Journal of Translational Medicine

RESEARCH

Open Access

CNV-mediated dysregulation of the ceRNA network mechanism revealed heterogeneity in diffuse and intestinal gastric cancers

Rongji Xu[1†], Danni He[1†], Rui Sun[1†], Jiaqi Zhou[2†], Mengyu Xin[1], Qian Liu[1], Yifan Dai[1], Houxing Li[1], Yujie Zhang[1], Jiatong Li[1], XinXin Shan[1], Yuting He[1], Borui Xu[1], Qiuyan Guo[3], Shangwei Ning[1*], Yue Gao[1*] and Peng Wang[1*]

Abstract

Background Gastric cancer (GC) is a highly heterogeneous tumour with high morbidity. Approximately 95% of GC cases are gastric adenocarcinomas, which are further categorized into two predominant subtypes: diffuse gastric cancer (DGC) and intestinal gastric cancer (IGC). These subtypes exhibit distinct pathophysiological and molecular characteristics, reflecting their unique tumorigenic mechanisms.

Method In this study, we employed a comprehensive approach to identify driver genes associated with DGC and IGC by focusing on copy number variation (CNV) genes within the competing endogenous RNA (ceRNA) network. The influence of driver CNV genes on the molecular, cellular, and clinical differences between DGC and IGC was subsequently analysed. Finally, therapeutic strategies for DGC and IGC were evaluated based on the status and functional pathways of the driver CNV genes.

Results A total of 17 and 22 driver CNV genes were identified in DGC and IGC, respectively. These genes drive subtype differences through the ceRNA network, resulting in alterations in the tumour microenvironment (TME). Based on these differences, personalized treatment strategies for DGC or IGC could be developed. Immune checkpoint inhibitors may be an effective treatment option in IGC. Additionally, DGC patients with homozygous deletion of PPIF might benefit from adjuvant chemotherapy, whereas those with high-level amplification of MTAP could respond to targeted therapy.

Conclusion Driver CNV genes were identified to reveal the underlying cause of heterogeneity in DGC and IGC. Furthermore, specific driver CNV genes were identified as potential therapeutic targets, facilitating personalized treatment.

Keywords Diffuse gastric cancer, Intestinal gastric cancer, CNV, CeRNA, Driver genes, Tumour heterogeneity

†Rongji Xu, Danni He, Rui Sun and Jiaqi Zhou these authors wish it to be known that, in their opinion, the first four authors should be regarded as joint First Authors.

*Correspondence:
Shangwei Ning
ningsw@ems.hrbmu.edu.cn
Yue Gao
gaoyue_hrm@163.com
Peng Wang
wpgqy@hrbmu.edu.cn
[1] College of Bioinformatics Science and Technology, Harbin Medical University, Harbin 150081, China
[2] The First Clinical School of Gansu University of Chinese Medicine, Lanzhou 730030, China
[3] The First Affiliated Hospital of Harbin Medical University, Harbin 150081, China

Introduction

Gastric cancer (GC) is a highly heterogeneous malignancy, with diffuse gastric cancer (DGC) and intestinal gastric cancer (IGC) representing the two predominant subtypes. These subtypes exhibit distinct pathophysiological and molecular characteristics. DGC is characterized by disorganized cellular architecture, reduced cell adhesion, and poor differentiation. In contrast, IGC is characterized by a more structured tubular or glandular organization, robust adhesion junctions, and a comparatively lower stromal component density [1]. Compared with IGC, DGC is associated with earlier onset, poorer prognosis, more aggressive progression, and a stronger familial predisposition [2, 3]. The pronounced heterogeneity among GC subtypes, encompassing variations in pathological, physiological, and molecular features, poses substantial challenges for tailoring precise therapeutic strategies. Therefore, elucidating the oncogenic mechanisms and inter-subtype differences is critical for advancing our understanding of GC progression. Such insights could facilitate the development of more targeted and effective therapeutic approaches, ultimately improving clinical outcomes.

Previous studies have suggested that molecular interactions may contribute to the observed tumour heterogeneity [4]. Recently, the competing endogenous RNA (ceRNA) hypothesis has garnered significant attention as a unifying framework for understanding the functional roles of long noncoding RNAs (lncRNAs), pseudogene transcripts, and circular RNAs [5]. Substantial evidence indicates that aberrant ceRNA expression disrupts homeostatic mechanisms, leading to dysregulation of ceRNA networks and consequent disruption of biological functions [6–8]. Advanced computational pipelines have been developed to identify molecular interactions, offering valuable insights into network construction and the identification of genetic markers associated with cancer progression [9–13]. In addition, studies leveraging these comprehensive networks have revealed novel correlations among molecular interactions, thereby underscoring their reliability in predicting potential relationships between noncoding RNAs (ncRNAs) and therapeutic drugs [14, 15]. However, few studies have focused on the impact of genomic alterations on the ceRNA network and thus on tumour heterogeneity. Genomic variations, such as copy number variations (CNVs), play pivotal roles in tumour development by affecting the expression of genes and collectively contribute to intratumour heterogeneity and critically impact prognosis [16, 17].

Therefore, the motivation for this work lies in the significant challenges posed by the heterogeneity of GC. Understanding the underlying causes of this heterogeneity is crucial for developing more effective and personalized treatment strategies. This study focused on the analysis of driver CNV genes affecting the associated ceRNA network to elucidate the discrepancies between DGC and IGC at multiple levels, as well as the underlying molecular mechanisms responsible for these differences. A total of 17 and 22 driver CNV genes were identified in DGC and IGC, respectively. These driver CNVs lead to significant alterations in the tumour microenvironment (TME), highlighting the molecular distinctions between the two subtypes. Based on the unique TME characteristics of each subtype, personalized therapeutic strategies have been proposed for DGC and IGC. Given that the majority of driver CNV genes were associated with immune cell infiltration, IGC may benefit from immunotherapy. On the other hand, DGC exhibited insensitivity to immunotherapy, potentially due to monocyte enrichment in its TME. DGC patients with homozygous deletion of the PPIF gene may respond favourably to chemotherapy, whereas targeted therapies could be effective in patients with high-level amplification of the MTAP gene. These findings offer a new perspective on precision medicine and individualized treatment for patients with gastric cancer and provide substantial support for future research and clinical practice.

Methods

Collection of expression, CNV, and clinical data

Gastric cancer data, including clinical data, CNV data, and gene expression data, were collated from The Cancer Genome Atlas (TCGA) (https://portal.gdc.cancer.gov/). For the clinical data, the samples of "Stomach Adenocarcinoma, Signet Ring Type" and "Stomach, Adenocarcinoma, Diffuse Type" were considered as DGC [18]. Moreover, the samples of "Stomach, Intestinal Adenocarcinoma, Not Otherwise Specified (NOS)", "Stomach, Intestinal Adenocarcinoma, Papillary Type", "Stomach, Intestinal Adenocarcinoma, Mucinous Type" and "Stomach, Intestinal Adenocarcinoma, Tubular Type" were categorized as IGC. Ultimately, clinical data from 85 DGC samples and 191 IGC samples were collected, including variables such as age, sex, and survival time. For CNV data generated using Affymetrix Genome-Wide Human SNP Array 6.0 platform, GISTIC 2.0 was applied to the CNV data at level 3 segmentation [19]. In this study, only homozygous deletions and high-level amplifications were considered reliable CNVs, whereas the remaining CNVs were deemed insignificant and excluded from further analysis. Gene expression profiles for mRNAs, lncRNAs, and miRNAs were obtained from Illumina HiSeq sequencing data (Illumina HiSeq_RNASeqV2 for mRNAs and lncRNAs and Illumina HiSeq_miRNASeq for miRNAs). All fragments per kilobase of transcript per million mapped reads (FPKM) data were log2-transformed

for subsequent analysis. Any mRNA, lncRNA, or miRNA was excluded if its nondetection frequency exceeded 30% or if its average expression level was less than 0.1. To ensure data completeness, only samples containing all three types of data (mRNA, lncRNA, and miRNA) were retained. In the end, 73 DGC samples and 163 IGC samples were included.

Collection of ceRNA interactions

A total of 423,975 miRNA-mRNA interactions and 35,459 miRNA-lncRNA interactions associated with gastric cancer (GC) were retrieved from the StarBase v2.0 database (https://rnasysu.com/encori/) [20]. Additionally, 22,286 miRNA-lncRNA interactions related to GC were obtained from the lnCeDB database (http://gyanxet-beta.com/lncedb) [21].

Identification of driver CNV genes

We developed a comprehensive method to identify genes that influence the heterogeneity of gastric cancer, where CNVs of these genes specifically affect the related ceRNA networks, thereby influencing the heterogeneity of subtypes (Fig. 1). First, genes with a high frequency of CNVs (> 8%) and a significant impact on gene expression were identified as candidate genes. The samples were subsequently stratified into two groups based on the CNV status of each candidate gene: CNV-positive and CNV-negative. For each group, ceRNA networks were constructed by integrating interactions among mRNAs, miRNAs, and lncRNAs while excluding interactions with an absolute correlation coefficient less than 0.25. A comparative analysis of the ceRNA networks between the two groups was then conducted. If a candidate gene-associated ceRNA network was exclusively observed in one group, the CNV of the candidate gene was inferred to influence the ceRNA network. Finally, genes with CNVs that significantly altered the ceRNA network were designated as driver CNV genes (Supplementary Table 1).

