ORIGINAL ARTICLE

A comprehensive survey of genomic alterations in gastric cancer reveals systematic patterns of molecular exclusivity and co-occurrence among distinct therapeutic targets

Niantao Deng,[1,2] Liang Kee Goh,[1,3,4] Hannah Wang,[1] Kakoli Das,[1] Jiong Tao,[1,5] Iain Beehuat Tan,[1,2,4] Shenli Zhang,[1] Minghui Lee,[6] Jeanie Wu,[6] Kiat Hon Lim,[7] Zhengdeng Lei,[8] Glenn Goh,[1] Qing-Yan Lim,[9] Angie Lay-Keng Tan,[1] Dianne Yu Sin Poh,[1] Sudep Riahi,[10] Sandra Bell,[10] Michael M Shi,[11] Ronald Linnartz,[11] Feng Zhu,[12] Khay Guan Yeoh,[12] Han Chong Toh,[4] Wei Peng Yong,[13] Hyun Cheol Cheong,[14] Sun Young Rha,[14] Alex Boussioutas,[15] Heike Grabsch,[16] Steve Rozen,[8] Patrick Tan[1,6,17,18]

See Commentary, p 638

► Additional materials are published online only. To view these files please visit the journal online (http://gut.bmj.com/content/61/5.toc).

For numbered affiliations see end of article.

Correspondence to
Dr Patrick Tan, Cancer and Stem Cell Biology Program, Duke-NUS Graduate Medical School, 8 College Road, Singapore 169857, Singapore; gmstanp@duke-nus.edu.sg

Accession number in NCBI GEO database: GSE31168.

Revised 9 January 2012
Accepted 10 January 2012
Published Online First
7 February 2012

ABSTRACT

Objective Gastric cancer is a major gastrointestinal malignancy for which targeted therapies are emerging as treatment options. This study sought to identify the most prevalent molecular targets in gastric cancer and to elucidate systematic patterns of exclusivity and co-occurrence among these targets, through comprehensive genomic analysis of a large panel of gastric cancers.

Design Using high-resolution single nucleotide polymorphism arrays, copy number alterations were profiled in a panel of 233 gastric cancers (193 primary tumours, 40 cell lines) and 98 primary matched gastric non-malignant samples. For selected alterations, their impact on gene expression and clinical outcome were evaluated.

Results 22 recurrent focal alterations (13 amplifications and nine deletions) were identified. These included both known targets (*FGFR2*, *ERBB2*) and also novel genes in gastric cancer (*KLF5*, *GATA6*). Receptor tyrosine kinase (RTK)/RAS alterations were found to be frequent in gastric cancer. This study also demonstrates, for the first time, that these alterations occur in a mutually exclusive fashion, with *KRAS* gene amplifications highlighting a clinically relevant but previously underappreciated gastric cancer subgroup. *FGFR2*-amplified gastric cancers were also shown to be sensitive to dovitinib, an orally bioavailable FGFR/VEGFR targeting agent, potentially representing a subtype-specific therapy for *FGFR2*-amplified gastric cancers.

Conclusion The study demonstrates the existence of five distinct gastric cancer patient subgroups, defined by the signature genomic alterations *FGFR2* (9% of tumours), *KRAS* (9%), *EGFR* (8%), *ERBB2* (7%) and *MET* (4%). Collectively, these subgroups suggest that at least 37% of gastric cancer patients may be potentially treatable by RTK/RAS directed therapies.

Significance of this study

What is already known about this subject?

- Gastric cancer patients with *ERBB2*-amplified tumours can clinically benefit from *ERBB2*-targeted therapies. Similar to *ERBB2*, several other molecularly targeted therapies are currently being evaluated in gastric cancer.
- Little is known regarding which molecular targets are concurrently expressed in the same gastric tumours, or independently in different tumours.
- Unlike other cancer types, activating mutations in *KRAS* are also rarely observed in gastric cancer.

What are the new findings?

- This study identified 22 recurrent genomic alterations in gastric cancer, comprising both known gastric cancer targets (*FGFR2*, *ERBB2*) and genes not previously reported to be amplified in gastric cancer (*KLF5*, *GATA6*).
- Genes related to RTK/RAS signalling, in particular *FGFR2*, *KRAS*, *ERBB2*, *EGFR* and *MET* are frequently amplified in gastric cancer in a mutually exclusive manner.
- *FGFR2*-amplified gastric cancers exhibited sensitivity to dovitinib, an orally bioavailable targeted therapy.
- KRAS amplifications, frequently observed in gastric cancer, are significantly associated with adverse prognosis.

Gastric adenocarcinoma, or gastric cancer is a leading cause of global cancer mortality with an overall 5-year survival rate of approximately 20%.[1 2] Particularly prevalent in many Asian countries,[3] most gastric cancer patients present at advanced disease stages and are treated by palliative chemotherapy, with median survival times of 11–12 months.[4] In addition to standard cytotoxic regimens, targeted therapies, which are small molecules or antibodies designed to disrupt the activity of

Significance of this study

How might it impact on clinical practice in the foreseeable future?

- Dovitinib may represent a subtype-specific therapy for *FGFR2*-amplified gastric cancers.
- KRAS genomic amplification status should be assessed in clinical trials involving therapies targeting upstream RTK.
- Genomic amplifications in RTK/RAS components define five distinct gastric cancer molecular subgroups, to which differing therapies can be allocated. In total, 37% of the gastric cancer population may be treatable by RTK/RAS targeting agents.

specific oncogenic signalling pathways, have recently emerged as a promising therapeutic strategy. In the recent ToGA trial,[4] trastuzumab, an anti-*HER2/ERBB2* targeting antibody, improved the overall survival of patients with HER2-positive tumours when combined with chemotherapy. However, because only 7–17% of gastric cancer patients are *HER2* positive (either gene amplification or overexpression) and thus suitable candidates for anti-*HER2* therapy,[5–7] further research is warranted to increase the population of gastric cancer patients for which targeted treatments are clinical options.

Reflecting this urgency, several other targeted therapies are currently undergoing preclinical and clinical testing in gastric cancer, directed against diverse oncogenic proteins including signalling receptors, histone deacetylases and cellular proteins.[8–10] However, because most of these targeted therapies were originally designed against proteins expressed or discovered in other cancers (eg, trastuzumab for breast cancer), in many cases surprisingly little is actually known either regarding the true prevalence of their oncogenic targets in primary gastric cancers, or if expression of these oncogenic targets is correlated with key clinico-pathological parameters such as patient outcome. As one example, the *FGFR2* receptor tyrosine kinase (RTK) has previously been proposed as a potential therapeutic target in gastric cancer.[11] However, most *FGFR2*-related studies in gastric cancer have been primarily restricted to in-vitro cultured cell lines,[12 13] and little data is available regarding the true prevalence of *FGFR2* gene amplification in primary gastric cancers particularly at the high-resolution genomic level. As such, a comprehensive and unbiased survey to identify the most prevalent molecular targets in gastric cancer could facilitate many aspects of gastric cancer translational research, for example, in focusing clinical trials efforts on those therapies that might benefit the greatest numbers of gastric cancer patients.

Besides identifying the most prevalent targets, recent findings have also highlighted the importance of determining if certain combinations of targets are expressed either independently from one another (ie, mutual exclusivity) or co-occurring in the same tumour. Knowledge of such 'inter-target relationships' (ITR) can shed critical insights into the signalling networks of a cancer cell, case examples being the mutual exclusivity of *KRAS* and *BRAF* activating mutations in colorectal cancer, and the exclusivity of *EGFR* and *KRAS* mutations in lung cancer.[14 15] Identifying ITR may also highlight promising drug combinations for combination therapy, and suggest rational molecular criteria for patient inclusion and exclusion in clinical trials. Recent studies exemplifying both the basic and clinical importance of ITR include *ERBB2* and *PIK3CA*, in which co-occurring *PIK3CA* mutations in *ERBB2*-positive breast cancers can modulate clinical responses to trastuzumab,[16] and *EGFR* and *MET*, in which clinical resistance to gefitinib in *EGFR*-mutated lung cancers can be caused by co-existing *MET* gene amplifications.[17]

In this study, we sought to identify the most prevalent molecular targets in gastric cancer and to elucidate their ITR. To achieve this aim, we performed, to our knowledge, the largest and most comprehensive survey of genomic copy number alterations in gastric cancer to date, profiling more than 230 gastric cancers (>190 primary tumours and 40 cell lines) on high resolution single nucleotide polymorphism (SNP) arrays containing over 1 million array probes.

MATERIALS AND METHODS

Patient samples were obtained from institutional tissue repositories of the participating centres. Primary gastric tumours were collected with approvals from the respective institutional research ethics review committees and with signed patient informed consent. 'Normal' (ie, non-malignant) samples used in this study refer to samples harvested from the stomach, from sites distant from the tumour and exhibiting no visible evidence of tumour or intestinal metaplasia/dysplasia upon surgical assessment. Clinicopathological information of these patients including age, disease stage, histological subtype, treatment and anatomical location, are included in supplementary table S1 (available online only). Only three patients received neo-adjuvant or preoperative chemotherapy before surgery. Gastric cancer cell lines were obtained from commercial sources (American Type Culture Collection, Japan Health Science Research Resource Bank) or from collaborators (Yonsei Cancer Centre, South Korea). Genomic DNA were extracted from flash-frozen tissues or cell pellets using a Qiagen genomic DNA extraction kit (Qiagen, Hilden, Germany), and profiled on Affymetrix SNP 6.0 arrays (Affymetrix, Santa Clara, California, USA) according to the manufacturer's specifications. The array data have been deposited into the National Centre for Biotechnology Information's Gene Expression Omnibus under accession number GSE31168. Tumour-specific genomic alterations were identified by normalising the primary gastric cancer profiles against the primary matched gastric normal samples. Analyses were performed using the genomic identification of significant targets in cancer (GISTIC) algorithm[18] using false discovery rate q-value thresholds of less than 0.25 for broad regions and less than 0.001 for focal regions, similar to those used in previous reports.[19–21]

Additional details, including methods associated with dimension reduction permutation (DRP), fluorescence in-situ hybridisation (FISH) assays, and functional assays, are presented in the supplementary materials (available online only).

RESULTS

Genomic landscape of Copy Number Alteration (CNA) in gastric cancer

We profiled genomic DNA samples from 193 primary gastric cancers, 98 primary matched gastric normal samples and 40 gastric cancer cell lines on Affymetrix SNP6 microarrays containing approximately 1.8 million probes with a median interprobe spacing of 680 bp. To identify tumour-specific genomic alterations and exclude regions of potential germ line copy number variation, we normalised the gastric cancer profiles against the matched gastric normal samples (see the Methods section and supplementary figure S1 (available online only) for representative profiles). On average, we observed approximately 150 genomic aberrations per gastric cancer, comprising a mixture of broad and focally altered regions. Frequently amplified broad chromosomal regions included 1q, 3q, 5p, 6p, 7pq, 8q, 12pq, 13q,

18pq, 19p, 20pq and 21p (frequencies 9.8–33.7%), and frequently deleted chromosomal regions included 3p, 4pq, 5q, 6q, 8p, 9p, 9q, 11q, 12p, 14q, 16q, 17p, 18p, 18q, 19p 21q and 22q (frequencies 7.8–13.0%) (figure 1A). These results are highly concordant with previous comparative genomic hybridisation (CGH/aCGH) studies of gastric cancer.[22–27]

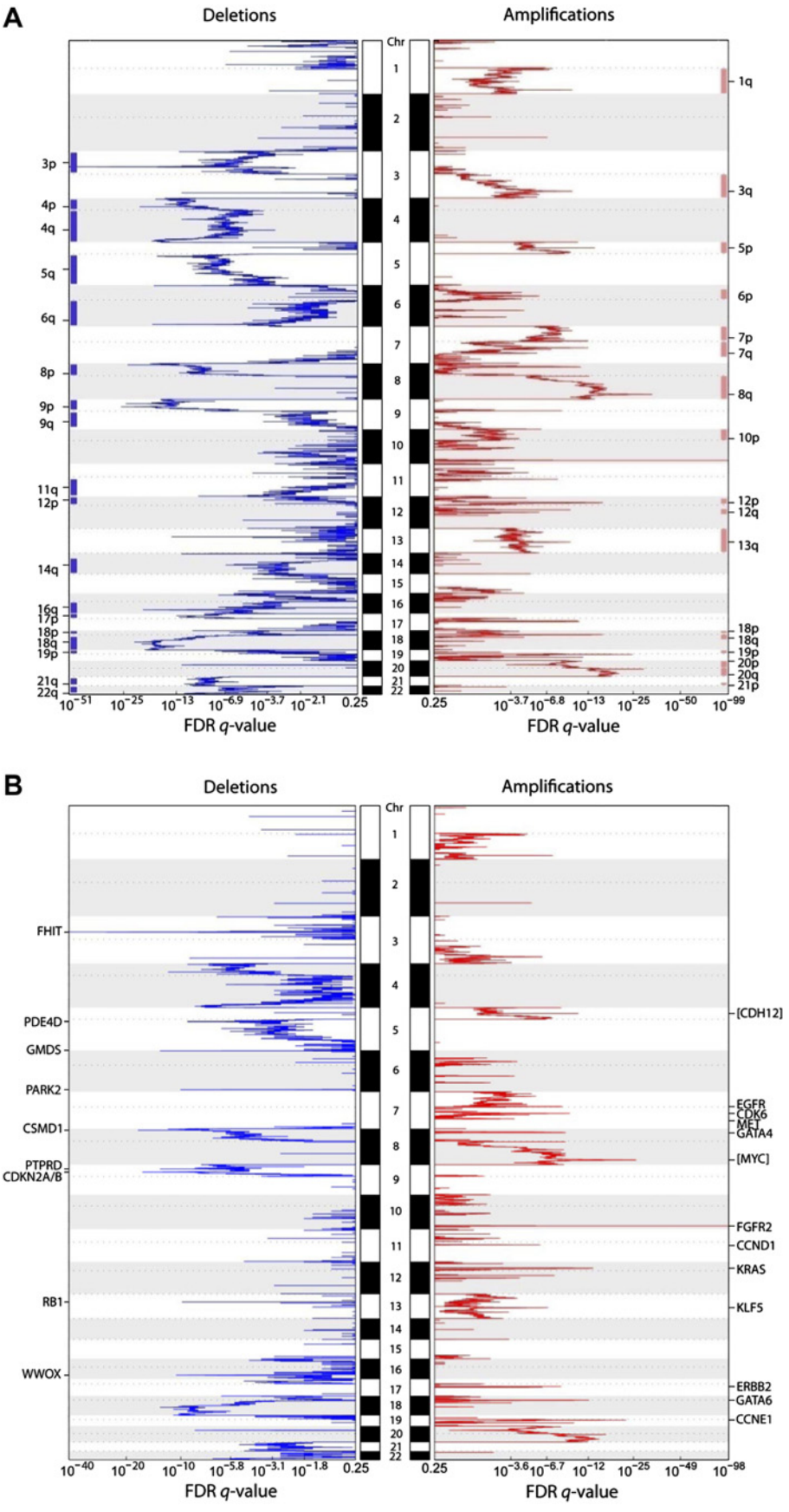


Figure 1 Broad and focal genomic alterations in gastric cancer. (A) Large-scale copy number alterations. The diagram shows a CNA plot where chromosomal regions of the 22 autosomes are represented on the y-axis, and genomic identification of significant targets in cancer (GISTIC) computed false discovery rate (FDR) q-values are on the x-axis. Chromosomal deletions are on the left (blue) and amplifications are on the right (red). Significantly altered regions of broad CNA are highlighted at the sides, as blue and red bars (GISTIC q value <0.25). (B) Focal alterations. Genes localised within the peaks of the focally altered regions are specified. Genes in square brackets are genes that lie immediately adjacent to the alteration peak (eg, *MYC*). Significantly altered focal events (GISTIC q-value <0.001) are highlighted at the sides and summarised in table 1.

Focal genomic alterations highlight 22 potential targets in gastric cancer

We identified 22 focal genomic alterations, defined as narrow regions (typically <100 kb) exhibiting high levels of copy number gain or loss (table 1). Among the amplified genes were several oncogenes previously known to be amplified in gastric cancer, including *EGFR*, *ERBB2/HER2* and *CCND1* (figure 1B).[6 28 29] Among the focally deleted genes in gastric cancer, we re-identified *FHIT*, *RB1*, *CDKN2A/B*, and *WWOX*, also previously known to be deleted in gastric cancer.[30–34] The re-discovery of these classic oncogenes and tumour suppressor genes supports the accuracy of the SNP6 array data. To validate the array data further, we performed *ERBB2* immunohistochemistry on 146 of the 193 cases (see supplementary figure S2, available online only), and confirmed a significant association between *ERBB2* copy number gain and *ERBB2* protein expression ($p<0.01$, Fisher's exact test, supplementary table S2, available online only).

Besides known genes, the analysis also revealed novel genes not previously reported in gastric cancer. These included genomic amplification of the transcription factors *GATA6* and *KLF5*, and somatic deletions in *PARK2*, *PDE4D*, *CSMD1* and *GMDS*. Recent data suggest that GATA factors in particular may play an oncogenic role in certain gastrointestinal cancers, for example, *GATA6* has been shown to be amplified in pancreatic cancer.[35] *PARK2* and *PDE4D* deletions have also recently been observed in glioblastoma and lung adenocarcinomas.[19 20] Using immunohistochemistry, we confirmed that one of these novel deleted genes, *CSMD1*, was downregulated or absent in approximately 40% of primary gastric cancers at the protein level, but was highly expressed in normal gastric epithelium (n=42; supplementary figure S3, available online only).

A network of non-random ITR define relationships between gastric cancer targets

A major goal of our study was to identify non-coincidental ITR between the 22 gastric cancer targets in a systematic, unbiased and statistically rigorous manner. We developed a statistical method called DRP for this purpose. Briefly, DRP identifies non-random ITR between targets by comparing the numbers of tumour samples exhibiting a particular ITR (associations between distinct alterations) against a null distribution of background ITR generated through random permutation. The supplementary information (available online only) provides a detailed description of the DRP method. Compared with other methods such as hierarchical clustering and correlation tests, DRP provides additional sensitivity in identifying ITR, without requiring a priori knowledge of specific gene functions (see supplementary figure S4, available online only).

