





Original Research

Genomic Instability is Widespread in Esophageal Squamous Dysplasia and Increases During the Progression to Cancer

Haiyin An^{1,2}, Xian Cheng¹, Liyan Xue³, Guiqi Wang⁴, Xiuli Zhu⁵, Junyi Li¹,
Ting Xiao^{1,*}, Shujun Cheng^{1,*}¹State Key Laboratory of Molecular Oncology, Department of Etiology and Carcinogenesis, National Cancer Center/National Clinical Research Center for Cancer/Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, 100021 Beijing, China²Department of Medical Oncology, Beijing Chest Hospital, Capital Medical University, Beijing Tuberculosis and Thoracic Tumor Research Institute, 101149 Beijing, China³Department of Pathology, National Cancer Center/National Clinical Research Center for Cancer/Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, 100021 Beijing, China⁴Department of Endoscopy, National Cancer Center/National Clinical Research Center for Cancer/Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, 100021 Beijing, China⁵Geneplus-Beijing Institute, 102206 Beijing, China*Correspondence: xiaot@cicams.ac.cn (Ting Xiao); chengshj@cicams.ac.cn (Shujun Cheng)

Academic Editor: Sung Eun Kim

Submitted: 13 May 2025 Revised: 7 September 2025 Accepted: 18 September 2025 Published: 31 October 2025

Abstract

Background: Research on the molecular progression of esophageal squamous dysplasia to cancer remains limited. The majority of prior studies have focused on morphological precancerous lesions sampled adjacent to tumors, and have relied primarily on the analysis of data from whole-exome sequencing. **Methods:** To investigate the development of esophageal squamous cell carcinoma (ESCC), whole genome analysis was conducted on 13 precancerous tissues and 15 ESCC tissues. Field effects were avoided by using biopsies of squamous dysplasia from patients without concurrent tumor, thereby allowing study of molecular alterations associated with the true precancerous state. **Results:** Our results revealed frequent copy number alterations (CNAs) and structural variants (SVs) in esophageal squamous dysplasia. These changes were also detected in ESCC, indicating that genomic instability markers such as CNAs and SVs occur at an early stage and persist throughout ESCC evolution. The detection of *TP53* mutations and *CASP8* deletions in both pre-malignant lesions and ESCC suggests they may be early driving events during esophageal carcinogenesis. Mutations in *MUC5B* were observed in 7.7% of precancerous lesions and 6.7% of ESCC. Moreover, these mutations were associated with a higher tumor mutational burden (TMB) and an immune “hot” tumor microenvironment. Apolipoprotein B mRNA-editing catalytic polypeptide-like (APOBEC) enzyme-associated mutational signatures were exclusively identified in ESCC and may further exacerbate genomic instability in the more advanced stages of tumorigenesis. Significantly higher ploidy alterations levels were detected in ESCC compared to squamous dysplasia. Moreover, the cohort that underwent local recurrence of dysplasia within two years had significantly elevated ploidy alterations levels compared to those with no long-term recurrence. These results indicate that elevated levels of aneuploidy and genomic instability were associated with tumor progression and local recurrence of dysplasia. **Conclusions:** Mutations in *TP53* and *MUC5B*, as well as deletion of *CASP8*, may be early driver events in carcinogenesis and could precede the emergence of the APOBEC mutation signature. Moreover, ploidy alterations confer a selective advantage to genomically unstable cells, thereby promoting their progression toward malignant transformation. Collectively, our results demonstrate that genomic instability is prevalent in precancerous lesions and intensifies during the late stages of tumor progression. Cells with a certain level of genomic instability appear to possess a competitive advantage for malignant transformation.

Keywords: precancerous conditions; genetic variation; esophageal squamous cell carcinoma

1. Introduction

Esophageal carcinoma ranks as the seventh most prevalent malignancy and the sixth leading cause of cancer-related mortality worldwide, accounting for over 540,000 annual deaths globally [1]. The two major pathological subtypes of esophageal cancer are esophageal squamous cell carcinoma (ESCC), and esophageal adenocarcinoma (EAC). They exhibit marked differences in epidemiology, molecular profiles, and oncogenic pathways [2]. ESCC is the predominant subtype in Asian populations, whereas

EAC is more common in Western countries [3,4]. ESCC is considered to be more analogous to other types of squamous cell carcinoma, whereas EAC exhibits notable similarities to gastric cancer [5]. Despite these differences, both subtypes follow a progressive carcinogenic sequence from precursor lesions to invasive carcinoma. The progression from normal epithelium to precursor lesions, and ultimately to invasive cancer, provides an invaluable model for studying the full trajectory of esophageal tumorigenesis. Uncovering the molecular mechanisms driving the transi-



tion from pre-invasive to invasive disease is critical for improving early detection and therapeutic strategies [6]. Over the past decade, significant progress has been made in understanding the mechanisms underlying the progression of Barrett's esophagus (BE), a metaplastic precursor lesion to EAC. This has resulted in significant advances in surveillance protocols of EAC [7–11]. However, in contrast to EAC, research on esophageal squamous dysplasia and its progression to ESCC remains limited, leaving significant gaps in our understanding of this pathway.

Although several studies have compared the genomic sequencing data of ESCC and adjacent precancerous lesions, many critical questions remain unanswered, particularly regarding molecular differences between the two histopathological stages. Liu Xi *et al.* [12] reported that dysplastic lesions adjacent to tumors exhibit similar alterations to ESCC in terms of both mutations and copy number alterations (CNAs). Chen Xixi *et al.* [13] found that “two-hit” events on *TP53* are rare in dysplasia samples from tumor-free patients, in contrast to dysplastic lesions that are adjacent to tumors. Chang Jiang *et al.* [14] inferred that *TP53* inactivation was one of the earliest steps in the malignant transformation of esophageal epithelium.

Most of the above studies focused on morphological precancerous lesions that were adjacent to tumors, and relied primarily on data from whole-exome sequencing (WES) for their analysis. However, it is important to note that due to field cancerization, dysplastic lesions adjacent to tumors are more likely to represent the final stages of cancer evolution. Consequently, it is not known to what extent these lesions accurately reflect earlier time points in the natural history of neoplastic transformation [7]. For instance, genetic alterations observed in precancerous lesions adjacent to tumors may differ from those found in precancerous lesions in patients who have not progressed to cancer [15]. Furthermore, many insights into neoplastic transformation can only be obtained through whole-genome analysis, and would not be seen with targeted exome sequencing [16]. Currently, there are few reports on the state of structural variants (SVs) in precursors to ESCC.

To address these gaps, we conducted whole-genome analysis on precancerous samples from patients without concurrent cancer, as well as on ESCC samples. By analyzing the genomic landscape of esophageal premalignant lesions, we detected molecular similarities and differences between dysplasia and cancer. Apart from finding some well-known CNAs known to act as drivers, our findings also revealed that SVs occur frequently in esophageal squamous dysplasia, indicating they emerge early in the evolution of ESCC. Moreover, our data showed that the level of ploidy alterations could distinguish esophageal squamous dysplasia from ESCC, and were also associated with the local recurrence of dysplasia. Based on these findings, we propose that elevated levels of aneuploidy occur relatively late in the progression of ESCC and exacerbate genomic

instability, thereby providing cells with a competitive advantage for malignant transformation.

2. Materials and Methods

2.1 Sample Collection

Two cohorts of fresh samples were obtained: pairs of dysplastic and normal samples from 13 patients without ESCC, and pairs of tumor and normal samples from 15 ESCC patients. None of the patients included in the study had received any treatment prior to surgery or biopsy. Sample collection was conducted following approval from the Institutional Review Board of the Cancer Hospital, Chinese Academy of Medical Sciences (Ethic Approval Number: 12-71/605). The study was carried out in accordance with the guidelines of the Declaration of Helsinki. A written consent was signed by the patients or their families/legal guardians.

2.2 Whole-genome Sequencing

Whole-genome sequencing was performed on all samples using the Illumina NOVA sequencing platform. The dataset included 13 dysplastic samples paired with morphologically normal esophageal epithelium, and 15 tumor samples paired with morphologically normal esophageal epithelium. The median sequencing depth achieved was 8.8X.

2.3 Identification of ESCC Driver Genes

To identify ESCC driver genes, we curated a comprehensive gene set by combining 200 pan-cancer and esophageal cancer driver genes reported in the literature [17] with 98 esophageal cancer driver genes downloaded from the IntOGen database [18]. This yielded a total of 253 ESCC-related driver genes for subsequent analysis.

2.4 Whole-genome Sequencing Analysis

Strelka software (version 2.9.10; Illumina, San Diego, CA, USA) was used to detect somatic mutations in the sequencing data. Biological hallmarks enrichment analyses were performed using the Metascape database. The facets tool (version 0.6.0) was used to analyze CNAs. We employed ichorCNA (version 2.0) to calculate ploidy alterations for each sample. Meerkat algorithm was used to detect SVs. GeneFuse [19] (version 0.6.1) was applied to detect target gene fusions by scanning FASTQ files.

