










OPEN ACCESS

Original research

# Spatial profiling of gastric cancer patient-matched primary and locoregional metastases reveals principles of tumour dissemination

Raghav Sundar <sup>1,2,3,4</sup> Drolaiz HW Liu,<sup>5</sup> Gordon GA Hutchins <sup>6</sup> Hayley L Slaney,<sup>6</sup> Arnaldo NS Silva <sup>6,7</sup> Jan Oosting <sup>8</sup> Jeremy D Hayden,<sup>9</sup> Lindsay C Hewitt,<sup>5</sup> Cedric CY Ng <sup>10</sup> Amrita Mangalvedhekar,<sup>2</sup> Sarah B Ng,<sup>11</sup> Iain BH Tan,<sup>11,12</sup> Patrick Tan <sup>2,3,11,13,14</sup> Heike I Grabsch <sup>5,6</sup>

► Additional material is published online only. To view, please visit the journal online (<http://dx.doi.org/10.1136/gutjnl-2020-320805>).

For numbered affiliations see end of article.

## Correspondence to

Prof. Heike I Grabsch, GROW - School for Oncology and Developmental Biology; Department of Pathology, Maastricht University Medical Centre+, Maastricht 6229 HX, Limburg, The Netherlands; [h.grabsch@maastrichtuniversity.nl](mailto:h.grabsch@maastrichtuniversity.nl) and Prof. Patrick Tan, Cancer and Stem Cell Biology Program, Duke-NUS Medical School, Singapore, Singapore; [gmstanp@duke-nus.edu.sg](mailto:gmstanp@duke-nus.edu.sg)

RS and DHL contributed equally.

PT and HIG are joint senior authors.

Received 1 February 2020  
Revised 2 November 2020  
Accepted 5 November 2020  
Published Online First  
23 November 2020



Watch Video  
[gut.bmj.com](http://gut.bmj.com)



© Author(s) (or their employer(s)) 2021. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

**To cite:** Sundar R, Liu DHW, Hutchins GGA, et al. *Gut* 2021;**70**:1823–1832.

## ABSTRACT

**Objective** Endoscopic mucosal biopsies of primary gastric cancers (GCs) are used to guide diagnosis, biomarker testing and treatment. Spatial intratumoural heterogeneity (ITH) may influence biopsy-derived information. We aimed to study ITH of primary GCs and matched lymph node metastasis (LN<sub>met</sub>).

**Design** GC resection samples were annotated to identify primary tumour superficial (PT<sub>sup</sub>), primary tumour deep (PT<sub>deep</sub>) and LN<sub>met</sub> subregions. For each subregion, we determined (1) transcriptomic profiles (NanoString 'PanCancer Progression Panel', 770 genes); (2) next-generation sequencing (NGS, 225 gastrointestinal cancer-related genes); (3) DNA copy number profiles by multiplex ligation-dependent probe amplification (MLPA, 16 genes); and (4) histomorphological phenotypes.

**Results** NanoString profiling of 64 GCs revealed no differences between PT<sub>sup1</sub> and PT<sub>sup2</sub> while 43% of genes were differentially expressed between PT<sub>sup</sub> versus PT<sub>deep</sub> and 38% in PT<sub>sup</sub> versus LN<sub>met</sub>. Only 16% of genes were differently expressed between PT<sub>deep</sub> and LN<sub>met</sub>. Several genes with therapeutic potential (eg *IGF1*, *PIK3CD* and *TGFB1*) were overexpressed in LN<sub>met</sub> and PT<sub>deep</sub> compared with PT<sub>sup</sub>. NGS data revealed orthogonal support of NanoString results with 40% mutations present in PT<sub>deep</sub> and/or LN<sub>met</sub> but not in PT<sub>sup</sub>. Conversely, only 6% of mutations were present in PT<sub>sup</sub> and were absent in PT<sub>deep</sub> and LN<sub>met</sub>. MLPA demonstrated significant ITH between subregions and progressive genomic changes from PT<sub>sup</sub> to PT<sub>deep</sub>/LN<sub>met</sub>.

**Conclusion** In GC, regional lymph node metastases are likely to originate from deeper subregions of the primary tumour. Future clinical trials of novel targeted therapies must consider assessment of deeper subregions of the primary tumour and/or metastases as several therapeutically relevant genes are only mutated, overexpressed or amplified in these regions.

## INTRODUCTION

Gastric cancer (GC) is a leading cause of mortality and morbidity worldwide.<sup>1</sup> Targeting of specific cancer-associated driver genes has improved outcomes in patients with lung, breast or colorectal cancer.<sup>2–6</sup> However, similar approaches have only proved modestly successful in patients with

## Significance of this study

### What is already known on this subject?

- Tumour heterogeneity of gastric cancer (GC) has been proposed as a major hurdle to developing novel targeted therapeutics.
- Genomic differences between primary and metastatic tumour sites have been reported as a potential barrier for precision oncology.
- Several previous studies have confirmed intertumoural heterogeneity in GC.

### What are the new findings?

- We studied intratumoural spatial heterogeneity in GC by annotating superficial and deep subregions of the primary tumour as well as tumour in matched lymph node metastasis (LN<sub>met</sub>) from GC resection specimens.
- By performing phenotypical, genomic and transcriptomic profiling, we found superficial subregion profiles to be significantly different compared with matched deep subregions and LN<sub>met</sub>. At the same time, profiles of matched deep subregions and metastatic lymph nodes exhibited greater similarity.
- Several therapeutically targetable genes were amplified, mutated and/or overexpressed in the deep subregions and/or LN<sub>met</sub> but not in the superficial subregions.

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

- Multiple superficial endoscopic biopsies are unlikely to reflect the whole breadth of intratumoural heterogeneity in GC.
- Future clinical trials of novel targeted therapies must consider assessment of the deeper subregions of the primary tumour and/or metastases.
- New understanding of the biology of metastatic dissemination of GC will allow advances in drug discovery and treatment.

GC. Tumour heterogeneity has been proposed as a major hurdle to developing novel therapeutics for patients with GC.<sup>7</sup> While *intertumoural*

heterogeneity (ie, between patients) has been described in GC in some detail,<sup>8,9</sup> *intratumoural* heterogeneity (ITH) (ie, within the same patient) is less well studied. Genomic differences between primary and metastatic GC sites have been recently reported as a potential barrier for precision oncology.<sup>10</sup> Clinically, cancer diagnosis and treatment choices are usually made by evaluation of endoscopic biopsies taken from the luminal aspects of the primary tumour ('superficial mucosa'). In contrast, material from deeper tumour regions or metastatic sites such as regional lymph node metastasis (LN<sub>met</sub>) or distant metastasis is often not easily accessible and therefore usually not investigated at the time of diagnosis.

In this study, we hypothesised that luminal and deep subregions of primary GCs exhibit systematic differences in cancer-related genes and pathways, and that these differences might also relate to the molecular profiles of patient-matched lymph node metastases. To investigate this, we adopted a multiassay strategy to profile various regions from primary GCs and matched lymph node metastases at the genomic, transcriptomic and histomorphological (phenotypic) levels.

## METHODS

### Patient samples

This cohort included patients diagnosed with gastric adenocarcinoma who underwent potentially curative surgery at the Department of Surgery, Leeds Teaching Hospital NHS Trust (Leeds, UK) between 1985 and 2003. Patients were included in the study if the adenocarcinoma invaded at least the muscularis propria, and lymph node metastases with sufficient tumour for RNA and DNA extraction were present. None of the patients underwent perioperative chemotherapy. Clinical and pathological data were retrieved from histopathology reports, electronic patient hospital records and the Northern and Yorkshire Cancer Registry.

Representative Haematoxylin/Eosin (H&E) stained sections from the resection specimen were scanned using an Aperio slide scanner (Aperio XT; Aperio Technologies, Vista, California, USA), reviewed, and multiple subregions were manually delineated in the primary tumour: two adjacent luminal subregions (primary tumour superficial (PT<sub>sup1</sub>) and PT<sub>sup2</sub>) extending up to 2 mm deep into the mucosa/superficial submucosa approximately equivalent to the depth that is reached by endoscopic biopsy, one deep tumour subregion (primary tumour deep (PT<sub>deep</sub>)) extending from 2 mm depth to the deepest point of tumour infiltration in the wall. The median size of the individual superficial region was 35 mm<sup>2</sup>, which is equivalent to the amount of tissue that can be obtained from approximately eight endoscopic biopsies. In addition, tumour was outlined in the lymph node with the largest contiguous metastatic tumour area (LN<sub>met</sub>) (figure 1).