We uncovered several significant ITR associated with the 22 gastric cancer targets. These target pairs were either amplified in a mutually exclusive manner in different tumours, or co-amplified in the same tumour (figure 2 and supplementary table S3, available online only). Functionally, the gastric cancer ITR tended to involve two specific target classes—genes related to RTK/RAS signalling, including *KRAS*, *FGFR2*, *ERBB2*, *EGFR* and *MET*, and genes related to transcription factor biology (*MYC*, *GATA4*, *GATA6* and *KLF5*). For example, tumours exhibiting *KRAS* amplifications were largely distinct from tumours exhibiting *ERBB2* or *FGFR2* amplification (p=0.02 and p=0.005 for

Table 1 Focal regions of CNA regions in gastric cancer

CNA	Chr	Start	End	Length (kb)	Cytoband	Q value	Genes in peak
Amplification							
1	10	123 336 181	123 337 713	1.5	10q26.13	3.9561E-99	***FGFR2***
2	8	128 628 340	128 670 251	41.9	8q24.21	7.984E-27	***[MYC]***
3	19	34 982 652	35 002 397	19.7	19q12	3.1439E-23	***CCNE1***
4	12	25 213 920	25 336 398	122.5	12p12.1	1.5713E-14	***KRAS***, *CASC1, LYRM5*
5	18	17 947 474	18 040 783	93.3	18q11.2	1.0616E-13	*GATA6*
6	5	21 377 838	21 406 308	28.5	5p14.3	9.501E-12	***[CDH12]***
7	7	91 921 079	92 111 471	190.4	7q21.2	2.0612E-10	***CDK6***, *PEX1, GATAD1, DKFZP564O0523, FAM133B*
8	8	11 346 688	11 659 701	313.0	8p23.1	9.0544E-10	*BLK, GATA4, C8orf13*
9	7	55 237 447	55 373 693	136.2	7p11.2	2.4109E-09	***EGFR***
10	17	35 102 118	35 136 335	34.2	17q12	3.8268E-09	***ERBB2***
11	13	72 528 937	72 770 614	241.7	13q22.1	1.4729E-07	*KLF5*
12	11	69 161 019	69 306 967	145.9	11q13.2	9.1737E-07	***CCND1***, *FGF4, FGF19, ORAOV1*
13	7	115 987 034	116 178 774	191.7	7q31.2	0.00012527	*CAV1*, ***MET***
Deletion							
1	3	60 447 451	60 472 964	25.5	3p14.2	3.4002E-41	***FHIT***
2	8	4 182 635	4 182 916	0.3	8p23.2	1.0797E-18	*CSMD1*
3	9	21 953 419	21 995 192	41.8	9p21.3	1.0299E-17	***CDKN2A, CDKN2B***
4	6	2 019 538	2 068 880	49.3	6p25.3	1.7756E-14	*GMDS*
5	16	77 269 209	77 293 232	24.0	16q23.1	5.4871E-12	*WWOX*
6	6	162 551 244	162 610 874	59.6	6q26	2.1056E-11	*PARK2*
7	13	47 806 677	47 809 375	2.7	13q14.2	3.3682E-11	***RB1***
8	5	58 436 441	58 569 237	132.8	5q11.2	1.6661E-10	*PDE4D*
9	9	9 524 063	9 675 303	151.2	9p23	1.2287E-09	*PTPRD*

Focal recurrent CNA (amplifications and deletions) identified by genomic identification of significant targets in cancer (GISTIC). Genes previously reported as oncogenes or tumour suppressor genes are highlighted in bold. Start and end indicates the boundary of the region identified. Length indicates size of each region identified. Q value represents the significance of the recurrent CNA region across all the gastric tumours. Genes in peak, genes covered by the corresponding region, a square bracket indicates that the gene lies immediately adjacent to the peak.

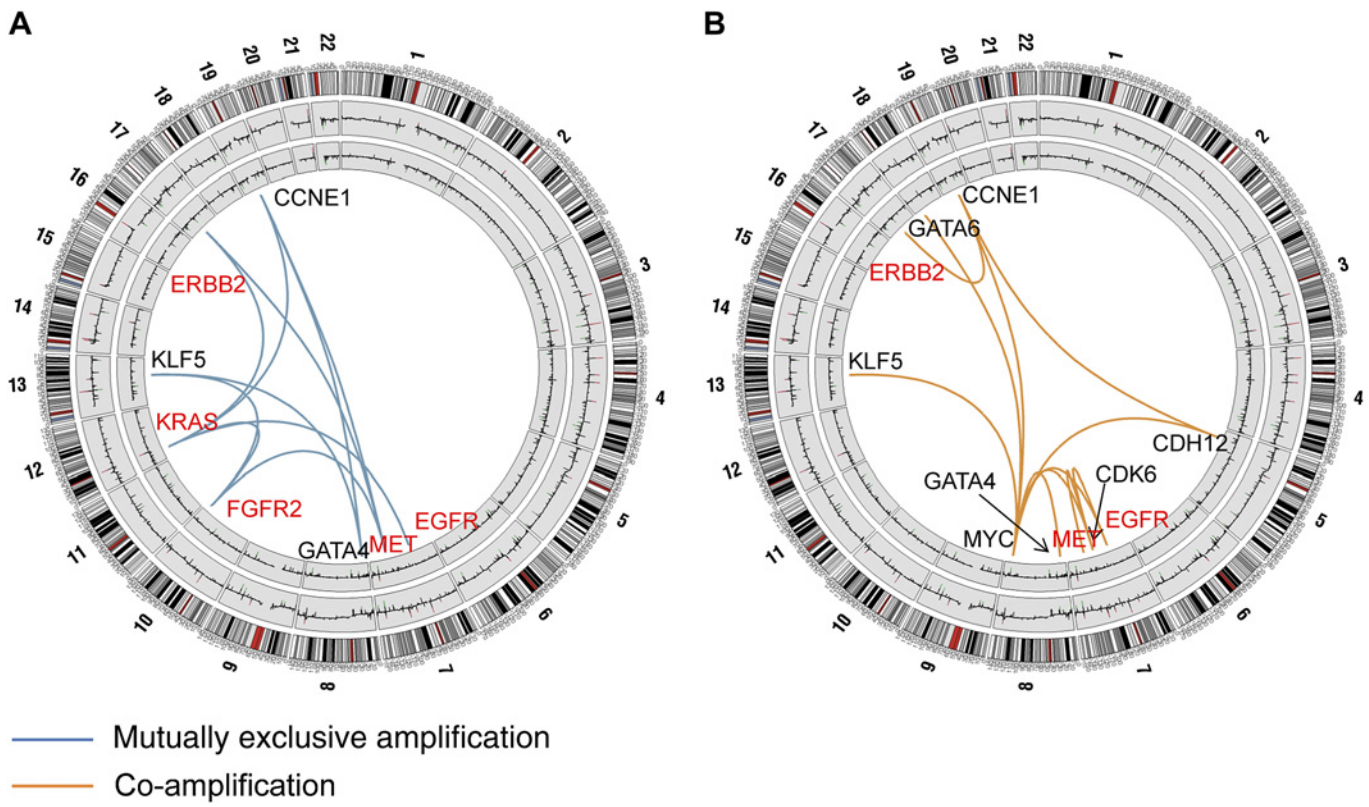


Figure 2 Mutually exclusive and co-amplified genomic alterations. (A) Focal regions exhibiting mutually exclusive patterns of genome amplification. Chromosomal diagrams were created using Circos software.[36] Circular tracks from outside to in: genomic positions by chromosomes (black lines are cytobands, red lines are centromeres); summarised CNA values in gastric tumours, summarised CNA values in normal gastric samples. Blue lines indicate pairs of focal regions (genes) exhibiting significant patterns of mutually exclusive genomic amplification identified by dimension reduction permutation (DRP) analysis ($p<0.05$; *EGFR/KRAS*, $p=0.05$). Genes involved in receptor tyrosine kinase (RTK)/RAS signalling are highlighted in red. (B) Focal regions exhibiting patterns of genomic co-amplification. Orange lines indicate pairs of focal regions (genes) exhibiting significant patterns of genomic co-amplification identified by DRP analysis ($p<0.05$). Genes involved in RTK/RAS signalling are highlighted in red. Supplementary table S3 (available online only) provides a complete list of significant mutually exclusive and co-alteration relationships for amplifications and deletions.

KRAS/ERBB2 and *KRAS/FGFR2*, respectively), while tumours exhibiting *MET* amplifications were distinct from tumours with *FGFR2* amplifications ($p=0.03$; figure 2A and supplementary table S3, available online only). Likewise, *GATA4*, *GATA6* and *KLF5* were significantly co-amplified with *MYC* (*KLF5*: $p=0.0005$; *GATA4*: $p=0.008$; *GATA6*: $p=0.01$), while *KLF5* and *GATA4* amplifications were mutually exclusive to one another ($p=0.01$).

Other notable ITR included a significant co-amplification interaction between *EGFR* and *MYC* ($p=0.002$) and between *ERBB2* and *CCNE1* ($p=0.05$) (figure 2B), a co-amplification pattern recently linked to trastuzumab resistance in breast cancer.[37] Taken collectively, these results support the existence of a complex functional network of ITR in gastric cancer. They provide evidence that instead of each target behaving independently from one another, the presence of one target in a gastric cancer is likely to exert a profound influence on the repertoire of other targets expressed in that same tumour.

Genomic alterations in RTK signaling genes—frequent, mutually exclusive and associated with patient survival in gastric cancer

Motivated by the clinical success of trastuzumab and the availability of other RTK-targeting drugs in the gastric cancer translational pipeline,[38] we decided to characterise the RTK genomic alterations and their impacts on patient outcome. A heat-map representation of the SNP array data confirmed that the four amplified RTK (*FGFR2*, *ERBB2*, *EGFR* and *MET*) were mutually exclusive to one another (figure 3A). In addition, *KRAS* genomic amplifications were also mutually exclusive to the other RTK (figure 3A), suggesting these five components may activate the same downstream pathway in gastric cancer (see supplementary figure S5, available online only). The *KRAS* amplifications are examined in more detail in the next section.

Taken collectively, RTK/RAS genomic amplifications occurred in approximately 37% of the entire gastric cancer cohort (figure 3B). The most frequently amplified RTK/RAS component was *FGFR2* (9.3%), followed by *KRAS* (8.8%), *EGFR* (7.7%) and *ERBB2* (7.2%). Of 72 tumours exhibiting amplification in at least one RTK/RAS component, 73.6% (53/72) exhibited amplification of only one component, and 26.4% (19/72) tumours exhibited high level amplification of one component with low level amplification of another. Only two tumours exhibited high level amplification of two RTK/RAS components (black arrows in figure 3A). Taken collectively, these results suggest that 37% of the gastric cancer population is thus potentially targetable by a RTK/RAS-directed therapy.

To assess the prognostic impact of RTK amplifications in gastric cancer, we performed a survival analysis comparing the clinical outcome of patients bearing tumours with RTK amplifications compared with patients with tumours lacking RTK amplification. In a univariate analysis, patients with RTK amplified tumours (*FGFR2*, *ERBB2*, *EGFR*, *MET*) experienced poor survival outcome compared with patients with RTK amplification-negative cancers ($p=0.01$, HR 1.636, 95% CI 1.101

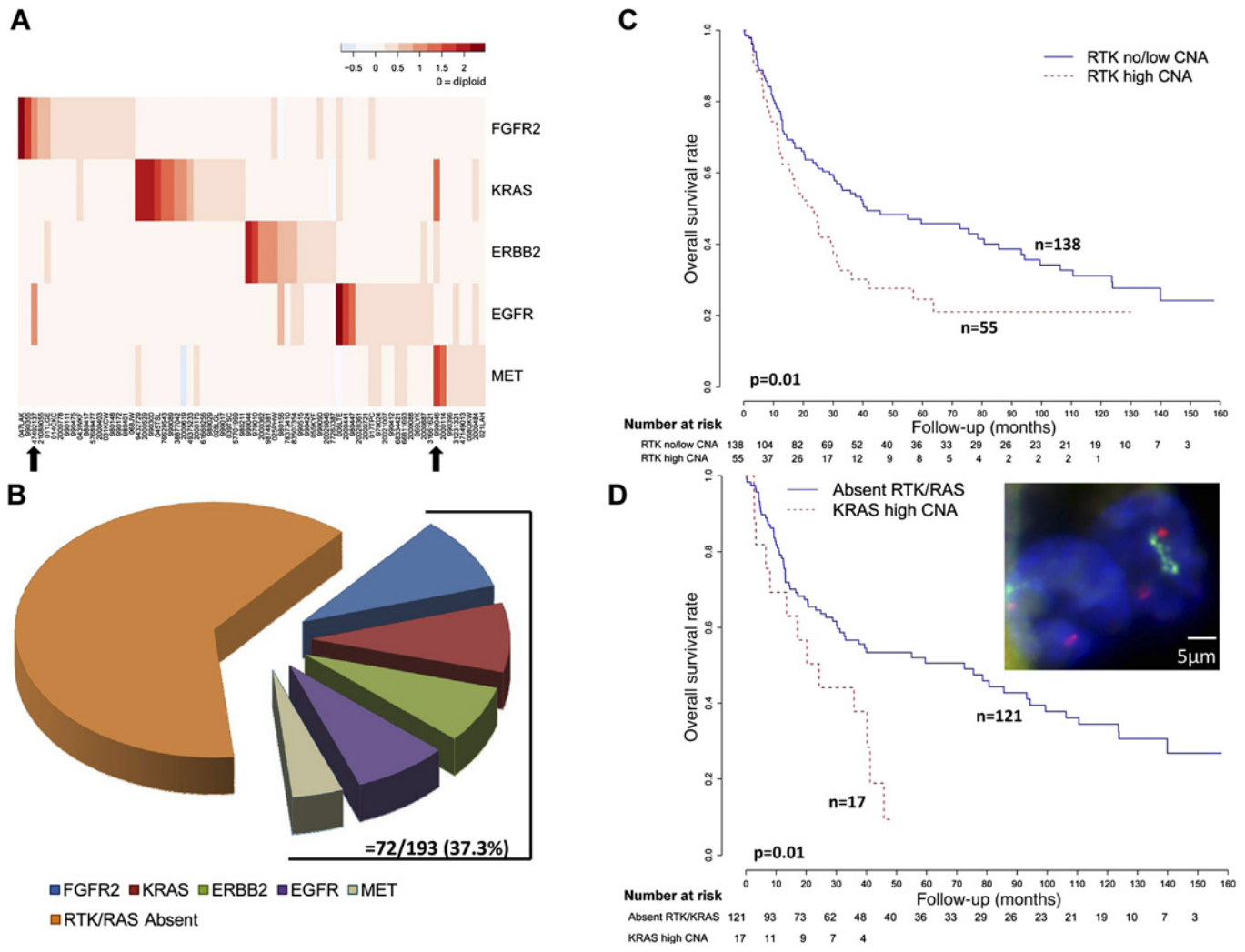


Figure 3 Genomic alterations of receptor tyrosine kinase (RTK)/RAS signalling components in gastric cancer. (A) Mutually exclusive amplification patterns of RTK/RAS signalling components. In the heat-map, each row represents a different RTK/RAS signalling component. Each column represents an individual tumour exhibiting RTK/RAS amplification (72 tumours). The red colour gradient (top right) highlights the degree of copy number amplification. Black arrows highlight two tumours exhibiting high level amplifications in two RTK/RAS components. (B) Overall frequency of RTK/RAS genomic alterations in gastric cancer. The pie chart displays the different gastric cancer subgroups exhibiting RTK/RAS amplification. Gastric cancers exhibiting at least one RTK/RAS amplification event comprise a collective 37% of the gastric cancer cohort analysed. (C) Kaplan–Meier survival analysis comparing outcomes of patients with tumours exhibiting RTK amplification (either *FGFR2*, *ERBB2*, *EGFR*, or *MET*) amplification to patients with tumours lacking RTK amplification. Patients with tumours exhibiting focal *KRAS* amplifications were included in analysis, and fall into the RTK low/no CNA group. Overall survival was used as the outcome metric. (D) Kaplan–Meier survival analysis comparing outcomes of patients with tumours exhibiting *KRAS* amplification (15 patients) to patients with non-RTK/*KRAS*-amplified tumours. Overall survival was used as the outcome metric. The inset photo displays a patient tumour (ID 49375233) with *KRAS* amplification confirmed by fluorescence in-situ hybridisation (FISH) analysis (blue, DAPI nuclear stain; green, KRAS FISH probe; red, centromere 12 probe).

to 2.432; figure 3C). Moreover, in multivariate Cox regression models including RTK amplification status, stage, grade and treatment status (surgery alone or 5-FU adjuvant chemo-radiation), RTK amplification status was shown to be an independent prognosis predictor (p=0.01, HR 1.966, 95% CI 1.180 to 3.279; see supplementary table S4a, available online only). The adverse prognosis of RTK-amplified gastric cancers was also largely independent of chromosomal instability (p=0.07), indicating that it is not a mere consequence of increased aneuploidy (see supplementary table S4a, available online only).[39]

To evaluate individual RTK, we performed a follow-up univariate Cox model analysis considering the four different amplified RTK (*FGFR2*, *ERBB2*, *EGFR* and *MET*) as independent factors. Patients with *ERBB2*-amplified tumours and *MET*-amplified tumours were found to exhibit the worst prognosis (*ERBB2*: p=0.0006, HR 2.824, 95% CI 1.558 to 5.119; *MET*: p=0.002, HR 2.744, 95% CI 1.190 to 6.327; see supplementary table S4b, available online only). The adverse prognostic impact of *ERBB2* amplification was also observed in a multivariate Cox model with adjustment for tumour stage and grade (see supplementary table S4c, available online only).[6 7] Therefore, among the four different RTK, *ERBB2* amplifications appear to exert the strongest prognostic impact in gastric cancer.

KRAS-genomic amplifications highlight a previously underappreciated gastric cancer subgroup

KRAS amplifications were frequently observed in our series, occurring in 9% of patients. This finding is of interest, because canonical activating mutations in *KRAS* at codons 12 and 13 are strikingly infrequent in gastric cancer, unlike other

gastrointestinal cancers (eg, colorectal and pancreatic cancer).[40 41] Confirming these earlier studies,[41] the *KRAS* mutation rate in our own series was extremely low—among 139 gastric cancers genotyped for *KRAS* codon 12 and 13 mutations, only one tumour exhibited a *KRAS* mutation (G13D in 069LYK). We thus hypothesised that *KRAS* genome amplification, rather than mutation, may represent a predominant mechanism for *KRAS* activation in gastric cancer.

To obtain additional evidence that *KRAS* genomic amplifications represent a distinct gastric cancer molecular subgroup, we performed a Kaplan–Meier survival analysis comparing outcomes of patients with *KRAS*-amplified samples versus patients with tumours lacking RTK or *KRAS* amplification. Patients with *KRAS*-amplified tumours exhibited significantly poorer prognosis (p=0.01, HR 2.158, 95% CI 1.172 to 3.971; figure 3D). Supporting the robustness of this survival association, similarly significant associations were observed when patients with *KRAS*-amplified tumours were compared against patients lacking *KRAS* amplification but irrespective of RTK amplification, or when the copy number threshold defining *KRAS* amplification was relaxed (p=0.06, HR 1.744, 95% CI 0.973 to 3.127; p=0.01, HR 1.665, 95% CI 1.114 to 2.488; see supplementary figure S6, available online only).

To benchmark the prognostic effect of *KRAS* amplification against other RTK, we applied a univariate Cox regression model consisting of all five genes. Similar to *ERBB2* and *MET* amplifications, gastric cancer patients with *KRAS* amplifications also exhibited significantly worse prognosis compared with patients with tumours lacking either RTK or *KRAS* amplifications (p=0.02, HR 2.116, 95% CI 1.155 to 6.940; see supplementary table S5a, available online only); however, this association may be related to tumour stage (p=0.2, HR 1.455, 95% CI 0.790 to 2.682; see supplementary table S5b, available online only).

Finally, to provide functional evidence that *KRAS* genomic amplification represents an important 'driver' event in *KRAS*-amplified gastric cancers, we performed genetic knockdown experiments. Small interfering RNA-mediated knockdown of *KRAS* in *KRAS* amplified and *KRAS*-mutated gastric cancer cell lines caused significant reductions in proliferation but not in *KRAS* wild-type lines, supporting an earlier report[41] (see supplementary figure S7, available online only). These results suggest that *KRAS* amplification in gastric cancer probably defines a specific subgroup of poor prognosis patients for which *KRAS* signalling in tumours is critical.

FGFR2 amplifications in gastric cancer: relationships to gene expression, clinical outcome and drug sensitivity

FGFR2 was being amplified in 9–10% of gastric cancers in our series (table 1). Consistent with *FGFR2* being the main driver of amplification in this locus, intersection of the amplification regions across 20 *FGFR2*-amplified tumours confirmed that *FGFR2* was the sole gene in this region exhibiting common copy number gain (figure 4A). Validating the SNP data, a quantitative PCR analysis using primers directed towards *FGFR2* confirmed that samples with high *FGFR2* qPCR values were associated with *FGFR2* amplification. (p=0.0006, Fisher's test; see supplementary figure S8, available online only). FISH analysis using BAC probes targeting *FGFR2* also confirmed *FGFR2* gene amplification in patient tumours and cell lines, relative to a centromere 10 probe (figure 4B).