2.5 Immune Infiltration-related Analysis and Functional Enrichment Analysis

Expression data, genomic data and associated clinical information for ESCC samples were obtained from the cancer genome atlas (TCGA) database. After retaining samples that included both sequencing data and clinical information, a total of 82 samples were selected for subsequent analysis. The R package “immunedeconv” and EPIC algorithm were employed to ensure accurate assessment of immune scoring. The Limma package in R software (version 4.0.3;

R Foundation for Statistical Computing, Vienna, Austria) was used to study the differentially expressed genes. We defined “ $p < 0.05$ and (fold change) > 1.5 ” as the threshold for screening differential mRNA expression. The Cluster-Profiler package in R was utilized to conduct an analysis of the Gene Ontology (GO) functions of potential genes and to perform enrichment analysis of the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways. Gene Set Enrichment Analysis (GSEA) was performed using SRplot [20].

2.6 Statistical Analysis

Comparisons between the two groups mentioned in the study was performed using the Wilcoxon rank-sum test. A two-sided p -value < 0.05 was considered statistically significant in all analyses.

3. Results

3.1 Clinical Characteristics of Enrolled Patients

This study enrolled 28 patients, consisting of 13 cases of pathologically confirmed and tumor-free esophageal squamous dysplasia, and 15 cases of early-stage ESCC. The clinical characteristics of the two cohorts are summarized in Tables 1,2, respectively. The dysplasia cohort comprised predominantly high-grade intraepithelial neoplasia (84.62%, 11/13), with 23.08% (3/13) exhibiting dual smoking/alcohol exposure. The ESCC cohort consisted of 8 cases with tumor-nodes-metastasis (TNM) stage I disease and 7 with TNM stage II disease. These samples represent relatively early-stage cancer and are closer to precursor lesions in terms of evolution, thereby providing a temporally compressed model to identify driver events that orchestrate the transition from intraepithelial neoplasia to early invasive carcinoma. Moreover, for the purposes of risk stratification, we conducted a prospective follow-up of the pre-cancerous lesion cohort for >7 years. Within two years following endoscopic submucosal dissection (ESD), 3 patients experienced a recurrence of intraepithelial neoplasia in proximity to the treatment site. In contrast, 7 patients demonstrated no recurrence of precursor during a follow-up period of >7 years (Table 1, Fig. 1A).

3.2 Mutation and Mutational Signature

The non-negative matrix factorization (NMF) method was used to identify mutational signatures. This revealed the squamous dysplasia and ESCC groups exhibited a similar mutational context (Fig. 1B,C). Notably, both groups showed a strong enrichment of Signature 5. According to the Catalogue of Somatic Mutations in Cancer (COSMIC) database, Signature 5 exhibits transcriptional strand bias for T $>$ C substitutions in an ApTpN context. Moreover, it occurs frequently in all cancer types and most cancer samples, although its etiology remains unknown. Signature 1 was also present in both groups and is associated with C $>$ T transitions. It correlates with patient age at cancer diagnosis, suggesting it results from an endoge-

Table 1. Baseline characteristics of enrolled patients with esophageal squamous dysplasia.

Clinical characteristic	Number of patients (%)
Gender	
Male	7 (53.85%)
Female	6 (46.15%)
Grade	
Low-grade intraepithelial neoplasia	2 (15.38%)
High-grade intraepithelial neoplasia	11 (84.62%)
Smoking history	
Yes	3 (23.08%)
No	10 (76.92%)
Drinking history	
Yes	3 (23.08%)
No	10 (76.92%)
Post-treatment follow-up	
Relapse of intraepithelial neoplasia	3 (23.08%)
No relapse	7 (53.85%)
Lost to follow-up	3 (23.08%)
Total	13

Due to rounding, percentages may not sum to 100%.

Table 2. Baseline characteristics of enrolled patients with ESCC.

Clinical characteristic	Number of patients (%)
Gender	
Male	11 (73.33%)
Female	4 (26.67%)
T stage	
T1	11 (73.33%)
T2	1 (6.67%)
T3	3 (20.00%)
N stage	
N0	12 (80.00%)
N1	3 (20.00%)
M stage	
M0	15 (100%)
TNM stage	
I	8 (53.33%)
II	7 (46.67%)
Smoking history	
Yes	7 (46.67%)
No	8 (53.33%)
Drinking history	
Yes	7 (46.67%)
No	8 (53.33%)
Total	15

ESCC, esophageal squamous cell carcinoma; TNM, tumor-nodes-metastasis.

nous mutational process initiated by spontaneous deamination of 5-methylcytosine during tumorigenesis. Signature 12 was present in both cohorts, although its etiology

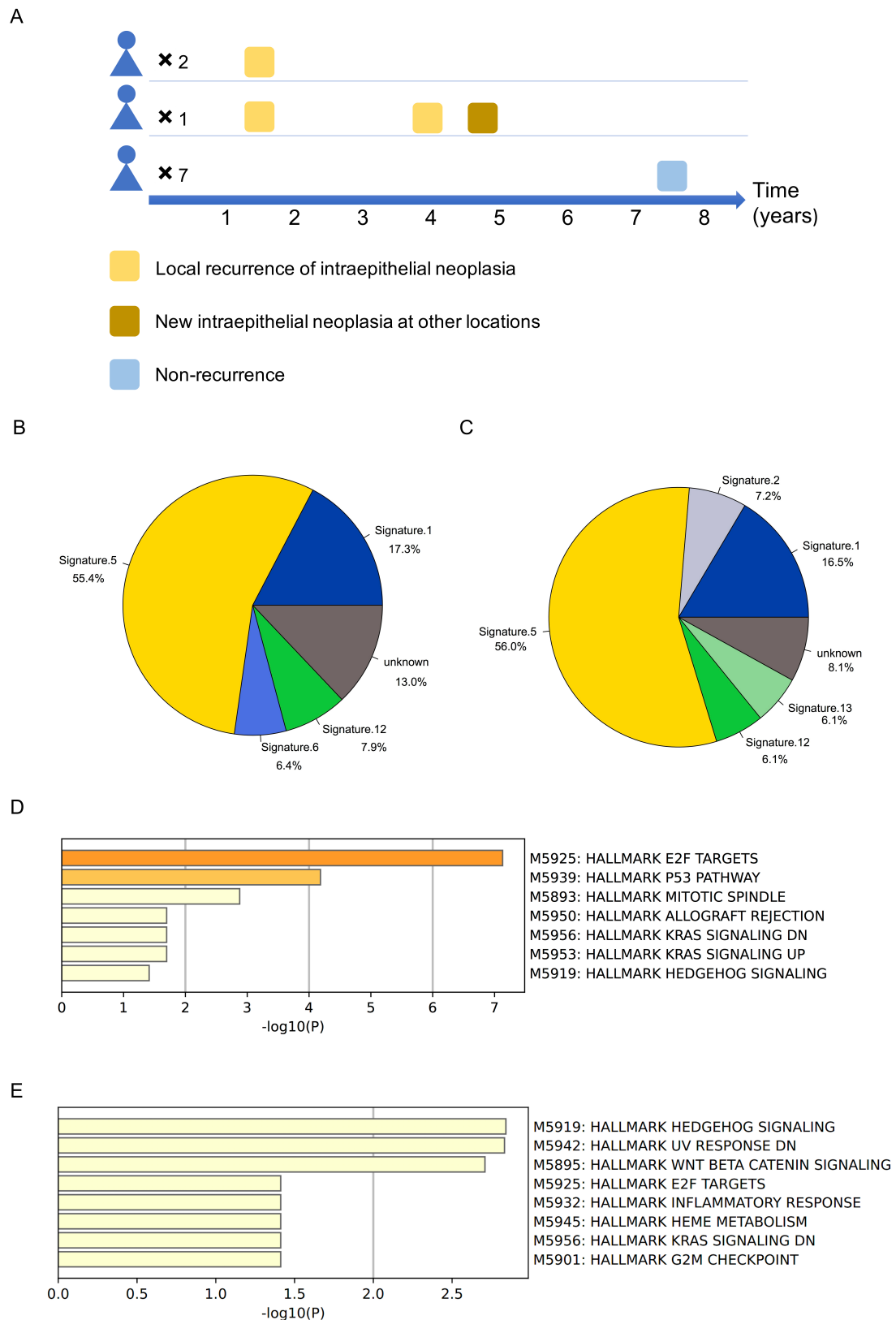


Fig. 1. Follow-up outcomes in the dysplasia group and mutation signatures, hallmarks associated with genes harboring nonsynonymous mutations in the two groups. (A) Follow-up outcomes in the dysplasia group. (B) Mutation signatures in the dysplasia group. (C) Mutation signatures in the ESCC group. (D) Hallmarks associated with genes harboring nonsynonymous mutations in the dysplasia group. (E) Hallmarks associated with genes harboring nonsynonymous mutations in the ESCC group. ESCC, esophageal squamous cell carcinoma.

remains unknown. In contrast, Signatures 2 and 13 were detected exclusively in ESCC samples. These signatures are related to the activity of the Apolipoprotein B mRNA-editing catalytic polypeptide-like (APOBEC) family of cytidine deaminases. The APOBEC mutagenesis pattern is well-documented in ESCC [21] and emerges late in the evolution of primary tumors [22], ultimately contributing to genomic instability [23]. This is consistent with our finding that the APOBEC mutation signature was present in the ESCC group, but not in squamous premalignant tissues in the absence of tumor. Signature 6 was found only in the precursor group, and comments in the COSMIC database suggest it is associated with defective DNA mismatch repair.

