Comprehensive analysis of 19q12 amplicon in human gastric cancers

Suet Yi Leung[1], Coral Ho[2], I-Ping Tu[3], Rui Li[4], Samuel So[4], Kent-Man Chu[5], Siu Tsan Yuen[1] and Xin Chen[2]

[1]Department of Pathology, The University of Hong Kong, Queen Mary Hospital, Pokfulam, Hong Kong; [2]Department of Biopharmaceutical Sciences, University of California, San Francisco, CA, USA; [3]The Institute of Statistical Science, Acdemia Sinica, Taipei, Taiwan, ROC; [4]Department of Surgery, Stanford University, Stanford, CA, USA and [5]Department of Surgery, The University of Hong Kong, Queen Mary Hospital, Pokfulam, Hong Kong

Amplification at 19q12 has been observed in multiple tumor types, while cyclin E1 (*CCNE1*) has been considered to be the key oncogene within this amplicon. We have previously applied cDNA microarray analysis to systematically characterize gene expression patterns of gastric tumor and nontumor samples. We identified a cluster of five tightly coregulated genes all located at chromosome 19q12, including *CCNE1*. We found that the 19q12 gene cluster is highly expressed in gastric tumors compared to nontumor gastric samples. Array based comparative genomic hybridization and real-time PCR was used to define the boundary of the 19q12 amplicon to a region of approximately 200 kb. Interestingly, we found that in some cases amplification at 19q12 was not associated with DNA copy number gain at *CCNE1*, suggesting that some other genes within the 19q12 amplicon may also have important function during gastric tumorigenesis. We found high expression of the 19q12 gene cluster to be statistically correlated with the cell proliferation gene signature. Using the SAM software, we identified a set of 577 genes whose expression levels positively correlated with the 19q12 gene cluster. GO term analysis revealed that this genelist is enriched with genes involved in cell cycle regulation and cell proliferation. In conclusion, expression array analysis combined with array comparative genomic hybridization and real-time PCR provides a new and powerful tool to identify clusters of genes which may be regulated by genomic DNA aberrations. In addition, our study indicates that amplification at 19q12 is associated with cell proliferation *in vivo*.
Modern Pathology (2006) **19,** 854–863. doi:10.1038/modpathol.3800593; published online 31 March 2006

Keywords: microarrays; array CGH; gastric cancer; amplification

Carcinogenesis is a multistep process involving many genetic and epigenetic events. The former can take the forms of amplification of oncogenes and deletion or mutation of tumor suppressor genes. Mutation of genes involved in maintaining genomic integrity results in catastrophic events, leading to rapid accumulation of chromosomal aberrations.[1] The occurrence of genomic instability is expected to affect global gene expression profiles, and provide a ground for selection of genetic traits that confer a growth advantage important for tumor progression. As a result, most human solid tumors are aneuploid, accumulating numerous gains and losses of chromosomal regions. In some cases, recurrent high levels of DNA copy number amplification can be observed in tumor samples. Among these amplicons many have well-characterized oncogenes, for example, *Myc* at 8q24, *ErbB2* at 17q12 and *cyclin D* at 11q13.

Amplification at chromsome 19q12 has been observed in multiple tumor types[2–5] (also see www.helsinki.fi/cmg/cgh_data). *Cyclin E1* (*CCNE1*), an E type cyclin, is located at 19q12, and has traditionally been considered to be the target of 19q12 amplicon. CCNE1 associates with CDK2 and CDK3 and directs the cell from the late G1 phase into the S phase. It also controls the initiation of DNA replication and centrosome duplication.[6,7] As unchecked cell growth and proliferation is the hallmark of human cancers, the CCNE1/CDK2 kinase system has been shown to be misregulated in multiple tumor types.[2,8] For example, CCNE1 is overexpressed in breast, ovarian, pancreatic, bladder, endometrial and colon cancers.[3–5,9–11] In many cases, the overexpression of CCNE1 is associated with amplification at the 19q12 locus.[2–5]

Gastric cancer is the second most common cause of cancer death worldwide.[12] Environmental and genetic factors are both important in gastric carcinogenesis. Gastritis, caused most commonly due to *Helicobacter pylori* (*H. pylori*) infection, is strongly associated with the development of gastric cancer.[13] Comparative genomic hybridization (CGH) has been applied to study DNA copy number variation among gastric cancers.[14–18] Like many other cancer types, *CCNE1* is found to be amplified at 19q12 and overexpressed in a subset of gastric cancer samples.[19–21] The biological and clinical significance of *CCNE1* amplification and overexpression remains controversial: while some studies reported the correlation of CCNE1 expression with poor patient survival,[20,22] others found no such correlation.[23]

Recently, we reported the systematic characterization of gene expression in a large-scale analysis of gastric adenocarcinomas *vs* non-neoplastic gastric mucosa using cDNA microarrays.[24,25] Hierarchical clustering of global expression profiles of human gastric cancers revealed a tightly coregulated cluster of five genes, all located at chromosome 19q12, including *CCNE1*. In this paper, we report an in depth analysis of this 19q12 gene cluster and demonstrate that high expression of the 19q12 gene cluster is associated with 19q12 locus amplification. Using array-based CGH (array CGH) and real-time PCR we further refine the boundary of 19a12 amplicon to approximately 200 kb along chromosome 19q. We found that high expression of the 19q12 gene cluster statistically correlates with the cell proliferation gene signature. In addition, we identified a set of 577 genes whose expression levels positively correlate with the 19q12 gene cluster. The study therefore indicates that amplification at 19q12 is associated with cell proliferation *in vivo*.

