**Gastroenterology 2017;153:657**–**673**

**REVIEWS AND**

**PERSPECTIVES**

**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. Fitzgerald1

*1Medical Research Council Cancer Unit, Hutchison/Medical Research Council Research Centre, University of Cambridge, Cambridge, UK; 2Cancer Epidemiology Program, Fred Hutchinson Cancer Research Center, Seattle, Washington; 3Department of Epidemiology, School of Public Health, University of Washington, Seattle, Washington; and*

*4Cancer 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 inva- sive cancer. These technologies led to identi**ﬁ**cation 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.

O

ver 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, aﬁgure 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 develop- ment 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 Sharma6). BE, which is found in approximately 1% to 2% of the general popula- tion, is a squamocolumnar metaplasia that develops in response to gastroesophageal reﬂux (GER). BE pathogen- esis involves a combination of anatomic (hiatus hernia), genetic, and lifestyle risk factors.6–10 Neoplastic trans- formation of NDBE usually occurs through progressive grades of dysplasia. Endoscopic treatment is recom- mended 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 endo- scopic 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 endo- scopic imaging technologies and sampling bias causes many dysplastic lesions to be missed. There is also low interob- server 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-gener- ation 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 translatingﬁndings from genomic analyses into screening, diagnostic, and therapeutic strategies.

**Abbreviations used in this paper: APOBEC, apolipoprotein B mRNA edit- ing enzyme catalytic polypeptide-like; BE, Barrett**’**s esophagus; BEACON, Barrett**’**s and Esophageal Adenocarcinoma Consortium; BMI, body mass index; EA, esophageal adenocarcinoma; GER, gastroesophageal re**ﬂ**ux; 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**

**REVIEWSAND**

**PERSPECTIVES**

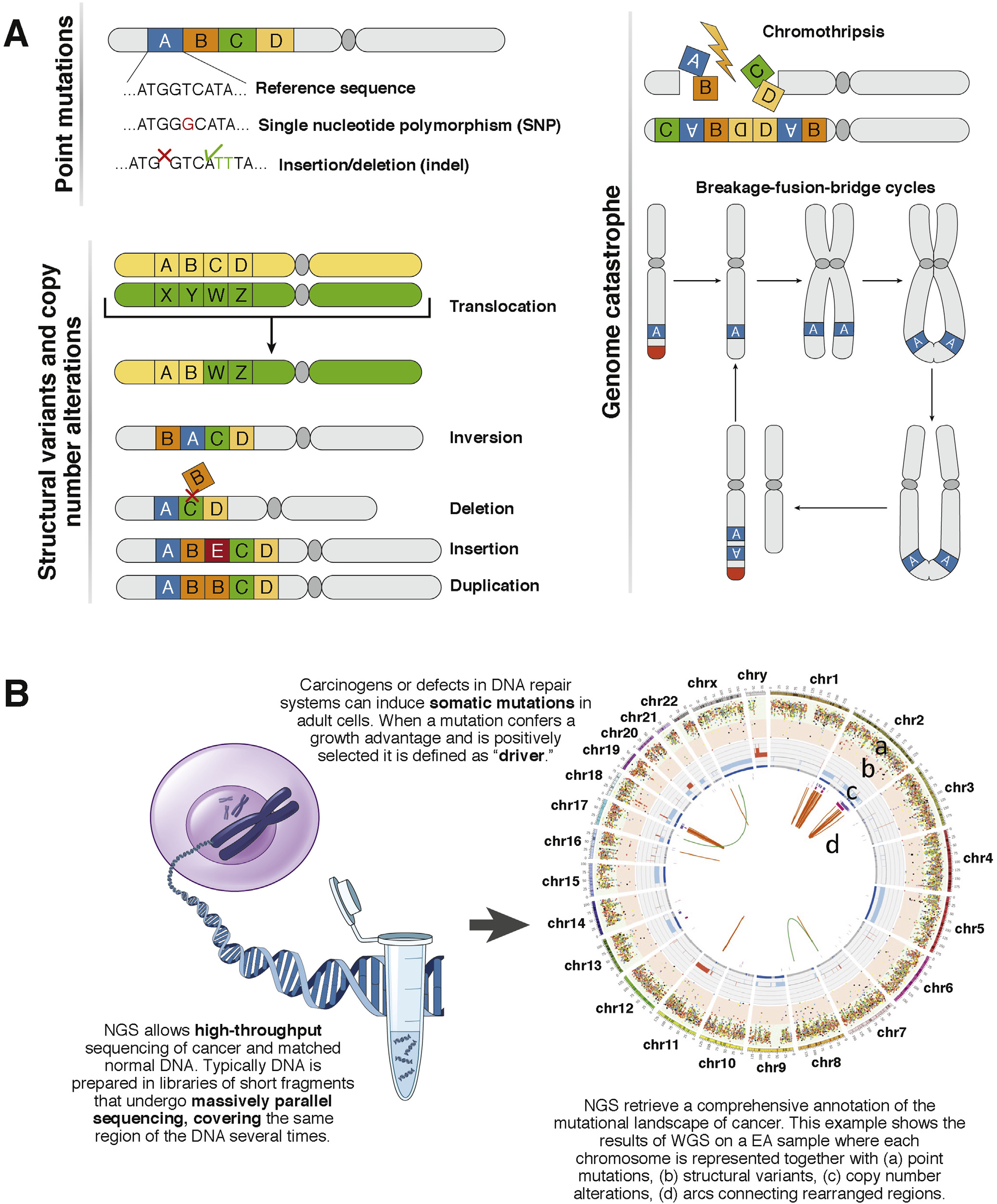
**658 Contino et al Gastroenterology Vol. 153, No. 3**