We adopted strict screening criteria to ensure the authenticity of the mutation results. Through stringent filtering standards, we identified 41 nonsynonymous single nucleotide variants (SNVs) in the dysplasia group and 53 in the ESCC group. Among these, *TP53* demonstrated the highest mutation frequency, occurring in 54% (7/13) of squamous dysplasia samples and 27% (4/15) of ESCC samples. This indicates that it serves as an early initiating event in the carcinogenesis of ESCC. Four genes (*TP53*, *NOTCH1*, *MUC5B*, and *MLL2*) were found to be mutated in both cohorts, suggesting they represent early events in the progression of ESCC. *MLL2*, an essential histone regulator gene, is frequently altered in ESCC [24]. *MUC5B* encodes structurally related mucin glycoproteins [25] and is known to facilitate cell-to-cell and cell-to-matrix interactions, as well as cell-autonomous signaling, thus promoting tumorigenesis and distant dissemination of tumor cells [26]. However, *MUC5B* is rarely reported in the mutational spectrum of cancer, suggesting it may represent a novel driver gene in ESCC progression. To further explore the role of *MUC5B*, we analyzed sequencing data from ESCC, head and neck squamous cell carcinoma (HNSCC), and lung squamous cell carcinoma (LSCC) obtained from the cBio Cancer Genomics Portal [27,28]. The mutation frequencies for *MUC5B* were 4% in ESCC, 7% in HNSCC, and 13% in LSCC (Fig. 2A and **Supplementary Fig. 1A,B**). Additionally, *MUC5B* mutation was correlated with higher tumor mutational burden (TMB) (nonsynonymous) in ESCC, HNSCC and LSCC (Wilcoxon test; $p = 0.007$, $p = 8.67 \times 10^{-6}$, $p = 4.63 \times 10^{-3}$, respectively, and $q = 0.058$, $q = 2.51 \times 10^{-3}$, $q = 0.106$, respectively) (Fig. 2B and **Supplementary Fig. 1C,D**). The mutation frequencies of recurrently mutated genes were also found to be higher in the *MUC5B*-mutated group than in the *MUC5B* wild-type group (Fig. 2C). Furthermore, in pan-cancer genomic datasets, the mutation frequencies of *MUC5B* range from 22.73% to 0.81%, and samples with *MUC5B* mutations exhibited significantly higher TMB (nonsynonymous) (Wilcoxon test; $p < 10^{-10}$, $q < 10^{-10}$) (Fig. 2D,E).

Since TMB is used as a biomarker for tumor immunotherapy, we conducted an in-depth analysis of

the relationship between *MUC5B* mutations and tumor-infiltrating immune cells in ESCC using the EPIC algorithm and the sequencing data in TCGA. The *MUC5B*-mutated group exhibited significantly higher infiltration levels of CD4+ T cells, CD8+ T cells, and endothelial cells compared to the *MUC5B* wild-type group (Wilcoxon test, $p = 0.026$, $p = 0.022$, $p = 0.027$, respectively) (Fig. 2F). This suggests the *MUC5B*-mutated group is characterized by an immune “hot” tumor microenvironment. To further explore the functional implications of *MUC5B* mutations, we analyzed the expression profile differences between mutant and wild-type groups, finding 252 upregulated and 134 downregulated genes in the *MUC5B*-mutated group. Enrichment analyses showed upregulated genes were mainly linked to phospholipid metabolism, cell adhesion, and platinum drug resistance (**Supplementary Fig. 2A,B**), while downregulated genes were associated with chromatin assembly pathway and rheumatic and autoimmune disease pathway (**Supplementary Fig. 2C,D**). Downregulation of genes enriched in chromatin assembly pathway suggests that core processes governing nucleosome formation and chromatin organization are attenuated in the *MUC5B*-mutant group, resulting in elevated genomic instability. Subsequently, we conducted GSEA analysis to characterize the KEGG pathways affected by the transcriptome change with the *MUC5B* mutation. We found 55 positively enriched pathways and 26 negatively enriched pathways with a q-value of less than 0.05. GSEA analysis further confirmed significant depletion of genome stability related pathways in the *MUC5B*-mutated group, including DNA replication, mismatch repair, homologous recombination, and base excision repair (**Supplementary Fig. 3**).

Next, we analyzed the biological hallmarks of genes harboring nonsynonymous mutations. Those harboring nonsynonymous mutations in the dysplasia group were mainly clustered in the hallmarks for *E2F* targets, *TP53* pathway, mitotic spindle, and *KRAS* and *HEDGEHOG* signaling (Fig. 1D). In addition to the cell cycle, *KRAS*, and *HEDGEHOG* signaling hallmarks observed in the precancerous group, nonsynonymous mutated genes in the ESCC group were also involved in *WNT* and inflammatory response hallmarks (Fig. 1E).

3.3 CNAs in Esophageal Squamous Precursor and Cancer

We next conducted CNA analysis to better decipher the genomic alterations during esophageal carcinogenesis. The median length of genome affected by CNAs in precancerous cohorts was 20.52 MB, including both gain and loss of copy number, while that of the ESCC cohort was 23.45 MB. Although the level of median copy number alteration was higher in ESCC than in esophageal squamous precursor, this was not significantly different (Wilcoxon test, $p = 0.82$) (Fig. 3A). Loss of heterogeneity (LOH) is ubiquitous in cancer, and studies have shown that LOH can contribute to the inactivation of some tumor suppressor genes,

