## SYSTEMATIC REVIEW

# Histological and mutational profile of diffuse gastric cancer: current knowledge and future challenges

José Garcia-Pelaez[1,2,3] (iD), Rita Barbosa-Matos[1,2,4] (iD), Irene Gullo[1,2,5,6] (iD), Fátima Carneiro[1,2,5,6] (iD) and Carla Oliveira[1,2,5] (iD)

1  i3S – Instituto de Investigação e Inovação em Saúde da Universidade do Porto,Portugal
2  IPATIMUP – Institute of Molecular Pathology and Immunology, University of Porto, Portugal
3  Doctoral Programme on Biomedicine, Faculty of Medicine, University of Porto, Portugal
4  Doctoral Programme on Cellular and Molecular Biotechnology Applied to Health Sciences (BiotechHealth), ICBAS – Institute of Biomedical Sciences Abel Salazar, University of Porto, Portugal
5  Department of Pathology, FMUP - Faculty of Medicine of the University of Porto, Portugal
6  Department of Pathology, CHUSJ – Centro Hospitalar Universitário São João, Porto, Portugal

Gastric cancer (GC) pathogenesis is complex and heterogeneous, reflecting morphological, molecular and genetic diversity. Diffuse gastric cancer (DGC) and intestinal gastric cancer (IGC) are the major histological types. GC may be sporadic or hereditary; sporadic GC is related to environmental and genetic low-risk factors and hereditary GC is caused by inherited high-risk mutations, so far identified only for the diffuse histotype. DGC phenotypic heterogeneity challenges the current understanding of molecular mechanisms underlying carcinogenesis. The definition of a DGC-specific mutational profile remains controversial, possibly reflecting the heterogeneity of DGC-related histological subtypes [signet-ring cell carcinoma (SRCC) and poorly cohesive carcinoma not otherwise specified (PCC-NOS)]. Indeed, DGC and DGC-related subtypes may present specific mutational profiles underlying the particularly aggressive behaviour and dismal prognosis of DGC vs IGC and PCC-NOS vs SRCC. In this systematic review, we revised the histological presentations, molecular classifications and approved therapies for gastric cancer, with a focus on DGC. We then analysed results from the most relevant studies, reporting mutational analysis data specifying mutational frequencies, and their relationship with DGC and IGC histological types, and with specific DGC subtypes (SRCC and PCC-NOS). We aimed at identifying histology-associated mutational profiles with an emphasis in DGC and its subtypes (DGC vs IGC; sporadic vs hereditary DGC; and SRCC vs PCC-NOS). We further used these mutational profiles to identify the most commonly affected molecular pathways and biological functions, and explored the clinical trials directed specifically to patients with DGC. This systematic analysis is expected to expose a DGC-specific molecular profile and shed light into potential targets for therapeutic intervention, which are currently missing.

# 1. Introduction – a general overview of gastric cancer morphology and mutational landscape

Gastric cancer (GC) is very heterogeneous from the morphological, genetic and molecular standpoints. The molecular heterogeneity is mirrored in the great variability of morphological phenotypes [1]. In this work, we revised the histological features, molecular classifications and approved therapies for GC, with a focus on DGC. We then explored a set of studies that reported mutational analysis data specifying mutational frequencies and their relationship with DGC and IGC histological types, as well as with specific subtypes of DGC, namely SRCC and poorly cohesive carcinoma not otherwise specified (PCC-NOS), to fill a gap in the current knowledge.

## 1.1. Main histopathological classifications of gastric cancer

A large number of histopathological classifications have been proposed over time [2–8]. The most commonly used gastric cancer classification is, in the Western countries, the one proposed by Laurén [3] and the World Health Organization (WHO) classification [8]. In Eastern countries, the most commonly used gastric cancer classification is the one issued by the Japanese Gastric Cancer Association (JGCA) [7], which is very similar to the WHO classification, 5th edition [9].

The Laurén classification [3] distinguishes two major types, intestinal gastric cancer (IGC) and diffuse gastric cancer (DGC). The former is constituted by tubular or papillary structures, while the latter is characterized by poorly cohesive and infiltrative tumour cells that may or may not have a signet-ring cell (SRC) morphology. Tumours presenting both intestinal and diffuse components are termed mixed carcinomas. Other carcinomas that do not fit in one of these subtypes are placed in the indeterminate category. Despite dating back to 1965, Laurén classification is still widely accepted and used, as it distinguishes subtypes with distinct epidemiologic settings, clinicopathologic profiles and biological behaviours. These subtypes also correspond to distinct tumour spreading patterns: while IGC tends to metastasize haematogenously to the liver, DGC usually disseminates through peritoneal surfaces. Mixed gastric cancer shows a poorer prognosis compared with IGC or DGC types [10,11] and a dual metastatic pattern (haematogenous metastases and peritoneal dissemination with lymph node metastases) [12], probably because of the cumulative adverse effect of the two components within a single tumour.

The classification of the World Health Organization (WHO), 5th edition (2019), [9] recognizes five major types of gastric adenocarcinoma (tubular, papillary, mucinous, poorly cohesive and mixed), as well as other rarer subtypes (gastric carcinoma with lymphoid stroma, hepatoid adenocarcinoma and related entities, micropapillary adenocarcinoma, gastric adenocarcinoma of fundic gland type, etc.).

## 1.2. The morphological heterogeneity of DGC

The DGC subtype in Laurén classification roughly corresponds to the poorly cohesive carcinomas (PCC) in the 2019 WHO classification [8].

DGC/PCC may show great morphological variability, between different tumours from distinct patients, but also within different areas of the same tumour (intertumour and intratumour morphological heterogeneity). The classical picture of DGC/PCC displays a tumour composed of poorly cohesive signet-ring cells (SRCs) (Fig. 1, left panel). SRCs are defined by the presence of an abundant mucin vacuole filling the cytoplasm and pushing the nucleus to the cell periphery [13]. Especially in early lesions (i.e., GCs limited to the mucosa and submucosa), different types of SRCs may be observed, which are distinguished by the nature of the cytoplasmic mucin [14,15]. The description of different SRC types may be considered merely
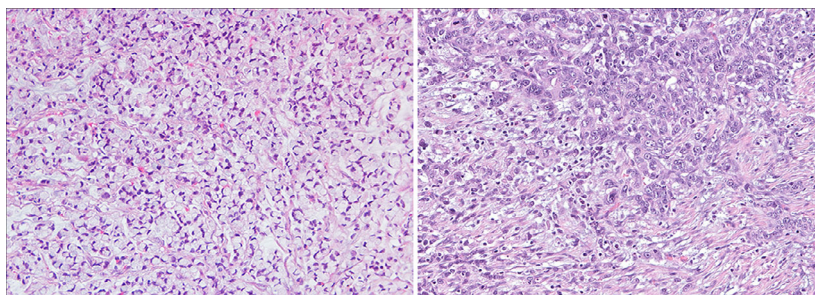


**Fig. 1.** Representative Images of a SRCC case (left) and a PCC-NOS case (right) highlighting the heterogeneity present in sporadic DGC.

academic, as it does not appear to influence the tumour prognosis. DGC/PCC may also show cells that do not present classic SRC morphology. In particular, advanced lesions infiltrating the gastric wall may show pleomorphic, bizarre and diffusely infiltrative neoplastic cells, sometimes with a lymphohistiocytic or sarcomatoid appearance, the latter representing the surrogate of epithelial-to-mesenchymal transition activation. The recognition of this subtype is of utmost importance, as it presents a poorer prognosis when compared to pure SRC carcinomas (SRCCs) [16]. Consistent with this observation, the new 2019 WHO classification stresses the importance of evaluating the morphological heterogeneity of DGC and distinguishing different PCC subtypes, based on the presence and quantity of SRCs [8]. Poorly cohesive carcinomas of the SRC type (SRCC) are composed predominantly or exclusively of SRCs and show a better prognosis, while non-SRC type, that is PCC-NOS, is composed (or show a predominant component) of poorly cohesive and cells, without a classic SRC morphology (Fig. 1, right panel).

## 1.3. Main molecular classifications of gastric cancer

The molecular and genomic analysis of large GC series has identified a large variety of oncogene activating mutations, tumour suppressor gene (TSG)-inactivating alterations, gene fusions, somatic copy number alterations (sCNAs) and other structural variations, as well as epigenetic and transcriptional changes [17]. Activating mutations in oncogenes such as *KRAS*, *BRAF* and *PIK3CA* have been reported in GC through single-gene analysis at least three decades ago [18–21] and remain important in GC using genome-wide strategies [22–24]. *TP53* is the most frequently mutated TSG in GC, often occurring in more than 50% of the cases, regardless of the experimental approach used, either target gene or genome-wide [22,25–27]. Somatic inactivating mutations in *CDH1* gene (a relevant TSG in GC) are known since the early 90's and remain a hallmark of the diffuse histological type [22,28–34].

Genome-wide strategies allowed to overcome the limitations of single-gene targeted approaches. Through next-generation sequencing (NGS) technologies, the knowledge on GC mutational burden and molecular signatures increased which culminated in the development of GC molecular classifications. These classifications based on genomics, transcriptomics and/or epigenomic profiles can be associated with the main GC histotypes (Table 1).

The first GC molecular classification was based exclusively on gene expression profiling and directly addressed differences in GC main histological types [35]. The second, although driven by gene expression profiling, was supported by gene-specific mutational analysis, chromosomal instability and DNA methylation analysis [36]. These studies also presented patients' response to treatment and disease outcomes according to molecular classification. In 2014, a game-changer GC molecular classification was published by The Cancer Genome Atlas (TCGA) consortium [22], based on a deep and integrative molecular profiling with array-based somatic copy number analysis, whole-exome sequencing, array-based DNA methylation profiling, messenger RNA sequencing (mRNA), microRNA (miRNA) sequencing and reverse-phase protein array (RPPA), microsatellite instability (MSI) testing and low-pass whole-genome sequencing. In 2015, the Asian Cancer Research Group (ACRG) generated comprehensive genomic and transcriptomic data sets from 300 GC cases, mainly from East Asia, producing the first GC classification based on integrated molecular data analysis and clinical outcomes [37]. In the two latter publications, there was an attempt to correlate the molecular subtypes with GC main histotypes, which did not provide a straightforward overlap (Table 1). Nevertheless, there is some correspondence between the two latter classifications, namely regarding tumours bearing mismatch repair deficiency (MSI), and those dominated by E-cadherin deficiency, genomic stability and epithelial-to-mesenchymal transition features (GS for TCGA and MSS/EMT for ACRG). The greater differences between TCGA and ACRG classifications relate to the lack of an EBV subgroup in the ACRG classification, which is molecularly defined in the TCGA study, and the heterogeneity of the CIN subgroup in TCGA, which becomes divided into two distinct subgroups in the ACRG study (Table 1).

## 1.4. Cost-effective strategies for molecular classification

Given the complexity and high cost of the experimental approaches needed to stratify GC according to TCGA and ACRG classifications, several studies proposed less costly approaches to reach similar endpoints (Fig. 2) [38–48].

In 2016, Setia *et al.* used 14 biomarkers in a GC cohort from the United States and Ahn *et al.* studied protein or mRNA expression of MLH1, E-cadherin, p53 and EBV in a South Korean population to derive five GC molecular subgroups and reproducing a combination of TCGA and ACRG classifications [38,39]. In 2016, *in situ* hybridization (ISH) or immunohistochemistry (IHC) were used to study MLH1, p53 and

**Table 1.** Overview of gastric cancer molecular classifications and relationship with main Laurén histotypes. ACRG, Asian Cancer Research Group; CIMP, CpG island methylation phenotype; CIN, chromosomal instability; CNA, copy number alteration; CSC, cancer stem cell; EBV, Epstein–Barr virus; EMT, epithelial-to-mesenchymal transition; GC, gastric cancer; GS, genomically stable; MSI, microsatellite instability; MSS, microsatellite stable; TCGA, The Cancer Genome Atlas; TKR, tyrosine kinase receptors. Lauren histological subtypes are highlighted in green (diffuse GC) and light blue (intestinal GC).

| Reference (sample size) | Methodology | Main features of molecular subtypes and relationship with Laurén histopathological classification | | | |
|---|---|---|---|---|---|
| [35] (n = 270) | Gene expression profiling | **G-DIF** (44%)<br>Cell proliferation<br>Fatty acid metabolism<br>Diffuse GC | **G-INT** (56%)<br>Cell adhesion<br>Carbohydrate and protein metabolism<br>Intestinal GC | | |
| [36] (n = 248) | Gene expression profiling, *TP53* mutation, CNA and ;DNA methylation analysis | **MESENCHYMAL**<br>Low *TP53* mutations<br>Low E-cadherin mRNA<br>CSC/EMT proprieties<br>Diffuse GC | **PROLIFERATIVE**<br>High *TP53* mutations<br>Genomic instability<br>Oncogene amplification<br>Intestinal GC (intestinal phenotype) | **METABOLIC**<br>Low *TP53* mutations<br>Normal gastric mucosa gene expression<br>Intestinal GC (gastric phenotype) | |
| [22] (n = 295) – TCGA cohort | Array-based somatic copy number analysis, whole-exome seq, array-based DNA methylation profiling, mRNA seq, microRNA seq and reverse-phase protein array, MSI testing. | **EBV** (9%)<br>EBV-CIMP<br>*CDKN2A* silencing<br>*PIK3A* mutations<br>*PD-L1/2* amplification<br>*JAK2* amplification | **MSI** (22%)<br>Gastric-CIMP<br>*MLH1* silencing<br>*PIK3A* mutations<br>*HER2/3* mutations<br>*EGFR* mutations<br>Intestinal GC | **GS** (20%)<br>*CDH1* mutations<br>*RHOA* mutations<br>*CLDN18-ARHGAP* fusion (RhoA-GTPase)<br>Diffuse GC | **CIN** (50%)<br>High *TP53* mutations<br>*TKR-RAS* amplification<br>Amplification of cell-cycle mediators<br>Intestinal GC |
| [37] (n = 251) – ACRG cohort | Whole-genome seq, gene expression profiling, copy number analysis, targeted re-sequencing. | **MSI** (23%)<br>*MLH1* loss<br>Hypermutation (KRAS, ARID1A, PIK3A)<br>Intestinal GC | **MSS/EMT** (15%)<br>*CDH1* loss<br>Diffuse GC<br>EBV+ cases included in MSS/TP53+ | **MSS/TP53-** (36%)<br>High *TP53* mutations<br>Genomic instability<br>Oncogene amplification<br>Intestinal GC | **MSS/TP53+** (26%)<br>Low *TP53* mutations<br>Intestinal GC |

