjuvant chemotherapy is representative of the genomic landscape of esophageal adenocarcinoma - In matched pre and post chemotherapy, analysis of SNVs in relation to allele-specific copy-number changes pinpoints the common ancestor to a point prior to chemotherapy.	Noorani et al, 2017[118]

structural variants in specific areas of the genome. These can be single events (chromothripsis) or repeated breakage–fusion–bridge cycles[100,101] (Figure 1*B*). There is evidence for chromothripsis in approximately 30% of EAs and breakage–fusion–bridge events in 25% of EAs. Genomic catastrophes could be a common mechanism through which preneoplastic lesions rapidly progress to invasive tumors[78,85,102–104] (Figures 1 and 4). Crises or punctuated equilibria (as opposed to gradual mutational accrual) can alter cell phenotypes and overcome the oncogene stress (cell cycle arrest or cellular senescence due to activation of an oncogene).

Furthermore, shattered chromosome segments not incorporated into the derivative chromosome can be linked to form a double-minute (circular) chromosome.[104] For example, *MYC*-containing double minutes have been convincingly described in chromothriptic EAs.[85,105] The second type of genomic catastrophes, breakage–fusion–bridge cycles, are related to telomere shortening, observed in advanced EAs.[85] Unprotected telomere ends and sister chromatids fuse and are subsequently torn apart during anaphase. This process can be repeated for several cell cycles resulting in inverted duplications and increased copy numbers of genes, including *KRAS*, *MDM2*, and *VEGFA.*[85]

Application to Therapy

Most cases of EAs present de novo without any previous diagnosis of BE. Standard treatment for EA remains chemoradiotherapy followed by surgery and only incremental gains in survival have been made in the past 20 years. Because loss of *TP53* is the most common mutation, it would make sense to try to restore its function as a therapeutic strategy. Several agents designed to increase P53 activity are in development, but their clinical efficacy has not yet been demonstrated (reviewed in Parrales and Iwakuma[106]).

Most trials of patients with EA have targeted receptor tyrosine kinases. However, trials are often performed without information on the expression level of the targeted receptor within the tumor. Recent sequencing data indicate that tumors from each patient have dysregulations in multiple receptor tyrosine kinases and their signaling pathways, so multiple agents could be required.[78,79,93] An approach to overcome the genomic heterogeneity of EA could be to identify broader pathways that may emerge by a combined analysis of low-frequency somatic variants and transcriptome. The clinical translation of such an approach is an open challenge due to the exponential complexity of cross-talk and primary or secondary resistance.[107,108] Mutational signature analysis is an alternative method for identifying therapeutic vulnerability in subgroups of patients. EA genomes can be grouped into 3 categories according to their dominant mutation signatures: C>A/T dominant (associated with age, S18-like), DNA damage repair impaired (BRCA), and mutagenic (predominantly S17A or S17B with a high mutation burden).[78] EAs with a DNA damage repair defect signature have significantly more defects in homologous recombination and chromosome segregation pathways and may therefore respond better to

DNA-damaging agents or photon irradiation in combination with inhibitors of PARP. On the other hand, EAs with a mutagenic signature have a higher neoantigen load and are characterized by infiltration of CD8+ T cells; these are more likely to respond to blockade of PDL1 and CTLA4, as observed in studies of patients with non–small-cell lung cancer and melanoma.[109] More preclinical studies are needed to test these strategies.

An area of active investigation is whether genomic alterations detected in endoscopic biopsies, Cytosponge, or peripheral blood samples can be used as a tool to monitor disease during treatment or surveillance. In peripheral blood samples, circulating tumor DNA has been used successfully to monitor response to therapy, as well as to identify the emergence of novel alterations conferring secondary resistance to targeted therapies[110,111] (Figure 5). In addition, genomic alterations indicating locoregional recurrence might be detected in biopsies and Cytosponge samples earlier than currently available diagnostic tests. Clinical trials are evaluating close follow-up as an alternative to surgery in patients with complete pathologic response to neoadjuvant radio-chemotherapy.[112] Early detection of recurrence could offer further margins for salvage surgery in those patients.

Future Directions

In the past few years, there have been intense international collaborative efforts to increase our understanding of genetic factors associated with BE and EA. These have produced many important and exciting findings. Genome-wide susceptibility scans have been performed on more than 6000 patients with BE and more than 4000 patients with EA. Although such sample sizes might seem large from the perspective of gastroenterology, they are dwarfed by patient collections for other human traits, such as obesity (∼300,000)[40] and breast cancer (∼62,000 cases).[113] Experience gained from studying those other conditions has shown that increasing study size brings greater ability to detect associations with rare genetic variants.