FGFR2 has previously been proposed as a potential therapeutic target in gastric cancer,[38] but little is known regarding the impact of *FGFR2* amplification on gene expression and other clinicopathological parameters. To investigate relationships between *FGFR2* gene amplification and *FGFR2* gene expression, we analysed gene expression profile data for 156 of the 193 gastric cancers analysed by SNP arrays in this study, which we have described in an earlier report.[42] *FGFR2*-amplified gastric cancers indeed exhibited significantly increased *FGFR2* gene expression levels (figure 4C and supplementary figure S9, available online only), when compared against a reference set of 100 normal gastric samples, or non-*FGFR2*-amplified tumours (Kruskal–Wallis test p=6.7e-9, Wilcoxon test p=1.7e-7 (vs normal) and p=1.9e-5 (vs non-*FGFR2*-amplified gastric cancers). In comparison, *ATE1* and *BRWD2*, two genes located adjacent to *FGFR2* exhibited less significant levels of copy number/gene expression correlation (p=0.004–0.3, relative to normals; supplementary figure S10, available online only), further supporting *FGFR2* as the major driver gene in this region.

Examining clinicopathological variables, *FGFR2*-amplified gastric cancers did not exhibit any significant associations with histology (Lauren's p=0.8, grade p=0.8 or tumour stage p=0.9) or patient survival (p=0.8, see supplementary table S4b, available online only). However, in an expanded gene expression dataset of 398 gastric tumours derived from four distinct cohorts of which the previous 156 gastric cancers form a subset (see supplementary information and supplementary table S6, available online only), high *FGFR2* expression (compared with normals, supplementary figure S11, available online only) was associated with poor survival outcome in a univariate analysis (p=0.01, HR 1.492, 95% CI 1.094 to 2.035; figure 4D). In a multivariate Cox regression model, samples with *FGFR2* high expression tended to exhibit borderline significance after adjusting for stage and grade (p=0.08, HR 1.321, 95% CI 0.966 to 1.807; see supplementary table S7, available online only). This result suggests that *FGFR2* overexpression in gastric cancer may be of prognostic relevance.

Dovitinib (TKI258) is an investigational multitargeting oral tyrosine kinase inhibitor with potent inhibitory activity against bFGF receptors 1, 2, 3, VEGF receptors 1, 2, 3, PDGFR and *c-KIT*.[43 44] In preclinical models, dovitinib has exhibited antitumour activity in *FGFR1*-amplified breast cancer,[45] and in several phase I clinical trials has shown good therapeutic profiles in human patients.[46 47] To test the potential efficacy of dovitinib in *FGFR2*-amplified gastric cancer, we treated *FGFR2*-amplified and non-amplified gastric cancer lines (figure 5A) with increasing dosages of dovitinib, to determine the GI50 concentration (the drug concentration required to cause 50% growth inhibition). We observed potent growth inhibitory activity of dovitinib specifically in *FGFR2*-amplified gastric cancer cell lines with GI50 dosages in the submicromolar range (KATO-III 0.12 μM; SNU-16 0.17 μM, figure 5B). Decreased phosphorylation of *FGFR2*, ERK and AKT was also observed after 1 h of dovitinib treatment (figure 5C). Besides inhibiting cell proliferation, dovitinib treatment also induced a significant decrease in soft-agar colony formation in *FGFR2*-amplified lines (KATO III p=0.002; SNU16 p=0.05; figure 5D and supplementary figure S12, available online only). In a cell death assay, dovitinib treatment induced apoptosis, measured by caspase 3/7 activation, in SNU-16 cells after 24 h of treatment, but not in KATO III cells (figure 5E). These results suggest that dovitinib treatment can inhibit several pro-oncogenic traits in *FGFR2*-amplified lines, but additional factors may be required for *FGFR2*-amplified cells to undergo apoptosis upon dovitinib treatment.

To evaluate the efficacy of dovitinib in an in-vivo model, we performed drug treatment experiments using an *FGFR2*-amplified primary human gastric cancer xenograft model, comparing

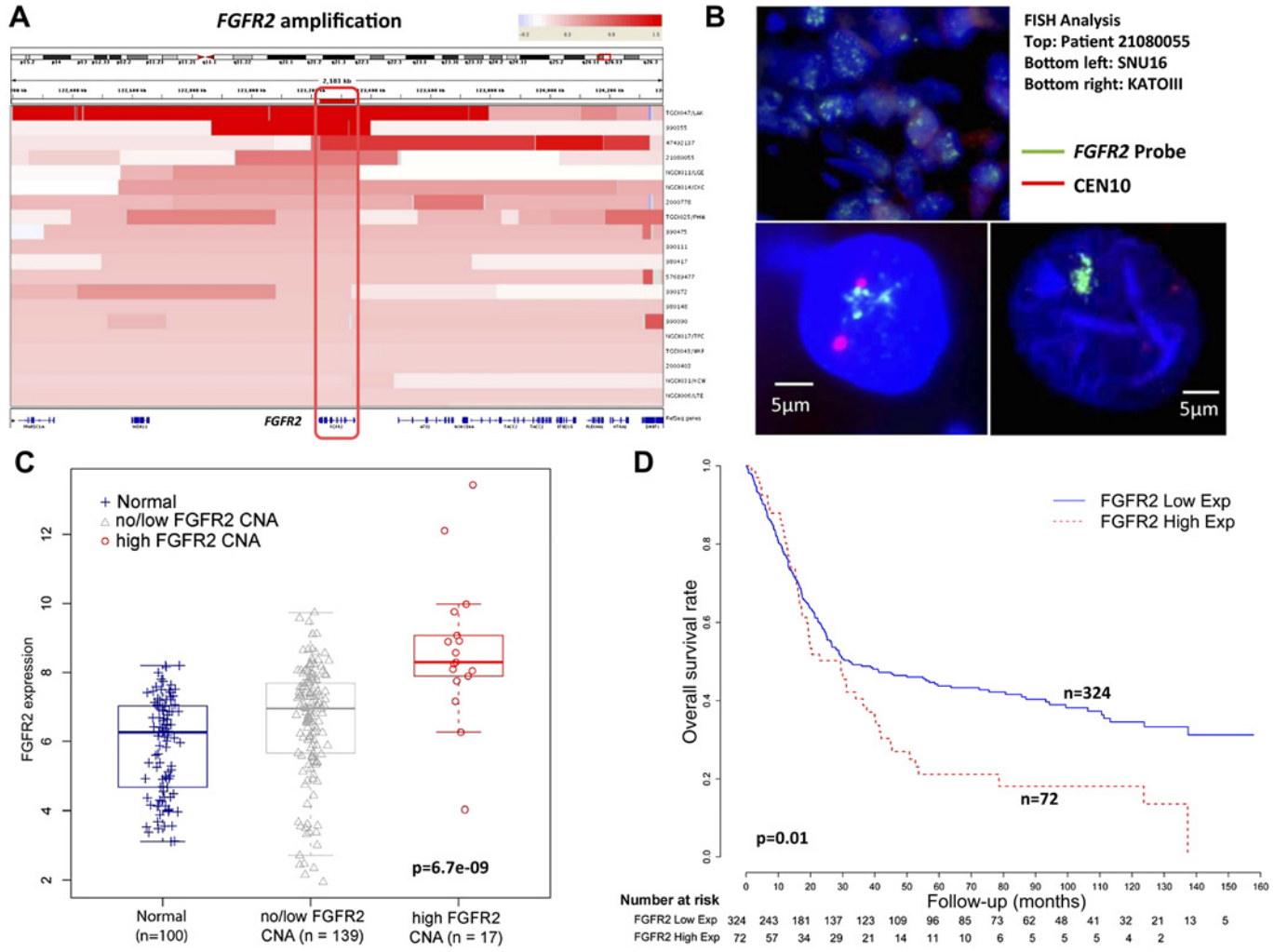


Figure 4 *FGFR2* gene amplification and messenger RNA expression in gastric cancer. (A) Heat-map showing the *FGFR2* gene amplification region in individual gastric cancer samples (20 tumours). Each row indicates one gastric cancer sample with the amplified region in red. Intensity of the red bar indicates the level of copy number amplification. Genes located in this region are shown at the bottom. The intersection of these amplified regions covers only the *FGFR2* gene (red box, gene outlined at bottom). (B) *FGFR2* genomic amplification confirmed by fluorescence in-situ hybridisation (FISH). The photo displays a patient tumour (ID 21080055) with *FGFR2* amplification and two *FGFR2*-amplified cell lines KATO-III and SNU16 confirmed by FISH analysis. Green signals indicate the *FGFR2* FISH probe, red signals probes to centremore 10. (C) *FGFR2* gene expression in clinical specimens. *FGFR2* gene expression was compared across three categories, each represented by a box-plot: non-malignant gastric tissues (normal) (n=100); tumours exhibiting no/low *FGFR2* CNA (n=139); and tumours exhibiting high *FGFR2* CNA (n=17). mRNA comparisons were based on 156 gastric cancers in which gene expression data were available, representing a subset of the 193 gastric cancers analysed by single nucleotide polymorphism arrays. *FGFR2* gene expression was inferred from Affymetrix microarrays (*FGFR2* probe 211401_s_at). *FGFR2* mRNA levels are significantly higher in samples with *FGFR2* high CNA compared with the other two categories (p=6.7e-9, Kruskal–Wallis test). Tumours exhibiting *FGFR2* amplification exhibit significantly increased *FGFR2* gene expression compared with tumours exhibiting no/low *FGFR2* CNA or non-malignant samples (p=1.9e-5 and 1.7e-7, Wilcoxon test). (D) Kaplan–Meier survival analysis comparing patients with tumours exhibiting high *FGFR2* gene expression, defined as twofold higher than the average *FGFR2* gene expression level in normal samples (72 tumours), with patients with tumours exhibiting low *FGFR2* gene expression (total 398 patients, the 156 patients analysed in figure 4C are a subset of these 398 patients). Overall survival was used as the outcome metric.

dovitinib responses with the positive control drug 5-FU. Mean tumour sizes of vehicle-treated mice reached 1163 mm^3 at day 25 post-treatment, while treatment with 5-FU at 20 mg/kg (qd × 5/week × 2 weeks, intraperitoneally) produced a reduced mean tumour size of 518 mm^3 (total growth inhibition 63%, p=0.08) after the same period. Importantly, treatment with dovitinib at 30 mg/kg and 50 mg/kg (qd × 25 days, by mouth) significantly inhibited tumour growth compared with vehicle-treated tumours (p=0.006 and 0.002, respectively), with final tumour sizes of 194 and 53 mm^3, respectively, at day 25 post-treatment (figure 5F). Dovitinib may thus represent a promising subtype-specific therapy for *FGFR2*-amplified gastric cancers.

DISCUSSION

Here we report a high-resolution genomic analysis of a large cohort of gastric cancer primary tumours and cell lines delineating the most prevalent molecular targets in this disease. While earlier reports analysing gastric cancer copy number alterations have largely analysed small patient populations or used low-resolution technologies (eg, chromosomal CGH),[22–26] these earlier studies were invaluable in benchmarking the reproducibility of our own data. For example, in a recent copy number analysis of 49 gastric cancers using Agilent 44k arrays,[27] concordant regions commonly identified in that study and ours include the frequent broad amplifications of chromosome 8 and

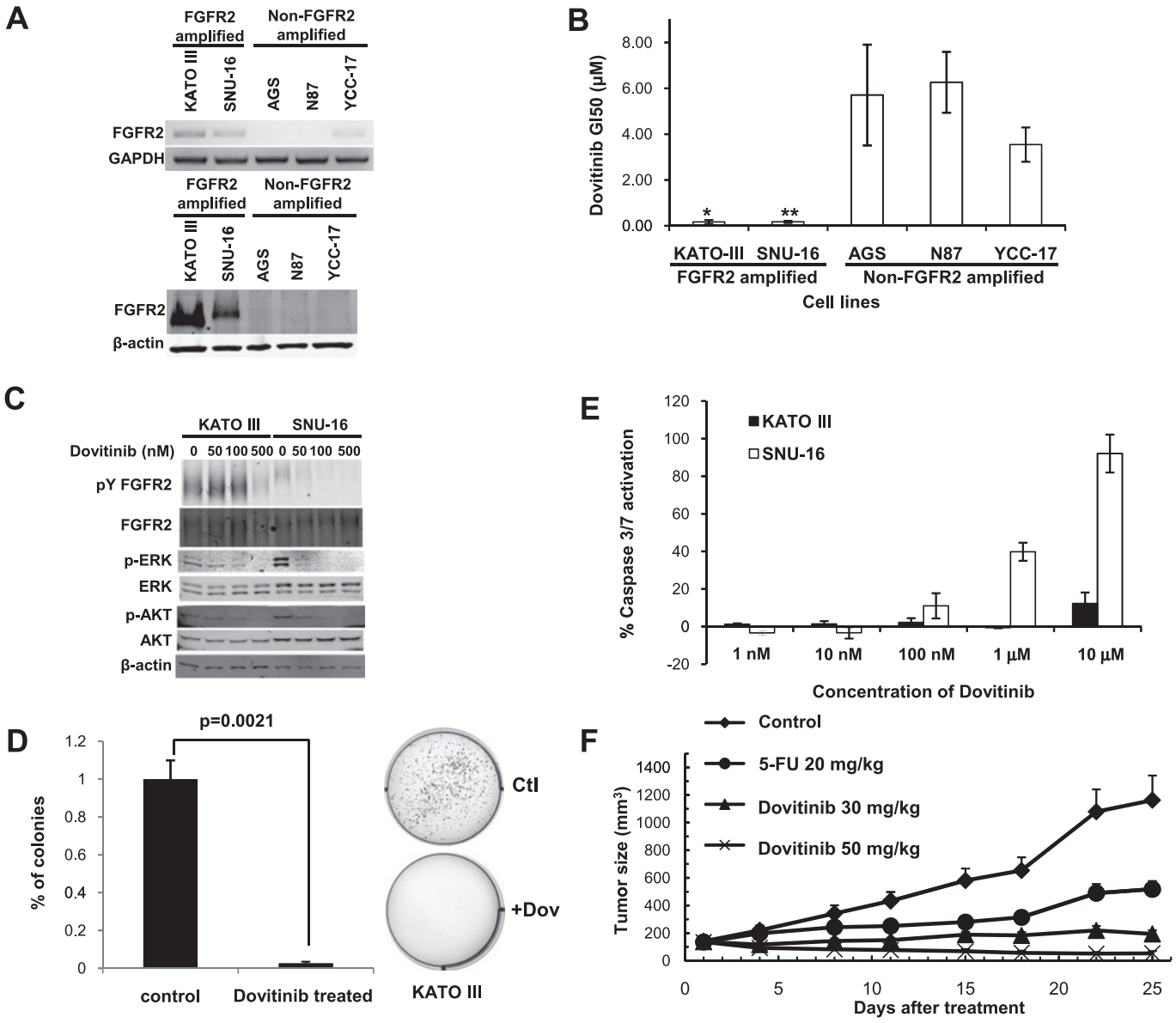


Figure 5 Sensitivity of *FGFR2*-amplified gastric cancer cell lines to dovitinib. (A) (Top) *FGFR2* reverse transcription PCR analysis of gastric cancer cell lines. Glyceraldehyde-3-phosphate dehydrogenase (GAPDH) was used as a loading control. (Bottom) *FGFR2* protein expression in lines. β-actin was used as a loading control. Cell lines KATOIII and SNU16 are observed to express elevated levels of *FGFR2* mRNA and protein. (B) Cell proliferation effects of dovitinib treatment. Dovitinib GI50 values for FGFR2-amplified and non-amplified cell lines. GI50, drug concentration required to cause 50% growth inhibition. GI50 values were calculated after 48 h dovitinib treatment. *$p<0.05$ compared with non-amplified lines. Results are a mean of three independent experiments. (C) Molecular effects of dovitinib treatment. Cells treated with dovitinib at 50 nM, 100 nM and 500 nM concentrations for 1 h. Lysates were immunoprecipitated with *FGFR2* anitbody MAB6841, and probed with 4G10 (phosphotyrosine detection) or MAB6841 for total *FGFR2*. Other antibodies included total and phospho-ERK, and total and phospho-AKT. Experiments were repeated a minimum of three independent times. (D) Dovitinib inhibits soft agar colony formation. *FGFR2*-amplified cells were treated with dovitinib at the GI50 concentration for each cell line (KATO-III 0.12 μM; SNU-16 0.17 μM) for 48 h, and soft-agar colony formation monitored over the subsequent 3–4 weeks. Data for KATO-III cells are provided, including representative colony plates. Similar results were observed for SNU16 (see supplementary figure S9, available online only). (E) Dovitinib induces caspase-3 activation. *FGFR2*-amplified cells were treated with increasing dovitinib concentrations, and apoptosis levels measured after 24 h using Caspase-Glo 3/7 assays. The y-axis represents the percentage of activation normalised against untreated controls. The results are a mean of triplicates ±SD. Experiments were repeated three independent times. (F) Dovitinib inhibits tumour growth in a human primary gastric cancer xenograft model bearing *FGFR2* gene amplification. The mean tumour size of the vehicle-treated mice reached 1163 mm^3 at day 25 post-treatment. Treatment with the positive control drug 5-FU at 20 mg/kg (qd × 5/week ×2 weeks, intraperitoneally) produced a mean tumour size of 518 mm^3 (total growth inhibition 63%, p=0.08) at the same time. Treatment with dovitinib at 30 mg/kg and 50 mg/kg (qd ×25 days, by mouth) significantly inhibited tumour growth compared with vehicle-treated animals, with a mean tumour size of 194 and 53 mm^3, respectively (p=0.006 and 0.002, respectively, at day 25 post-treatment).

20, losses of chromosome 16 and amplified genes such as *ERBB2*, *EGFR*, *GATA4*, *MYC*, *KRAS* and *CCNE1*. However, reflecting the increased size (193 vs 49) and resolution (44 K vs 1.8 million SNP probes) of our study, we also detected amplifications of chromosome 18 and deletions of chromosome 6q, which were not detected in earlier work.[22–27]

Using GISTIC, we identified 22 recurrently altered regions in gastric cancer that are likely to represent the most prevalent

molecular targets. For several of these targets, we further confirmed the SNP array results using a variety of orthogonal methodologies, including immunohistochemistry, FISH and qPCR. A survey of genes in the 22 altered regions revealed that they could be broadly partitioned into three major functional categories: RTK/RAS signalling (*FGFR2*, *KRAS*, *ERBB2*, *EGFR*, *MET*); transcriptional regulation (*MYC*, *GATA4*, *GATA6*, *KLF5*) and cell cycle control (*CCND1*, *CCNE1*, *CDK6*, *CDKN2A/B*, *RB*). As expected, many of these genes were already known to be associated with genomic alterations in gastric cancer.[6 11 28 29] Critically, however, our analysis also identified several novel genes not previously known to be amplified or deleted in gastric cancer. For example, we observed for the first time frequent deletions of *PARK2*, a E3 ubiquitin ligase, in gastric cancer.[48] Mutations in *PARK2* have been associated with early-onset Parkinson's disease,[49] and more recently *PARK2* mutations and deletions have been observed in other cancers.[50] Another novel altered gastric cancer gene was *CSMD1*, a gene of uncertain function but that has been proposed as a tumour suppressor in breast cancer.[51] Using immunohistochemistry, we confirmed that up to 40% of gastric cancers can exhibit *CSMD1* protein loss or reduced expression. Addressing the functions of these novel altered genes, given their frequency of alteration in gastric cancer, will probably be an important goal of future research work. In addition, our study also highlights interesting therapeutic opportunities—for example, the cyclin-dependent kinase *CDK6* was frequently amplified in our series, and small molecule-targeted inhibitors of CDK have been developed.[52]

A notable finding in this study was that *GATA4*, *GATA6* and *KLF5* are frequently amplified in gastric cancer. Notably, *GATA4* amplifications in gastric cancer have also been observed by other groups.[53] Intriguingly, when compared against genes identified as amplified in other comparable copy number studies from glioblastoma, lung cancer and multiple cancer types,[19–21] it appears that amplification of these three genes appears to be restricted to either gastric cancer or to other cancers related to gastrointestinal tract origin. It is possible that these genes may represent 'lineage-specific' oncogenes, a recently described class of cancer genes that enhance oncogenesis by reactivating lineage-specific survival mechanisms normally operative only in early embryonic development.[54] Examples of lineage survival oncogenes include *MITF* in melanoma, *TITF1/NKX2.1* in lung cancer[55 56] and *SOX2* in oesophageal and lung cancers.[57] Indeed, *GATA6* has recently been proposed to function as an amplified lineage-survival oncogene in pancreatic cancer,[35 58] and *KLF5* has been shown to be expressed during early development in the cardiovascular system and gastrointestinal tract epithelium in the proliferating zone of intestinal crypts.[59 60] These transcription factors may reflect the existence of an underlying transcriptional regulatory programme important for the maintenance of the gastric cancer phenotype. Interestingly, a recent genomic study from our group reported the discovery of two gastric cancer subtypes (G-INT and G-DIF) with distinct gene expression, clinical outcome and chemotherapy response features.[61] We have since discovered that G-DIF gastric cancers appear to be significantly enriched in *GATA6* gene amplifications (Fisher's exact test, p=0.04), suggesting that *GATA6* may be associated with a specific molecular subtype of gastric cancer. From a therapeutic perspective, transcription factors are commonly regarded as 'undruggable'. It is possible, however, that some of these transcription factors may regulate the expression of key genes that are pharmacologically targetable. For example, *BCL2* has been described as a target of the *MITF* transcription factor frequently amplified in melanoma,[62] and *BCL2* inhibitor drugs are available. Such a strategy may represent one method to target amplified transcription factors indirectly.