Experimental methods

Patient Samples, RNA Preparation, cDNA Microarray and Hierarchical Clustering

Samples of tumor and normal gastric mucosa were collected from gastrectomy specimens from Department of Surgery, Queen Mary Hospital, The University of Hong Kong. Tissues were frozen in liquid nitrogen within half an hour after they were resected. Non-neoplastic mucosa from stomach was dissected free of muscle, and histologically confirmed to be tumor free by frozen section. Total RNA was extracted using Trizol (Invitrogen, Carlsbad, CA, USA). A detailed description of DNA microarray procedure has been previously reported.[24,25] Raw data are available at both the Stanford Microarray Database (http://genome-www5.stanford.edu/MicroArray/SMD/) and the ArrayExpress online data repository (http://www.ebi.ac.uk/arrayexpress, accession number: E-SMDB-2). The 6688 cDNA clones, representing approxi-

mately 5200 unique genes for hierarchical clustering, were generated as described.[24,25]

Genomic DNA Extraction

Genomic DNA was extracted from gastric tumor samples using QiaAmp kit (Qiagen, Valencia, CA, USA). In all cases, DNA was isolated from adjacent blocks from the same tumors that RNA was extracted and used for microarray analysis.

Array CGH and Analysis

The arrays used in the study were prepared and hybridized as described previously.[26] In brief, human 1.14 arrays were obtained from the UCSF Cancer Center Array Core (http://cc.ucsf.edu/microarray/). The arrays consisted of 2463 bacterial artificial chromosome (BAC) clones that covered the human genome at 1.5 Mb resolution. For hybridization, 1 μg of tumor DNA and 1 μg of gender-matched reference DNA (isolated from normal donor lymphocytes) was labeled by random priming using Cy3-dCTP and Cy5-dCTP, respectively, and hybridized to the arrays. Three single-color intensity images (DAPI, Cy3 and Cy5) were collected from each array using a charge-coupled device camera.

UCSF SPOT software[27] was used to automatically segment the spots based on DAPI images, perform local background correction and to calculate various measurement parameters, including $\log_2$ ratios of the total integrated Cy3 and Cy5 intensities for each spot. A second custom program, SPROC (http://jainlab.ucsf.edu/Downloads.html) was used to associate clone identities and a mapping information file with each spot, so that the data could be plotted relative to the position of the BACs. Chromosomal aberrations were classified as a gain when the normalized $\log_2$ Cy3/Cy5 ratio was higher than 0.225; and as a loss when the ratio was lower than -0.225. Steep copy number changes with the graph showing a peak rather than a plateau, and a minimal normalized $\log_2$ Cy3/Cy5 ratio of 0.9 or higher, were classified as amplifications. Likewise, $\log_2$ Cy3/Cy5 ratio of -0.8 or lower were classified as homozygous deletions.

Quantitative Real-Time PCR to Determine DNA Copy Number Changes

Sybergreen-based real-time PCR was performed to assay the DNA copy number variation among nine loci at 19q12. House keeping gene *GAPDH* was used as the reference control. The detailed method of quantitative real-time PCR is the same as previously described.[28] Assays were carried out using the software supplied with the ABI 7900 (Applied Biosystems, Foster city, CA, USA). Control samples were selected based on array CGH results: gastric

cancers with 19q12 amplication demonstrated by array CGH (for example HKG91T and HKG92T) were chosen as positive control, whereas gastric cancers showed no 19q12 amplication by array CGH (HKG81T), as well as DNA from normal blood samples were used as negative control. All samples were assayed as triplicate and the values were averaged. DNA copy number was calculated as $2^{-\Delta\Delta C_t}$. Copy numbers of two blood samples from healthy donors were averaged and normalized to 2, which was subsequently used to normalize the copy numbers of gastric cancer samples. The primers were designed using Primer Expression program and the sequences are available in Supplementary Table 1. The efficiency of primers was tested to be between 90 and 100%.

Statistical Analysis

To identify gene expression signatures associated with the 19q12 gene cluster, expression values (normalized $\log_2$ (red/green) ratio) of five clones in the cluster were averaged (19q12 cluster expression vector). The Pearson's correlation coefficient was then calculated for every clone in the whole cluster (comprised of 6688 clones) against the 19q12 cluster expression vector. To smooth the correlation curve, the moving average of 21 clones was calculated and plotted against the main cluster (along the vertical axis) as described.[29] The significant threshold for the moving average correlations was set by permuting the data 10 000 times. For each permutation, we permuted the 19q12 cluster expression vector, constructed the corresponding smoothed correlation curve and picked the maximum of the curve. We used the α-percentile of the 10 000 maximum values to determine the threshold of α significant level for the moving average correlations. We also applied the significance analysis of microarrays (SAM) method[30] and used the 19q12 cluster expression vector as a continuous variable to identify genes that were significantly statistically correlated with the 19q12 cluster expression vector. Gene ontology categories were analyzed by GO-TermFinder.[31]

Results

19q12 Gene Cluster Revealed by Global Expression Profiling

We previously reported the global gene expression pattern of human gastric tumors, including 90 tumors, 14 lymph nodes and 22 nontumor gastric mucosa.[24,25] A hierarchical clustering algorithm was used to group 126 gastric tissue samples according to their gene expression patterns as delineated by 6688 cDNA clones (Figure 1a). This analysis revealed prominent clusters of coexpressed genes, which included: a set of clusters that appeared to reflect intrinsic characteristics of proliferating tumor

cells (ie, a proliferation and beta-catenin cluster); normal gastric mucosal cells (ie, normal cluster); differentiation (ie, intestinal metaplasia cluster), stromal cells (ie, ECM cluster), and that of infiltrating lymphocytes (ie, leukocyte cluster). Interestingly, we also noticed several gene clusters that were comprised of genes located at close range on chromosomes (ie, Erbb2 cluster).

One such cluster contained five genes (POP4, C19or12, RMP, CCNE1 and UQCRFS1), all located at chromosome 19q12 position 34.3–35.2 Mb (Figure 1b and Table 1). We refer to this cluster as the 19q12 gene cluster. To further illustrate the correlation between these five genes along chromosome 19, we retrieved genes which are located at chromosome 19 from the main cluster as shown in Figure 1a, and organized the genes along their chromosomal positions. We constructed a correlation map of chromosome 19 using 20 genes upstream of the five 19q12 cluster genes and 20 genes downstream (Figure 2). This displayed pairwise correlations between expression patterns of the genes and allows visual identification of groups of adjacent genes with similar gene expression patterns. Figure 2 depicts the group of five correlated 19q12 genes as a red block centered on the diagonal of the correlation matrix. The data further illustrated that these five genes located at 19q12 showed highly correlated gene expression patterns.