**Germline Variations and Susceptibility**

*Family Studies*

Evidence that germline mutations contribute to devel-

opment of EA originated from reports of familial

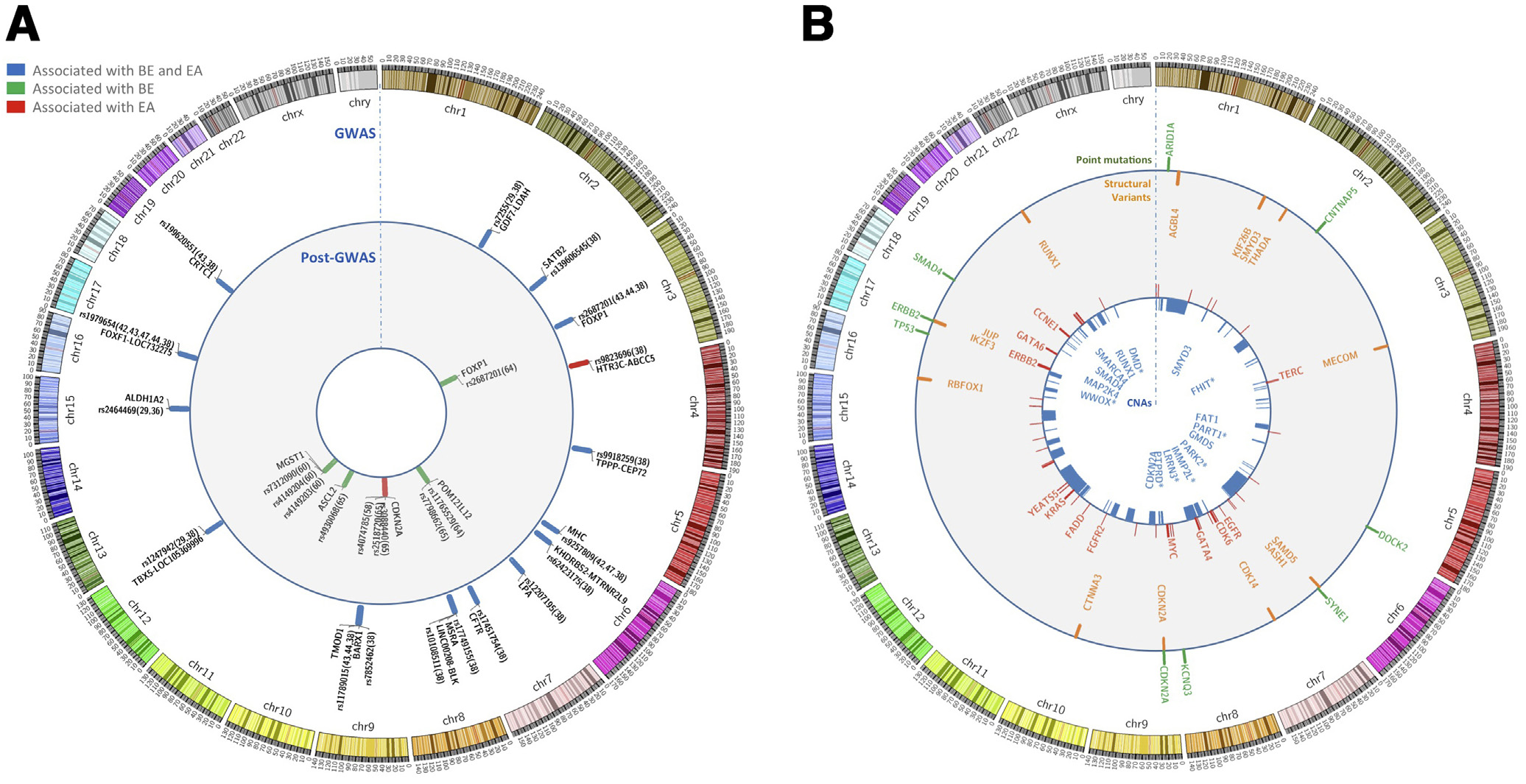


aggregation of this cancer19 and BE.20–23 Orloff et al24 per- formed linkage analyses comparing 21 concordant affected sibling pairs (42 siblings with BE and/or EA) and 11 discordant sibling pairs using a 100K single nucleotide polymorphism (SNP) set. Subsequent ﬁne-mapping of

**REVIEWS AND**

**PERSPECTIVES**

**September 2017 Evolving Genomic Landscape of BE and EA 659**



**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 conﬁr- matory 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 al91 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 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 candi- date genes for validation, performed in an independent set

of 58 persons with BE or EA. Variants in *MSR1*, on chro- mosome 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 asso- ciations have not been identiﬁed.

The extent to which BE or EA (including adenocarci- noma 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 un- affected 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 esti- mated for a 50-year-old woman, but applied to a much lower baseline risk (0.5%.) However, when the discrimina- tion 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 con- tained 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, struc- tural 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.

**REVIEWSAND**

**PERSPECTIVES**

**660 Contino et al Gastroenterology Vol. 153, No. 3**

and the strength and higher prevalence of the other estab- lished 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 anal- ysis, Ek et al27 estimated that 25% (standard error [SE], 5%; 1-sided *P*¼.0000002) of EA cases and 35% of BE cases (SE, 6%; 1-sided *P* ¼ .000000001) were determined by the composite effect of many common mutations of small indi- vidual relative risk.28 Furthermore, they demonstrated substantial polygenic overlap between EA and BE, indicating that shared genes inﬂuence the development of the 2 dis-

orders. No other studies have reported on the EA genetic variance explained, nor on the overlap between EA and BE.