**Algorithm strategy and molecular groups.**

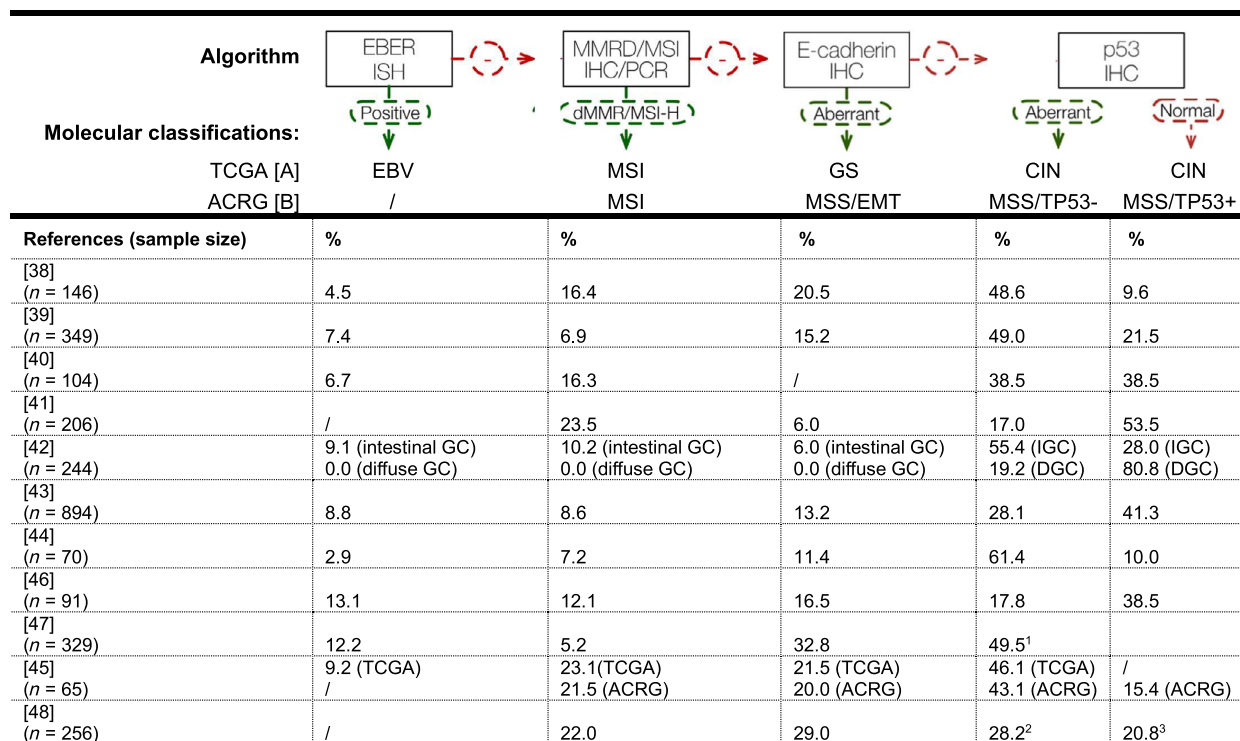| | EBER ISH (Positive) | MMRD/MSI IHC/PCR (dMMR/MSI-H) | E-cadherin IHC (Aberrant) | p53 IHC (Aberrant) | p53 IHC (Normal) |
|---|---|---|---|---|---|
| **TCGA [A]** | EBV | MSI | GS | CIN | CIN |
| **ACRG [B]** | / | MSI | MSS/EMT | MSS/TP53- | MSS/TP53+ |
| **References (sample size)** | % | % | % | % | % |
| [38] (*n* = 146) | 4.5 | 16.4 | 20.5 | 48.6 | 9.6 |
| [39] (*n* = 349) | 7.4 | 6.9 | 15.2 | 49.0 | 21.5 |
| [40] (*n* = 104) | 6.7 | 16.3 | / | 38.5 | 38.5 |
| [41] (*n* = 206) | / | 23.5 | 6.0 | 17.0 | 53.5 |
| [42] (*n* = 244) | 9.1 (intestinal GC) 0.0 (diffuse GC) | 10.2 (intestinal GC) 0.0 (diffuse GC) | 6.0 (intestinal GC) 0.0 (diffuse GC) | 55.4 (IGC) 19.2 (DGC) | 28.0 (IGC) 80.8 (DGC) |
| [43] (*n* = 894) | 8.8 | 8.6 | 13.2 | 28.1 | 41.3 |
| [44] (*n* = 70) | 2.9 | 7.2 | 11.4 | 61.4 | 10.0 |
| [46] (*n* = 91) | 13.1 | 12.1 | 16.5 | 17.8 | 38.5 |
| [47] (*n* = 329) | 12.2 | 5.2 | 32.8 | 49.5[1] | |
| [45] (*n* = 65) | 9.2 (TCGA) / | 23.1 (TCGA) 21.5 (ACRG) | 21.5 (TCGA) 20.0 (ACRG) | 46.1 (TCGA) 43.1 (ACRG) | / 15.4 (ACRG) |
| [48] (*n* = 256) | / | 22.0 | 29.0 | 28.2[2] | 20.8[3] |

**Fig. 2.** Cost-effective strategies for molecular classification. [A] TCGA, The Cancer Genome Atlas [22]; [B] ACRG, Asian Cancer Research Group [37]; CIN, chromosomal instability; EBER ISH, EBV-encoded small RNA *in situ* hybridization; EBV, Epstein–Barr virus; EMT, epithelial-to-mesenchymal transition; GC, gastric cancer; GS, genomically stable; IHC, immunohistochemistry; MMRD, mismatch repair deficiency; MSI, microsatellite instability; MSS, microsatellite stability; PCR, polymerase chain reaction. [1]CIN subtype defined by the presence of aneuploidy, using DNA flow cytometry; [2]MSS/TP53- subtype defined by low p21 protein expression; [3]MSS/TP53+ subtype defined by high p21 protein expression.

EBV, but excluded the analysis of MSS/EMT (ACRG) or GS (TCGA) [40]. In 2018, GC patients were classified using IHC based on molecular biomarkers to derive four molecular subgroups, which excluded EBV, and supported this classification demonstrating a significant and independent correlation between IHC subtype and overall survival, clinical outcome, SRC phenotype, tumour grade and vessel invasion [41].

In parallel, in 2018, ISH and IHC were performed in a cohort mainly composed of IGC to refine previous GC classifications [42]. In 2019, 894 consecutive GC Korean patients were used to validate previous classifications, with IHC- and PCR-based approaches and confirming the association between molecular subtypes and patients' overall survival [43]. More recently, in 2020, an immunophenotypic classification, based on IHC and ISH, was used to evaluate MLH1, p53, HER2, E-cadherin and EBV in whole sections from surgical GC specimens [44]. Another similar approach was used to apply the IHC/ISH molecular algorithm to a cohort of GC patients from Chile [46]. In this

study, the authors used targeted NGS sequencing to characterize the four molecular subtypes, identifying *FGFR2* and *KRAS* gene amplification as potential actionable targets in the EMT-like subgroup [46]. Wang *et al.* [45] studied a GC Chinese cohort to conclude that it would be possible to reproduce the TCGA and ACRG molecular classifications only by using IHC. In the two most recent cost-effective strategies for GC molecular classification, Zhao *et al.* and Tsai *et al.* used alternative surrogate biomarkers to define MSS/TP53 subgroups and CIN subtype, based on p21 protein expression and DNA content assessed by DNA flow cytometry, respectively [47,48].

## 1.5. Clinical outcomes, molecular features, prognosis, and approved and promising therapies for GC molecular subgroups

The research effort that followed the publication of the TCGA and ACRG molecular classifications led not only to the development of cost-effective strategies

to stratify GC patients according to these classifications (Fig. 2), but also to validate the molecular components from different molecular subgroups. This further allowed correlating molecular subgroups or alterations with clinical outcomes and response to therapy (Table 2). In Table 2, the possible correlations between TCGA and ACRG subgroups are presented, supported by shared molecular features. This information allows rethinking the application of former approved therapies and novel promising ones.

One of the most evident gaps in GC therapy is the lack of targeted therapies for DGC patients, which mainly reside in the TCGA GS group and in the MSS/EMT ACRG group (Table 2). This group of patients, which represents nearly 20% of all GC and those with the poorest prognosis, is currently treated with conventional chemotherapy but responds poorly [49] (Table 2). Immune checkpoint inhibitors did not prove as an alternative in these patients, as MSS/EMT cases from the ACRG group also showed poor treatment response [37]. Promising alternative therapies such as anti-CLDN18.2 antibody, targeting the *CLDN18-ARHGAP26* gene fusion, which is enriched in DGC cases [22,50,51] and intraperitoneal administration of anticancer agents [52], are just emerging treatments for this molecular subgroup, but results are still pending (Table 2).

The remaining GC molecular subgroups are less problematic regarding prognosis and available approved therapies or emerging ones (Table 2). The subgroups EBV-positive from TCGA and MSI from TCGA and ACRG show the best prognosis and are more likely to benefit from immune checkpoints inhibitors. In fact, MSI tumours are already selected for anti-PD1 treatment [53–55]. Patients bearing EBV-positive tumours are part of the ACRG MSS/TP53+ subgroup, present intermediate prognosis and are also often PD-L1-positive, turning them also into good candidates for immune checkpoint inhibition. Indeed, a phase II study on pembrolizumab showed the ORR to be not only 85.7% in MSI-H gastric cancers, but also more important to be 100% in EBV-related GC [56]. The TCGA CIN and ACRG MSS/TP53-subgroups present intermediate prognosis and the best survival benefit from adjuvant chemotherapy [37]. Given the high frequency of oncogene amplifications, tumours from this subgroup may also be treated with the approved anti-HER2 therapy [57], and promising treatments include inhibitors of the RTK-RAS pathway, DNA repair molecules and ATR (Table 2).

Diffuse gastric cancer has long been recognized as an independent GC histopathological entity; however, along the time, DGC-associated subtypes have been recognized and deeply characterized at the morphological and molecular levels. Despite this recognition, the most

two comprehensive GC molecular classifications (TCGA and ACRG) have been blind to these subtypes, and DGC as a whole, mainly cluster in the TCGA GS group and in the MSS/EMT ACRG group. These two molecular subgroups, besides sharing a dominant GC histotype, also share mutations in DGC-associated genes. Despite the increasing knowledge about GC histopathological and molecular subtypes, the integration of these two aspects and its clinical utility remains poorly explored.

In summary, DGC and its subtypes remain a major problem in terms of prognosis, response to conventional chemotherapy and a still arid ground for novel therapeutic interventions.

## 1.6. Aim

After recognizing a knowledge gap regarding molecular features potentially useful for clinical outcome prediction and therapy design in DGC, our aim was to identify histology-associated mutational profiles for this GC histotype and its subtypes. We further used these mutational profiles to identify their most affected molecular pathways and biological functions, and explored the clinical trials specifically designed for DGC patients. This systematic analysis is expected to expose a DGC-specific molecular profile and shed light into potential targets for therapeutic intervention, which are currently missing.

## 2. Methodology

We collected available literature from 1998 to 2020, in order to get an overall view of the morphological and molecular features of sporadic GC and its molecular classifications. Peer-reviewed published articles were searched with the help of 'PubMed' database using the following keywords in different combinations: 'mutational landscape, mutational signature, gastric cancer, diffuse-type gastric cancer, poorly cohesive carcinoma and signet-ring cell carcinoma'. Only peer-reviewed published articles written in English language, irrespective of their study design and type (case reports, case series, original research articles, reviews, systematic reviews, meta-analyses, cohort studies, clinical trials), were consulted, and in the context of the present study, we selected only articles with relevant data and best fitting our aim.

### 2.1. Data collection for recurrent gene analysis

For the recurrent gene analysis, published studies or public cohorts reporting mutational frequencies and

**Table 2.** Representative genomic subtypes, common molecular features, clinical outcomes and corresponding therapeutic approaches. NA, not available; The Cancer Genome Atlas (TCGA); Asian Cancer Research Group (ACRG); Genomically stable (GS); Microsatellite Stable/Epithelial-Mesenchymal transition (MSS/EMT); Microsatellite instable (MSI); Helicobacter pylori (H. pylori); Vascular endothelial growth factor receptor 2 tyrosine kinase (VEGFR2); Programmed-cell-death-1-receptor (PD-1); Programmed death-ligand 1 (PD-L1); Receptor Tyrosine Kinase (RTK); Human epidermal growth factor receptor 2 (HER2); Epidermal growth factor receptor(EGFR); Fibroblast Growth Factor Receptor 2 (FGFR2); Poly (ADP-ribose) polymerase (PARP); ataxia telangiectasia and Rad-3 related protein kinase (ATR); References: [A] [22]; [B] [37]; [C] [49] [D] [83]; [E] [84]; [F] [85]; [G] [86]; [H] [87] (e-updated in 2019); [I] [52]; [J] [88]; [K] [89]; [L] [53]; [M] [90]; [N] [91]; [O] [92]; [P] [57]; [Q] [93]; [R] [56].

| TCGA molecular classification [A] | | | ACRG molecular classification [B] | | | Shared features and therapeutic approach | | |
|---|---|---|---|---|---|---|---|---|
| Subtype frequency main location predominant histology | Clinical outcome/ prognosis | Molecular features | Subtype frequency main location predominant histology | Clinical outcome/ prognosis | Molecular features | Shared molecular features | Approved therapeutic strategies | Promising targets/ strategies |
| GS 20% NA DGC | Diagnosed at early age. Worst overall survival. Less benefit from adjuvant chemotherapy [C]. Poor response to immune checkpoint inhibitors [D][E][F]. | High frequency of *CDH1* (37%) and *RHOA* (15%) mutations. *CLDN18-ARHGAP* fusion. Elevated expression of angiogenesis-related pathways [A]. | MSS/EMT 15% NA DGC | Poor prognosis and high recurrence frequency with tendency to develop at earlier age. Majority of cases diagnosed at stages III/IV [B]. | Loss of E-cadherin expression. Low mutation rate [B]. | GS and MSS/EMT *CDH1* and *RHOA* mutations | Standard chemotherapy: combinations of fluoropyrimidines and platinum derivatives [G][H]. | Cell junctions: CLDN18.2 antibody (phase III clinical trials: NCT03504397, NCT03653507); intraperitoneal administration of anticancer agents (phase III clinical trial: UMIN000005930 [I]; (phase II clinical trial NCT04034251); anti-VEGFR2 treatment using ramucirumab and apatinib [J]; [K] and other angiogenesis inhibitors (phase III clinical trial NCT02773524) [L][M] |
| MSI 22% antrum IGC | Diagnosed at older ages; associated with *H. pylori* infection and intestinal metaplasia [A]. Better prognosis | CpG island methylator phenotype (*MLH1* silencing): hypermutation phenotype frequently targeting genes such as *TP53, KRAS,* | MSI 23% antrum IGC | Good prognosis [B]. | *MLH1* gene silencing; commonly mutated genes: *ARID1A, MTOR, KRAS, PIK3CA, ALK, PTEN.* | MSI Hypermutated and hypermethylated tumours. Both enriched in IGCs. | Immune checkpoint inhibitors: Anti-PD-1 [L][O][P][R]. | Immune checkpoint inhibitors: PD-1/PD-L1 pathway: (phase III clinical trials: NCT02872116, NCT02746796, NCT03019588, |

