Published in final edited form as:
Gastroenterology. 2011 August ; 141(2): 476–485.e11. doi:10.1053/j.gastro.2011.04.042.

Intrinsic Subtypes of Gastric Cancer, Based on Gene Expression Pattern, Predict Survival and Respond Differently to Chemotherapy

Iain Beehuat Tan[1,2,3], **Tatiana Ivanova**[4], **Kiat Hon Lim**[5], **Chee Wee Ong**[6], **Niantao Deng**[3], **Julian Lee**[4], **Sze Huey Tan**[19], **Jeanie Wu**[4], **Ming Hui Lee**[4], **Chia Huey Ooi**[3], **Sun Young Rha**[8], **Wai Keong Wong**[9], **Alex Boussioutas**[10], **Khay Guan Yeoh**[11], **Jimmy So**[12], **Wei Peng Yong**[6], **Akira Tsuburaya**[13], **Heike Grabsch**[14], **Han Chong Toh**[1], **Steven Rozen**[3], **Jae Ho Cheong**[15], **Sung Hoon Noh**[15], **Wei Kiat Wan**[5], **Jaffer A. Ajani**[16], **Ju-Seog Lee**[17], **Manuel Salto Tellez**[6,18], and **Patrick Tan**[3,4,6,19]

[1] Department of Medical Oncology, National Cancer Centre Singapore, Singapore

[2] National University of Singapore Graduate School of Integrative Sciences and Engineering, National University of Singapore, Singapore

[3] Duke-National University of Singapore Graduate Medical School, Singapore

[4] Division of Cellular and Molecular Research, National Cancer Centre Singapore, Singapore

[5] Department of Pathology, Singapore General Hospital, Singapore

[6] Cancer Science Institute, National University of Singapore, Singapore

[7] Clinical Trials and Epidemiological Sciences, National Cancer Centre Singapore, Singapore

[8] Department of Internal Medicine, Yonsei Cancer Centre, South Korea

[9] Department of General Surgery, Singapore General Hospital, Singapore

[10] Cancer Genomics and Biochemistry Laboratory, Peter MacCallum Cancer Centre, Australia

[11] Department of Medicine, National University Health System, Singapore

Correspondence: Patrick Tan, 8 College Road. Singapore 169857. Republic of Singapore. Tel: 65 65161924, Fax: 65 62265294, gmstanp@duke-nus.edu.sg.

Conflicts of interest: The authors disclose no conflicts.

Author Contributions:
Iain Beehuat Tan: study concept and design; acquisition of data; analysis and interpretation of data; drafting of the manuscript; statistical analysis;
Tatiana Ivanova, Sze Huey Tan, Kiat Hon Lim, Chee Wee Ong, Niantao Deng, Chia Huey Ooi, Steven Rozen, Wei Kiat Wan, Manuel Salto Tellez: acquisition of data; analysis and interpretation of data;
Julian Lee: study concept and design
Jeanie Wu, Ming Hui Lee: acquisition of data
Sun Young Rha, Wai Keong Wong, Alex Boussioutas, Jimmy So, Sung Hoon Noh, Jaffer Ajani, Ju-Seog Lee, Jae Ho Cheong: technical and material support
Khay Guan Yeoh: Obtained funding
Wei Peng Yong, Akira Tsuburaya, Heike Grabsch, Han Chong Toh: technical or material support, critical revision of the manuscript for important intellectual content
Patrick Tan: study concept and design; critical revision of the manuscript for important intellectual content; obtained funding; study supervision.

[12] Department of Surgery, National University Health System, Singapore

[13] Department of Gastrointestinal Surgery, Kanagawa Cancer Centre, Japan

[14] Dept of Pathology and Tumour Biology, Leeds Institute for Molecular Medicine, United Kingdom

[15] Department of Surgery, Yonsei University College of Medicine, South Korea

[16] Department of GI Oncology, MD Anderson Cancer Centre, USA

[17] Department of Systems Biology, Division of Cancer Medicine, MD Anderson Cancer Centre, USA

[18] Department of Pathology, National University Health System, Singapore

[19] Genome Institute of Singapore, Singapore

Abstract

BACKGROUND & AIMS—Gastric cancer (GC) is a heterogeneous disease comprising multiple subtypes that each have distinct biological properties and effects in patients. We sought to identify new, intrinsic subtypes of GC by gene expression analysis of a large panel of GC cell lines. We tested if these subtypes might be associated with differences patient survival times and responses to various standard-of-care cytotoxic drugs.

METHODS—We analyzed gene expression profiles for 37 GC cell lines to identify intrinsic GC subtypes. These subtypes were validated in primary tumors from 521 patients in 4 independent cohorts, where the subtypes were determined by either expression profiling or subtype-specific immunohistochemical markers (LGALS4, CDH17). In vitro sensitivity to 3 chemotherapy drugs (5-FU, cisplatin, oxaliplatin) was also assessed.

RESULTS—Unsupervised cell line analysis identified 2 major intrinsic genomic subtypes (G-INT and G-DIF), that had distinct patterns of gene expression. The intrinsic subtypes, but not subtypes based on Lauren's histopathologic classification, were prognostic of survival, based on univariate and multivariate analysis in multiple patient cohorts. The G-INT cell lines were significantly more sensitive to 5-FU and oxaliplatin, but more resistant to cisplatin, than the G-DIF cell lines. In patients, intrinsic subtypes were associated with survival time following adjuvant, 5-FU based therapy.

CONCLUSIONS—Intrinsic subtypes of GC, based on distinct patterns of expression, are associated with patient survival and response to chemotherapy. Classification of GC based on intrinsic subtypes might be used to determine prognosis and customize therapy.

Keywords

Microarray analysis; pharmacogenomics; mRNA; stomach; carcinogenesis

Introduction

Gastric adenocarcinoma (gastric cancer, GC) is the second leading cause of global cancer mortality and 4th most common cancer worldwide[1]. Most GC patients present with late stage disease with an overall 5-year survival of about 20%[2]. A wealth of clinical, molecular, and pathological data suggests that GC is a heterogeneous disease. Objective response rates to conventional chemotherapeutic regimens range from 20–40%[3], indicating that individual GCs can exhibit a range of responses when treated identically. Canonical oncogenic pathways such as E2F, K-RAS, p53, and Wnt/β-catenin signalling are also known to be deregulated with varying frequencies in GC[4, 5], suggesting a high degree of molecular

heterogeneity. However, despite evidence that GCs can exhibit striking inter-individual differences in disease aggressiveness[6], histopathologic features[7], and responses to therapy[8], most GC patients today are managed alike with a “one size fits all” approach resulting in markedly diverse clinical outcomes. Approaches capable of classifying heterogeneous populations of GC patients into biologically and clinically homogenous subgroups are thus urgently required, such that GC patient prognoses can be accurately predicted, and clinical decisions made based on the underlying biology of each subgroup.

Reflecting this urgency, several classification systems for GC have been reported over the decades. In 1965, Lauren described two main subtypes of GC, intestinal and diffuse, on the basis of microscopic features observed in gastric tumors[7]. Since then, several other GC histopathological classifications have since been developed, such as the WHO[9], Ming[10], and Goseki[11] systems, and more recently, molecular classifications based on immunohistochemistry, gene expression profiles[12–15], and proteomic technologies[16]. However, to date, none of these GC classification systems been shown to provide reliable independent prognostic information, nor have they been able to suggest specific treatment options for patients.

Most previously-described GC classification systems have principally focused on the characterization of primary tumors, which are known to contain many distinct cell types including tumor cells, fibroblastic/desmoplastic stroma, blood vessels, and immune cells. Given this high level of tissue complexity, we reasoned that subtle variations in these diverse cell types, both across and within-tumors, could cause differences in interpretation between observers, and ultimately pose difficulties for standardization across different centres. In this study, we pursued an alternative strategy where we initially focused not on primary GCs, but on a diverse panel of GC cell lines. We hypothesized that since cancer cell lines are devoid of other cell types besides cancer cells, any genomic differences detected in cell lines should be by nature tumor-centric and thereby “intrinsic” to the underlying biology of the GC cancer cell. Investigating a large panel of GC cell lines, we identified a genomic expression signature clearly defining two major intrinsic subgroups of GC. Importantly, we validated these intrinsic subgroups in primary tumors from 4 independent GC cohorts. The, intrinsic subtypes proved capable of providing independent prognostic information. *In vitro* and *in vivo* evidence also suggests that GCs belonging to different intrinsic subtypes may respond differently to various standard-of-care chemotherapies.