The average expression levels of the 19q12 gene cluster in 90 gastric tumors and 22 nontumor gastric mucosa samples is shown in Figure 3. We found that the 19q12 gene cluster is statistically significantly highly expressed in tumor samples compared with nontumor gastric mucosa ($P = 3.7 \times 10^{-5}$). Among the 90 gastric tumor samples, 10 showed expression of 19q12 gene cluster at very high levels (we used $\log_2$ (average array expression level) > 0.8 as the cut off for high expression). The data indicated that there might be amplification at 19q12 in these 10 gastric tumor samples.

19q12 Amplicon by Array CGH

Multiple genes located at 19q12 showing similar overexpression in a subset of gastric cancer samples suggest that the 19q12 amplicon may contain a rather large region of multiple genes. To investigate whether high expression of the 19q12 gene cluster is due to amplification at 19q12, we applied array CGH in a subset of gastric cancer samples (all of them has been used in expression array studies). In all cases, gastric cancer samples without high expression of 19q12 gene cluster do not show 19q amplification. A representative image of DNA copy number on chromosome 19 of a gastric tumor without high expression of 19q12 gene cluster is shown in Figure 4a. On the other hand, seven of the seven gastric cancers with high expression of 19q12 gene cluster showed either amplification ($n = 6$) or a one copy
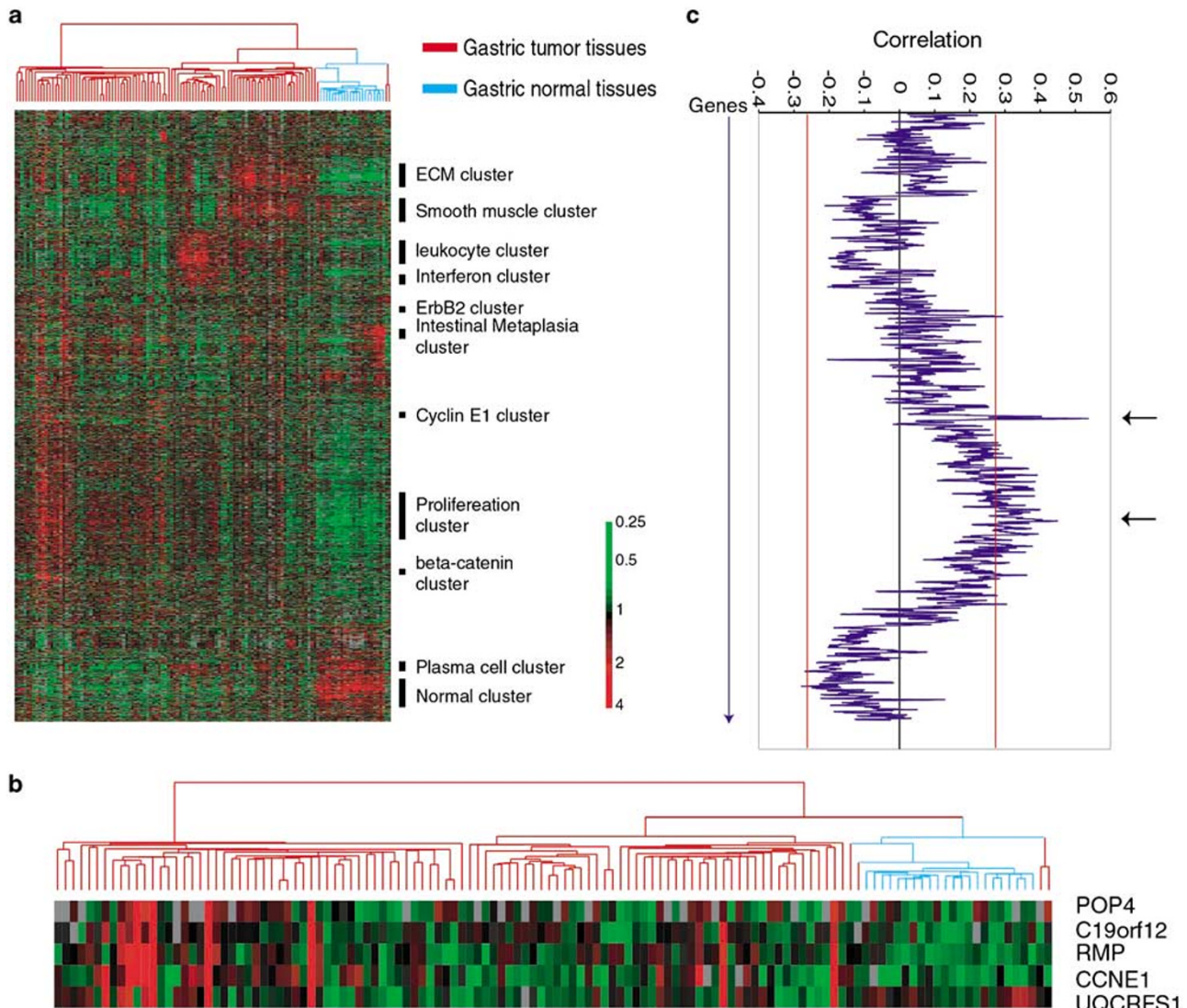
Figure 1 Expression patterns of human gastric tissues. (**a**) Hierarchical clustering of 6688 cDNA clones (corresponding to 5200 unique genes) in 126 gastric tissue samples, including gastric tumors ($n = 90$), metastasis lymph nodes ($n = 14$) and nontumor gastric mucosa ($n = 22$). The figure is presented in table format, where each row represents a gene and each column represents a sample. The relative abundance of each transcript centered across all the samples is depicted according to the color scale shown at the lower right corner. Gray indicates missing or excluded data. Specific gene clusters identified in the gastric tissue expression profiles are annotated to the right. See previous publication[24,25] for a detailed analysis of gene expression patterns in human gastric cancers. (**b**) Expanded view of the 19q12 gene cluster. (**c**) Correlation of expression levels between 19q12 gene cluster and global gene expression signature in human gastric cancer samples. Genes are organized along the *X*-axis the same order as in the clustering (**a**). The Pearson's correlation coefficient was calculated for every clone in the whole cluster (comprised of 6688 clones) against the 19q12 cluster expression vector. The correlation values (the blue line) are plotted as moving average of 21 genes (along the vertical axis). The red lines represent thresholds for the 2.5% one-sided tail probability under 10 000 permutations. Arrow indicates regions of statistically significant correlation. Part of the figure is the same as figures published previously[25] with the permission to use from Molecular Biology of Cell.