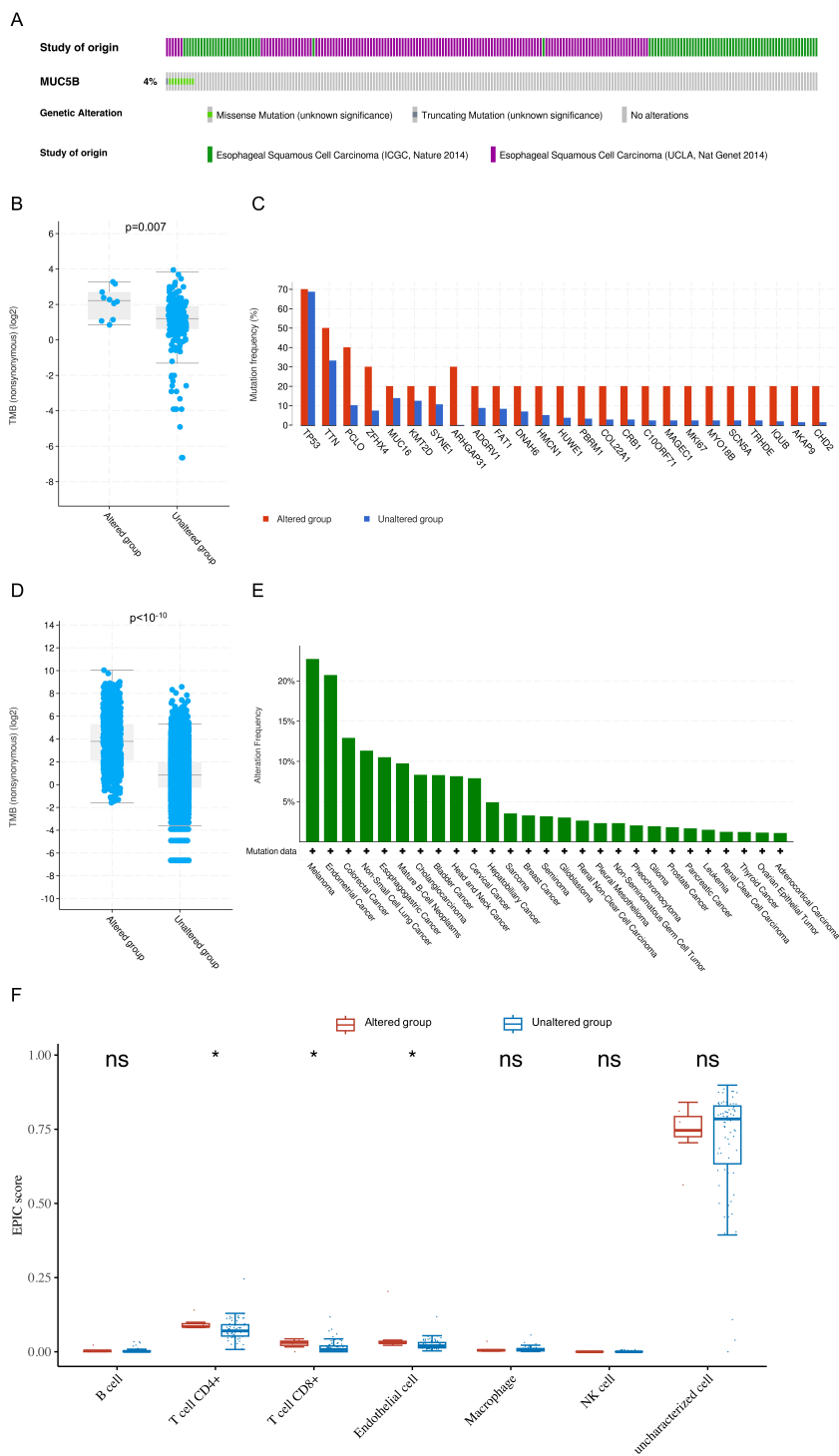


Fig. 2. Mutation frequency of *MUC5B* in ESCC and pan-cancer datasets, the relationship between *MUC5B* mutations and TMB and the relationship between *MUC5B* mutations and tumor-infiltrating immune cells. (A) Mutation frequency of *MUC5B* in ESCC datasets from the cBio Cancer Genomics Portal. (B) Comparison of TMB between *MUC5B*-mutated and *MUC5B* wild-type groups in ESCC datasets from the cBio Cancer Genomics Portal. (C) Comparison of the mutation frequency of recurrently mutated genes between *MUC5B*-mutated and *MUC5B* wild-type cohorts in ESCC datasets from the cBio Cancer Genomics Portal. (D) Comparison of TMB between *MUC5B*-mutated and *MUC5B* wild-type groups in pan-cancer datasets from the cBio Cancer Genomics Portal. (E) Mutation frequency of *MUC5B* in pan-cancer datasets from the cBio Cancer Genomics Portal. (F) Comparison of tumor-infiltrating immune cells between *MUC5B*-mutated and wild-type groups in ESCC datasets from TCGA. * $p < 0.05$. ns means no significant difference. TMB, tumor mutational burden; TCGA, the cancer genome atlas.

thereby affecting the occurrence and progression of cancer. We found that esophageal precancerous lesions and ESCC shared a similar LOH burden (Wilcoxon test, $p = 0.16$) (Fig. 3B), suggesting that LOH already occurs frequently in precancerous lesions.

We also analyzed segmented CNAs that were frequently amplified or deleted. Recurrent CNA profiles demonstrated both overlapping and divergent patterns between premalignant and malignant lesions (Fig. 3C). Focal CNA analysis identified 8q amplification as a shared event in both premalignant lesions and malignant tissues. This segment, reported as a hotspot in ESCC, contains the cancer-related gene *MYC*, which plays a critical role in regulating cell proliferation. These findings suggest that amplification of the 8q segment may be a key driving event in carcinogenesis. In addition, we identified several alterations in chromosomal segments in esophageal squamous precancerous lesions that have been rarely reported in previous studies. For instance, a 7q segment containing the driver gene *BRAF* was found to be amplified in precancerous lesions. *BRAF* is involved in regulation of the *MAPK/ERK* signaling pathway, which affects cell division, differentiation and secretion. Moreover, we identified frequent deletions that affect segments in 4p, 9q (containing the driver gene *NOTCH1*), and 17q. The deletion in 4p has been frequently reported in ESCC [29], while LOH of 17q has also been identified in ESCC [30].

3.4 SV Landscape in the Dysplasia and ESCC Cohorts

SVs play a mysterious role in cancer through their ability to move blocks of adjacent genes simultaneously, or to generate gene fusion, thereby contributing to concurrent oncogenic events [31,32]. However, the landscape of SVs in esophageal squamous dysplasia remains unclear. To address this, we conducted SV analysis through the Meerkat computational tool [33,34]. A median of 6300 SVs per genome (range 3684–8308) were identified in the dysplasia cohort, and 3616 (range 2751–5184) in the ESCC cohort, with no statistically significant difference between the two (Wilcoxon test, $p = 0.41$). Five SV types were detected, consisting of duplications/intra-chromosomal translocations, deletions/intra-chromosomal translocations, inversions and two types of inter-chromosomal translocations (type 1 and type 2) [34] (Fig. 4A). Genome-wide SV analysis revealed that inter-chromosomal translocation was the predominant SV type in both precancerous and cancerous tissues.

To further study the potential impact of the large number of SVs on biological processes, these were mapped at the gene level and the frequently rearranged driver genes in esophageal cancer were analyzed. No significant difference in the frequency of SVs that disrupt ESCC driver genes was observed between the dysplasia and tumor cohorts (Wilcoxon test, $p = 0.46$) (Fig. 4B).

We next analyzed and characterized the spectrum of driver genes with a high frequency of SVs in both groups (Fig. 4C). *CASP8*, a recently identified novel driver gene associated with ESCC [35], exhibited the highest frequency of SVs in both the dysplasia and tumor groups, predominantly in the form of deletions. While deletions in *CASP8* have been reported in ESCC [36], the involvement of *CASP8* in the precursor stage of esophageal lesions has not been previously documented. *CASP8*, a key initiator of death receptor-mediated apoptosis, also plays a critical role in suppressing *RIPK3-MLKL*-dependent necroptosis [37,38]. This gene is essential for preventing tissue damage during both embryonic development and adulthood [37], and its deletion can lead to ileocolitis and gut barrier dysfunction [39]. *KANSL1* was also frequently rearranged in both cohorts. Amplifications and rearrangements of *KANSL1* were previously reported in ovarian cancer, and it could serve as a potential biomarker and therapeutic target for immune response modulation [40]. *KANSL1* was also found to be enriched in tumors with genome doubling events [41]. *ESR1* is frequently affected by deletion-type SVs in both precancerous lesions and tumor samples. High frequencies of *ESR1* SVs were reported in inflammatory reflux esophagitis, Barrett's metaplasia, dysplasia, and EAC [42].

SVs have been shown to mediate oncogenic gene fusion events. The present analysis identified 169 fusion genes that were shared between premalignant and malignant esophageal lesions. Among these, the kinase-associated *SND1-BRAF* fusion, previously reported in thyroid carcinomas [43], was detected in both premalignant (1/13, 7.7%) and malignant (1/15, 6.7%) samples.

3.5 Ploidy Alterations

Ploidy alterations were also compared between the dysplasia and cancer groups. The average ploidy was found to be 2.025 in precursor lesions and 2.31 in the ESCC group (Fig. 5A), demonstrating a significantly higher degree of genomic instability in the cancer group (Wilcoxon test, $p = 0.029$). Importantly, the cohort that underwent local recurrence of dysplasia within two years had significantly elevated level of ploidy alterations compared to those with no long-term recurrence (Wilcoxon test, $p = 0.033$). This finding suggests that higher level of ploidy alterations may serve as a predictive marker for an increased risk of recurrence (Fig. 5B).

4. Discussion

The current study comprehensively analyzed the whole genomic landscape of esophageal squamous precancerous lesions and ESCC. Field effects, which encompass tumor-host interactions and gene-environment interactions [44], can remodel the neighboring microenvironment of tumor cells. Tumorous dysplasia exhibits a more malignant molecular phenotype compared to non-tumorous dysplasia, with a similar pattern and burden of genomic alteration to