**Table 2.** (Continued).

| TCGA molecular classification [A] | | | ACRG molecular classification [B] | | | Shared features and therapeutic approach | | |
|---|---|---|---|---|---|---|---|---|
| Subtype frequency main location predominant histology | Clinical outcome/ prognosis | Molecular features | Subtype frequency main location predominant histology | Clinical outcome/ prognosis | Molecular features | Shared molecular features | Approved therapeutic strategies | Promising targets/ strategies |
| | than the MSS subtype [C] | ARID1A, PIK3CA, ERBB3, PTEN, CHRD and HLA-B with nucleotide substitutions; and RNF43, B2 M and NF1 with insertions/ deletions [A][N]. | | | Overexpression of PD-L1; T-cell infiltration [B]. | | | NCT03615326, NCT03189719, NCT02625610, NCT04152889 [L][M]. |
| EBV + | | | MSS/TP53+ | | | EBV+ and MSS/ TP53+ | | |
| 9% Gastric fundus or body Gastric cancer with lymphoid stroma | Good prognosis [C]. | EBV infection; CpG island methylation; CCDKN2A silencing; JAK2, CD274, PDCD1LG2, ERBB2, PD-L1 and PD-L2 amplification; mutations in PIK3CA (80%), ARID1A (55%) and BCOR (23%) [A]. | 26% NA IGC | Intermediate prognosis [B]. | Mutations in ARID1A, PIK3CA, SMAD4, APC [B]. | EBV infection; PIK3CA mutations | Standard chemotherapy: combinations of fluoropyrimidines and platinum derivatives [G][H]. | Immune checkpoint inhibitors: PD-1/PD-L1 pathway: (phase III clinical trials NCT02872116, NCT02746796, NCT03019588, NCT03615326, NCT03189719, NCT02625610, NCT04152889) and (phase II clinical trials NCT03755440, NCT03257163 and NCT04202601) [L][M] [P][Q][R]. |

**Table 2.** (Continued).

| TCGA molecular classification [A] | | | ACRG molecular classification [B] | | | | Shared features and therapeutic approach | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Subtype frequency main location predominant histology | Clinical outcome/ prognosis | Molecular features | Subtype frequency main location predominant histology | Clinical outcome/ prognosis | Molecular features | Shared molecular features | Approved therapeutic strategies | Promising targets/ strategies |
| CIN | | | MSS/TP53- | | | CIN and MSS/TP53- | | |
| 50% Gastro-oesophageal junction/cardia IGC | Poorer overall survival than EBV subtype but better than GS. Benefit from adjuvant chemotherapy [C]. Poor response to immune checkpoint inhibitors [E][F]. | RTK-RAS pathway genes amplification namely HER2, VEGFA and FGFR2; TP53 (71%) mutations [A] [O]. | 36% NA IGC | Intermediate prognosis and recurrence rates compared with the other subtypes. Survival benefit with adjuvant chemotherapy [B]. | Genomic instability; TP53 mutations; recurrent focal amplification of RTKs and cell cycle modulators; ERBB2 (HER2) amplification[B]. | High frequency of TP53 mutations. Alterations in RTK-RAS pathway genes. Both enriched in IGCs | Standard chemotherapy: Combinations of fluoropyrimidines and platinum derivatives [G][H]. Trastuzumab for HER2-positive GCs [L][P]. | RTK-RAS pathway: EGFR, HER2 and HER4 inhibition (phase II/III clinical trial NCT03130790); EGFR overexpression (phase III clinical trial NCT01813253); FGFR2 amplification (phase III clinical trial NCT0369452 and phase II clinical trial NCT03694522); anti-VEGFR2 treatment using ramucirumab and apatinib [J][K], and other angiogenesis inhibitors (phase III clinical trial NCT02773524); DNA repair: PARP inhibitors (phase II clinical trials NCT03427814, NCT02678182) TP53 mutants: ATR inhibitor (phase II clinical trial NCT03641313) [L][M]. |

their relationship with GC histological types and subtypes of interest (IGC, DGC, SRCC and PCC-NOS) were searched for, selected upon compliance with the following criteria and refined by manual curation:

1 Studies from CBIOPORTAL (https://www.cbioportal.org/) encompassing 'gastric/ stomach adenocarcinoma' cancer NGS data;
2 'PubMed' deposited peer-reviewed articles, published between 2014 and 2020, reporting on data produced by targeted sequencing and/or whole-exome sequencing and/or whole-genome sequencing, and according to different MESH term combinations:

  a (mutational landscape OR mutational signature) AND (diffuse-type gastric carcinoma OR diffuse-type gastric cancer OR gastric adenocarcinoma OR gastric cancer OR poorly cohesive carcinoma) NOT review.
  b (genomic variation) AND (diffuse-type gastric cancer) NOT review.
  c (target sequencing OR whole-exome sequencing OR whole-genome sequencing) AND (diffuse-type gastric cancer) NOT review.

Manual curation was further performed on all publications selected with the above criteria, aiming at finding those presenting histological classification for each tumour, either according to Laurén [3] for comparison between DGC and IGC, or specifically mentioning comparisons between DGC subtypes, and not restricted to Laurén classification, namely if mentioning SRCC and PCC-NOS. Moreover, to be selected for the analysis, each publication should also specify either the set of mutated genes per case, or the gene mutation frequency among histological types or subtypes, as well as the total number of cases analysed within the materials (main or supplementary).

Whenever gene mutation frequencies per histological type or subtypes were not provided, their calculation was performed based on the formula:

$$\text{Mutated frequency of gene X} = \frac{\text{N}^\circ \text{of cases of a specific histological subtype mutated for gene X}}{\text{Total N}^\circ \text{of cases of a specific histological subtype}}$$

According to criteria 1, seven different entries were identified; from these, only three data sets were used due to the availability of Laurén histological classification and specific gene mutation frequencies. According to criteria 2a, there were 162 results from which five publications were selected after manual curation. Criteria 2b retrieved 11 results, from which two had already been found with previous criteria and eight did not comply with our criteria after manual curation; therefore, only one publication was selected. Criteria 2c retrieved 29 results, from which four had already been found with previous criteria and 24 did not comply with our criteria after manual curation. Accordingly, a single new publication was selected based on criteria 2c. Overall, 3 data sets and 7 original articles were selected (Fig. 3).

All remaining articles not fulfilling the above criteria were excluded.

## 2.2. Criteria for recurrent gene analysis

To identify relevant mutated genes, we established two selection steps: gene recurrence and gene classification.

### 2.2.1. Gene Recurrence

Only genes found to be mutated in at least three studies/resources in a frequency higher than 5% were selected for further analysis (Fig. 3, Analysis I). Due to the limited number of studies available for the analysis of SRCC and PCC-NOS, all genes with an available mutation frequency were used for this part of the study (Fig. 3, Analysis II).

### 2.2.2. Gene Classification

Mutated genes, for each histological type or subtype, were allocated to three classes: specific mutated genes, transversally mutated genes and enriched mutated genes.

Specific mutated genes were those exclusive of a single histological type or subtype.

Transversally mutated genes were those for which the difference between histological types or subtypes was, on average, lower than twofold. If the mutation frequency of a gene was reported in one study only, hampering average calculation, the absolute mutation frequency was used.
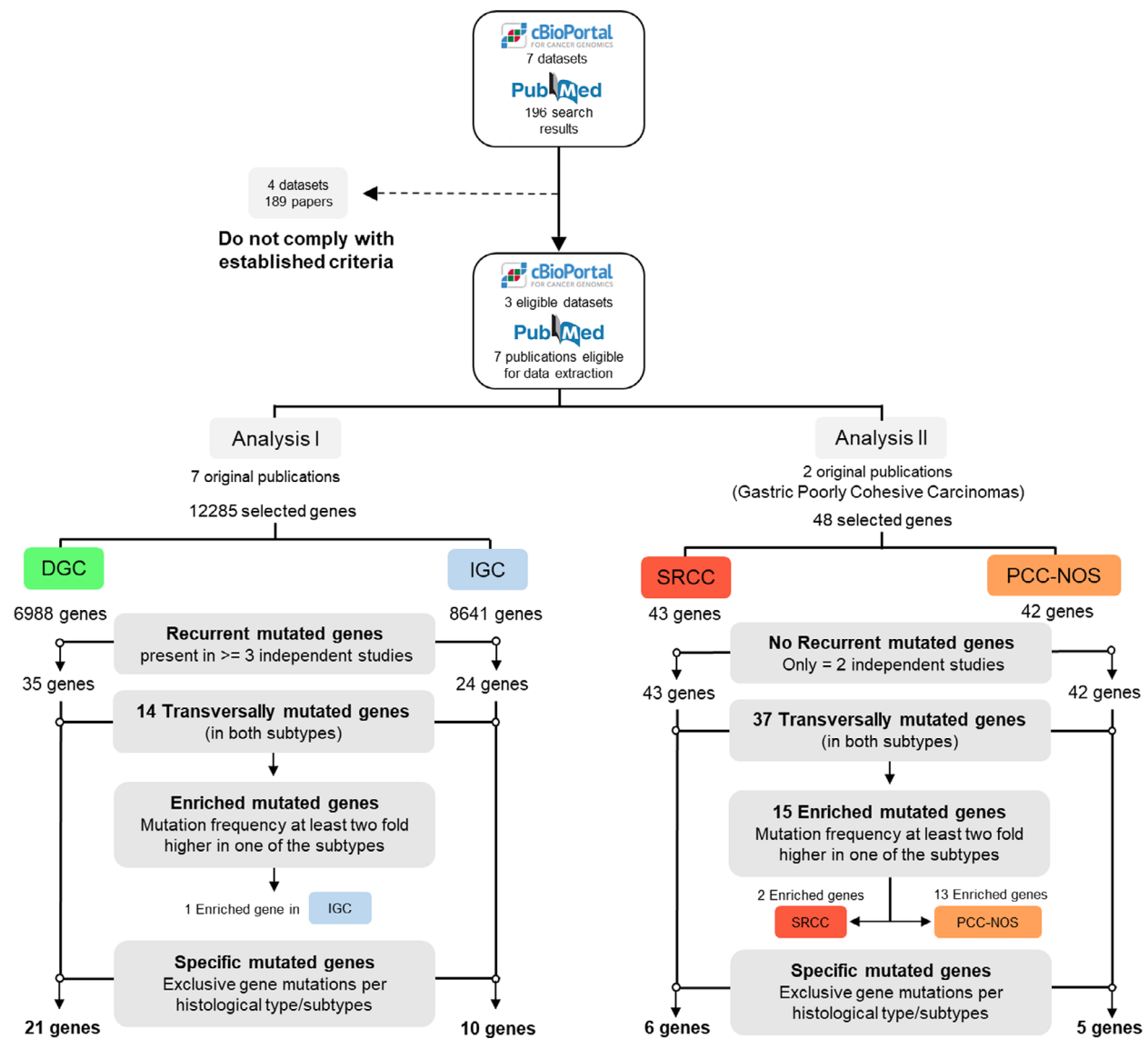
Enriched mutated genes were those for which difference between histological types or subtypes was, on average, higher than twofold. If the mutation frequency of a gene was reported in one study only, hampering average calculation, the absolute mutation frequency was used.

**Fig. 3.** Data collection and research strategy for recurrent gene analysis.

## 2.3. Identification of biological processes and molecular function associated with recurrent gene analysis

Specific mutated genes, transversally mutated genes and enriched mutated genes were used to feed Jvenn [58], an interactive diagram viewer, and allowed identifying genes overlapping between distinct GC histological types and subtypes through Venn Diagrams.

Recurrently mutated genes in DGC and IGC were compared to define the specific mutated genes for each histological type. The remaining DGC-associated genes were considered transversally DGC- and IGC-mutated

genes, as no enrichment was seen in any of them. For the SRCC vs PCC-NOS analysis, specific mutated genes per subtype were obtained from a single report.

Further, the output of the Venn diagrams was used to analyse Gene Ontology (GO) terms, biological processes and molecular functions associated with each gene set using Enrichr [59].

## 2.4. Search for clinical trials directed to DGC

Relevant clinical trials directed specifically to DGC patients were selected by searching the website www.clinicaltrials.gov (/clinicaltrials.gov/), and using as

query terms 'diffuse gastric cancer', which retrieved 24 trials; 'signet-ring cell carcinoma' with 19 trials; and 'gastric poorly cohesive carcinoma' with no results. After identifying only those associated with GC, 12 studies were selected. Only two clinical trials (NCT01285557 – DIGEST; NCT03977220 – NOR-DICA) were specifically directed to DGC patients, and one was directed to patients bearing SRCC (NCT01717924 – ADCI002 Study). Beyond these, and given the promising results recently obtained using anti-claudin 18.2 antibodies, we manually added the MONO study where almost 50% of patients enrolled had DGC (NCT01197885 – MONO) [60]. In total, 13 clinical trials were selected since they were directed to DGC patients.

For these, a deeper analysis on therapeutic strategy, candidate drugs and molecular targets was performed.

## 2.5. Study limitations

There are several limitations in this study, namely the discrepancies between mutational profiles and frequencies among cohorts, which may be due to several factors: the methodological approach applied by different research groups (sequencing methods and pipeline analyses); the heterogeneity of GC (intra- and inter-tumoral); and the sample size of the cohorts under analysis (specifically in the studies focused on SRCC and PCC-NOS). Another issue is the lack of diversity in the geographical origin of the cohorts used in the selected studies for analysis, which reflects data available in the literature. However, the reproducibility of findings across the various molecular classifications (e.g. TCGA and ACRG) indicates that relevant GC-associated molecular defects likely override population differences. Our analysis was restricted to mutational data, leaving behind transcriptomics and epigenetics, which can provide valuable data. Copy number variation analysis would have been a strength in this study; however, there are very few publications listing CNVs per case or their relative frequency to allow an accurate analysis. In addition, recent data suggested that copy number gains tended to occur more commonly in IGC than in DGC that was the focus of this work [61]. In addition, the comparison of sporadic and hereditary DGC suffered from scarce data published so far on somatic profiles in HDGC. Notwithstanding, we believe that the information herein provided constitutes a good basis for further studies aiming at a better understanding of the molecular pathogenesis of DGC and its clinical implications regarding prognosis and therapy.

## 3. Results

Strategies for gene mutation detection evolved from single-gene analysis to the analysis of multigene panels, whole exomes or even whole genomes. This created the opportunity to derive multigene profiles that are valuable to correlate molecular defects with specific histological subtypes.

