

Accepted Manuscript

Identification of Molecular Subtypes of Gastric Cancer with Different Responses to PI3-Kinase Inhibitors and 5-Fluorouracil

Zhengdeng Lei, Iain Beehuat Tan, Kakoli Das, Niantao Deng, Hermioni Zouridis, Sharon Pattison, Clarinda Chua, Zhu Feng, Yeoh Khay Guan, Chia Huey Ooi, Tatiana Ivanova, Shenli Zhang, Minghui Lee, Jeanie Wu, Anna Ngo, Sravanthy Manesh, Elisabeth Tan, Bin Tean Teh, Jimmy Bok Yan So, Liang Kee Goh, Alex Boussioutas, Tony Kiat Hon Lim, Horst Flotow, Patrick Tan, Steven G. Rozen

PII: S0016-5085(13)00722-1
DOI: [10.1053/j.gastro.2013.05.010](https://doi.org/10.1053/j.gastro.2013.05.010)
Reference: YGAST 58438

To appear in: *Gastroenterology*
Accepted Date: 9 May 2013

Please cite this article as: Lei Z, Tan IB, Das K, Deng N, Zouridis H, Pattison S, Chua C, Feng Z, Guan YK, Ooi CH, Ivanova T, Zhang S, Lee M, Wu J, Ngo A, Manesh S, Tan E, Teh BT, Yan So JB, Goh LK, Boussioutas A, Hon Lim TK, Flotow H, Tan P, Rozen SG, Identification of Molecular Subtypes of Gastric Cancer with Different Responses to PI3-Kinase Inhibitors and 5-Fluorouracil, *Gastroenterology* (2013), doi: 10.1053/j.gastro.2013.05.010.

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

All studies published in *Gastroenterology* are embargoed until 3PM ET of the day they are published as corrected proofs on-line. Studies cannot be publicized as accepted manuscripts or uncorrected proofs.



Manuscript Number: Gastro 12-01903R1

Identification of Molecular Subtypes of Gastric Cancer with Different Responses to PI3-Kinase Inhibitors and 5-Fluorouracil

Short title: Differing Drug Responses in Gastric Cancer Subtypes

Zhengdeng Lei^{1,2}, Iain Beehuat Tan^{2,3,8}, Kakoli Das², Niantao Deng^{2,3}, Hermioni Zouridis², Sharon Pattison¹², Clarinda Chua⁵, Zhu Feng⁵, Yeoh Khay Guan^{5,6}, Chia Huey Ooi², Tatiana Ivanova⁹, Shenli Zhang², Minghui Lee⁹, Jeanie Wu⁹, Anna Ngo¹⁵, Sravanthy Manesh¹⁵, Elisabeth Tan^{1,2}, Bin Tean Teh^{2,10}, Jimmy Bok Yan So⁷, Liang Kee Goh², Alex Boussioutas^{11,12,13,14}, Tony Kiat Hon Lim¹⁶, Horst Flotow¹⁵, Patrick Tan^{2,4,9,17}, Steven G. Rozen^{1,2}

¹Centre for Computational Biology

²Cancer and Stem Cell Biology Program

Duke-NUS Graduate Medical School, Singapore, 169857, Singapore

³NUS Graduate School for Integrative Sciences and Engineering,

⁴Cancer Science Institute of Singapore, Yong Loo Lin School of Medicine,

⁵Department of Medicine, National University Health System,

⁶Department of Gastroenterology and Hepatology, National University Health System,

⁷Department of Surgery, National University Health System,

National University of Singapore, Singapore, 119074, Singapore

⁸Department of Medical Oncology

⁹Cellular and Molecular Research

¹⁰NCCS-VARI Translational Research Laboratory

National Cancer Centre Singapore, Singapore, 169612, Singapore

¹¹Department of Medicine, Royal Melbourne Hospital, University of Melbourne, Parkville, 3050,

Victoria, Australia

¹²Sir Peter MacCallum Department of Oncology, Peter MacCallum Cancer Centre, University of Melbourne, East Melbourne, 3002, Victoria, Australia

¹³Department of Gastroenterology, Royal Melbourne Hospital, Parkville, 3050, Victoria, Australia

¹⁴Department of Gastroenterology, Western Hospital, Footscray, 3011, Victoria, Australia

¹⁵The Experimental Therapeutics Centre, A*STAR, Singapore, 138669, Singapore

¹⁶Department of Pathology, Singapore General Hospital, Singapore, 169608, Singapore

¹⁷Genome Institute of Singapore, A*STAR, Singapore, 138672, Singapore

ACCEPTED MANUSCRIPT

Short title: Differing Drug Responses in Gastric Cancer Subtypes

Grant Support: This work was supported by the Duke-NUS Signature Research Programs, funded by the Singapore Agency for Science, Technology, and Research and the Singapore Ministry of Health; by the Singapore National Medical Research Council, Translational Clinical Research Grant NMRC/TCR/001/2007; by the Singapore National Research Foundation (NRF) and Ministry of Education (MOE), grant CL2008-07; and by the Singapore Biomedical Research Council, grant 05/1/31/19/423.

Abbreviations: BFRM, Bayesian Factor Regression Modelling; *CD24*, cluster of differentiation 24; *CD44*, cluster of differentiation 44; CHC, consensus hierarchical clustering; CHC_IFS, CHC with iterative feature selection; CI, confidence interval; CNA, copy number alteration; CSC, cancer stem cell; EMT, epithelial-mesenchymal transition; FDR, false discovery rate; 5-FU, 5-fluorouracil; GC-Class, Gastric Cancer Classifier; G-DIF, genomic diffuse; G-INT, genomic intestinal; GO, gene ontology; HR, hazard ratio; IC50, half maximal inhibitory concentration; IFS, iterative feature selection; KEGG, Kyoto Encyclopedia of Genes and Genomes; NTP, nearest template prediction; mTOR, mammalian target of rapamycin; PI3K, phosphatidylinositol 3-kinase; REMARK, REporting recommendations for tumor MARKer prognostic studies; SHH, sonic hedgehog; SPEM, spasmodic-polypeptide-expressing metaplasia.

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

Steven G. Rozen, 8 College Road. Singapore 169857, Singapore. 65 98573213, Fax: 65 62218625, Email: steve.rozen@duke-nus.edu.sg

Disclosures: The authors disclose the following: As mandated by our funders, a patent application covering the gene-expression-based classification of gastric adenocarcinoma has been filed by Exploit Technologies Pte Ltd, a technology transfer arm of the Agency for Science, Technology and Research (A*STAR), Singapore.

Microarray Profilings: All microarray data are available at the NCBI Gene Expression Omnibus repository (<http://www.ncbi.nlm.nih.gov/projects/geo/>):

- (1) Gene expression, Singapore Batch A: GSE15459
- (2) Gene expression, Singapore Batch B: GSE34942
- (3) Gene expression, Australian cohort: GSE35809
- (4) Gene expression, gastric cancer cell lines: GSE22183
- (5) DNA copy number, Singapore cohort: GSE31168
- (6) DNA methylation, Singapore cohort: GSE30601
- (7) Gene expression reported in reference¹: GSE13861

Author Contributions: Z.L. performed most of the data analysis and methodology development and drafted the manuscript and most of the figures and supplementary information. S.G.R. and P.T. designed and supervised the project and helped write and edit the manuscript, figures, and supplementary information. N.D., H.Z., performed additional data analysis and contributed supplementary information. I.B.T. coordinated collection of clinical information on patients and aided with interpretation. C.H.O. provided bioinformatics support. T.K.H.L. coordinated gathering of histopathological information and advised on interpretation of histological sections. K.D., T.I., S.Z., M.L., J.W., A.N., S.M., and E.T. performed laboratory experiments. S.P., C.C., F.Z., K.G.Y., and J.B.Y.S. helped obtain the clinical information. A.B. coordinated expression profiling of Australian tumors and the collection of clinical and histopathological information. B.T.T. and L.K.G. provided scientific advice and coordination. H.F., A.N. and S.M. designed and carried out all aspects of the high throughput drug screening.