For now, however, several conclusions seem reasonable. BE and EA each have sizable heritable components, estimated at approximately 35% and 25%, respectively; heritability is conferred by a combination of many genetic factors that each increase risk by a small amount. Almost all of the genetic variants discovered have been associated with BE and EA, indicating that EA arises via a metaplasia–dysplasia–carcinoma pathway.

Annotations of the genes indicate that aberrations in the musculature of the foregut, perhaps during embryogenesis, likely contribute to GER. GER is associated with development of BE and EA. Understanding the underlying functional mechanisms through which these variants act to increase risk will be necessary for these discoveries to yield practical utility.

The value of exploring genomic data using a variety of approaches beyond straightforward association testing has been demonstrated already, but much work remains to be done. More powerful methods for analyzing pathways and detecting gene–environment interactions need to be explored for germline as well as somatic variants. Well-annotated epidemiologic datasets with careful measurement of environmental and phenotypic factors will be essential to achieve these aims. For example, exploring the possible role of DNA repair genes in esophageal carcinogenesis would be enhanced by analyzing associations with genotype separately among ever smokers and never smokers. Functional studies will be needed to characterize the downstream effects of genetic variants identified through the association studies currently under way. In the absence of valid animal models for BE or EA, mechanistic data will be crucial for identifying potential targets for treatment or chemoprevention.

Genomes of BE and EA cells are both highly mutated and heterogeneous and contain well-defined mutational signatures. Analyses of germline and somatic mutations indicate that these disorders have a common etiology. Several pathways are involved in the progression from BE to EA, but the final common feature is that of an abnormal copy number profile with large-scale structural variants, and with amplifications and complex rearrangements occurring to a variable extent. In some instances, disease progression can be very slow, with a cumulative number of point mutations in tumor suppressor genes; other tumors have punctuated evolution resulting in rapid progression.

To generate a coherent picture of genomic alterations that accumulate during progression of EA, genomes of EA must be integrated with analyses of the epigenome, transcriptome, and proteome and compared with other cancers. Results of pancancer analysis across cancer genome sequencing consortia are being released as we speak. Significant effort will be required to investigate the causes and consequences of mutations, especially with respect to rearrangements and the functional consequences of putative driver mutations.

There are many opportunities to translate the findings from genetic susceptibility and somatic sequencing studies to the clinic (Figure 5).[114] Functional studies of susceptibility loci, to identify causal variants and the carcinogenic mechanisms, may eventually lead to the discovery of biomarkers of risk and targeted approaches to prevention and treatment. Data on susceptibility loci could be combined with clinicodemographic factors to generate a risk score, to identify individuals who might benefit from targeted screening for BE. Screening tests for BE could use newer noninvasive sampling methods more suitable for primary care. However, rather than relying solely on histopathologic assessment, inclusion of biomarkers emerging from genome-wide data could help objectively determine whether the epithelium is genomically unstable.

The most important criterion for determining the success of such tools is whether they accurately predict cancer risk; this needs to be assessed in large prospective studies. Ideally the term dysplasia would be superceded by a molecular readout detailing *TP53* mutations and copy number alterations, so that a risk-benefit profile for endoscopic therapy can be determined. For patients with invasive cancer, increasing our understanding of the genomic landscape could improve tumor subclassification and lead to personalized therapy. Tests might soon be available to detect DNA shed from tumors into blood; these could be used to monitor response to therapy or identify emergent clones for therapeutic targeting.[110] With the advent of affordable and rapid sequencing technologies, we can move from discovery science into the clinic.

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.2017.07.007.