Table 1 Five genes in the CCNE1 cluster

Gene symbol	Gene name	Cytoband	Position
POP4	POP4 (processing of precursor, *Saccharomyces cerevisiae*) homolog	19q12	chr19:34,789,052-34,791,982
C19orf12	Chromosome 19 open reading frame 12	19q12	chr19:34,883,915-34,897,989
RMP	RPB5-mediated protein	19q12	chr19:35,125,264-35,198,451
CCNE1	Cyclin E1	19q12	chr19:34,994,751-35,007,054
UQCRFS1	Ubiquinol-cytochrome *c* reductase, Rieske iron–sulfur polypeptide 1	19q12	chr19:34,395,802-34,395,897

Position refers to the genomic location of cDNA clones retrieved from the UCSC genome browser based on the May 2004 assembly.
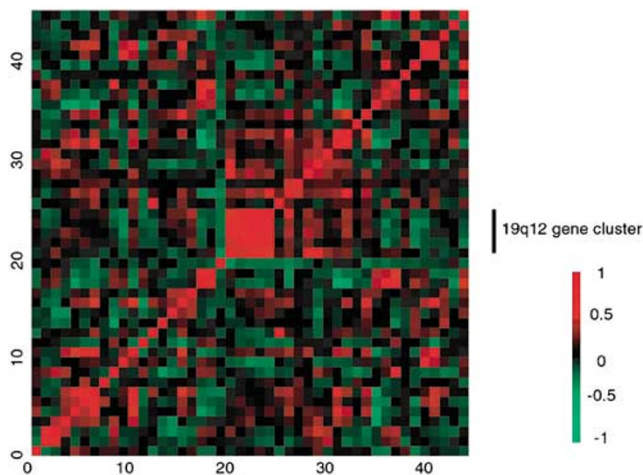
Figure 2 Correlation matrix of 45 genes at chromosome 19 (20 genes upstream of 19q12 gene cluster and 20 genes downstream of 19q12 cluster). All the pairwise correlations between gene expression profiles of 126 gastric samples of the chromosome 19 genes are displayed as a pseudocolor map in which color and intensity of each element depict direction and degree of correlation, respectively. Genes are ordered along both axes based on their genomic position along the chromosome 19.
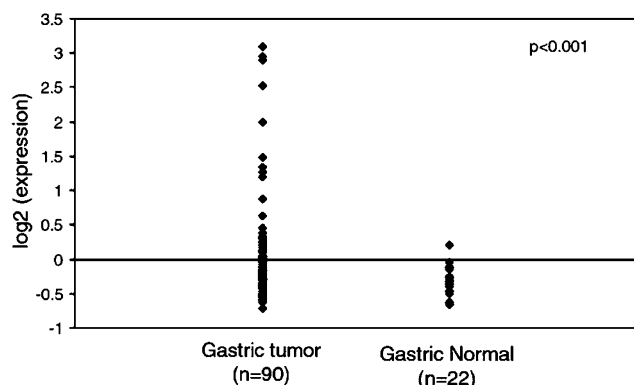

Figure 3 Expression of 19q12 gene cluster in gastric cancer ($n = 90$) and normal ($n = 22$) samples revealed by cDNA micro-arrays. P-value was calculated using Student's t-test.

Figure 4 Representative array CGH images of chromosome 19 for gastric cancer samples without (**a**) or with (**b**) 19q12 amplification.

Table 2 Five BAC clones located at 19q12 were most frequently amplified in gastric cancer samples

BAC clone	CytoBand	Position (Kb)
RP11-206O6	19q12	33 315
RP11-152P7	19q12	33 418
RP11-46I12	19q12	34 302
RP11-110J19	19q12	34 465
CTB-25O22[a]	19q12	35 007

[a]Minimal overlap BAC clone for 19q12 amplicon.

gain ($n = 1$) at 19q12 region. A representative image of 19q12 amplicon assayed by array CGH is shown in Figure 4b. We compared the BAC clones at 19q12 in these gastric cancer samples and identified a set of five BAC clones which were most frequently amplified (Table 2). Particularly, BAC clone CTB25O22 showed amplification or gain in all seven samples. The five most frequently amplified BAC clones are mapped from 33.3 to 35.0 Mb, which is consistent with the position of the five clones in the 19q12 gene cluster (Table 1).