Of major clinical significance was the observation that genes related to RTK/RAS signalling are frequently altered and mutually exclusive to one another in gastric cancer. First, because numerous targeted inhibitors directed against various components of the RTK/RAS pathway are already in clinical testing,[4 9] these results raise the possibility that a substantial proportion (37% of gastric cancers) may be potentially targetable by a RTK/RAS-directed therapy. In essence, this finding dramatically increases the population of gastric cancer patients for which targeted treatments could be considered. Second, the mutually exclusive nature of these RTK/RAS alterations strongly suggests that the majority of gastric cancers are likely to have only a single RTK/RAS driver oncogene, thereby greatly simplifying the challenge of defining which RTK/RAS targeted inhibitor compound to allocate to which patient population. In terms of clinical trials, the mutually exclusive nature of the RTK/RAS alterations also renders it technically feasible to implement a multibiomarker-based trial,[63] in which multiple targeted compounds are tested in different biomarker-defined populations within a single trial design, as has been recently described for non-small-cell lung cancer (BATTLE trial).[64] Third, these results suggest that a much larger proportions of gastric cancers may be reliant on RTK/RAS signalling than previously appreciated, particularly if one notes that in this study alternative mechanisms of RTK/RAS activation were not considered, and for certain gastric cancers the presence of non-malignant cells may have reduced the sensitivity of RTK/RAS alteration detection. For example, in a recent kinome sequencing study, kinases related to MAPK signalling, a pathway downstream of *KRAS*, were identified as being the most significantly altered in gastric cancer.[65] Another alternative mechanism of RTK/RAS activation may also involve gene fusions, in which we recently described RAF-related gene rearrangements in gastric cancer.[66] Taken collectively, we believe that our finding that 37% of gastric cancers exhibit a RTK/RAS alteration should best be regarded as a lower limit, and are consistent with the notion that RTK/RAS signalling is a dominant oncogenic pathway in gastric cancer.

In our series, *FGFR2* was amplified at frequencies comparable to *ERBB2*, providing one of the first assessments of *FGFR2* gene amplification in primary gastric cancers. Interestingly, the smallest common peak of *FGFR2* amplification in the gastric cancers appears to centre around a 1.5 kb region in *FGFR2* intron 2, which overlaps a SNP locus associated with breast cancer susceptibility.[67] It is intriguing to consider whether the process of genomic amplification might also bias the expression of the *FGFR2* gene towards transcript isoforms (IIIc) that are pro-oncogenic.[68] We also found that in preclinical assays, dovitnib, a *VEGFR/FGFR2* inhibitor, can potently inhibit the growth of *FGFR2*-amplified gastric cancer cell lines and xenografts. In breast cancer, dovitinib has been found to exert effects primarily in *FGFR1*-amplified breast cancers, suggesting the importance of *FGFR*-related genome amplification in predicting dovitinib response.[69] *FGFR2* is thus likely to represent an attractive therapeutic target in gastric cancer. However, one question not addressed by our data is whether gastric cancers that lack *FGFR2* amplification, but nevertheless express *FGFR2*, will also be dovitnib responsive, as we also observed that a significant number of *FGFR2* copy-neutral tumours also exhibited elevated *FGFR2* expression levels relative to matched normal tissues, indicating that other mechanisms besides gene amplification can

also cause *FGFR2* upregulation in tumours. Notably, a recent study showed that *FGFR2* inhibition can potentially reverse chemoresistance in OCUM-2M gastric cancer cells, which are also *FGFR2* copy-number amplified.[70] We are currently addressing these questions by conducting a biopsy-mandated phase I/II trial at our centre, evaluating the efficacy of dovitinib in *FGFR2*-amplified and *FGFR2*-expressing gastric cancer samples.

Finally, our results highlight *KRAS* amplification (rather than *KRAS* mutation) as a prevalent event in gastric cancer. While *KRAS* amplifications have been reported in other cancers (eg, lung),[71] these observations have been largely anecdotal, with emphasis directed towards more conventional codon 12 and 13 activating mutations. Consistent with *KRAS* activating as an important driver gene in amplified samples, patients in our series with *KRAS*-amplified gastric cancers exhibited poor prognosis, and in vitro, *KRAS*-amplified gastric cancer lines were sensitive to *KRAS* silencing, similar to *KRAS* mutated lines. The high frequency of *KRAS* amplifications in gastric cancer is probably a major reason why *KRAS* activating mutations are strikingly infrequent in gastric cancer.[41] However, the exact mechanisms underlying this striking tissue-specific preference for *KRAS* amplification remain to be elucidated. Nevertheless, given recent data demonstrating that *KRAS*-mutated colon cancers are resistant to anti-*EGFR* therapies,[72] and that *KRAS*-amplified tumours may be resistant to MEK1/2 inhibitors,[73] our findings strongly suggest that testing *KRAS* amplification status in tumours should be fully considered in any trials evaluating RTK-targeting compounds in gastric cancer.

In conclusion, our results provide for the first time a detailed molecular map of genomic alterations in gastric cancer, which has revealed several promising targets for subtype-specific therapies. Classifying gastric cancer patients by these signature genomic alterations may facilitate patient allocations to the most appropriate clinical trials, thereby maximising patient participation in combatting this lethal disease.

Author affiliations
[1]Cancer and Stem Cell Biology Program, Duke-NUS Graduate Medical School, Singapore
[2]NUS Graduate School for Integrative Sciences and Engineering, National University of Singapore, Singapore
[3]Saw Swee Hock School of Public Health, National University of Singapore, Singapore
[4]Division of Medical Oncology, National Cancer Centre, Singapore
[5]Department of Physiology, National University of Singapore, Singapore
[6]Cellular and Molecular Research, National Cancer Centre, Singapore
[7]Department of Pathology, Singapore General Hospital, Singapore
[8]Neuroscience and Behavioral Disorders, Duke-NUS Graduate Medical School, Singapore
[9]School of Biological Sciences, Nanyang Technological University, Singapore
[10]Section of Ophthalmology and Neuroscience, Leeds Institute for Molecular Medicine, Leeds, UK
[11]Novartis Oncology, East Hanover, New Jersey, USA
[12]Department of Medicine, National University Health System, Singapore
[13]National Cancer Institute Singapore, National University Health System, Singapore
[14]Department of Internal Medicine, Yonsei Cancer Centre, Yonsei, South Korea
[15]Cancer Genomics and Biochemistry Laboratory, Peter MacCallum Cancer Centre, Melbourne, Australia
[16]Department of Pathology and Tumour Biology, Leeds Institute for Molecular Medicine, Leeds, UK
[17]Cancer Science Institute of Singapore, National University of Singapore, Singapore
[18]Genome Institute of Singapore, Singapore

Contributors ND, LKG and PT wrote the paper. ND, LKG, HW, KD, JT, SZ, IBT, ZL, GG, HG and PT analysed the data. HW, KD, JT, SZ, ML, JW, GG, QYL, ALKT, DYSP and SR conducted the experiments. KHL, MMS, RL, FZ, KGY, HCT, WPY, HCC, SYR and AB contributed data and reagents. LKG, SB, HG, SR and PT supervised the research. The first two authors contributed equally to this study.

Funding This study was supported by NMRC grants TCR/001/2007, BMRC 10/1/24/19/655, BMRC-NMRC 10/1/33/19/676 and core grants from Duke-National University of Singapore and the Cancer Sciences Institute of Singapore to PT. This work was also supported by an ASCO grant to IBT, a Singhealth talent development grant to IBT and Priscilla Ng and a Khoo discovery award (KDP/2008/0002 and KDP/2009/0006) to LKG.

Competing interests MMS and RL are employees of Novartis Pharmaceuticals Corporation. All other authors declare that they have no competing interests.

Patient consent Obtained.

Ethics approval Ethics approval was provided by the Institutional Research Ethics Review Committees of Singapore Health Services and the National University Hospital System.

Provenance and peer review Not commissioned; externally peer reviewed.

REFERENCES

1. **Brenner H,** Rothenbacher D, Arndt V. Epidemiology of stomach cancer. *Methods Mol Biol* 2009;**472**:467–77.
2. **Hartgrink HH,** Jansen EP, van Grieken NC, *et al*. Gastric cancer. *Lancet* 2009;**374**:477–90.
3. **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.
4. **Bang YJ,** Van Cutsem E, Feyereislova A, *et al*. Trastuzumab in combination with chemotherapy versus chemotherapy alone for treatment of HER2-positive advanced gastric or gastro-oesophageal junction cancer (ToGA): a phase 3, open-label, randomised controlled trial. *Lancet* 2010;**376**:687–97.
5. **Hofmann M,** Stoss O, Shi D, *et al*. Assessment of a HER2 scoring system for gastric cancer: results from a validation study. *Histopathology* 2008;**52**:797–805.
6. **Tanner M,** Hollmén M, Junttila TT, *et al*. Amplification of HER-2 in gastric carcinoma: association with topoisomerase II alpha gene amplification, intestinal type, poor prognosis and sensitivity to trastuzumab. *Ann Oncol* 2005;**16**:273–8.
7. **Gravalos C,** Jimeno A. HER2 in gastric cancer: a new prognostic factor and a novel therapeutic target. *Ann Oncol* 2008;**19**:1523–9.
8. **Weichert W,** Röske A, Gekeler V, *et al*. Association of patterns of class I histone deacetylase expression with patient prognosis in gastric cancer: a retrospective analysis. *Lancet Oncol* 2008;**9**:139–48.
9. **Yap TA,** Olmos D, Brunetto AT, *et al*. Phase I trial of a selective c-MET inhibitor ARQ 197 incorporating proof of mechanism pharmacodynamic studies. *J Clin Oncol* 2011;**29**:1271–9.
10. **Moser C,** Lang SA, Stoeltzing O. Heat-shock protein 90 (Hsp90) as a molecular target for therapy of gastrointestinal cancer. *Anticancer Res* 2009;**29**:2031–42.
11. **Turner N,** Grose R. Fibroblast growth factor signalling: from development to cancer. *Nat Rev Cancer* 2010;**10**:116–29.
12. **Kunii K,** Davis L, Gorenstein J, *et al*. FGFR2-amplified gastric cancer cell lines require FGFR2 and Erbb3 signaling for growth and survival. *Cancer Res* 2008;**68**:2340–8.
13. **Takeda M,** Arao T, Yokote H, *et al*. AZD2171 shows potent antitumor activity against gastric cancer over-expressing fibroblast growth factor receptor 2/keratinocyte growth factor receptor. *Clin Cancer Res* 2007;**13**:3051–7.
14. **Ding L,** Getz G, Wheeler DA, *et al*. Somatic mutations affect key pathways in lung adenocarcinoma. *Nature* 2008;**455**:1069–75.
15. **Rajagopalan H,** Bardelli A, Lengauer C, *et al*. Tumorigenesis: RAF/RAS oncogenes and mismatch-repair status. *Nature* 2002;**418**:934.
16. **Berns K,** Horlings HM, Hennessy BT, *et al*. A functional genetic approach identifies the PI3K pathway as a major determinant of trastuzumab resistance in breast cancer. *Cancer Cell* 2007;**12**:395–402.
17. **Bean J,** Brennan C, Shih JY, *et al*. MET amplification occurs with or without T790M mutations in EGFR mutant lung tumors with acquired resistance to gefitinib or erlotinib. *Proc Natl Acad Sci U S A* 2007;**104**:20932–7.
18. **Beroukhim R,** Getz G, Nghiemphu L, *et al*. Assessing the significance of chromosomal aberrations in cancer: methodology and application to glioma. *Proc Natl Acad Sci U S A* 2007;**104**:20007–12.
19. **Weir BA,** Woo MS, Getz G, *et al*. Characterizing the cancer genome in lung adenocarcinoma. *Nature* 2007;**450**:893–8.
20. **Cancer Genome Atlas Research Network.** Comprehensive genomic characterization defines human glioblastoma genes and core pathways. *Nature* 2008;**455**:1061–8.
21. **Beroukhim R,** Mermel CH, Porter D, *et al*. The landscape of somatic copy-number alteration across human cancers. *Nature* 2010;**463**:899–905.
22. **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.
23. **Peng DF,** Sugihara H, Mukaisho K, *et al*. Alterations of chromosomal copy number during progression of diffuse-type gastric carcinomas: metaphase- and array-based comparative genomic hybridization analyses of multiple samples from individual tumours. *J Pathol* 2003;**201**:439–50.
24. **Tada M,** Kanai F, Tanaka Y, *et al*. Prognostic significance of genetic alterations detected by high-density single nucleotide polymorphism array in gastric cancer. *Cancer Sci* 2010;**101**:1261–9.
25. **Tsukamoto Y,** Uchida T, Karnan S, *et al*. Genome-wide analysis of DNA copy number alterations and gene expression in gastric cancer. *J Pathol* 2008;**216**:471–82.

26. **Kimura Y,** Noguchi T, Kawahara K, *et al*. Genetic alterations in 102 primary gastric cancers by comparative genomic hybridization: gain of 20q and loss of 18q are associated with tumor progression. *Mod Pathol* 2004;**17**:1328—37.
27. **Rossi E,** Klersy C, Manca R, *et al*. Correlation between genomic alterations assessed by array comparative genomic hybridization, prognostically informative histologic subtype, stage, and patient survival in gastric cancer. *Hum Pathol* 2011;**42**:1937—45.
28. **Bizari L,** Borim AA, Leite KR, *et al*. Alterations of the CCND1 and HER-2/neu (ERBB2) proteins in esophageal and gastric cancers. *Cancer Genet Cytogenet* 2006;**165**:41—50.
29. **Hirono Y,** Tsugawa K, Fushida S, *et al*. Amplification of epidermal growth factor receptor gene and its relationship to survival in human gastric cancer. *Oncology* 1995;**52**:182—8.
30. **Xiao YP,** Wu DY, Xu L, *et al*. Loss of heterozygosity and microsatellite instabilities of fragile histidine triad gene in gastric carcinoma. *World J Gastroenterol* 2006;**12**:3766—9.
31. **Aqeilan RI,** Kuroki T, Pekarsky Y, *et al*. Loss of WWOX expression in gastric carcinoma. *Clin Cancer Res* 2004;**10**:3053—8.
32. **He XS,** Su Q, Chen ZC, *et al*. Expression, deletion [was deleton] and mutation of p16 gene in human gastric cancer. *World J Gastroenterol* 2001;**7**:515—21.
33. **Schneider BG,** Pulitzer DR, Brown RD, *et al*. Allelic imbalance in gastric cancer: an affected site on chromosome arm 3p. *Genes Chromosomes Cancer* 1995;**13**:263—71.
34. **Lee TL,** Leung WK, Chan MW, *et al*. Detection of gene promoter hypermethylation in the tumor and serum of patients with gastric carcinoma. *Clin Cancer Res* 2002;**8**:1761—6.
35. **Kwei KA,** Bashyam MD, Kao J, *et al*. Genomic profiling identifies GATA6 as a candidate oncogene amplified in pancreatobiliary cancer. *PLoS Genet* 2008;**4**: e1000081.
36. **Krzywinski M,** Schein J, Birol I, *et al*. Circos: an information aesthetic for comparative genomics. *Genome Res* 2009;**19**:1639—45.
37. **Scaltriti M,** Eichhorn PJ, Cortés J, *et al*. Cyclin E amplification/overexpression is a mechanism of trastuzumab resistance in HER2+ breast cancer patients. *Proc Natl Acad Sci U S A* 2011;**108**:3761—6.
38. **Asaoka Y,** Ikenoue T, Koike K. New targeted therapies for gastric cancer. *Expert Opin Invest Drugs* 2011;**20**:595—604.
39. **Sanchez-Perez I,** Garcia Alonso P, Belda Iniesta C. Clinical impact of aneuploidy on gastric cancer patients. *Clin Transl Oncol* 2009;**11**:493—8.
40. **Lievre A,** Bachet JB, Le Corre D, *et al*. KRAS mutation status is predictive of response to cetuximab therapy in colorectal cancer. *Cancer Res* 2006;**66**:3992—5.
41. **Mita H,** Toyota M, Aoki F, *et al*. A novel method, digital genome scanning detects KRAS gene amplification in gastric cancers: involvement of overexpressed wild-type KRAS in downstream signaling and cancer cell growth. *BMC Cancer* 2009;**9**:198.
42. **Ooi CH,** Ivanova T, Wu J, *et al*. Oncogenic pathway combinations predict clinical prognosis in gastric cancer. *PLoS Genet* 2009;**5**:e1000676.
43. **Trudel S,** Li ZH, Wei E, *et al*. CHIR-258, a novel, multitargeted tyrosine kinase inhibitor for the potential treatment of t(4;14) multiple myeloma. *Blood* 2005;**105**:2941—8.
44. **Lee SH,** Lopes de Menezes D, Vora J, *et al*. In vivo target modulation and biological activity of CHIR-258, a multitargeted growth factor receptor kinase inhibitor, in colon cancer models. *Clin Cancer Res* 2005;**11**:3633—41.
45. **Dey JH,** Bianchi F, Voshol J, *et al*. Targeting fibroblast growth factor receptors blocks PI3K/AKT signaling, induces apoptosis, and impairs mammary tumor outgrowth and metastasis. *Cancer Res* 2010;**70**:4151—62.
46. **Sarker D,** Molife R, Evans TR, *et al*. A phase I pharmacokinetic and pharmacodynamic study of TKI258, an oral, multitargeted receptor tyrosine kinase inhibitor in patients with advanced solid tumors. *Clin Cancer Res* 2008;**14**:2075—81.
47. **Ocio EM,** Mateos MV, Maiso P, *et al*. New drugs in multiple myeloma: mechanisms of action and phase I/II clinical findings. *Lancet Oncol* 2008;**9**:1157—65.
48. **Bedford L,** Lowe J, Dick LR, *et al*. Ubiquitin-like protein conjugation and the ubiquitin-proteasome system as drug targets. *Nat Rev Drug Discov* 2011;**10**:29—46.
49. **Healy DG,** Falchi M, O'Sullivan SS, *et al*. Phenotype, genotype, and worldwide genetic penetrance of LRRK2-associated Parkinson's disease: a case—control study. *Lancet Neurol* 2008;**7**:583—90.
50. **Veeriah S,** Taylor BS, Meng S, *et al*. Somatic mutations of the Parkinson's disease-associated gene PARK2 in glioblastoma and other human malignancies. *Nat Genet* 2010;**42**:77—82.
51. **Kamal M,** Shaaban AM, Zhang L, *et al*. Loss of CSMD1 expression is associated with high tumour grade and poor survival in invasive ductal breast carcinoma. *Breast Cancer Res Treat* 2010;**121**:555—63.
52. **Lapenna S,** Giordano A. Cell cycle kinases as therapeutic targets for cancer. *Nat Rev Drug Discov* 2009;**8**:547—66.
53. **Weiss MM,** Kuipers EJ, Postma C, *et al*. Genomic alterations in primary gastric adenocarcinomas correlate with clinicopathological characteristics and survival. *Cell Oncol* 2004;**26**:307—17.
54. **Garraway LA,** Sellers WR. Lineage dependency and lineage-survival oncogenes in human cancer. *Nat Rev Cancer* 2006;**6**:593—602.
55. **Garraway LA,** Widlund HR, Rubin MA, *et al*. Integrative genomic analyses identify MITF as a lineage survival oncogene amplified in malignant melanoma. *Nature* 2005;**436**:117—22.
56. **Kwei KA,** Kim YH, Girard L, *et al*. Genomic profiling identifies TITF1 as a lineage-specific oncogene amplified in lung cancer. *Oncogene* 2008;**27**:3635—40.
57. **Bass AJ,** Watanabe H, Mermel CH, *et al*. SOX2 is an amplified lineage-survival oncogene in lung and esophageal squamous cell carcinomas. *Nat Genet* 2009;**41**:1238—42.
58. **Alvarez H,** Opalinska J, Zhou L, *et al*. Widespread hypomethylation occurs early and synergizes with gene amplification during esophageal carcinogenesis. *PLoS Genet* 2011;**7**:e1001356.
59. **Ohnishi S,** Laub F, Matsumoto N, *et al*. Developmental expression of the mouse gene coding for the Kruppel-like transcription factor KLF5. *Dev Dyn* 2000;**217**:421—9.
60. **Conkright MD,** Wani MA, Anderson KP, *et al*. A gene encoding an intestinal-enriched member of the Kruppel-like factor family expressed in intestinal epithelial cells. *Nucleic Acids Res* 1999;**27**:1263—70.
61. **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—85; 485 e1—11.
62. **McGill GG,** Horstmann M, Widlund HR, *et al*. Bcl2 regulation by the melanocyte master regulator Mitf modulates lineage survival and melanoma cell viability. *Cell* 2002;**109**:707—18.
63. **Printz C.** BATTLE to personalize lung cancer treatment. Novel clinical trial design and tissue gathering procedures drive biomarker discovery. *Cancer* 2010;**116**:3307—8.
64. **Kim ES,** Herbst RS, Wistuba II, *et al*. The BALLTE trial: personalizing therapy for lung cancer. *Cancer Discovery* 2011;**1**:44—53.
65. **Zang ZJ,** Ong CK, Cutcutache I, *et al*. Genetic and structural variation in the gastric cancer kinome revealed through targeted deep sequencing. *Cancer Res* 2011;**71**:29—39.
66. **Palanisamy N,** Ateeq B, Kalyana-Sundaram S, *et al*. Rearrangements of the RAF kinase pathway in prostate cancer, gastric cancer and melanoma. *Nat Med* 2010;**16**:793—8.
67. **Hunter DJ,** Kraft P, Jacobs KB, *et al*. A genome-wide association study identifies alleles in FGFR2 associated with risk of sporadic postmenopausal breast cancer. *Nat Genet* 2007;**39**:870—4.
68. **Katoh Y,** Katoh M. FGFR2-related pathogenesis and FGFR2-targeted therapeutics (Review). *Int J Mol Med* 2009;**23**:307—11.
69. **Andre F,** *et al*. A multicenter, open-label phase II trial of dovitinib, an FGFR1 inhibitor, in FGFR1 amplified and non-amplified metastatic breast cancer. *J Clin Oncol* 2011;**29**:(ASCO meeting abstract):508.
70. **Qiu H,** Yashiro M, Zhang X, *et al*. A FGFR2 inhibitor, Ki23057, enhances the chemosensitivity of drug-resistant gastric cancer cells. *Cancer Lett* 2011;**307**:47—52.
71. **Wagner PL,** Stiedl AC, Wilbertz T, *et al*. Frequency and clinicopathologic correlates of KRAS amplification in non-small cell lung carcinoma. *Lung Cancer* 2011;**74**:118—23.
72. **Van Cutsem E,** Köhne CH, Hitre E, *et al*. Cetuximab and chemotherapy as initial treatment for metastatic colorectal cancer. *N Engl J Med* 2009;**360**:1408—17.
73. **Little AS,** Balmanno K, Sale MJ, *et al*. Amplification of the driving oncogene, KRAS or BRAF, underpins acquired resistance to MEK1/2 inhibitors in colorectal cancer cells. *Sci Signal* 2011;**4**:ra17.