ABSTRACT

BACKGROUND & AIMS: Almost all gastric cancers are adenocarcinomas, which have considerable heterogeneity among patients. We sought to identify subtypes of gastric adenocarcinomas with particular biological properties and responses to chemotherapy and targeted agents.

METHODS: We compared gene expression patterns among 248 gastric tumors; using a robust method of unsupervised clustering, consensus hierarchical clustering with iterative feature selection, we identified 3 major subtypes. We developed a classifier for these subtypes and validated it in 70 tumors from a different population. We identified distinct genomic and epigenomic properties of the subtypes. We determined drug sensitivities of the subtypes in primary tumors using clinical survival data, and of cell lines through high-throughput drug screening.

RESULTS: We identified 3 subtypes of gastric adenocarcinoma, called proliferative, metabolic, and mesenchymal. Tumors of the proliferative subtype had high levels of genomic instability, TP53 mutations, and DNA hypomethylation. Cancer cells of the metabolic subtype were more sensitive to 5-fluorouracil than the other subtypes. Furthermore, in 2 independent groups of patients, those with tumors of the metabolic subtype appeared to have greater benefits from 5-fluorouracil treatment. Tumors of the mesenchymal subtype contain cells with features of cancer stem cells, and cell lines of this subtype are particularly sensitive to PI3K–AKT–mTOR inhibitors in vitro.

CONCLUSIONS: Based on gene expression patterns, we classified gastric cancers into 3 subtypes, and validated these in an independent set of tumors. The subgroups have differences in

molecular and genetic features and response to therapy; this information might be used to select specific treatment approaches for patients with gastric cancer.

Keywords: stomach cancer;GC-Class;personalized cancer treatment;cancer classification

ACCEPTED MANUSCRIPT

INTRODUCTION

Like many other kinds of cancer, gastric adenocarcinomas are heterogeneous and arise from a multitude of genetic and epigenetic alterations, and the underlying oncogenic mechanisms are poorly understood.² There are differences between patients in tumor aggressiveness, histopathological features, and responses to therapy.³ However, most patients are still treated with “one-size-fits-all” approaches, and overall five-year survival is only ~27% in the United States.⁴ Consequently, there is an urgent need for a robust classification of gastric cancer that provides insight into oncogenic mechanisms and predicts treatment response.

There have been a few efforts to use microarray-based gene expression profiling to discover new molecular subtypes of gastric cancer by unsupervised hierarchical clustering.^{5, 6} However, these studies relied on relatively small numbers of tumors ($n = 22$ and 47). In addition, we recently carried out unsupervised hierarchical clustering on gastric cancer cell lines (as contrasted with surgically removed tumors) and found two subtypes that were differentiated by the levels of 171 transcripts.⁷ These transcripts then served as the basis for constructing a classifier that we applied to the expression profiles of surgically removed gastric tumors. A strength of this study was that the initial unsupervised hierarchical clustering was based not on primary tumors, which comprise mixtures of malignant and non-malignant cells, but rather on gastric cancer cell lines, with no admixture of non-malignant cells. However, this cell-line based approach may not have captured the full diversity of gastric cancer subtypes. For example, it has proved difficult to derive immortalized cell lines from primary gastric tumors, and almost all gastric cancer cell lines available to date have been derived from metastases or ascites rather than primary sites.⁸

Thus, the present study had two major goals (Figure 1): (1) to develop a robust classification of primary gastric adenocarcinomas based on a relatively large number, 248, of tumors and validate this classification in 70 additional tumors from a separate population, and (2)

to determine the biological, clinical, and potential therapeutic characteristics of the classification's subtypes.

MATERIALS AND METHODS

Overall Design of this Study

Figure 1 presents the overall design of this study.

Supplementary Files

Supplementary files in addition to Supplementary Information are available at <http://dl.dropbox.com/u/62547840/GCSubtyping/index.html>.

Patients and Tumors

Singaporean patients were recruited at the National Cancer Centre and hospitals of the National Healthcare Group. Australian patients were recruited at the Peter MacCallum Cancer Centre in Melbourne. All patients gave written informed consent, and all tissue samples were collected with approvals from respective ethics committees. Tumors were macro dissected. Supplementary Table S1 summarizes the information associated with the patients and tissue samples. This was a retrospective study. We enrolled all consented patients from whom primary gastric tumors were available in the participating center's tissue repositories or pathology archives. Supplementary Materials and Methods provides details, in accordance with REMARK guidelines⁹, on patient treatment, tumor sample handling, and methods for assessing mRNA levels and copy number alterations (CNAs).

Consensus Clustering in Combination with Iterative Feature Selection

We carried out consensus hierarchical clustering (CHC) with iterative feature selection (CHC_IFS). Here we focus on CHC, and describe CHC_IFS in Supplementary Figures S1 and S2.

CHC is a resampling based procedure that repeatedly samples a subset of the tumors and then uses hierarchical clustering to find intrinsic groupings. CHC records the proportion of resamplings in which each pair of tumors was in the same cluster, their “consensus index”.¹⁰ Figure 2A shows the matrix of all consensus indices after CHC_IFS. Each element in this “consensus matrix” is the consensus index for one pair of tumors. In an ideal matrix, all consensus indices would be 1 or 0, indicating that each pair of tumors always or never, respectively, clustered together during the resampling. Figure 2B is close to ideal, as most pairs of tumors have consensus indices near 1 or 0.

In addition, as part of validation of the clustering, we carried out consensus *k*-means clustering using the Bioconductor “ConsensusClusterPlus” package with the clusterAlg parameter set to “km” (Supplementary Table S2 and source code “CHC_IFS.R” in supplementary file SourceCodes.docx).

The GC-Class Classifier for Gastric Cancer Subtypes

We used limma¹¹ to obtain the top differentially expressed probe sets between all three pairs of subtypes, using FDR (false discovery rate) < 0.001 and absolute fold change > 1.5 as thresholds (supplementary file NTP_pairwise_features.xlsx lists these probe sets). These differentially expressed genes and the associated moderated t-statistics generated by limma served as features and weights in the three Nearest Template Predictors¹² (NTPs) that make up a subtype predictor called GC-Class. GC-Class determines the subtype of a sample as follows: If two of the three constituent predictors make the same classification with at least one FDR < 0.05 , then GC-Class uses that classification. Otherwise, if neither FDR is < 0.05 or if all three constituent NTP classifications are different, then GC-Class considers the test sample to be unclassifiable.

RESULTS

Three Subtypes of Gastric Cancer

We assembled a collection of 248 gene-expression profiles by combining our previously reported data¹³ with new data from 56 gastric adenocarcinomas. We dealt with the issue of batch effects using the ComBat algorithm¹⁴ as described in Supplementary Materials and Methods, and we then used CHC_IFS to discover intrinsic subtypes among these tumors (Materials and Methods). This CHC_IFS approach addressed two important issues in unsupervised clustering. One issue is avoidance of the creation of clusters that stem from accidental, non-generalizable characteristics of the specific data set used. CHC avoids this by carrying out multiple rounds of hierarchical clustering on random subsets of the samples.^{10, 15} Thus, CHC can determine if the number of clusters and the assignment of samples to clusters are robust, i.e., stable across multiple re-sampled data sets. The second important issue is the identification of the transcripts that are the most informative with respect to classifying gastric tumors. CHC_IFS addresses this issue by using multiple rounds of CHC, each followed by selection of the transcripts (“features”) showing the largest differences across groups (Supplementary Figure S1). CHC_IFS also significantly improved clustering stability (Supplementary Figure S2).