### Histomorphological phenotype classification

The histological phenotype of all cases was classified by a pathologist without knowledge of associated clinicopathological information or results from RNA/DNA investigations. Every subregion (PT<sub>sup1</sub>, PT<sub>sup2</sub>, PT<sub>deep</sub> and LN<sub>met</sub>) was assessed for the presence of three main histological phenotypes: poorly cohesive, non-poorly cohesive and mucinous phenotypes. For poorly cohesive cancers, the presence of signet-ring cells was recorded according to the recently published consensus as <10%, 10%–90% or ≥90% signet-ring cells.<sup>11</sup> Signet-ring cells were defined as cells with an eccentric nucleus and optically clear cytoplasmic mucin on the H&E-stained slides.

### Quantification of tumour and stroma content per subregion

The tumour and stroma content was quantified in all cases for each primary tumour subregion and LN<sub>met</sub> separately using the same tissue section and the same outlines as used for DNA and RNA extractions. The stereology tool in the MIM software (HeteroGenius, Leeds, UK) was used to take a random systematic sample of 300 points per subregion as described previously.<sup>12</sup> The tissue category (tumour epithelium vs stroma (anything in between epithelial tumour cells) vs non-informative (normal epithelium, artefacts and non-interpretable tissue)) was determined for every point. The tumour:stroma ratio per subregion was calculated by dividing the number of tumour points by the number of stroma points.

### DNA and RNA extraction

DNA and RNA were extracted from the subregions separately. DNA was extracted using the QIAmp DNA Micro Kit (Qiagen, Hilden, Germany) as previously described.<sup>13</sup> DNA concentration was measured by ND-100 Spectrophotometer (Labtech International) and samples were diluted using Tris-EDTA buffer. RNA was extracted using the Qiagen RNeasy Kit (Qiagen) according to the instructions of the manufacturer.

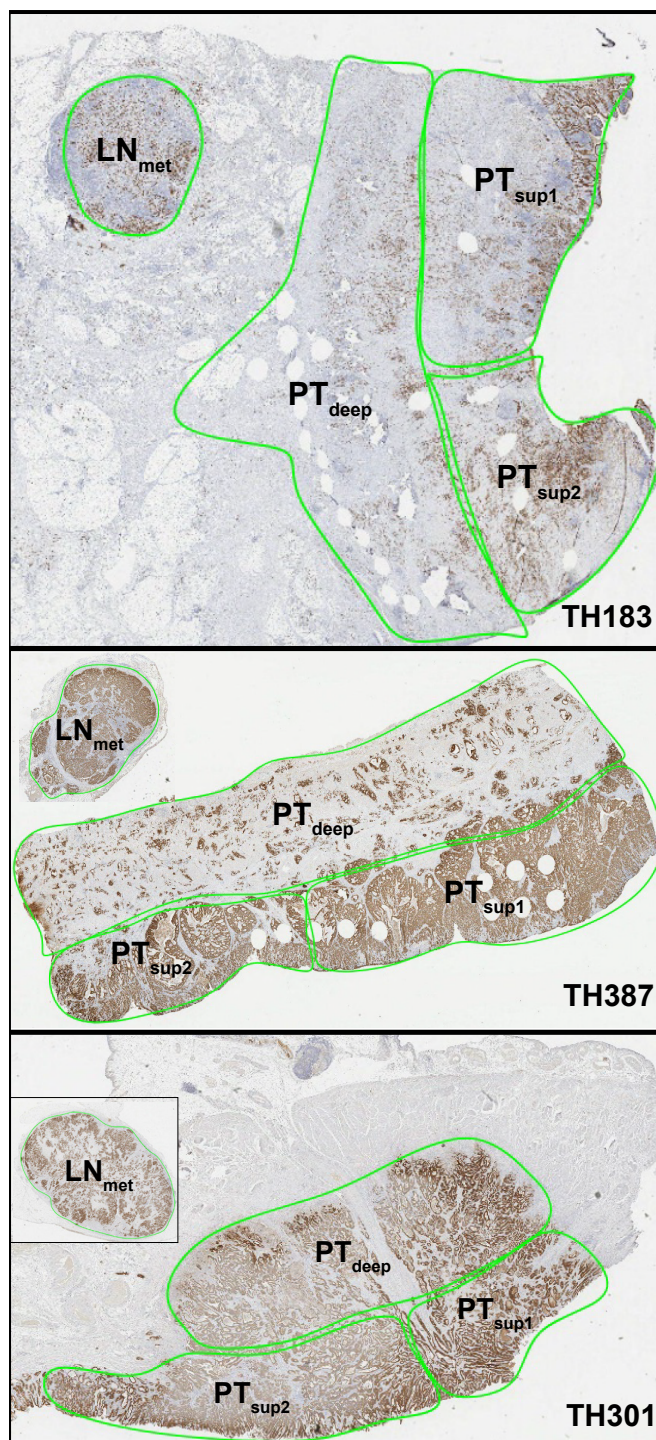
### NanoString analysis

The 'PanCancer Progression Panel' NanoString nCounter Reporter CodeSet (Nanostring Technologies, USA) was used for this study. This panel includes 770 genes covering a range of cancer-related pathways, including angiogenesis, extracellular matrix (ECM) remodelling, epithelial-to-mesenchymal transition and metastasis (online supplemental table S1). Pathway annotations were based on NanoString panel designation, which was curated from publicly available data sources.<sup>14–16</sup> Raw counts were normalised using the geometric mean of the internal positive control probes included in the CodeSet using the vendor-provided nCounter software (nSolver). ITH was quantified by calculating the arithmetic mean of the standard deviation (SD) of gene expression between subregions per gene for each tumour. GCs with mean SD > 50th centile were classified as ITH<sub>high</sub>, all others were classified as ITH<sub>low</sub>.

Quantitative reverse transcriptase-PCR (qRT-PCR) validation of NanoString transcriptomic findings was performed for selected genes (*TGFB*, *CLDN4*, *CDH1*, *FGF18* and *CEACAM6*) using RNA extracted from the different tumour subregions. One-step qRT-PCR was performed using the KAPA SYBR FAST One step universal kit (Sigma, Thermofisher USA) and gene-specific primers (sequences shown in online supplemental table S2). Triplicate runs were performed from each subregion. To calculate fold changes of gene expression between subregions, the delta-delta cycle threshold ( $\Delta\Delta CT$ ) method was used.<sup>17</sup> *GAPDH* was used as endogenous control for normalisation. Comparisons for  $\Delta\Delta CT$  were made using LN<sub>met</sub> as reference, and relative quantification was calculated as  $2^{-\frac{\Delta\Delta CT}{\Delta\Delta CT}}$ . For each tumour, pairwise comparisons between PT<sub>sup</sub>, PT<sub>deep</sub> and LN<sub>met</sub> were performed and the direction of fold change was compared with the NanoString results (eg, the value of PT<sub>sup</sub> greater than PT<sub>deep</sub> in both qRT-PCR and NanoString for a particular GC).

### Next-generation sequencing (NGS) analysis

DNA was sheared to 200 base pairs using a sonicator and cleaned-up with 1.4X Agencourt AMPure XP beads (Beckman Coulter, A63881). Libraries were generated using the KAPA Hyper Prep Kit (Roche, KK8504). Hybridisation capture was done using an IDT xGen Custom Panel of 225 genes (online



**Figure 1** Pathology selection of tumour subregions. Blocks with highest tumour content per area were selected from the resection specimen and subregions were manually delineated in the primary tumour by a pathologist: two adjacent luminal subregions (PT<sub>sup1</sub> and PT<sub>sup2</sub>) extending approximately 2 mm deep into the mucosa/superficial submucosa equivalent to the depth that is reached by endoscopic biopsy, one deep tumour subregion (PT<sub>deep</sub>) extending from 2 mm depth to the deepest point of tumour infiltration in the wall. The median size of the individual superficial region was 35 mm<sup>2</sup>, which is equivalent to the amount of tissue that can be obtained from approximately eight endoscopic biopsies. In addition, tumour was outlined in the lymph node with the largest contiguous metastatic tumour area (LN<sub>met</sub>). Cytokeratin-stained images are provided in this figure for illustrative purposes only to facilitate visualisation of tumour cells (3,3'-diaminobenzidine (DAB) brown). LN<sub>met</sub>, lymph node metastasis; PT<sub>deep</sub>, primary tumour deep; PT<sub>sup</sub>, primary tumour superficial.

supplemental table S3) and reagents as per manufacturer's instruction (xGen Hybridization and Wash Kit, IDT 1080577). Sequencing was performed on an Illumina NextSeq 550 System Mid-Output Kit (2×150 paired).