A Comprehensive Survey of Genomic Alterations in Gastric Cancer Reveals Systematic Patterns of Molecular Exclusivity and Co-Occurrence among Distinct Therapeutic Targets

(Supplementary Information)

Niantao Deng*[,1,2], Liang Kee Goh*[,1,3,4], Hannah Wang[1], Kakoli Das[1], Jiong Tao[1,5], Iain Beehuat Tan[1,2,4], Shenli Zhang[1], Minghui Lee[6], Jeanie Wu[6], Kiat Hon Lim[7], Zhengdeng Lei[8], Glenn Goh[1], Qing-Yan Lim[9], Angie Lay-Keng Tan[1], Dianne Yu Sin Poh[1], Sudep Riahi[10], Sandra Bell[10], Michael M. Shi[11], Ronald Linnartz[11], Feng Zhu[12], Khay Guan Yeoh[12], Han Chong Toh[4], Wei Peng Yong[13], Hyun Cheol Cheong[14], Sun Young Rha[14], Alex Boussioutas[15], Heike Grabsch[16], Steve Rozen[8], Patrick Tan[1,6,17,18,**]

[1]Cancer and Stem Cell Biology Program, Duke-NUS Graduate Medical School, Singapore
[2]NUS Graduate School for Integrative Sciences and Engineering/[3]Saw Swee Hock School of Public Health, National University of Singapore, Singapore
[4]Division of Medical Oncology, National Cancer Centre, Singapore
[5]Department of Physiology, National University of Singapore, Singapore
[6]Cellular and Molecular Research, National Cancer Centre, Singapore
[7]Dept of Pathology, Singapore General Hospital, Singapore
[8]Neuroscience and Behavioral Disorders, Duke-NUS Graduate Medical School, Singapore
[9]School of Biological Sciences, Nanyang Technological University, Singapore
[10]Section of Ophthalmology and Neuroscience, Leeds Institute for Molecular Medicine, Leeds, England
[11]Novartis Oncology, East Hanover, New Jersey, USA
[12]Department of Medicine/[13]National Cancer Institute Singapore, National University Health System, Singapore
[14]Department of Internal Medicine, Yonsei Cancer Centre, South Korea
[15]Cancer Genomics and Biochemistry Laboratory, Peter MacCallum Cancer Centre, Australia
[16]Department of Pathology and Tumour Biology, Leeds Institute for Molecular Medicine, Leeds, England
[17]Cancer Science Institute of Singapore, National University of Singapore, Singapore
[18]Genome Institute of Singapore, Singapore
*These authors contributed equally to this study
** **Correspondence to**
Patrick Tan
Cancer and Stem Cell Biology Program, Duke-NUS Graduate Medical School, 8 College Road, Singapore, 169857;
Tel: (65) 6516 1783; Fax: (65) 6221 2402
gmstanp@duke-nus.edu.sg

Keyword: copy number alterations, receptor tyrosine kinases, mutual exclusivity, targeted therapies

Supplementary Items

Materials and methods

Supplementary Text

Supplementary Tables

Supplementary Figures

Figure S11: *FGFR2* Overexpression in GCs Relative to Normal Gastric Samples

Figure S12: Inhibition of Soft Agar Colony Growth by Dovitinib (SNU-16)

Materials and Methods

Clinical Samples and Cell Lines

Primary gastric samples were obtained from the Singapore Health Services (SingHealth) and the National University Hospital System (NUHS) tissue repositories, with signed informed patient consent and approvals from the respective institutional Research Ethics Review Committees. Clinical information was collected with Institutional Review Board approval. There was no pre-specified sample size calculation since this is a hypothesis generating discovery study. Clinical characteristics of patients analyzed in this study are presented in Supplementary Table S1. GC cell lines were obtained from commercial sources (American Type Culture Collection, Japan Health Science Research Resource Bank) or from collaborators (Yonsei Cancer Centre, S. Korea).

DNA and RNA Extraction

Genomic DNA was extracted from flash-frozen tissues and cells using a Qiagen genomic DNA extraction kit. Total RNAs was extracted using Trizol (Invitrogen, CA), digested with RNase free DNase (RQ1 DNase, Promega), and subsequently purified using an RNeasy Mini kit (Qiagen,CA).

Copy Number Profiling and GISTIC Analysis

Genomic DNAs from gastric tumors and matched non-malignant gastric tissues (normal) were hybridized on Affymetrix SNP6 genotyping arrays and processed as follows:
Step 1) Normalization: Raw SNP6 CEL files were processed using Affymetrix Genotyping Console 4.0. A reference file was first created from the SNP6 CEL files of normal gastric samples (98 samples). The 193 tumor SNP6 CEL files were then normalized against this normal reference file.
Step 2) Segmentation: Copy number segmentation data was produced using the Circular Binary Segmentation (CBS) algorithm using the R package *DNAcopy* [1] for both tumor and normal gastric samples. The p value cutoff for detecting a change-point was 0.01, with a permutation number of 10000.

3) GISTIC Analysis: The GISTIC algorithm [2] was used to identify genomic regions with recurrent copy number alterations. GISITC was applied to the CBS-segmented files of tumors, and filtered through a CNV (copy number variation) file constructed from the segmented data of normal samples to identify somatic tumor-specific CNAs. GISTIC reports regions of interest with an associated q-value, which is obtained by multiple hypotheses correction. Genomic regions with q-value<0.25 for broad regions and q-value<0.001 for focal regions were considered significant. Proportions of CNA for individual normal and tumor sample was defined as: size of CBS regions with CNA per sample divided by the sum of all autosome lengths. Chromosomal instability values for GCs were estimated by the number of cytobands exhibiting CNA for each sample, calculated by averaging the CBS segmented value for each cytoband.
The SNP6 copy number data has been deposited into the National Centre for Biotechnology Information's (NCBI) Gene Expression Omnibus (GEO) website, series accession number GSE31168.The Reviewer link is http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?token=tbqtnokgaooucjo&acc=GSE31168

DRP: Identification of Mutually Exclusive and Co-Altered CNAs

To identify significant relationships between regions of frequent CNA, we implemented a dimension reduction permutation (DRP) statistical algorithm adapted from a previous study analyzing patterns of somatic DNA mutations in tumor [3]. To determine the significance of any specific mutually exclusive (ME) or co-alteration (CA) interaction, we compared the numbers of samples exhibiting a particular ME or CA interaction against a null distribution of interactions obtained by randomly permuting the genomic alterations across samples and genes (100,000 permutations), while taking into consideration the prevalence of genomic alterations. Essentially, for each permutation, we constrained the number of samples with genomic alterations and the number of genes exhibiting alterations within each sample to be similar to the original data. Empirical p-values of <0.05 were considered significant. An in-depth description of the DRP methodology is presented in Text S1, and the DRP software can be downloaded from http://research.duke-nus.edu.sg/papers/DRP.zip.

FISH and Immunohistochemical Analysis

KRAS and *FGFR2* FISH was performed using BAC clones obtained from the BACPAC resources center (CHORI, Oakland, CA USA). BAC DNA was labeled using a Bioprime DNA labeling kit (Invitrogen, Carlsbad, CA, USA). FISH was performed on metaphase spreads (cell lines) or on FFPE sections after deparaffinization (clinical specimens). Target DNA probes were labeled using spectrum green and control probes in spectrum orange (centromeric CEP probes for chromosomes 10 and 12) (Abbott Molecular Inc, Des Plaines, IL, USA). Hybridized slides were counterstained with DAPI and analyzed using a Olympus BX50 fluorescence microscope. Nuclei were scored for amplification by comparing signals from internal controls (CEP probes) against target gene signals (*KRAS*, and *FGFR2*). For *ERBB2* immunohistochemistry, we analyzed 146 of the 193 tumors, representing all cases for which we were able to obtain full sections. The remaining 47 cases were not analyzed for a variety of reasons, including failure to retrieve the samples due to historical storage arrangements (archival samples are stored off-site at our center) and insufficient material due to exhaustion of the FFPE blocks (small tumors). Sections of archival formalin-fixed, paraffin-embedded tissue (3 μm) were placed on slides coated with poly-L-lysine. After deparaffinisation and blocking of endogenous peroxidase, *ERBB2* immunostaining was performed using rabbit anti-human c-erbB-2 oncoprotein as primary antibody (Dako Corp, Carpinteria, CA, USA) at 1/100 dilution. Binding of the primary antibody was revealed by means of the Dako Quick-Staining, Labelled Streptavidin–Biotin System (Dako), followed by the addition of diaminobenzidine as a chromogen. *ERBB2* immunoreactivity was evaluated by an experienced pathologist (LKH) according to the scoring system of [4]. CSMD1 immunohistochemistry was performed on full sections as described in [5]. Tumors were scored by two independent observers (HG, SB) and classified as *CSMD1* present (> 25% positive positive tumour cells) or *CSMD1* Absent/Reduced (<= 25% positive tumor cells).

DNA Sequencing, Mutation Genotyping and Quantitative PCR

DNA products corresponding to the coding regions of target genes were amplified by PCR and were subjected to cycle sequencing using the BigDye Terminator v3.1 Cycle

Sequencing Kit (Applied Biosystems, Foster City, CA, USA) on a 3730xl DNA Analyzer (Applied Biosystems, Foster City, CA, USA). *KRAS* mutation genotyping was performed by both Sanger sequencing (139 GCs) and mass-spectrometry based genotyping (Sequenom MassARRAY) (94 GCs). Reference sequences were obtained from the Ensembl Genome Browser database. Quantitative real-time PCR was performed on an ABI 7900 HT instrument using *FGFR2* intron 2 primers. Reaction mixes consisted of 5ul SYBR green PCR master mix (ABI), 1ul *FGFR2/LINE1* primers, 20ng (0.5ul) of genomic DNA template in a final reaction volume of 10ul. All experiments were performed in triplicate. *FGFR2* cycle thresholds were normalized to the *LINE1* repeat element from the same samples, as an endogenous control. Normal human genomic DNA was chosen as the calibrator and for each analysis a negative control was also prepared using all reagents except DNA template.

Gene Expression Analysis

Of the 193 tumors profiled on Affymetrix SNP6 arrays (Affymetrix, Santa Clara, CA, USA), 156 tumors had corresponding gene expression data available along with 100 normal gastric samples on Affymetrix U133P2 arrays (this cohort is analyzed in Figure 4C). Additional details of the gene expression data set are presented in [6] and are publicly available at GEO under accession number GSE15460. To analyze *FGFR2* mRNA survival associations in Figure 4D, we analyzed a combined GC gene expression data set of 398 tumors. The 156 patients analyzed in Figure 4C form a subset of the 398 patients. To establish this combined data set, we combined gene expression data from GSE15460 and three other GC cohorts from Singapore (U133AB), Australia (AU) and the University of Leeds, UK (UK). Clinical information for these gene expression data sets is provided in Table S6. Briefly, individual arrays were normalized using the MAS5 algorithm, and batch effects removed using the COMBAT algorithm [7].

Clinico-Pathologic Correlation Analysis

Survival curves were estimated using the Kaplan-Meier method, with the duration of survival measured from the date of surgery to date of death or last follow-up visit. Overall survival was used as the outcome metric. Patients who were still alive or lost to

follow-up at time of analysis were censored at their last date of follow up. Univariate and multivariate survival analysis was performed using the Cox proportional hazards regression model. Besides genetic factors (e.g. *FGFR2*, *KRAS*), other clinical factors considered in the multivariate model included grade and stage which were also significant in univariate analysis. Associations with other clinical variables were performed using the Fisher Exact Test, at a significance threshold of $p<0.05$.

Reverse Transcription-PCR (RT-PCR) and Western Blotting Analysis

For mRNA analysis, equal quantities of RNA were reverse transcribed using SuperScript III Reverse Transcriptase enzyme and oligo$(dT)_{20}$ primers (Invitrogen). RT-PCR was performed with forward primers to *FGFR2* exon 8 (5'- GTGCTTGGCGGGTAATTCTA-3') and reverse primers to exon 9 (5'-TACGTTTGGTCAGCTTGTGC -3'). *GAPDH* was used as a loading control (forward primer (5'- GTGCTTGGCGGGTAATTCTA-3'); reverse primer (5'-TCCACCACCCTGTTGCTGTA-3'). For protein analysis, cells were harvested in lysis buffer (0.3M NaCl, 0.05M Tris-HCl pH8, 0.5% NP40, 0.1% SDS, Protease Inhibitor (Roche, Mannhein, Germany) and Halt Phosphatase Inhibitor Cocktail (Pierce, Rockford, IL, USA)). *FGFR2* immunoprecipitation was performed by incubating lysates with MAB6841 (R&D Systems, Minneapolis, MN, USA) for 4 hrs at room temperature; followed by incubation with protein A/G agarose beads (Pierce, Rockford, IL, USA) overnight at 4°C. After washing, 4X SDS loading buffer was added and the mixture was boiled at 95°C for 5 minutes. Antibodies against p-ERK, ERK, p-AKT, AKT and Caspase-3(8G10) were obtained from Cell Signaling Technology (Cell Signaling Technology, Danvers,. MA, USA). Other antibodies include 4G10 phosphotyrosine antibody (Upstate Biotechnology, Lake Placid, NY, USA β-actin (Millipore, Billerica, MA, USA) or α-tubulin (Cell Signaling Technologies, Danvers, MA, USA) were used as loading controls. Blots were incubated with DyLight Fluorescence secondary antibodies (Thermo Scientific) and imaged using LI-COR Odyssey. Experiments were repeated a minimum of three independent times.

Cell Proliferation Assays and Drug Treatments

Cell proliferation assays were performed using the CellTiter 96® AQ_{ueous} One Solution Assay kit (Promega) and the plates were measured using a PerkinElmer plate reader. Each assay was performed in triplicate, and the results were averaged over three independent experiments. Dovitinib was provided by Drs. D. Graus-Porta and C. Garcia-Echeverria (Novartis Institutes for Biomedical Research, Basel, Switzerland). GC cells were seeded in 96-well plates 24 hours prior to Dovitinib treatment. On the day of drug treatment, CellTiter reagent was added to one plate of cells to provide a measurement of the cell population at the time of drug addiction (T_z). Five serial 10-fold dilution mixtures of Dovitinib, beginning with a maximum concentration of 10^{-5} M, were added to the respective wells. The final DMSO concentration in the wells did not exceed 0.1% (v/v). GI50 values for Dovitinib, representing the concentration at which 50% cell growth inhibition is achieved for 48 hours of treatment, were computed using the GI50 calculation formula at http://dtp.nci.nih.gov/branches/btb/ivclsp.html.

Cell Death and Colony Formation Assays

Caspase 3/7 assays were performed using the Caspase-Glo® 3/7 Assay kit (Promega, WI, USA) and the plates were measured using a Tecan plate reader. Three independent experiments were performed and each assay was performed in triplicate. GC cells were seeded in 96-well black plates and treated with Dovitinib using the same method as the cell proliferation assays. For colony formation assays, base layers of 0.5% Gum Agar in 1x McCoy's 5A and 10% FBS were poured into 6-well plates and allowed to harden at 4°C. After siRNA transfection, overexpression, or drug treatment, 50 000 cells/well were seeded in complete media plus agar mixture at 42°C and seeded on top of the solidified base layer. Plates were incubated at 37°C in for 3-4 weeks, during which plates were fed drop-wise with complete media. After 3-4 weeks, plates were photographed using the Kodak GL 200 System (EpiWhite illumination). Each assay was performed in triplicate, and the results were averaged over three independent experiments.

Xenograft assays

Efficacy of dovitinib was evaluated and compared to the positive control drug 5-FU in a primary human gastric cancer xenograft model (n= 10 in each group). This tumor model

was derived from a primary gastric cancer from Chinese ethnicity and is confirmed with *FGFR2* gene amplification (26 copies of *FGFR2* by SNP6.0 array). Tumor fragments from stock mice inoculated with selected primary human gastric cancer tissues were harvested and used for inoculation into Balb/c nude mice. Each mouse was inoculated subcutaneously at the right flank with primary human gastric tumor fragment (2-3 mm in diameter) for tumor development. Treatments were started at day 24 after tumor inoculation when the average tumor size reached about 150 mm^3.