### 3.1. DGC vs IGC: distinctive mutational profiles

The data compilation performed herein and depicted in Fig. 4 shows the mutational landscape across DGC cohorts retrieved from nine independent studies, mainly Asian, and obtained with different technical approaches (four targeted sequencing; four WES, two WGS) (Data S1, Table S1). We next retrieved data from six independent studies of IGC, mostly also of Asian origin, having a similar degree of detailed information, depicted in Fig. 5. The mutational landscape of IGC obtained with different technological approaches (three targeted sequencing; three WES, two WGS) (Data S1, Table S1).

To derive the mutational landscape from DGC, we performed a recurrent gene analysis, as described in the methodology section that provided three mutated gene sets, both for DGC and for IGC: specific mutated genes, transversally mutated genes and enriched mutated genes.

### 3.2. DGC-specific mutated genes

This analysis confirmed *CDH1* genetic variants as the hallmark of the DGC histological type [31,33,34], as *CDH1* is the only gene found to be mutated in all DGC data sets and in none of the IGC data sets (specific DGC-mutated gene) (Figs 4 and 5). *CDH1* variants were found on average in 25% of DGC cases, ranging from 12% to 38% among studies (Fig. 4). Besides *CDH1*, another 20 genes were found to be mutated specifically in DGC, in at least three studies (Fig. 4). *TTN* gene mutations were found on average in 40% of DGC cases, and described in three studies, making this gene the top-ranking specific DGC-mutated gene, while *RHOA* and *CTNNB1* were found to be mutated in seven and five studies, on average in 10% and 7% of the cases, respectively. All other specific DGC-mutated genes were found in three or four studies (Fig. 4). Seven of these were mutated on average in > 12% of DGC. *FAT3* (15%) and *KMT2C* (12%) have a previously reported association with DGC; *CSMD1* (21%), *OBSCN* (15%), *RYR2* (13%) and *TAF1* (12%) have

| Information source | TCGA Firehouse Legacy | CBIPORTAL OncoSG, 2018 | PMID: 29622765 | PMID: 31171626 | PMID: 24816253 | PMID: 24816255 | PMID: 28873240 | PMID: 25583476 | PMID: 31164161 | Average mutation frequency |
|---|---|---|---|---|---|---|---|---|---|---|
| Origin of samples | Several population | China | Korea | USA, China, Japan | Hong Kong | Tokyo | Korea | China | China | |
| Cohort Specification | DGC | DGC | DGC | DGC | DGC | DGC | Gastric Poorly Cohesive Carcinoma | DGC | DGC | |
| Number of samples | n = 72 | n = 34 | n = 23 | n = 21 | n = 29 | n = 87 | n = 91 | n = 155 | n = 50 | |
| Genes | Frequency of mutated samples per study | | | | | | | | | |
| **Specific DGC-mutated genes** | | | | | | | | | | |
| TTN | 39% | 41% | 39% | – | – | – | – | – | – | 40% |
| CSMD1 | 29% | 21% | 13% | – | – | – | – | – | – | 21% |
| OBSCN | 23% | 9% | 13% | – | – | – | – | – | – | 15% |
| CDH1 | 21% | 12% | 30% | 38% | 28% | 33% | 20% | 12% | 30% | 25% |
| RYR2 | 16% | 9% | 13% | – | – | – | – | – | – | 13% |
| KMT2C | 14% | 6% | – | 29% | – | 1% | – | – | – | 12% |
| MUC6 | 14% | 6% | – | – | 7% | – | – | – | – | 9% |
| FAT3 | 11% | 15% | 18% | – | – | – | – | – | – | 15% |
| APOB | 10% | 9% | – | 10% | – | – | – | – | – | 9% |
| CHL1 | 10% | 9% | 9% | – | – | – | – | – | – | 9% |
| SETBP1 | 10% | 9% | – | 5% | – | – | – | – | – | 8% |
| DOCK2 | 10% | 6% | 18% | – | – | – | – | – | – | 11% |
| THSD7B | 7% | 15% | – | – | 14% | – | – | – | – | 12% |
| RHOA | 7% | 12% | 9% | 0% | 10% | 25% | 17% | 3% | – | 10% |
| CACNA1E | 7% | 12% | 13% | – | – | – | – | – | – | 11% |
| ABCA13 | 7% | 9% | 13% | – | – | – | – | – | – | 10% |
| TAF1 | 7% | 9% | – | 19% | – | – | – | – | – | 12% |
| NCOR2 | 7% | 6% | 13% | – | – | – | – | – | – | 9% |
| TGFBR2 | 7% | 6% | 9% | – | – | – | – | – | – | 7% |
| CTNNB1 | 6% | 6% | – | – | – | 1% | 22% | 1% | – | 7% |
| ERBB3 | 6% | 6% | – | – | – | – | 11% | 2% | – | 6% |
| **Transversally DGC- and IGC-mutated genes** | | | | | | | | | | |
| TP53 | 37% | 41% | 18% | 57% | 45% | 34% | 41% | – | 54% | 41% |
| FAT4 | 23% | 15% | – | – | – | 5% | – | – | – | 14% |
| ARID1A | 23% | 9% | – | – | 10% | 15% | 20% | – | – | 15% |
| SPTA1 | 21% | 3% | – | 14% | 7% | – | – | – | – | 11% |
| PIK3CA | 19% | – | – | – | – | 5% | 22% | 10% | – | 14% |
| GLI3 | 14% | 9% | – | – | 14% | – | – | – | – | 12% |
| LRP2 | 9% | 9% | – | – | – | – | – | – | 6% | 8% |
| RIMS2 | 9% | 9% | 23% | – | 10% | – | – | – | – | 13% |
| DCLK1 | 9% | 6% | – | – | 10% | – | – | – | – | 8% |
| APC | 7% | 3% | – | – | – | 2% | 28% | 6% | 0% | 8% |
| KRAS | 6% | – | – | 10% | 3% | 1% | 12% | 3% | – | 6% |
| SMAD4 | 4% | 6% | – | 14% | 10% | – | 13% | 3% | 10% | 9% |
| ATM | 4% | 6% | – | 10% | – | – | 12% | – | 14% | 9% |
| PTEN | 4% | 3% | – | 19% | 0% | 3% | 17% | – | 10% | 8% |

**Fig. 4.** Mutational landscape of DGC subtype. Nine sources were used: two CBIOPORTAL data sets and seven references. The origin, number of samples, reported study identification and other cohorts' characteristics are indicated in the heading of the table. The mutated genes of each subtype were divided into two main classifications: transversally mutated and specific genes. The specific mutated genes for DGC were associated with the green colour and are listed in the upper part of the table. In the lower part of the table, the transversally mutated genes in both DGC and IGC are presented. Mutation frequency was represented in a colour scale. Green was appointed to genes with mutation frequencies between 0% and 9% (less frequently mutated). Yellow was used for mutation frequencies between 10% and 19%. Light orange to mutation frequencies between 20% and 39%. A variation between dark orange and red was associated with mutation frequencies increasing from 40% to a maximum of 57% (frequently mutated). The average mutation frequency calculated for each DGC recurrent gene is present in the right column, and a similar colour scale was applied to distinguish the genes with higher mutation frequencies.

| Information source | CBIOPORTAL, UHK, 2011 | CBIPORTAL OncoSG, 2018 | PMID: 24816253 | PMID: 24816255 | PMID: 31164161 | PMID: 25583476 | Average mutation frequency |
|---|---|---|---|---|---|---|---|
| **Origin of samples** | **Hong Kong** | **China** | **Hong Kong** | **Tokyo** | **China** | **China** | |
| **Cohort Specification** | IGC | IGC | IGC | IGC | IGC | IGC | |
| **Number of samples** | *n* = 18 | *n* = 75 | *n* = 41 | *n* = 51 | *n* = 47 | *n* = 139 | |
| **Genes** | Frequency of mutated samples per study | | | | | | |
| **Specific IGC-mutated genes** BRCA2 | 17% | 7% | – | – | 4% | 9% | 9% |
| ERBB2 | 11% | 12% | 12% | – | – | 6% | 10% |
| AKAP9 | 11% | 8% | – | – | 26% | – | 15% |
| DLC1 | 11% | 9% | – | 8% | – | – | 9% |
| MACF1 | 11% | 9% | – | 8% | – | – | 9% |
| RNF43 | 11% | 11% | 7% | – | – | – | 10% |
| NRG1 | 6% | 5% | – | – | – | 6% | 6% |
| CTNNA2 | 6% | 11% | 15% | – | – | – | 10% |
| GSR | 6% | 9% | – | – | 6% | – | 7% |
| PKHD1 | 6% | 12% | – | 6% | – | – | 8% |
| **Transversally DGC- and IGC-mutated genes** TP53 | 44% | 49% | 76% | 43% | 72% | – | **57%** |
| PTEN | 33% | 1% | 10% | 2% | 13% | – | 12% |
| ARID1A | 28% | 16% | 10% | 12% | – | – | 16% |
| SPTA1 | 22% | 15% | 12% | – | – | – | 16% |
| APC | 17% | 9% | – | 10% | 11% | 9% | 11% |
| PIK3CA | 17% | 7% | – | 14% | – | 11% | 12% |
| FAT4 | 17% | 8% | – | 6% | – | – | 10% |
| LRP2* | 17% | 13% | – | – | 21% | – | 17% |
| DCLK1 | 11% | 7% | 15% | – | – | – | 11% |
| GLI3 | 11% | 11% | 10% | – | – | – | 11% |
| KRAS | 6% | 9% | 12% | 12% | – | 7% | 9% |
| SMAD4 | 6% | 5% | 2% | – | 11% | 5% | 6% |
| ATM | 6% | 11% | – | – | 9% | – | 8% |
| RIMS2 | 6% | 13% | 17% | – | – | – | 12% |

**Fig. 5.** Mutational landscape of IGC subtype. Six sources were used: two CBIOPORTAL data sets and four publications. The origin, number of samples, reported study identification and other cohorts' characteristics are indicated in the heading of the table. The mutated genes of each subtype were divided into two main classifications: transversally mutated and specific genes. The specific mutated genes for IGC were associated with the blue colour and are listed in the upper part of the Table. In the lower part of the table are presented the transversally mutated genes in both DGC and IGC. Mutation frequency was represented in a colour scale. Green was appointed to genes with mutation frequencies between 0% and 9% (less frequently mutated). Yellow was used for mutation frequencies between 10% and 19% and light orange for mutation frequencies between 20% and 39%. A variation between dark orange and red was associated with mutation frequencies increasing from 40% to a maximum of 76% (frequently mutated). The average mutation frequency calculated for each IGC recurrent gene is present in the right column, and a similar colour scale was applied to distinguish the genes with higher mutation frequencies. * LPR2 is enriched in IGC.

been previously associated with GC but not DGC, and *THSD7B* (12%) with no previous association with GC.

There are other specific DGC-mutated genes depicted in Fig. 4 that present on average < 12% of mutations that we used for downstream analyses.

### 3.3. IGC-specific mutated genes

We next focused on the mutational landscape of IGC from six studies, to identify the set of specific IGC-

mutated genes (Fig. 5). No single gene was commonly mutated across all IGC data sets; however, *BRCA2* and *ERBB2* were specific IGC-mutated genes (9% and 10% on average, respectively), according to four out of six studies.

Additionally, *AKAP9*, *DLC1*, *MACF1*, *RNF43*, *NRG1*, *CTNNA2*, *GSR* and *PKHD1* were also identified as specific IGC-mutated genes. *AKAP9* (found in 3/6 studies) was the one with the highest average mutation frequency (15%).

## 3.4. Transversally DGC- and IGC- mutated genes

We then ranked the transversally DGC- and IGC-mutated genes across 3 studies or more (Figs 4 and 5). These were *TP53, FAT4, ARID1A, SPTA1, PIK3CA, GLI3, LRP2, RIMS2, DCLK1, APC, KRAS, SMAD4, ATM* and *PTEN* (Figs 4 and 5). From these 14 genes, a single gene was classified as enriched mutated gene in IGC (*LRP2*), as its mutation frequency was >twofold higher in IGC (17%) as compared to DGC (8%) (Figs 4 and 5). All remaining genes displayed similar mutation frequencies (<twofold difference between histotypes) in both DGC and IGC.

*TP53* and *ARID1A* genes were the most frequently transversally mutated genes across DGC (41% and 15%, respectively) and IGC (57% and 16%, respectively) data sets (Figs 4 and 5). Besides these two

important GC driver genes, *APC* and *PTEN* tumour suppressor driver genes, as well as *KRAS* and *PIK3CA* driver oncogenes [62], were also transversally mutated in both GC histotypes with frequencies below 15% across data sets (Figs 4 and 5).

## 3.5. DGC vs IGC: biological processes and molecular functions associated with mutational profiles

The comparison of molecular functions and biological terms derived from the mutational landscapes of specific DGC genes vs specific IGC genes (Fig. 6) shows that the recurrence and high mutation frequency of *CDH1* and *RHOA* associate DGC with cell adhesion and calcium dependence. On the other hand, specific IGC-mutated genes are mainly associated with tyrosine
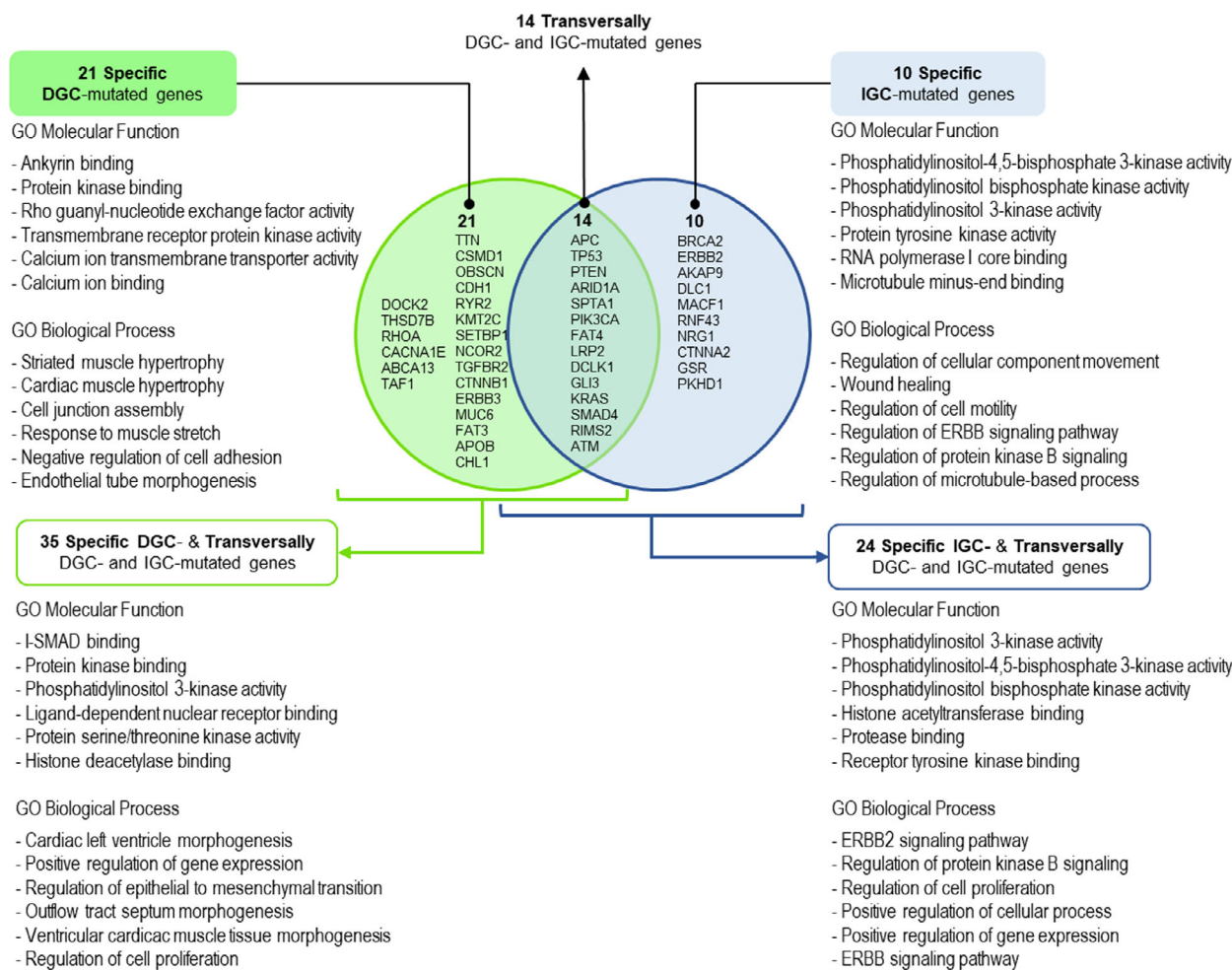