Validation of driver CNV genes through external datasets and methods

To evaluate the consistency of our method with other approaches, we reanalysed our data using previously published methodologies and performed comparative analyses. First, as described by Xu et al., the chi-square test was used to construct CNV-related ceRNA networks [22]. Next, as demonstrated by Ding et al., we integrated CNV data with gene expression alterations to identify CNV-driven lncRNA-associated ceRNAs [23]. Finally, following the method of Wang et al., a multivariate multiple regression model was applied to investigate whether specific CNV events regulate the expression of ceRNA axes [24]. The three methods mentioned above were used to validate the stability of our approach by applying other techniques to our data. To further assess the robustness of the driver CNV genes, we applied our method to identify these genes in three independent public GC datasets and performed a comparative analysis. Independent validation datasets, including GSE51575 [25], GSE26899 [26], and GSE62717 [27] (Supplementary Table 2), were obtained from the GEO database (https://www.ncbi.nlm.nih.gov/geo/). Specifically, in the GSE51575 and GS26899 datasets, differentially expressed genes (DEGs) between DGC/IGC and normal samples were identified. These DEGs were then subjected to a hypergeometric test to assess their overlap with the driver CNV genes of DGC or IGC, thereby validating the robustness of these genes. Additionally, in the GS62717 dataset, CNV genes in DGC or IGC samples were selected and compared with the driver CNV genes of DGC or IGC using the hypergeometric test. Furthermore, we cross-referenced our identified CNV genes with GC-related gene sets from the GeneCards database to explore their potential relevance in GC pathogenesis [28]. Previous studies have classified DGC into two subtypes based on transcriptomic analysis: the intestinal-like (INT) subtype and the core diffuse-type (COD) subtype [29]. The corresponding expression matrices were obtained from the GSE113255 dataset (Supplementary Table 2), which included 44 INT samples and 55 COD samples.

Construction and visualisation of networks

The protein–protein interaction (PPI) networks of driver CNVs were constructed using the STRING database (https://cn.string-db.org/) [30]. And the visualisation of the ceRNA networks was performed using Cytoscape v3.9.1.

Functional analysis of driver CNV genes

To further investigate the biological processes influenced by the driver CNV genes, the R package clusterProfiler (v4.12.0) was used to associate these genes with cancer-related pathways and characteristics. A total of 50 cancer-related features were curated from the MsigDB database (https://www.gsea-msigdb.org/gsea/msigdb) [31].

Calculation of immune cell infiltration and immune evasion scores

To verify the relationship between different GC subtypes and the TME, CIBERSORT was used to predict the infiltration of immune cells. Furthermore, TIDE (http://tide.dfci.harvard.edu/login/) was used to calculate the immune evasion score of each sample based on the normalized expression matrix [32]. Specifically, the

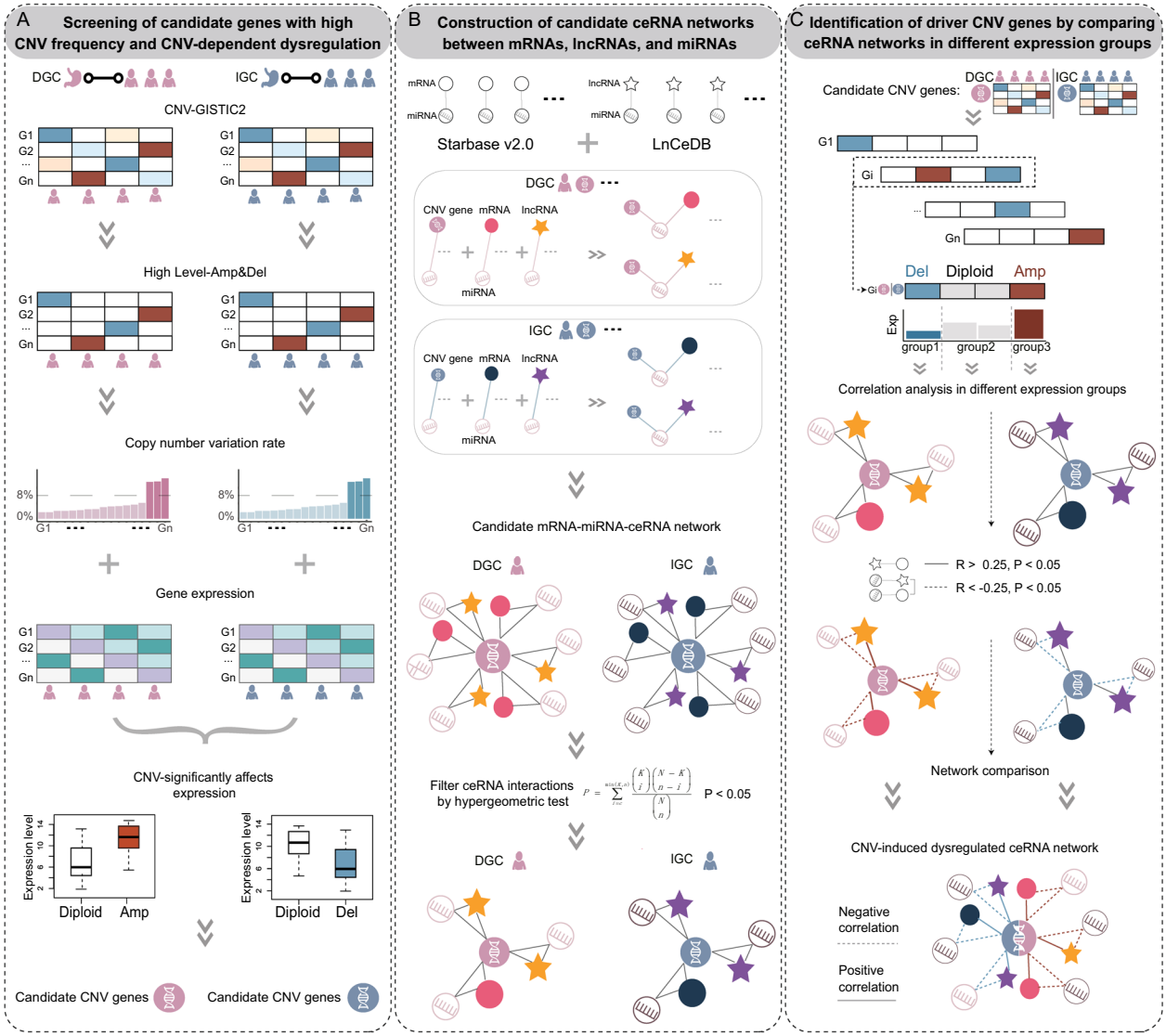

Fig. 1 Flowchart for the identification of driver CNV genes. **A** Selection of candidate CNV genes with high-frequency copy number variations (CNVs) and a significant impact on gene expression. **B** Construction of candidate ceRNA networks based on interactions among mRNAs, miRNAs, and lncRNAs. **C** Identification of driver CNV genes inferred to modulate the ceRNA network

normalization method involved subtracting the average expression value of each gene across all samples from its expression value in each sample. The normalized expression matrix was then uploaded to the website to obtain the immune evasion score, T-cell dysfunction score, and T-cell exclusion score.

Consensus clustering based on ICR-related genes

We used 5 clustering methods (hclust, kmeans, skmeans, pam, and mclust) to infer potential stable consensus subgroups of 2–6 clusters. The most stable partitions were selected from all methods based on an inspection of the membership matrix.

Identification and analysis of GC cell subpopulations

Single-cell data for GC were obtained from the GEO database (GSE183904, Supplementary Table 2) [33]. First, genes that were expressed in fewer than 5 cells and cells that expressed fewer than 300 genes were excluded. Then, cells with more than 200 and fewer than 2500 expressed genes and a mitochondrial content less than 0.05 were excluded. Finally, mitochondrial genes beginning with "MT" were excluded. The final expression profiles

included 20,974 genes from 9,619 cells in 6 DGC samples and 22,758 genes from 29,860 cells in 14 IGC samples. The R package Seurat (v4.4.0) was used to identify the cell subpopulations, whereas CellMarker 2.0 (http://bio-bigdata.hrbmu.edu.cn/CellMarker/) [34] was used for the annotation of these cell subpopulations. CIBERSORTx (https://cibersortx.stanford.edu/upload.php) was used to estimate the proportion of cell subpopulations in each sample.

Calculation of driver CNV gene activity

The R package AUCell (v1.18.0) was used for calculations to assess the activity of driver CNV genes across different cell types, with gene activity levels quantified by AUC values. In general, a positive correlation was observed between activity and the AUC value.

Construction of risk score models

Considering the synergistic impact of multiple driver CNV genes on the prognosis of INT and COD patients, a risk score model was constructed based on CNV to estimate patient prognosis.

$$Riskscore = \sum_{i=1}^{17} e_i \times cnv_i,$$

where e_i represents the expression of the driver CNV gene and cnv_i represents the CNV status of the driver CNV gene (−2 or 2). An elevated risk score was indicative of a worse prognosis.

Evaluation of resistance to targeted drugs

We collected data on drugs that are highly resistant or highly sensitive in DGC and IGC from published studies [35] and evaluated their efficacy using the half maximal inhibitory concentration (IC50) values. Specifically, the IC50 values of DGC samples for different targeted drugs were calculated using the R package OncoPredict (v1.2) based on the expression data of driver CNV genes. The IC50 refers to the concentration of a drug that results in 50% inhibition of a specific biological or biochemical function based on the dose–response curve. It is commonly used to assess a drug's pharmacodynamic activity, selectivity, toxicity, and resistance. A lower IC50 value typically indicates greater efficacy in cancer inhibition.