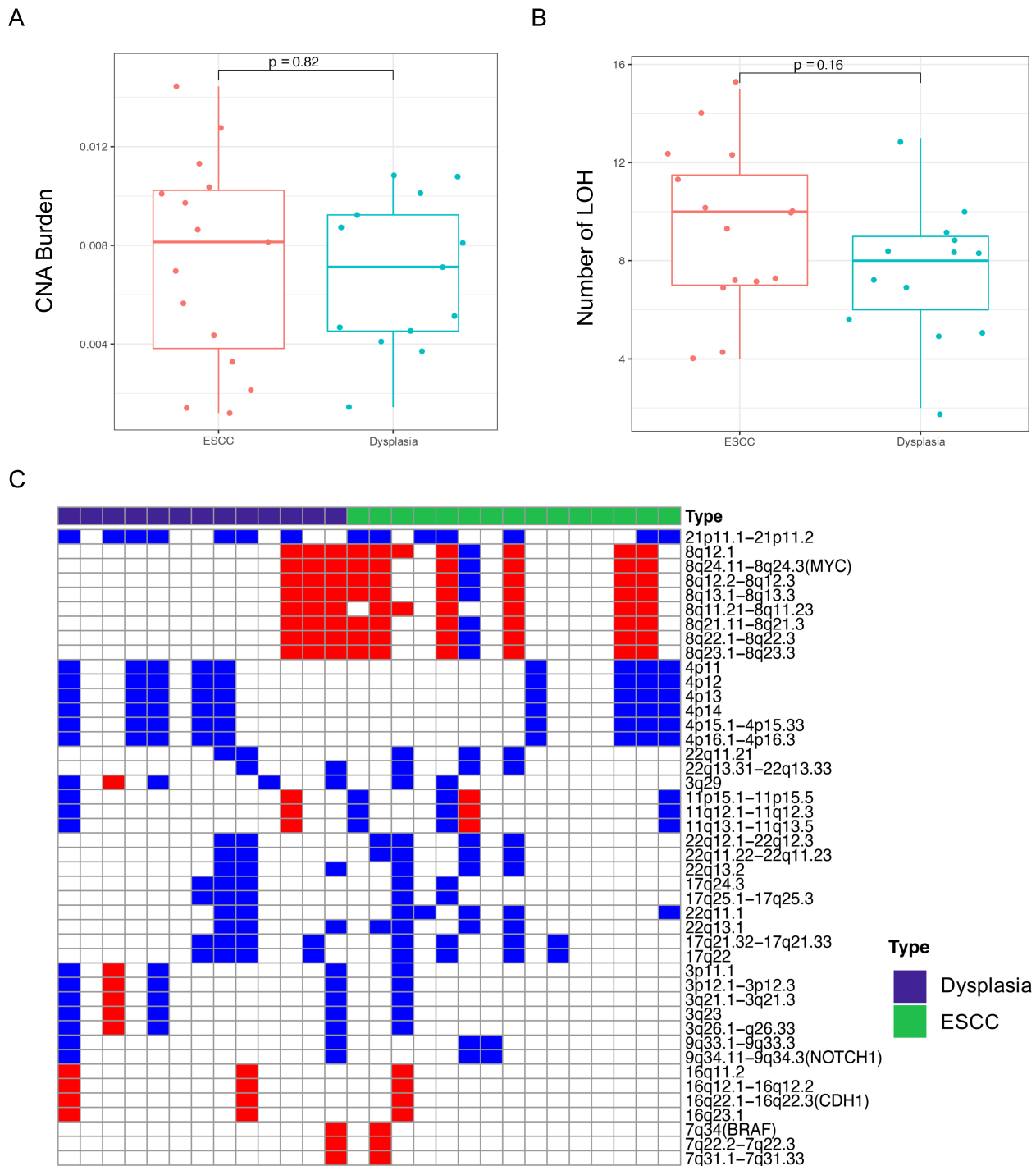


Fig. 3. CNA profiling across the dysplasia and ESCC groups. (A) Comparison of the CNA burden between ESCC and dysplasia samples. (B) Comparison of the number of LOH events between dysplasia and ESCC samples. (C) Frequently amplified and deleted segmented CNAs in the dysplasia and ESCC groups. Red indicates copy number amplification, and blue indicates copy number deletion. CNA, copy number alteration; LOH, loss of heterogeneity.

that of tumor tissue [12]. Such tumor-adjacent precancerous lesions may already have acquired some transformative potential [13]. Once the cancerous tissue is removed and suppression of the dominant ESCC clone is lifted, the lesion may rapidly progress into carcinoma. However, among precursor cases from patients with no sign of ESCC, 33%

progressed or recurred, while others regressed or remained static [45]. To reduce the effects of field cancerization and to conduct risk stratification, we analyzed biopsies of squamous dysplasia obtained from patients without concurrent tumor, as well as performing long-term follow-up of these patients.

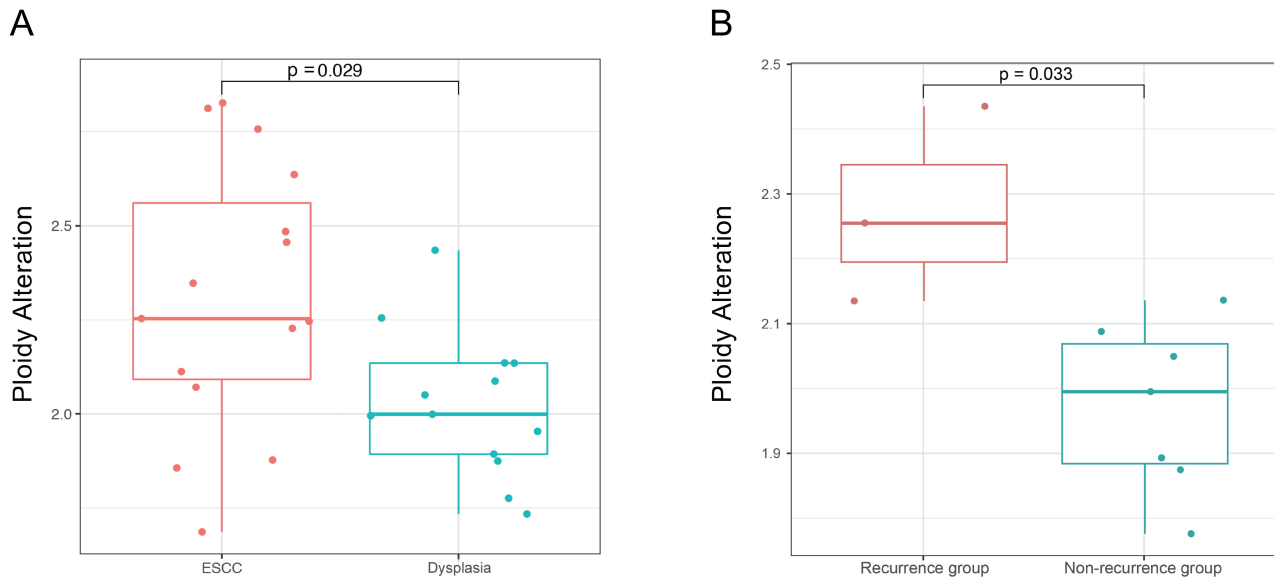


Fig. 5. Comparison of ploidy alterations between dysplasia and ESCC samples and comparison of ploidy alterations between recurrence group and non-recurrence group. (A) Comparison of ploidy alterations between dysplasia and ESCC samples. (B) Comparison of ploidy alterations between recurrence group and non-recurrence group in the dysplasia cohort.

Our analysis revealed that precursor lesions without tumors harbor frequent CNAs and SVs, which differs from some previous reports on esophageal carcinogenesis. Our use of whole-genome sequencing generates a more comprehensive genomic profile compared to targeted or exome sequencing approaches, potentially explaining the discrepancy with earlier studies. The inherent heterogeneity of dysplasia may also contribute to the differences observed between studies. Despite the comparable burden of CNAs and SVs between precancerous lesions and ESCC, the overall genomic alteration burden was generally higher in ESCC. This result suggests that chromosomal instability, a hallmark of cancer, progressively accumulates during carcinogenesis. Similar to our finding, a large burden of canonical CNAs were also detected in the hyperplasia/metaplasia stages of the LSCC progression model, with their frequency also increasing during the progression of lesions [46].

Our results showed a high mutation frequency of *TP53* in precancerous lesions. Chang Jiang *et al.* [14] reported that the frequency of *TP53* biallelic inactivation increases dramatically in early precancerous lesion stage and *TP53* inactivation leads to CNAs in *TP53* biallelic loss clones. These results suggested that *TP53* may be an early initiator of carcinogenesis. Additionally, we found the *MUC5B* gene was mutated in 7.7% of precancerous lesions and 6.7% of ESCC. This gene is implicated in cell-to-cell and cell-to-matrix interactions. Interestingly, mutations in *MUC5B* have been associated with increased TMB in ESCC, HNSCC, LSCC, and pan-cancer datasets, suggesting they may represent a novel early event in carcinogenesis and a biomarker of genomic instability. More-

over, our analysis revealed that samples with *MUC5B* mutations were associated with immunologically “hot” tumors, indicating a potentially heightened responsiveness to immunotherapy. Functional enrichment analysis of differentially expressed genes indicates that the *MUC5B*-mutant group is more prone to invasive metastasis and platinum-based chemoresistance and exhibits increased genomic instability owing to attenuated chromatin assembly. Due to the deficiency in homologous recombination function in the *MUC5B*-mutant group, these patients may derive therapeutic benefit from polyadenosine-diphosphate-ribose polymerase (PARP) inhibitors. Meanwhile, DNA-repair defects in the mismatch repair pathway within the *MUC5B*-mutant cohort are also biomarkers for guiding the use of immune-checkpoint inhibitors [47,48]. Concurrently, individuals harboring *MUC5B* mutations may be more sensitive to novel treatment agents that target genomic instability, including *KIF18A* inhibitors, *p53*-reactivating agents and *PLK4* inhibitors, all of which are presently being tested in clinical trials. Furthermore, the APOBEC mutagenesis signature, a known contributor to genomic instability, was detected exclusively in ESCC but was absent in precancerous lesions. This observation indicates the APOBEC mutagenesis pattern may represent a relatively late event in the carcinogenic process, potentially conferring invasive capabilities to neoplastic cells.

The recent advances in genomic research have focused increasing attention towards the study of genomic alterations in morphologically normal cells. Li Ruoyan *et al.* [49] reported that normal esophageal tissues already harbor CNAs. These are predominantly characterized by whole-chromosomal amplifications of chromosomes 3, 5, and 7.