References

1. Hur C, Miller M, Kong CY, et al. Trends in esophageal adenocarcinoma incidence and mortality. Cancer 2013; 119:1149–1158.
2. National Cancer Institute. Cancer of the Esophagus—SEER Stat Fact Sheets. Cancer Stat 2015. Available at: http://seer.cancer.gov/statfacts/html/esoph.html. Accessed April 13, 2017.
3. Cancer Resarch UK. Oesophageal cancer incidence statistics. Available at: cancerresearchuk.org/cancer-info/cancerstats/. Accessed April 13, 2017.
4. Hirst J, Smithers BM, Gotley DC, et al. Defining cure for esophageal cancer: analysis of actual 5-year survivors following esophagectomy. Ann Surg Oncol 2011; 18:1766–1774.
5. van Hagen P, Hulshof MCCM, van Lanschot JJB, et al. Preoperative chemoradiotherapy for esophageal or junctional cancer. N Engl J Med 2012;366:2074–2084.
6. Anaparthy R, Sharma P. Progression of Barrett oesophagus: role of endoscopic and histological predictors. Nat Rev Gastroenterol Hepatol 2014;11:525–534.
7. Vaughan TL, Fitzgerald RC. Precision prevention of oesophageal adenocarcinoma. Nat Rev Gastroenterol Hepatol 2015;12:243–248.
8. Spechler SJ, Souza RF. Barrett's esophagus. N Engl J Med 2014;371:836–845.
9. Spechler SJ, Fitzgerald RC, Prasad GA, et al. History, molecular mechanisms, and endoscopic treatment of Barrett's esophagus. Gastroenterology 2010;138: 854–869.
10. Ronkainen J, Aro P, Storskrubb T, et al. Prevalence of Barrett's esophagus in the general population: an endoscopic study. Gastroenterology 2005;129:1825–1831.
11. Phoa KN, van Vilsteren FGI, Weusten BL, et al. Radiofrequency ablation vs endoscopic surveillance for patients with Barrett esophagus and low-grade dysplasia. JAMA 2014;311:1209–1217.
12. Shaheen NJ, Sharma P, Overholt BF, et al. Radiofrequency ablation in Barrett's esophagus with dysplasia. N Engl J Med 2009;360:2277–2288.
13. Rastogi A, Puli S, El-Serag HB, et al. Incidence of esophageal adenocarcinoma in patients with Barrett's esophagus and high-grade dysplasia: a meta-analysis. Gastrointest Endosc 2008;67:394–398.
14. NICE. Clinical Guidelines [CG106] Barrett 's oesophagus: ablative therapy. 2010. Available at: https://www.nice.org.uk/guidance/cg106. Accessed April 13, 2017.
15. Shaheen NJ, Falk GW, Iyer PG. ACG clinical guideline: diagnosis and management of Barrett's esophagus. Am J Gastroenterol 2016;111:30–50.
16. Corley DA, Mehtani K, Quesenberry C, et al. Impact of endoscopic surveillance on mortality from Barrett's esophagus-associated esophageal adenocarcinomas. Gastroenterology 2013;145:312–319.e1.
17. Spechler SJ, Sharma P, Souza RF, et al. American Gastroenterological Association technical review on the management of Barrett's esophagus. Gastroenterology 2011;140:18–52.
18. Qiao Y, Hyder A, Bae SJ, et al. Surveillance in patients with Barrett's esophagus for early detection of esophageal adenocarcinoma: a systematic review and meta-analysis. Clin Transl Gastroenterol 2015;6:e131.
19. Jochem VJ, Fuerst PA, Fromkes JJ. Familial Barrett's esophagus associated with adenocarcinoma. Gastroenterology 1992;102:1400–1402.
20. Crabb DW, Berk MA, Hall TR, et al. Familial gastroesophageal reflux and development of Barrett's esophagus. Ann Intern Med 1985;103:52–54.
21. Chak A, Faulx A, Kinnard M, et al. Identification of Barrett's esophagus in relatives by endoscopic screening. Am J Gastroenterol 2004;99:2107–2114.
22. Chak A, Lee T, Kinnard MF, et al. Familial aggregation of Barrett's oesophagus, oesophageal adenocarcinoma, and oesophagogastric junctional adenocarcinoma in Caucasian adults. Gut 2002;51:323–328.
23. Sun X, Elston R, Barnholtz-Sloan J, et al. A segregation analysis of Barrett's esophagus and associated adenocarcinomas. Cancer Epidemiol Biomarkers Prev 2010; 19:666–674.
24. Orloff M, Peterson C, He X, et al. Germline mutations in MSR1, ASCC1, and CTHRC1 in patients with Barrett esophagus and esophageal adenocarcinoma. JAMA 2011;306:410–419.
25. Sun X, Elston R, Falk GW, et al. Linkage and related analyses of Barrett's esophagus and its associated adenocarcinomas. Mol Genet Genomic Med 2016;4:407–419.
26. Sun X, Elston RC, Barnholtz-Sloan JS, et al. Predicting Barrett's esophagus in families: an Esophagus Translational Research Network (BETRNet) model fitting clinical data to a familial paradigm. Cancer Epidemiol Biomarkers Prev 2016;25:727–735.
27. 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.
28. Yang J, Lee SH, Goddard ME, et al. GCTA: a tool for genome-wide complex trait analysis. Am J Hum Genet 2011;88:76–82.
29. Palles C, Chegwidden L, Li X, et al. Polymorphisms near TBX5 and GDF7 are associated with increased risk for Barrett's esophagus. Gastroenterology 2015;148: 367–378.
30. Cook MB, Corley DA, Murray LJ, et al. Gastroesophageal reflux in relation to adenocarcinomas of the esophagus: a