**Fig. 6.** Integrative analysis of specific- and transversally mutated genes in DGC or IGC. GO terms for molecular functions and biological processes are depicted for the specific genes lists and for the overall set of genes of each subtype (specific- and transversally mutated genes analysed together). The analysis generated using Enrichr and the associated GO terms number is available in Data S2.

kinase activity and related signalling pathways, encompassing some promising therapeutic targets (Table 2).

As tumour cells most likely bear histotype-specific mutations together with histotype nonspecific mutations, equivalently frequent, we combined the specific DGC-mutated gene set with the transversally DGC- and IGC-mutated gene set (all genes listed in Fig. 4) and compared the resulting gene list with its IGC counterpart (all genes in Fig. 5) regarding molecular functions and biological terms. This integrated analysis of specific and transversally mutated genes, which is depicted in Fig. 6, turned DGC closer to IGC, with proliferation and tyrosine kinase activity and associated signalling pathways appearing as an output, also in DGC. However, regulation of epithelial-to-mesenchymal transition was still a feature of DGC only. This analysis also revealed the involvement of histone modifications associated with both histological types, histone deacetylase related to DGC and histone acetyltransferase (HAT) to IGC. Details on all analyses are presented in Fig. 6 and Data S2.

### 3.6. SRCC vs PCC-NOS: distinctive mutational profiles

We then wanted to understand whether the mutational landscape of DGC is common across DGC histological subtypes, or if each subtype (SRCC and PCC-NOS) has its unique molecular profile.

For this, we compared the set of genes mutated in SRCC and in PCC-NOS, described in two Asian studies and obtained with different technical approaches (one targeted sequencing; one WES) (Data S1, Table S1).

### 3.7. SRCC vs PCC-NOS transversally mutated genes

There were 22 transversally mutated genes across SRCC and PCC-NOS, from which four genes (*CDH1*, *TP53*, *ATM* and *KRAS)* were found in two independent data sets per histological subtype (Fig. 7). *TP53* was the most frequently mutated gene in both DGC subtypes. The remaining transversally mutated genes in SRCC and PCC-NOS were found in a single study, and none of them was mutated >twofold when comparing SRCC and PCC-NOS (Fig. 7).

### 3.8. SRCC- vs PCC-NOS-enriched mutated genes

Only two genes (*CREBBP* and *MAP2K4*) were found enriched in SRCC (27% for both) as compared to PCC-NOS (10% for both) (Fig. 7). On the other hand, 13 mutated genes were found enriched in PCC-NOS as compared to SRCC. From these, *PTEN, SMAD4, SETD2, PIK3CA, BRAF* and *RPL5* mutations were found in at least 20% of PCC-NOS as compared to a maximum of 9% in SRCC (Fig. 7). Most PCC-NOS enriched mutated genes either had tumour suppressor activity or were associated with histone markers of transcription repression.

### 3.9. SRCC-specific mutated genes

Six genes (*CDKN2A*, *POLQ*, *SETBP1*, *SOX9*, *TNFAIP3* and *ZFHX3*) were found to be exclusive of SRCC (SRCC-specific mutated genes) (Fig. 7). While *CDKN2A* was found mutated in 27% of the cases, the remaining five genes were found mutated in < 10% of the SRCC.

### 3.10. PCC-NOS-specific mutated genes

Five genes (*RHOA*, *BRCA2*, *JAK3*, *BRCA1* and *MSH6*), were found exclusively mutated in PCC-NOS (PCC-NOS-specific mutated genes) (Fig. 7). While *RHOA*, *BRCA2* and *JAK3* were found mutated in > 20% of the cases, *BRCA1* and *MSH6* genes were found mutated in 10% of PCC-NOS cases.

This mutation frequency-focused analysis revealed more similarities than differences between SRCC and PCC-NOS, regarding the set of genes affected in both subtypes. However, some very specific features were also found per subtype, likely supporting their morphological differences.

### 3.11. Insights into similarities and differences between SRCC and PCC-NOS mutational profiles

We compared the mutation landscapes of DGC, IGC, SRCC and PCC-NOS aiming to address particular mutational landscapes of SRCC and PCC-NOS. Nine genes were shared by DGC, IGC, PCC-NOS and

**Fig. 7.** Mutational landscape of SRCC and PCC-NOS. For this analysis, two original publications were used. The origin, number of samples, reported study identification and other cohorts' characteristics are indicated in the heading of the table. The mutated genes of each subtype were divided into three main classifications: enriched, transversally mutated and specific genes. Each mutation frequency was associated with a colour scale. Green was appointed to genes with mutation frequencies between 0% and 9% (less frequently mutated). Yellow was used for mutation frequencies between 10% and 19% and light orange for mutation frequencies between 20% and 39%. A variation between dark orange and red was associated with mutation frequencies increasing from 40% to a maximum of 66% (frequently mutated).

| Information source | PMID: 28873240 | PMID: 31171626 | PMID: 28873240 | PMID: 31171626 | |
|---|---|---|---|---|---|
| Origin of samples | Korea | USA, China, Japan | Korea | USA, China, Japan | Recurrence |
| Number of samples | *n* = 32 | *n* = 11 | *n* = 59 | *n* = 10 | |
| Cohort Specification | SRCC | SRCC | PCC-NOS | PCC-NOS | |
| Genes | Frequency of mutated samples per study | | | | |
| **Transversally** mutated genes — TP53 | 25% | 64% | 66% | 50% | Frequencies reported in both studies |
| CDH1 | 16% | 45% | 22% | 30% | |
| ATM | 6% | 9% | 15% | 10% | |
| KRAS | 6% | 9% | 15% | 10% | |
| PCC-NOS-**enriched** mutated genes — PTEN | 6% | 9% | 22% | 30% | |
| SMAD4 | 3% | 9% | 19% | 20% | |
| PCC-NOS-**specific** mutated genes — RHOA | 0% | 0% | 25% | 0% | |
| **Transversally** mutated genes — APC | 19% | – | 32% | – | Frequencies reported in a single study |
| KIT | 19% | – | 29% | – | |
| ARID1A | 19% | – | 27% | – | |
| CTNNB1 | 19% | – | 24% | – | |
| EGFR | 16% | – | 29% | – | |
| KDR | 16% | – | 20% | – | |
| PDGFRB | 16% | – | 19% | – | |
| ERBB3 | 13% | – | 10% | – | |
| SHH | 13% | – | 8% | – | |
| IGF1R | 9% | – | 17% | – | |
| KMT2C | – | 36% | – | 20% | |
| MLH1 | – | 18% | – | 30% | |
| TAF1 | – | 18% | – | 20% | |
| SPTA1 | – | 18% | – | 10% | |
| APOB | – | 9% | – | 10% | |
| GNAS | – | 9% | – | 10% | |
| MET | – | 9% | – | 10% | |
| SPTAN1 | – | 9% | – | 10% | |
| PCC-NOS-**specific** mutated genes — BRCA2 | – | 0% | – | 20% | |
| JAK3 | – | 0% | – | 20% | |
| BRCA1 | – | 0% | – | 10% | |
| MSH6 | – | 0% | – | 10% | |
| SETD2 | – | 9% | – | 30% | |
| PCC-NOS-**enriched** mutated genes — PIK3CA | 9% | – | 24% | – | |
| BRAF | 6% | – | 24% | – | |
| RPL5 | – | 9% | – | 20% | |
| VHL | 9% | – | 19% | – | |
| GLI1 | 9% | – | 19% | – | |
| HDAC9 | 9% | – | 19% | – | |
| ERBB4 | 9% | – | 19% | – | |
| EZH2 | 6% | – | 17% | – | |
| PDGFRA | 6% | – | 17% | – | |
| PTCH1 | 3% | – | 17% | – | |
| SRCC-**specific** mutated genes — CDKN2A | – | 27% | – | 0% | |
| POLQ | – | 9% | – | 0% | |
| SETBP1 | – | 9% | – | 0% | |
| SOX9 | – | 9% | – | 0% | |
| TNFAIP3 | – | 9% | – | 0% | |
| ZFHX3 | – | 9% | – | 0% | |
| SRCC-**enriched** mutated genes — CREBBP | – | 27% | – | 10% | |
| MAP2K4 | – | 27% | – | 10% | |
| Cohort Specification | SRCC | SRCC | PCC-NOS | PCC-NOS | |
| Information source | PMID: 28873240 | PMID: 31171626 | PMID: 28873240 | PMID: 31171626 | |

**A   SRCC:**
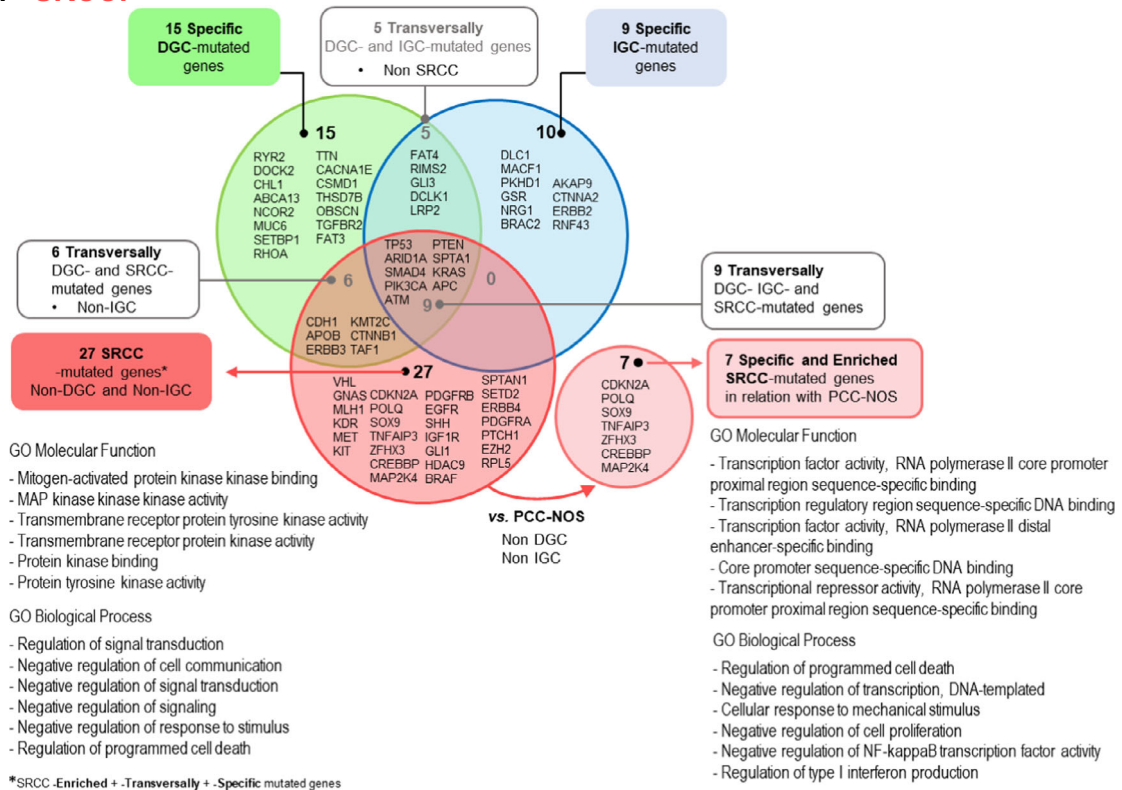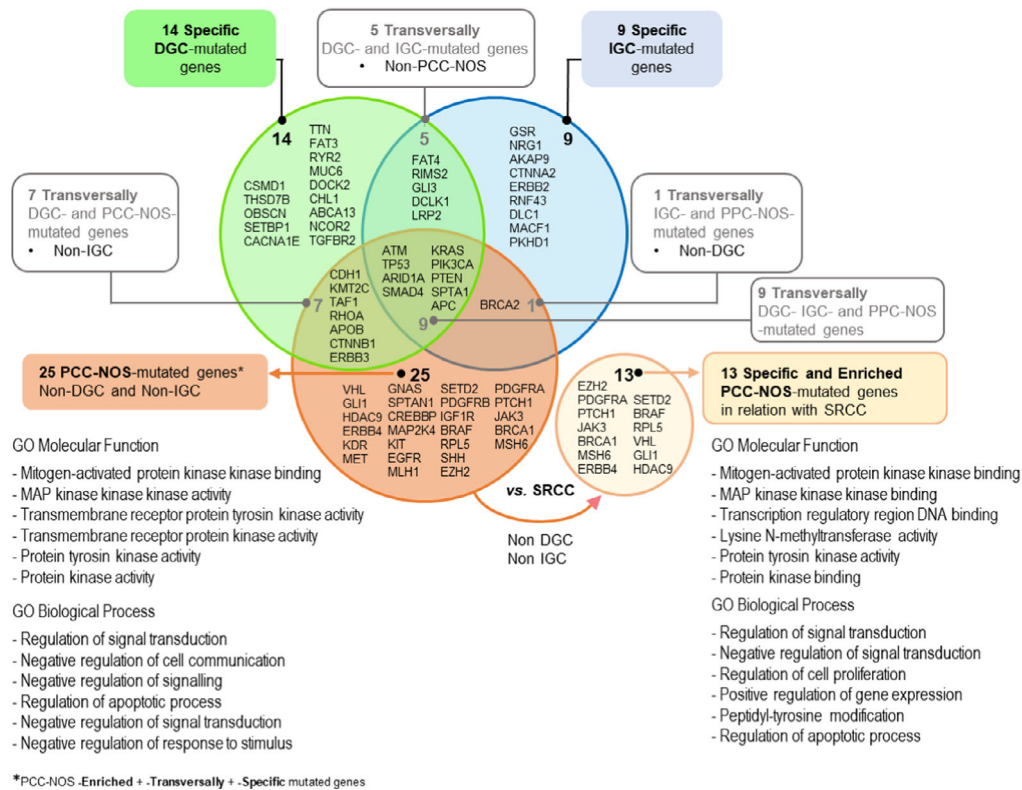


**B   PCC-NOS:**

**Fig. 8.** Specific and transversally mutated genes in SRCC (A) and PCC/NOS (B) in comparison with DGC and IGC. SRCC-associated genes were linked to a red colour, while PCC-NOS-associated genes are related to the orange colour. In both Fig. 8 A, B, the circle outside of the Venn diagram contains the specific genes for SRCC and PCC-NOS, respectively. GO terms for molecular functions and biological processes are depicted for the mutated genes associated with each subtype that were not shared with DGC nor IGC, as well for the SRCC/PCC-NOS-specific genes. The analysis generated using *Enrichr* and the associated GO terms number is available in Data S2.