Statistics

All comparative analyses were conducted using the hypergeometric test, t-test, and Wilcoxon rank-sum test, as appropriate. Pearson correlation analysis was used to assess the relationship between driver CNV genes and immune checkpoint signature genes (Supplementary Table 3). Moreover, the same method was used to assess the correlation between driver CNV genes and senescence signature genes (Supplementary Table 4) [36]. Survival analyses were conducted using the Kaplan–Meier method, and differences were tested using the log-rank test. $P < 0.05$ was considered statistically significant. All the statistical analyses were performed in the R4.2.3 software environment.

Results

Distinct driver CNV genes influence divergent developmental mechanisms in GC subtypes

The comparison and identification of driver CNV genes and their ceRNA networks in DGC and IGC facilitate an understanding of the developmental mechanisms of GC subtypes. Here, a comprehensive method was developed to identify driver CNV genes in different GC subtypes based on the dysregulated ceRNA network (Fig. 1). A total of 17 and 22 driver CNV genes were found in DGC and IGC, respectively (Fig. 2A, Supplementary Table 1). To validate the robustness of these genes, we first used other published methods to re-identify key genes. Our findings revealed that most GC subtype-specific CNV genes identified by our method were validated by other methods (Fig. 2B). Next, using a hypergeometric test, we demonstrated that the driver CNV genes identified by our method were significantly enriched among the differentially expressed and CNV-altered genes from three independent public GC datasets (Fig. 2B, C). Furthermore, we observed a significant overlap between the driver CNV genes and the experimentally validated GC-related genes in the GeneCards database (Fig. 2B, C). Notably, many of these CNV genes were identified as potential drug targets for GC, suggesting their possible roles in GC pathogenesis and their potential therapeutic relevance (Fig. 2B).

Interestingly, the number of candidate genes in DGC was significantly lower than that in IGC at each step of identification of the driver CNV genes (Fig. 2A). The number of DGC-specific mRNAs, miRNAs, and lncRNAs was relatively high across the entire ceRNA network (Fig. 2D). These findings indicated the presence of genetic and biological differences between DGC and IGC. The majority of driver CNV genes in DGC exhibited high-level amplification, except for MTAP and PPIF. Similarly, MTAP, ITGAV, IRF2, and CDKN2B were homozygously deleted in IGC samples, whereas the remaining driver CNV genes exhibited high-level amplification (Fig. 2E). These findings highlight genetic and biological differences between DGC and IGC, with high-level amplification being a dominant phenomenon in both subtypes.

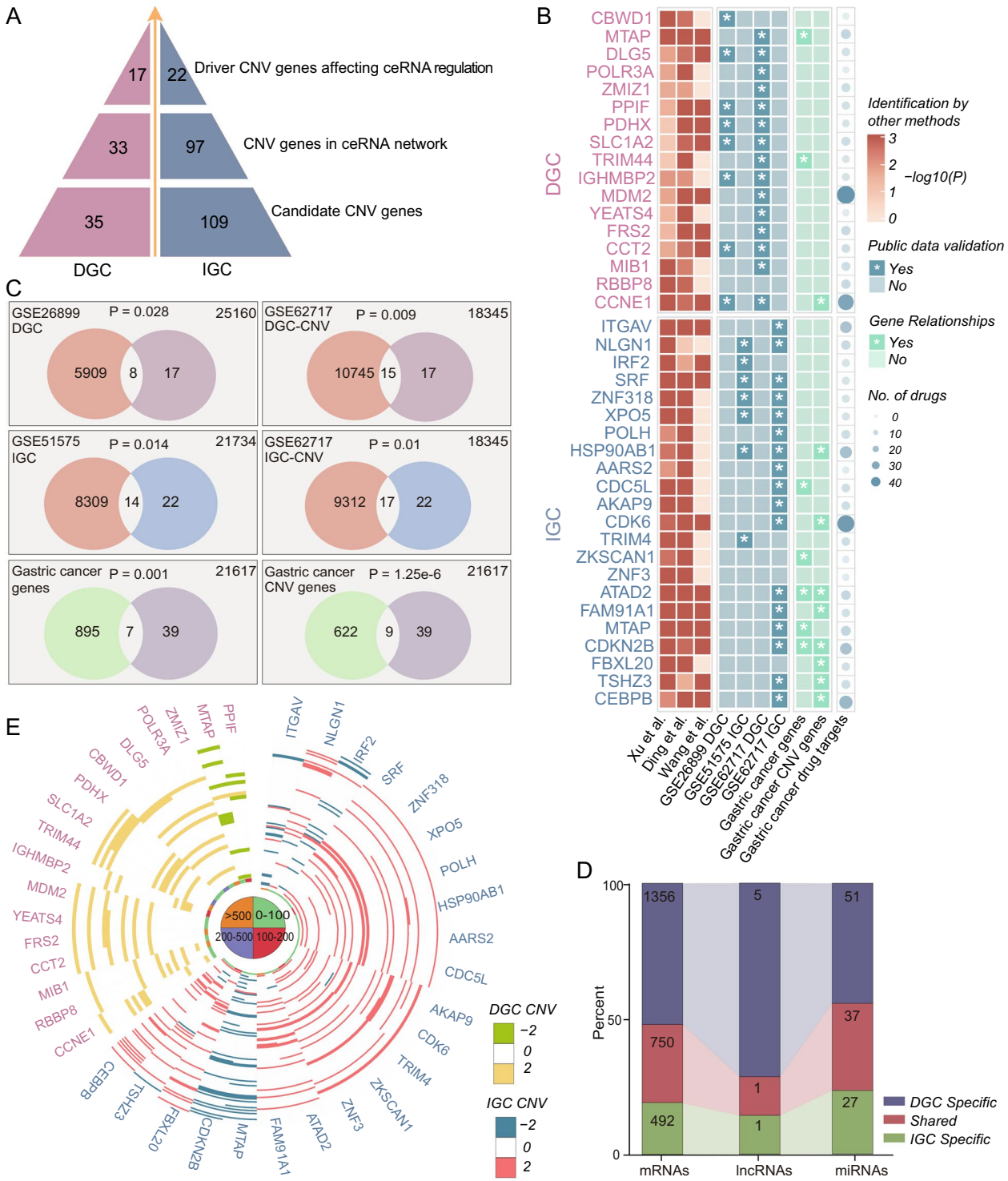

Fig. 2 Multilevel landscape of DGC and IGC. **A** Number of candidate genes identified at each step of the screening process. **B** Validation of driver CNV genes using published methods and independent datasets. **C** Overlap between driver CNV genes identified from independent public datasets and those identified in this study. **D** Distribution of driver CNV genes and ceRNAs in DGC and IGC. **E** Comparison of shared and unique ceRNA interactions in DGC and IGC

Driver CNV genes affect the TME by targeting the corresponding ceRNA networks

Despite their distinct pathophysiology and clinical features, DGC and IGC share interconnected mechanisms. A PPI network of driver CNV genes was constructed based on the STRING database, revealing interactions among their encoded proteins (Fig. 3A). The association between DGC and IGC was represented in the ceRNA network in two ways. One way was that the same driver CNV genes were identified in both DGC and IGC. For example, MTAP could influence the ceRNA network in IGC through hsa-miR-30a-3p and affect the ceRNA network in DGC via hsa-miR-195-5p (Fig. 3B). The other way was that the same miRNA regulated distinct driver CNV genes in DGC and IGC. For example, hsa-miR-155-5p targeted ZKSCAN1 in IGC and TRIM44 in DGC (Fig. 3C). Numerous physiological and pathological processes, including the inhibition of viral infection and tumour proliferation, are mediated by hsa-miR-155-5p [37]. Similarly, several different driver CNV genes of IGC and DGC were found to be regulated by identical miRNAs (Supplementary Fig. 1). These findings illustrated the complexity of the ceRNA network and its potential significance in tumour development. In addition, enrichment analysis of the driver CNV genes in IGC and DGC revealed both common and subtype-specific pathways (Fig. 3D). Notably, the IGC-specific pathway "IL6_JAK_STAT3_SIGNALING" and the DGC-specific pathway "Wnt_BETA_CATENIN_SIGNALING" were both immune related [38, 39], suggesting that differences in the tumour microenvironment (TME) between DGC and IGC are driven by distinct driver CNV genes.

IGC is more sensitive to immunotherapy than DGC

The investigation of the differences in the TME between DGC and IGC provides a theoretical basis for understanding their distinct biological behaviours and potential therapeutic strategies. CIBERSORT was used to quantify the infiltration of immune cells in DGC and IGC, revealing significant differences in the infiltration scores of distinct immune cell types between the two groups (Fig. 4A). Furthermore, the immune cell types identified through CIBERSORT were classified into four functional groups based on a previous study [40], and these group-based differences in immune cell infiltration scores between DGC and IGC remained consistent (Fig. 4B). Notably, compared with DGC, IGC presented a greater number of driver CNV genes associated with immune cell infiltration scores (Fig. 4C), suggesting that IGC is more likely to benefit from immune checkpoint therapies. Compared with IGC, DGC had significantly greater T-cell dysfunction scores, T-cell exclusion scores, and TIDE scores, indicating a less favourable immune microenvironment for immunotherapy efficacy (Fig. 4D). Additionally, the immune rejection constant (ICR), a modular immune gene signature associated with immunotherapy responsiveness in cancer [41], was used to stratify IGC samples into three immune subtypes (Fig. 4E). Among the 22 driver CNV genes identified in IGC, IRF2 was significantly upregulated in the ICR-high subtype, whereas HSP90AB1, AARS2, ZKSCAN1, and FAM91A1 were markedly downregulated (Fig. 4F). The ICR-high subtype, characterized as a "hot tumour," exhibited elevated expression of immune regulatory genes (Fig. 4G) and was more likely to respond positively to immunotherapy. Therefore, the combination of high IRF2 expression and low expression of HSP90AB1, AARS2, ZKSCAN1, and FAM91A1 may serve as a predictive biomarker for improved prognosis in IGC patients receiving immune checkpoint inhibition therapy [42, 43].