pooled analysis from the Barrett's and Esophageal Adenocarcinoma Consortium (BEACON). PLoS One 2014;9:e103508.

31. Corley DA, Kubo A, Levin TR, et al. Abdominal obesity and body mass index as risk factors for Barrett's esophagus. Gastroenterology 2007;133:34–41.
32. Edelstein ZR, Farrow DC, Bronner MP, et al. Central adiposity and risk of Barrett's esophagus. Gastroenterology 2007;133:403–411.
33. Hoyo C, Cook MB, Kamangar F, et al. Body mass index in relation to oesophageal and oesophagogastric junction adenocarcinomas: a pooled analysis from the International BEACON Consortium. Int J Epidemiol 2012;41:1706–1718.
34. Romero Y, Cameron AJ, Locke GR 3rd, et al. Familial aggregation of gastroesophageal reflux in patients with Barrett's esophagus and esophageal adenocarcinoma. Gastroenterology 1997;113:1449–1456.
35. Cameron AJ, Lagergren J, Henriksson C, et al. Gastroesophageal reflux disease in monozygotic and dizygotic twins. Gastroenterology 2002;122:55–59.
36. 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.
37. Mohammed I, Cherkas LF, Riley SA, et al. Genetic influences in gastro-oesophageal reflux disease: a twin study. Gut 2003;52:1085–1089.
38. Gharahkhani P, Tung J, Hinds D, et al. Chronic gastroesophageal reflux disease shares genetic background with esophageal adenocarcinoma and Barrett's esophagus. Hum Mol Genet 2016;25:828–835.
39. Apalasamy YD, Mohamed Z. Obesity and genomics: role of technology in unraveling the complex genetic architecture of obesity. Hum Genet 2015;134:361–374.
40. **Locke AE, Kahali B, Berndt SI**, et al. Genetic studies of body mass index yield new insights for obesity biology. Nature 2015;518:197–206.
41. Thrift AP, Shaheen NJ, Gammon MD, et al. Obesity and risk of esophageal adenocarcinoma and Barrett's esophagus: a mendelian randomization study. J Natl Cancer Inst 2014:106.
42. **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.
43. 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.
44. Becker J, May A, Gerges C, et al. Supportive evidence for FOXP1, BARX1, and FOXF1 as genetic risk loci for the development of esophageal adenocarcinoma. Cancer Med 2015;4:1700–1704.
45. **Gharahkhani P, Fitzgerald RC, Vaughan TL**, et al. Genome-wide association studies in oesophageal adenocarcinoma and Barrett's oesophagus: a large-scale meta-analysis. Lancet Oncol 2016;17:1363–1373.
46. Zehir A, Benayed R, Shah RH, et al. Mutational landscape of metastatic cancer revealed from prospective clinical sequencing of 10,000 patients. Nat Med 2017;23:703–713.
47. 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.
48. Pauwels A, Blondeau K, Dupont LJ, et al. Mechanisms of increased gastroesophageal reflux in patients with cystic fibrosis. Am J Gastroenterol 2012;107:1346–1353.
49. Houghton LA, Lee AS, Badri H, et al. Respiratory disease and the oesophagus: reflux, reflexes and microaspiration. Nat Rev Gastroenterol Hepatol 2016;13:445–460.
50. Schrijver I, Pique L, Graham S, et al. The spectrum of CFTR variants in nonwhite cystic fibrosis patients: implications for molecular diagnostic testing. J Mol Diagnostics 2016;18:39–50.
51. Meltzer SJ. Leaky transporters and sphincters in Barrett's oesophagus? Lancet Oncol 2016;17:1336–1337.
52. **Agrawal N, Jiao Y, Bettegowda C**, et al. Comparative genomic analysis of esophageal adenocarcinoma and squamous cell carcinoma. Cancer Discov 2012;2:899–905.
53. **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.
54. **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.
55. Barrett MT, Sanchez CA, Prevo LJ, et al. Evolution of neoplastic cell lineages in Barrett oesophagus. Nat Genet 1999;22:106–109.
56. Blount PL, Galipeau PC, Sanchez CA, et al. 17p allelic losses in diploid cells of patients with Barrett's esophagus who develop aneuploidy. Cancer Res 1994;54:2292–2295.
57. Barrett MT, Sanchez CA, Galipeau PC, et al. Allelic loss of 9p21 and mutation of the CDKN2/p16 gene develop as early lesions during neoplastic progression in Barrett's esophagus. Oncogene 1996;13:1867–1873.
58. Buas MF, Levine DM, Makar KW, et al. Integrative post genome-wide association analysis of CDKN2A and TP53 SNPs and risk of esophageal adenocarcinoma. Carcinogenesis 2014;35:2740–2747.
59. Reid BJ, Li X, Galipeau PC, et al. Barrett's oesophagus and oesophageal adenocarcinoma: time for a new synthesis. Nat Rev Cancer 2010;10:87–101.
60. Buas MF, He Q, Johnson LG, et al. Germline variation in inflammation-related pathways and risk of Barrett's oesophagus and oesophageal adenocarcinoma [published online ahead of print August 2, 2016]. Gut. http://dx.doi.org/10.1136/gutjnl-2016-311622.
61. Buas MF, Onstad L, Levine DM, et al. MiRNA-related SNPs and risk of esophageal adenocarcinoma and Barrett's esophagus: post genome-wide association analysis in the BEACON Consortium. PLoS One 2015;10:e0128617.
62. Ek WE, Lagergren K, Cook M, et al. Polymorphisms in genes in the androgen pathway and risk of Barrett's