Text S1: Dimension Reduction Permutation (DRP): Identification of Mutually Exclusive and Co-Altered CNAs

Non-random associations between distinct genomic alterations (co-associated or mutually exclusive) may suggest synergistic or antagonistic biological event in carcinogenesis. To compute the significance of these associations, a dimension reduction permutation (DRP) algorithm was developed. It was adapted from a previous study analyzing patterns of somatic DNA mutations in tumor [3]. To determine the significance of any pair of mutually exclusive or co-altered CNAs, we used permutation testing, taking into consideration the prevalence of genomic alterations. Since we are testing for associations regardless of the level of alterations (i.e. focal or broad), we assigned each gene to either an amplification or deletion status, based on the mean aggregation of log ratio signals of all probes within each gene. To maintain a similar prevalence of genomic alterations observed in the original data, the number of samples with genomic alterations and the number of genes exhibiting the alterations were maintained in the permutations. Suppose the matrix is represented as genes (row) x samples (column). DRP permutes the genomic alterations by row or by column progressively, depending on which number of rows or columns is smaller. Permutations can start from the top row or the left column of the matrix while maintaining the marginal counts for genomic alterations in genes and samples to be similar to the original data. In effect, for each permutation, the algorithm traverses iteratively from top left to bottom right of the matrix, each time reducing the dimension by multiple numbers of rows and columns – hence the name Dimension Reduction Permutation. For each permutation, the number of samples with co-altered (N_{CA}) and mutually exclusive CNA (N_{ME}) was then recorded for each pair of genes and then compared with original data on co-altered (O_{CA}) and mutually exclusive genes (O_{ME}) respectively. Frequencies were summarized for co-altered ($N_{CA}>=O_{CA}$) and mutually exclusive associations ($N_{ME}>=O_{ME}$). Empirical p-values were then computed against these frequencies under the null hypothesis.

Supplementary Tables

Table S1. Clinical Characteristics of the GC Patient Cohort.

This table provides clinical data for 193 patients analyzed by Affymetrix SNP6 arrays. Stage categories were based on the AJCC 6th edition classification**.** 3 patients received neoadjuvant therapy, and of 131 patients where subsequent treatment information was available, 28 patients received 5-FU chemoradiation as adjuvant therapy.

	GC Samples (193)
Age	
Range	23-92
Mean,S.D	64.2, 12.6
Gender	
Male	123
Female	70
Lauren Classification	
Intestinal	99
Diffuse	73
Mixed/Others	21
Anatomical Location*	
Gastro-oesophageal junction	9
Cardia	13
Body	24
Greater Curve	17
Lesser Curve	37
Pylorus	12
Antrum	22
Incisura	2
Grade	
Undifferentiated	2
Poorly differentiated	117
Moderately differentiated	67
Well differentiated	5
Unknown	2
Stage	
1	32
2	26
3	71
4	64

*This is only for 136 patients where location information was reliably recorded.

Table S2: Concordance Table between *ERBB2* SNP6 and *ERBB2* IHC

9 of 132 (6.8%) *ERBB2* copy number neutral tumors exhibit ERBB2 protein expression (IHC 1-3+), while 8 of 13 (61.5%) tumors with *ERBB2* copy number gain also exhibit ERBB2 protein expression ($p<0.01$, Fisher's exact test).

	ERBB2 Immunohistochemistry				
ERBB2 SNP 6 Copy Number	Positive staining	0	1+	2+	3+
Loss (logRatio<-0.2)	0 / 1 (0 %)	1	0	0	0
Neutral (-0.2 < logRatio < 0.2)	9 / 132 (6.8%)	123	2	3	4
Gain (logRatio > 0.2)	8 / 13 (61.5%)	5	2	3	3

Table S3: DRP Analysis of Mutually Exclusive and Co-Amplification Interactions

This table lists all significant mutually exclusive (ME) and co-occurring (CO) interactions for a pair of genes ('Gene1' and 'Gene2'). The columns are : '#Gene1' and '#Gene2' are the observed frequency of amplification for each pair of genes. '#Both' indicates the observed number of cases of coamplification for this pair of genes, and '#OnlyOne' indicates the observed number of cases for amplification in only one of this pair of genes. '#BothExp' and '#OnlyOneExp' are the expected results from the DRP permutation for coamplification cases and non-coamplification cases. 'PvalueME' and 'PvalueCO' are the empirical pvalues for ME and CO interactions. 'QvalueME' and 'QvalueCO' are converted Storey's qvalue. Gene pairs related to RTK/RAS signaling are highlighted. Significant ME interactions are at the top of the list, while significant CO interactions are at the bottom.

Gene1	Gene2	#Gene1	#Gene2	#Both	#OnlyOne	#BothExp	#OnlyOneExp	PvalueME	QvalueME	PvalueCO	QvalueCO
FGFR2	*KLF5*	22	22	0	44	5.765	32.470	**0.001**	0.118	0.999	0.999
GATA4	*KLF5*	23	22	1	43	5.840	33.320	**0.010**	0.464	0.990	0.999
KRAS	*ERBB2*	21	17	1	36	5.191	27.619	**0.018**	0.464	0.982	0.999
FGFR2	*MET*	22	14	1	34	4.681	26.637	**0.028**	0.464	0.972	0.999
CCNE1	*MET*	23	14	1	35	4.692	27.615	**0.028**	0.464	0.972	0.999
ERBB2	*MET*	17	14	1	29	4.558	21.884	**0.031**	0.464	0.969	0.999
CCNE1	*GATA4*	23	23	2	42	5.904	34.193	**0.042**	0.470	0.958	0.999
CCNE1	*KRAS*	23	21	2	40	5.744	32.513	**0.048**	0.470	0.952	0.999
FGFR2	*KRAS*	22	21	2	39	5.696	31.607	**0.049**	0.470	0.951	0.999
KRAS	*EGFR*	21	21	2	38	5.634	30.733	0.052	0.470	0.948	0.999
GATA4	*ERBB2*	23	17	2	36	5.248	29.503	0.070	0.578	0.930	0.999
GATA6	*CDH12*	25	14	2	35	4.687	29.625	0.105	0.602	0.895	0.999
CCND1	*MET*	24	14	2	34	4.692	28.616	0.106	0.602	0.895	0.999
CDH12	*CCND1*	14	24	2	34	4.694	28.611	0.106	0.602	0.894	0.999
GATA4	*MET*	23	14	2	33	4.685	27.631	0.106	0.602	0.894	0.999

CCNE1	*CDK6*	23	26	3	43	6.034	36.933	0.112	0.602	0.888	0.999
GATA6	*KLF5*	25	22	3	41	5.917	35.167	0.119	0.602	0.881	0.999
FGFR2	*CCNE1*	22	23	3	39	5.826	33.349	0.127	0.602	0.873	0.999
KRAS	*CCND1*	21	24	3	39	5.791	33.418	0.131	0.602	0.869	0.999
FGFR2	*EGFR*	22	21	3	37	5.700	31.599	0.137	0.602	0.863	0.999
CDH12	*MET*	14	14	2	24	4.316	19.368	0.139	0.602	0.861	0.999
CDK6	*ERBB2*	26	17	3	37	5.282	32.435	0.181	0.741	0.820	0.999
FGFR2	*ERBB2*	22	17	3	33	5.232	28.536	0.187	0.741	0.813	0.999
CCNE1	*CCND1*	23	24	4	39	5.963	35.074	0.249	0.763	0.751	0.999
GATA4	*CCND1*	23	24	4	39	5.962	35.075	0.250	0.763	0.750	0.999
FGFR2	*CDH12*	22	14	3	30	4.690	26.620	0.258	0.763	0.742	0.999
CDH12	*GATA4*	14	23	3	31	4.684	27.633	0.260	0.763	0.740	0.999
GATA6	*MET*	25	14	3	33	4.684	29.632	0.260	0.763	0.740	0.999
KRAS	*CDH12*	21	14	3	29	4.670	25.660	0.262	0.763	0.738	0.999
CDH12	*EGFR*	14	21	3	29	4.669	25.663	0.263	0.763	0.737	0.999
KRAS	*GATA6*	21	25	4	38	5.817	34.366	0.268	0.763	0.732	0.999
EGFR	*ERBB2*	21	17	4	30	5.204	27.592	0.367	0.960	0.633	0.976
FGFR2	*CDK6*	22	26	5	38	5.950	36.101	0.431	0.960	0.569	0.909
KLF5	*CCND1*	22	24	5	36	5.888	34.224	0.440	0.960	0.561	0.909
EGFR	*CCND1*	21	24	5	35	5.786	33.427	0.459	0.960	0.541	0.909
GATA4	*EGFR*	23	21	5	34	5.750	32.500	0.466	0.960	0.534	0.909
KRAS	*GATA4*	21	23	5	34	5.735	32.529	0.468	0.960	0.532	0.909
CDH12	*CDK6*	14	26	4	32	4.689	30.622	0.472	0.960	0.528	0.909
KRAS	*MET*	21	14	4	27	4.663	25.674	0.474	0.960	0.526	0.909
EGFR	*KLF5*	21	22	5	33	5.698	31.604	0.475	0.960	0.525	0.909
CDH12	*ERBB2*	14	17	4	23	4.559	21.882	0.499	0.982	0.501	0.909
GATA6	*ERBB2*	25	17	5	32	5.273	31.455	0.561	0.982	0.439	0.897
ERBB2	*CCND1*	17	24	5	31	5.259	30.482	0.565	0.982	0.435	0.897
ERBB2	*KLF5*	17	22	5	29	5.231	28.539	0.568	0.982	0.432	0.897
GATA6	*CDK6*	25	26	6	39	6.201	38.598	0.568	0.982	0.432	0.897

CDK6	*GATA4*	26	23	6	37	6.052	36.896	0.596	0.982	0.404	0.897
FGFR2	*MYC*	22	46	6	56	6.023	55.955	0.604	0.982	0.397	0.897
CCNE1	*GATA6*	23	25	6	36	6.010	35.979	0.604	0.982	0.396	0.897
GATA6	*GATA4*	25	23	6	36	6.007	35.985	0.606	0.982	0.394	0.897
CDK6	*KLF5*	26	22	6	36	5.951	36.098	0.615	0.982	0.385	0.897
KRAS	*CDK6*	21	26	6	35	5.839	35.322	0.637	0.999	0.363	0.897
KRAS	*KLF5*	21	22	6	31	5.694	31.612	0.664	1.000	0.336	0.897
KLF5	*MET*	22	14	5	26	4.687	26.626	0.685	1.000	0.315	0.895
MYC	*ERBB2*	46	17	6	51	5.289	52.423	0.740	1.000	0.260	0.789
FGFR2	*GATA6*	22	25	7	33	5.922	35.157	0.780	1.000	0.220	0.714
GATA6	*EGFR*	25	21	7	32	5.815	34.370	0.797	1.000	0.203	0.685
CCNE1	*EGFR*	23	21	7	30	5.734	32.533	0.810	1.000	0.190	0.664
CDH12	*KLF5*	14	22	6	24	4.677	26.646	0.846	1.000	0.154	0.560
EGFR	*MET*	21	14	6	23	4.678	25.644	0.847	1.000	0.153	0.560
CDK6	*CCND1*	26	24	8	34	6.125	37.751	0.866	1.000	0.134	0.529
GATA6	*CCND1*	25	24	8	33	6.065	36.870	0.873	1.000	0.127	0.524
MYC	*KRAS*	46	21	8	51	5.894	55.211	0.892	1.000	0.108	0.466
FGFR2	*CCND1*	22	24	8	30	5.881	34.237	0.895	1.000	0.105	0.466
FGFR2	*GATA4*	22	23	8	29	5.847	33.307	0.898	1.000	0.102	0.466
CCNE1	*KLF5*	23	22	8	29	5.838	33.324	0.898	1.000	0.102	0.466
MYC	*CCND1*	46	24	9	52	6.222	57.556	0.932	1.000	0.068	0.365
MYC	*MET*	46	14	7	46	4.695	50.610	0.938	1.000	0.062	0.351
CCNE1	*ERBB2*	23	17	8	24	5.247	29.505	0.950	1.000	0.051	0.306
CCNE1	*CDH12*	23	14	8	21	4.676	27.649	0.982	1.000	**0.019**	0.140
MYC	*GATA6*	46	25	11	49	6.306	58.388	0.988	1.000	**0.012**	0.111
MYC	*GATA4*	46	23	11	47	6.124	56.752	0.991	1.000	**0.009**	0.090
MYC	*CCNE1*	46	23	11	47	6.129	56.741	0.991	1.000	**0.009**	0.090
CDK6	*EGFR*	26	21	11	25	5.849	35.303	0.995	1.000	**0.005**	0.062
MYC	*EGFR*	46	21	12	43	5.896	55.209	0.999	1.000	**0.001**	0.023
CDK6	*MET*	26	14	10	20	4.694	30.613	0.999	1.000	**0.001**	0.015

MYC	*CDH12*	46	14	10	40	4.698	50.603	0.999	1.000	**0.001**	0.015
MYC	*KLF5*	46	22	13	42	6.021	55.959	0.999	1.000	**5.00E-04**	0.015
MYC	*CDK6*	46	26	15	42	6.375	59.249	1.000	1.000	**1.00E-04**	0.005

Table S4a: Multivariate analysis comparing RTK amplification status with tumor stage, grade, adjuvant treatment and genome instability
(Outcome: overall survival, relative to patients lacking RTK amplification).

Model 1 (Predictors: RTK Amp, Stage ,Grade and Adjuvant Treatment)	**Hazard Ratio (95% CI)**	**P-value**
RTK Amp vs RTK Absent	1.966 (1.180, 3.279)	**0.01**
Stage 2 vs Stage 1	2.329 (0.867, 6.254)	0.09
Stage 3 vs Stage 1	6.522 (2.712, 15.686)	**2.8E-05**
Stage 4 vs Stage 1	8.576 (3.280, 22.425)	**1.2E-05**
Poorly Differentiated vs Moderately to well Differentiated	1.058 (0.642, 1.741)	0.8
Surgery alone vs Surgery + 5 FU	0.951 (0.556, 1.628)	0.3
Model 2 (Predictors: RTK Amp and Genomic Instability*)	**Hazard Ratio (95% CI)**	**P-value**
RTK Amp vs RTK Absent	1.495 (0.970, 2.304)	0.07
High CNA vs Low CNA	1.228 (0.823, 1.833)	0.3

Significant p-values are shown in bold type. *Genomic Instability was inferred based on the number of copy number altered cytobands for each tumor sample (methods).

Table S4b: Univariate analysis analyzing the prognostic impact of individual RTK amplifications
(Outcome: overall survival, relative to patients lacking RTK amplifications)

Model 3 (Predictors: RTK Amp vs RTK Absent)	**Hazard Ratio (95% CI)**	**P-value**
EGFR Amp vs RTK Absent	1.179 (0.589, 2.360)	0.6
ERBB2 Amp vs RTK Absent	2.824 (1.558, 5.119)	**0.0006**
FGFR2 Amp vs RTK Absent	1.098 (0.549, 2.196)	0.8
MET Amp vs RTK Absent	2.744 (1.190, 6.327)	**0.002**

Significant p-values are shown in bold type.

Table S4c: Multivariate analysis comparing individual RTK amplification status with tumor stage and grade
(Outcome: overall survival, relative to patients lacking RTK amplifications)

Model 4 (Predictors: RTK Amp, Stage and Grade)	**Hazard Ratio (95% CI)**	**P-value**
EGFR Amp vs RTK Absent	1.160 (0.570, 2.360)	0.7
ERBB2 Amp vs RTK Absent	3.691 (1.985, 6.863)	**3.7E-05**
FGFR2 Amp vs RTK Absent	1.227 (0.609, 2.471)	0.6
MET Amp vs RTK Absent	1.358 (0.564, 3.269)	0.5
Stage2 vs Stage 1	1.968 (0.816, 4.744)	0.1
Stage3 vs Stage 1	4.969 (2.325, 10.621)	**3.5E-05**
Stage4 vs Stage 1	8.414 (3.887, 18.213)	**6.5E-08**
Poorly Differentiated vs Moderately to well Differentiated	0.996 (0.665, 1.491)	1.0

Significant p-values are shown in bold type.

Table S5a: Univariate analysis of prognostic associations for individual RTK/*KRAS* amplifications
(Outcome: overall survival, relative to patients lacking RTK or *KRAS* amplifications)

Model 1 (Predictors: RTK/*KRAS* Amp vs RTK/KRAS Absent)	**Hazard Ratio (95% CI)**	**P-value**
EGFR Amp vs RTK/*KRAS* Absent	1.306 (0.647, 2.638)	0.5
ERBB2 Amp vs RTK/*KRAS* Absent	3.141 (1.714, 5.756)	**0.0002**
FGFR2 Amp vs RTK/*KRAS* Absent	1.217 (0.603, 2.453)	0.6
MET Amp vs RTK/*KRAS* Absent	2.993 (1.291, 6.940)	**0.01**
KRAS Amp vs RTK/*KRAS* Absent	2.116 (1.155, 3.879)	**0.02**

Significant p-values are highlighted in bold type.

Table S5b: Multivariate analysis comparing *KRAS* and RTK Amplifications with tumor stage and grade

Model 2 (Predictors: RTK/*KRAS* Amp, Stage and Grade)	**Hazard Ratio (95% CI)**	**P-value**
EGFR Amp vs RTK/*KRAS* Absent	1.231 (0.600, 2.528)	0.6
ERBB2 Amp vs RTK/*KRAS* Absent	3.909 (2.082, 7.340)	**2.2E-05**
FGFR2 Amp vs RTK/*KRAS* Absent	1.296 (0.639, 2.631)	0.5
MET Amp vs RTK/*KRAS* Absent	1.440 (0.594, 3.493)	0.4
KRAS Amp vs RTK/*KRAS* Absent	1.455 (0.790, 2.682)	0.2
Stage2 vs Stage 1	1.935 (0.802, 4.670)	0.1
Stage3 vs Stage 1	4.786 (2.230, 10.269)	**5.8E-05**
Stage4 vs Stage 1	8.053 (3.702, 17.515)	**1.4E-07**
Poorly Differentiated vs Moderately to well Differentiated	1.012 (0.675, 1.517)	1.0

Significant p-values are highlighted in bold type.

Table S6. Clinical Characteristics of GC Patient Cohorts Used in Gene Expression Analysis

	SG U133A (51)	SG U133B (248)	AU(70)	UK(29)
Age				
range	38-86	23-92	32-85	53-84
mean,S.D	64.0, 11.2	65.4, 12.5	65.5, 12.5	71.7, 9.11
Gender				
Male	33	161	48	16
Female	18	87	22	13
Lauren classification				
Intestinal	27	138	34	20
Diffuse	11	86	30	6
Mixed	13	24	6	3
Grade				
Moderate to well differentiated	20	96	24	13
Poorly differentiated	30	149	46	15
Unknown	1	3	0	1
Stage				
1	10	40	13	6
2	11	43	16	4
3	15	88	33	15
4	12	76	8	4
Unknown	3	1	0	0

Table S7: Multivariate analysis comparing high *FGFR2* gene expression (>2-fold mean level in normal gastric tissues) with tumor stage and grade

(Outcome: overall survival, relative to patients with low *FGFR2* expression (< 2-fold mean level in normal gastric tissues)

Model 1 (Predictors: *FGFR2* Expression, Stage and Grade)	**Hazard Ratio (95% CI)**	**P-value**
FGFR2 High Expression vs *FGFR2* Low Expression	1.321 (0.966, 1.807)	0.08
Stage 2 vs Stage 1	1.643 (0.924, 2.922)	0.09
Stage 3 vs Stage 1	4.593 (2.807, 7.514)	**1.3e-09**
Stage 4 vs Stage 1	8.440 (5.009, 14.221)	**1.1e-15**
Poorly Differentiated vs Moderately to well Differentiated	0.942 (0.718, 1.235)	0.7

Significant p-values are highlighted in bold type.