Driver gene CNVs in monocytes contribute to immunotherapeutic resistance in DGC

Modifications at the cellular level typically reflect the effects of immune checkpoint therapy. Comparing the differences at the single-cell level between IGC and DGC may reveal the mechanisms underlying DGC resistance to immune therapy, providing valuable insights for further screening and optimizing treatment strategies for DGC. A total of 10 and 9 cell clusters were identified in DGC and IGC based on single-cell data from the GEO database (Fig. 5A, B). Interestingly, the DGC driver CNV gene CCT2 and the IGC driver CNV gene HSP90AB1 were highly expressed across a range of cell clusters. In contrast, the DGC driver CNV gene PPIF was specifically expressed in monocytes, whereas the IGC driver CNV gene ATAD2 was specifically highly expressed in gastric isthmus cells (Fig. 5C, D).

To accurately assess the enrichment of driver CNV genes in DGC or IGC cells, AUCell was used to calculate the activity of these genes across different cell types. Among them, the activity of DGC driver CNV genes was notably greater in DGC monocytes (Fig. 5E). Monocytes potentially influence the effectiveness of tumour immunotherapy through immune evasion mechanisms [44]. In our study, we further performed enrichment analysis of DEGs between DGC and IGC monocytes. These genes are involved in processes such as antigen presentation and processing and proton transmembrane transport (Fig. 5F). Moreover, the AUCell scores of DGC monocytes were significantly greater than those of other cell types (Fig. 5G). These findings suggest that the driver CNV genes in DGC influence monocytes in the TME, thereby affecting the immune response and ultimately limiting the effectiveness of immunotherapy in DGC. Additionally, the significant difference in AUCell scores

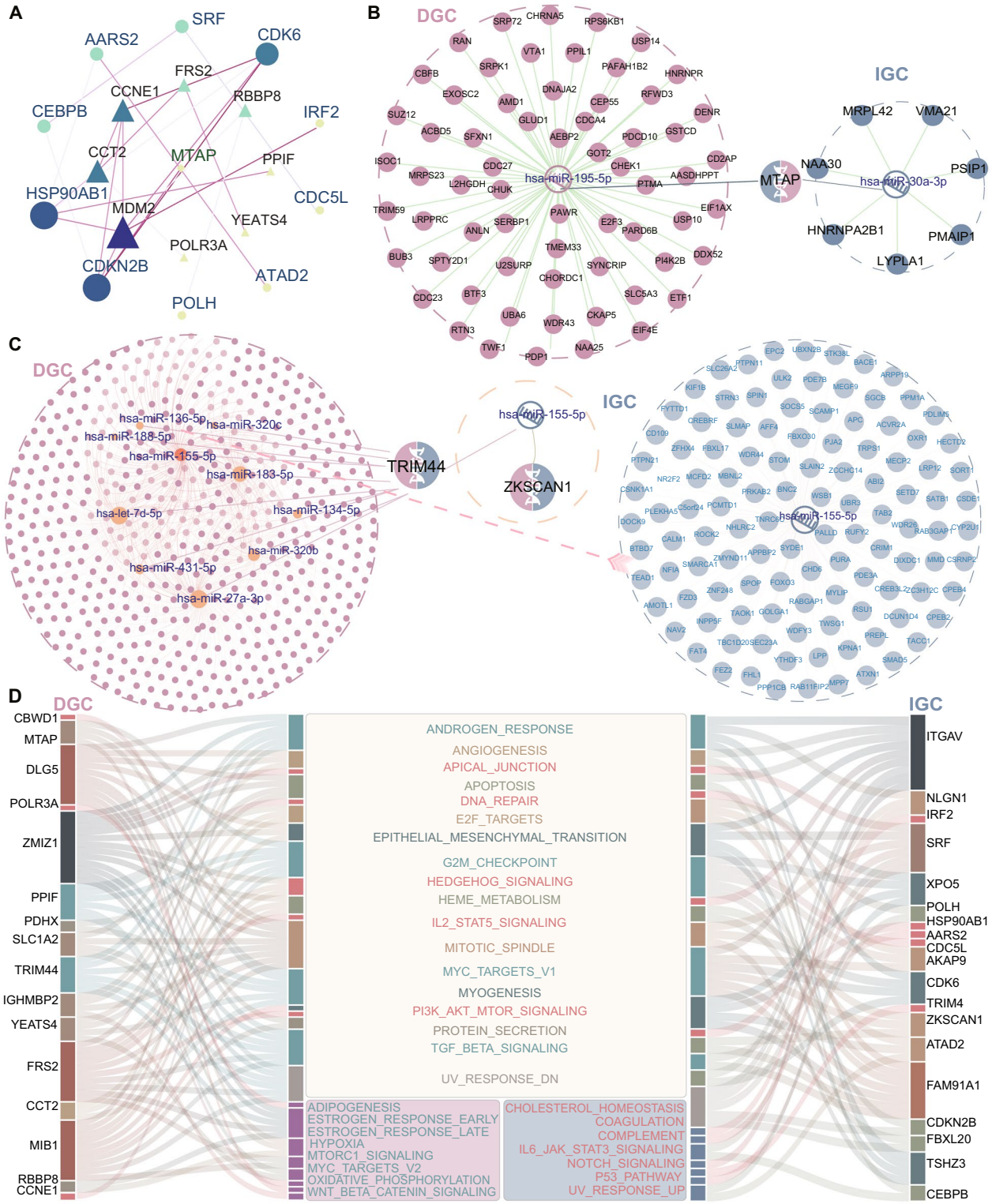

Fig. 3 Regulatory roles of driver CNV genes in the ceRNA network. **A** The PPI network of driver CNV genes in DGC and IGC. **B** CeRNA network regulated by a shared driver CNV gene in DGC and IGC. **C** Hsa-miR-155-5p regulates distinct driver CNV genes in DGC and IGC. **D** Enrichment analysis of driver CNV genes in DGC and IGC

between B cells and other cell types in IGC (Fig. 5H) suggests that B cells might play a distinct role in immune responses within IGC.

DGC might benefit from adjuvant chemotherapy or targeted therapy

Next, we identified suitable treatments for DGC patients through analysis of the mechanism of DGC anti-immune checkpoint therapy. First, 178 immune checkpoint related genes were manually gathered (Supplementary Table 3), and they were screened for immune checkpoint genes that correlated with driver CNV genes (Fig. 6A, Supplementary Figs. 2 and 3). The DGC driver CNV gene ZMIZ1 was found to be negatively correlated with PD-L1 expression and to be targeted by the greatest number of miRNAs within its dysregulated ceRNA network. Although PD-L1 expression does not necessarily indicate the efficacy of immunotherapy, it typically indicates a high level of PD-L1 antibody targets expressed in the tumour, which is correlated with comparatively better efficacy [45]. Therefore, the high level of amplification of ZMIZ1 in DGC was identified as a potential factor contributing to the poor immunotherapeutic efficacy of DGC.

Furthermore, identifying efficacious therapeutic strategies for DGC is critical. A previous study classified DGC into two subtypes: the INT and the COD. The COD was responsive to immunotherapy and engaged in EMT-related processes, whereas the INT subtype was associated with DNA repair and the cell cycle and was responsive to adjuvant chemotherapy. In our study, the ceRNA network mediated by PPIF was significantly enriched for the DNA repair pathway, which was consistent with the INT subtype, whereas the ceRNA network mediated by SLC1A2 was significantly enriched for the TGF-beta signaling pathway, which was consistent with the COD subtype (Fig. 6B). Moreover, a significant difference in PPIF expression was observed between INT and COD subtypes, however, no variation in SLC1A2 expression was detected (Fig. 6C). Furthermore, a risk score based on CNV was constructed, which revealed that the COD subtype had a worse prognosis than the INT subtype, as indicated by a higher risk score (Fig. 6D). Consequently, DGC patients with homozygous deletion of PPIF might benefit from adjuvant chemotherapy.

To develop effective targeted therapeutic strategies, identifying potential targets that are specifically overexpressed in DGC is essential. Previous studies have identified drugs that are highly resistant or highly sensitive in DGC and IGC (Fig. 6E), with their effectiveness evaluated by IC50 values. By comparing the differences in IC50 values across samples with varying expression levels of driver CNV genes, precise and personalized treatment strategies can be provided for DGC patients. Specifically, the expression of MDM2 exhibited a significant positive correlation with the IC50 values of afatinib, AZD4547, and trametinib, indicating a worsened drug response with increased MDM2 levels (Fig. 6F). Similarly, a significant negative correlation was observed between the IC50 value of ibrutinib and MTAP expression, suggesting that increased MTAP levels were associated with increased medication efficacy (Fig. 6F).

Overall, targeted therapy, immunotherapy, and chemotherapy have facilitated the implementation of precision medicine for DGC and IGC on a personalized basis. Our research has enhanced the efficacy and relevance of treatment by facilitating the development of individualized treatment regimens.