The present study also detected amplification of 7q in both dysplasia and ESCC, suggesting that gain of 7q may represent an early genomic event and play a role in the accumulation of genomic instability during the progression to ESCC.

SVs are well-known hallmarks of cancer, affecting a larger fraction of the genome than point mutations by increasing or deleting gene copy numbers. These frequent genomic rearrangements can result in the amplification of oncogenes or the disruption of tumor suppressor genes, thus contributing to tumorigenesis. In our study, *CASP8*, a recently identified driver gene, exhibited frequent deletions in premalignant lesions, indicating that SVs of this gene may be early driver events in the development of ESCC. In HNSCC, the *CASP8* inactivating mutations have been reported to confer cells the ability of resistance to death receptor [50]. However, recent studies have reported that the function of *CASP8* in cancer represents a complex double-edged sword, and its specific role in ESCC remains to be further elucidated [51,52]. While most genome-scale rearrangements are thought to be neutral, some give rise to oncogenic gene fusions [32]. Gene fusions are potential targets for personalized therapeutics, and may also serve as potential prognostic and diagnostic biomarkers [53]. The current study presented the gene fusion landscape of esophageal squamous precursors and carcinoma. However, further studies are needed to determine whether these gene fusions are drivers or passengers.

Our study found that a high level of ploidy alterations can effectively distinguish between the two histologically different cohorts, suggesting that it may be an adaptive mechanism enabling cell survival under selective pressure. Aneuploidy is closely associated with increased genomic instability [54], thus aligning with the observation that ESCC is characterized by high levels of genomic instability [36]. Similarly, the transition from carcinoma *in situ* to LSCC is associated with a high level of chromosomal instability, which was identified as an early marker in this progression [55,56]. Previous research comparing the genomic landscapes of early-stage and late-stage ovarian high-grade serous carcinoma (HGSC) also observed a significant increase in overall ploidy in late-stage tumors compared to their early-stage counterparts [57]. Moreover, genomic instability and aneuploidy may facilitate metastatic progression by inducing epithelial-to-mesenchymal transition (EMT), thereby promoting tumor invasion [58]. High levels of aneuploidy and genomic instability also correlate with the expression of immune evasion markers [59] and with poor prognosis [58,60]. Our results suggest that a high level of ploidy confers cells with the ability to outcompete the original diploid dysplastic cells. This may result in phenotypic changes, such as the ability to infiltrate the basement membrane, which can be recognized as a distinct pathological change by histological examination.

Long-term follow up in our study revealed that patients with precancerous lesions characterized by elevated ploidy alterations displayed a greater risk of recurrence following ESD treatment. This result provides further evidence that intraepithelial neoplasia characterized by a high level of ploidy alterations has an increased risk for malignancy. Furthermore, the identification of individuals with high-risk dysplasia may facilitate more targeted and effective endoscopic surveillance and monitoring strategies. In summary, ploidy alterations appear to be closely associated with malignant transformation and with local recurrence of esophageal dysplasia.

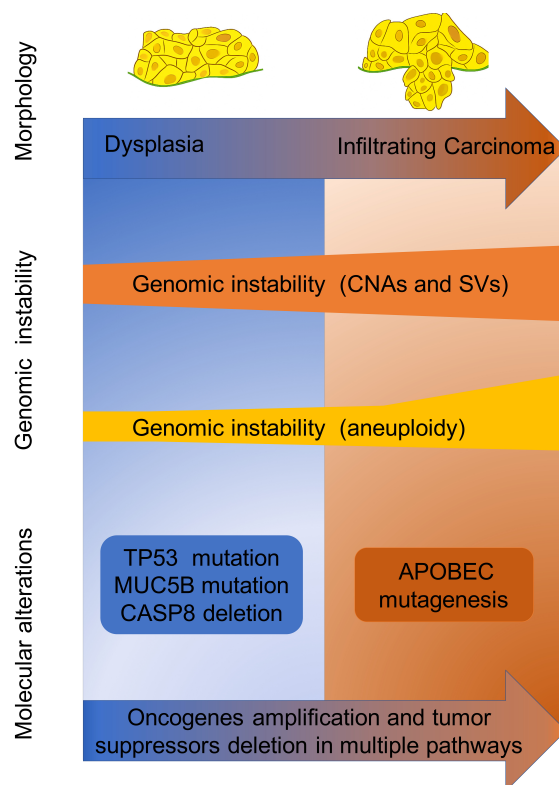


Fig. 6. Preliminary model of the evolution from dysplasia to ESCC. The schematic diagram of esophageal precancerous lesion and cancerous tissue in Fig. 6 was generated by ChatGpt (version 5.0; OpenAI, San Francisco, CA, USA).

Our comprehensive molecular profiling revealed distinct temporal phases in ESCC tumorigenesis. It also identified critical early-stage alterations that initiate malignant transformation, followed by subsequent transitional genomic events that are associated with tumor progression and phenotypic evolution. However, the findings of our study are subject to several limitations. First, due to the very small tissue volume of dysplasia samples, direct validations using these specimens were not feasible in the present study. And due to the limited samples number, we may not be able to capture a broader spectrum of genomic variations. We therefore plan to accrue a prospec-

tive clinical cohort of premalignant samples to replicate our findings. Second, although the current study includes long-term clinical follow-up validation and confirmation through genomic and transcriptomic analyses in independent cohorts, it is limited by the absence of molecular experiments. The functional effects of *MUC5B* mutations and *CASP8* deletions still require mechanistic molecular experiments in further studies. Third, our next-generation sequencing analysis was restricted to the genomic and transcriptomic level, without the incorporation of epigenetic data. To overcome these constraints, subsequent investigations should employ multi-omics strategies that combine whole-genome, transcriptome, epigenome, and proteome analyses across a well-curated sample cohort.

5. Conclusions

In summary, this study comprehensively characterized the genomic alterations in ESCC and precancerous lesions. The cancer hallmark of genomic instability was prevalent even in the precancerous stages, and increased further during the transition from precursor to invasive cancer (Fig. 6). Mutations in *TP53* and *MUC5B*, as well as the deletion of *CASP8*, may represent early driver events in carcinogenesis that precede emergence of the APOBEC mutagenesis pattern. In addition, ploidy alterations confer a selective advantage to genomically unstable cells, promoting their progression toward malignant transformation and local recurrence. These findings provide new insights into the genomic evolution of ESCC and highlight potential avenues for early detection and targeted intervention in esophageal carcinogenesis.

Availability of Data and Materials

The datasets used and analyzed during the current study are available from the corresponding author on reasonable request.

Author Contributions

HA performed the advanced bioinformatics analysis and wrote the manuscript. XC helped to verify the analysis results. GW participated in the experimental design and provided the dysplasia specimens used in this research. LX provided ESCC specimens for the research and conducted pathological diagnosis of all the samples. XZ performed basic bioinformatics analysis of data. JL assisted with literature retrieval and proofread the manuscript. SC took the lead in designing the study and overseeing the experimental framework. TX played a critical role in data analysis and drafting the manuscript. All authors contributed to editorial changes in the manuscript. All authors read and approved the final manuscript. All authors have participated sufficiently in the work and agreed to be accountable for all aspects of the work.

Ethics Approval and Consent to Participate

This study was performed in accordance with the Declaration of Helsinki and was approved by the Institutional Review Board of Cancer Hospital, Chinese Academy of Medical Sciences (Ethical Approval Number: 12-71/605). A written consent was signed by the patients or their families/legal guardians.

Acknowledgment

We are grateful to Professor Kaitai Zhang and Lin Feng from the Cancer Hospital, Chinese Academy of Medical Sciences for their guidance on the analytical approach. We are grateful to Haonan Gu for his assistance in collecting the clinical follow-up data. We sincerely thank all the researchers and study participants for their contributions.

Funding

This work was supported by the National Natural Science Foundation of China (No. 82372718), the National Key R&D Program of China (No. 2023YFC3503205), and the Chinese Academy of Medical Sciences Innovation Fund for Medical Sciences (No. 2023-I2M-2-004).

Conflict of Interest

The authors declare no conflict of interest. Xiuli Zhu is an employee of Geneplus-Beijing Institute, the judgments in data interpretation and writing were not influenced by this relationship.

Declaration of AI and AI-Assisted Technologies in the Writing Process

During the preparation of this work we used ChatGpt (version 5.0; OpenAI, San Francisco, CA, USA) to generate a schematic diagram of esophageal precancerous lesion and cancerous tissue in Fig. 6. After using this tool, we reviewed and edited the content as needed and takes full responsibility for the content of the publication.

Supplementary Material

Supplementary material associated with this article can be found, in the online version, at <https://doi.org/10.31083/FBL41107>.