Sequence reads were processed using the bcbio-nextgen pipeline V.1.1.1a0 with default settings, including adapter trimming, quality filtering and read mapping on the reference genome hg38 with BWA V.0.7.17-r11880 and duplicate read marking with

GATK MarkDuplicates. Quality statistics and on-target analysis of coverage were performed using samtools V.1.9 and bedtools V.2.27.1. Variants were called using VarDict, and somatic variants per patient were annotated using Ensembl VEP V.23. Only variants with moderate or high impact, with maximum population variant frequency of  $>0.00001$  and number of alternate supporting reads of  $>10$  were considered for further analyses.

### Multiplex ligation-dependent probe amplification (MLPA) assay

Details of the MLPA assay, including probemix development, sensitivity against single-nucleotide polymorphisms and fluorescent in situ hybridisation using GC cell lines, have been previously described.<sup>18</sup> The MLPA probemix P458-B1 (MRC-Holland, NL) included the genes *EGFR*, *ERBB2*, *FGFR2*, *MET*, *TOP2A*, *KRAS*, *MYC*, *PTP4A3*, *PIK3CA*, *KLF5*, *CCNE1*, *GATA6*, *FGFR1*, *GATA4*, *CDK6* and *CCND1*.

The MLPA analysis methodology has been previously described.<sup>19, 20</sup> In short, analyses were performed using the MLPA intermethod,<sup>19, 20</sup> whereby samples were normalised per batch using samples processed in each batch. Quality control ensured that samples with low overall intensity, with a large difference in intensity between short probes and long probes, with low intensity of denaturation controls, or high within-gene variation, defined as the average of the SD of log-transformed values were excluded.

Final values were calculated by averaging the peak height of each probe and then averaging the results of replicates. Copy number thresholds were set according to previously published studies<sup>18, 21</sup> as follows: a gene copy number ratio of less than 0.80 was categorised as ‘loss’, between  $>0.80$  and 1.30 as ‘normal’, between  $>1.30$  and 2.0 as ‘low amplification’, between  $>2.0$  and 5.0 as ‘high amplification’ and  $>5.0$  as ‘very high amplification’. Pairwise comparisons were made within each GC between subregions. If the copy number aberrations were the same in the subregions, concordance was established (eg, normal for  $PT_{sup}$  and normal for  $PT_{deep}$  for a particular gene in a GC were classified as concordant), while discordance was defined as a different value for the comparison (eg, ‘loss’ or ‘amplification’ for  $LN_{met}$  and normal for  $PT_{sup}$  for a particular gene in a GC was classified as discordant).

### Statistical analyses

Categorical data such as sex, pT (depth of invasion) and pN (lymph node status) were compared using Fisher’s exact test. Comparison of NanoString RNA expression values between  $LN_{met}$ ,  $PT_{sup}$  and  $PT_{deep}$  was performed using non-parametric paired Wilcoxon signed-rank test with false discovery rate (FDR) corrections for multiple hypothesis testing. An adjusted p value of  $<0.05$  was considered statistically significant. Unsupervised hierarchical clustering of samples was done using principal component analysis and the ggbiplot package. After calculating the principal components of the explained variance when centred and scaled, a mean centre point for each cluster was calculated for generation of ellipses. Kaplan-Meier plots and log-rank statistics were used for survival analysis. All statistical analyses were performed using R V.3.6.1 (R Foundation for Statistical Computing, Vienna, Austria).

## RESULTS

For this study, up to 64 primary GCs and matched regional lymph node metastases from patients with locally advanced resectable GC were included per assay. Each patient had several tumour

subregions analysed: primary tumour from the luminal (superficial) surface, which was subdivided into two parts ( $PT_{sup1}$  and  $PT_{sup2}$ ); primary tumour from deep regions ( $PT_{deep}$ ); and tumour regions in regional lymph nodes ( $LN_{met}$ ). Based on quality of DNA/RNA extracted from each sample, different numbers of samples were available for the various assays (online supplemental table S4). All 64 (100%) GCs had clinical and histological phenotyping data available. The median age was 71 years (range: 39–91 years), and 73% were male. Most GCs were of intestinal type (92%); six (9%) GCs were Epstein-Barr virus positive; and five (8%) GCs were mismatch repair (MMR) protein deficient (online supplemental table S5). The estimated median overall survival was 1.02 years (95% CI 0.80 to 1.58).

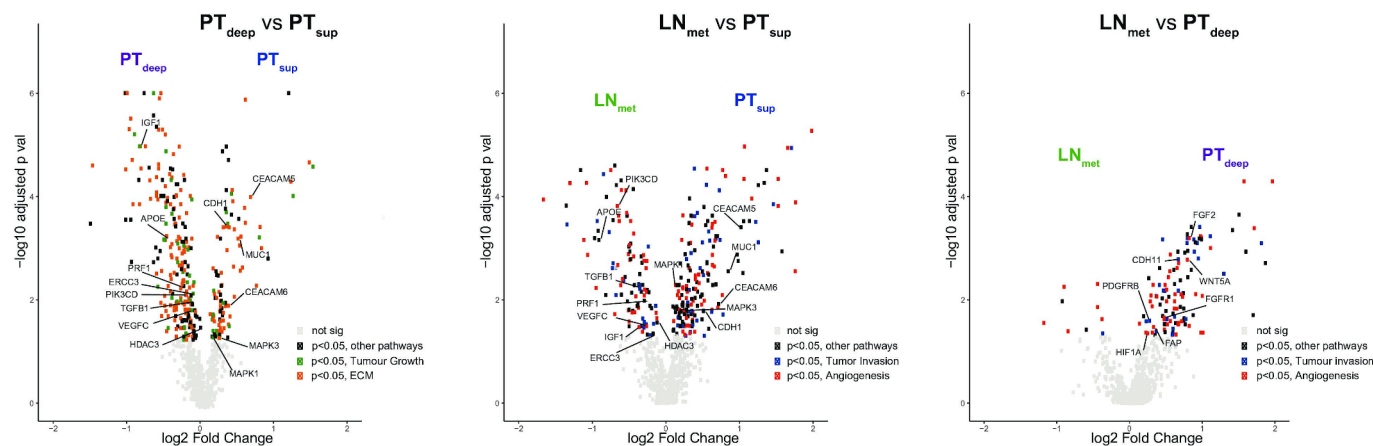
### Differential expression of genes between spatial subregions in primary tumor and matched metastatic lymph node

After quality control, NanoString data were available from  $PT_{sup1}$  (63 (98%)),  $PT_{sup2}$  (54 (84%)),  $PT_{deep}$  (57 (89%)) and  $LN_{met}$  (51 (80%)) GCs, respectively (online supplemental tables S4 and S6). Transcriptomic profiles consisting of 770 genes were compared for differential expression between subregions. Between samples that had matched  $LN_{met}$  and samples that did not, we did not observe any significant clinicopathological differences at the transcriptomic level (online supplemental figure S1). When  $PT_{sup1}$  and  $PT_{sup2}$  were compared within the same patient, none of the 770 genes was significantly differently expressed after false discovery rate (FDR) correction for multiplicity (online supplemental figure S2). Furthermore, there was a strong correlation of the transcriptomic profiles between the two adjacent luminal GC subregions (Spearman’s  $R=0.95$ ,  $p<0.0001$ ) (online supplemental figure S3). For further analyses of NanoString data,  $PT_{sup1}$  values were designated  $PT_{sup}$  and compared with  $PT_{deep}$  subregion and  $LN_{met}$  samples.

Differential gene expression analysis between  $PT_{deep}$  and  $PT_{sup}$  subregions identified 331 (43%) genes which were significantly different between the two subregions (figure 2A). Genes mapping to the ‘ECM’ and ‘tumour growth’ pathways were the genes most differentially expressed between  $PT_{sup}$  and  $PT_{deep}$  subregions (online supplemental table S7).

Differential gene expression analysis between  $PT_{sup}$  and  $LN_{met}$  subregions identified 291 (38%) genes as significantly different between the two subregions (figure 2B and online supplemental table S8).  $PT_{sup}$  had 186 (24%) genes overexpressed compared with  $LN_{met}$ , while  $LN_{met}$  had 105 (14%) genes overexpressed compared with  $PT_{sup}$  (table 1). Genes mapping to the ‘angiogenesis’ and ‘tumour invasion’ pathways were the most commonly differentially expressed genes between  $PT_{sup}$  and  $LN_{met}$ . Approximately 10% of genes were found to overlap between the comparison of  $PT_{sup}$  versus  $PT_{deep}$  and  $PT_{sup}$  versus  $LN_{met}$  (81 genes were overexpressed in  $PT_{sup}$  compared with  $PT_{deep}$  and  $LN_{met}$ , while 69 genes were overexpressed in  $PT_{deep}$  and  $LN_{met}$  compared with  $PT_{sup}$ ). Several of these genes were potential therapeutic targets, such as *APOE*, *ERCC3*, *HDAC3*, *IGF1*, *PIK3CD*, *PRF1*, *TGFB1* and *VEGFC* (higher in  $PT_{deep}$  and  $LN_{met}$ ) and *CDH1*, *CEACAM5*, *CEACAM6*, *MAPK1*, *MAPK3*, *MUC1* and *WNT5A* (higher in  $PT_{sup}$ ). Quantitative RT-PCR was performed on selected genes and validated the NanoString gene expression findings (online supplemental figure S4).