Driver CNV genes lead to age- or sex-specific survival differences in IGC and DGC

Genomic and cellular alterations ultimately result in modified clinical characteristics. Therefore, the differences in clinical characteristics between DGC and IGC patients were further investigated (Fig. 7A). Notably, IGC patients who were less than or equal to 60 years of age exhibited significantly prolonged survival compared with that in IGC patients who were greater than 60 years of age (Fig. 7B), suggesting that aging has an impact on the prognosis of IGC. A total of 125 genes and their corresponding status associated with senescence were subsequently assembled (Supplementary Table 4), and the correlations between driver CNV genes of IGC and these senescence genes were calculated (Fig. 7C). In the case of gene types such as growth factors and protease inhibitors, there was a strong correlation between driver CNV genes and senescence genes ($R > 0.5$), implying that IGC driver CNV genes may play an important role in gene expression, tissue development, and cell growth within the regulatory network.

(See figure on next page.)

Fig. 4 Quantitative analysis of tumour-infiltrating immune cells. **A** Boxplot of the differences in the infiltration of 22 immune cell types between DGC and IGC ("***": $P < 0.001$; "**": $P < 0.01$; "*": $P < 0.05$). **B** Differences in the infiltration of the 4 major immune cell types. **C** Heatmap of the correlation between the expression of driver CNV genes and the degree of immune cell infiltration in DGC and IGC. **D** Differences in TIDE scores between DGC and IGC. **E** Consistent clustering of IGC based on ICR-related genes. **F** Differential expression of driver CNV genes in the ICR subtype. **G** Heatmap of immune-related gene expression patterns across ICR subtypes

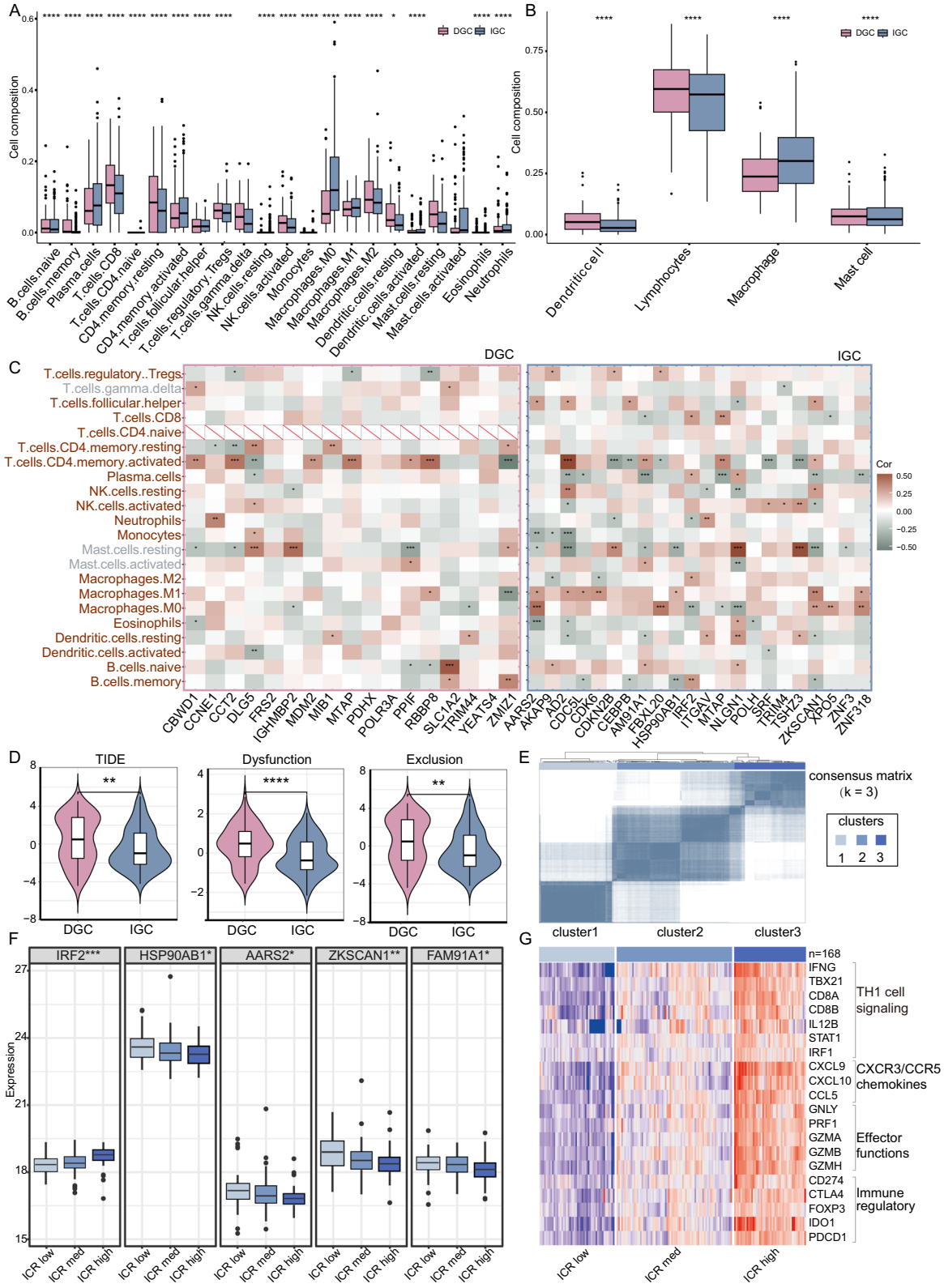

Fig. 4 (See legend on previous page.)

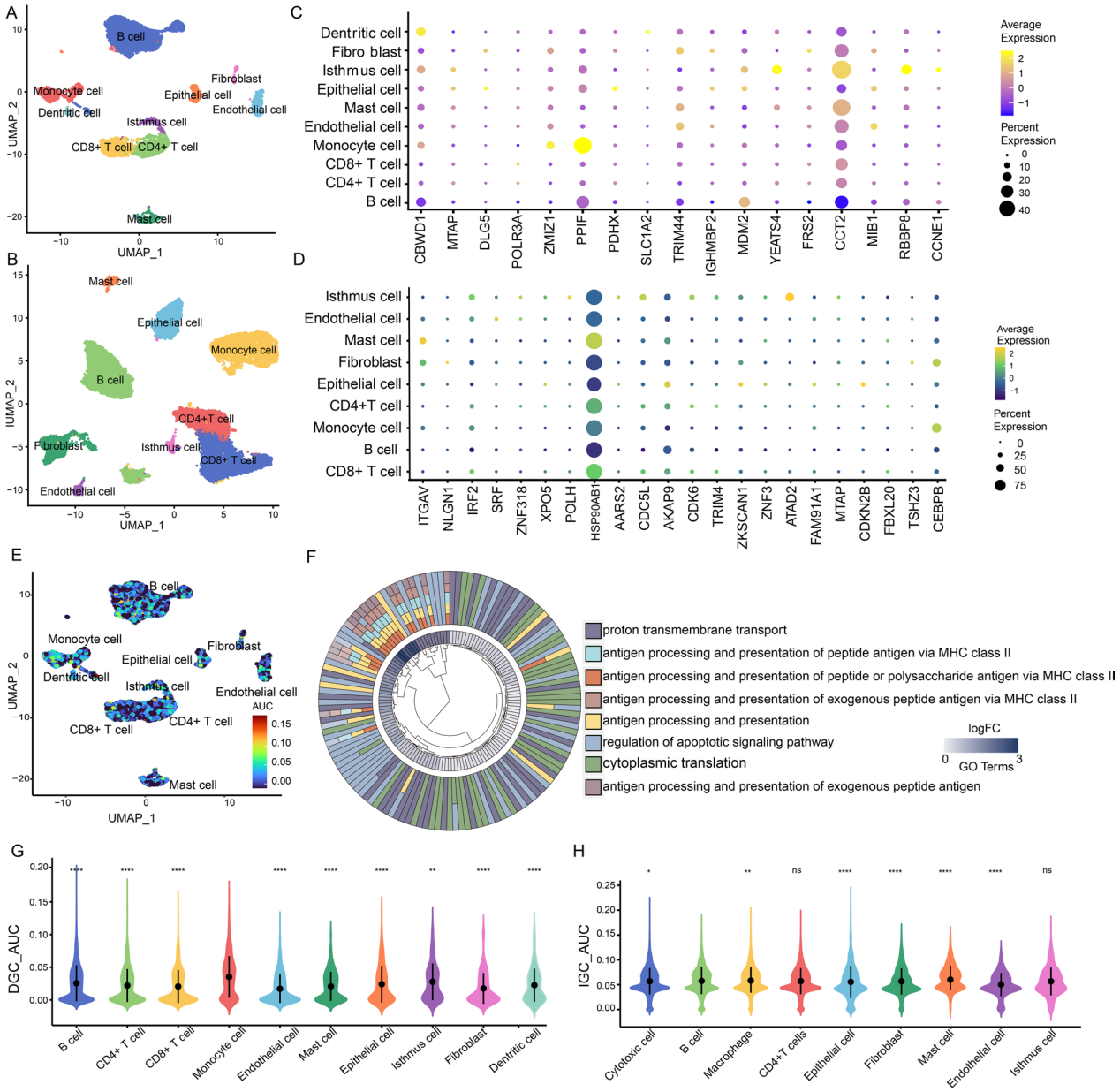

Fig. 5 Single-cell level analysis of driver CNV genes. **A**, **B** Dimensionality reduction (UMAP) of cells in DGC (**A**) and IGC (**B**). **C**, **D** Average expression and proportion of driver CNV genes in DGC (**C**) and IGC (**D**) among different cell clusters. **E** Distribution of AUC values in DGC cell clusters. **F** LogFC values (inside) and enrichment analyses (outside) of DEGs between DGC and IGC in monocytes. **G**, **H** Box plots of AUC values for driver CNV genes in DGC (**G**) and IGC (**H**) among cell clusters ("****": $P < 0.001$, "**": $P < 0.01$)

(See figure on next page.)