MATERIALS AND METHODS

GC Cell Lines

GC cell lines were obtained from commercial sources or collaborators and cultured as recommended (Supplementary Info). Cell proliferation assays were performed using a tetrazolium compound-based colorimetric method (Supplementary Methods).

Patient Cohorts and Clinical Characteristics

Four independent patient cohorts were analyzed (n=521). Cohort 1 (SG): 200 patients, National Cancer Centre Singapore, Singapore; Cohort 2 (AU): 70 patients, Peter MacCallum Cancer Centre, Australia; Cohort 3 (YG): 65 patients, Yonsei University, South Korea; Cohort 4 (TMA): 186 patients, National Healthcare Group, Singapore. Primary tumors were collected from institutional tissue repositories and pathology archives with approvals from Research Ethics Review Committees and signed patient informed consent. There was no pre-specified sample size calculation since this is a hypothesis generating discovery study. Clinical characteristics of the four cohorts are in accordance with REMARK guidelines[17] and presented in Table 1. Clinical information was available for all patients except 3 patients

in the SG cohort. Cohorts 1–3 (SG/AU/YG) comprise gene expression profiles of primary GCs, while cohort 4 (TMA) comprises tumor sections on a tissue microarray.

Gene Expression Profiling (GC Cell lines and Primary Tumors)

GC cell lines and patient cohorts 1 and 2 were profiled using Affymetrix Human Genome U133 plus Genechips (HG-U133 Plus 2.0, Affymetrix). Patient cohort 3 was profiled using Illumina Human-6 v2 Expression Beadchips. Primary microarray data is available in the GEO database (GSE 15460 and GSE13861).

Histology and Immunohistochemistry

Two independent pathologists (LKH, WWK) performed central pathologic review on cohort 1 samples blinded to the genomic classification. Immunohistochemical studies using LGALS4 and CDH17 antibodies were performed on a tissue microarray of 186 GC patients (cohort 4), and staining intensities determined by a pathologist blinded to the clinical data (MST). Photomicrographs, details of staining patterns and grading scales are provided in the Supplementary Information.

Bioinformatics and Statistical Analysis

Bioinformatic analyses were performed using R. Raw affymetrix datasets were preprocessed with quantile normalization using RMA (package Affy). We filtered the Gastric cancer cell line using the nsFilter function from the Genefilter package on Bioconductor. The R package LIMMA was used for feature selection. Enrichment of functional annotations in the gene expression data were performed using EASE software (http://apps1.niaid.nih.gov/david/). Statistical significance was determined using the Fisher exact score and EASE score. For patient cohorts, preprocessing of cohort 1 and 2 (Affymetrix) was performed with Refplus while preprocessing of cohort 3 (Illumina) was performed with quantile normalization and the average signal intensity used for summarization. Nearest Template Prediction[18,19] was performed using Genepattern[20]. The R package e1071 was used for support vector machine (SVM) learning and classification. Correlation with clinico-pathologic parameters and survival analysis were performed using SPSS software (version 16, Chicago). Survival curves were estimated using the Kaplan-Meier method and the duration of survival was measured from the date of surgery to date of death or last follow-up visit. Cancer-specific survival (CSS) was used as the outcome metric, with deaths due to cancer was regarded as an event. Patients who are still alive, died from other causes or lost to follow-up at time of analysis were censored at their last date of follow up. Univariable and multivariable survival analyses were performed using the Cox proportional hazards regression model. The test of interaction between the genomic subtypes and therapy was performed with the null hypothesis of treatment equivalence within the subtypes and the alternative hypothesis was of differential treatment efficacy in the subtypes[21]. Two-sided p-values less than 0.05 were considered statistically significant.

Further details of bioinformatics and statistical analyses are provided in the Supplementary Information.

RESULTS

Genomic Analysis of GC Cell Lines Reveals Two Major Intrinsic Subclasses

We performed gene expression profiling for a panel of 37 GC cell lines. To identify pervasive and thereby “intrinsic” gene expression differences across the cell lines, we analyzed the expression data using four different unsupervised and unbiased clustering techniques (hierarchical clustering, silhouette plot (SP) analysis[22], nonnegative matrix factorization (NMF)[23], and principal components analysis (PCA)). Two major intrinsic

subtypes were identified by hierarchical clustering (Figure 1A). The robustness of the subtypes was further verified by SP, NMF, and PCA analysis (Figure 1B and Supplementary Figure S1). For reasons that will become apparent in later sections, we will henceforth refer to these two intrinsic subtypes as Genomic intestinal (G-INT) and Genomic Diffuse (G-DIF).

The Intrinsic Subtypes are Associated with Highly Distinctive Gene Expression Patterns

To analyze gene expression differences between the intrinsic subtypes, we used LIMMA (Linear models for microarray data)[24], a modified t-test incorporating the Benjamini Hochberg multiple correction technique. We identified a genomic signature of 171 genes distinguishing the G-INT and G-DIF intrinsic subtypes (FDR<0.002; Figure 1C and Supplementary Table S1). We attempted to refine this signature by searching for potentially redundant features among the 171 gene set. Comparing the correlation coefficients of the 171 genes to one another, we found that only 2 of the 171 genes exceeded a pre-defined correlation threshold of 0.88. Given this lack of redundancy, we performed further analysis using the entire 171 gene set. Using Expression Analysis Systematic Explorer (EASE) (http://david.abcc.ncifcrf.gov/ease/ease.jsp), we identified enriched biological themes within genes expressed in either subtype Genes up-regulated in the G-INT subtype were related to carbohydrate and protein metabolism (*FUT2*) and cell adhesion (*LGALS4*, *CDH17*), while cell proliferation (*AURKB*) and fatty acid metabolism (*ELOVL5*) functional annotations were enriched in the G-DIF subtype (within system FDR < 0.01, Supplementary Methods and Supplementary Table S2). The two intrinsic subtypes, G-INT and G-DIF, are thus associated with highly distinctive gene expression patterns and biological pathways.

The Intrinsic Subtypes are Recurrently Observed in Primary Tumors

We mapped the intrinsic 171-gene genomic signature onto primary tumors in two independent cohorts of GC patients (SG and AU), collectively totaling 270 patients. We used 2 classification algorithms (Nearest Template Prediction and a support vector machine classifier). Concordance between the 2 classification systems (SVM and NTP) was 94–96% in the SG and AU cohorts with 88% of samples identified by NTP at an FDR of < 0.05. These results suggest the 171 gene set can robustly classify primary tumors into G-INT and G-DIF sub-classes. Due to its methodological simplicity and applicability to single samples without requiring a corresponding training dataset[19], the NTP classifications were used for subsequent analyses. Specifically, 114 samples in the SG cohort and 38 samples in the AU cohort were classified as G-INT (Figures 2A&B) (Supplementary Table S3).

The Intrinsic Subtypes are Partially Associated with Lauren's Histopathologic Classification

We next investigated their associations with clinical-pathologic parameters. The intrinsic subtypes were found to be significantly associated with Lauren's intestinal and diffuse subtypes respectively in the SG (p=0.002) and AU cohorts (p=0.003), hence their name (G-INT and G-DIF). Besides Lauren's, the intrinsic subtypes were also related to tumor grade (Supplementary Table S3).

Although we named the intrinsic subtypes G-INT and G-DIF due to their associations with Lauren's histopathology, it is worth noting that the overall concordance between the intrinsic genomic subtypes and Lauren's histopathology was only 64%. Thus, the two classifications should more appropriately be regarded as related but distinct. Specifically, 91 of 134 Lauren's intestinal cases were classified at G-INT, and 64 of 106 Lauren's diffuse cases were classified as G-DIF (Figures 2A&B). These discrepancies are unlikely to be due to inter-pathologist differences alone, as pathologic review in the SG cohort had been performed by 2 independent pathologists blinded to the genomic classification

(Representative H&E slides of discordant tumors are also presented in figures 2C & 2D). Rather, it is possible that the intrinsic genomic signature may capture salient features of the tumor that are less obvious to discern by light microscopy.