Differential gene expression analysis between  $PT_{deep}$  and  $LN_{met}$  subregions identified 127 (16%) genes as significantly different between the two subregions (figure 2C and online supplemental table S9). Of these, 117 (15%) genes were overexpressed in  $PT_{deep}$  compared with 10 (1%) overexpressed genes in  $LN_{met}$



**Figure 2** Volcano plots comparing differential gene expression between subregions. (A) Volcano plot of 770 genes from the NanoString PanCancer progression panel comparing differential gene expression between  $PT_{deep}$  and  $PT_{sup}$ . The x-axis is the  $\log_2$  fold change of gene expression between  $PT_{sup}$  and  $PT_{deep}$ . The y-axis is the  $-\log_{10}$  adjusted p value results (false discovery rate correction). Genes of interest have been annotated within the plot. Grey dots are genes which are similarly expressed in  $PT_{deep}$  and  $PT_{sup}$ . Green dots are genes which are statistically significant and belong to the 'tumour growth' pathway. Orange dots are genes which are statistically significant and belong to the 'ECM remodelling' pathway. Black dots are genes which are statistically significant and belong to other pathways. In general, genes which are at the extreme ends of the volcano are highly differentially expressed in that particular subregion compared with the other and are of greatest clinical interest. (B) Volcano plot of genes comparing  $LN_{met}$  and  $PT_{sup}$ . Red dots are genes which are statistically significant and belong to the 'angiogenesis' pathway. Blue dots are genes which are statistically significant and belong to the 'tumour invasion' pathway. Black dots are genes which are statistically significant and belong to other pathways. Grey dots are genes which are similarly expressed in  $LN_{met}$  and  $PT_{sup}$ . (C) Volcano plot of genes compared between  $LN_{met}$  and  $PT_{deep}$ . Red dots are genes which are statistically significant and belong to the angiogenesis pathway. Blue dots are genes which are statistically significant and belong to the tumour invasion pathway. Black dots are genes which are statistically significant and belong to other pathways. Grey dots are genes which are similarly expressed in  $LN_{met}$  and  $PT_{deep}$ . ECM, extracellular matrix;  $LN_{met}$ , lymph node metastasis;  $PT_{deep}$ , primary tumour deep;  $PT_{sup}$ , primary tumour superficial.

(table 1). Similar to the analysis comparing  $PT_{sup}$  and  $LN_{met}$ , most of the differentially expressed genes belonged to the angiogenesis and tumour invasion pathways. Genes such as *FGF2*, *WNT5A*, *FGFR1*, *FAP*, *HIF1A*, *PDGFRB* and *CDH11* were overexpressed in  $PT_{deep}$ . The lower number of genes with differential expression between  $PT_{deep}$  and  $LN_{met}$  suggests that the RNA expression profile of the tumour in  $LN_{met}$  is likely more similar to the deep subregion of the primary tumour than to superficial subregions.

We classified samples into  $ITH_{high}$  and  $ITH_{low}$  based on the mean SD of gene expression between subregions. We found no significant relationship between ITH and clinicopathological characteristics (online supplemental tables S10 and S11 and online supplemental figure S5). We found that genes previously used to classify intertumoural heterogeneity of GC such as *KRAS*, *TP53*, *SMAD4* and *WARS* showed low ITH (below the 10th centile).<sup>8 9 22</sup> Certain genes with therapeutic and clinical relevance such as *CEACAM6*, *CEACAM5* and *EGF* had the highest ITH (above the 90th centile) (online supplemental table S12).<sup>23</sup>

In this study, we selected tissue blocks for analysis based on highest tumour content. Subregions were annotated for DNA and RNA extraction and morphometric and phenotypic assessment by a senior GI pathologist (figure 1). To address the possibility that differences in 'contaminating' normal tissue might confound our results, we morphometrically quantified the tumour:stroma ratio per subregion (online supplemental table S13) and compared it between subregions in the same patient. The median tumour:stroma ratio was 0.45 (IQR 0.32–0.60). There was no difference in the tumour:stroma ratio between  $PT_{sup}$  and  $LN_{met}$  (median 0.51 ( $PT_{sup}$ ) vs 0.47 ( $LN_{met}$ ),  $p=0.49$ ), whereas the tumour:stroma ratio between  $PT_{sup}$  and  $PT_{deep}$  was significantly different (median 0.35 ( $PT_{deep}$ ),  $p<0.001$ ). However, it is important to note that we observed a significant overlap in highly expressed genes common to both  $PT_{deep}$  and  $LN_{met}$  regions compared with  $PT_{sup}$  subregions, which is more consistent with cancer cell-related expression. These included several well-known oncogenes and other cancer-related genes

**Table 1** Number of genes in the NanoString panel with significant differences in matched paired analyses of tumour subregions

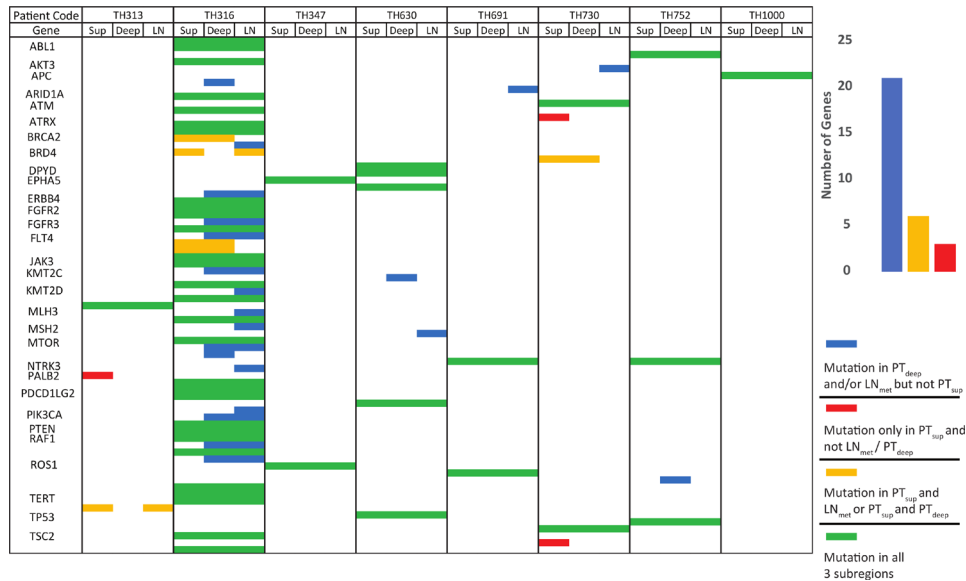
$LN_{met}$ versus $PT_{sup}$ versus $PT_{deep}$	GCs compared (n)	Genes sig different, n (% of 770)	A>B n (% of 770)	A>1.5×B n (% of 770)	B>A n (% of 770)	B>1.5×A n (% of 770)
$PT_{sup1}$ versus $PT_{sup2}$	54	0	0	0	0	0
$PT_{deep}$ versus $PT_{sup}$	57	331 (43)	231 (30)	45 (6)	100 (13)	11 (1)
$LN_{met}$ versus $PT_{sup}$	50	291 (38)	105 (14)	35 (5)	186 (24)	54 (7)
$LN_{met}$ versus $PT_{deep}$	46	127 (16)	10 (1)	5 (<1)	117 (15)	75 (10)

A is the first subregion in the corresponding row of the first column ( $LN_{met}$  vs  $PT_{sup}$  vs  $PT_{deep}$ ).

B is the second subregion in the corresponding row of the first column ( $LN_{met}$  vs  $PT_{sup}$  vs  $PT_{deep}$ ).

For example, in the first row, A= $PT_{sup1}$  and B= $PT_{sup2}$ .

GC, gastric cancer;  $LN_{met}$ , lymph node metastasis;  $PT_{deep}$ , primary tumour deep;  $PT_{sup}$ , primary tumour superficial; sig, significant.