Fig. 6 Therapeutic targets of DGC. **A** Correlations of immune checkpoint-related genes with driver CNV genes in DGC and IGC and miRNAs associated with driver CNV genes. The left column represents the 5 common immune checkpoints. In the middle column, blue represents the driver CNV genes in IGC, and brown represents the driver CNV genes in DGC. The right column represents the miRNAs associated with each driver CNV gene. **B** The ceRNA networks regulated by specific driver CNV genes in DGC or IGC are functionally identical to those in the INT or COD subtypes. **C** Boxplots of differences in the expression of specific driver CNV genes between INT and COD subtypes. **D** Boxplots of differences in risk scores between INT and COD subtypes. **E** Heatmap of drug sensitivity between DGC and IGC. Drugs are clustered according to their known target. **F** Comparison of IC50 values for drugs in DGC samples with different driver CNV gene expression levels ("**": $P < 0.01$, "*": $P < 0.05$)

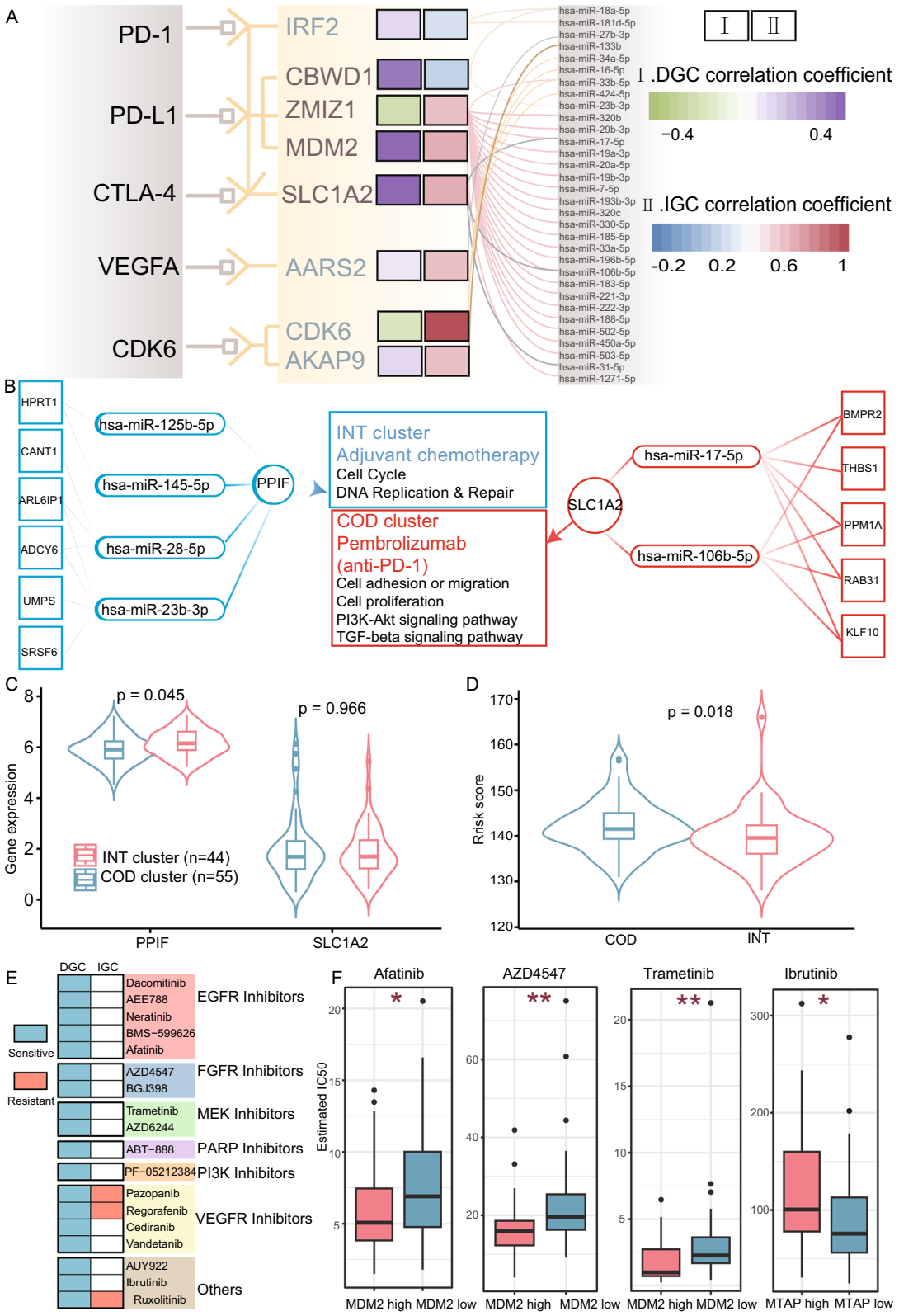

Fig. 6 (See legend on previous page.)

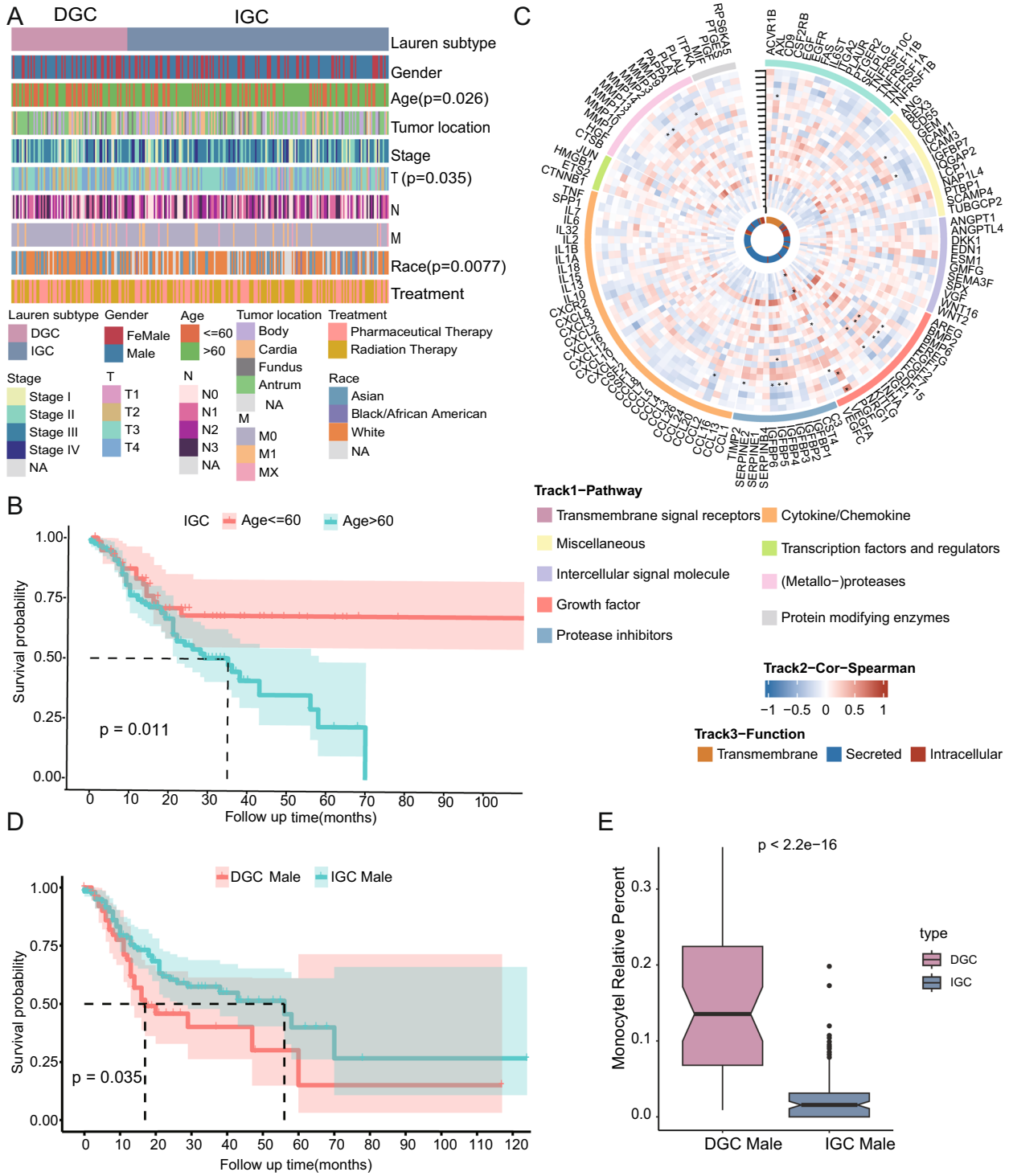

Fig. 7 Differences in clinical characteristics between DGC and IGC patients. **A** Clinical characteristics of DGC and IGC patients. P-values were calculated by the two-sided Fisher's exact test, and values less than 0.05 indicate a difference between DGC and IGC. **B** Survival analysis of IGC samples from different age groups. **C** Enrichment analysis of senescence genes (outside), heatmap of the correlation between driver CNV genes and senescence genes in the IGC (middle, "*": $P < 0.05$), and the status of senescence genes (inside). **D** Survival analysis of male DGC and IGC samples. **E** Differences in the proportions of monocytes between male DGC and IGC samples

In addition, the prognosis for male patients with DGC was worse than that for those with IGC (Fig. 7D), which might be attributed to the complex interplay between cellular characteristics and clinical characteristics. Accordingly, we observed that the proportion of monocytes in male DGC patients was significantly greater than that in male IGC patients (Fig. 7E). As a result, the expression of driver CNV genes enriched in monocytes of DGC was increased, which significantly affects the associated ceRNA networks, ultimately resulting in a poorer prognosis.