The Intrinsic Subtypes are Independently Prognostic of Patient Survival

We next asked if the intrinsic subtypes could be used to define patient subgroups exhibiting differences in survival outcomes. Using cancer-specific survival as the outcome metric, patients with G-DIF cancers had worse survival outcomes compared to patients with G-INT tumors in the SG and AU cohorts (cohort 1: HR 1.78, 95%CI: 1.19–2.64, p=0.004; cohort 2: HR 1.73, 95%CI: 0.92–3.26, p=0.09) and also in a combined analysis (HR: 1.79, 95% CI: 1.28–2.51, p=0.001, Figure 3A). In contrast, Lauren's classification was not prognostic (p=0.23). Further supporting the prognostic relevance of the intrinsic subtypes, in discordant cases, patients with G-INT but diffuse type cancers exhibited superior survival compared to patients with G-DIF but intestinal type cancers (HR 1.83, 95%CI: 1.02–3.30, p=0.04, Figure 3B).

In a multivariate analysis (Table 2), the intrinsic subtypes remained prognostic ($p<0.001$) even after accounting for other interacting factors such as Lauren's classes and grade. The intrinsic subtypes were also prognostic after accounting for other variables that were also prognostic in univariate analysis (stage, margin status and gender) (p=0.005).

The Intrinsic Subtypes are Prognostic in an Independent Patient Cohort Profiled by a Different Microarray Platform

To further determine the general applicability of the intrinsic subclasses, we then applied the intrinsic genomic signature onto a third GC patient cohort (YG) profiled on a different microarray platform (Illumina Human-6 v2 Expression Beadchip). Of the 65 patients, 35 were classified as G-INT by NTP (heatmap provided in Supplementary Figure S3A). Similar to the SG and AU cohorts, patients with G-INT tumors had superior overall survival compared to patients with G-DIF tumors in the YG cohort (HR 3.3, 95%CI: 1.03–10.53, p=0.04) (Supplementary Figure S3B), while Lauren's classes was not prognostic (p=0.23).

G-INT Patients Identified by Immunohistochemical Markers Exhibit Improved Survival Outcomes

To assess if a panel of immunohistochemical markers might also be used to identify the intrinsic subtypes and its relation to survival outcomes, we then analyzed an independent tissue microarray (TMA) cohort (cohort 4) of 186 GC patients. We selected 2 G-INT markers (LGALS4 and CDH17) meeting the criteria of high gene expression in G-INT cell lines and tumors, and for which commercial immunohistochemical markers were available. We classified the TMA tumors based on their intensity of LGALS4 and CDH17 staining (CDH17 (> 1+) and LGALS4 (>2+)), using intensity cutoffs determined by a pathologist blinded to the clinical data (see Supplementary Information). To confidently distinguish between G-INT and G-DIF cancers, we specifically compared the 2-marker positive group (G-INT) to the 2-marker negative group (G-DIF). Among the 186 tumors, 75 were classified as G-INT (both markers positive), 44 as G-DIF (neither marker positive) and 67 were equivocal (one marker positive). Patients with G-DIF tumors classified by IHC exhibited worse outcomes than G-INT tumors classified by IHC (Hazard ratio, adjusted for stage: 1.95, 95%CI: 1.13–3.38, p=0.02) (Supplementary Figure S4), while Lauren was once again not prognostic (p=0.33).

The Intrinsic Subtypes Exhibit Distinct *in vitro* Responses to Chemotherapy

Of the 37 cell lines, 28 cell lines (11 G-INT and 17 G-DIF) had growth characteristics suitable for i*n vitro* drug sensitivity testing (see Supplementary Methods). 5-FU, oxaliplatin and cisplatin are drugs presently employed in the adjuvant and 1st line palliative treatment of GC. We treated the 28 cell lines with increasing concentrations of these drugs. G-INT cell lines were significantly more sensitive to 5-FU (p=0.04) and oxaliplatin (p=0.02) *in vitro*, while G-DIF cell lines were more sensitive to cisplatin (p=0.03) (Figure 4, see legend for mean drug concentrations). The *in vitro* dosages used are comparable to therapeutic ranges observed in human patients based on pharmacokinetic analysis[25–27] (grey lines in Figure 4). These results point to differential *in vitro* sensitivities of G-INT cell lines to 5-FU and oxaliplatin, and G-DIF cell lines to cisplatin.

G-INT Patients may Derive Differential Benefit from 5-FU Treatment

Information regarding use of adjuvant 5 Fluorouracil chemoradiation were available from 2 gene expression cohorts (1 &2) and the TMA cohort (cohort 4). Decisions regarding adjuvant therapy in these cohorts were based upon existing knowledge at the point of diagnosis, patient's general health status, risk factors for relapse especially disease stage, treatment related toxicities and patient preference.

Patients with advanced stage disease were more likely to receive adjuvant treatment (p=0.03), however no significant differences were observed in prescribing 5-FU therapy between the intrinsic subtypes either across all stages (p=0.27) or within each stage (p~0.4–0.8) (Supplementary Table S3). We performed a statistical test for interaction that was specifically adjusted for stage, to evaluate if the intrinsic subtypes might exhibit differential benefit with 5-FU chemoradiation in our patient cohorts.

We observed a significant interaction between the intrinsic subtypes and benefit with 5-FU based chemoradiation (Table 3), suggesting that patients with G-INT tumors may derive differential benefit from adjuvant 5-FU based therapy. Specifically, the test for interaction by Cox proportional hazards regression was p=0.002 (combined analysis), gene expression (p=0.03) and TMA cohorts (p=0.02). The stage adjusted hazard ratio of death due to cancer for surgery alone compared to adjuvant 5-FU therapy was 1.68 (p=0.06 for G-INT tumors and 0.90 (p=0.67) for G-DIF tumors. Table 3 presents the interactions for the combined analysis, while the gene expression and TMA cohorts are separately presented in Supplementary Table 4.

DISCUSSION

In this study, we report the discovery of two genomic subtypes of GC using profiles initially derived from GC cell lines. Since cancer cell lines are devoid of stroma, vasculature and immune cells, we reasoned that comparing signatures between cell lines would be more likely to reflect intrinsic differences between tumor cells, minimizing potentially confounding effects from neighboring non-cancer tissues. The validity of the cell line based approach is supported by similar studies in other solid tumors[28, 29], where major patterns of tumor heterogeneity have also been shown to be recapitulated by cell lines. We acknowledge that this approach does have caveats - for example, it may miss rare subtypes represented by only one or two lines, or subtypes for which it is difficult to derive immortalized cell lines. The use of cell lines is also unlikely to fully recapitulate the biology of tumor/microenvironment interactions, which can also influence disease prognosis and treatment response. Nevertheless, we believe the intrinsic subtypes discovered in this study represents an important first step in establishing a basic foundational taxonomy of GC, one to which additional layers of genomic complexity can be subsequently added.

One notable aspect of our study was that unlike previous comparative molecular studies for GC[30, 31], we chose to use unsupervised approaches for subclass discovery. We did so because a) the major distinctions in the molecular heterogeneity of GC might be unrelated to presently known classification systems or phenotypes, and b) using current classification systems, reproducibility among pathologists is only about 70%[32, 33] and we were concerned that this lack of inter-observer concordance might compromise supervised analysis. Testing several different prediction algorithms, we confirmed that the intrinsic subtypes exhibited stable and reproducible classification performance in cell lines and primary tumors, thus demonstrating that the subtypes are statistically robust. Using a strict filtering criteria (FDR<0.002), we then identified a genomic classifier of 171 genes exhibiting differential regulation between the subtypes. Biological curation of the classifier confirmed that the intrinsic subtypes are associated with very different gene expression features, cellular processes and biological pathways. Taken collectively, these finding are consistent with the intrinsic subtypes being very distinct and possibly representing distinct lineages.