However, Palles et al29 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 al38 esti- mated heritability based on genotype arrays and reported that 7% of the variance in GER symptoms could be explained by genetic factors. Furthermore, they found evi- dence 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/m2 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 Well- come 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 (n¼2390) in addition to BE cases (n¼3175), was con- ducted 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 pre- viously 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* ¼ .00000008). The eighth, on chromosome 9q22, narrowly missed this threshold (*P* ¼ .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 devel-

opment of BE and EA, and the fact that hiatal hernia substantially predisposes to GER, the ﬁndings identify mechanisms by which these variants might affect develop- ment 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 devel-

opment and differentiation, associated with these conditions.45

The large meta-analysis identiﬁed an intriguing associ- ation 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 char- acterized 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

**REVIEWS AND**

**PERSPECTIVES**

**September 2017 Evolving Genomic Landscape of BE and EA 661**

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 malig- nancies)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 asso- ciated 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 pre- dicted 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 associ- ated with reduced risks of progression. Expression of one of the variants (rs3088440) in cell lines indicated that it re- duces 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 associ- ated 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 asso- ciations, 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 dif- ferences in statistical associations across strata of exposure to those factors (eg, BMI, sex). Using the well-annotated

BEACON GWAS, Dai et al64 examined the ﬁrst 7 SNPs identiﬁed as associated with BE or EA at the genome-wide

level of signiﬁcance for interactions with BMI, GER symp- toms, 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*interaction ¼.0005; false discovery rate¼0.042.)

Dai et al65 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 al66 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, muta- tions 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 Inter- national Cancer Genome Consortium (http://icgc.org). New data are being added every day.

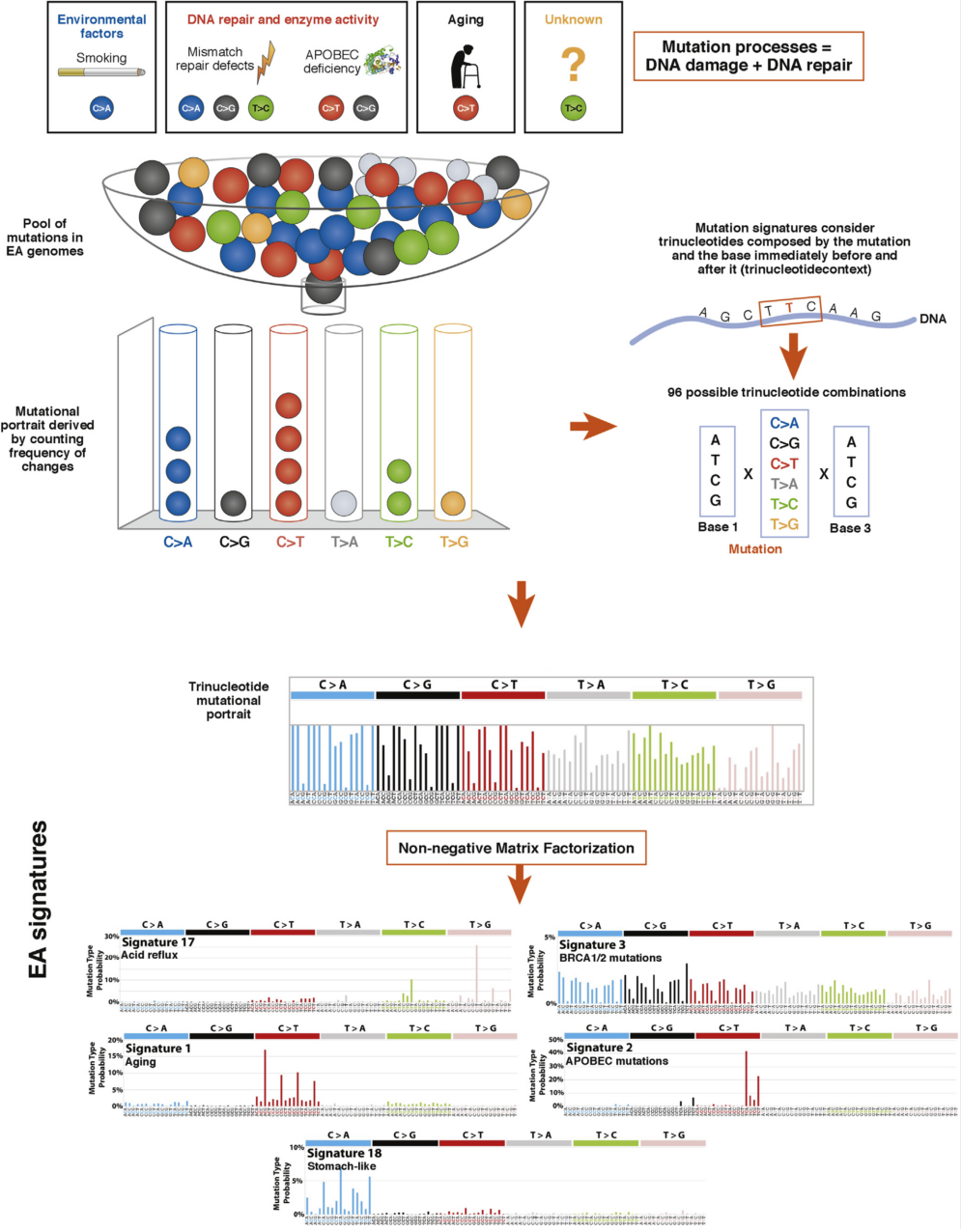
*Progression from Premalignant BE to EA*