CHC_IFS provided strong support for the presence of three subtypes (clusters) of tumors (Figures 2A and Supplementary Figure S3A). The subtype boundaries were highly significant (all pairwise adjusted p-values $< 10^{-30}$, Supplementary Materials and Methods) as assessed by the statistical significance of clustering (SigClust) procedure,^{16, 17} which examines whether clusters could be merely artifacts of sampling.

For further studies, we focused on 201 tumors (termed SG201) that were the most representative of the clusters (Figure 2B, Supplementary Figure S3B, and Supplementary Materials and Methods), an approach that has been used previously.¹⁷ CHC_IFS on SG201 yields almost perfectly stable clustering (Figure 2B).

As a further validation of the robustness of the three subtypes, we carried out a second clustering approach, consensus *k*-means clustering with iterative feature selection. This approach generated cluster assignments that were 99.5% identical to the CHC_IFS assignments (Supplementary Table S2) and strong evidence for three clusters (Supplementary Figure S4). Thus, multiple analyses support the existence of three robust subtypes of gastric adenocarcinomas in the Singapore cohort.

A Subtype Classifier and Validation of Subtypes in a Second Group of Tumors

To investigate the reproducibility and utility of the three-subtype classification in an independent group of 70 gastric adenocarcinomas from Australian patients (Supplementary Table S1), we developed a classifier, “GC-Class” (Materials and Methods). After confirming GC-Class’s accuracy in five-fold cross validation in SG201 (accuracy 97%, source code in supplementary file SourceCodes.docx, subsection SG201_5Fold_CV.R), we then proceeded as follows. First, we used GC-Class to assign each of the Australian tumors to one of the three subtypes. Second, we used CHC_IFS to co-cluster SG201 with the Australian tumors (Supplementary Figure S5). The concordance between the GC-class-determined subtype and the clusters by CHC_IFS was 94.3% (Supplementary Table S3), demonstrating the reproducibility and generality of the three-subtype classification and the accuracy of GC-Class.

Characteristic Gene Expression and Functional Annotation of Subtypes

To gain insight into the biological characteristics of each of the subtypes, we first determined which transcripts are significantly up-regulated in one subtype compared to the other two subtypes. For this we used limma,¹¹ using the thresholds of FDR < 0.001 and fold change > 1.5 (gene lists in supplementary file GC_GeneSignatures_and_Annotation.xlsx). We submitted the list of up-regulated genes for each subtype to DAVID^{18, 19} to identify enriched gene sets among the KEGG biological pathways and GO (Gene Ontology) collections (Figure 3A, Supplementary

Materials and Methods, supplementary file GC_GeneSignatures_and_Annotation.xlsx). Genes high in the first subtype were overrepresented in the following gene sets: KEGG focal adhesion (FDR = 2.1×10^{-18}), KEGG extracellular-matrix-receptor interaction (FDR = 3.2×10^{-9}), and GO cell adhesion (FDR = 4.1×10^{-18}). For reasons that we discuss below, we termed this subtype “mesenchymal”. We termed the second subtype “proliferative” because it was characterized by gene sets related to the cell cycle: KEGG cell cycle (FDR = 4.0×10^{-16}), KEGG DNA replication (FDR = 5.8×10^{-7}), and 13 GO gene sets related to cell cycle (FDRs < 10^{-16}). We termed the third subtype “metabolic” because it was characterized by gene sets from several KEGG metabolism pathways (FDRs between 4.6×10^{-6} and 0.023) as well as GO digestion (FDR = 4.21×10^{-15}). Although the metabolic subtype showed expression of genes characteristic of normal gastric mucosa (GO digestion), we show below that this was not because metabolic-subtype tumors had more contaminating normal mucosa, and in fact some tumor cells express these genes.

We also used an additional approach to investigate how the activities of a few specific cancer-related pathways varied across the three subtypes: We used Bayesian Factor Regression Modelling (BFRM, Supplementary Materials and Methods),²⁰ which models observed gene expression levels as consequences of an underlying latent factor that can be viewed as a “pathway activity”. BFRM starts with an initial “seed gene set”—genes that have expression levels (partly) governed by a particular pathway—and then generates a regression model in which these genes’ expression levels are a function of the latent factor (i.e., the pathway activity). BFRM then refines the model by adding further genes that also appear to be associated with the latent factor.

BFRM revealed dramatic and significant differences in pathway activity among the three subtypes (Figure 3B, supplementary file PathwayActivity.xlsx), even though BFRM does not use subtype information in its analysis. We found that the first subtype had high activity of the epithelial-mesenchymal transition (EMT) pathway, which led us to term it “mesenchymal”. Consistent with this, this subtype had high mRNA levels of *CDH2* (N-cadherin) and low levels of *CDH1* (E-cadherin), that are characteristic of mesenchymal cells. This subtype also had high

activities of cancer stem cell (CSC) pathways,^{21, 22} which we discuss in detail below. The mesenchymal subtype is also associated with the p53, TGF- β (transforming growth factor beta), VEGF (vascular endothelial growth factor), NF κ B (nuclear factor kappa-light-chain-enhancer of activated B cells), mTOR (mammalian target of rapamycin), and SHH (sonic hedgehog) pathways. The proliferative subtype is associated with high activities for several oncogenic pathways: E2F, MYC, and RAS. Finally, the metabolic subtype showed high activity for a pathway related to a particular kind of gastric metaplasia termed “SPEM” (spasmolytic-polypeptide-expressing metaplasia), which has been proposed as an intermediate step in the development of gastric adenocarcinoma.^{23, 24} This suggests that progression from SPEM might be a hallmark of metabolic-subtype gastric cancer.

Important Clinical, Genomic, and Biological Differences among the Three Subtypes

To further assess the biological and clinical relevance of the classification scheme, we investigated differences across the three subtypes in multiple characteristics. In terms of basic clinical and pathological characteristics, both the Singapore and Australian cohorts exhibited significant differences with respect to the Lauren histopathological classification²⁵, level of cellular differentiation (“grade”) and the G-INT/G-DIF classification that we developed previously based on cell-line gene expression⁷ (Supplementary Table S4). For example, the mesenchymal subtype is strongly associated with the Lauren diffuse type and the proliferative subtype is strongly associated with the Lauren intestinal type (Supplementary Table S4).

TP53 Mutation and Genomic Instability Are Common in the Proliferative Subtype

There were previously generated data regarding *TP53* mutations,²⁶ copy number alterations (CNAs),²⁷ ENREF_27 and DNA methylation²⁸ ENREF_28 for most of the SG201 tumors. We examined how these characteristics varied across the three subtypes.

Compared to the other two subtypes, the proliferative subtype is characterized by more frequent mutations in the *TP53* mutation hot-spot regions (exons 4 through 9) ($p = 3.16 \times 10^{-3}$, Supplementary Table S5) and by much more extensive copy number amplification (Figure 3C). We also found that gastric adenocarcinomas form two well-differentiated groups with respect to the overall number of cytobands affected by CNA (copy number alteration) and by the level of CNA (Supplementary Figure S6). One group, termed high-CNA, has both more extensive CNA and more extreme levels of CNA than the low-CNA group. The mesenchymal-subtype is significantly enriched for low-CNA tumors, and the proliferative-subtype is significantly enriched for high-CNA tumors (Figure 3D). High-CNA enrichment in the proliferative subtype is primarily due to copy number gains (Supplementary Figure S7). Although loss of *TP53* function is associated with genomic instability, analysis of the relationship between *TP53* mutation status and the extent of CNA suggests that the more extensive CNA in proliferative-subtype tumors is not simply the result of *TP53* mutation, because proliferative subtype tumors have more CNA among both *TP53* wild-type and mutated tumors (Supplementary Figure S8 and Supplementary Table S6).

Consistent with the generally higher levels of CNA in proliferative-subtype tumors, we found that they are enriched for CNAs in six out of 22 previously reported regions of recurrent CNA.²⁷ ENREF 27 Four of these regions involve amplifications of the oncogenes *CCNE1*, *MYC*, *ERBB2* and *KRAS*, and two regions involve deletions of the genes *PDE4D*, *PTPRD* (Supplementary Table S7).