The clinical relevance of the intrinsic subclasses is supported by the finding that it can act as an independent predictor of clinical survival in multiple patient cohorts, even after controlling for tumor stage. In this regard, it is interesting to contrast the intrinsic subclasses against Lauren's histological subtypes (intestinal and diffuse), to which the intrinsic classes are partially associated. Intestinal cancers are classically characterized by glandular differentiation on a background of gastric atrophy or intestinal metaplasia, while diffuse cancers typically appear as rows of single mononuclear "signet ring" cells with little cell adhesion. These apparently distinct features, however, are not always discernable in clinical samples where inter-observer variation and unclassifiable or "mixed" subtypes are not uncommonly reported. In our study, patients stratified by Lauren's histopathology did not exhibit significantly different survival outcomes, while patients discordant between the intrinsic subclasses and Lauren's exhibited survival patterns that support the intrinsic genomic taxonomy. Our data thus suggests that the intrinsic subclasses can provide information about the predominant lineage in GC samples that may not be precisely distinguished by morphology, and that this information is clinically relevant.

Besides gene expression, we also employed two genes in the classifier (LGALS4 and L1-Cadherin (CDH17)) as immunohistochemical markers for the G-INT intrinsic subtype. LGALS4 and CDH17 have been previously reported to be differentially regulated across subsets of gastric tumors[14] and cell line[34], and expressed in intestinal metaplasia[35, 36]. CDH17 was recently reported as a prognostic factor in early-stage GC[37], a marker of poor prognosis in another study[36], and a potential therapeutic target in experimental models[38]. In our study, we specifically chose to compare the 2-marker positive group to the 2-marker negative group to confidently distinguish between the G-INT and G-DIF cancers. Interestingly, we have also observed that the one-third of 1-marker positive patients also appeared to exhibit an improved survival trend compared to the 2-marker negative group (CDH17, p=0.08 adjusted for stage; LGALS4, p=0.07 adjusted for stage). This result suggests that some of the 1-marker positive cancers may also be G-INT cancers as well (Supplementary Figure S4). At this stage, it is likely premature to definitely conclude that LGALS4 and CDH17 necessarily represent the best markers for the G-INT subtype. Identifying the optimal combination of immunohistochemical markers to distinguish between the intrinsic subtypes should thus be a focus for future research.

In vitro, G-INT lines were more sensitive to 5-FU and oxaliplatin than G-DIF cell lines, but were also more resistant to cisplatin. Although the absolute magnitude of these *in vitro* differential sensitivities appears modest (3–5 fold), these differences could still prove clinically meaningful given the relatively small therapeutic windows associated with cytotoxic chemotherapy. Indeed, supporting the clinical relevance of these differences, we

observed a significant interaction between the intrinsic subtypes and differential benefit from adjuvant 5-FU therapy in retrospective patient cohorts (Table 3 and Supplementary Table S4). These findings suggest that in addition to patient prognosis, the intrinsic subtypes could potentially be used to guide treatment selection. However, we also emphasize that our results at this stage should be interpreted in the context of an early discovery study, requiring further validation efforts. For example, while we observed a trend of stage-adjusted survival benefit in G-INT patients due to 5-FU adjuvant therapy, this benefit was not statistically significant relative to surgery alone (Table 3 (combined, p=0.06) and Supp Table S4 (gene expression, p=0.18; TMA, p=0.11). There may be multiple reasons for this – compared to large Phase III trials such as INT-0116[39], which evaluated adjuvant 5-FU based chemoradiation in GC, our patient cohorts are smaller in size, non-randomized, and being retrospective in nature, quality of surgery and treatment regimens were also not strictly enforced. As such, the interaction of intrinsic subclass with treatment response should definitely be further validated in additional studies (see below). Nevertheless, given the association the intrinsic subtypes with Lauren's, it is worth noting that in INT-0116[40], a ten-year update subgroup analyses revealed that all GC subsets benefited from 5-FU therapy *except* for cases with diffuse histology. Moreover, in JCOG 9912[41] which established S-1 monotherapy as a first-line palliative chemotherapy option in Japan, benefit of irinotecan/cisplatin over 5-FU based monotherapy was observed in diffuse but not intestinal GCs. Our findings in this study are thus consistent with exploratory subgroup analysis of these two large GC clinical trials.

Our hypothesis-generating study requires further validation studies. Most immediately, the interaction between the intrinsic subtypes and treatment outcomes will be retrospectively validated in larger patient cohorts including tissue from a randomized phase 3 Japanese trial in the first line palliative treatment of advanced gastric cancer (PI: Akira Tsuburaya). Future prospective research efforts will focus on validating the predictive value of the GC intrinsic subclasses in the metastatic setting. In most countries, combination therapy with a fluropyrimidine/platinum combination therapy with or without epirubicin remains standard first line palliative chemotherapy**(3)**. We have thus initiated a phase 2 study, Genomic Guided Gastric cancer (3G) trial with mandatory pre-treated tumor biopsies from which patients will be allocated to S1/oxaliplatin or S1/cisplatin based upon their genomic profiles (NCT01100801, http://clinicaltrialsfeeds.org).

In conclusion, the quest for a clinically meaningful GC taxonomy has motivated several attempts to classify GC on the basis of clinical, histologic, and molecular features. With the exception of tumor stage however, few previously-proposed classification strategies have been shown to provide additional prognostic or predictive value. It is possible that the intrinsic subtypes reported here in this study may represent a promising step in establishment of a clinically relevant genomic taxonomy of GC.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Grant Support: Grants to PT from BMRC 05/1/31/19/423, NMRC grant TCR/001/2007, and a Duke-NUS core grant. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. Cohort 3 was funded by NIH R01 to JSL. JAA is supported by Program grant from UT MDACC, Park, Dallas, Cantu, and Smith families, Kevin and Sultan funds, and Rivercreek and Schecter Private foundations.

Abbreviations

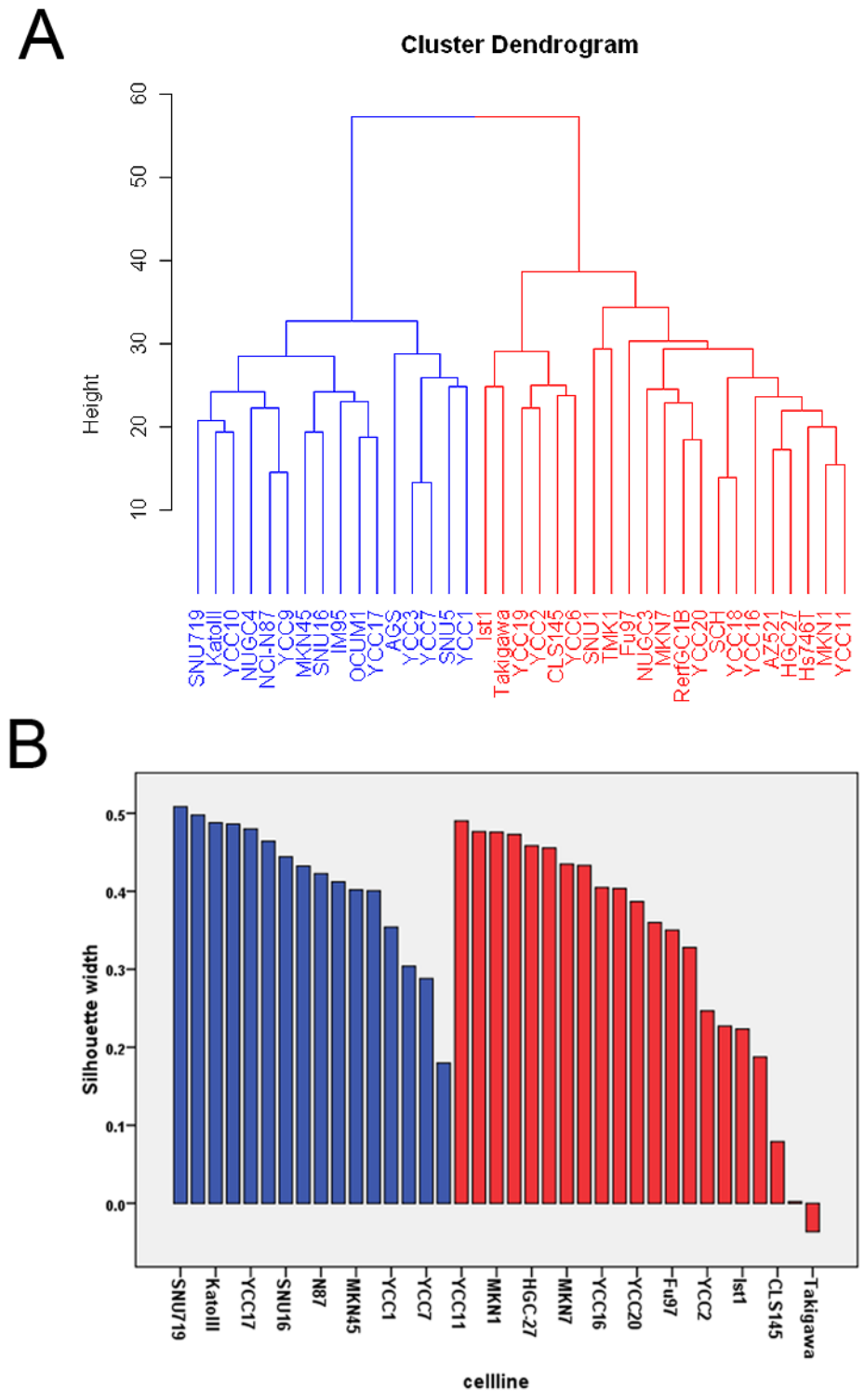
5-FU	5-Fluorouracil
cDNA	complementary Deoxyribonucleic acid
CI	confidence interval
GC	Gastric Cancer
HR	Hazard Ratio
SVM	Support Vector Machine