SRCC. These are well-known cancer driver genes with crucial functions related to cell proliferation, migration, DNA repair or angiogenesis [63]. Some are well-known tumour suppressors such as *TP53*, *ATM*, *APC*, *ARID1A*, *SMAD4* and *PTEN* or oncogenes such as *PIK3CA* and *KRAS* (Fig. 8A,B).

SRCC shares a set of six mutated genes with PCC-NOS, which are also found in DGC but not in IGC (*CDH1*, *APOB*, *ERBB3*, *KMT2C*, *CTNNB1*, *TAF1*) (Fig. 8A,B). Among these six genes is the hallmark gene *CDH1* (Fig. 8A,B). There are no genes shared between SRCC and IGC, and *BRCA2* is the only mutated gene shared by PCC-NOS and IGC (Fig. 8B). These data likely identify the mutated genes that are common in DGC, SRCC and PCC-NOS, in keeping with the fact that the two latter are subtypes of the former. SRCC and PCC-NOS diverge in morphology and also in their mutational profile. We found seven genes that are SRCC-specific and enriched genes (*CDKN2A*, *POLQ*, *SOX9*, *TNFAIP3*, *ZFHX3*, *CREBBP* and *MAP2K4*) in relation to PCC-NOS (Fig. 8A). Both in the restricted set of seven specific and enriched SRCC-mutated genes and in a larger set of 27 SRCC-mutated genes (Fig. 8A), an association with programmed cell death is always identified among GO biological processes, representing a novel feature of SRCC. Moreover, the seven genes that differentiate SRCC from all other GC histotypes or subtypes are mainly associated with transcriptional activity (TOP6 ranking GO molecular functions) and cell proliferation and immune or inflammatory response (TOP6 GO biological processes) (Fig. 8A).

We found 25 mutated genes that differentiate PCC-NOS from DGC and IGC, and from these, 13 are PCC-NOS-specific and enriched genes (*EZH2*, *SETD2*, *PDGFRA*, *BRAF*, *PTCH1*, *RPL5*, *JAK3*, *VHL*, *BRCA1*, *GLI1*, *MSH6*, *HDAC9* and *ERBB4*) in relation to SRCC (Fig. 8B). This combination of mutated genes likely represents a distinctive feature of PCC-NOS, which despite being mainly associated with MAPK pathway and protein kinase activity, as for the main GC histotypes, it is specifically associated with transcription regulatory region DNA binding (TOP6 ranking GO molecular functions). This feature brings PCC-NOS closer to SRCC (Fig. 8B).

In summary, this analysis supports a divergent evolution of SRCC and PCC-NOS mutation profiles and highlights particular molecular pathways and functions from each of these two DGC subtypes.

## 3.12. Therapies specifically directed to DGC do not consider DGC mutational landscape

We searched for ongoing and recently closed clinical trials specifically mentioning the inclusion of patients with DGC or DGC subtypes to understand whether any of these trials used as therapy target, the genes or pathways herein identified has preferentially involved in DGC and its subtypes.

We found 13 clinical trials using the search criteria described in the methodology (Table S2). In total, we found four trials directed to DGC patients, two completed (DIGEST and MONO) and two currently recruiting or about to start recruitment (NORDICA and ADCI002 Study) (Table 3).

The DIGEST phase III study (NCT01285557) compared the use of fluorouracil (5-FU) combined with cisplatin with the use of the latter plus S-1 in 361 patients with metastatic DGC (Table 3). This trial has been completed, and results showed that there was no significant difference in the outcome of the patients comparing both approaches [64].

The MONO study (NCT01197885), a phase II clinical trial, examined the outcome of zolbetuximab (monoclonal antibody against CLDN18.2: IMAB362) monotherapy in a series of patients with recurrent or refractory, locally advanced or metastatic and CLDN18.2-positive gastric, gastro-oesophageal junction, or oesophageal adenocarcinoma (Table 3). From the 54 patients enrolled in this trial, 22 were advanced DGC patients with CLDN18.2 expression in ≥ 50% of tumour cells. Zolbetuximab was well-tolerated and exhibited antitumour activity with a clinical benefit rate of 23% [60]; whether patients benefiting from this targeted therapy were DGC patients is unknown.

The ADCI002 phase II/III clinical trial (NCT01717924), that is currently recruiting, was designed to address the inherent chemoresistance of SRCC. It evaluates primary surgery vs primary chemotherapy followed by surgery, using epirubicin for topoisomerase II inhibition combined with cisplatin and 5-FU [65].

**Table 3.** Details of clinical trials registered at ClinicalTrials.gov directed to DGC patients.Green was used to highlight clinical trials that specifically mention DGC; red was used to highlight clinical trials that specifically mention SRCC; and white was used to highlight a specific clinical trial with promising results based on CLDN18.2 expression (Table 2).

| ClinicalTrials.gov Identifier | Title of the clinical trial | Strategy | Molecular target | Status | Publications | Clinical trial outcome |
|---|---|---|---|---|---|---|
| NCT03977220 | Nab-paclitaxel combined with S-1 treating diffuse type of stage gastric cancer as adjuvant setting (NORDICA) | Evaluate paclitaxel for microtubules stabilization and S-1 (oral derivative of 5-FU) | Microtubules and conventional chemotherapy targets | Not yet recruiting | No publications available | Not available |
| NCT01717924 | Evaluation of Surgery vs Primary Chemotherapy in Resectable Signet-ring Cell Gastric Adenocarcinoma (ADCI002) (Phase II/III) | Compare primary surgery vs primary chemotherapy followed by surgery. epirubicin for topoisomerase II inhibition; cisplatin; 5-fluorouracil | Topoisomerase II and conventional chemotherapy targets | Recruiting | [65] | Not available |
| NCT01285557 | Diffuse Gastric and Esophagogastric Junction Cancer S-1 Trial (DIGEST) (Phase III) | Evaluate the safety and efficacy of S-1 and cisplatin compared with 5-FU and cisplatin | Conventional chemotherapy targets | Completed | [64] | No significant difference has been found in the outcome of the patients comparing both approaches |
| NCT01197885 | Efficacy and Safety Study of Multiple Doses of IMAB362 in Patients with Advanced Gastroesophageal Cancer (MONO) (Phase IIa) | Evaluate the safety and efficacy of IMAB362 used as monoclonal antibody against CLDN18.2 (single agent) in GC patients (22/54 DGC) | CLDN18.2 | Completed | [60] | Zolbetuximab was well-tolerated and exhibited antitumour activity with a clinical benefit rate for patients of 23%; whether these were DGC patients was not disclosed |

The NORDICA phase I study (NCT03977220), which is specifically directed to DGC patients, is not yet recruiting and aims at using microtubules stabilization agent (paclitaxel) combined with S-1 (Table 3).

None of the trials described above used as target for therapy, the specific mutations or mutation profiles and associated molecular functions herein found to differentiate DGC and its subtypes from each other and from IGC.

### 3.13. Sporadic DGC and hereditary DGC (HDGC): somatic mutational profiles

We tried to understand whether DGC arising in different settings (hereditary or sporadic) would share a similar mutational profile, and whether DGC arising in the HDGC context would mimic a mutational profile closer to that of SRCC.

The so-called HDGC syndrome is caused mainly by germline *CDH1* and rarely by *CTNNA1* pathogenic or likely pathogenic variants and is dominated by the development of early-onset DGC, mainly SRCC [66–68]. However, the somatic mutational landscape of DGC/SRCC arising in this setting is, to date, fairly unknown.

We could only find two studies reporting the somatic mutational landscape of HDGC tumours [67,69]. The germline causative defect in one family affected *CDH1* [69] and in the other family affected *CTNNA1* [67].

We compared the list of somatic mutated genes found in these HDGC cases with that found in sporadic DGC cases (Fig. 4). The mutational profile of the *CTNNA1* germline mutated tumour [67] encompassed somatic mutations in *RHOA*, a specific DGC and PCC-NOS gene (Fig. 7), and in *ARID1A* and *PIK3CA*, genes transversally mutated in both DGC and IGC (Fig. 7). This suggests that, independently of the cancer setting (hereditary vs diffuse), or histological type (DGC vs IGC), GC developing in this *CTNNA1* variant germline carrier presents a

nonspecific combination of somatic events. By contrast, for the *CDH1* germline mutated tumour, somatic mutations were described in *TTN* and *CDH1* by Funakoshi *et al* [69], which are both specific DGC-mutated genes (this study, Fig. 6); being *CDH1* was also considered transversally mutated in SRCC and PCC-NOS (Fig. 7).

To better understand whether there are real differences between the mutational profiles of sporadic and hereditary DGC, or whether HDGC cases are closer to a particular DGC subtype, additional studies need to be performed.

## 4. Discussion

In the present study, we started by revising GC histological and molecular classifications, highlighting associated relevant molecular landscapes, clinical outcomes and currently approved or promising therapies. The state of the art on all these aspects is presented in the introduction section. The main outcome of this literature revision was the identification of an asymmetry, regarding prognosis, therapy targets and available effective therapies (both conventional and targeted), in TCGA GS and ACRG MSS/EMT, mainly encompassing DGC, in comparison with other molecular subgroups.

To address this gap in knowledge, and shed light into potential targets for therapeutic intervention, we explored a set of studies reporting on multigene mutational analysis to identify DGC-associated mutational profiles, molecular pathways and biological functions. We chose this approach, as mutational profiles are rather stable genetic events, easy to detect, interpret and less costly as compared to RNA profiles, methylation patterns or other epigenomic approaches. On the other hand, several tumour-specific mutations constitute well-proven, highly sensitive and specific predictive biomarkers of response to selective targeted therapies and are now an indispensable laboratory tool for therapeutic decisions [70].

Our systematic analysis confirms that *TP53, ARID1A, SMAD4, PIK3CA, PTEN, KRAS, APC, ATM* and *SPTA1* are driver and central genes in gastric tumorigenesis, and transversal to the main GC histotypes and DGC subtypes [62].

Also, herein we identified specific sets of genes that define the main molecular pathways classically involved in either DGC or IGC. However, tumours emerge and evolve due to a mutational profile that can be shared with other tumour types. We could find that the combination of all recurrently mutated genes either in DGC or in IGC triggered the involvement of similar molecular pathways regardless of being histotype-specific or common to both histotypes. This may explain why conventional therapies, and even some approved targeted therapies, demonstrate clinical benefit regardless of the GC histological type.

We further showed that the most recurrently mutated and specific genes in DGC, across many studies, were *CDH1* and *RHOA*. This result is reflected in the findings from the TCGA GS and ACRG MSS/ EMT subgroups, which are dominated by DGC. According to both studies, this subgroup presents the worst prognosis within GC molecular subtypes, resistance to conventional chemotherapy and so far, have no approved targeted therapies [22,37,49]. Interestingly, a recent study reported that late-onset DGCs presented higher mutation frequencies in *RHOA* and less frequent mutations in *CDH1* in comparison with early-onset DGCs [31,71].

Although none of the ongoing clinical trials directed to DGC patients use these two proteins as targets for therapy (Table 3), it is tempting to speculate whether their mutations would perform well as therapy targets or predictors of therapy responsiveness. E-cadherin inactivation is a highly unlikely druggable target mechanism, and it has been shown to cause multidrug resistance. By contrast, E-cadherin disruption is also known to induce upregulation of the anti-apoptotic protein BCL2, and activation of Rho-Rock pathway, which may be targetable (reviewed in Ref. [72]). Taking advantage of the pro-survival context in E-cadherin deficient DGC, treatment using pro-apoptotic drugs may be feasible. Indeed, preclinical treatment with pro-apoptotic BH3-only mimetics combined with chemotherapy or endocrine treatment regimens had a positive effect in breast cancer [73,74]. Preclinical studies also suggested that pharmacological inhibition of Rho-Rock signalling may be a clinically relevant target in cancer [75]. Other genes specifically mutated in DGC may also be interesting to consider when rethinking biomarkers for prognosis and therapy in patients with DGC. An example is the *TTN* gene, whose mutations were found specifically in DGC and on average in 40% of the cases. Although not previously related specifically to DGC, *TTN* was found as one of the top five hub genes of a specific co-expression module positively correlated with GC pathological tumour and lymph node stages [76]. In another study, *TTN* gene mutations were correlated with poor prognosis and predicted tumour mutation burden and immunotherapy efficacy in GC [77]. Other examples are *FAT3* and *KMT2C*, herein found mutated on average in 15% and 12% of DGC, respectively, and found previously associated with this cancer

histotype [76,78,79]. Namely, *KMT2C* loss was shown to promote epithelial-to-mesenchymal transition and was associated with poor overall survival.

Although our study has not addressed the role of epigenetic changes in DGC, these events are also considered significant carcinogenic factors, and epigenetic regulation has been claimed as a promising molecular target for therapy in GC precision oncology [17,80]. Herein, we have shown that histone modifications appear not only in the TOP6 molecular functions from DGC, but also in those from IGC, and when specific and transversally mutated genes are considered. This finding reinforces the need of studies focusing specifically on the epigenetics profiles of DGC and its subtypes, with the aim of finding potential targets for precision oncology.