19q12 Amplification in Human Gastric Cancers Assay by Real-Time PCR

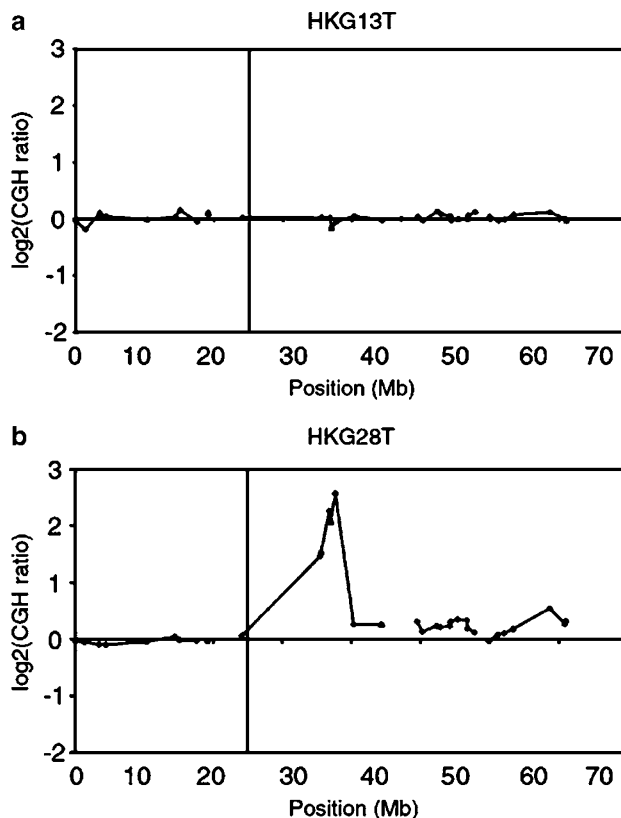While array CGH analysis showed DNA copy number variations among tumor samples on a genome-wide basis, it only provided limited resolution when applied to the mapping of the boundary of amplicon. This is mainly due to the fact that there are on average 1.5 Mb distance between each BAC clone printed on the array. To further map the 19q12 amplicon, we identified genes located at 19q12 (Figure 5a). We performed real-time PCR to assay the genomic DNA copy number variation of nine genes at 19q12 in 10 gastric cancers with high expression of 19q12 gene cluster, as well as two control gastric cancer samples without high expression of 19q12 gene cluster. A composite figure showing all of the 10 gastric cancers with high 19q expression and the two with low expression is available as Supplementary Figure 1. Figure 5b
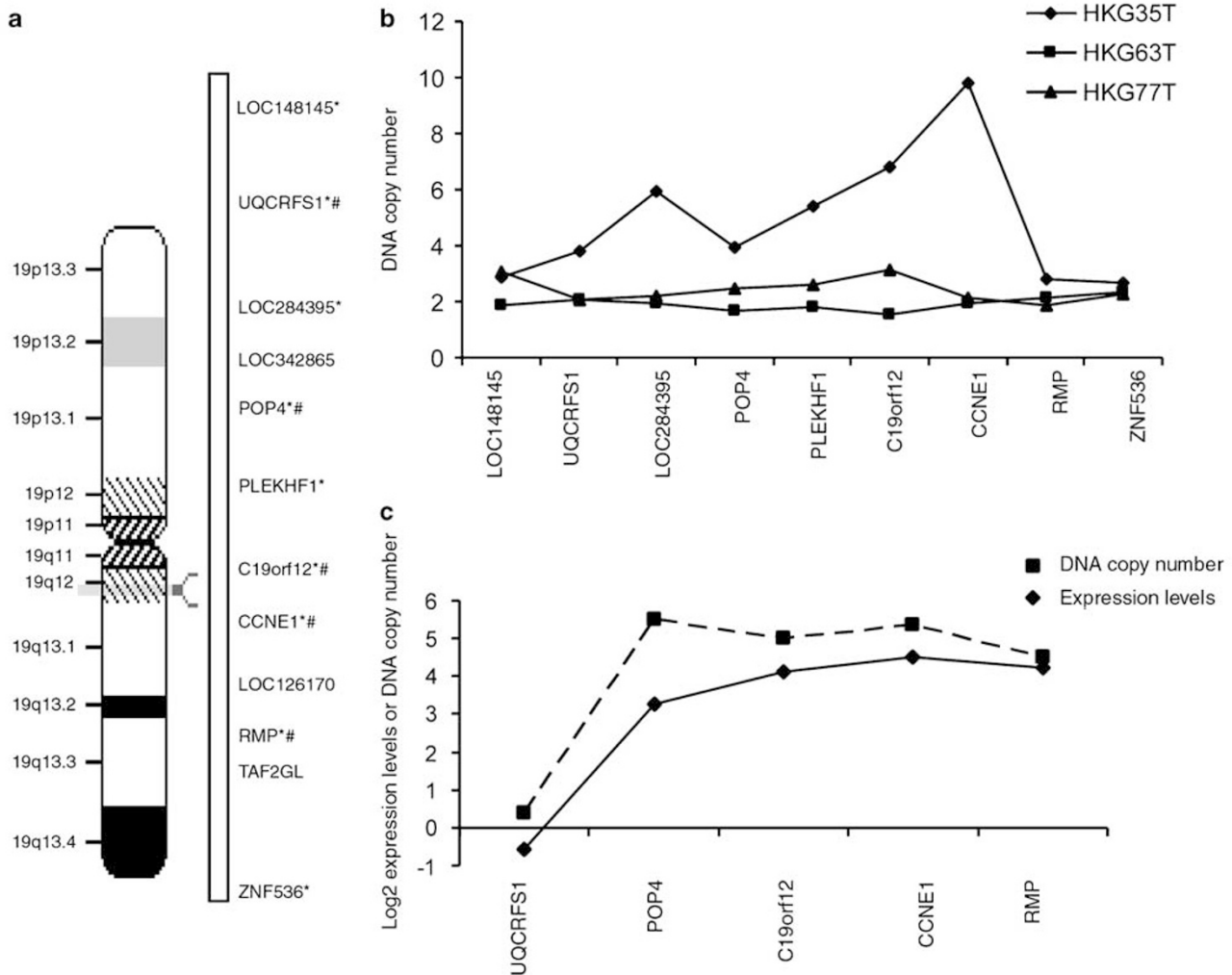
Figure 5 Mapping of 19q12 amplicon using real-time quantitative PCR. (**a**) Genes located at 19q12. #Genes in 19q12 cluster; and *genes assayed by real-time PCR. (**b**) DNA copy number assayed by real-time PCR in three gastric cancer samples. HKG35T: a gastric cancer sample with 19q12 amplification; HKG77T: a gastric cancer sample with one copy gain at 19q12, note the lack of DNA copy number change at *CCNE1* locus; HKG63T: a gastric cancer sample with normal DNA copy number at 19q12. (**c**) Gene expression level parallels genomic DNA amplication at 19q12 in HKG92T gastric tumor sample.

show representative images from gastric cancer samples with or without high expression of 19q12 gene cluster. Real-time PCR mapped the most frequently amplified region (defined as genes with showed copy number >3 in more than 80% of the samples) to be from *POP4* to *CCNE1* (roughly 34.79–35.00 Mb) among chromosome 19q12, with locus at C19orf12 being amplified in all 10 samples tested.

We compared the DNA copy number assayed by real-time PCR and gene expression levels assayed by microarrays, and found in most cases the expression levels paralleled with the genomic DNA copy numbers (Supplementary Figure 2). Figure 5c illustrates such a correlation in gastric tumor sample HKG92T.

