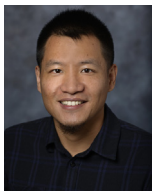
Gastroenterology 2018;154:374–389

Genomic and Epigenomic Aberrations in Esophageal Squamous Cell Carcinoma and Implications for Patients



De-Chen Lin[1] Ming-Rong Wang[2] H. Phillip Koeffler[1,3,4]

[1]Department of Medicine, Cedars-Sinai Medical Center, Los Angeles, California; [2]State Key Laboratory of Molecular Oncology, National Cancer Center/Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China; [3]Cancer Science Institute of Singapore, National University of Singapore, Singapore; and [4]National University Cancer Institute, National University Hospital Singapore, Singapore

Esophageal squamous cell carcinoma (ESCC) is a common malignancy without effective therapy. The exomes of more than 600 ESCCs have been sequenced in the past 4 years, and numerous key aberrations have been identified. Recently, researchers reported both inter- and intratumor heterogeneity. Although these are interesting observations, their clinical implications are unclear due to the limited number of samples profiled. Epigenomic alterations, such as changes in DNA methylation, histone acetylation, and RNA editing, also have been observed in ESCCs. However, it is not clear what proportion of ESCC cells carry these epigenomic aberrations or how they contribute to tumor development. We review the genomic and epigenomic characteristics of ESCCs, with a focus on emerging themes. We discuss their clinical implications and future research directions.

Keywords: Esophageal Cancer; Genomics; Epigenomics.

Esophageal cancer is the sixth leading cause of cancer-related mortality worldwide,[1,2] with more than 480,000 new cases diagnosed yearly.[2] More than 80% of esophageal cancers are esophageal squamous cell carcinomas (ESCCs).[3] With the advent of new biochemical technologies (particularly next-generation sequencing), hundreds of ESCCs have been profiled in a comprehensive and unbiased manner using whole-exome sequencing (WES) in the past 4 years. Our understanding of the genomic features of this cancer has therefore greatly advanced.

Genomic analyses provide evidence for pervasive abnormalities in the ESCC epigenome: epigenetic regulators themselves are frequently altered by genetic changes. Although the ESCC epigenome per se has not been characterized as extensively as its genome, many epigenetic dysregulations have been recently discovered, and their biologic significance determined. However, these advances in our understanding of the molecular alterations in ESCC have not been translated to the bedside.

There is no sensitive method for early detection of ESCC, so more than half of patients with these tumors have distal metastases at the time of diagnosis; only 10% to 20% survive for 5 years.[3] Several tumor types have genetic alterations that can be targeted therapeutically, such as $HER2^{+}$ breast tumors, $EGFR^{+}$ lung tumors, and $BRAF^{+}$ melanomas, but such actionable alterations have not been identified in ESCCs. We review genomic and epigenomic aberrations found in ESCC, and discuss how these findings have increased our understanding of ESCC pathogenesis and progression. These acquired lesions might be used as biomarkers or targeted by therapeutic agents.

Somatic Alterations in ESCC Genomes

Copy Number Alterations and Structural Rearrangements

Chromosome aneuploidy and arm-level aberrations in ESCCs, discovered using traditional approaches (such as karyotyping and fluorescence in situ hybridization), have been summarized elsewhere.[4] We focus on focal copy number alterations (CNAs) and structural rearrangements identified by high-resolution and high-throughput methods,

Abbreviations used in this paper: **ADAR, adenosine deaminases acting on RNA; APOBEC, apolipoprotein B mRNA editing enzyme, catalytic polypeptide-like; BCH, basal cell hyperplasia; CIS, carcinoma in situ; CNA, copy number alteration; ESCC, esophageal squamous cell carcinoma; FGFR, fibroblast growth factor receptor; HGIEN, high-grade IEN; HNSCC, head and neck squamous cell carcinoma; IEN, intraepithelial neoplasia; KLF5, Krüppel-like factor 5; LGIEN, low-grade IEN; LINE-1, long interspersed nuclear element-1; LUSC, squamous cell carcinoma of the lung; NRF2, nuclear factor, erythroid 2 like 2; SMG, significantly mutated gene; WES, whole-exome sequencing.**

Most current article

0016-5085/$36.00
https://doi.org/10.1053/j.gastro.2017.06.066

including single-nucleotide polymorphism microarrays and next-generation sequencing.