**Figure 3** NGS analysis of spatial heterogeneity in GC. Heatmap of DNA mutations detected by performing NGS of 225 genes per subregion on eight GCs. GCs are represented into columns, with subregions forming subcolumns within each column in the order of PT<sub>sup</sub>, PT<sub>deep</sub> and LN<sub>met</sub>. Genes where mutations were detected in at least one subregion of one GC are represented in rows. Genes which had mutations called in all three subregions (ie, no ITH) were coded green; genes which had mutational calls in PT<sub>sup</sub> but not PT<sub>deep</sub> or LN<sub>met</sub> were coded in red; those with mutational calls in either PT<sub>deep</sub> and/or LN<sub>met</sub> but not PT<sub>sup</sub> were coded in blue; and PT<sub>sup</sub> and PT<sub>deep</sub> but not LN<sub>met</sub> or PT<sub>sup</sub> and LN<sub>met</sub> but not PT<sub>deep</sub> were coded in yellow. Bar graph shows prevalence of mutations detected in PT<sub>deep</sub> and/or LN<sub>met</sub> (blue) but not PT<sub>sup</sub> is significantly higher than any other combination of spatial ITH. ITH, intratumoural heterogeneity; LN<sub>met</sub>, lymph node metastasis; NGS, next-generation sequencing; PT<sub>deep</sub>, primary tumour deep; PT<sub>sup</sub>, primary tumour superficial.

commonly expressed in both PT<sub>deep</sub> and LN<sub>met</sub> regions compared with PT<sub>sup</sub> (eg, *TGFB1*, *PIK3CD*, *IGF1* and *VEGFC*). These findings suggest that differences in tumour content proportions are unlikely to have played a major role in the transcriptomic differences between subregions.

### DNA spatial heterogeneity in gastric cancer

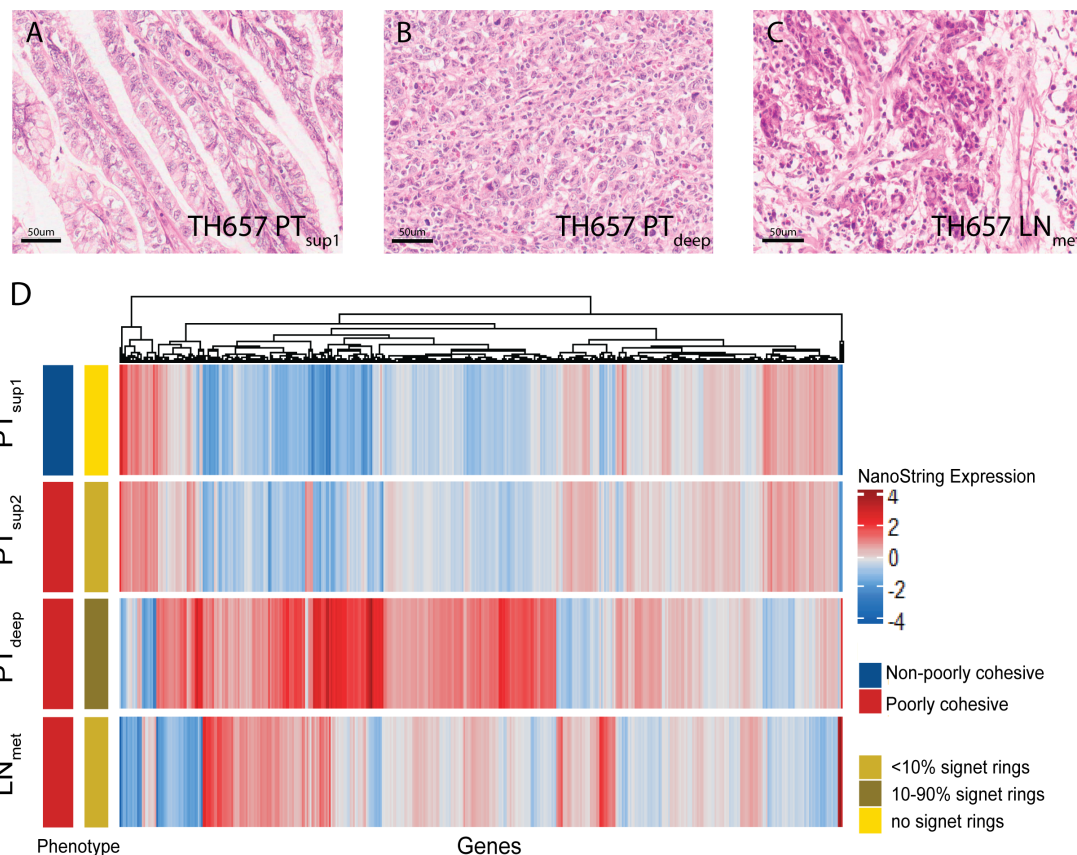
An NGS panel of 225 genes commonly altered in GI cancers was analysed in eight GCs (three subregions per sample: PT<sub>sup</sub>, PT<sub>deep</sub> and LN<sub>met</sub>). Mutations were detected in one to six genes per subregion, except for GC case TH316, where 33–44 mutations per subregion were detected (note that TH316 is an MMR-deficient tumour). Of the 225 genes included in the panel, mutations were detected in 53 (24%) genes. *TP53* mutations were the most commonly detected mutations and present in 50% of GCs. Of these 53 genes, 21 (40%) genes had mutations in PT<sub>deep</sub> and/or LN<sub>met</sub> but not in PT<sub>sup</sub>. Conversely, only three (6%) genes had mutations in PT<sub>sup</sub> but not in PT<sub>deep</sub> and LN<sub>met</sub> (figure 3). For 44 genes, at least one GC had mutations detected in all three subregions. In these, allele frequency values were compared between subregions. In 33 (75%) genes, the allele frequency of the specified gene in the LN<sub>met</sub> subregion was higher than PT<sub>sup</sub> (online supplemental table S14). These findings from the NGS mutation analysis provide orthogonal support of the NanoString data for spatial ITH between subregions, particularly between PT<sub>sup</sub> compared with PT<sub>deep</sub> and/or LN<sub>met</sub>.

Apart from studying ITH of single-base DNA mutations, we also studied gene copy number variation by analysing MLPA subregion data. Of the 64 samples used for the NanoString analysis, 41 (64%) had sufficient DNA quantity and quality to perform the MLPA analysis. There were no significant clinical differences between cases with MLPA data and those without (online supplemental tables S4 and S5). Furthermore, to reduce ‘missing-data bias’, we included additional 20 GCs for MLPA

profiling to have a total of 61 GC for MLPA analysis (additional sample clinical data available in online supplemental table S15). Ninety-one per cent of the 16 genes across 61 samples per three subregions were categorised as ‘normal’ copy number. The highest frequency of gene copy number aberrations was found in *CCNE1* (15%), *GATA4* (16%) and *MYC* (15%), range of gene copy number aberrations 2%–16% (online supplemental table S16). Pairwise comparisons of gene copy number values were made between subregions to study ITH. Out of 976 comparisons (61 samples × 16 genes) for each subregion pairing, the highest discordance was seen between PT<sub>sup</sub> and LN<sub>met</sub> (PT<sub>sup</sub> vs LN<sub>met</sub> 112 (11%) discordant comparisons); followed by PT<sub>deep</sub> versus LN<sub>met</sub> (93 (10%) discordant comparisons) and finally PT<sub>sup</sub> versus PT<sub>deep</sub> (75 (8%) discordant comparisons), consistent with progressive genomic changes from PT<sub>sup</sub> to LN<sub>met</sub>.

### Phenotypical spatial heterogeneity in gastric cancer

In total, 228 subregions from 64 GCs had phenotypical subtyping performed. Sixty-one (27%) subregions were classified as poorly cohesive, 158 (69%) as non-poorly cohesive and 9 (4%) as mucinous. Of the 61 poorly cohesive subregions, 49 (80%) contained signet-ring cells. When intratumoural subregions were compared, only eight GCs (13%) demonstrated different phenotypical classifications between subregions (online supplemental table S17). There was little correlation between phenotypic and transcriptomic ITH. Subregions with the same phenotype showed transcriptomic differences, while in some cases, subregions with different phenotypes appeared to have similar transcriptomic profiles. We selected one GC to illustrate this phenomenon (figure 4). In the majority of the cases which had no phenotypical ITH (n=56, 88%), there was a variable degree of transcriptomic ITH. We selected two GCs with the no phenotypical ITH to illustrate one GC with low transcriptomic



**Figure 4** Phenotypical and transcriptomic heterogeneity in gastric cancer. TH657 is an example of a tumour with phenotypical heterogeneity. (A–C) H&E images of the tumour subregions. (A) Superficial subregion (PT<sub>sup1</sub>) which demonstrates a non-poorly cohesive subtype, with no signet-rings. (B) Deep subregion (PT<sub>deep</sub>), which is poorly cohesive and has 10%–90% signet-rings. (C) Metastatic lymph node (LN<sub>met</sub>) which has <10% signet-rings and is poorly cohesive. (D) Heatmap of NanoString expression profile of TH657 by various spatial subregion. NanoString gene expression is represented in columns, scaled. Blue to red denotes gene expression, with blue implying low gene expression and red implying high gene expression. There is little correlation between phenotypical and transcriptomic heterogeneity. For example, PT<sub>sup1</sub> and PT<sub>sup2</sub> are phenotypically different (PT<sub>sup1</sub> is non-poorly cohesive, while PT<sub>sup2</sub> is poorly cohesive) but have similar transcriptomic profiles. However, PT<sub>sup2</sub> and PT<sub>deep</sub> have the same phenotype, but significantly different transcript profiles. LN, lymph node; LN<sub>met</sub>, lymph node metastasis; PT, primary tumour; PT<sub>deep</sub>, primary tumour deep; PT<sub>sup</sub>, primary tumour superficial.