Supplementary Figures

Figure S1: Copy Number PRofiles of matched gastric tumor and non-malignant samples

Three representative paired primary GC tumor/normal samples are shown (IDs 2000068, 57689477 and 980021). The x-axis represents chromosomes 1 to 22 and chromosomes X and Y, y-axis represents the extent of copy number amplifications/deletions. The proportion of CNAs for each sample are indicated respectively as a percentage of the whole genome.

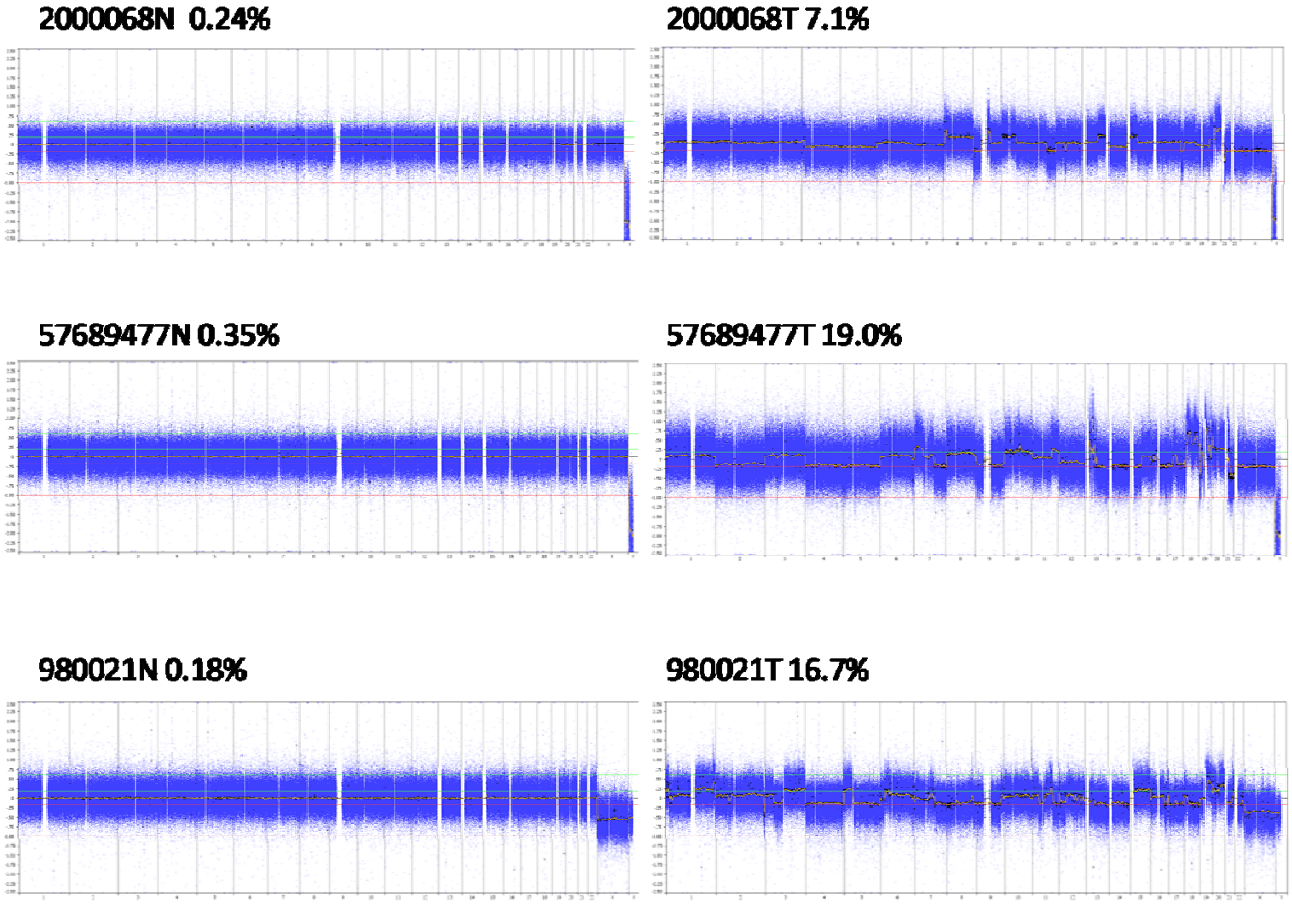


Figure S2: *ERBB2* Copy Number and Protein Expression in GC.

Two primary GCs are shown (IDs 970010 (A,B) and 2000472 (C,D)). (A) Tumor 970010 is predicted to exhibit *ERBB2* copy number amplification. The top graph represents a segment of Chromosome 17 where ERBB2 resides. The *ERBB2* region is marked by yellow boundaries. The y-axis represents the extent of copy number amplification. The bottom graph is a close up of the region, where the ERBB2 gene is marked by a red box. (B) Immunohistochemical (IHC) analysis of ERBB2 reveals high ERBB2 protein expression (IHC 3+) in 970010. (C) Tumor 2000472 is predicted to show normal/neutral ERBB2 copy number levels. Boundaries of the yellow and red boxes are the same as in (A). (D) IHC analysis of ERBB2 reveals absence of ERBB2 protein expression (IHC 0) in 2000472.

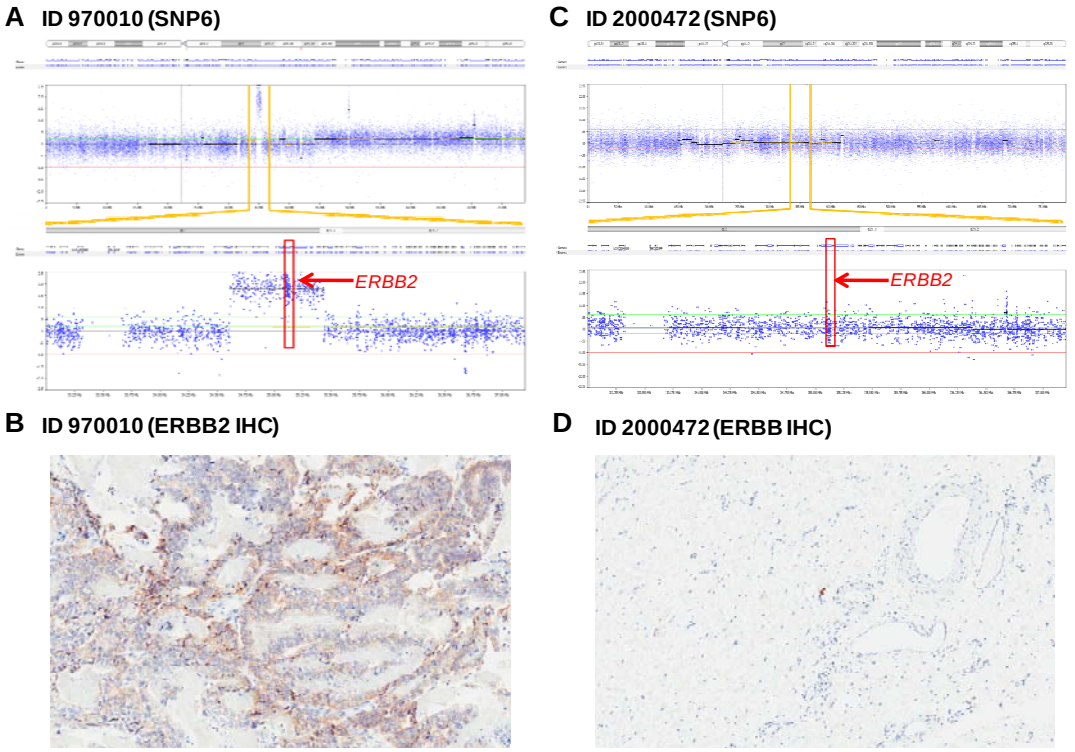


Figure S3: *CSMD1* Expression in GC

Full sections of GCs (n=42) were subjected to CSMD1 immunohistochemistry. (A) CSMD1 expression in normal gastric epithelium (black triangle) and loss of expression in intestinal metaplasia (blue triangle). (B) Loss of CSMD1 expression in a diffuse-type GC. Staining in adjacent normal gastric epithelial (black triangle) cells and within endothelial cells within the tumor serves as a positive internal control. (C) Strong membranous CSMD1 staining in an intestinal-type GC. Approximately 40% of GCs show absent or reduced CSMD1 expression relative to normal gastric epithelium.

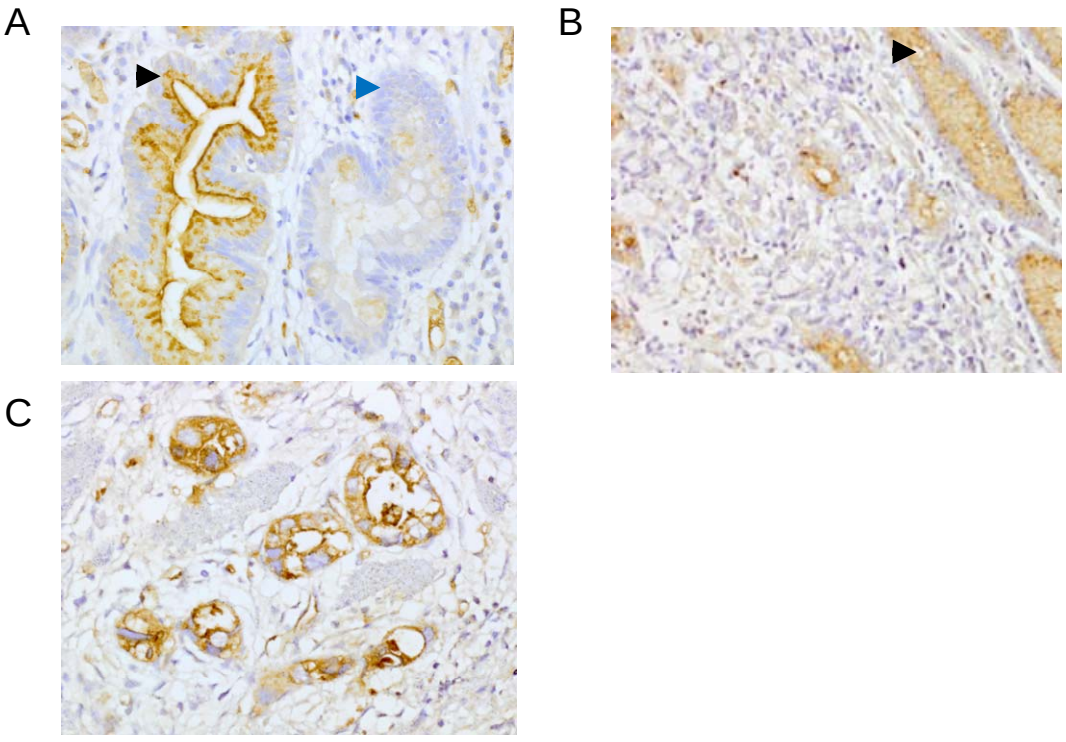


Figure S4. Hierarchical clustering of GCs using genes exhibiting recurrent focal amplifications

In the heatmap, each row represents a different focally amplified amplified gene from the highest recurrent regions (Table 1 in Main Text). Each column represents an individual tumor exhibiting amplifications of these genes (total 113 tumors). The red color gradient (top right) highlights the degree of copy number amplification. Hierarchical clustering was performed both row and column-wise. The highlighted region identified *ERBB2* and *CCNE1*, which exhibit a significant co-amplification pattern as identified by DRP.

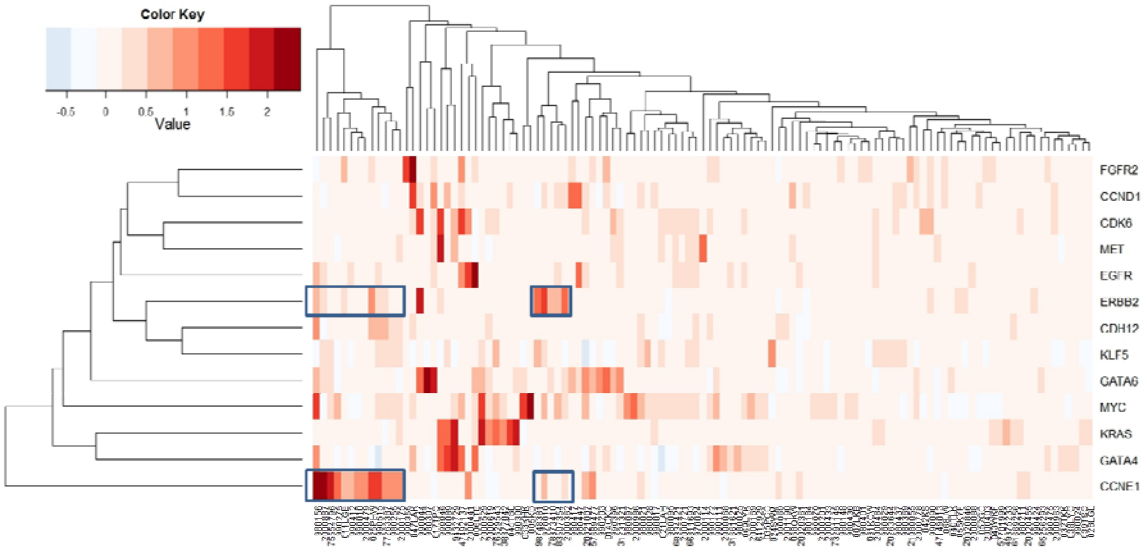


Figure S5: Network Diagram Showing Relationship of RTK Signaling to RAS

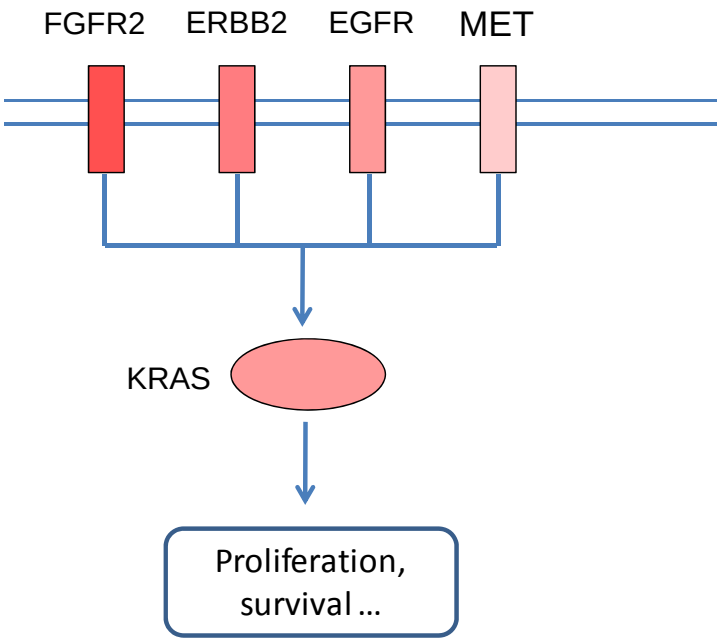


Figure S6: Kaplan-Meier Survival Analysis based on *KRAS* Copy Number Status
A) KM survival graph comparing outcomes of patient with tumors exhibiting KRAS amplification against patients with no/low KRAS CNA irrespective of RTK amplification status. The 17 KRAS-amplified patients correspond to the same patients identified in the Figure 3A heat-map presented in the Main Text.
B) KM survival graph comparing outcomes of patients with tumors exhibiting high KRAS copy number, defined as the top 25% of patients exhibiting a high SNP6 logRatio (high KRAS LR) vs the remaining 75% of patients (low KRAS LR).

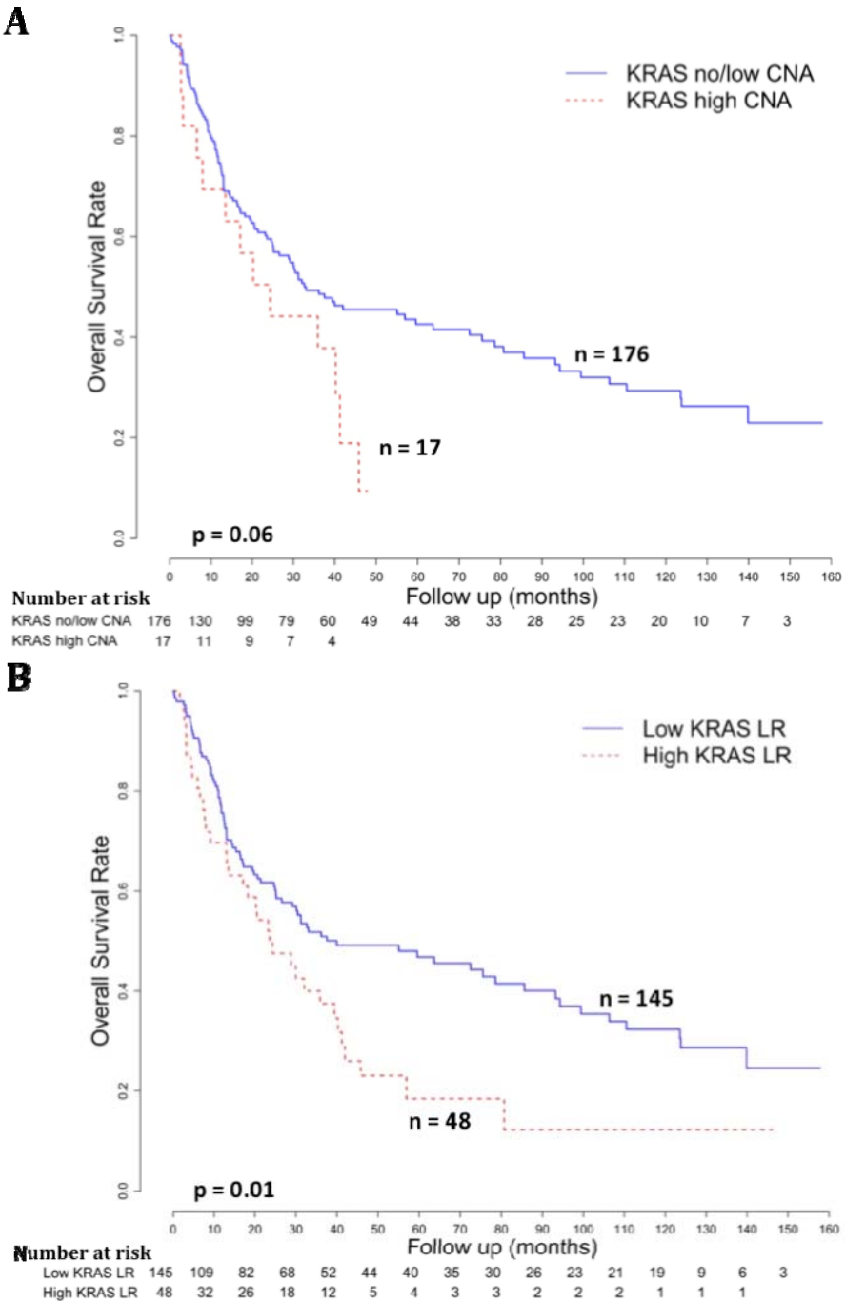


Figure S7: Phenotypic Effects of *KRAS* siRNA Knockdown in *KRAS*-amplified, Mutated and Wild-type GC Lines

KRAS siRNAs or Control Scrambled siRNAs were applied to four GC cell lines – YCC1 and MKN1 (KRAS-amplified), AGS (KRAS-mutated; G12D), and TMK1 (KRAS non-amplified and wild-type). For each cell line, KRAS knockdown was confirmed at the protein level (Western blots – not treated (--), scrambled siRNA (Ctl), *KRAS* siRNA (KRAS)). Cell proliferation was measured 48-96 h after knockdown, comparing *KRAS* siRNA-treated cells to control siRNA treated cells (Numbers above bars are p-values comparing KRAS siRNA vs control siRNA treated cells). Significant reductions in cell proliferation are observed in KRAS-amplified and KRAS-mutated lines (P<0.05), but no significant effects are seen in wild-type TMK1 cells. Similar effects were observed with two non-overlapping *KRAS* siRNAs. All experiments were repeated a minimum of three independent times.