The current systematic analysis also provides, to our knowledge the first combined analysis on SRCC and PCC-NOS mutational profiles, considering two independent studies. This revealed *CDH1*, *TP53*, *ATM* and *KRAS* mutations as recurrent and common events of both DGC subtypes, again supporting the role of GC driver genes, and *RHOA* mutations as a specificity of PCC-NOS [16]. In terms of more specific mutational landscapes, both sets of divergent mutated genes between overall DGC and SRCC and between overall DGC and PCC-NOS are associated with TRK pathways, cell–cell communication and apoptosis, although involving different genes. However, and interestingly, the set of genes that distinguish SRCC and PCC-NOS from DGC overall and from each other associates with transcriptional regulation and gene expression control, besides TRK pathways and apoptosis, particularly the SRCC subtype. Indeed, *CREBBP* and *MAP2K4* genes, whose mutations are enriched in SRCC as compared to PCC-NOS (Fig. 7), are involved in coupling chromatin remodelling to transcription factor recognition and transcription regulation, respectively [81,82]. A loss-of-function mutation in the *CBP*/*CREBBP* gene, which encodes a HAT, is common to many cancers, including GC [82], and *MAP2K4* loss of expression has been reported in gastric adenocarcinoma associated with poor survival [81]. Given that a great deal of transcriptional control involves epigenetic regulation, epigenetic-related treatments may hold promise for SRCC and PCC-NOS precision oncology [80], besides the above-mentioned possibilities described for DGC overall.

To the best of our knowledge, this is the first systematic analysis describing the distinctive mutational profiles of DGC and its subtypes, resourcing to a recurrent analysis, to address the gap in potential treatment options for this GC histotype. By exploring the literature regarding clinical trials directed to the particularities of DGC, we concluded that these are only a few and rarely specifically designed for DGC patients. This may explain the delay in finding treatment options for these patients. Because we were not able to find more promising results than those obtained for anti-Claudin18.2 targeted therapy, we herein highlight a set of recurrent mutant genes and associated pathways that may be explored for novel therapeutic designs in DGC and its subtypes. The interpretation of past clinical trials in the light of tumour morphology can also guide therapeutic development by identifying subsets of patients with better response to a certain treatment strategy. Exploring the mutational landscape of DGC subtypes can further identify new and more adequate druggable targets with therapeutic implications.

Finally, improving the knowledge on somatic molecular/genomic players in sporadic DGC may also be valuable to guide future treatments in the HDGC setting, which also remains poorly studied from the somatic standpoint and without chemotherapeutic options.

# Acknowledgements

## Conflict of interest

The authors declare no conflict of interest.

## Author contributions

CO and FC conceived and supervised the study; CO, RB-M and JG-P designed the study; RB-M, JG-P and IG acquired the data; RB-M, JG-P, IG, FC and CO analysed and interpreted the data; all authors drafted the manuscript; FC, IG and CO edited and reviewed the manuscript; CO and FC obtained funding. All authors critically revised the manuscript for important intellectual content. All authors have read and agreed to the published version of the manuscript.

## References

1 Gullo I, Carneiro F, Oliveira C & Almeida GM (2018) Heterogeneity in gastric cancer: from pure morphology to molecular classifications. *Pathobiology* **85**, 50–63.

2 Carneiro F, Seixas M & Sobrinho-Simões M (1995) New elements for an updated classification of the carcinomas of the stomach. *Pathol Res Pract* **191**, 571–584.

3 Lauren P (1965) 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* **64**, 31–49.

4 Ming SC (1977) Gastric carcinoma. A pathobiological classification. *Cancer* **39**, 2475–2485.

5 Nakamura K, Sugano H & Takagi K (1968) Carcinoma of the stomach in incipient phase: its histogenesis and histological appearances. *Gan* **59**, 251–258.

6 Solcia E, Klersy C, Mastracci L, Alberizzi P, Candusso M, Diegoli M, Tava F, Riboni R, Manca R & Luinetti O (2009) A combined histologic and molecular approach identifies three groups of gastric cancer with different prognosis. *Virchows Arch* **455**, 197–211.

7 Japanese Gastric Cancer Association (2017) Japanese Classification of Gastric Carcinoma, 15th edn. Kanehara Shuppan, Tokyo.

8 WHO Classification of Tumours Editorial Board (2019) Digestive system tumours, 5th edn. WHO classification of tumours series, Lyon: International Agency for Research on Cancer.

9 Carneiro F, Fukayama M, Grabsch H & Yasui WGA (2019) WHO Classification of Tumours Editorial Board. Digestive system tumours, pp. 85–95, WHO classification of tumours series, Lyon: International Agency for Research on Cancer.

10 Park HK, Lee KY, Yoo MW, Hwang TS & Han HS (2016) Mixed carcinoma as an independent prognostic factor in submucosal invasive gastric carcinoma. *J Korean Med Sci* **31**, 866–872.

11 Zheng HC, Li XH, Hara T, Masuda S, Yang XH, Guan YF & Takano Y (2008) Mixed-type gastric carcinomas exhibit more aggressive features and indicate the histogenesis of carcinomas. *Virchows Arch* **452**, 525–534.

12 Carneiro F (1997) Classification of gastric carcinomas. *Curr Diagn Pathol* **4**, 51–59.

13 Mariette C, Carneiro F, Grabsch HI, van der Post RS, Allum W & de Manzoni G (2019) Consensus on the pathological definition and classification of poorly cohesive gastric carcinoma. *Gastric Cancer* **22**, 1–9.

14 Sugihara H, Hattori T, Fukuda M & Fujita S (1987) Cell proliferation and differentiation in intramucosal and advanced signet ring cell carcinomas of the human stomach. *Virchows Arch A Pathol Anat Histopathol* **411**, 117–127.

15 Gullo I, van der Post RS & Carneiro F (2021) Recent advances in the pathology of heritable gastric cancer syndromes. *Histopathology* **78**, 125–147.

16 Kwon CH, Kim YK, Lee S, Kim A, Park HJ, Choi Y, Won YJ, Park DY & Lauwers GY (2018) Gastric poorly cohesive carcinoma: a correlative study of mutational signatures and prognostic significance based on histopathological subtypes. *Histopathology* **72**, 556–568.

17 Ooi WF, Xing M, Xu C, Yao X, Ramlee MK, Lim MC, Cao F, Lim K, Babu D, Poon LF *et al.* (2016) Epigenomic profiling of primary gastric adenocarcinoma reveals super-enhancer heterogeneity. *Nat Commun* **7**, 12983.

18 Miki H, Ohmori M, Perantoni AO & Enomoto T (1991) K-ras activation in gastric epithelial tumors in Japanese. *Cancer Lett* **58**, 107–113.

19 Oliveira C, Pinto M, Duval A, Brennetot C, Domingo E, Espín E, Armengol M, Yamamoto H, Hamelin R, Seruca R *et al.* (2003) BRAF mutations characterize colon but not gastric cancer with mismatch repair deficiency. *Oncogene* **22**, 9192–9196.

20 Samuels Y, Wang Z, Bardelli A, Silliman N, Ptak J, Szabo S, Yan H, Gazdar A, Powell SM, Riggins GJ *et al.* (2004) High frequency of mutations of the PIK3CA gene in human cancers. *Science* **304**, 554.

21 Velho S, Oliveira C, Ferreira A, Ferreira AC, Suriano G, Schwartz S Jr, Duval A, Carneiro F, Machado JC, Hamelin R *et al.* (2005) The prevalence of PIK3CA mutations in gastric and colon cancer. *Eur J Cancer* **41**, 1649–1654.

22 Cancer Genome Atlas Research Network (2014) Comprehensive molecular characterization of gastric adenocarcinoma. *Nature* **513**, 202–209.

23 Hu BL, Xie MZ, Li KZ, Li JL, Gui YC & Xu JW (2019) Genome-wide analysis to identify a novel distant metastasis-related gene signature predicting survival in patients with gastric cancer. *Biomed Pharmacother* **117**, 109159.

24 Pirini F, Noazin S, Jahuira-Arias MH, Rodriguez-Torres S, Friess L, Michailidi C, Cok J, Combe J, Vargas G, Prado W *et al.* (2017) Early detection of gastric cancer using global, genome-wide and IRF4, ELMO1, CLIP4 and MSC DNA methylation in endoscopic biopsies. *Oncotarget* **8**, 38501–38516.

25 Cui J, Yin Y, Ma Q, Wang G, Olman V, Zhang Y, Chou WC, Hong CS, Zhang C, Cao S *et al.* (2015) Comprehensive characterization of the genomic alterations in human gastric cancer. *Int J Cancer* **137**, 86–95.

26 Hanazono K, Natsugoe S, Stein HJ, Aikou T, Hoefler H & Siewert JR (2006) Distribution of p53 mutations in esophageal and gastric carcinomas and the relationship with p53 expression. *Oncol Rep* **15**, 821–824.

27 van Beek E, Hernandez JM, Goldman DA, Davis JL, McLaughlin K, Ripley RT, Kim TS, Tang LH, Hechtman JF, Zheng J *et al.* (2018) Rates of TP53 Mutation are Significantly Elevated in African American Patients with Gastric Cancer. *Ann Surg Oncol* **25**, 2027–2033.

28 Becker KF, Atkinson MJ, Reich U, Becker I, Nekarda H, Siewert JR & Höfler H (1994) E-cadherin gene mutations provide clues to diffuse type gastric carcinomas. *Cancer Res* **54**, 3845–3852.

29 Berx G, Becker KF, Höfler H & van Roy F (1998) Mutations of the human E-cadherin (CDH1) gene. *Hum Mutat* **12**, 226–237.

30 Bustos-Carpinteyro AR, Oliveira C, Sousa A, Oliveira P, Pinheiro H, Carvalho J, Magaña-Torres MT, Flores-Miramontes MG, Aguilar-Lemarroy A, Jave-Suárez LF *et al.* (2019) CDH1 somatic alterations in Mexican patients with diffuse and mixed sporadic gastric cancer. *BMC Cancer* **19**, 69.

31 Cho SY, Park JW, Liu Y, Park YS, Kim JH, Yang H, Um H, Ko WR, Lee BI, Kwon SY *et al.* (2017) Sporadic early-onset diffuse gastric cancers have high frequency of somatic CDH1 alterations, but low frequency of somatic RHOA mutations compared with late-onset cancers. *Gastroenterology* **153**, 536–549.e26.

32 Moridnia A, Tabatabaiefar MA, Zeinalian M, Minakari M, Kheirollahi M & Moghaddam NA (2019) Novel variants and copy number variation in CDH1 gene in Iranian patients with sporadic diffuse gastric cancer. *J Gastrointest Cancer* **50**, 420–427.

33 Nemtsova MV, Kalinkin AI, Kuznetsova EB, Bure IV, Alekseeva EA, Bykov II, Khorobrykh TV, Mikhaylenko DS, Tanas AS, Kutsev SI *et al.* (2020) Clinical relevance of somatic mutations in main driver genes detected in gastric cancer patients by next-generation DNA sequencing. *Sci Rep* **10**, 504.

34 Wang SC, Yeu Y, Hammer STG, Xiao S, Zhu M, Hong C, Clemenceau JR, Yoon LY, Nassour I, Shen J *et al.* (2020) Hispanic/Latino patients with gastric adenocarcinoma have distinct molecular profiles including a high rate of germline CDH1 variants. *Cancer Res* **80**, 2114–2124.

35 Tan IB, Ivanova T, Lim KH, Ong CW, Deng N, Lee J, Tan SH, Wu J, Lee MH, Ooi CH *et al.* (2011) Intrinsic subtypes of gastric cancer, based on gene expression pattern, predict survival and respond differently to chemotherapy. *Gastroenterology* **141**, 476–485.e11.

36 Lei Z, Tan IB, Das K, Deng N, Zouridis H, Pattison S, Chua C, Feng Z, Guan YK, Ooi CH *et al.* (2013) Identification of molecular subtypes of gastric cancer with different responses to PI3-kinase inhibitors and 5-fluorouracil. *Gastroenterology* **145**, 554–565.

37 Cristescu R, Lee J, Nebozhyn M, Kim KM, Ting JC, Wong SS, Liu J, Yue YG, Wang J, Yu K *et al.* (2015) Molecular analysis of gastric cancer identifies subtypes associated with distinct clinical outcomes. *Nat Med* **21**, 449–456.

38 Setia N, Agoston AT, Han HS, Mullen JT, Duda DG, Clark JW, Deshpande V, Mino-Kenudson M, Srivastava A, Lennerz JK *et al.* (2016) A protein and mRNA expression-based classification of gastric cancer. *Modern Pathol* **29**, 772–784.

39 Ahn S, Lee SJ, Kim Y, Kim A, Shin N, Choi KU, Lee CH, Huh GY, Kim KM, Setia N *et al.* (2017) High-throughput protein and mRNA expression-based classification of gastric cancers can identify clinically distinct subtypes, concordant with recent molecular classifications. *Am J Surg Pathol* **41**, 106–115.

40 Gonzalez RS, Messing S, Tu X, McMahon LA & Whitney-Miller CL (2016) Immunohistochemistry as a surrogate for molecular subtyping of gastric adenocarcinoma. *Hum Pathol* **56**, 16–21.

41 Diaz Del Arco C, Estrada Munoz L, Molina Roldan E, Ceron Nieto MA, Ortega Medina L, Gomez G, de Las Heras S & Fernandez Acenero MJ (2018) Immunohistochemical classification of gastric cancer based on new molecular biomarkers: a potential predictor of survival. *Virchows Arch* **473**, 687–695.

42 Birkman EM, Mansuri N, Kurki S, Algars A, Lintunen M, Ristamaki R, Sundstrom J & Carpen O (2018) Gastric cancer: immunohistochemical classification of molecular subtypes and their association with clinicopathological characteristics. *Virchows Arch* **472**, 369–382.

43 Koh J, Lee KW, Nam SK, Seo AN, Kim JW, Kim JW, Park DJ, Kim HH, Kim WH & Lee HS (2019) Development and validation of an easy-to-implement, practical algorithm for the identification of molecular subtypes of gastric cancer: prognostic and therapeutic implications. *Oncologist* **24**, e1321–e1330.

44 Di Pinto F, Armentano R, Arborea G, Schena N, Donghia R & Valentini AM (2020) Are immunohistochemical markers useful in phenotypic gastric cancer classification? *Oncology* **98**, 566–574.

45 Wang Q, Xie Q, Liu Y, Guo H, Ren Y, Li J & Zhao Q (2020) Clinical characteristics and prognostic significance of TCGA and ACRG classification in gastric cancer among the Chinese population. *Mol Med Rep* **22**, 828–840.

46 Pinto MP, Córdova-Delgado M, Retamal IN, Muñoz-Medel M, Bravo ML, Durán D, Villanueva F, Sanchez C, Acevedo F, Mondaca S *et al.* (2020) A molecular stratification of Chilean gastric cancer patients with potential clinical applicability. *Cancers* **12**, 1863.