ITH (ITH<sub>low</sub>) and one GC with extensive transcriptomic ITH (ITH<sub>high</sub>) (figure 5).

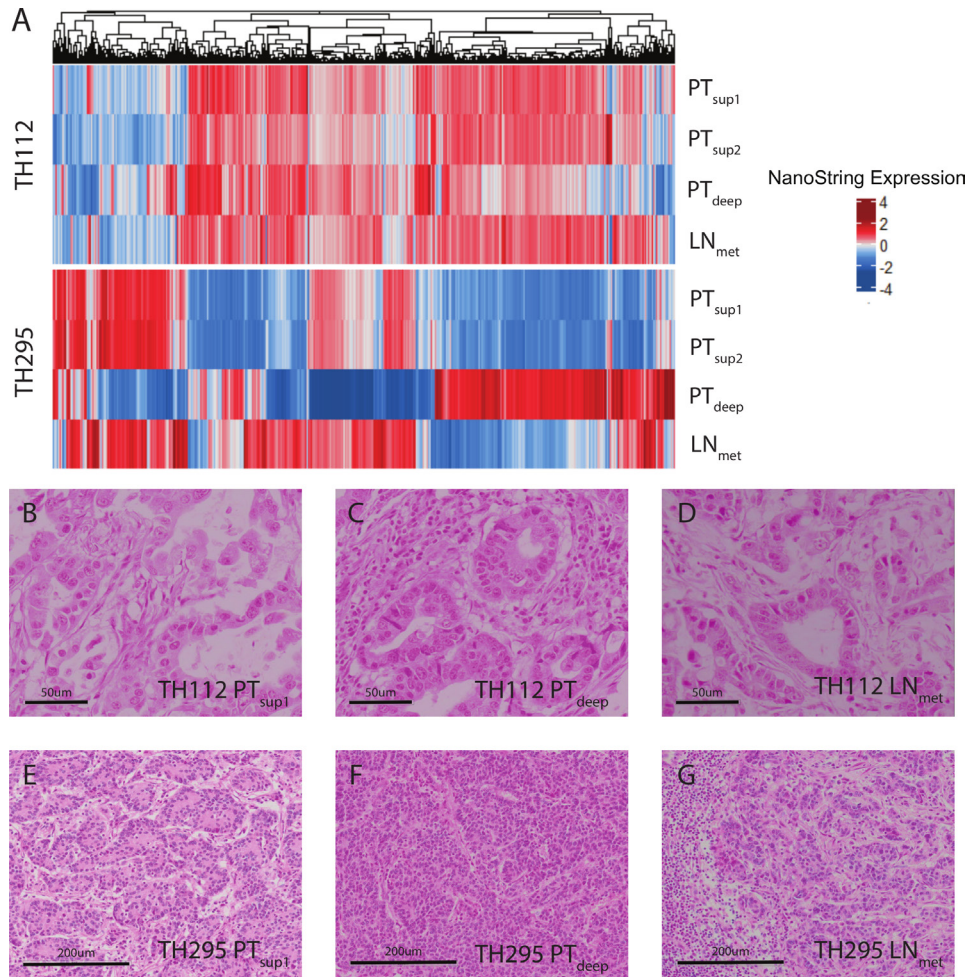
## DISCUSSION

Gastric cancer has long been recognised as a heterogeneous group of adenocarcinomas. Classifying GC into more homogenous subtypes might allow the development of subtype-specific treatments under the overarching umbrella of ‘precision oncology’. To increase ‘coverage’ of potential intratumour heterogeneity, one could consider increasing the number of endoscopic biopsies beyond the current standard of six to eight.<sup>24 25</sup> It has been shown that HER2-positive GC cells are more frequent in the luminal tumour surface compared with deeper layers and that at least five tumour containing endoscopic biopsies are needed for reliable HER2 testing.<sup>26</sup> Similar spatial heterogeneity has been demonstrated when measuring FGFR2 copy number in GC.<sup>27</sup> However, whether a larger number of endoscopic biopsies would also be sufficient to address ITH for other DNA or RNA-based biomarkers, as well as for the histological phenotype, is currently unknown. A recent study suggested that circulating cell-free DNA measured in the blood detects mutations not identified in multiple endoscopic biopsies in GC.<sup>28</sup> Spatial annotation of tumours and biopsy samples has only infrequently

been considered in clinical biomarker tests or research analyses in the past.

In this study, we annotated surgical resection samples and demarcated multiple superficial (luminal/mucosal) and deep regions of primary GC, as well as tumour in matched lymph node metastases, and analysed these subregions individually using multiple platforms such as RNA expression analyses, NGS mutation analysis, DNA copy number and histomorphology. Other studies have demonstrated a relationship between molecular characteristics and histological phenotype in GC but did not include a spatial perspective.<sup>9 19</sup> In oesophageal adenocarcinoma, whole-genome sequencing of matched primary tumour, LN<sub>met</sub> and distant metastatic sites suggests subclonal mutations found in distant metastases were found in spatially distinct areas of the primary tumour.<sup>29</sup> While differentiation between superficial and deep aspects of the tumour was not performed, this study, as well as another similar study performed in colorectal cancer, confirms a distinct model of metastasis with diverse trajectories for lymph node and distant organ dissemination.<sup>30</sup>

Our results yielded several findings of potential clinical and biological relevance. First, we observed minimal differences in gene expression between superficial subregions from the same GC. This suggests that increasing the number of endoscopic



**Figure 5** Phenotype and NanoString gene expression of select GC. (A) Heatmap of NanoString gene expression of TH112 (ITH<sub>low</sub>) and TH295 (ITH<sub>high</sub>). (B–D) are H&E images of TH112 subregions (B=PT<sub>sup1</sub>, C=PT<sub>deep</sub> and D=LN<sub>met</sub>). (E–G) H&E images of TH295 subregions (E=PT<sub>sup1</sub>, F=PT<sub>deep</sub> and G=LN<sub>met</sub>). Phenotypically, both samples do not have significant heterogeneity between subregions—all subregions for both TH112 and TH295 were classified as non-poorly cohesive, no signet-ring cells and no mucin. In the heatmap (A), TH112 (top four rows) is an example of a tumour with low transcriptomic ITH in this cohort, while TH295 (bottom four rows) is an example of high transcriptomic ITH. The heatmap demonstrates obvious heterogeneity in gene expression between TH295 PT<sub>deep</sub> and the rest of the subregions TH295 LN<sub>met</sub>, PT<sub>sup1</sub> and PT<sub>sup2</sub>. In contrast, for TH112, there are minimal differences in gene expression between the LN<sub>met</sub>, PT<sub>sup1</sub>, PT<sub>sup2</sub> and PT<sub>deep</sub>. ITH, intratumoural heterogeneity; LN<sub>met</sub>, lymph node metastasis; PT<sub>deep</sub>, primary tumour deep; PT<sub>sup</sub>, primary tumour superficial

tumour biopsies may not be sufficient to fully characterise intratumour heterogeneity. Next, we observed significant differences in gene expression, mutations and amplification between superficial subregions, deep subregions and LN<sub>met</sub>. Several genes are differentially upregulated in the deeper aspect of the tumour and LN<sub>met</sub>. Interestingly, fewer differences were detected in gene overexpression between the deep subregion of the tumour and the LN<sub>met</sub>. This observation is consistent with a model where tumour cells in the regional LN<sub>met</sub> are more likely to originate from the deep tumour subregion than from the mucosal (superficial) subregion. Based on these findings, biomarker analyses performed on superficial biopsies alone may miss important genomic alterations that occur in the deep subregions and/or metastatic sites which drive tumour oncogenesis and may be potentially targetable. Our results are consistent with the principles of clonal evolution of GI tumours from superficial to deep to metastases which were described over three decades ago by Fearon and Vogelstein.<sup>31</sup> Yet, most of these previous studies were conducted by analysing and comparing tumours with different depths of invasion. The present study is one of the

first to describe this phenomenon within the same tumour in GC. Importantly, from the differential gene expression analysis, genes mapping to angiogenesis and tumour invasion pathways were enriched in the PT<sub>deep</sub> and LN<sub>met</sub> subregions, providing a plausible biological rationale for our findings. In a previous small study (n=15 GCs) of matched primary tumour and lymph node metastases, whole-exome sequencing data were analysed, and mutational abundance was compared. This study suggested that lymph node metastases arise from specific subclonal regions of the primary tumour.<sup>32</sup> These findings can be corroborated with our deeper analyses by spatial subregions.