esophagus and esophageal adenocarcinoma. Int J Cancer 2016;138:1146–1152.

63. **Lagergren K, Ek WE**, Levine D, et al. Polymorphisms in genes of relevance for oestrogen and oxytocin pathways and risk of Barrett's oesophagus and oesophageal adenocarcinoma: a pooled analysis from the BEACON Consortium. PLoS One 2015;10:e0138738.

64. Dai JY, de Dieu Tapsoba J, Buas MF, et al. A newly identified susceptibility locus near FOXP1 modifies the association of gastroesophageal reflux with Barrett's esophagus. Cancer Epidemiol Biomarkers Prev 2015; 24:1739–1747.

65. Dai JY, de Dieu Tapsoba J, Buas MF, et al. Constrained score statistics identify genetic variants interacting with multiple risk factors in Barrett's esophagus. Am J Hum Genet 2016;99:352–365.

66. Lee E, Stram DO, Ek WE, et al. Pleiotropic analysis of cancer risk loci on esophageal adenocarcinoma risk. Cancer Epidemiol Biomarkers Prev 2015;24:1801–1803.

67. Prevo LJ, Sanchez CA, Galipeau PC, et al. p53-mutant clones and field effects in Barrett's esophagus 1. 1999;59:4784–4787.

68. Wong DJ, Barrett MT, Stöger R, et al. p16 INK4a Promoter is hypermethylated at a high frequency in esophageal adenocarcinomas. Cancer Res 1997;57:2619–2622.

69. 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 2014;7:114–127.

70. **Stachler MD, Taylor-Weiner A**, Peng S, et al. Paired exome analysis of Barrett's esophagus and adenocarcinoma. Nat Genet 2015;47:1047–1055.

71. Galipeau PC, Prevo LJ, Sanchez CA, et al. Clonal expansion and loss of heterozygosity at chromosomes 9p and 17p in premalignant esophageal (Barrett's) tissue. J Natl Cancer Inst 1999;91:2087–2095.

72. Reid BJ, Haggitt RC, Rubin CE, et al. Barrett's esophagus. Correlation between flow cytometry and histology in detection of patients at risk for adenocarcinoma. Gastroenterology 1987;93:1–11.

73. Levine DS, Reid BJ, Haggitt RC, et al. Correlation of ultrastructural aberrations with dysplasia and flow cytometric abnormalities in Barrett's epithelium. Gastroenterology 1989;96:355–367.