References

- [1] Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, *et al.* Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA: A Cancer Journal for Clinicians*. 2021; 71: 209–249. <https://doi.org/10.3322/caac.21660>.
- [2] Siewert JR, Ott K. Are squamous and adenocarcinomas of the esophagus the same disease? *Seminars in Radiation Oncology*. 2007; 17: 38–44. <https://doi.org/10.1016/j.semradonc.2006.09.007>.
- [3] Torre LA, Bray F, Siegel RL, Ferlay J, Lortet-Tieulent J, Jemal

- A. Global cancer statistics, 2012. CA: A Cancer Journal for Clinicians. 2015; 65: 87–108. <https://doi.org/10.3322/caac.21262>.
- [4] Zhong L, Li H, Chang W, Ao Y, Wen Z, Chen Y. TP53 Mutations in Esophageal Squamous Cell Carcinoma. *Frontiers in Bioscience (Landmark Edition)*. 2023; 28: 219. <https://doi.org/10.31083/j.fbl2809219>.
- [5] Cancer Genome Atlas Research Network, Analysis Working Group: Asan University, BC Cancer Agency, Brigham and Women's Hospital, Broad Institute, Brown University, *et al*. Integrated genomic characterization of oesophageal carcinoma. *Nature*. 2017; 541: 169–175. <https://doi.org/10.1038/nature20805>.
- [6] Pashayan N, Pharoah PDP. The challenge of early detection in cancer. *Science (New York, N.Y.)*. 2020; 368: 589–590. <https://doi.org/10.1126/science.aaz2078>.
- [7] Katz-Sumercorn AC, Jammula S, Frangou A, Peneva I, O'Donovan M, Tripathi M, *et al*. Multi-omic cross-sectional cohort study of pre-malignant Barrett's esophagus reveals early structural variation and retrotransposon activity. *Nature Communications*. 2022; 13: 1407. <https://doi.org/10.1038/s41467-022-28237-4>.
- [8] Stachler MD, Taylor-Weiner A, Peng S, McKenna A, Agoston AT, Odze RD, *et al*. Paired exome analysis of Barrett's esophagus and adenocarcinoma. *Nature Genetics*. 2015; 47: 1047–1055. <https://doi.org/10.1038/ng.3343>.
- [9] Killcoyne S, Fitzgerald RC. Evolution and progression of Barrett's oesophagus to oesophageal cancer. *Nature Reviews. Cancer*. 2021; 21: 731–741. <https://doi.org/10.1038/s41568-021-00400-x>.
- [10] Iyer PG, Chak A. Surveillance in Barrett's Esophagus: Challenges, Progress, and Possibilities. *Gastroenterology*. 2023; 164: 707–718. <https://doi.org/10.1053/j.gastro.2023.01.031>.
- [11] Paulson TG, Galipeau PC, Oman KM, Sanchez CA, Kuhner MK, Smith LP, *et al*. Somatic whole genome dynamics of precancer in Barrett's esophagus reveals features associated with disease progression. *Nature Communications*. 2022; 13: 2300. <https://doi.org/10.1038/s41467-022-29767-7>.
- [12] Liu X, Zhang M, Ying S, Zhang C, Lin R, Zheng J, *et al*. Genetic Alterations in Esophageal Tissues From Squamous Dysplasia to Carcinoma. *Gastroenterology*. 2017; 153: 166–177. <https://doi.org/10.1053/j.gastro.2017.03.033>.
- [13] Chen XX, Zhong Q, Liu Y, Yan SM, Chen ZH, Jin SZ, *et al*. Genomic comparison of esophageal squamous cell carcinoma and its precursor lesions by multi-region whole-exome sequencing. *Nature Communications*. 2017; 8: 524. <https://doi.org/10.1038/s41467-017-00650-0>.
- [14] Chang J, Zhao X, Wang Y, Liu T, Zhong C, Lao Y, *et al*. Genomic alterations driving precancerous to cancerous lesions in esophageal cancer development. *Cancer Cell*. 2023; 41: 2038–2050.e5. <https://doi.org/10.1016/j.ccell.2023.11.003>.
- [15] Spira A, Yurgelun MB, Alexandrov L, Rao A, Bejar R, Polyak K, *et al*. Precancer Atlas to Drive Precision Prevention Trials. *Cancer Research*. 2017; 77: 1510–1541. <https://doi.org/10.1158/0008-5472.CAN-16-2346>.
- [16] ICGC/TCGA Pan-Cancer Analysis of Whole Genomes Consortium. Pan-cancer analysis of whole genomes. *Nature*. 2020; 578: 82–93. <https://doi.org/10.1038/s41586-020-1969-6>.
- [17] Bailey MH, Tokheim C, Porta-Pardo E, Sengupta S, Bertrand D, Weerasinghe A, *et al*. Comprehensive Characterization of Cancer Driver Genes and Mutations. *Cell*. 2018; 173: 371–385.e18. <https://doi.org/10.1016/j.cell.2018.02.060>.
- [18] Perez-Llamas C, Gundem G, Lopez-Bigas N. Integrative cancer genomics (IntOGen) in Biomart. *Database: the Journal of Biological Databases and Curation*. 2011; 2011: bar039. <https://doi.org/10.1093/database/bar039>.
- [19] Chen S, Liu M, Huang T, Liao W, Xu M, Gu J. GeneFuse: detection and visualization of target gene fusions from DNA sequencing data. *International Journal of Biological Sciences*. 2018; 14: 843–848. <https://doi.org/10.7150/ijbs.24626>.
- [20] Tang D, Chen M, Huang X, Zhang G, Zeng L, Zhang G, *et al*. SRplot: A free online platform for data visualization and graphing. *PLoS One*. 2023; 18: e0294236. <https://doi.org/10.1371/journal.pone.0294236>.
- [21] Moody S, Senkin S, Islam SMA, Wang J, Nasrollahzadeh D, Cortez Cardoso Penha R, *et al*. Mutational signatures in esophageal squamous cell carcinoma from eight countries with varying incidence. *Nature Genetics*. 2021; 53: 1553–1563. <https://doi.org/10.1038/s41588-021-00928-6>.
- [22] de Bruin EC, McGranahan N, Mitter R, Salm M, Wedge DC, Yates L, *et al*. Spatial and temporal diversity in genomic instability processes defines lung cancer evolution. *Science (New York, N.Y.)*. 2014; 346: 251–256. <https://doi.org/10.1126/science.1253462>.
- [23] Zou J, Wang C, Ma X, Wang E, Peng G. APOBEC3B, a molecular driver of mutagenesis in human cancers. *Cell & Bioscience*. 2017; 7: 29. <https://doi.org/10.1186/s13578-017-0156-4>.
- [24] Song Y, Li L, Ou Y, Gao Z, Li E, Li X, *et al*. Identification of genomic alterations in oesophageal squamous cell cancer. *Nature*. 2014; 509: 91–95. <https://doi.org/10.1038/nature13176>.
- [25] Roy MG, Livraghi-Butrico A, Fletcher AA, McElwee MM, Evans SE, Boerner RM, *et al*. Muc5b is required for airway defence. *Nature*. 2014; 505: 412–416. <https://doi.org/10.1038/nature12807>.
- [26] Ganguly K, Rauth S, Marimuthu S, Kumar S, Batra SK. Unravelling mucin domains in cancer and metastasis: when protectors become predators. *Cancer Metastasis Reviews*. 2020; 39: 647–659. <https://doi.org/10.1007/s10555-020-09896-5>.
- [27] Cerami E, Gao J, Dogrusoz U, Gross BE, Sumer SO, Aksoy BA, *et al*. The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. *Cancer Discovery*. 2012; 2: 401–404. <https://doi.org/10.1158/2159-8290.CD-12-0095>.
- [28] Gao J, Aksoy BA, Dogrusoz U, Dresdner G, Gross B, Sumer SO, *et al*. Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. *Science Signaling*. 2013; 6: p11. <https://doi.org/10.1126/scisignal.2004088>.
- [29] Cheng C, Cui H, Zhang L, Jia Z, Song B, Wang F, *et al*. Genomic analyses reveal FAM84B and the NOTCH pathway are associated with the progression of esophageal squamous cell carcinoma. *GigaScience*. 2016; 5: 1. <https://doi.org/10.1186/s13742-015-0107-0>.
- [30] Hu N, Wang C, Ng D, Clifford R, Yang HH, Tang ZZ, *et al*. Genomic characterization of esophageal squamous cell carcinoma from a high-risk population in China. *Cancer Research*. 2009; 69: 5908–5917. <https://doi.org/10.1158/0008-5472.CA.N-08-4622>.
- [31] Inaki K, Liu ET. Structural mutations in cancer: mechanistic and functional insights. *Trends in Genetics: TIG*. 2012; 28: 550–559. <https://doi.org/10.1016/j.tig.2012.07.002>.
- [32] Tubio JMC. Somatic structural variation and cancer. *Briefings in Functional Genomics*. 2015; 14: 339–351. <https://doi.org/10.1093/bfpg/elv016>.
- [33] Cheng C, Zhou Y, Li H, Xiong T, Li S, Bi Y, *et al*. Whole-Genome Sequencing Reveals Diverse Models of Structural Variations in Esophageal Squamous Cell Carcinoma. *American Journal of Human Genetics*. 2016; 98: 256–274. <https://doi.org/10.1016/j.ajhg.2015.12.013>.
- [34] Yang L, Luquette LJ, Gehlenborg N, Xi R, Haseley PS, Hsieh CH, *et al*. Diverse mechanisms of somatic structural variations in human cancer genomes. *Cell*. 2013; 153: 919–929. <https://doi.org/10.1016/j.cell.2013.04.010>.
- [35] Cui Y, Chen H, Xi R, Cui H, Zhao Y, Xu E, *et al*. Whole-genome