47 Tsai JH, Jeng YM, Chen KH, Lee CH, Yuan CT & Liau JY (2020) An integrative morphomolecular classification system of gastric carcinoma with distinct clinical outcomes. *Am J Surg Pathol* **44**, 1017–1030.

48 Zhao C, Feng Z, He H, Zang D, Du H, Huang H, Du Y, He J, Zhou Y & Nie Y (2020) Protein expression-based classification of gastric cancer by immunohistochemistry of tissue microarray. *PLoS One* **15**, e0238836.

49 Sohn BH, Hwang JE, Jang HJ, Lee HS, Oh SC, Shim JJ, Lee KW, Kim EH, Yim SY, Lee SH *et al.* (2017) Clinical significance of four molecular subtypes of gastric cancer identified by the cancer genome atlas project. *Clin Cancer Res* **23**, 4441–4449.

50 Shu Y, Zhang W, Hou Q, Zhao L, Zhang S, Zhou J, Song X, Zhang Y, Jiang D, Chen X *et al.* (2018) Prognostic significance of frequent CLDN18-ARHGAP26/6 fusion in gastric signet-ring cell cancer. *Nat Commun* **9**, 2447.

51 Nakayama I, Shinozaki E, Sakata S, Yamamoto N, Fujisaki J, Muramatsu Y, Hirota T, Takeuchi K, Takahashi S, Yamaguchi K *et al.* (2019) Enrichment of CLDN18-ARHGAP fusion gene in gastric cancers in young adults. *Cancer Sci* **110**, 1352–1363.

52 Ishigami H, Fujiwara Y, Fukushima R, Nashimoto A, Yabusaki H, Imano M, Imamoto H, Kodera Y, Uenosono Y, Amagai K *et al.* (2018) Phase III trial comparing intraperitoneal and intravenous paclitaxel plus S-1 versus cisplatin plus S-1 in patients with gastric cancer with peritoneal metastasis: PHOENIX-GC trial. *J Clin Oncol* **36**, 1922–1929.

53 Selim JH, Shaheen S, Sheu WC & Hsueh CT (2019) Targeted and novel therapy in advanced gastric cancer. *Exp Hematol Oncol* **8**, 25.

54 Fashoyin-Aje L, Donoghue M, Chen H, He K, Veeraraghavan J, Goldberg KB, Keegan P, McKee AE & Pazdur R (2019) FDA approval summary: pembrolizumab for recurrent locally advanced or metastatic gastric or gastroesophageal junction adenocarcinoma expressing PD-L1. *Oncologist* **24**, 103–109.

55 Kang YK, Boku N, Satoh T, Ryu MH, Chao Y, Kato K, Chung HC, Chen JS, Muro K, Kang WK *et al.* (2017) Nivolumab in patients with advanced gastric or gastro-oesophageal junction cancer refractory to, or intolerant of, at least two previous chemotherapy regimens (ONO-4538-12, ATTRACTION-2): a randomised, double-blind, placebo-controlled, phase 3 trial. *Lancet* **390**, 2461–2471.

56 Kim ST, Cristescu R, Bass AJ, Kim KM, Odegaard JI, Kim K, Liu XQ, Sher X, Jung H, Lee M *et al.* (2018) Comprehensive molecular characterization of clinical responses to PD-1 inhibition in metastatic gastric cancer. *Nat Med* **24**, 1449–1458.

57 Bang YJ, Van Cutsem E, Feyereislova A, Chung HC, Shen L, Sawaki A, Lordick F, Ohtsu A, Omuro Y, Satoh T *et al.* (2010) Trastuzumab in combination with chemotherapy versus chemotherapy alone for treatment of HER2-positive advanced gastric or gastro-oesophageal junction cancer (ToGA): a phase 3, open-label, randomised controlled trial. *Lancet* **376**, 687–697.

58 Bardou P, Mariette J, Escudié F, Djemiel C & Klopp C (2014) jvenn: an interactive Venn diagram viewer. *BMC Bioinformatics* **15**, 293.

59 Chen EY, Tan CM, Kou Y, Duan Q, Wang Z, Meirelles GV, Clark NR & Ma'ayan A (2013) Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool. *BMC Bioinformatics* **14**, 128.

60 Türeci O, Sahin U, Schulze-Bergkamen H, Zvirbule Z, Lordick F, Koeberle D, Thuss-Patience P, Ettrich T, Arnold D, Bassermann F *et al.* (2019) A multicentre, phase IIa study of zolbetuximab as a single agent in patients with recurrent or refractory advanced adenocarcinoma of the stomach or lower oesophagus: the MONO study. *Ann Oncol* **30**, 1487–1495.

61 Jin D-H, Park S-E, Lee J, Kim K-M, Kim S, Kim D-H & Park J (2015) Copy number gains at 8q24 and 20q11-q13 in gastric cancer are more common in intestinal-type than diffuse-type. *PLoS One* **10**, e0137657.

62 Chandrashekar P, Ahmadinejad N, Wang J, Sekulic A, Egan JB, Asmann YW, Kumar S, Maley C & Liu L (2019) Somatic selection distinguishes oncogenes and tumor suppressor genes. *Bioinformatics* **36**, 1712–1717.

63 Kandoth C, McLellan MD, Vandin F, Ye K, Niu B, Lu C, Xie M, Zhang Q, McMichael JF, Wyczalkowski MA *et al.* (2013) Mutational landscape and significance across 12 major cancer types. *Nature* **502**, 333–339.

64 Ajani JA, Abramov M, Bondarenko I, Shparyk Y, Gorbunova V, Hontsa A, Otchenash N, Alsina M, Lazarev S, Feliu J *et al.* (2017) A phase III trial

comparing oral S-1/cisplatin and intravenous 5-fluorouracil/cisplatin in patients with untreated diffuse gastric cancer. *Ann Oncol* **28**, 2142–2148.

65 Piessen G, Messager M, Le Malicot K, Robb WB, Di Fiore F, Guilbert M, Moreau M, Christophe V, Adenis A & Mariette C (2013) Phase II/III multicentre randomised controlled trial evaluating a strategy of primary surgery and adjuvant chemotherapy versus peri-operative chemotherapy for resectable gastric signet ring cell adenocarcinomas - PRODIGE 19 - FFCD1103 - ADCI002. *BMC Cancer* **13**, 281.

66 Guilford P, Hopkins J, Harraway J, McLeod M, McLeod N, Harawira P, Taite H, Scoular R, Miller A & Reeve AE (1998) E-cadherin germline mutations in familial gastric cancer. *Nature* **392**, 402–405.

67 Majewski IJ, Kluijt I, Cats A, Scerri TS, de Jong D, Kluin RJ, Hansford S, Hogervorst FB, Bosma AJ, Hofland I *et al.* (2013) An α-E-catenin (CTNNA1) mutation in hereditary diffuse gastric cancer. *J Pathol* **229**, 621–629.

68 Blair VR, McLeod M, Carneiro F, Coit DG, D'Addario JL, van Dieren JM, Harris KL, Hoogerbrugge N, Oliveira C, van der Post RS *et al.* (2020) Hereditary diffuse gastric cancer: updated clinical practice guidelines. *Lancet Oncol* **21**, e386–e397.

69 Funakoshi T, Miyamoto S, Kakiuchi N, Nikaido M, Setoyama T, Yokoyama A, Horimatsu T, Yamada A, Torishima M, Kosugi S(2019) Genetic analysis of a case of Helicobacter pylori-uninfected intramucosal gastric cancer in a family with hereditary diffuse gastric cancer. *Gastric Cancer* **22**, 892–898.

70 Malone ER, Oliva M, Sabatini PJB, Stockley TL & Siu LL (2020) Molecular profiling for precision cancer therapies. *Genome Med* **12**, 8.

71 Ishikawa M, Hayashi H, Sakamoto N, Tanaka S & Nishihara H (2019) Cancer gene profiling explores the possible precision medicine for diffuse-type gastric adenocarcinoma. *Med Oncol* **37**, 10.

72 Christgen M & Derksen P (2015) Lobular breast cancer: molecular basis, mouse and cellular models. *Breast Cancer Res* **17**, 16.

73 Vaillant F, Merino D, Lee L, Breslin K, Pal B, Ritchie ME, Smyth GK, Christie M, Phillipson LJ, Burns CJ *et al.* (2013) Targeting BCL-2 with the BH3 mimetic ABT-199 in estrogen receptor-positive breast cancer. *Cancer Cell* **24**, 120–129.

74 Oakes SR, Vaillant F, Lim E, Lee L, Breslin K, Feleppa F, Deb S, Ritchie ME, Takano E, Ward T *et al.* (2011) Sensitization of BCL-2–expressing breast tumors to chemotherapy by the BH3 mimetic ABT-737, *Proc Natl Acad Sci USA* **109**, 2766–2771.

75 Ying H, Biroc SL, Li W-W, Alicke B, Xuan J-A, Pagila R, Ohashi Y, Okada T, Kamata Y & Dinter H (2006) The Rho kinase inhibitor fasudil inhibits tumor progression in human and rat tumor models. *Mole Cancer Therap* **5**, 2158–2164.

76 Wang H, Shen L, Li Y & Lv J (2020) Integrated characterisation of cancer genes identifies key molecular biomarkers in stomach adenocarcinoma. *J Clin Pathol* **73**, 579–586.

77 Yang Y, Zhang J, Chen Y, Xu R, Zhao Q & Guo W (2020) MUC4, MUC16, and TTN genes mutation correlated with prognosis, and predicted tumor mutation burden and immunotherapy efficacy in gastric cancer and pan-cancer. *Clin Transl Med* **10**, e155.

78 Luo T, Zhao J, Lu Z, Bi J, Pang T, Cui H, Yang B, Li W, Wang Y, Wu S *et al.* (2018) Characterization of long non-coding RNAs and MEF2C-AS1 identified as a novel biomarker in diffuse gastric cancer. *Transl Oncol* **11**, 1080–1089.

79 Cho S-J, Yoon C, Lee JH, Chang KK, Lin J-X, Kim Y-H, Kook M-C, Aksoy BA, Park DJ, Ashktorab H *et al.* (2018) *KMT2C* mutations in diffuse-type gastric adenocarcinoma promote epithelial-to-mesenchymal transition. *Clin Cancer Res* **24**, 6556–6569.

80 Ho SWT & Tan P (2019) Dissection of gastric cancer heterogeneity for precision oncology. *Cancer Sci* **110**, 3405–3414.

81 Cunningham SC, Kamangar F, Kim MP, Hammoud S, Haque R, Iacobuzio-Donahue C, Ashfaq R, Kern SE, Maitra A, Heitmiller RE *et al.* (2006) MKK4 status predicts survival after resection of gastric adenocarcinoma. *Arch Surg* **141**, 1095–1099.discussion 1100.

82 Ogiwara H, Sasaki M, Mitachi T, Oike T, Higuchi S, Tominaga Y & Kohno T (2016) Targeting p300 addiction in CBP-deficient cancers causes synthetic lethality by apoptotic cell death due to abrogation of MYC expression. *Cancer Discov* **6**, 430–445.

83 Garattini SK, Basile D, Cattaneo M, Fanotto V, Ongaro E, Bonotto M, Negri FV, Berenato R, Ermacora P, Cardellino GG *et al.* (2017) Molecular classifications of gastric cancers: Novel insights and possible future applications. *World J Gastrointest Oncol* **9**, 194–208.

84 Sundar R, Huang KK, Qamra A, Kim KM, Kim ST, Kang WK, Tan ALK, Lee J & Tan P (2019) Epigenomic promoter alterations predict for benefit from immune checkpoint inhibition in metastatic gastric cancer. *Ann Oncol* **30**, 424–430.

85 Refolo MG, Lotesoriere C, Messa C, Caruso MG & D'Alessandro R (2020) Integrated immune gene expression signature and molecular classification in gastric cancer: new insights. *J Leukoc Biol* **108**, 633–646.

86 Orditura M, Galizia G, Sforza V, Gambardella V, Fabozzi A, Laterza MM, Andreozzi F, Ventriglia J, Savastano B, Mabilia A *et al.* (2014) Treatment of gastric cancer. *World J Gastroenterol* **20**, 1635–1649.

87 Smyth EC, Verheij M, Allum W, Cunningham D, Cervantes A & Arnold D (2016) Gastric cancer: ESMO

Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Ann Oncol* 27, v38–v49.

88 Wilke H, Muro K, Van Cutsem E, Oh SC, Bodoky G, Shimada Y, Hironaka S, Sugimoto N, Lipatov O, Kim TY *et al.* (2014) Ramucirumab plus paclitaxel versus placebo plus paclitaxel in patients with previously treated advanced gastric or gastro-oesophageal junction adenocarcinoma (RAINBOW): a double-blind, randomised phase 3 trial. *Lancet Oncol* 15, 1224–1235.

89 Li J, Qin S, Xu J, Guo W, Xiong J, Bai Y, Sun G, Yang Y, Wang L, Xu N *et al.* (2013) Apatinib for chemotherapy-refractory advanced metastatic gastric cancer: results from a randomized, placebo-controlled, parallel-arm, phase II trial. *J Clin Oncol* 31, 3219–3225.

90 Cartwright E, Athauda A & Chau I (2020) Emerging precision therapies for gastric cancer. *Exp Rev Precis Med Drug Dev* 5, 299–311.

91 Riquelme I, Saavedra K, Espinoza JA, Weber H, García P, Nervi B, Garrido M, Corvalán AH, Roa JC & Bizama C (2015) Molecular classification of gastric cancer: towards a pathway-driven targeted therapy. *Oncotarget* 6, 24750–24779.

92 Tirino G, Pompella L, Petrillo A, Laterza MM, Pappalardo A, Caterino M, Orditura M, Ciardiello F, Galizia G & De Vita F (2018) What's New in Gastric Cancer: the therapeutic implications of molecular classifications and future perspectives. *Int J Mol Sci* 19, 2659.

93 Ignatova E, Seriak D, Fedyanin M, Tryakin A, Pokataev I, Menshikova S, Vakhabova Y, Smirnova K, Tjulandin S & Ajani JA (2020) Epstein-Barr virus-associated gastric cancer: disease that requires special approach. *Gastric Cancer* 23, 951–960.

## Supporting information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

**Data S1**. Full data on gene mutation frequencies obtained from the selected studies (sheet 1), reporting all mutated genes found for DGC cases (sheet 2), IGC cases (sheet 3) and SRCC/PCC-NOS (sheet 4).

**Data S2**. GO TERMS, Biological Processes and Molecular Functions retrieved by Enrichr for each genes list: Transversally + Specific DGC/IGC, Specific DGC/IGC genes, Transversally + Enriched + Specific SRCC genes, Specific SRCC genes, Transversally + Enriched + Specific PCC-NOS genes, and Specific PCC-NOS genes.

**Table S1**. Sample size and details of sequencing methodologies used in the selected publications for analysis.

**Table S2**. Details of Clinical trial registered at ClinicalTrials.gov directed to gastric cancer and often including DGC patients.