74. **Weaver JMJ, Ross-Innes CS**, Fitzgerald RC. The "-omics" revolution and oesophageal adenocarcinoma. Nat Rev Gastroenterol Hepatol 2014;11:19–27.

75. **Weaver JMJ, Ross-Innes CS, Shannon N**, et al. Ordering of mutations in preinvasive disease stages of esophageal carcinogenesis. Nat Genet 2014;46:837–843.

76. **Kandoth C, McLellan MD**, Vandin F, et al. Mutational landscape and significance across 12 major cancer types. Nature 2013;502:333–339.

77. Martinez P, Timmer MR, Lau CT, et al. Dynamic clonal equilibrium and predetermined cancer risk in Barrett's oesophagus. Nat Commun 2016;7:12158.

78. **Secrier M, Li X**, N de Silva, et al. Mutational signatures in esophageal adenocarcinoma define etiologically distinct subgroups with therapeutic relevance. Nat Genet 2016; 48:1131–1141.

79. Paterson AL, Shannon NB, Lao-Sirieix P, et al. A systematic approach to therapeutic target selection in oesophago-gastric cancer. Gut 2012;62:1415–1424.

80. **Dulak AM, Schumacher SE**, van Lieshout J, et al. Gastrointestinal adenocarcinomas of the esophagus, stomach, and colon exhibit distinct patterns of genome instability and oncogenesis. Cancer Res 2012;72: 4383–4393.

81. Galipeau PC, Cowan DS, Sanchez CA, et al. 17p (p53) allelic losses, 4N (G2/tetraploid) populations, and progression to aneuploidy in Barrett's esophagus. Proc Natl Acad Sci U S A 1996;93:7081–7084.

82. Kadri SR, Lao-Sirieix P, O'Donovan M, et al. Acceptability and accuracy of a non-endoscopic screening test for Barrett's oesophagus in primary care: cohort study. BMJ 2010;341:c4372.

83. Ross-Innes CS, Debiram-Beecham I, O'Donovan M, et al. Evaluation of a minimally invasive cell sampling device coupled with assessment of trefoil factor 3 expression for diagnosing Barrett's esophagus: a multi-center case-control study. PLoS Med 2015;12:1–19.

84. **Ross-Innes CS, Chettouh H, Achilleos A**, et al. Risk stratification of Barrett's oesophagus using a non-endoscopic sampling method coupled with a biomarker panel: a cohort study. Lancet Gastroenterol Hepatol 2016;3:1–9.

85. **Nones K, Waddell N, Wayte N**, et al. Genomic catastrophes frequently arise in esophageal adenocarcinoma and drive tumorigenesis. Nat Commun 2015;5:1–9.

86. Helleday T, Eshtad S, Nik-Zainal S. Mechanisms underlying mutational signatures in human cancers. Nat Rev Genet 2014;15:585–598.

87. Alexandrov LB, Nik-Zainal S, Wedge DC, et al. Deciphering signatures of mutational processes operative in human cancer. Cell Rep 2013;3:246–259.

88. Nik-Zainal S, Alexandrov LB, Wedge DC, et al. Mutational processes molding the genomes of 21 breast cancers. Cell 2012;149:979–993.

89. Tomasetti C, Marchionni L, Nowak MA, et al. Only three driver gene mutations are required for the development of lung and colorectal cancers. Proc Natl Acad Sci U S A 2015;112:118–123.

90. Dong C, Guo Y, Yang H, et al. iCAGES: integrated CAncer GEnome Score for comprehensively prioritizing cancer driver genes in personal genomes. Genome Med 2016;8:1–22.

91. Kim JJ, Bowlby R, Mungall AJ, et al. Integrated genomic characterization of oesophageal carcinoma. Nature 2017;541:169–175.

92. Paterson AL, Donovan MO, Provenzano E, et al. Characterization of the timing and prevalence of receptor tyrosine kinase expression changes in oesophageal carcinogenesis. J Pathol 2013;230:118–128.

93. Kim J, Fox C, Peng S, et al. Preexisting oncogenic events impact trastuzumab sensitivity in ERBB2-amplified gastroesophageal adenocarcinoma. J Clin Invest 2014; 124:5145–5158.

94. Saldivar JC, Miuma S, Bene J, et al. Initiation of genome instability and preneoplastic processes through loss of Fhit expression. PLoS Genet 2012;8:e1003077.