Matched primary tumour, lymph node and distant metastasis studies of other tumour types (eg, colorectal and lung cancer) report similar findings.<sup>33–37</sup> In a study of melanoma lymph node biopsies, spatial transcriptomics revealed that the lymph node region closest to the primary tumour displayed expression patterns that reflected the tumour microenvironment of the primary tumour.<sup>38</sup> ITH has also been assessed by other groups in gastro-oesophageal adenocarcinoma.<sup>10,39</sup> However, the principal difference between those studies and



ours is spatial mapping performed prior to analysis, demarcating superficial and deep aspects of the tumour with subsequent molecular analysis defined by these spatial regions. Our NGS analysis of ITH in GC included one MMR-deficient tumour. To the best of our knowledge, this is the first description of mutational spatial heterogeneity in an MMR-deficient GC. Given the therapeutic implications of MMR deficiency and treatment with immune checkpoint blockade,<sup>40 41</sup> studies are emerging on ITH as a potential mechanism of resistance to this form of therapy.<sup>42</sup>

Limitations of the present study include the time period during which these samples were accrued from and the analysis being limited only to primary tumours and LN<sub>met</sub> and did not include distant metastatic sites. Our samples were collected from an era before perioperative or adjuvant chemotherapy was administered routinely for GC, and thus, a potential relationship of ITH with chemotherapy cannot be investigated in the current study. However, one might also consider the absence of chemotherapy an advantage as expression profiles, genomic profiles and phenotypical appearances are not altered by preoperative treatment. A retrospective cohort study of spatial heterogeneity in GC has demonstrated poorer survival for those with higher ITH.<sup>43</sup> However, from our study, it is unclear if tumour biology driven by ITH can be altered with chemotherapy or if this is a factor that drives resistance to systemic therapy, which must be considered in future studies. As the inclusion criteria for the study required tumours which were metastatic to the lymph node and that had undergone surgical resection, earlier stage, lymph node-negative tumours and advanced, surgically inoperable tumours were not included. Thus, we do not know whether the observed ITH is specific to tumours with lymph node metastases or can also be observed in primary GC without lymph node metastases. Furthermore, while we acknowledge that analyses of more than one LN<sub>met</sub> from each GC for ITH may produce slightly different results on an individual-patient basis, our results are generally consistent across the cohort, where randomly selected LNs with high tumour cell content were selected for analysis. Moreover, a majority of our samples were of intestinal subtype (>90%), which rendered our study population relatively homogenous with respect to clinicopathological characteristics. While this is not necessarily a limitation, it could explain the lack of correlation of ITH with these clinicopathological features and perhaps limits the generalisability of our findings to other histological subtypes of GC. Discordance between phenotypical and transcriptomic heterogeneity could be due to several reasons, such as the phenotype being driven by genes not included in the NanoString panel or post-translational changes, which would not be detected by NanoString.<sup>44 45</sup>

In conclusion, our study has highlighted clinically significant genomic, transcriptomic and phenotypical heterogeneity in GC defined by spatial subregions. Several therapeutic targets appear to be overexpressed, mutated or amplified only in the deep subregions/metastatic lymph nodes. We have elicited potential biological mechanisms of tumour metastases from deeper subregions of the tumour to lymph nodes. Endoscopists and pathologists must consider the potentially limited value of multiple superficial endoscopic biopsies for the assessment of presence or absence of a biomarker that determines patient treatment and consider investigating deeper parts of the primary tumour and/or metastatic lymph nodes for this purpose. Future studies and clinical trials in GC should take this spatial heterogeneity into consideration.

#### Author affiliations

- <sup>1</sup>Department of Haematology-Oncology, National University Cancer Institute Singapore, National University Health System, Singapore
- <sup>2</sup>Cancer and Stem Cell Biology Program, Duke-NUS Medical School, Singapore
- <sup>3</sup>Yong Loo Lin School of Medicine, National University of Singapore, Singapore
- <sup>4</sup>The N.1 Institute for Health, National University of Singapore, Singapore
- <sup>5</sup>Department of Pathology, GROW - School for Oncology and Developmental Biology, Maastricht University Medical Center+, Maastricht, The Netherlands
- <sup>6</sup>Division of Pathology and Data Analytics, Leeds Institute of Medical Research at St James's, University of Leeds, Leeds, UK
- <sup>7</sup>Department of Surgery, University of Cambridge, Cambridge University Hospitals, Addenbrookes, Cambridge, UK
- <sup>8</sup>Department of Pathology, Leiden University Medical Center, Leiden, Zuid-Holland, The Netherlands
- <sup>9</sup>Department of Upper Gastrointestinal Surgery, Institute of Oncology, Leeds Teaching Hospitals NHS Trust, Leeds, UK
- <sup>10</sup>Laboratory of Cancer Epigenome, Department of Medical Sciences, National Cancer Centre Singapore, Singapore
- <sup>11</sup>Genome Institute of Singapore, Agency for Science, Technology and Research (A\*STAR), Singapore
- <sup>12</sup>Division of Medical Oncology, National Cancer Centre Singapore, Singapore
- <sup>13</sup>SingHealth/Duke-NUS Institute of Precision Medicine, National Heart Centre Singapore, Singapore
- <sup>14</sup>Cancer Science Institute of Singapore, National University of Singapore, Singapore

**Twitter** Jeremy D Hayden @jezhayden

**Contributors** Conceptualisation: GH, PT and HG. Data curation: RS, DL, GH, HS, ANSA, JO, JH, LH, CCYN, AMN, SN, IT, PT and HG. Formal analysis: RS, AMN, SN and HG. Funding acquisition: GH, PT and HG. Methodology: RS, SN, PT and HG. Project administration: RS, HG and PT. Resources: GH, PT and HG. Supervision: HG, PT and HG. Visualisation: RS and DL. Writing, original draft: RS. Writing, review and editing: GH, HG and PT. Approval of final version of manuscript: all authors.

**Funding** RS is supported by a National Medical Research Council (NMRC) Fellowship, Singapore. PT is supported by Duke-NUS Medical School and the Genome Institute of Singapore, Agency for Science, Technology and Research. This work was also supported by National Medical Research Council grants (OF-LCG18May-0023, NR13NMR1110M and NMRC/STaR/0026/2015). This study was supported by Pathological Society Career Development Fellowship (CDF2015/01, PI: GH) and Academy of Medical Sciences Starter Grant for Clinical Lecturers (SGCL14/ GH, PI: GH).

**Disclaimer** The funders of the study had no role in study design, data collection, analysis, interpretation or writing of the manuscript. HG and PT had full access to all data in the study and as corresponding author had final responsibility for the decision to submit for publication.

**Competing interests** PT has stock and other ownership interests in Tempus Healthcare, research funding from Kyowa Hako Kirin and Thermo Fisher Scientific, and patents/other intellectual property through the Agency for Science and Technology Research, Singapore (all outside the submitted work). RS has received honoraria from Bristol-Myers Squibb, Lilly and MSD, has advisory activity with Bristol-Myers Squibb, Eisai and AstraZeneca, received research funding from Paxman Coolers and has received travel grants from AstraZeneca, Roche and Taiho Pharmaceutical (all outside the submitted work). IBHT has received consultation fees for MSD, Taiho, Roche, BMS, Bayer, Amgen and Merck Serono, and research funding from MSD and Taiho (all outside the submitted work). HIG has received honoraria from MSD (outside of the submitted work).

**Patient consent for publication** Not required.

**Ethics approval** The study was approved by the Leeds Research Ethics Committee (LREC no. CA01/122).

**Provenance and peer review** Not commissioned; externally peer reviewed.

**Data availability statement** All data relevant to the study are included in the article or uploaded as supplementary information. Data are available upon reasonable request: Please contact PI Heike Grabsch.