Interestingly we found that HKG77, which clearly showed one copy gain at 19q12, has normal copy number at the *CCNE1* locus (Figure 5b). We also analyzed the DNA copy number of *CCNE1* in

HKG77T at a different position using a different primer set. We consistently found no evidence of DNA copy number gain or amplification of *CCNE1* in HKG77T (data not shown). The data therefore suggest that DNA copy number gain at 19q12 is not necessarily associated with amplification at *CCNE1* locus.

Gene Expression Patterns Associated with the 19q12 Gene Cluster

In an attempt to identify features in the global gene expression cluster that may be associated with expression profile of the 19q12 gene cluster, we calculated correlations for each gene expression vector with the average expression vector of 19q12 gene cluster among the 126 gastric samples. The resulting correlation curve, plotted as moving

Table 3 Top 20 clones identified by SAM to be statistically significantly correlated with the expression of the 19q12 gene cluster

Accession number	GeneSymbol	Cell cycle regulated
AA450020	RMP	No
N73011	c19orf21	No
T54121	CCNE1	Yes
R67616	POP4	No
AA448184	UQCRFS1	No
AA026682	TOP2A	Yes
AA460115	ODC1	No
AI341712	FLJ12528	No
H95638	ILF2	Yes
AA504348	TOP2A	Yes
AA188768	WHIP	No
AA862529	SMT3H1	No
AI734240	TOP2A	Yes
AI362799	D21S2056E	Yes
AA160498	MRPS18B	Yes
AA278629	DDA3	Yes
AA670347	MTX1	No
AA485978	KIAA1688	No
AA699361	SF3B4	No
AA022472	HTCD37	No

The status of genes being regulated by cell cycle in Hela cells is retrieved from the publication by Whitfield et al.[32]

Table 4 GO terms that are nonrandomly enriched (corrected P-value <0.05) with the products of the 577 genes whose expression levels are correlated with expression of the 19q12 gene cluster in gastric cancer samples

Gene ontology term	GO category	Corrected P-value
Cell cycle	Biological processes	4.71E-10
Mitotic cell cycle	Biological processes	2.94E-09
M phase	Biological processes	1.05E-08
Nuclear division	Biological processes	2.23E-08
Mitosis	Biological processes	9.93E-08
M phase of mitotic cell cycle	Biological processes	1.46E-07
Cell proliferation	Biological processes	3.38E-05
Regulation of cell cycle	Biological processes	0.00026
Cytokinesis	Biological processes	0.00118
Regulation of mitosis	Biological processes	0.00656
DNA metabolism	Biological processes	0.01362
Mitotic checkpoint	Biological processes	0.01436
Chromosome, pericentric region	Molecular functions	0.01714

Corrected P-value is calculated by the GO-TermFinder program.

averages (window size = 21 genes), are displayed in Figure 1c. We observed two positively correlated peaks, one representing the 19q12 gene cluster itself, and the other representing the cell proliferation cluster (arrows in Figure 1c). Similar results were obtained when we used a moving average window size of 11 or when we only used the data from 90 gastric cancer tissues. The cell proliferation gene cluster is comprised of genes whose functions are required for cell cycle progression and whose expression levels correlate with cellular proliferation rates. To determine whether the observed correlation is statistically significant, we permuted the data set and calculated the correlation. The red lines in Figure 1c represent the thresholds which are derived by the permutation test as described in the section 'Statistical Analysis' for the 2.5% one-sided tail probability. Clearly the two peaks observed are statistically significantly correlated with the 19q12 gene cluster.

To further identify gene expression signatures that may be regulated by the 19q12 gene cluster, we applied SAM to the data set using the 19q12 gene cluster expression value as a continuous variable vector. A delta value of 2.02 was selected, and the program identified 577 positively correlated genes with the median false significant number of 0.46 (<0.1%). We named this gene list as '19q12 gene cluster correlated genes.' The top 20 genes are shown in Table 3 and the complete gene list is available as Supplementary Table 2. The top five genes are genes of the 19q12 gene cluster itself. Careful examination of the rest of the genes revealed that many of them have important functions during

cell cycle progression and cell proliferation, for example, topoisomerase II alpha (TOP2A), CDC2, centromere protein A (CENPA), translin (TSN), tubulin beta 2 (TUBB2) and BUB1.

Our analysis of the Gene Ontology annotations for the 19q12 gene cluster correlated genes in gastric cancer indicated nonrandom enrichment in a variety of biological process categories, including cell cycle, nuclear division, cytokinesis and cell proliferation (Table 4). Furthermore, expression levels of many of these 19q12 gene cluster correlated genes are regulated during cell cycle. We downloaded the 1134 cell cycle-regulated genes previously published[32] and found that in the main cluster in Figure 1a, 440 clones (6.58%) of the total 6688 clones are cell cycle regulated. On the other hand, 127 clones (22.0%) of the total 577 19q12 gene cluster-correlated genes are cell cycle regulated. This difference is high statistically significant using the hypergeometric test ($P < 0.001$). The data therefore indicate that expression of the 19q12 gene cluster is associated with cell proliferation gene expression signature in vivo in gastric tumors.

Discussion

We have shown in this study that by combining microarray expression data, based array CGH and quantitative real-time PCR, we can identify small regions of chromosomal amplification and the resulting genes that are upregulated in gastric cancers due to the amplification. Furthermore, we can associate on a genomic scale the genes that are correlated with the amplification and support a role of the amplification of 19q12 and cell proliferation in vivo.