To summarize, alterations in samples at various levels ultimately influence clinical outcomes. Therefore, comprehending the molecular mechanisms underlying these variations is critical to facilitate the development of personalized therapeutic strategies.

Discussion

Previous studies have shown that CNV plays a vital role in cancer development. However, the impact of CNV on the ceRNA network and its relationship with tumour heterogeneity, especially in the context of DGC and IGC, has not been thoroughly investigated. Therefore, investigating genomic variation to identify driver CNV genes of different GC subtypes will generate novel findings, which is conducive to further revealing the pathogenesis of GC. To fill this gap, comprehensive techniques were used to identify specific driver CNV genes in DGC and IGC. By analysing the differences and associations between DGC and IGC at the genomic, transcriptomic, and cellular levels, potential molecular mechanisms responsible for the heterogeneity of GC were illustrated. For example, genes such as MTAP and PPIF in DGC and IGC showed different CNV patterns, which may lead to different biological behaviours of the two subtypes. This knowledge can potentially lead to the identification of specific driver CNV genes as potential therapeutic targets, thereby providing novel results for personalized treatment and improving the prognosis of GC patients.

A total of 17 and 22 identical or different driver CNV genes in DGC and IGC, respectively, suggested biological and genetic similarities or differences between the two gastric cancer subtypes. These differences might lead to alterations in cancer proliferation, clinical characteristics, and prognosis. Notably, the DGC-specific pathway "Wnt_BETA_CATENIN_SIGNALING" and the IGC-specific pathway "IL6_JAK_STAT3_SIGNALING" were highly correlated with immunity. Therefore, an analysis of immune cell infiltration in DGC and IGC was conducted, which revealed that the majority of immune cells were significantly different between the two subtypes. Moreover, a strong correlation was observed between immune cell infiltration and driver CNV genes in IGC, suggesting that IGC patients might benefit from immunotherapy. Furthermore, PPIF was activated in monocytes, which have been associated with immune evasion, confirming that immunotherapy was not applicable to DGC. To identify potential therapeutic strategies for DGC, DGC samples were classified into two subtypes based on previous studies. The INT subtype was enriched in DNA repair-related pathways and was deemed suitable for adjuvant chemotherapy. Moreover, DGC patients with high-level amplification of MTAP might benefit from targeted therapy.

There is a growing body of literature that uses multiomics data to investigate cancer heterogeneity [46–48], and our study contributes novel and pertinent insights to this field. By integrating transcriptomic and genomic data, we identified driver CNV genes associated with GC subtypes and identified personalized therapeutic strategies based on the relationship between these genes and therapeutic targets. Moreover, the screening process for driver CNV genes could be used to identify specific genes in other cancer types, facilitating the study of other types of cancer. Nevertheless, there are several limitations to this study. One limitation is that although the treatment response was assessed using IC50 values, the lack of real data could bias our results. Future studies should include more comprehensive treatment response data to validate the identified therapeutic targets. Additionally, our study is limited to the analysis of transcriptomic, genomic, and immune cell infiltration data. The incorporation of more data, such as epigenomic and proteomic data, could provide a more comprehensive understanding of the heterogeneity of GC [49, 50].

Conclusion

In conclusion, our multiomics analysis revealed that different gastric cancer subtypes exhibit different molecular mechanisms and TMEs. The driver CNV genes identified through our analyses have prognostic implications and might serve as potential therapeutic targets for immunotherapy or targeted therapy. These findings provide new insights into the basic biology of gastric cancer and may have an impact on the development of personalized therapies for gastric cancer patients.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12967-025-06222-x.

Supplementary file 1

Acknowledgements
Not applicable

Author contributions
Conceptualization, P.W., Y.G. and S.N.; methodology, R.X.; validation, D.H., R.X., Y.Z., J.L., Q.G. and X.S.; formal analysis, J.Z., R.S., Y.H. and B.X.; resources, M.X., Q.L., Y.D. and H.L; writing-original draft preparation, D.H.; writing-review and editing, P.W. and Y.G.; visualization, R.X. and D.H.; supervision, P.W.. All authors have read and agreed to the published version of the manuscript.

Funding
This research was funded by The National Natural Science Foundation of China (82373408, 32370718); the Heilongjiang Provincial Natural Science Foundation (YQ2024C040); The Postdoctoral Science Foundation of China (2023T160172).

Availability of data and materials
The datasets supporting the conclusions of this article are available in the TCGA database (The Cancer Genome Atlas, https://portal.gdc.cancer.gov/), Starbase v2.0 database (https://rnasysu.com/encori/), lncCeDB database (http://gyanxet-beta.com/lncedb), GEO database (https://www.ncbi.nlm.nih.gov/geo/), MsigDB database (https://www.gsea-msigdb.org/gsea/msigdb).

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare that they have no competing interests.

Received: 13 November 2024 Accepted: 11 February 2025
Published online: 11 March 2025

References

1. Lauren P. The two histological main types of gastric carcinoma: diffuse and so-called intestinal-type carcinoma. an attempt at a histo-clinical classification. Acta Pathol Microbiol Scand. 1965;64:31–49.
2. Fukamachi H, Kim SK, Koh J, Lee HS, Sasaki Y, Yamashita K, Nishikawaji T, Shimada S, Akiyama Y, Byeon SJ, et al. A subset of diffuse-type gastric cancer is susceptible to mTOR inhibitors and checkpoint inhibitors. J Exp Clin Cancer Res. 2019;38(1):127.
3. Shen L, Shan YS, Hu HM, Price TJ, Sirohi B, Yeh KH, Yang YH, Sano T, Yang HK, Zhang X, et al. Management of gastric cancer in Asia: resource-stratified guidelines. Lancet Oncol. 2013;14(12):e535-547.
4. Carter B, Zhao K. The epigenetic basis of cellular heterogeneity. Nat Rev Genet. 2021;22(4):235–50.
5. Thomson DW, Dinger ME. Endogenous microRNA sponges: evidence and controversy. Nat Rev Genet. 2016;17(5):272–83.
6. Guo Q, Liu Q, He D, Xin M, Dai Y, Sun R, Li H, Zhang Y, Li J, Kong C, et al. LnCeCell 20: an updated resource for lncRNA-associated ceRNA networks and web tools based on single-cell and spatial transcriptomics sequencing data. Nucleic Acids Res. 2025;53(1):D107–15.
7. Wang P, Guo Q, Qi Y, Hao Y, Gao Y, Zhi H, Zhang Y, Sun Y, Zhang Y, Xin M, et al. LncACTdb 3.0: an updated database of experimentally supported ceRNA interactions and personalized networks contributing to precision medicine. Nucleic Acids Res. 2022;50(D1):D183–9.
8. Wang P, Guo Q, Hao Y, Liu Q, Gao Y, Zhi H, Li X, Shang S, Guo S, Zhang Y, et al. LnCeCell: a comprehensive database of predicted lncRNA-associated ceRNA networks at single-cell resolution. Nucleic Acids Res. 2021;49(D1):D125–33.
9. Liu L, Wei Y, Zhang Q, Zhao Q. SSCRB: predicting circRNA-RBP interaction sites using a sequence and structural feature-based attention model. IEEE J Biomed Health Inform. 2024;28(3):1762–72.
10. Zhang L, Yang P, Feng H, Zhao Q, Liu H. Using network distance analysis to predict lncRNA-miRNA interactions. Interdis Sci Comput Life Sci. 2021;13(3):535–45.
11. Wang W, Zhang L, Sun J, Zhao Q, Shuai J. Predicting the potential human lncRNA-miRNA interactions based on graph convolution network with conditional random field. Briefings Bioinf. 2022. https://doi.org/10.1093/bib/bbac463.
12. Chen Z, Zhang L, Sun J, Meng R, Yin S, Zhao Q. DCAMCP: A deep learning model based on capsule network and attention mechanism for molecular carcinogenicity prediction. J Cell Mol Med. 2023;27(20):3117–26.
13. Xie J, Xu P, Lin Y, Zheng M, Jia J, Tan X, Sun J, Zhao Q. LncRNA-miRNA interactions prediction based on meta-path similarity and Gaussian kernel similarity. J Cell Mol Med. 2024;28(19): e18590.
14. Yin S, Xu P, Jiang Y, Yang X, Lin Y, Zheng M, Hu J, Zhao Q. Predicting the potential associations between circRNA and drug sensitivity using a multisource feature-based approach. J Cell Mol Med. 2024;28(19): e18591.
15. Yang X, Sun J, Jin B, Lu Y, Cheng J, Jiang J, Zhao Q, Shuai J. Multi-task aquatic toxicity prediction model based on multi-level features fusion. J Adv Res. 2024;68:477–89.
16. Yuan P, Huang PC, Martin TK, Chappell TM, Kolomiets MV. Duplicated copy number variant of the maize 9-lipoxygenase ZmLOX5 improves 9,10-KODA-mediated resistance to fall armyworms. Genes. 2024;15(4):401.
17. Ping Y, Zhou Y, Hu J, Pang L, Xu C, Xiao Y. Dissecting the functional mechanisms of somatic copy-number alterations based on dysregulated ceRNA networks across cancers. Mol Ther Nucleic Acids. 2020;21:464–79.
18. Moore JL, Davies AR, Santaolalla A, Van Hemelrijck M, Maisey N, Lagergren J, Gossage JA, Kelly M, Baker CR: Clinical relevance of the tumor location-modified laurén classification system for gastric cancer in a western population. (1534–4681 (Electronic)).
19. Mermel CH, Schumacher SE, Hill B, Meyerson ML, Beroukhim R, Getz G. GISTIC2.0 facilitates sensitive and confident localization of the targets of focal somatic copy-number alteration in human cancers. Genome Biol. 2011;12(4):R41.
20. Li JH, Liu S, Zhou H, Qu LH, Yang JH. starBase v2.0: decoding miRNA-ceRNA, miRNA-ncRNA and protein-RNA interaction networks from large-scale CLIP-Seq data. Nucleic Acids Res. 2014;42:D92-97.
21. Das S, Ghosal S, Sen R, Chakrabarti J. lnCeDB: database of human long noncoding RNA acting as competing endogenous RNA. PLoS ONE. 2014;9(6): e98965.
22. Xu G, Wang H, Zhuang Y, Lin Q, Li Y, Cai Z, Lin G, Liu W. Identification of a ceRNA network driven by copy number variations in esophageal cancer. J Nippon Med School. 2023;90(6):426–38.
23. Ding J, Huang M, Huang B, Peng X, Wu G, Peng C, Zhang H, Mao C, Wu X. Identification of a dysregulated ceRNA network modulated by copy number variation-driven lncRNAs in lung squamous cell carcinoma. Environ Mol Mutagen. 2022;63(7):351–61.
24. Wang P, Li X, Gao Y, Guo Q, Ning S, Zhang Y, Shang S, Wang J, Wang Y, Zhi H, et al. LnCeVar: a comprehensive database of genomic variations that disturb ceRNA network regulation. Nucleic Acids Res. 2020;48(D1):D111-d117.
25. Kim SY, Park C, Kim HJ, Park J, Hwang J, Kim JI, Choi MG, Kim S, Kim KM, Kang MS. Deregulation of immune response genes in patients with Epstein-Barr virus-associated gastric cancer and outcomes. Gastroenterology. 2015;148(1):137-147.e139.
26. Oh SC, Sohn BH, Cheong JH, Kim SB, Lee JE, Park KC, Lee SH, Park JL, Park YY, Lee HS, et al. Clinical and genomic landscape of gastric cancer with a mesenchymal phenotype. Nat Commun. 2018;9(1):1777.
27. Cristescu R, Lee J, Nebozhyn M, Kim KM, Ting JC, Wong SS, Liu J, Yue YG, Wang J, Yu K, et al. Molecular analysis of gastric cancer identifies subtypes associated with distinct clinical outcomes. Nat Med. 2015;21(5):449–56.
28. Rappaport N, Fishilevich S, Nudel R, Twik M, Belinky F, Plaschkes I, Stein TI, Cohen D, Oz-Levi D, Safran M, et al. Rational confederation of genes and diseases: NGS interpretation via GeneCards, MalaCards and VarElect. Biomed Eng Online. 2017;16(Suppl 1):72.
29. Kim SK, Kim HJ, Park JL, Heo H, Kim SY, Lee SI, Song KS, Kim WH, Kim YS. Identification of a molecular signature of prognostic subtypes in diffuse-type gastric cancer. Gastric Cancer. 2020;23(3):473–82.
30. Szklarczyk D, Gable AL, Nastou KC, Lyon D, Kirsch R, Pyysalo S, Doncheva NT, Legeay M, Fang T, Bork P, et al. The STRING database in 2021: customizable protein-protein networks, and functional characterization of user-uploaded gene/measurement sets. Nucleic Acids Res. 2021;49(D1):D605-d612.