DNA hypomethylation may have a role in promoting chromosomal instability,²⁹⁻³² and in fact, in the proliferative subtype, sites that are hypomethylated (relative to normal mucosa) constitute a significantly higher proportion of aberrantly methylated CpGs than in other two subtypes (Figure 3E, Supplementary Materials and Methods). This may be linked to the high levels of CNA this subtype. We also observed that in the mesenchymal subtype, a significantly higher proportion of aberrantly methylated CpGs are hypermethylated.

The Mesenchymal Subtype Has Cancer-Stem-Cell-Like Properties

Mesenchymal-subtype gastric adenocarcinomas show cancer-stem-cell-like properties³³ in four respects. First, BFRM pathway activity analysis demonstrated that mesenchymal-subtype cancers are significantly associated with the activities of cancer-stem-cell (CSC) pathways. These pathways were derived from (1) transcripts characteristic of prostate CSCs (Supplementary Table S8)²¹ and (2) transcripts characteristic of purified CD44+ breast CSCs (Supplementary Table S9).²² Second, mesenchymal-subtype gastric tumors had high *CD44* and low *CD24* expression compared to the other two subtypes (*CD44*, $p = 1.17 \times 10^{-5}$; *CD24*, $p = 3.39 \times 10^{-9}$; Supplementary Figure S9A). Selection for CD44+/CD24- cells has been used to isolate CSCs in breast cancer,³⁴ pancreatic cancer^{36,37} and gastric cancer.³⁸ High *CD44* and low *CD24* levels are also observed in pancreatic ductal adenocarcinomas of the quasi-mesenchymal subtype (Supplementary Figure S9B),³⁹ which is similar to the mesenchymal subtype of gastric cancer based on highly expressed genes (Supplementary Table S10). Third, the mesenchymal subtype is associated with poorly differentiated gastric cancers (Supplementary Table S4), and maintenance of an undifferentiated state is an essential characteristic of CSCs. Fourth, the set of genes harboring hypermethylated CpG sites in mesenchymal-subtype gastric cancer overlap significantly with the genes down-regulated in hepatocellular carcinoma cells that have hepatic stem cell properties⁴⁰ ($p = 1.10 \times 10^{-5}$, Supplementary Materials and Methods).

No Strong Differences in Survival across the Three Subtypes

We then examined whether there were differences in survival among the three subtypes. We analyzed data for patients from the Singapore and Australian cohorts for whom survival information was available. Kaplan-Meier analysis of cancer-specific survival indicated no significant survival difference among the three subtypes (Singapore, $p = 0.310$; Australia, $p = 0.322$; Supplementary Figure S10). Further analysis with univariate and multivariate Cox

proportional hazards regression likewise detected no significant differences in cancer-specific or disease-free survival among the three subtypes, except that proliferative-subtype patients had worse disease-free survival in multivariate analysis ($p = 0.031$, Supplementary Tables S11, S12, S13, and S14). As expected, higher TNM stages were associated with worse outcomes.

Patients with Metabolic-Subtype Tumors Benefited from 5-Fluorouracil Treatment

Many of the patients in the Singapore and Australian cohorts were treated with 5-fluorouracil (5-FU) in addition to surgery, and we compared the survival of these patients to those treated with surgery alone. The clinical decision to treat with 5-FU was based on solely on clinical consideration of multiple factors, including the patient's general health, disease stage, treatment-related toxicities, and patient preference.

In each of the Singapore and Australian cohorts analyzed separately, metabolic-subtype patients treated with 5-FU fared better than those treated with surgery alone, as follows. In the Singapore cohort, log-rank tests showed better disease-free survival ($p = 0.012$, Supplementary Figure S11C and Supplementary Table S15) and a trend for better cancer-specific survival among those treated with 5-FU ($p = 0.057$, Figure 4C, Table 1). (It was not possible to carry out Cox proportional hazards analysis as there were no deaths among 5-FU treated Singapore patients with metabolic-subtype tumors.) In the Australian cohort, after adjusting for TNM stage, metabolic-subtype patients treated with 5-FU had significantly better cancer-specific and disease-free survival (Table 1 and Supplementary Table S15) than those treated with surgery alone. In addition, Cox proportional hazard analysis of the combined Singapore and Australian patients showed a significant interaction between the metabolic-subtype and 5-FU treatment in cancer-specific and disease-free survival ($p = 8.53 \times 10^{-3}$ and $p = 1.22 \times 10^{-3}$, respectively, Supplementary Tables S16 and S17), providing further evidence that metabolic-subtype patients benefited from 5-FU therapy.

We note that, in patients with mesenchymal and proliferative subtypes, survival was worse among those treated with 5-FU (Figure 4A, B). We believe this was because patients with more advanced TNM stages were more likely to receive 5-FU (Bonferroni-adjusted p-values < 0.02, Fisher's exact test, Supplementary Table S18). However, even after adjusting for TNM stage, among Singapore patients with mesenchymal or proliferative-subtype tumors, we saw no evidence that 5-FU treatment affected cancer-specific or disease-free survival (Table 1 and Supplementary Table S15). Among Australian patients, after adjusting for TNM stage, there is no evidence that 5-FU treatment improved survival of proliferative-subtype patients (Table 1 and Supplementary Table S15). Mesenchymal-subtype patients treated with 5-FU had better disease-free survival (hazard ratio 0.172, $p = 6.66 \times 10^{-3}$, Supplementary Table S15).

We also used cell lines as experimental models to validate the preferential sensitivity of the metabolic subtype to 5-FU: We used GC-Class to predict the subtypes of 28 cell lines for which 5-FU GI50 values (drug concentration at which 50% growth inhibition is achieved) had previously been determined⁷ (Supplementary Table S19). Consistent with our observations of better survival among 5-FU-treated metabolic-subtype patients, metabolic-subtype cells were significantly more sensitive to 5-FU (adjusted $p = 4.41 \times 10^{-3}$ for metabolic versus proliferative, Supplementary Figure S12).

Mesenchymal-Subtype Gastric Cancer Cell Lines Are Sensitive to Compounds that Target the PI3K–AKT–mTOR Pathway

To search more widely for potential drugs effective against the three gastric-cancer subtypes, we carried out a high throughput screen to characterize the sensitivity of 23 gastric cancer cell lines to 158 targeted inhibitors that have been used in clinical trials or that are approved for therapeutic application. We focused on the 29 compounds that had low half maximal inhibitory concentration (IC50) values (< 100 nM) in at least one cell line (Supplementary Table S20). Among these compounds, seven target the PI3K–AKT–mTOR pathway (Table 2), which regulates cellular

metabolism, proliferation, and survival. Mesenchymal-subtype cell lines are significantly more sensitive to six of these seven compounds: the IC50 values for the mesenchymal-subtype cells were significantly lower than for cells in the non-mesenchymal subtypes (FDR < 0.2, Wilcoxon test, Table 2 and Supplementary Figure S13). Abnormal activation of the PI3K–AKT–mTOR pathway is important to the initiation and maintenance of many human tumors.⁴¹ This pathway is also a key regulator of metabolic activities in tumor cells, and it promotes angiogenesis. For example, BEZ235, a dual inhibitor of PI3K and mTOR, has anti-angiogenic properties.⁴² The finding that mesenchymal-subtype gastric cancer cell lines are sensitive to PI3K–AKT–mTOR pathway inhibitors is consistent with the high activation of the mTOR pathway in these tumors (indicated by an arrow in Figure 3B). This finding is also consistent with the CSC-like characteristics of mesenchymal tumors, as CSCs in glioblastoma and prostate cancer are also preferentially sensitive to PI3K–AKT–mTOR inhibitors.^{43, 44}

DISCUSSION

Using CHC_IFS, we found three well-defined subtypes of gastric adenocarcinoma: mesenchymal, proliferative, and metabolic. We developed a predictor, GC-Class, for these subtypes and showed that the three-subtype classification could be applied to an additional set of 70 gastric tumors. Table 3 summarizes the biological and clinical characteristics of each of the three subtypes. Notably, we found that patients with metabolic-subtype tumors benefited preferentially from 5-FU treatment (Figure 4, Supplementary Figure S11, Table 1, and Supplementary Tables 15-17) and that metabolic-subtype cell lines were more sensitive than other cell lines to 5-FU (Supplementary Figure S12). We also found that mesenchymal-subtype cells resemble CSCs, and, consistent with this resemblance, are preferentially sensitive to PI3K–AKT–mTOR inhibitors (Table 2). Thus, the PI3K–AKT–mTOR pathway could be an effective drug target in mesenchymal-subtype tumors.