Compared with aneuploidy and arm-level alterations, focal CNAs have a higher probability of making changes to genes that provide cancer cells with a proliferative advantage.[5,6] The most frequent high-level amplifications and homozygous deletions involve 11q13.2-q13.3 and 9p21.3 (locus of *CDKN2A* and *CDKN2B*), respectively (Table 1). Copy number gains increase levels of mRNAs transcribed from more than 80% of genes situated at 11q13.[7] These include oncogenes, such as *CPT1A*, *ANO1*, *ORAOV1*, *CCND1*, *FGF3, FGF4, FGF19*, *CTTN*, and *MIR548K*; their contributions to the malignant phenotype of ESCC cells have been verified.[8–11] The genomic material between 11q13.2 and 11q13.3 is nonuniformly increased, with the most recurrent peak spanning *CCND1.*[12]

CCND1 is amplified by breakage fusion bridge cycles, which occurs under conditions of chromosomal instability.[13] This gene is often coamplified with its neighboring oncogene *CTTN;* its product promotes migration of ESCC cells.[14] Additional recurrent focal amplifications in ESCC include those at 8p11.23 (*FGFR1*), 8q24.21 (*MYC*), 7p11.2 (*EGFR*), 12p12.1 (*KRAS*), 12q15 (*MDM2*), 3q26 (*TP63* and *PRKCI*), 3q26.32–q26.33 (*SOX2* and *PIK3CA*), and 14q13.3 (*NKX2-1*). Other frequent homozygous deletions contain 2q22.1–q22.2 (*LRP1B*), 9p24.1 (*PTPRD*), and 3p14.2 (*FHIT*). All of these genes have been validated as drivers of development of ESCC or other cancers.[12,13,15–27]

Recently, researchers also have performed high-resolution characterization of structural rearrangements in ESCCs, including intrachromosome insertions, inversions and duplications, and interchromosome translocations. These have been identified by WGS and associated mathematic analyses.

A total of 62 ESCC genomes have been characterized using the WGS platform, and more than 1000 alterations were identified as structural rearrangements.[8,13,28,29] Most structural rearrangements are not likely to have pathogenic potential, but a few recurrent ones might be candidate driver events, such as frequent structural breakpoints affecting the *KCNB2* gene,[8] or an in-frame fusion between *TRAPPC9* and *CLVS1*. This fusion results from chromothripsis-associated rearrangements, in which thousands of clustered rearrangements occur during a single event in confined genomic regions. Other possible drivers of ESCC development include an in-frame fusion between *EIF3E* and *RAD51B*, caused by interchromosomal translocations.[13] The tumor-promoting effects of TRAPPC9 and RAD51B have been reported in other types of cancers,[30–32] so these rearrangements could have important biological effects in ESCC cells. Additional structural rearrangements involving potential oncogenes included *MYBL2* duplication, fusions of *RUNX1T1–PHACTR1*, *MAML2–TTC28*, *ASXL1–RNF170*, and *FGF19–SHANK2.*[29] Despite the discovery of these interesting rearrangements, the overall number of WGS-profiled ESCC samples is small. Large-scale, more uniformly processed WGSs are needed to identify the genomic rearrangements that contribute to development of ESCC.

Somatic Mutations

Sequence analyses of ESCC tumor–germline paired exomes[6,8,12,29,33–35] identified a total of 22 mutation-associated driver genes, also known as significantly mutated genes (SMGs). Here the significance refers to the functional relevance of the somatic variant, which is measured by bioinformatic methods modeling molecular characteristics of driver and passenger mutations, including rates of silent vs nonsilent mutations, mutation spectrum, gene expression level, and DNA replication time.[36] ESCCs share multiple SMGs exclusively with squamous cell carcinoma of the lung (LUSC)[37] and head and neck squamous cell carcinoma (HNSCC).[38] Some of these regulate squamous cell differentiation, such as *ZNF750* and *NOTCH1*. These findings indicate that genetic aberrations could have similar oncogenic potential in cells of similar origins, with similar gene expression profiles and cell differentiation programs. Therefore, neoplasms that arise from developmentally similar lineages are more molecularly alike than neoplasms that arise from different cell lineages but are located within the same organ. This concept has important clinical implications because tumors with similar molecular features can share biomarkers and therapeutic targets. Design of clinical trials and biomarker development for ESCC may benefit from the studies in LUSC and HNSCC, and vice versa.

Like many other cancer types, ESCCs contain prevalent mutations in *TP53*. Other driver mutations are much less frequent in ESCCs (found in fewer than 20% of samples). Thirteen of the 22 SMGs are consistently identified across different cohorts, and most have been associated with development of other tumor types (Table 2). Of the remaining SMGs, 5 have been validated in functional experiments that confirmed their biological effect, indicating that these genes are drivers of ESCC development.

A total of 10 SMGs have additional genomic or epigenomic aberrations in ESCCs, suggesting that their activation or inactivation are important for development of ESCC. Mutations in *FAM135B,*[8] *EP300,*[33,34] and *TET2*[34] have been associated with shorter times of patient survival. Nevertheless, the discovery of ESCC SMGs, particularly those mutated at low frequency, is far from saturated because of insufficient statistical power. Specifically, studies of ESCCs typically sequenced the exomes of 100 to 150 tumor-germline pairs, whereas mathematical analysis estimated that 1000 to 2000 samples are needed to identify with confidence SMGs mutated in 2% to 3% of the population, taking into account the background mutational rate of ESCC.[36]

Deconvolution of the complex mutational spectrum with mathematical algorithms has increased our understanding of the mutation process in cancer cells. Endogenous (such as spontaneous deamination of 5-methylcytosines) and