References

1. Kamangar F, Dores GM, Anderson WF. Patterns of cancer incidence, mortality, and prevalence across five continents: defining priorities to reduce cancer disparities in different geographic regions of the world. J Clin Oncol. 2006; 24:2137–50. [PubMed: 16682732]
2. Hartgrink HH, Jansen EP, van Grieken NC, van de Velde CJ. Gastric cancer. Lancet. 2009; 374:477–90. [PubMed: 19625077]
3. Clinical Guidelines in Oncology: Gastric cancer. National Comprehensive Cancer Network; 2009.
4. Suzuki T, Yasui W, Yokozaki H, Naka K, Ishikawa T, Tahara E. Expression of the E2F family in human gastrointestinal carcinomas. Int J Cancer. 1999; 81:535–8. [PubMed: 10225440]
5. Cheng XX, Wang ZC, Chen XY, Sun Y, Kong QY, Liu J, Gao X, Guan HW, Li H. Frequent loss of membranous E-cadherin in gastric cancers: A cross-talk with Wnt in determining the fate of beta-catenin. Clin Exp Metastasis. 2005; 22:85–93. [PubMed: 16132582]
6. Chau I, Ashley S, Cunningham D. Validation of the Royal Marsden hospital prognostic index in advanced esophagogastric cancer using individual patient data from the REAL 2 study. J Clin Oncol. 2009; 27:e3–4. [PubMed: 19470917]
7. Lauren P. The Two Histological Main Types of Gastric Carcinoma: Diffuse and So-Called Intestinal-Type Carcinoma. an Attempt at a Histo-Clinical Classification. Acta Pathol Microbiol Scand. 1965; 64:31–49. [PubMed: 14320675]
8. Wagner AD, Grothe W, Haerting J, Kleber G, Grothey A, Fleig WE. Chemotherapy in advanced gastric cancer: a systematic review and meta-analysis based on aggregate data. J Clin Oncol. 2006; 24:2903–9. [PubMed: 16782930]
9. Jass JR, Sobin LH, Watanabe H. The World Health Organization's histologic classification of gastrointestinal tumors. A commentary on the second edition. Cancer. 1990; 66:2162–7. [PubMed: 2171747]
10. Ming SC. Gastric carcinoma. A pathobiological classification. Cancer. 1977; 39:2475–85. [PubMed: 872047]
11. Goseki N, Takizawa T, Koike M. Differences in the mode of the extension of gastric cancer classified by histological type: new histological classification of gastric carcinoma. Gut. 1992; 33:606–12. [PubMed: 1377153]
12. Tay ST, Leong SH, Yu K, Aggarwal A, Tan SY, Lee CH, Wong K, Visvanathan J, Lim D, Wong WK, Soo KC, Kon OL, Tan P. A combined comparative genomic hybridization and expression microarray analysis of gastric cancer reveals novel molecular subtypes. Cancer Res. 2003; 63:3309–16. [PubMed: 12810664]
13. Kim B, Bang S, Lee S, Kim S, Jung Y, Lee C, Choi K, Lee SG, Lee K, Lee Y, Kim SS, Yeom YI, Kim YS, Yoo HS, Song K, Lee I. Expression profiling and subtype-specific expression of stomach cancer. Cancer Res. 2003; 63:8248–55. [PubMed: 14678982]
14. Chen X, Leung SY, Yuen ST, Chu KM, Ji J, Li R, Chan AS, Law S, Troyanskaya OG, Wong J, So S, Botstein D, Brown PO. Variation in gene expression patterns in human gastric cancers. Mol Biol Cell. 2003; 14:3208–15. [PubMed: 12925757]

15. Boussioutas A, Li H, Liu J, Waring P, Lade S, Holloway AJ, Taupin D, Gorringe K, Haviv I, Desmond PV, Bowtell DD. Distinctive patterns of gene expression in premalignant gastric mucosa and gastric cancer. Cancer Res. 2003; 63:2569–77. [PubMed: 12750281]
16. Lee HS, Cho SB, Lee HE, Kim MA, Kim JH, Park do J, Kim JH, Yang HK, Lee BL, Kim WH. Protein expression profiling and molecular classification of gastric cancer by the tissue array method. Clin Cancer Res. 2007; 13:4154–63. [PubMed: 17634543]
17. McShane LM, Altman DG, Sauerbrei W, Taube SE, Gion M, Clark GM. Reporting recommendations for tumor marker prognostic studies (REMARK). J Natl Cancer Inst. 2005; 97:1180–4. [PubMed: 16106022]
18. Hoshida Y, Villanueva A, Kobayashi M, Peix J, Chiang DY, Camargo A, Gupta S, Moore J, Wrobel MJ, Lerner J, Reich M, Chan JA, Glickman JN, Ikeda K, Hashimoto M, Watanabe G, Daidone MG, Roayaie S, Schwartz M, Thung S, Salvesen HB, Gabriel S, Mazzaferro V, Bruix J, Friedman SL, Kumada H, Llovet JM, Golub TR. Gene expression in fixed tissues and outcome in hepatocellular carcinoma. N Engl J Med. 2008; 359:1995–2004. [PubMed: 18923165]
19. Hoshida Y. Nearest template prediction: a single-sample-based flexible class prediction with confidence assessment. PLoS One. 2010; 5:e15543. [PubMed: 21124904]
20. Reich M, Liefeld T, Gould J, Lerner J, Tamayo P, Mesirov JP. GenePattern 2.0. Nat Genet. 2006; 38:500–1. [PubMed: 16642009]
21. Simon R. Patient subsets and variation in therapeutic efficacy. Br J Clin Pharmacol. 1982; 14:473–82. [PubMed: 7138732]
22. Rousseeuw PJ. Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. J Comput Appl Math. 1987; 20:53–65.
23. Lee DD, Seung HS. Learning the parts of objects by non-negative matrix factorization. Nature. 1999; 401:788–91. [PubMed: 10548103]
24. Smyth GK. Linear models and empirical Bayes methods for assessing differential expression in microarray experiments. Statistical Applications in Genetics and Molecular Biology. 2004; 3:Article 3.
25. Saif MW, Choma A, Salamone SJ, Chu E. Pharmacokinetically guided dose adjustment of 5-fluorouracil: a rational approach to improving therapeutic outcomes. J Natl Cancer Inst. 2009; 101:1543–52. [PubMed: 19841331]
26. Ikeda K, Terashima M, Kawamura H, Takiyama I, Koeda K, Takagane A, Sato N, Ishida K, Iwaya T, Maesawa C, Yoshinari H, Saito K. Pharmacokinetics of cisplatin in combined cisplatin and 5-fluorouracil therapy: a comparative study of three different schedules of cisplatin administration. Jpn J Clin Oncol. 1998; 28:168–75. [PubMed: 9614438]
27. Graham MA, Lockwood GF, Greenslade D, Brienza S, Bayssas M, Gamelin E. Clinical pharmacokinetics of oxaliplatin: a critical review. Clin Cancer Res. 2000; 6:1205–18. [PubMed: 10778943]
28. Neve RM, Chin K, Fridlyand J, Yeh J, Baehner FL, Fevr T, Clark L, Bayani N, Coppe JP, Tong F, Speed T, Spellman PT, DeVries S, Lapuk A, Wang NJ, Kuo WL, Stilwell JL, Pinkel D, Albertson DG, Waldman FM, McCormick F, Dickson RB, Johnson MD, Lippman M, Ethier S, Gazdar A, Gray JW. A collection of breast cancer cell lines for the study of functionally distinct cancer subtypes. Cancer Cell. 2006; 10:515–27. [PubMed: 17157791]
29. Hoshida Y, Toffanin S, Lachenmayer A, Villanueva A, Minguez B, Llovet JM. Molecular classification and novel targets in hepatocellular carcinoma: recent advancements. Semin Liver Dis. 2010; 30:35–51. [PubMed: 20175032]
30. Jinawath N, Furukawa Y, Hasegawa S, Li M, Tsunoda T, Satoh S, Yamaguchi T, Imamura H, Inoue M, Shiozaki H, Nakamura Y. Comparison of gene-expression profiles between diffuse- and intestinal-type gastric cancers using a genome-wide cDNA microarray. Oncogene. 2004; 23:6830–44. [PubMed: 15273739]
31. Meireles SI, Cristo EB, Carvalho AF, Hirata R Jr, Pelosof A, Gomes LI, Martins WK, Begnami MD, Zitron C, Montagnini AL, Soares FA, Neves EJ, Reis LF. Molecular classifiers for gastric cancer and nonmalignant diseases of the gastric mucosa. Cancer Res. 2004; 64:1255–65. [PubMed: 14973074]