Beyond the findings reported here, we also found evidence in another report¹ that suggests that the sensitivity of the metabolic subtype to 5-FU is a general phenomenon. Although the number of patients studied was small, reanalysis of the reported data reveals a trend for metabolic-subtype patients to benefit from 5-FU treatment. Among the patients treated with 5-FU (or a derivative, or in combination with other drugs), only one of the nine metabolic-subtype patients died, and none relapsed. In contrast, eight of the 20 mesenchymal-subtype patients died and seven relapsed, while six of the 20 proliferative-subtype patients died and five relapsed (Supplementary Figure S14).

The preferential sensitivity of metabolic-subtype gastric cancers to 5-FU probably stems from the significantly lower expression of both thymidylate synthase (*TS*) and dihydropyrimidine dehydrogenase (*DPD*) in this subtype compared to the other two subtypes (Supplementary Figure S15). In colorectal cancers, low levels of *TS* and *DPD* transcripts are associated with favorable response to 5-FU,⁴⁵ and in gastric cancers, high levels of *TS* transcripts are associated with poor response.⁴⁶ The mechanisms by which low levels of *TS* and *DPD* predispose to 5-FU sensitivity while high levels of either gene predispose to 5-FU resistance appear to be the following. First, 5-FU acts mainly through inhibition of *TS*, so high levels of *TS* presumably lead to resistance to 5-FU.^{47, 48} Second, *DPD* is rate-limiting in 5-FU degradation. Thus, high levels of *DPD* presumably lead to rapid 5-FU degradation and, as a consequence, resistance to 5-FU.

The genes characteristically high in the metabolic subtype include genes that are expressed in normal stomach mucosa (Supplementary Table S21). Could the metabolic subtype comprise tumors with especially high proportions of contaminating normal mucosa? Several lines of evidence suggest instead that the gene signature of the metabolic subtype arises primarily from tumor cells. First, examination of hematoxylin-and-eosin-stained slides from the three subtypes does not suggest a higher proportion of normal mucosa in metabolic-subtype tumors (Supplementary Table S4). Second, metabolic-subtype tumors do not have generally lower tumor content, lower TNM stages, or smaller sizes (Supplementary Table 4). Third, gastric cancer cell

lines express genes characteristic of normal gastric lineages, and the mRNA levels of these genes are consistently higher in metabolic-subtype cell lines (Supplementary Figure S16). This is consistent with the well-established immunohistochemical observation that many gastric cancers express genes that are highly expressed in normal gastric mucosa.⁴⁹ Furthermore, previous reports indicate that tumors expressing such genes have low thymidylate synthase expression and high response to 5-FU,⁵⁰ which are also characteristics of metabolic-subtype cancers. Indeed, we confirmed the presence of mucin 5AC (MUC5AC) protein in tumor cells primarily (though not only) in metabolic subtype tumors (Supplementary Figure S17). MUC5AC is expressed in normal gastric epithelium, some forms of intestinal metaplasia⁵¹, and some gastric adenocarcinomas.⁴⁹ Thus, we propose the hypothesis that metabolic subtype tumors involve not a higher proportion of normal mucosa, but rather tumor cells in which gene expression is generally closer to expression in normal mucosa. This would be consistent with the observation that there are fewer genes that are upregulated in metabolic-subtype tumors than in the other subtypes (Figure 3A, red rectangle on the lower right). This would also be consistent with the theory that gastric tumors arising via SPEM are more likely to express genes also expressed in normal gastric mucosa,^{24, 52} since we observed high SPEM pathway activity in metabolic-subtype tumors (Figure 3B).

In summary, there is substantial evidence that the molecular classification of gastric cancers reported here is reproducible and biologically and therapeutically meaningful. There are numerous molecular differences among the three subtypes. In terms of clinical treatment, there are two promising findings: one is that 5-FU has been particularly effective against metabolic-subtype tumors, and the second is that drugs targeting the PI3K–AKT–mTOR pathway may be particularly effective against mesenchymal-subtype cancers. Thus, if confirmed and extended in future studies, the classification of gastric adenocarcinomas reported here could guide development of therapies tailored to the molecular subtypes.

REFERENCES

1. Cho JY, Lim JY, Cheong JH, et al. Gene expression signature-based prognostic risk score in gastric cancer. *Clin Cancer Res* 2011;17:1850-7.
2. Yokozaki H, Yasui W, Tahara E. Genetic and epigenetic changes in stomach cancer. *Int Rev Cytol* 2001;204:49-95.
3. Dicken BJ, Bigam DL, Cass C, et al. Gastric adenocarcinoma: review and considerations for future directions. *Ann Surg* 2005;241:27-39.
4. Siegel R, Naishadham D, Jemal A. Cancer statistics, 2012. *CA Cancer J Clin* 2012;62:10-29.
5. Tay ST, Leong SH, Yu K, et al. A combined comparative genomic hybridization and expression microarray analysis of gastric cancer reveals novel molecular subtypes. *Cancer Res* 2003;63:3309-16.
6. Kim B, Bang S, Lee S, et al. Expression profiling and subtype-specific expression of stomach cancer. *Cancer Res* 2003;63:8248-55.
7. Tan IB, Ivanova T, Lim KH, et al. Intrinsic subtypes of gastric cancer, based on gene expression pattern, predict survival and respond differently to chemotherapy. *Gastroenterology* 2011;141:476-485 e11.
8. Suzuki T, and Sekiguchi, M. Gastric cancer In Masters J. R. W., & Palsson, B. eds., *Human cell culture*. Dordrecht: Kluwer Academic Publishers, 1999.
9. McShane LM, Altman DG, Sauerbrei W, et al. REporting recommendations for tumour MARKer prognostic studies (REMARK). *Eur J Cancer* 2005;41:1690-6.
10. Monti S, Tamayo P, Mesirov J, et al. Consensus clustering: a resampling-based method for class discovery and visualization of gene expression microarray data. *Machine Learning* 2003;52:91-118.
11. Smyth GK. Linear models and empirical bayes methods for assessing differential expression in microarray experiments. *Stat Appl Genet Mol Biol* 2004;3:Article3.
12. Hoshida Y. Nearest template prediction: a single-sample-based flexible class prediction with confidence assessment. *PLoS One* 2010;5:e15543.
13. Ooi CH, Ivanova T, Wu J, et al. Oncogenic pathway combinations predict clinical prognosis in gastric cancer. *PLoS Genet* 2009;5:e1000676.
14. Johnson WE, Li C, Rabinovic A. Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostatistics* 2007;8:118-127.
15. Wilkerson MD, Hayes DN. ConsensusClusterPlus: a class discovery tool with confidence assessments and item tracking. *Bioinformatics* 2010;26:1572-3.