95. Iliopoulos D, Guler G, Han SY, et al. Roles of FHIT and WWOX fragile genes in cancer. Cancer Lett 2006; 232:27–36.
96. LeTallec B, Millot G, Blin M, et al. Common fragile site profiling in epithelial and erythroid cells reveals that most recurrent cancer deletions lie in fragile sites hosting large genes. Cell Rep 2013;4:420–428.
97. Durkin SG, Glover TW. Chromosome fragile sites. Annu Rev Genet 2007;41:169–192.
98. Paterson AL, Weaver JMJ, Eldridge MD, et al. Mobile element insertions are frequent in oesophageal adenocarcinomas and can mislead paired-end sequencing analysis. BMC Genomics 2015;16:473.
99. Helman E, Lawrence MS, Stewart C, et al. Somatic retrotransposition in human cancer revealed by whole-genome and exome sequencing. Genome Res 2014; 24:1053–1063.
100. Korbel JO, Campbell PJ. Criteria for inference of chromothripsis in cancer genomes. Cell 2013;152:1226–1236.
101. **Garsed DW, Marshall OJ, Corbin VDA**, et al. The architecture and evolution of cancer neochromosomes. Cancer Cell 2014;26:653–667.
102. Baca SC, Prandi D, Lawrence MS, et al. Punctuated evolution of prostate cancer genomes. Cell 2013; 153:666–677.
103. **Zhang C-Z, Spektor A, Cornils H**, et al. Chromothripsis from DNA damage in micronuclei. Nature 2015;522: 179–184.
104. **Zhang C-Z, Leibowitz ML**, Pellman D. Chromothripsis and beyond: rapid genome evolution from complex chromosomal rearrangements. Genes Dev 2013; 27:2513–2530.
105. Storlazzi CT, Fioretos T, Surace C, et al. MYC-containing double minutes in hematologic malignancies: evidence in favor of the episome model and exclusion of MYC as the target gene. Hum Mol Genet 2006;15:933–942.
106. Parrales A, Iwakuma T. Targeting oncogenic mutant p53 for cancer therapy. Front Oncol 2015;5:1–13.
107. Niroula A, Vihinen M. Variation interpretation predictors: principles, types, performance, and choice. Hum Mutat 2016;37:579–597.
108. **Samyn B, Samyn B, Sergeant K**, et al. Pathway and network analysis of cancer genomes. Nat Methods 2015; 2:1–6.
109. **Mcgranahan N, Furness AJS, Rosenthal R**, et al. Clonal neoantigens elicit T cell immunoreactivity and sensitivity to immune checkpoint blockade. Science 2016;351:1463–1469.
110. **Bettegowda C, Sausen M, Leary RJ**, et al. Detection of circulating tumor DNA in early- and late-stage human malignancies. Sci Transl Med 2014;6. 224ra24–224ra24.
111. Siravegna G, Mussolin B, Buscarino M, et al. Clonal evolution and resistance to EGFR blockade in the blood of colorectal cancer patients. Nat Med 2015; 21:795–801.
112. Noordman BJ, Shapiro J, Spaander MC, et al. Accuracy of detecting residual disease after cross neoadjuvant chemoradiotherapy for esophageal cancer (preSANO trial): rationale and protocol. JMIR Res Protoc 2015;4:e79.
113. Michailidou K, Beesley J, Lindstrom S, et al. Genome-wide association analysis of more than 120,000 individuals identifies 15 new susceptibility loci for breast cancer. Nat Genet 2015;47:373–380.
114. Vaughan TL. From genomics to diagnostics of esophageal adenocarcinoma. Nat Genet 2014;46:806–807.
115. Bunting SF, Nussenzweig A. End-joining, translocations and cancer. Nat Rev Cancer 2013;13:443–454.
116. An O, Dall'Olio GM, Mourikis TP, et al. NCG 5.0: updates of a manually curated repository of cancer genes and associated properties from cancer mutational screenings. Nucleic Acids Res 2016;44:D992–D999.
117. Murugaesu N, Wilson GA, Birkbak NJ, et al. Tracking the genomic evolution of esophageal adenocarcinoma through neoadjuvant chemotherapy. Cancer Discov 2015;5:821–832.
118. **Noorani A, Bornschein J, Lynch AG**, et al. A comparative analysis of whole genome sequencing of esophageal adenocarcinoma pre- and post-chemotherapy. Genome Res 2017;27:902–912.