There was a reasonable expectation that sequencing the genomes of BE or EA tissues would identify somatic alter- ations 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 pro- gression to invasive cancer. In our current model, dysplasia progresses to invasive EA via early loss of *CDKN2A*, emer- gence 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 pro- gression can be nonlinear.74,75

**REVIEWSAND**

**PERSPECTIVES**

**662 Contino et al Gastroenterology Vol. 153, No. 3**



**REVIEWS AND**

**PERSPECTIVES**

**September 2017 Evolving Genomic Landscape of BE and EA 663**

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 adja- cent 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 alterationsﬁtsinto the linear multistep process of BE pro- gression, 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 life- time 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 resec- tion 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 sam- pling is critical, because BE is a polyclonal disease and endoscopic biopsies have inherent sampling bias. Fortu- nately, 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 Cyto- sponge 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

=

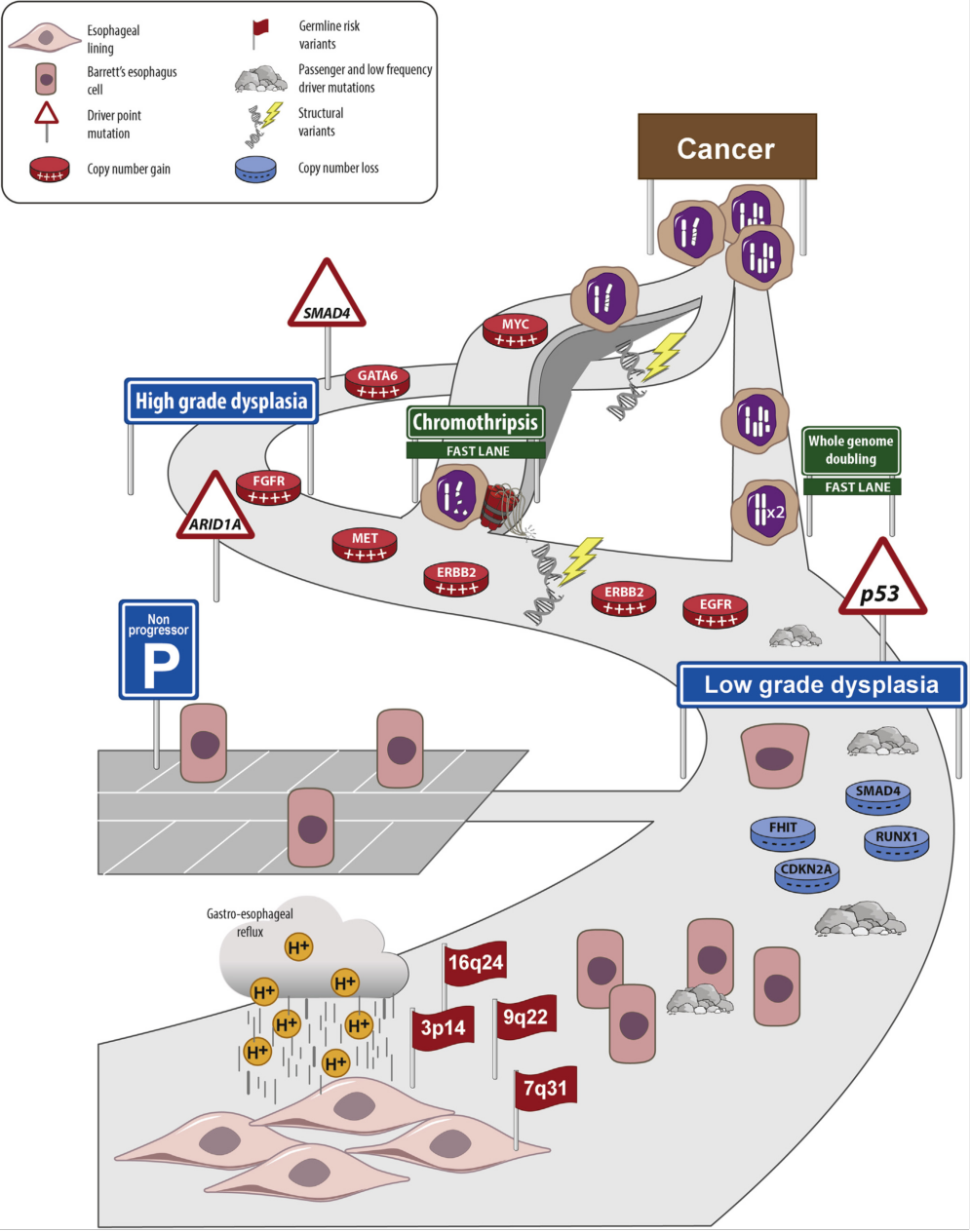
**Figure 3.** Mutational signatures of tumors. (*A*) Mutational processes are biological activities (eg, aging, smoking, UV light exposure, unknown carcinogens) that generate patterns of mutations (mutational signatures) through a damage of the DNA sequence and its attempt to repair it by DNA repair 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 al87 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