16. Liu Y, Hayes DN, Nobel A, et al. Statistical significance of clustering for high-dimension, low-sample size data. *Journal of the American Statistical Association* 2008;103:1281-1293.
17. **Verhaak RG, Hoadley KA**, Purdom E, et al. Integrated genomic analysis identifies clinically relevant subtypes of glioblastoma characterized by abnormalities in *PDGFRA*, *IDH1*, *EGFR*, and *NF1*. *Cancer Cell* 2010;17:98-110.
18. Dennis G, Jr., Sherman BT, Hosack DA, et al. DAVID: database for annotation, visualization, and integrated discovery. *Genome Biol* 2003;4:P3.
19. Huang DW, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nature Protocols* 2009;4:44-57.
20. Carvalho CM, Chang J, Lucas JE, et al. High-dimensional sparse factor modeling: applications in gene expression genomics. *Journal of the American Statistical Association* 2008;103:1438-1456.
21. Birnie R, Bryce SD, Roome C, et al. Gene expression profiling of human prostate cancer stem cells reveals a pro-inflammatory phenotype and the importance of extracellular matrix interactions. *Genome Biol* 2008;9:R83.
22. Shipitsin M, Campbell LL, Argani P, et al. Molecular definition of breast tumor heterogeneity. *Cancer Cell* 2007;11:259-73.
23. Lee HJ, Nam KT, Park HS, et al. Gene expression profiling of metaplastic lineages identifies CDH17 as a prognostic marker in early stage gastric cancer. *Gastroenterology* 2010;139:213-25 e3.
24. Weis VG, Goldenring JR. Current understanding of SPEM and its standing in the preneoplastic process. *Gastric Cancer* 2009;12:189-97.
25. 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.
26. Cheng LL, Itahana Y, Lei ZD, et al. *TP53* genomic status regulates sensitivity of gastric cancer cells to the histone methylation inhibitor 3-deazaneplanocin A (DZNep). *Clin Cancer Res* 2012.
27. Deng N, Goh LK, Wang H, et al. A comprehensive survey of genomic alterations in gastric cancer reveals systematic patterns of molecular exclusivity and co-occurrence among distinct therapeutic targets. *Gut* 2012;61:673-84.
28. **Zouridis H, Deng N**, Ivanova T, et al. Methylation subtypes and large-scale epigenetic alterations in gastric cancer. *Sci Transl Med* 2012;4:156.
29. Gaudet F, Hodgson JG, Eden A, et al. Induction of tumors in mice by genomic hypomethylation. *Science* 2003;300:489-92.

30. Eden A, Gaudet F, Waghmare A, et al. Chromosomal instability and tumors promoted by DNA hypomethylation. *Science* 2003;300:455.
31. Suzuki K, Suzuki I, Leodolter A, et al. Global DNA demethylation in gastrointestinal cancer is age dependent and precedes genomic damage. *Cancer Cell* 2006;9:199-207.
32. Wilson AS, Power BE, Molloy PL. DNA hypomethylation and human diseases. *Biochim Biophys Acta* 2007;1775:138-62.
33. Visvader JE, Lindeman GJ. Cancer stem cells in solid tumours: accumulating evidence and unresolved questions. *Nat Rev Cancer* 2008;8:755-68.
34. Al-Hajj M, Wicha MS, Benito-Hernandez A, et al. Prospective identification of tumorigenic breast cancer cells. *Proc Natl Acad Sci U S A* 2003;100:3983-8.
35. Yu F, Yao H, Zhu P, et al. let-7 regulates self renewal and tumorigenicity of breast cancer cells. *Cell* 2007;131:1109-23.
36. Li C, Heidt DG, Dalerba P, et al. Identification of pancreatic cancer stem cells. *Cancer Res* 2007;67:1030-7.
37. Li C, Lee CJ, Simeone DM. Identification of human pancreatic cancer stem cells. *Methods Mol Biol* 2009;568:161-73.
38. Takaishi S, Okumura T, Tu S, et al. Identification of gastric cancer stem cells using the cell surface marker CD44. *Stem Cells* 2009;27:1006-20.
39. **Collisson EA, Sadanandam A**, Olson P, et al. Subtypes of pancreatic ductal adenocarcinoma and their differing responses to therapy. *Nat Med* 2011;17:500-3.
40. Yamashita T, Ji J, Budhu A, et al. EpCAM-positive hepatocellular carcinoma cells are tumor-initiating cells with stem/progenitor cell features. *Gastroenterology* 2009;136:1012-24.
41. Wymann MP, Zvebil M, Laffargue M. Phosphoinositide 3-kinase signalling--which way to target? *Trends Pharmacol Sci* 2003;24:366-76.
42. Schnell CR, Stauffer F, Allegrini PR, et al. Effects of the dual phosphatidylinositol 3-kinase/mammalian target of rapamycin inhibitor NVP-BEZ235 on the tumor vasculature: implications for clinical imaging. *Cancer Res* 2008;68:6598-607.
43. Dubrovska A, Elliott J, Salamone RJ, et al. Combination therapy targeting both tumor-initiating and differentiated cell populations in prostate carcinoma. *Clinical Cancer Research* 2010;16:5692-5702.
44. **Eyler CE, Foo WC**, LaFiura KM, et al. Brain cancer stem cells display preferential sensitivity to Akt inhibition. *Stem Cells* 2008;26:3027-36.
45. Danenberg K, Salonga D, Park JM, et al. Dihydropyrimidine dehydrogenase and thymidylate synthase gene expressions identify a high percentage of colorectal tumors responding to 5-fluorouracil. *Proc Am Soc Clin Oncol* 1998;17:258a.

46. Lenz HJ, Leichman CG, Danenberg KD, et al. Thymidylate synthase mRNA level in adenocarcinoma of the stomach: a predictor for primary tumor response and overall survival. *J Clin Oncol* 1996;14:176-82.
47. Longley DB, Harkin DP, Johnston PG. 5-fluorouracil: mechanisms of action and clinical strategies. *Nat Rev Cancer* 2003;3:330-8.
48. Schilsky RL. Biochemical and clinical pharmacology of 5-fluorouracil. *Oncology (Williston Park)* 1998;12:13-8.
49. Kushima R, Vieth M, Borchard F, et al. Gastric-type well-differentiated adenocarcinoma and pyloric gland adenoma of the stomach. *Gastric Cancer* 2006;9:177-84.
50. Tajima Y, Shimoda T, Nakanishi Y, et al. Association of gastric and intestinal phenotypic marker expression of gastric carcinomas with tumor thymidylate synthase expression and response to postoperative chemotherapy with 5-fluorouracil. *J Cancer Res Clin Oncol* 2003;129:683-90.
51. Reis CA, David L, Correa P, et al. Intestinal metaplasia of human stomach displays distinct patterns of mucin (MUC1, MUC2, MUC5AC, and MUC6) expression. *Cancer Res.* 1999;59:1003-1007.
52. Goldenring JR, Nomura S. Differentiation of the gastric mucosa III. Animal models of oxyntic atrophy and metaplasia. *Am J Physiol Gastrointest Liver Physiol* 2006;291:G999-1004.

FIGURE LEGENDS**Figure 1. Study design****Figure 2. Three subtypes of gastric cancer**

CHC_IFS strongly supported three subtypes of gastric adenocarcinoma as indicated by the consensus matrices for all Singapore tumors (A), and for a selected set of highly representative tumors, SG201(Supplementary Materials and Methods) (B). Supplementary Figure S3 provides evidence that there are three subtypes.

Figure 3. Systematic differences among the three subtypes

(A) Gene signatures and KEGG and GO annotations for the three subtypes.

(B) Pathway activity. Each column represents one tumor sample. Each row represents pathway activities computed by BFRM from the pathway seed gene set indicated at the right. Arrow indicates row for an mTOR pathway. Supplementary file PathwayActivity.xlsx provides the numerical pathway activities and the sources of the seed gene sets.

(C) The number of cytobands with CNA by subtype (p-values by Kruskal-Wallis test).

(D) The numbers of tumors in each subtype falling in the low or high CNA group (Bonferroni-adjusted p-values by hypergeometric tests).