Genomic analyses using microarrays has provided us global programs of gene expression patterns of human cancers.[33,34] Analysis of gene expression programs along chromosomes can help reveal unidirectional changes in expression of a large number of adjacently located genes, which is generally named as regional expression bias.[35–37] As expected, a majority of the detectable regional gene expression biases in tumor samples also coincide with chromosomal aberrations.[38–40] It is therefore feasible to infer cytogenetic abnormalities by examination of high-density gene expression data. However, these analysis methods focus on the whole chromosome, and it is easy to detect DNA copy number gains or losses along the chromosomes that occur frequently in tumor cells. Amplification or deletion events which occur at lower frequency (eg, <10 or 15%) and at a narrow region (<1 Mb) will be difficult to detect with these analysis methods. In this paper, we have shown that by simply using hierarchical clustering, we were able to identify several clusters of coregulated genes which may represent amplification events in tumor cells. Besides the 19q12 gene cluster which has been detailed in this paper, there is also a cluster at 17q12, which includes *ErBB2*, *GRB7* and *MLN64*; a cluster at 20q13, which includes *ZNF217*, *ATP5E* and *GNAS*; a cluster at 7q22, which includes *ATP5J2*, *CPSF4* and *PDAP1*; and a cluster at Xq28, which includes *ARD1*, *IRAK1* and *UBL4* (data not shown). Therefore, mapping the gene clusters along the chromosomes may help to identify narrow and less frequent gene amplification or deletion events in tumor samples.

In this paper, we focused our analysis of 19q21 gene cluster, it is important to further expand the study to other commonly amplified regions in gastric cancers. In addition, correlation of expression arrays and array CGH data will clearly assist to identify genes whose expression levels are altered because of the DNA copy number gain or loss, and possibly help to identify novel oncogenes and tumor suppressor genes in gastric cancers. The analysis between gene expression and DNA copy number variation is currently in progress and will be present separately (Leung *et al*, unpublished data).

While current studies of 19q12 amplicon have focused on *CCNE1*, we found that amplification at 19q12 is not always necessarily associated with DNA copy number gain at *CCNE1* locus. The study therefore suggests that other genes within this amplicon may also have potential functions in tumor development. For example, RMP has been found to be expressed in small cell lung carcinoma cells and Reed–Sternberg cells of Hodgkin's disease.[41] It binds to RNA polymerase II subunit 5 and functions as a transcriptional corepressor.[42,43] UQCRFS1 is also known as the complex III of the mitochondrial respiratory chain.[44] This complex passes electrons from reduced ubiquinol to cytochrome *c* during the process of synthesis of ATP. The amplification or overexpression of *UQCRFS1* may therefore play an important role for dividing and proliferating cancer cells. The amplification of *UQCRFS1* has also been reported in breast and ovarian cancers.[45,46] Clearly, it is important to further evaluate the functions of these genes in gastric cancer development and determine the extent to which their gene expression is regulated by DNA amplification. In addition, these genes within 19q12 amplicon may also serve as treatment targets. For example, an antibody against POP4, C19orf12, UQCRFS1 or RMP may be useful to treat gastric cancer patients who have 19q12 amplication.

In summary, our study shows that expression array analysis combined with array CGH and real-time PCR provides a new and powerful tool to identify clusters of genes which may be regulated by genomic DNA aberrations. In addition, our study indicates that amplification at 19q12 is associated with cell proliferation *in vivo*.

Acknowledgements

References

1 Lengauer C, Kinzler KW, Vogelstein B. Genetic instabilities in human cancers. Nature 1998;396: 643–649.

2 Schraml P, Bucher C, Bissig H, *et al.* Cyclin E overexpression and amplification in human tumours. J Pathol 2003;200:375–382.

3 Tsuda H, Bandera CA, Birrer MJ, *et al.* Cyclin E amplification and overexpression in clear cell adenocarcinoma of the ovary. Oncology 2004;67:291–299.

4 Richter J, Wagner U, Kononen J, *et al.* High-throughput tissue microarray analysis of cyclin E gene amplification and overexpression in urinary bladder cancer. Am J Pathol 2000;157:787–794.

5 Cassia R, Moreno-Bueno G, Rodriguez-Perales S, *et al.* Cyclin E gene (CCNE) amplification and hCDC4 mutations in endometrial carcinoma. J Pathol 2003; 201:589–595.

6 Koff A, Giordano A, Desai D, *et al.* Formation and activation of a cyclin E-cdk2 complex during the G1 phase of the human cell cycle. Science 1992;257:1689–1694.

7 Zariwala M, Liu J, Xiong Y. Cyclin E2, a novel human G1 cyclin and activating partner of CDK2 and CDK3, is induced by viral oncoproteins. Oncogene 1998;17: 2787–2798.

19q12 amplicon in gastric cancers
SY Leung et al

862

8 Akli S, Keyomarsi K. Cyclin E and its low molecular weight forms in human cancer and as targets for cancer therapy. Cancer Biol Ther 2003;2:S38–S47.

9 Keyomarsi K, Tucker SL, Buchholz TA, et al. Cyclin E and survival in patients with breast cancer. N Engl J Med 2002;347:1566–1575.

10 Al-Aynati MM, Radulovich N, Ho J, et al. Over-expression of G1-S cyclins and cyclin-dependent kinases during multistage human pancreatic duct cell carcinogenesis. Clin Cancer Res 2004;10:6598–6605.

11 Sutter T, Dansranjavin T, Lubinski J, et al. Over-expression of cyclin E protein is closely related to the mutator phenotype of colorectal carcinoma. Int J Colorectal Dis 2002;17:374–380.

12 Parkin DM, Pisani P, Ferlay J. Estimates of the world-wide incidence of 25 major cancers in 1990. Int J Cancer 1999;80:827–841.

13 Parsonnet J, Friedman GD, Vandersteen DP, et al. Helicobacter pylori infection and the risk of gastric carcinoma. N Engl J Med 1991;325:1127–1131.

14 Sakakura C, Mori T, Sakabe T, et al. Gains, losses, and amplifications of genomic materials in primary gastric cancers analyzed by comparative genomic hybridization. Genes Chromosomes Cancer 1999;24: 299–305.

15 Koo SH, Kwon KC, Shin SY, et al. Genetic alterations of gastric cancer: comparative genomic hybridization and fluorescence in situ hybridization studies. Cancer Genet Cytogenet 2000;117:97–103.

16 Wu MS, Chang MC, Huang SP, et al. Correlation of histologic subtypes and replication error pheno-type with comparative genomic hybridization in gastric cancer. Genes Chromosomes Cancer 2001;30: 80–86.

17 Kim YH, Kim NG, Lim JG, et al. Chromosomal alterations in paired gastric adenomas and carcinomas. Am J Pathol 2001;158:655–662.