**REVIEWSAND**

**PERSPECTIVES**

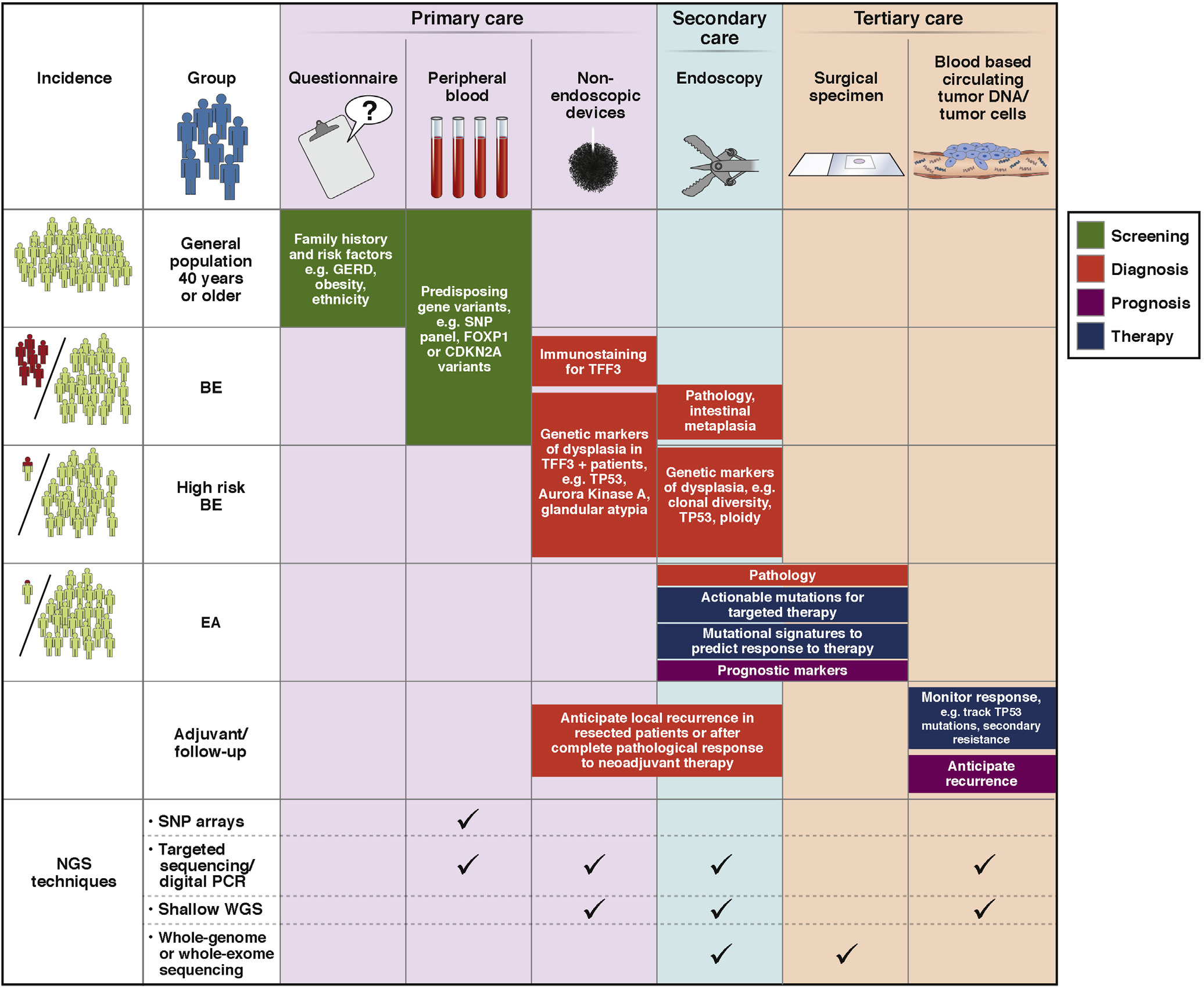
**664 Contino et al Gastroenterology Vol. 153, No. 3**



**REVIEWS AND**

**PERSPECTIVES**

**September 2017 Evolving Genomic Landscape of BE and EA 665**



**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 immuno- staining) 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 posi- tivity), 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 cohort84 (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 (eachﬂ*ag* 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 chro- mothripsis and whole-genome doubling, can occur at any stage and dramatically accelerate progression of BE.

**REVIEWSAND**

**PERSPECTIVES**

**666 Contino et al Gastroenterology Vol. 153, No. 3**

**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 prolif- eration, and the inﬂammatory response. Although no mutagen has been convincingly proven to cause EA, carci- nogenesis is believed to involve acid and bile reﬂux. Little is known about the mechanisms by which these luminal con- stituents might cause DNA damage; inherited mismatch

repair gene deﬁciencies are not commonly observed.78

One method to identify and classify mutational pro- cesses 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 aging86 (Figure 3). Alexandrov et al87 created a catalogue of these signatures, using a non-negative matrix factorization algo- rithm. 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 treatment78 (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 muta- tions 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 los- ses 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 targets79,92,93 (Figures 2 and 4, Supplementary Table 2).

Rearrangements are variably distributed in the genomes

of EA samples. Nones et al85 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 sig- niﬁcance is unclear. For instance, the fragile histidine triad gene (*FHIT* or *FRA3B*) and WW domain containing oxido- reductase 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 varia- tions could be a common mechanism of recurrent mutation in EA. *RUNX1*, a gene translocated in acute myeloid leuke- mia, 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