(E) The numbers of aberrantly methylated CpGs that are hyper- and hypomethylated for each subtype (Bonferroni-adjusted p-values by hypergeometric tests).

Figure 4. Cancer specific survival of patients with and without 5-FU treatment

Kaplan-Meier plots for patients treated with surgery alone versus surgery plus 5-FU for each of the three subtypes. “K-M log-rank p” refers to the p-value of the log-rank test. “Cox p” refers to

the p-value from Cox proportional hazards models that include TNM stage as a covariate (Table 1). Only cancer-related deaths are treated as events.

ACCEPTED MANUSCRIPT

Table 1. Cox proportional hazards models of cancer-specific survival as a function of 5-FU treatment and TNM stage for each of the three subtypes in each population

Model	Covariate	Value	Hazard Ratio (95% CI)	p-value
Singapore, Mesenchymal	5-FU	N	1	
		Y	1.062 (0.469, 2.407)	0.885
	TNM stage	I	1	
		II	4.799 (0.545, 42.294)	0.158
		III	17.460 (2.149, 141.874)	7.46x10 ⁻³
IV		36.870 (3.915, 347.224)	1.62x10 ⁻³	
Singapore, Proliferative	5-FU	N	1	
		Y	2.105 (0.879, 5.043)	0.095
	TNM stage	I	1	
		II	0.851 (0.053, 13.664)	0.909
		III	8.028 (1.032, 62.431)	0.047
IV		26.062 (2.984, 227.603)	3.19x10 ⁻³	
Singapore, Metabolic	(As there were no deaths among 5-FU treated patients, Cox regression is not available)			
Australia, Mesenchymal	5-FU	N	1	
		Y	0.242 (0.056, 1.045)	0.057
	TNM stage	I/II	1	
III/IV		16.381 (2.968, 90.403)	1.33x10 ⁻³	
Australia, Proliferative	5-FU	N	1	
		Y	9.979 (1.947, 51.160)	5.80x10 ⁻³
	TNM stage	I/II	1	
III/IV		0.176 (0.037, 0.840)	0.029	
Australia, Metabolic	5-FU	N	1	
		Y	0.090 (0.010, 0.813)	0.032
	TNM stage	I/II	1	
III/IV		136.325 (6.766, 2,746.938)	1.34x10 ⁻³	

Note: N = No, Y = Yes. Non-cancer deaths were treated as censored. Hazard ratios were computed by the coxph function in the R package “survival”. In addition to subtype, we included TNM stage, age, grade, and tumor site in initial models, because these variables were significant in univariate analysis in either the Singapore or Australian cohort (Supplementary Table S11). Variables age, grade and tumor site were subsequently omitted because they were not significant in the initial models.

Table 2. Chemosensitivity differences among the three subtypes of gastric cancer cell lines.

Compound, Batch	p-value Mes vs. Non-mes	FDR Mes vs. Non-mes	Targeted pathway	Sensitive in subtype
BEZ235, Batch A	3.46E-04	0.011	PI3K–mTOR	Mes
BEZ235, Batch B	0.011	0.068	PI3K–mTOR	Mes
PI-103 hydrochloride	6.61E-03	0.068	PI3K	Mes
PIK-75	0.010	0.068	PI3K	Mes
PI-103	0.020	0.107	PI3K	Mes
ZSTK474	0.046	0.184	PI3K	Mes
GSK690693	0.038	0.176	Akt	Mes
MK2206	0.103	0.328	Akt	Mes

Note: p-values by Wilcoxon tests for possibly tied observations. FDRs obtained by R function p.adjust.

See Supplementary Figure S14 and Supplementary Table 20 for additional details.

Mes = Mesenchymal subtype.

BEZ235 was tested in two batches.

Table 3. Characteristics of the three subtypes of gastric adenocarcinoma. (Please refer to Supplementary Table S4 for details.)

	Mesenchymal	Proliferative	Metabolic
5-FU effect on patient survival	No effect in Singapore cohort; beneficial in Australian cohort	No effect	Beneficial
Chemosensitivity in cell lines	PI3K–AKT–mTOR inhibitors	-	5-FU
KEGG pathways associated with up-regulated genes	Focal adhesion, ECM-receptor interaction	Cell cycle, DNA replication	Metabolic processes
Gene Ontology biological processes associated with up-regulated genes	Cell adhesion, vasculature development, cell motility, angiogenesis	M phase, mitotic cell cycle	Digestion, secretion
Pathway activation determined by BFRM	EMT, TGF- β , VEGF, NF κ B, mTOR, SHH, and CSC	E2F, MYC, and RAS	SPEM (spasmolytic polypeptide-(TFF2)-expressing-metaplasia)
Grade	High	Low	-
TNM stage	No significant difference among the three subtypes		
Tumor size	No significant differences among the three subtypes		
Tumor content	Tendency for mesenchymal < metabolic < proliferative (details in Supplementary Table S2)		
Age \pm SD	62.51 \pm 12.33	66.42 \pm 11.78	63.28 \pm 15.71
Lauren classification (%)			
Diffuse	58.2	17.3	40.6
Intestinal	29.9	73.6	53.6
Mixed	11.9	9.1	5.8
Classification in reference ⁷ (%)			
G-INT	7.5	71.2	84.3
G-DIF	92.5	28.8	15.7
Characteristic copy number alteration	Low CNA	High CNA	-
Amplified Genes	-	<i>CCNE1, MYC, ERBB2, KRAS</i>	-
Aberrantly methylated CpGs (%)			
Hypermethylated	84.6	57.8	76.1
Hypomethylated	15.4	42.2	23.9
Characteristic aberrant methylation	Hypermethylation	Hypo-methylation	-
Frequency of <i>TP53</i> mutation	Low	High	Low