18 Weiss MM, Kuipers EJ, Postma C, et al. Genomic profiling of gastric cancer predicts lymph node status and survival. Oncogene 2003;22:1872–1879.

19 Akama Y, Yasui W, Yokozaki H, et al. Frequent amplification of the cyclin E gene in human gastric carcinomas. Jpn J Cancer Res 1995;86:617–621.

20 Ahn MJ, Kim BH, Jang SJ, et al. Expression of cyclin D1 and cyclin E in human gastric carcinoma and its clinicopathologic significance. J Korean Med Sci 1998;13:513–518.

21 Hofler H, Becker KF. Molecular mechanisms of carcinogenesis in gastric cancer. Recent Results Cancer Res 2003;162:65–72.

22 So JB, Samarasinghe K, Raju GC. Expression of cell-cycle regulators p27 and cyclin E correlates with survival in gastric carcinoma patients. J Surg Res 2000;94:56–60.

23 Chetty R, Sitti CW. Cyclin E immunoexpression in gastric cancer does not correlate with clinicopatho-logical parameters. Histopathology 2003;42:66–69.

24 Leung SY, Chen X, Chu KM, et al. Phospholipase A2 group IIA expression in gastric adenocarcinoma is associated with prolonged survival and less frequent metastasis. Proc Natl Acad Sci USA 2002;99:16203–16208.

25 Chen X, Leung SY, Yuen ST, et al. Variation in gene expression patterns in human gastric cancers. Mol Biol Cell 2003;14:3208–3215.

26 Patil MA, Gutgemann I, Zhang J, et al. Array-based comparative genomic hybridization reveals recurrent chromosomal aberrations and Jab1 as a potential target for 8q gain in hepatocellular carcinoma. Carcino-genesis 2005;26:2050–2057.

27 Jain AN, Tokuyasu TA, Snijders AM, et al. Fully automatic quantification of microarray image data. Genome Res 2002;12:325–332.

28 Ginzinger DG, Godfrey TE, Nigro J, et al. Measure-ment of DNA copy number at microsatellite loci using quantitative PCR analysis. Cancer Res 2000;60:5405–5409.

29 Whitney AR, Diehn M, Popper SJ, et al. Individua-lity and variation in gene expression patterns in human blood. Proc Natl Acad Sci USA 2003;100: 1896–1901.

30 Tusher VG, Tibshirani R, Chu G. Significance analysis of microarrays applied to the ionizing radia-tion response. Proc Natl Acad Sci USA 2001;98: 5116–5121.

31 Boyle EI, Weng S, Gollub J, et al. GO::TermFinder—open source software for accessing gene ontology information and finding significantly enriched gene ontology terms associated with a list of genes. Bioinformatics 2004;20:3710–3715.

32 Whitfield ML, Sherlock G, Saldanha AJ, et al. Identification of genes periodically expressed in the human cell cycle and their expression in tumors. Mol Biol Cell 2002;13:1977–2000.

33 Chung CH, Bernard PS, Perou CM. Molecular portraits and the family tree of cancer. Nat Genet 2002;32 (Suppl):533–540.

34 Holloway AJ, van Laar RK, Tothill RW, et al. Options available—from start to finish—for obtaining data from DNA microarrays II. Nat Genet 2002;32(Suppl): 481–489.

35 Crawley JJ, Furge KA. Identification of frequent cytogenetic aberrations in hepatocellular carcinoma using gene-expression microarray data. Genome Biol 2002;3:RESEARCH0075.

36 Husing J, Zeschnigk M, Boes T, et al. Combining DNA expression with positional information to detect functional silencing of chromosomal regions. Bioinfor-matics 2003;19:2335–2342.

37 van Eijk R, Oosting J, Sieben N, et al. Visualization of regional gene expression biases by microarray data sorting. Biotechniques 2004;36:592–594, 596.

38 Harding MA, Arden KC, Gildea JW, et al. Functional genomic comparison of lineage-related human bladder cancer cell lines with differing tumorigenic and metastatic potentials by spectral karyotyping, com-parative genomic hybridization, and a novel method of positional expression profiling. Cancer Res 2002;62: 6981–6989.

39 Pollack JR, Sorlie T, Perou CM, et al. Microarray analysis reveals a major direct role of DNA copy number alteration in the transcriptional program of human breast tumors. Proc Natl Acad Sci USA 2002;99:12963–12968.

40 Furge KA, Dykema KJ, Ho C, et al. Comparison of array-based comparative genomic hybridization with gene expression-based regional expression biases to identify genetic abnormalities in hepatocellular carci-noma. BMC Genomics 2005;6:67.

41 Van Leuven F, Torrekens S, Moechars D, et al. Molecular cloning of a gene on chromosome 19q12 coding for a novel intracellular protein: analysis of expression in human and mouse tissues and in human tumor cells, particularly Reed–Sternberg

Modern Pathology (2006) 19, 854–863

cells in Hodgkin disease. Genomics 1998;54: 511–520.

42 Wei W, Gu JX, Zhu CQ, *et al.* Interaction with general transcription factor IIF (TFIIF) is required for the suppression of activated transcription by RPB5-mediating protein (RMP). Cell Res 2003;13: 111–120.

43 Dorjsuren D, Lin Y, Wei W, *et al.* RMP, a novel RNA polymerase II subunit 5-interacting protein, counteracts transactivation by hepatitis B virus X protein. Mol Cell Biol 1998;18:7546–7555.

44 Pennacchio LA, Bergmann A, Fukushima A, *et al.* Structure, sequence and location of the UQCRFS1 gene for the human Rieske Fe-S protein. Gene 1995;155: 207–211.

45 Ohashi Y, Kaneko SJ, Cupples TE, *et al.* Ubiquinol cytochrome c reductase (UQCRFS1) gene amplification in primary breast cancer core biopsy samples. Gynecol Oncol 2004;93:54–58.

46 Kaneko SJ, Gerasimova T, Smith ST, *et al.* CA125 and UQCRFS1 FISH studies of ovarian carcinoma. Gynecol Oncol 2003;90:29–36.

Supplementary Information accompanies the paper on the Modern Pathology website (http://www.nature.com/modpathol).