32. Palli D, Bianchi S, Cipriani F, Duca P, Amorosi A, Avellini C, Russo A, Saragoni A, Todde P, Valdes E, et al. Reproducibility of histologic classification of gastric cancer. Br J Cancer. 1991; 63:765–8. [PubMed: 2039701]
33. Shibata A, Longacre TA, Puligandla B, Parsonnet J, Habel LA. Histological classification of gastric adenocarcinoma for epidemiological research: concordance between pathologists. Cancer Epidemiol Biomarkers Prev. 2001; 10:75–8. [PubMed: 11205493]
34. Ji J, Chen X, Leung SY, Chi JT, Chu KM, Yuen ST, Li R, Chan AS, Li J, Dunphy N, So S. Comprehensive analysis of the gene expression profiles in human gastric cancer cell lines. Oncogene. 2002; 21:6549–56. [PubMed: 12226758]
35. Dong W, Yu Q, Xu Y. Altered expression of a Li-cadherin in gastric cancer and intestinal metaplasia. Dig Dis Sci. 2007; 52:536–42. [PubMed: 17226075]
36. Ito R, Oue N, Yoshida K, Kunimitsu K, Nakayama H, Nakachi K, Yasui W. Clinicopathological significant and prognostic influence of cadherin-17 expression in gastric cancer. Virchows Arch. 2005; 447:717–22. [PubMed: 16044349]
37. Lee HJ, Nam KT, Park HS, Kim MA, Lafleur BJ, Aburatani H, Yang HK, Kim WH, Goldenring JR. Gene expression profiling of metaplastic lineages identifies CDH17 as a prognostic marker in early stage gastric cancer. Gastroenterology. 2010; 139:213–25. e3. [PubMed: 20398667]
38. Liu QS, Zhang J, Liu M, Dong WG. Lentiviral-mediated miRNA against liver-intestine cadherin suppresses tumor growth and invasiveness of human gastric cancer. Cancer Sci. 2010; 101:1807–12. [PubMed: 20500517]
39. Macdonald JS, Smalley SR, Benedetti J, Hundahl SA, Estes NC, Stemmermann GN, Haller DG, Ajani JA, Gunderson LL, Jessup JM, Martenson JA. Chemoradiotherapy after surgery compared with surgery alone for adenocarcinoma of the stomach or gastroesophageal junction. N Engl J Med. 2001; 345:725–30. [PubMed: 11547741]
40. Macdonald JS. Chemoradiation of resected gastric cancer: A 10 year follow-up of the phase III trial INT0116 (SWOG 9008). J Clin Oncol. 2009; 27 abst 4515.
41. Boku N, Yamamoto S, Fukuda H, Shirao K, Doi T, Sawaki A, Koizumi W, Saito H, Yamaguchi K, Takiuchi H, Nasu J, Ohtsu A. Fluorouracil versus combination of irinotecan plus cisplatin versus S-1 in metastatic gastric cancer: a randomised phase 3 study. Lancet Oncol. 2009; 10:1063–9. [PubMed: 19818685]



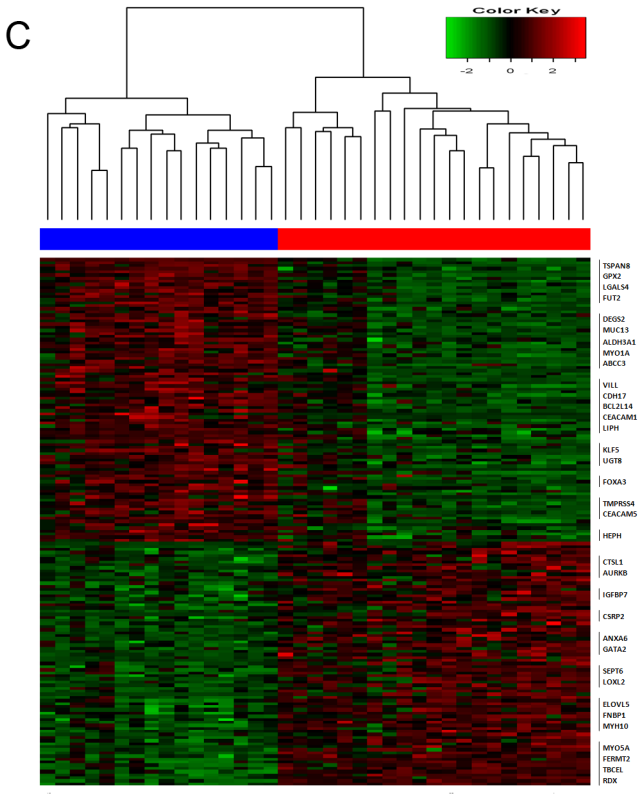
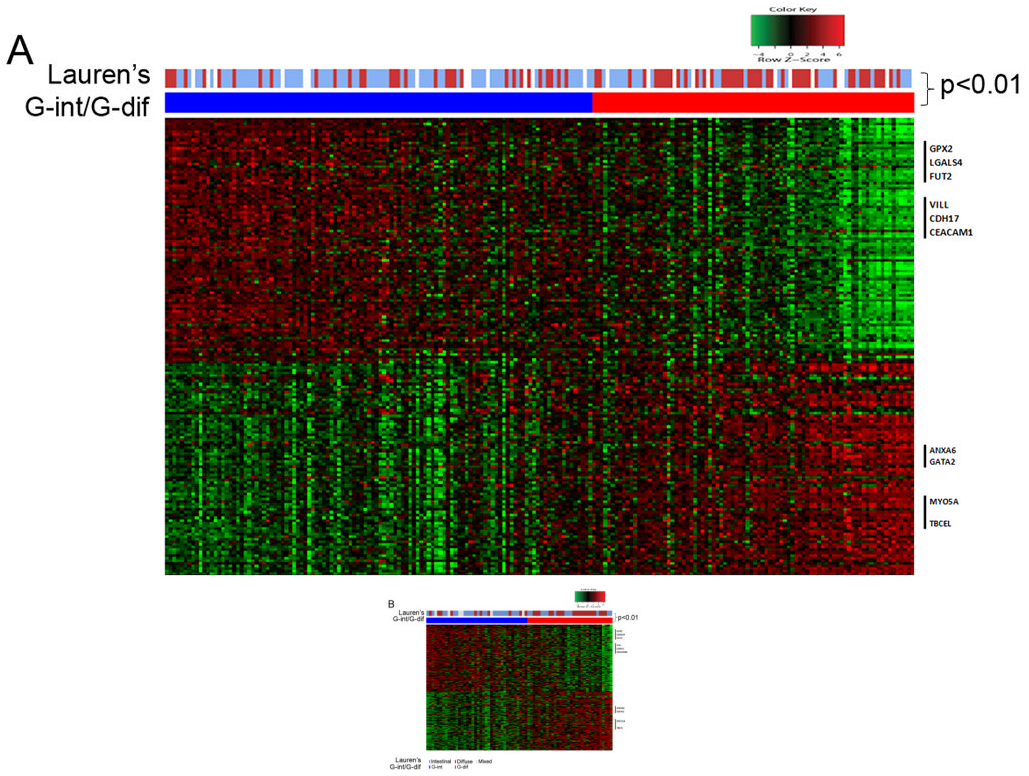
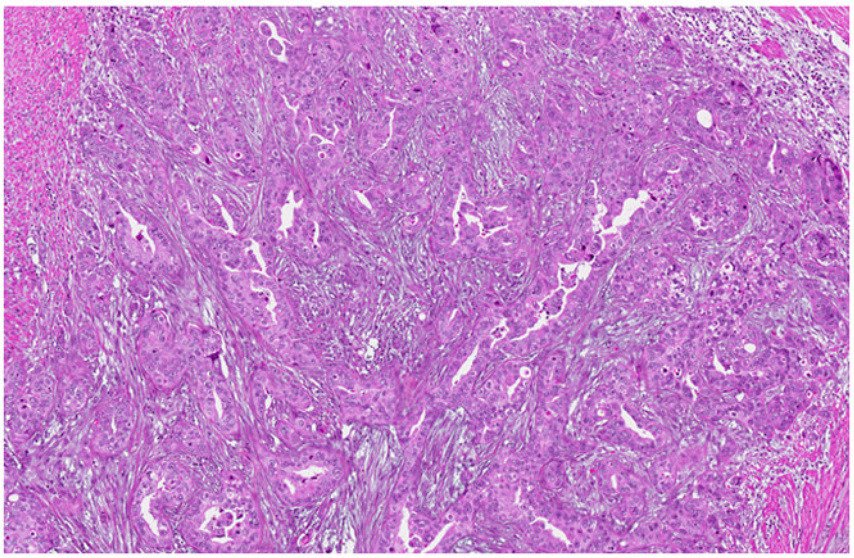