- sequencing of 508 patients identifies key molecular features associated with poor prognosis in esophageal squamous cell carcinoma. *Cell Research*. 2020; 30: 902–913. <https://doi.org/10.1038/s41422-020-0333-6>.
- [36] Cui H, Zhou Y, Wang F, Cheng C, Zhang W, Sun R, *et al.* Characterization of somatic structural variations in 528 Chinese individuals with Esophageal squamous cell carcinoma. *Nature Communications*. 2022; 13: 6296. <https://doi.org/10.1038/s41467-022-33994-3>.
- [37] Fritsch M, Günther SD, Schwarzer R, Albert MC, Schorn F, Werthenbach JP, *et al.* Caspase-8 is the molecular switch for apoptosis, necroptosis and pyroptosis. *Nature*. 2019; 575: 683–687. <https://doi.org/10.1038/s41586-019-1770-6>.
- [38] Li X, Li F, Zhang X, Zhang H, Zhao Q, Li M, *et al.* Caspase-8 auto-cleavage regulates programmed cell death and collaborates with RIPK3/MLKL to prevent lymphopenia. *Cell Death and Differentiation*. 2022; 29: 1500–1512. <https://doi.org/10.1038/s41418-022-00938-9>.
- [39] Kaden-Volynets V, Günther C, Zimmermann J, Beisner J, Becker C, Bischoff SC. Deletion of the Casp8 gene in mice results in ileocolitis, gut barrier dysfunction, and malassimilation, which can be partially attenuated by inulin or sodium butyrate. *American Journal of Physiology. Gastrointestinal and Liver Physiology*. 2019; 317: G493–G507. <https://doi.org/10.1152/ajpgi.00297.2018>.
- [40] Fejzo MS, Chen HW, Anderson L, McDermott MS, Karlan B, Konecny GE, *et al.* Analysis in epithelial ovarian cancer identifies KANSL1 as a biomarker and target gene for immune response and HDAC inhibition. *Gynecologic Oncology*. 2021; 160: 539–546. <https://doi.org/10.1016/j.ygyno.2020.11.008>.
- [41] Quinton RJ, DiDomizio A, Vittoria MA, Kotýnková K, Ticas CJ, Patel S, *et al.* Whole-genome doubling confers unique genetic vulnerabilities on tumour cells. *Nature*. 2021; 590: 492–497. <https://doi.org/10.1038/s41586-020-03133-3>.
- [42] Eads CA, Lord RV, Kurumboor SK, Wickramasinghe K, Skinner ML, Long TI, *et al.* Fields of aberrant CpG island hypermethylation in Barrett’s esophagus and associated adenocarcinoma. *Cancer Research*. 2000; 60: 5021–5026.
- [43] Chu YH, Wirth LJ, Farahani AA, Nosé V, Faquin WC, Dias-Santagata D, *et al.* Clinicopathologic features of kinase fusion-related thyroid carcinomas: an integrative analysis with molecular characterization. *Modern Pathology: an Official Journal of the United States and Canadian Academy of Pathology, Inc.* 2020; 33: 2458–2472. <https://doi.org/10.1038/s41379-020-0638-5>.
- [44] Lochhead P, Chan AT, Nishihara R, Fuchs CS, Beck AH, Giovannucci E, *et al.* Etiologic field effect: reappraisal of the field effect concept in cancer predisposition and progression. *Modern Pathology: an Official Journal of the United States and Canadian Academy of Pathology, Inc.* 2015; 28: 14–29. <https://doi.org/10.1038/modpathol.2014.81>.
- [45] Wen D, Zhang L, Wang X, Li Y, Ma C, Liu X, *et al.* A 5.5-year surveillance of esophageal and gastric cardia precursors after a population-based screening in China. *Journal of Gastroenterology and Hepatology*. 2015; 30: 1720–1725. <https://doi.org/10.1111/jgh.13040>.
- [46] Luo B, Luo Y, Chen S, Yang T, Liu B, Liao X, *et al.* Real-Time Evolutionary Landscape of the Bronchial Epithelium and Corresponding Dynamic Immune Cell Alterations in Lung Squamous Cell Carcinogenesis. *Advanced Science (Weinheim, Baden-Württemberg, Germany)*. 2025; 12: e13256. <https://doi.org/10.1002/adv.202413256>.
- [47] Guan J, Li GM. DNA mismatch repair in cancer immunotherapy. *NAR Cancer*. 2023; 5: zcad031. <https://doi.org/10.1093/narcan/zcad031>.
- [48] Incorvaia L, Bazan Russo TD, Gristina V, Perez A, Brando C, Mujacic C, *et al.* The intersection of homologous recombination (HR) and mismatch repair (MMR) pathways in DNA repair-defective tumors. *NPJ Precision Oncology*. 2024; 8: 190. <https://doi.org/10.1038/s41698-024-00672-0>.
- [49] Li R, Di L, Li J, Fan W, Liu Y, Guo W, *et al.* A body map of somatic mutagenesis in morphologically normal human tissues. *Nature*. 2021; 597: 398–403. <https://doi.org/10.1038/s41586-021-03836-1>.
- [50] Li C, Egloff AM, Sen M, Grandis JR, Johnson DE. Caspase-8 mutations in head and neck cancer confer resistance to death receptor-mediated apoptosis and enhance migration, invasion, and tumor growth. *Molecular Oncology*. 2014; 8: 1220–1230. <https://doi.org/10.1016/j.molonc.2014.03.018>.
- [51] Contadini C, Ferri A, Cirotti C, Stupack D, Barilà D. Caspase-8 and Tyrosine Kinases: A Dangerous Liaison in Cancer. *Cancers*. 2023; 15: 3271. <https://doi.org/10.3390/cancers15133271>.
- [52] Mandal R, Barrón JC, Kostova I, Becker S, Strebhardt K. Caspase-8: The double-edged sword. *Biochimica et Biophysica Acta. Reviews on Cancer*. 2020; 1873: 188357. <https://doi.org/10.1016/j.bbcan.2020.188357>.
- [53] Kumar-Sinha C, Kalyana-Sundaram S, Chinnaiyan AM. Landscape of gene fusions in epithelial cancers: seq and ye shall find. *Genome Medicine*. 2015; 7: 129. <https://doi.org/10.1186/s13073-015-0252-1>.
- [54] Santaguida S, Richardson A, Iyer DR, M’Saad O, Zasadil L, Knouse KA, *et al.* Chromosome Mis-segregation Generates Cell-Cycle-Arrested Cells with Complex Karyotypes that Are Eliminated by the Immune System. *Developmental Cell*. 2017; 41: 638–651.e5. <https://doi.org/10.1016/j.devcel.2017.05.022>.
- [55] Teixeira VH, Pipinikas CP, Pennycuik A, Lee-Six H, Chandrasekharan D, Beane J, *et al.* Deciphering the genomic, epigenomic, and transcriptomic landscapes of pre-invasive lung cancer lesions. *Nature Medicine*. 2019; 25: 517–525. <https://doi.org/10.1038/s41591-018-0323-0>.
- [56] Denisov EV, Schegoleva AA, Gervas PA, Ponomaryova AA, Tashireva LA, Boyarko VV, *et al.* Premalignant lesions of squamous cell carcinoma of the lung: The molecular make-up and factors affecting their progression. *Lung Cancer (Amsterdam, Netherlands)*. 2019; 135: 21–28. <https://doi.org/10.1016/j.lungcan.2019.07.001>.
- [57] Cheng Z, Mirza H, Ennis DP, Smith P, Morrill Gavarró L, Sokota C, *et al.* The Genomic Landscape of Early-Stage Ovarian High-Grade Serous Carcinoma. *Clinical Cancer Research: an Official Journal of the American Association for Cancer Research*. 2022; 28: 2911–2922. <https://doi.org/10.1158/1078-0432.CCR-21-1643>.
- [58] Lukow DA, Sheltzer JM. Chromosomal instability and aneuploidy as causes of cancer drug resistance. *Trends in Cancer*. 2022; 8: 43–53. <https://doi.org/10.1016/j.trecan.2021.09.002>.
- [59] Davoli T, Uno H, Wooten EC, Elledge SJ. Tumor aneuploidy correlates with markers of immune evasion and with reduced response to immunotherapy. *Science (New York, N.Y.)*. 2017; 355: eaf8399. <https://doi.org/10.1126/science.aaf8399>.
- [60] Ben-David U, Amon A. Context is everything: aneuploidy in cancer. *Nature Reviews. Genetics*. 2020; 21: 44–62. <https://doi.org/10.1038/s41576-019-0171-x>.