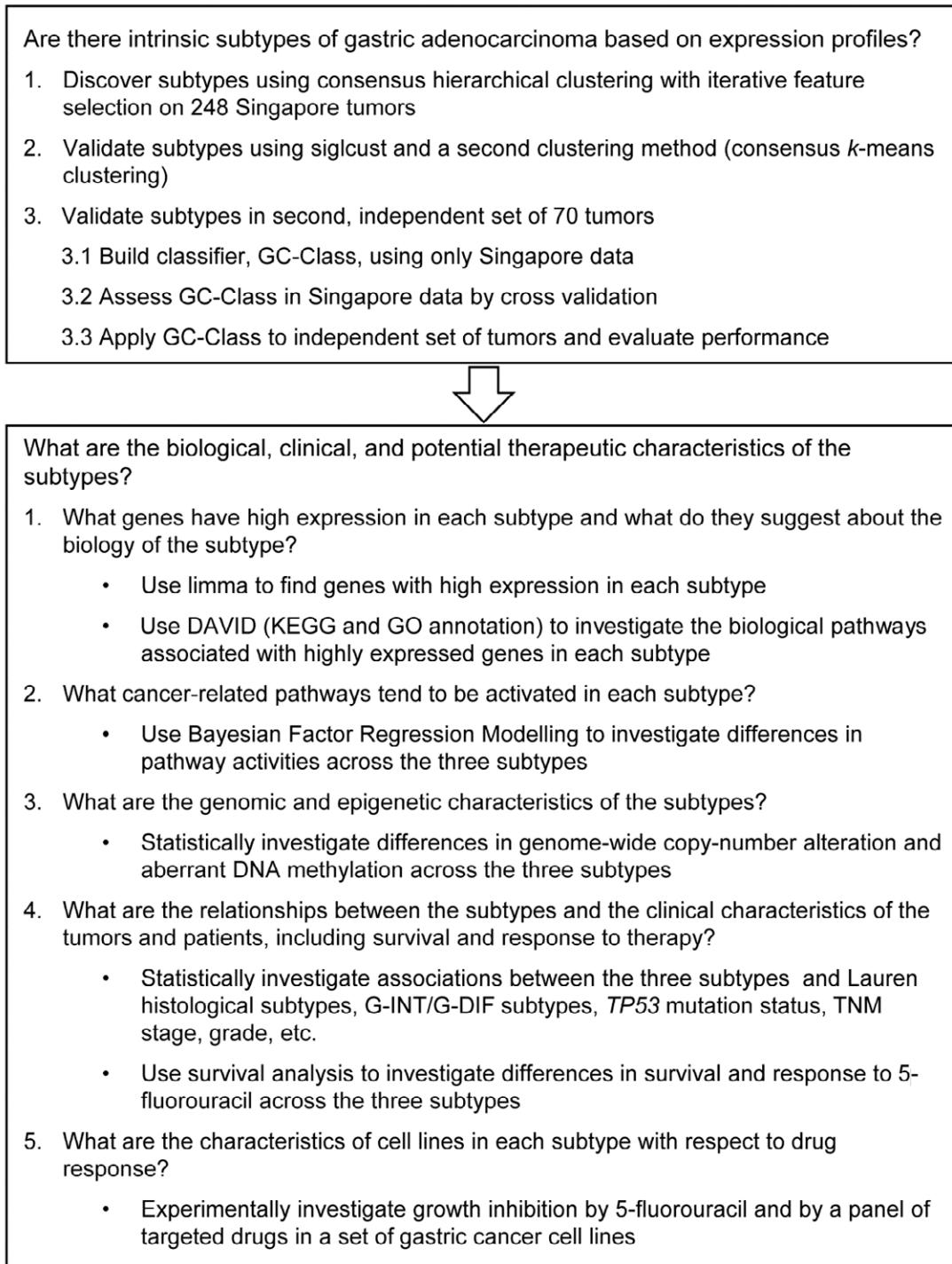
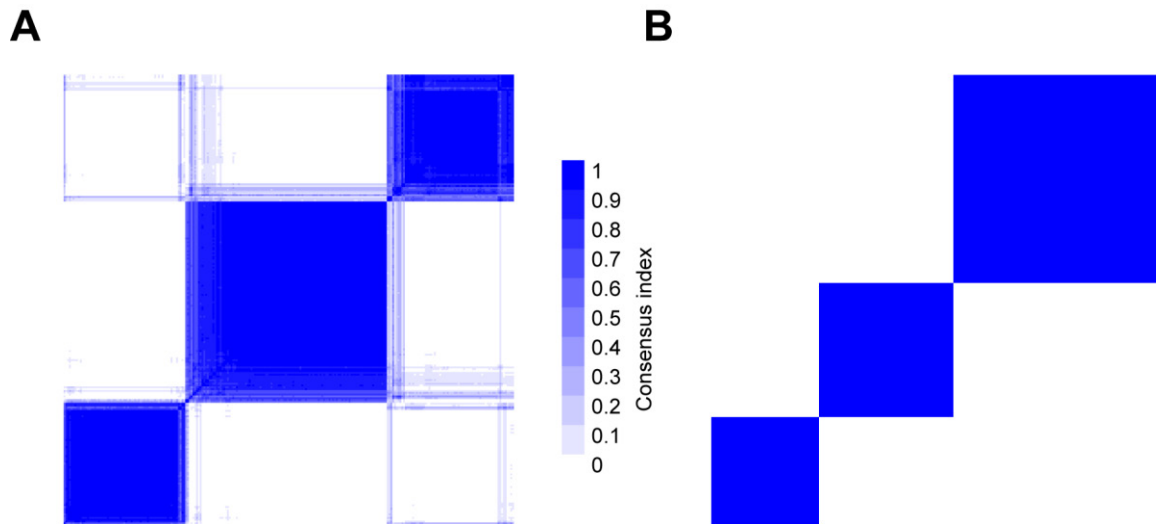
Figure 1. Study design

Figure 2. Three subtypes of gastric cancer

ACCEPTED MANUSCRIPT

Figure 3. Systematic differences among the three subtypes

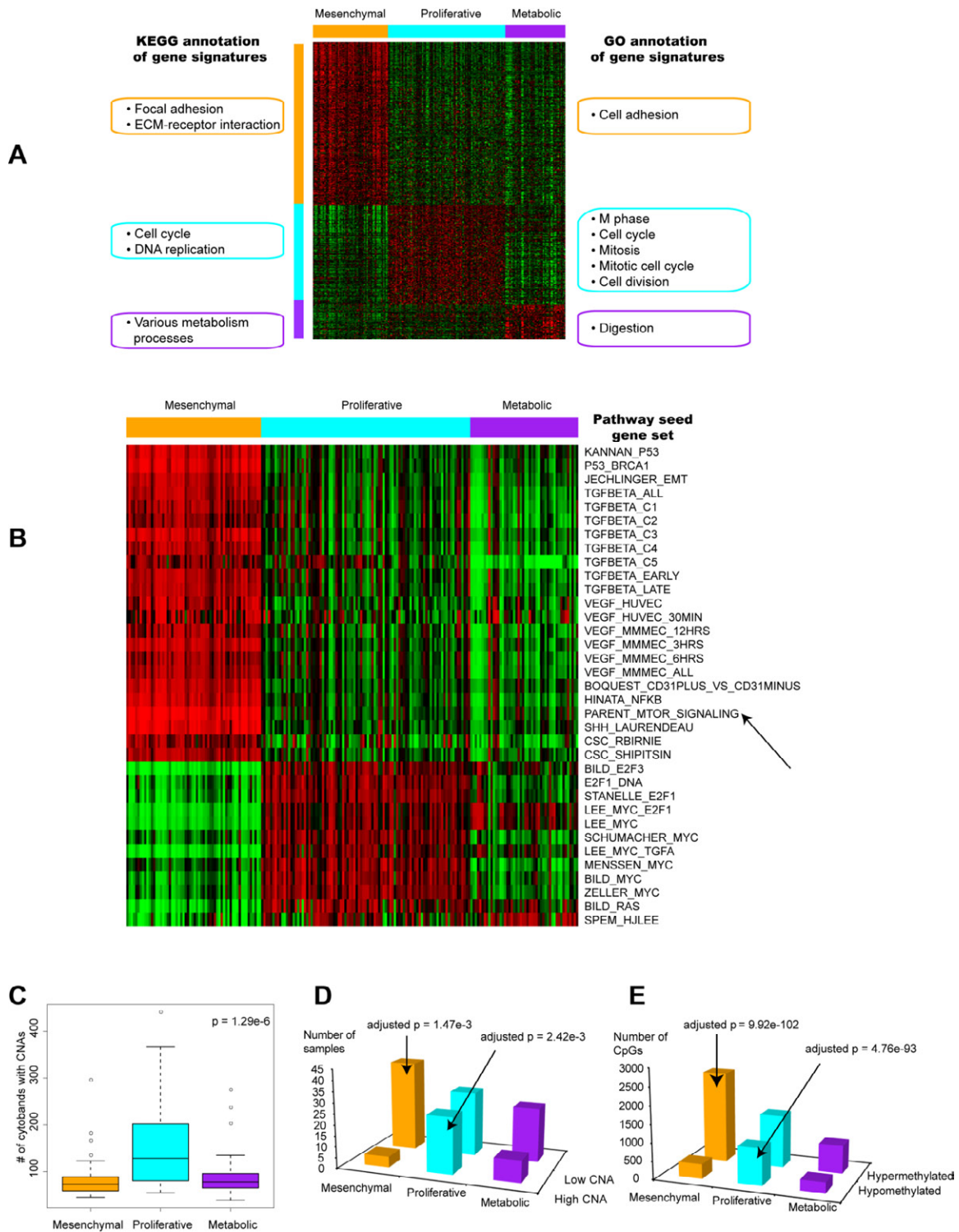


Figure 4. Cancer-specific survival of patients with and without 5-FU treatment in the Singapore and Australian cohorts