Figure 1. Unsupervised clustering of GCCLs reveals 2 major intrinsic subtypes

(A) Hierarchical dendrogram depicting clustering of 37 GCCLs into G-INT (blue) and G-DIF (red); height: squared euclidean distances between cluster means. (B) Silhouette widths of individual cell lines when classified in 2 clusters. Silhouette width: a measure for each sample of membership of within it s own class against that of another class. (C) heat map of expression of 171 genes arranged by hierarchal clustering of cell lines (columns) and expression difference for each gene between G-INT and G-DIF as measured by the t-test statistic (rows).



C



D

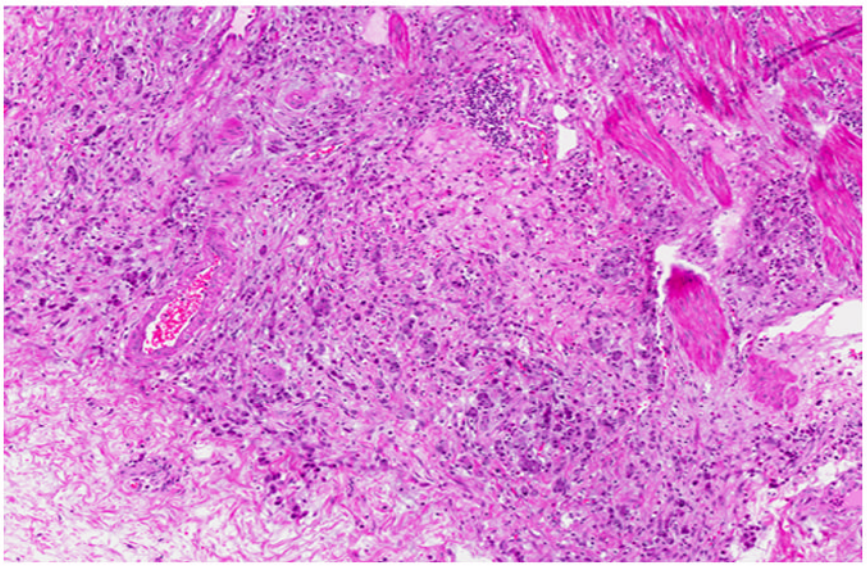
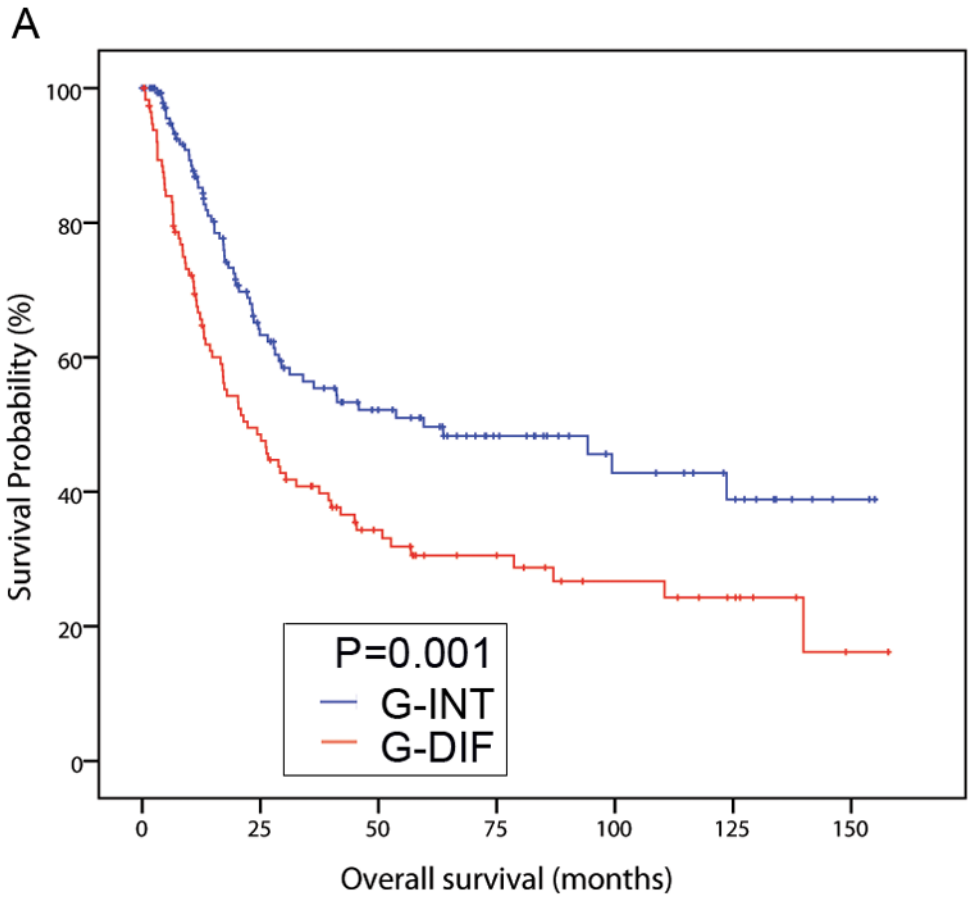


Figure 2. Associations of intrinsic subtypes with Lauren's classification in primary GCs

Heat map of gene expression in (A) SG and (B) AU cohorts arranged by strength of association (columns) and expression difference for each gene between G-INT and G-DIF as measured by the t-test statistic (rows). 1st row label: Laurens class; light blue: intestinal, brown: diffuse, white: mixed. 2nd row label: Intrinsic classes: blue: G-INT, red: G-DIF. C. Representative H&E of (C) G-DIF/intestinal cancer and (D) G-INT/Diffuse cancer.