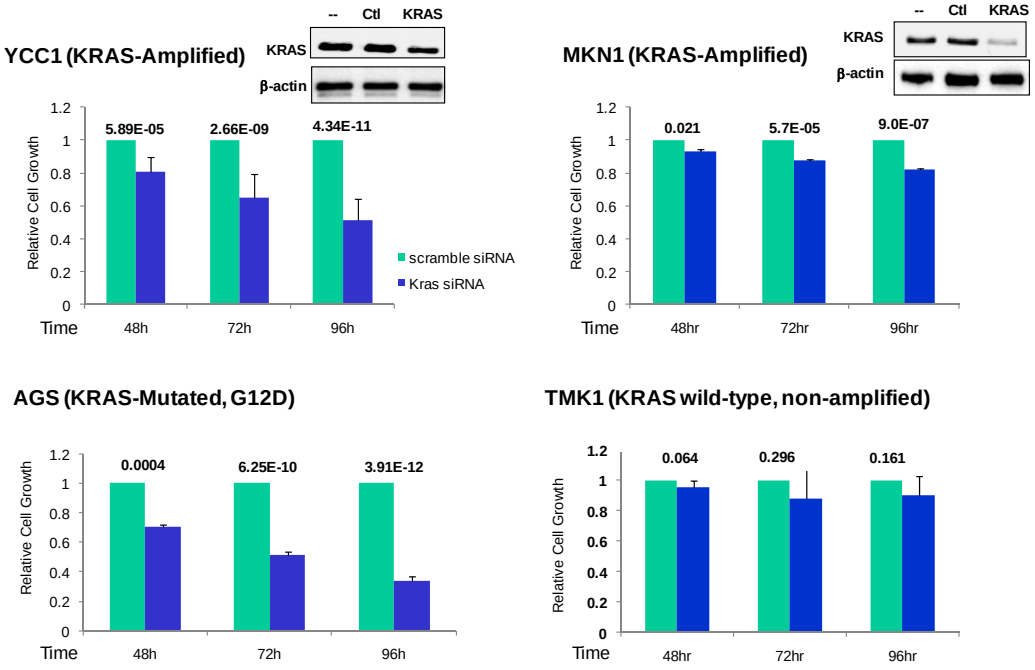


Figure S8: qPCR Analysis of *FGFR2* Amplification in GC

Quantitative PCR of genomic DNA from 63 GC primary tumors, performed using *FGFR2* primers flanking the GISTIC identified amplification peak in intron 2. A) The X-axis shows samples classified into three categories - normal (black), tumors without *FGFR2* amplification (grey), and tumors with *FGFR2* amplification (red, including samples with high copy number level (Figure 3A) and intron 2 copy number). The Y-axis indicates the qPCR DNA level. The horizontal broken black line indicates the cutoff for qPCR amplification. A Fisher exact test shows that samples with high *FGFR2* qPCR values are associated with *FGFR2* amplification (p = 0.0006). Samples were internally normalized against a LINE1 control. B) An X-Y scatter plot of *FGFR2* qPCR values and *FGFR2* copy number based on SNP arrays. x-axis indicates qPCR value and y-axis represents the copy number logRatio. Red, orange and grey colored samples represent high CNA (Figure 3A), focal high CNA (intron 2) and no/low CNA samples respectively. The Spearman correlation is 0.84, showing a positive correlation between *FGFR2* copy number and qPCR values (p < 2.2e-16)

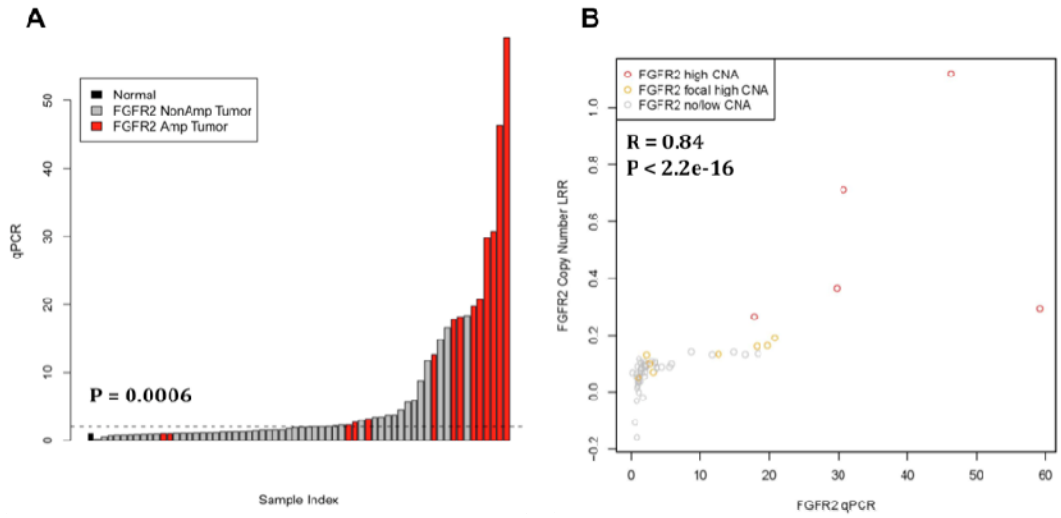


Figure S9. Scatter plot of gene expression and copy number for FGFR2

The figure shows an XY scatter plot of *FGFR2* gene expression and *FGFR2* copy number. x-axis - log2 transformed mRNA expression values; y-axis - copy number logRatio. Red, grey and blue colored samples represent high CNA, low/no CNA, and normal samples respectively. Spearman correlation value is indicated as R = 0.38, with p value = 3.3e-7.

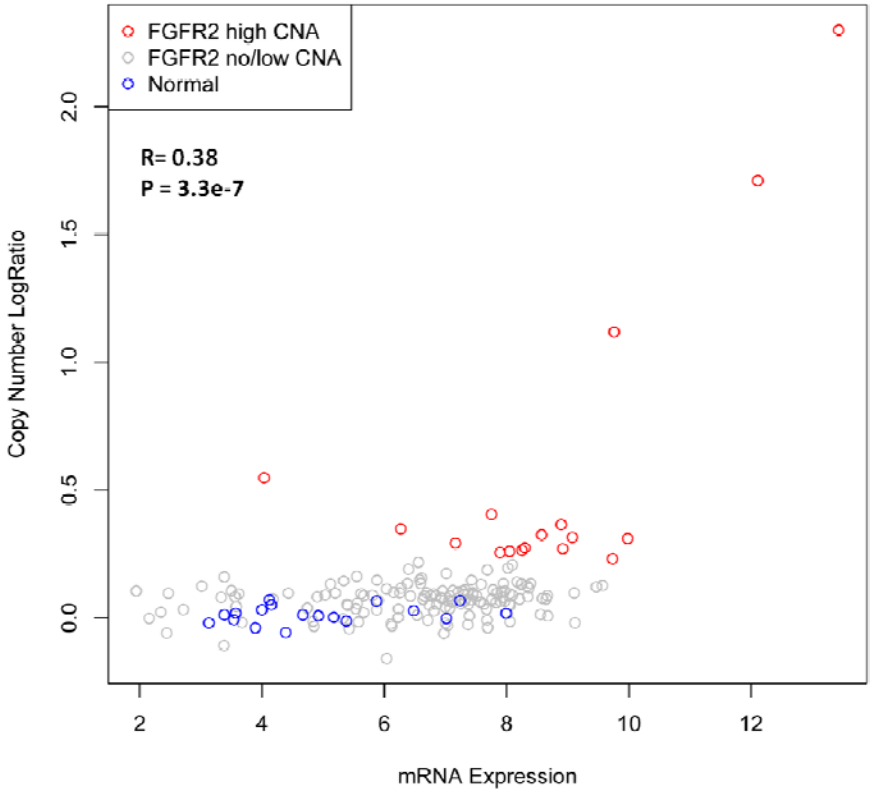


Figure S10: Relationship between Copy Number and Gene Expression for ATE1 and BRWD2, Genes Adjacent to FGFR2

Primary GCs exhibiting genomic amplification of the *FGFR2* locus were also assessed for relationships between copy number status and gene expression in A,C) *ATE1* (upstream of *FGFR2*) and B,D) *BRWD2* (downstream of *FGFR2*). For each gene, mRNA expression was compared across three categories, each represented by a box-plot - non-malignant gastric tissues (normal) (n =100 for A,B, n =18 samples with available copy number information for C,D), tumors exhibiting no/low *FGFR2* gene locus CNA (n = 139), and tumors exhibiting high *FGFR2* gene locus CNA (n = 17). *ATE1* and *BRWD2* expression was inferred from Affymetrix microarrays (*ATE1* 234584_s; *BRWD2* probe 218090_s_at).

A) *ATE1* expression levels in amplified tumors are observed to be significantly higher than normal samples (P=0.004, Wilcoxon test, underlined). However this significance level is weaker than that observed for *FGFR2* (p=1.7e-7, see Main Text).

B) *BRWD2* expression levels in amplified tumors are not significantly higher than normal samples (P=0.3, Wilcoxon test, underlined).

C) XY scatter plot of *ATE1* expression with copy number information. Spearman correlation R is 0.16 with p value = 0.04.

D) XY scatter plot of *BRWD2* expression with copy number information. Spearman correlation R is 0.16 with p value = 0.04.

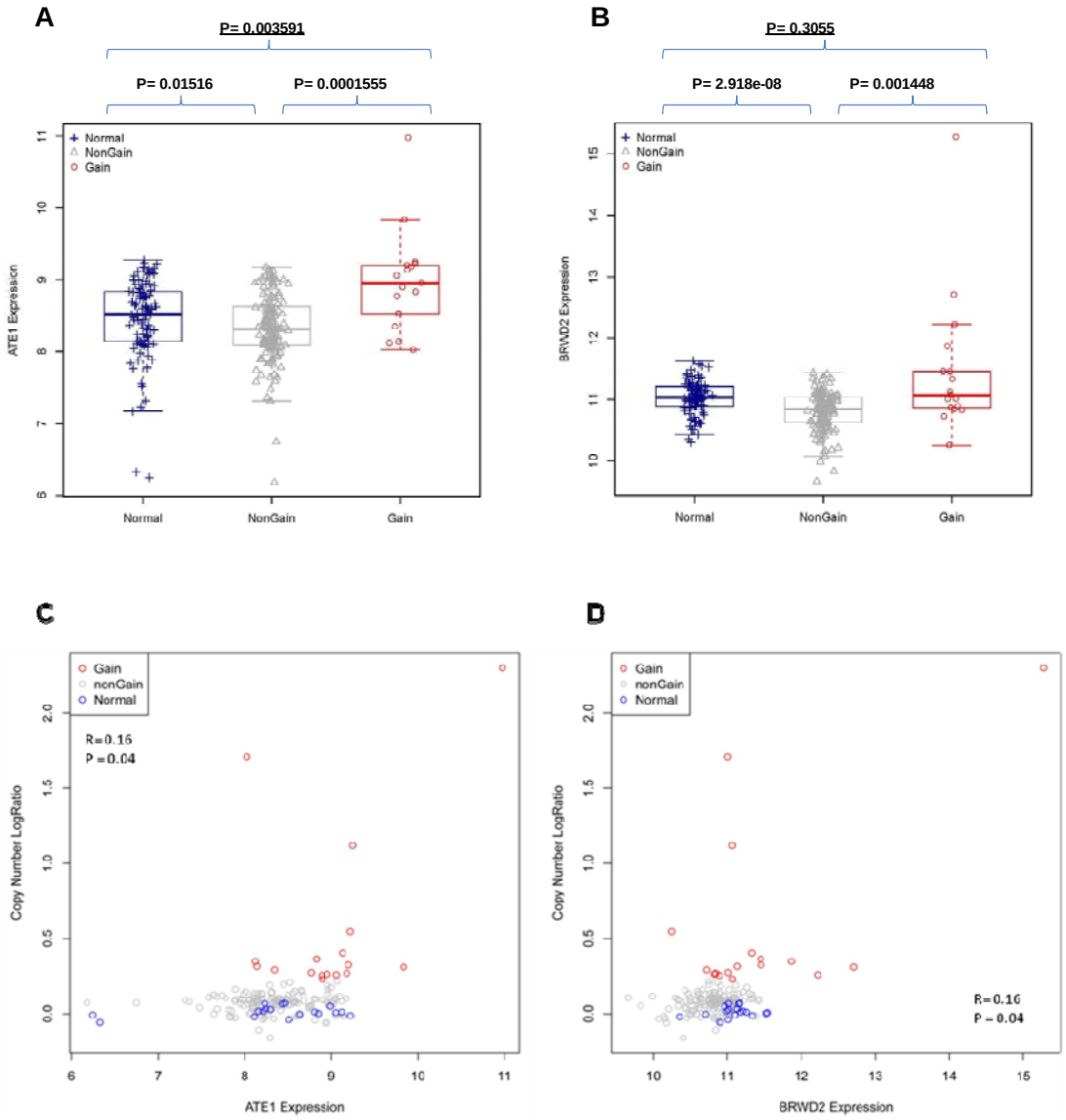

A
P= 0.003591
P= 0.01516
P= 0.0001555
Normal
NonGain
Gain
ATE1 Expression
B
P= 0.3055
P= 2.918e-08
P= 0.001448
BRWD2 Expression
C
Gain
nonGain
Normal
R=0.16
P=0.04
Copy Number LogRatio
ATE1 Expression
D
BRWD2 Expression

Figure S11: *FGFR2* Overexpression in GCs Relative to Normal Gastric Samples

The graph depicts 236 normal gastric tissues and 399 primary gastric tumors, arranged along the x-axis in ascending order of their *FGFR2* expression level. FGFR2 gene expression levels were inferred using Affymetrix microarrays (FGFR2 probe 211401_s_at). At the cut-off threshold level of >2x the average level in normal tissues (dotted line), approximately 18% of gastric tumors exhibit high *FGFR2* levels (marked in red).

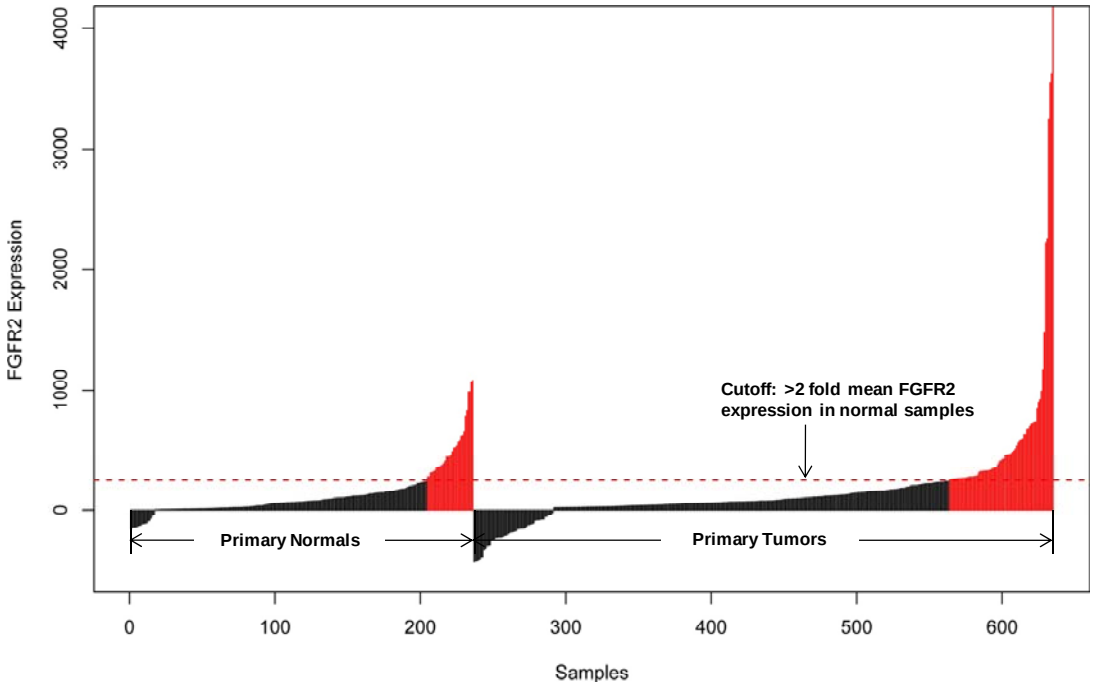
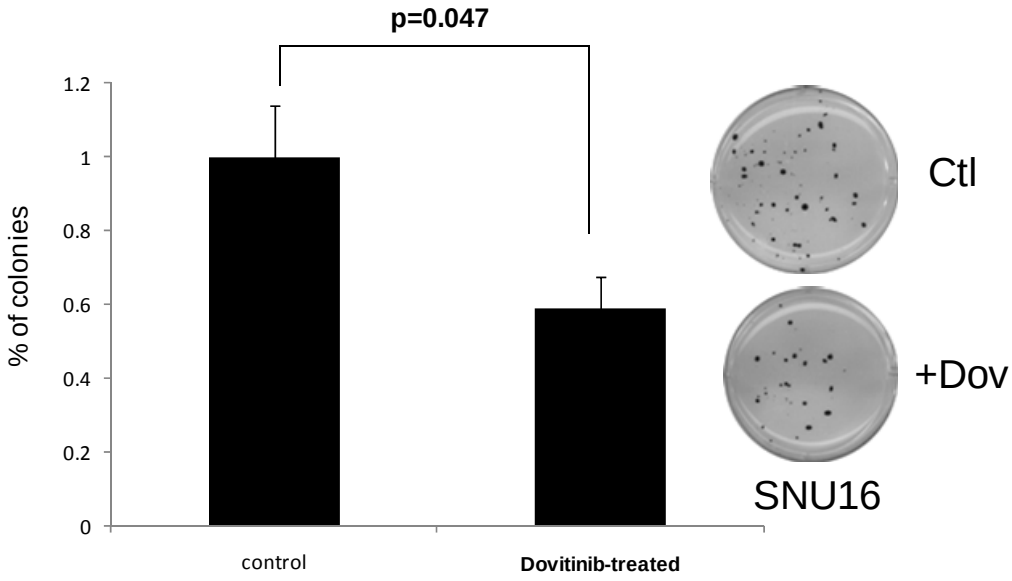


Figure S12: Inhibition of Soft Agar Colony Growth by Dovitinib (SNU-16)

FGFR2-amplified SNU16 cells were treated with dovitinib at the GI50 concentration (0.17uM) for 48 hrs, and soft-agar colony formation monitored over the subsequent 3-4 weeks. Representative plates are shown (Ctl : mock treated, + Dov : Dovitinib treated). Bar graphs depict results from a minimum of three independent experiments.



References

1. Olshen, A.B., et al., *Circular binary segmentation for the analysis of array-based DNA copy number data.* Biostatistics, 2004. **5**(4): p. 557-72.
2. Beroukhim, R., et al., *Assessing the significance of chromosomal aberrations in cancer: methodology and application to glioma.* Proc Natl Acad Sci U S A, 2007. **104**(50): p. 20007-12.
3. Ding, L., et al., *Somatic mutations affect key pathways in lung adenocarcinoma.* Nature, 2008. **455**(7216): p. 1069-75.
4. Hofmann, M., et al., *Assessment of a HER2 scoring system for gastric cancer: results from a validation study.* Histopathology, 2008. **52**(7): p. 797-805.
5. Kamal, M., et al., *Loss of CSMD1 expression is associated with high tumour grade and poor survival in invasive ductal breast carcinoma.* Breast Cancer Res Treat, 2010. **121**(3): p. 555-63.
6. Ooi, C.H., et al., *Oncogenic pathway combinations predict clinical prognosis in gastric cancer.* PLoS Genet, 2009. **5**(10): p. e1000676.
7. Johnson, W.E., C. Li, and A. Rabinovic, *Adjusting batch effects in microarray expression data using empirical Bayes methods.* Biostatistics, 2007. **8**(1): p. 118-27.