Author names in bold designate shared co-first authorship.

Received February 11, 2017. Accepted July 11, 2017.

Reprint requests

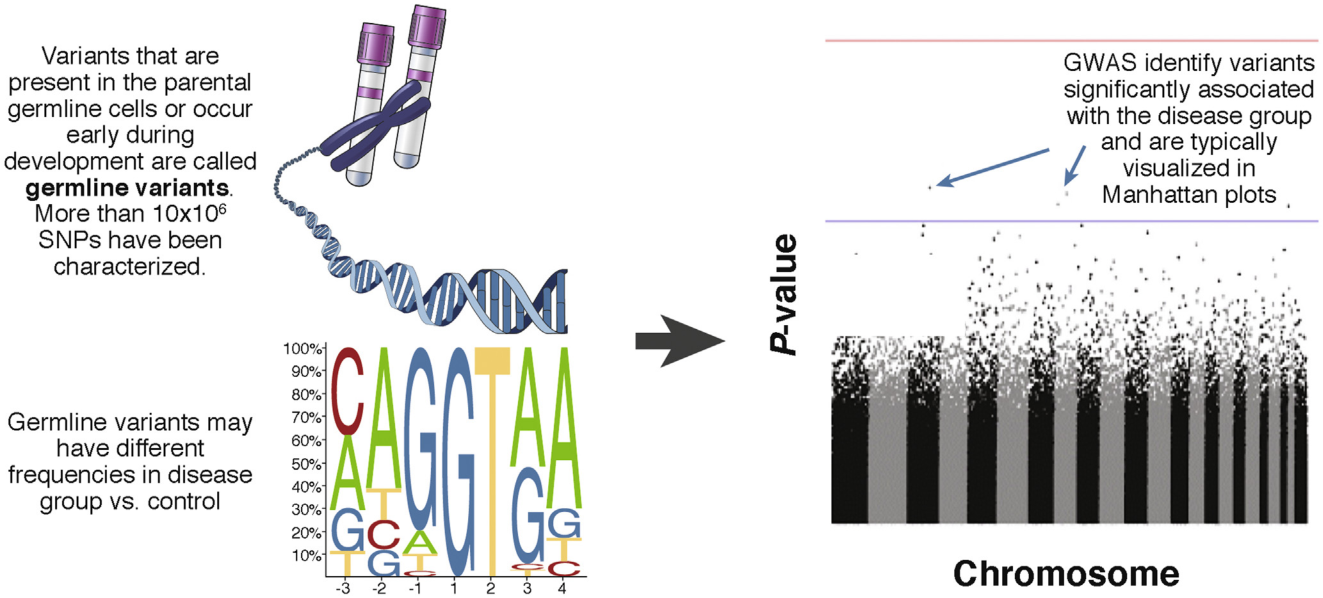
Address requests for reprints to: Gianmarco Contino, MD, PhD, Medical Research Council Cancer Unit, Hutchison/Medical Research Council Research Centre, Box 197 Cambridge Biomedical Campus, University of Cambridge, Cambridge, UK CB2 0XZ. e-mail: gc502@mrc-cam.ac.uk; fax: +44 (0)1223 763241.

Conflicts of interest

R.C.F. is named on patents pertaining to the Cytosponge and associated assays that have been licensed by the Medical Research Council to Covidien (now Medtronic). All the authors declare no other competing interests.

Funding

G.C. is a National Institute for Health Research (NIHR) Lecturer as part of an NIHR professorship grant to R.C.F. T.L.V. is supported in part by National Institutes of Health (NIH) K05CA124911 and NIH R21 CA197502. D.W. is supported by a Research Fellowship (APP1058522) from the National Health and Medical Research Council of Australia. R.C.F. is funded by an NIHR professorship (RG67258), the Medical Research Council (RG74187), the Biomedical Research Centre and the Experimental Cancer Medicine Centre, and a program grant from Cancer Research UK (RG66287).


Variants that are present in the parental germline cells or occur early during development are called **germline variants**. More than 10x10^6 SNPs have been characterized.
Germline variants may have different frequencies in disease group vs. control
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
-3
-2
-1
1
2
3
4
GWAS identify variants significantly associated with the disease group and are typically visualized in Manhattan plots
P-value
Chromosome

Supplementary Figure 1. GWAS analysis of germline DNA (from peripheral blood) can identify variants associated with the development of a condition.