**Supplemental material** This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

**Open access** This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which

permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

#### ORCID iDs

Raghav Sundar <http://orcid.org/0000-0001-9423-1368>

Gordon GA Hutchins <http://orcid.org/0000-0002-1707-4415>

Arnaldo NS Silva <http://orcid.org/0000-0003-2779-9572>

Jan Oosting <http://orcid.org/0000-0001-8894-1742>

Cedric CY Ng <http://orcid.org/0000-0002-7150-4674>

Patrick Tan <http://orcid.org/0000-0002-0179-8048>

Heike I Grabsch <http://orcid.org/0000-0001-9520-6228>

#### REFERENCES

- Bray F, Ferlay J, Soerjomataram I, et al. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* 2018;68:394–424.
- Sundar R, Chénard-Poirier M, Collins DC, et al. Imprecision in the era of precision medicine in non-small cell lung cancer. *Front Med* 2017;4:39.
- Sundar R, Hong DS, Kopetz S, et al. Targeting BRAF-Mutant Colorectal Cancer: Progress in Combination Strategies. *Cancer Discov* 2017;7:558–60.
- Wang J, Xu B. Targeted therapeutic options and future perspectives for HER2-positive breast cancer. *Signal Transduction and Targeted Therapy* 2019;4:34.
- Sundar R, Tan IBH, Chee CE. Negative predictive biomarkers in colorectal cancer: PRESSING ahead. *J Clin Oncol* 2019;37:3066–8.
- Kopetz S, Grothey A, Yaeger R, et al. Encorafenib, Binimetinib, and Cetuximab in BRAF V600E-Mutated Colorectal Cancer. *N Engl J Med* 2019;381:1632–43.
- Sundar R, Tan P. Genomic analyses and precision oncology in gastroesophageal cancer: forwards or backwards? *Cancer Discov* 2018;8:14–16.
- Cristescu R, Lee J, Nebozhyn M, et al. Molecular analysis of gastric cancer identifies subtypes associated with distinct clinical outcomes. *Nat Med* 2015;21:449–56.
- Cancer Genome Atlas Research Network. Comprehensive molecular characterization of gastric adenocarcinoma. *Nature* 2014;513:202–9.
- Pectasides E, Stachler MD, Derks S, et al. Genomic heterogeneity as a barrier to precision medicine in gastroesophageal adenocarcinoma. *Cancer Discov* 2018;8:37–48.
- Mariette C, Carneiro F, Grabsch HI, et al. Consensus on the pathological definition and classification of poorly cohesive gastric carcinoma. *Gastric Cancer* 2019;22:1–9.
- Aoyama T, Hutchins G, Arai T, et al. Identification of a high-risk subtype of intestinal-type Japanese gastric cancer by quantitative measurement of the luminal tumor proportion. *Cancer Med* 2018;7:4914–23.
- Hutchins G, Southward K, Handley K, et al. Value of mismatch repair, KRAS, and BRAF mutations in predicting recurrence and benefits from chemotherapy in colorectal cancer. *J Clin Oncol* 2011;29:1261–70.
- Jassal B, Matthews L, Viteri G, et al. The reactome pathway knowledgebase. *Nucleic Acids Res* 2020;48:D498–503.
- Ashburner M, Ball CA, Blake JA, et al. Gene ontology: tool for the unification of biology. the gene ontology Consortium. *Nat Genet* 2000;25:25–9.
- Ogata H, Goto S, Sato K, et al. Kegg: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res* 1999;27:29–34.
- Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods* 2001;25:402–8.
- Silva ANS, Coffa J, Menon V, et al. Frequent coamplification of receptor tyrosine kinase and downstream signaling genes in Japanese primary gastric cancer and conversion in matched lymph node metastasis. *Ann Surg* 2018;267:114–21.
- Hewitt LC, Saito Y, Wang T, et al. Kras status is related to histological phenotype in gastric cancer: results from a large multicentre study. *Gastric Cancer* 2019;22:1193–203.
- van Eijk R, Eilers PHC, Natté R, et al. MLPAinter for MLPA interpretation: an integrated approach for the analysis, visualisation and data management of multiplex ligation-dependent probe amplification. *BMC Bioinformatics* 2010;11:67.
- Moelans CB, Holst F, Hellwinkel O, et al. Esr1 amplification in breast cancer by optimized RNase fish: frequent but low-level and heterogeneous. *PLoS One* 2013;8:e84189.
- Cheong J-H, Yang H-K, Kim H, et al. Predictive test for chemotherapy response in resectable gastric cancer: a multi-cohort, retrospective analysis. *Lancet Oncol* 2018;19:629–38.
- Toh J, Hoppe MM, Thakur T, et al. Profiling of gastric cancer cell-surface markers to achieve tumour-normal discrimination. *BMJ Open Gastroenterol* 2020;7:e000452.
- Ajani JA, Barthel JS, Bekaii-Saab T, et al. Gastric cancer. *J Natl Compr Canc Netw* 2010;8:378–409.
- Rüschoff J, Hanna W, Bilous M, et al. Her2 testing in gastric cancer: a practical approach. *Mod Pathol* 2012;25:637–50.
- Tominaga N, Gotoda T, Hara M, et al. Five biopsy specimens from the proximal part of the tumor reliably determine HER2 protein expression status in gastric cancer. *Gastric Cancer* 2016;19:553–60.
- Su X, Zhan P, Gavine PR, et al. Fgfr2 amplification has prognostic significance in gastric cancer: results from a large international multicentre study. *Br J Cancer* 2014;110:967–75.
- Chao J, Lee J, Kim K, et al. A pilot study of baseline spatial genomic heterogeneity in primary gastric cancers using Multi-Region endoscopic sampling. *Front Oncol* 2020;10:225.
- Noorani A, Li X, Goddard M, et al. Genomic evidence supports a clonal diaspora model for metastases of esophageal adenocarcinoma. *Nat Genet* 2020;52:74–83.
- Reiter JG, Hung W-T, Lee I-H, et al. Lymph node metastases develop through a wider evolutionary bottleneck than distant metastases. *Nat Genet* 2020;52:692–700.
- Fearon ER, Vogelstein B. A genetic model for colorectal tumorigenesis. *Cell* 1990;61:759–67.
- Lee HH, Kim SY, Jung ES, et al. Mutation heterogeneity between primary gastric cancers and their matched lymph node metastases. *Gastric Cancer* 2019;22:323–34.
- Xie F, Zhang Y, Mao X, et al. Comparison of genetic profiles among primary lung tumor, metastatic lymph nodes and circulating tumor DNA in treatment-naïve advanced non-squamous non-small cell lung cancer patients. *Lung Cancer* 2018;121:54–60.
- Chen Y, Mao B, Peng X, et al. A comparative study of genetic profiles of key Oncogenesis-Related genes between primary lesions and matched lymph nodes metastasis in lung cancer. *J Cancer* 2019;10:1642–50.
- Ullintz PJ, Greenon JK, Wu R, et al. Lymph node metastases in colon cancer are polyclonal. *Clin Cancer Res* 2018;24:2214–24.
- Gundem G, Van Loo P, Kremeyer B, et al. The evolutionary history of lethal metastatic prostate cancer. *Nature* 2015;520:353–7.
- Hatogai K, Fujii S, Kitano S, et al. Relationship between the immune microenvironment of different locations in a primary tumour and clinical outcomes of oesophageal squamous cell carcinoma. *Br J Cancer* 2020;122:413–20.
- Thrane K, Eriksson H, Maaskola J, et al. Spatially resolved transcriptomics enables dissection of genetic heterogeneity in stage III cutaneous malignant melanoma. *Cancer Res* 2018;78:canres.0747.2018–9.
- Murugaesu N, Wilson GA, Birkbak NJ, et al. Tracking the genomic evolution of esophageal adenocarcinoma through neoadjuvant chemotherapy. *Cancer Discov* 2015;5:821–31.
- Le DT, Durham JN, Smith KN, et al. Mismatch repair deficiency predicts response of solid tumors to PD-1 blockade. *Science* 2017;357:409–13.
- Sundar R, Smyth EC, Peng S, et al. Predictive biomarkers of immune checkpoint inhibition in gastroesophageal cancers. *Front Oncol* 2020;10:763.
- von Loga K, Woolston A, Punta M, et al. Extreme intratumour heterogeneity and driver evolution in mismatch repair deficient gastro-oesophageal cancer. *Nat Commun* 2020;11:139.
- Chao J, Bedell V, Lee J, et al. Association between spatial heterogeneity within nonmetastatic gastroesophageal adenocarcinomas and survival. *JAMA Netw Open* 2020;3:e203652.
- Virk RKA, Wu W, Almossalha LM, et al. Disordered chromatin packing regulates phenotypic plasticity. *Sci Adv* 2020;6:eaax6232.
- Prabakaran S, Lippens G, Steen H, et al. Post-Translational modification: nature's escape from genetic imprisonment and the basis for dynamic information encoding. *Wiley Interdiscip Rev Syst Biol Med* 2012;4:565–83.