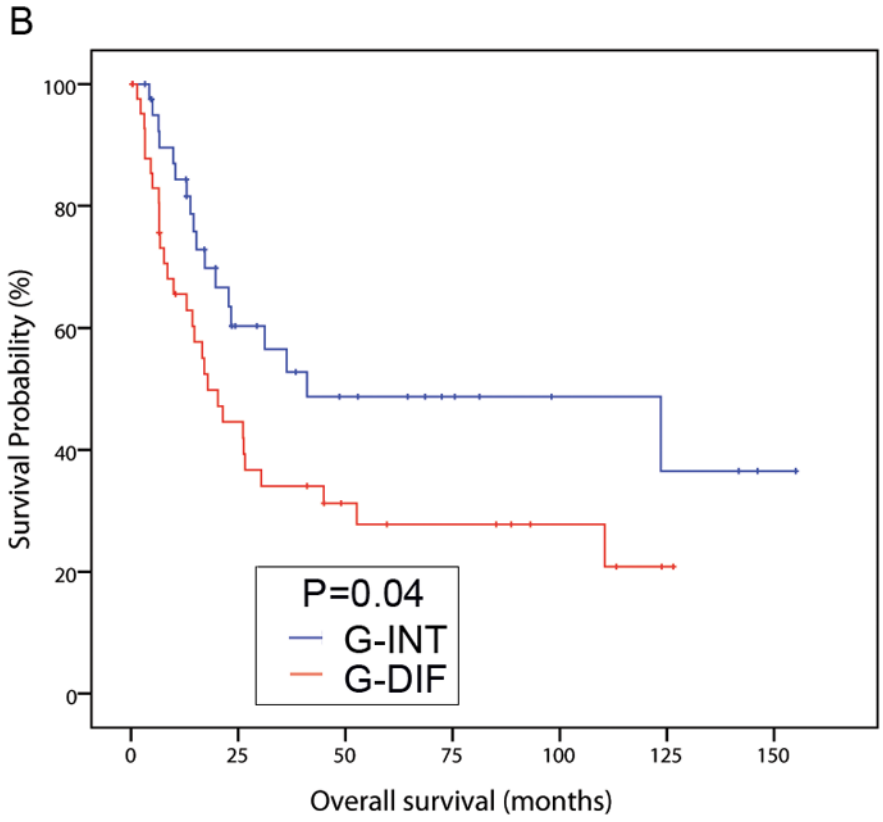


Figure 3. Intrinsic genomic subclasses are prognostic

Kaplan-Meier plots of survival in (A) all patients (HR: 1.79, 95% CI: 1.28–2.51, p=0.001) and (B) when the intrinsic classification and Lauren's classes are discordant (HR 1.83, 95%CI: 1.02–3.30, p=0.04).

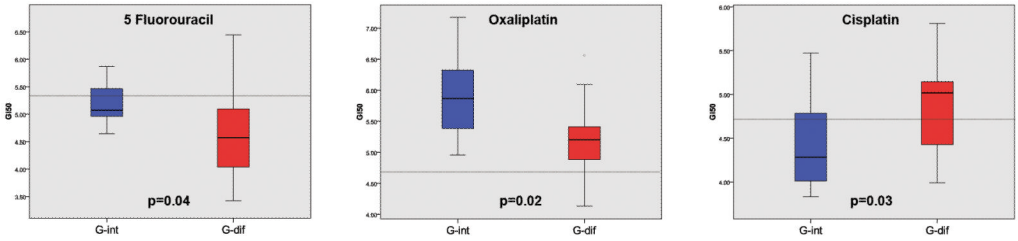


Figure 4. ***In vitro*** **chemosensitivity of G-INT and G-DIF cell lines**

GI-50 values of 11 G-INT and 17 G-DIF cell lines upon treatment with 5-FU, oxaliplatin and cisplatin. GI-50s refer to the drug concentration at which 50% growth inhibition is achieved. (y-axis: GI-50 enumerated in negative $\log_{10}$). The horizontal grey lines represent the therapeutic concentration patients are exposed to based on pharmacokinetic data[25–27]. Mean GI-50 concentrations for G-INT and G-DIF cell lines respectively: 5FU: 5.20 μM, 23.22 μM; Cisplatin: 38.61 μM, 13.35 μM; Oxaliplatin: 1.33 μM, 5.49 μM.

Table 1

Clinical Characteristics of Patient Cohorts

Clinical information is available for all but 3 patients in the SG cohort. Median follow-up for patients still alive for the 4 cohorts are 33, 56, 39 and 36 months respectively.

	SG (n=197)	AU (n=70)	YG (n=65)	TMA (n=186)
Age				
range	23–92	32–85	32–83	31–87
mean, S.D	64.6, 13.1	65.5, 12.5	61.0, 11.5	65.8, 11.7
Gender				
Male	128	48	46	128
Female	69	22	19	58
Lauren's				
Intestinal	100	34	22	97
Diffuse	76	30	31	46
Mixed	21	6	12	43
Grade				
Moderate to well differentiated	72	24	40	52
Poorly differentiated	125	46	25	134
Stage				
1	31	13	12	12
2	32	16	2	68
3	72	33	35	57
4	62	8	16	49
Adjuvant 5-FU based therapy (in eligible patients)				
Yes	36	28	Not available	19
No	123	31		70
Surgical Margins				
Negative	169	66	Not available	162
Positive	28	4		24

Table 2
Multivariable Cox proportional hazards models

Model (1) incorporates G-INT/G-DIF classes together with Lauren's classes and histological grade which were found to be associated with G-INT/G-DIF subtypes. Patients with mixed histology were excluded from Model (1), Model (2) incorporates all variables found to be prognostic on univariate analysis. Statistically significant results are in bold.

Model (1): Factors interacting with G-INT/G-DIF subtypes			
		Univariate, HR (95% CI), p value	**Multivariable, HR (95% CI), p value**
G-INT/G-DIF	G-INT	1.00	1.00
	G-DIF	**1.95 (1.36–2.78), p<0.001**	**1.92 (1.32–2.78), p<0.001**
Grade	Moderate/Well differentiated	1.00	1.00
	Poor/undifferentiated	1.41 (0.98–2.04), p=0.07	1.40 (0.85–2.31), p=0.19
Lauren's	Intestinal	1.00	1.00
	Diffuse	1.24 (0.87–1.76), p=0.23	0.81 (0.50–1.32), p=0.40

Model (2): Factors affecting survival in univariate analysis			
		Univariate, HR (95% CI), p value	**Multivariable, HR (95% CI), pvalue**
G-INT/G-DIF	G-INT	1.00	1.00
	G-DIF	**HR: 1.79, (1.28–2.51), p=0.001**	**1.63 (1.16–2.29), p=0.005**
Gender	Male	**1.45 (1.01–2.08), p=0.05**	1.00 (0.69–1.47), p=0.98
	Female	1.00	1.00
Margins	Negative	1.00	1.00
	Positive	**1.83 (1.16–2.90), p=0.01**	1.56 (0.98–2.49), p=0.06
Stage	Stage 1	1.00	
	Stage 2	**4.40 (1.49–12.99), p=0.01**	**4.39 (1.48–12.97), p=0.01**
	Stage 3	**11.99 (4.35–33.04), p<0.001**	**12.29 (4.45–33.98), p<0.001**
	Stage 4	**30.13 (10.78–84.22), p<0.001**	**28.56 (10.14–80.43), p<0.001**

Table 3

Interaction between the G-INT and G-DIF subtypes and benefit from 5-FU based adjuvant treatment.Cox proportional hazards regression for survival was used to evaluate interactions between the intrinsic subtypes and 5-FU adjuvant treatment, in patients eligible for adjuvant 5-FU based therapy. Hazard ratios are adjusted for stage.

	G-INT (deaths/N)	G-DIF (deaths/N)	HR (95%CI), p value (G-INT: HR=1.0)	p value for interaction
Adjuvant 5-FU based-treatment	20/45 (44%)	29/38 (76%)	2.71 (1.52–4.85), p=0.001	
Surgery alone	49/136 (36%)	48/86 (56%)	1.37 (0.92 – 2.05), p=0.12	**P=0.002**
HR (95%CI), p value (5-FU based therapy, HR=1)	1.68 (0.98–2.88), p=0.06	0.90 (0.56–1.45), p=0.67		