31. Liberzon A, Birger C, Thorvaldsdóttir H, Ghandi M, Mesirov JP, Tamayo P. The molecular signatures database (MSigDB) hallmark gene set collection. Cell Syst. 2015;1(6):417–25.
32. Jiang P, Gu S, Pan D, Fu J, Sahu A, Hu X, Li Z, Traugh N, Bu X, Li B, et al. Signatures of T cell dysfunction and exclusion predict cancer immunotherapy response. Nat Med. 2018;24(10):1550–8.
33. Kumar V, Ramnarayanan K, Sundar R, Padmanabhan N, Srivastava S, Koiwa M, Yasuda T, Koh V, Huang KK, Tay ST, et al. Single-cell atlas of lineage states, tumor microenvironment, and subtype-specific expression programs in gastric cancer. Cancer Discov. 2022;12(3):670–91.
34. Hu C, Li T, Xu Y, Zhang X, Li F, Bai J, Chen J, Jiang W, Yang K, Ou Q, et al. Cell marker 2.0: an updated database of manually curated cell markers in human/mouse and web tools based on scRNA-seq data. Nucleic Acids Res. 2023;51(1):870–6.
35. Sa JK, Hong JY, Lee IK, Kim JS, Sim MH, Kim HJ, An JY, Sohn TS, Lee JH, Bae JM, et al. Comprehensive pharmacogenomic characterization of gastric cancer. Genome medicine. 2020;12(1):17.
36. Saul D, Kosinsky RL, Atkinson EJ, Doolittle ML, Zhang X, LeBrasseur NK, Pignolo RJ, Robbins PD, Niedernhofer LJ, Ikeno Y, et al. A new gene set identifies senescent cells and predicts senescence-associated pathways across tissues. Nat Commun. 2022;13(1):4827.
37. Antonova E, Hambikova A, Shcherbakov D, Sukhov V, Vysochanskaya S, Fadeeva I, Gorshenin D, Sidorova E, Kashutina M, Zhdanova A, et al. Determination of common microRNA biomarker candidates in stage IV melanoma patients and a human melanoma cell line: a potential anti-melanoma agent screening model. Int J Mol Sci. 2023;24(11):9160.
38. Long Y, Wu J, Shen Y, Gan C, Zhang C, Wang G, Jing J, Zhang C, Pan W. CAPG is a novel biomarker for early gastric cancer and is involved in the Wnt/β-catenin signaling pathway. Cell death discovery. 2024;10(1):15.
39. Liu D, Yang F, Zhang T, Mao R. Leveraging a cuproptosis-based signature to predict the prognosis and drug sensitivity of cutaneous melanoma. J Transl Med. 2023;21(1):57.
40. Newman AM, Liu CL, Green MR, Gentles AJ, Feng W, Xu Y, Hoang CD, Diehn M, Alizadeh AA. Robust enumeration of cell subsets from tissue expression profiles. Nat Methods. 2015;12(5):453–7.
41. Roelands J, Kuppen PJK, Ahmed EI, Mall R, Masoodi T, Singh P, Monaco G, Raynaud C, de Miranda N, Ferraro L, et al. An integrated tumor, immune and microbiome atlas of colon cancer. Nat Med. 2023;29(5):1273–86.
42. Kriegsman BA, Vangala P, Chen BJ, Meraner P, Brass AL, Garber M, Rock KL. Frequent loss of IRF2 in cancers leads to immune evasion through decreased MHC class i antigen presentation and increased PD-L1 expression. J Immunol. 2019;203(7):1999–2010.
43. Lai J, Tong C, Chien JR. Ovarian serous carcinoma with a novel HSP90AB1 mutation in a patient with synchronous primary fallopian tube serous carcinoma. Anticancer Res. 2021;41(9):4417–22.
44. Mandula JK, Chang S, Mohamed E, Jimenez R, Sierra-Mondragon RA, Chang DC, Obermayer AN, Moran-Segura CM, Das S, Vazquez-Martinez JA, et al. Ablation of the endoplasmic reticulum stress kinase PERK induces paraptosis and type I interferon to promote anti-tumor T cell responses. Cancer Cell. 2022;40(10):1145-1160.e1149.
45. Gao X, Zuo S. Immune landscape and immunotherapy of hepatocellular carcinoma: focus on innate and adaptive immune cells. Clin Exp Med. 2023;23(6):1881–99.
46. Fu Y, Tao J, Liu T, Liu Y, Qiu J, Su D, Wang R, Luo W, Cao Z, Weng G, et al. Unbiasedly decoding the tumor microenvironment with single-cell multiomics analysis in pancreatic cancer. Mol Cancer. 2024;23(1):140.
47. Mao X, Tang X, Ye J, Xu S, Wang Y, Liu X, Wu Q, Lin X, Zhang M, Liu J, et al. Multi-omics profiling reveal cells with novel oncogenic cluster, TRAP1(low)/CAMSAP3(low), emerge more aggressive behavior and poor-prognosis in early-stage endometrial cancer. Mol Cancer. 2024;23(1):127.
48. Zhou Y, Mo S, Cui H, Sun R, Zhang W, Zhuang X, Xu E, Li H, Cheng Y, Meng Y, et al. Immune-tumor interaction dictates spatially directed evolution of esophageal squamous cell carcinoma. Natl Sci Rev. 2024;11(5):nwae150.
49. Ogbeide S, Giannese F, Mincarelli L, Macaulay IC. Into the multiverse: advances in single-cell multiomic profiling. Trends Genet. 2022;38(8):831–43.
50. Gohil SH, Iorgulescu JB, Braun DA, Keskin DB, Livak KJ. Applying high-dimensional single-cell technologies to the analysis of cancer immunotherapy. Nat Rev Clin Oncol. 2021;18(4):244–56.

Publisher's Note