Am J Transl Res 2016;8(11):4994-5004
www.ajtr.org /ISSN:1943-8141/AJTR0036837

Original Article

Detection of gene copy number alterations in DCIS and invasive breast cancer by QM-FISH

Aifeng Pan[1], Yawei Zhou[2], Kun Mu[1], Yansong Liu[3], Feifei Sun[1], Peng Li[4], Li Li[1]

[1]Department of Pathology, Shandong University School of Medicine, 44 Wenhua Xi Road, Jinan 250012, Shandong, P. R. China; [2]Division of Hematology, Jinan Central Hospital, 105 Jiefang Road, Jinan 250013, Shandong, P. R. China; [3]Department of Breast Surgery, Shandong Cancer Hospital and Institute, 440 Jiyan Road, Jinan 250117, Shandong, P. R. China; [4]Department of Pathology, Shandong University Qilu Hospital, 107 Wenhua Xi Road, Jinan 250012, Shandong, P. R. China

Received July 29, 2016; Accepted October 16, 2016; Epub November 15, 2016; Published November 30, 2016

Abstract: The exact roles of copy number alteration (CNA) in initiation, progression and immunotherapy of breast cancer and the genomic alterations behind progression from ductal carcinoma in situ (DCIS) to invasive carcinoma remain unknown. Quantitative multi-gene fluorescence in situ hybridization (QM-FISH) opens a possibility of large scale genomic analysis of specific deletions and amplifications with high-resolution at one cell level. We detected CNAs of 30 genes using QM-FISH and analyzed their association with clinicopathological parameters and patients' outcomes in 66 breast cancers with synchronous invasive carcinoma and DCIS. The copy numbers of 30 genes in DCIS and the invasive area in all tumors were compared. The results revealed some recurrent CNAs including amplifications of MDMx, CCNE2, HER2 and deletions in Chek1, p53, Rb1 with a frequency of over 20%. By comparing the CNAs in invasive tumors and co-occurring DCIS, the similarity of chromosomal instability (CIN) in both components was visualized. Some co-occurrence patterns of CNAs of 30 genes were observed. The study also demonstrated higher frequencies of occurrence of CNAs in aneuploidy tumors, high grade tumors and tumors with high proliferation index. Higher CNAs were also found in death patients. Overall, we uncovered some frequently occurring gene aberrations out of 30 genes and synchronous pre-invasive lesions share majority of CNAs with invasive breast cancer. Moreover QM-FISH is a powerful technique to detect CNAs of multi-genes and give more information on co-occurrence of CNAs.

Keywords: Breast cancer, ductal carcinoma in situ, copy number alteration, quantitative multi-gene fluorescence in situ hybridization, chromosomal instability

Introduction

Breast cancer is a complex disease characterized by accumulation of genetic changes in a large number of genes, including point mutations, rearrangements, deletions, and amplifications. In the last 10 years many different multi-gene assays provide possibility in clinical practice to stratify patients for prognostic predictions and selecting treatments [1, 2]. However, the complexity of breast cancer is not fully reflected by the genes or biomarkers used currently for clinical purpose.

The genomic copy number alterations (CNAs) can alter the ability of the cancer cell to proliferate, survive, and spread in the host and therefore is now recognized as a characteristic of most breast cancers [3, 4]. The comprehensive characterization of CNAs in cancer genomes is critical for understanding breast carcinogenesis and for development of targeted therapies for individual cancer patients [5]. Curtis et al conducted an integrated analysis of copy number in a large numbers of patients and developed a novel molecular stratification of the breast cancer population [6]. It provides a definitive framework for understanding how gene copy number aberrations affect gene expression in breast cancer and reveals novel subgroups that should be the target of future investigation and tumor therapy. However there remains much more unknown about the roles of CNAs in initiation and progression of breast cancer. Quantitative multi-gene fluorescence in situ hybridization (QM-FISH) is a cytological

Table 1. CNAs of 30 genes in primary breast cancer (n=66)

Genes	N	Frequency (%)
Amplifications		
MDMx	36	54.5
PIK3CA	5	7.6
JARID2	2	3.0
FGFR1	10	15.2
CKS1a	19	28.8
CCNE2	23	34.8
c-myc	18	27.3
Cyclin D1	12	18.2
MDM2	2	3.03
IGF1R	16	24.2
HER2	15	22.7
TBX2	3	4.5
STK6	5	7.6
Deletions		
PAX7	7	10.6
FHIT	13	19.7
MAD4	3	4.5
MAP3K7	9	13.6
WTAP	4	6.1
LZTS1	5	7.6
p16	14	21.2
Chek1	27	41.0
LATS2	10	15.2
Rb1	25	37.9
Nek9	17	25.8
MAPK3	2	3.03
CDH1	23	34.5
p53	30	45.5
MADH4	10	15.2
TPTE	13	19.7
Chek2	25	37.9

technique that targets for cytological analysis and enables the detection of multiple chromosomal changes in single cells [7]. We developed the protocol involving in optimization of fixation and signal-washing methods for tumor samples in Anders Zetterberg Lab in CCK, Sweden. Up to 32 genes can be quantified synchronously in the same individual tumor cell nuclei, making QM-FISH a high-resolution, accurate and powerful tool for large scale clinical analysis. Herein, QM-FISH is the technique of choice to quantify copy number changes of genes or chromosomal regions in clinical samples. Meanwhile QM-FISH procedure gives the chance to analyze the relationship of CNAs of multi-genes at the single cell level and to avoid intra-tumor genomic heterogeneity to the largest extent.

Based on the ROMA data of breast cancer tumors obtained by Hick et al [8], 30 genes located at "hot-spots" were chosen for further analyses including some well-known genes and a few new genes. We detected gene CNAs in breast cancers by QM-FISH, and analyzed their association with clinicopathologic parameters and outcomes. We addressed the following questions: (1) to identify a panel of specific disease-associated CNA loci using QM-FISH, (2) to determine whether there are some features in the genomes of tumor cells that correlate with invasion by comparing CNAs of 30 genes in synchronous ductal carcinoma in situ (DCIS) and invasive cancer, (3) to better understand which and how CNAs might combine, and (4) to find some valuable markers for predicting prognosis of patients with breast cancer.

Results

CNAs of genes

All tumors had at least 1 chromosomal aberration out of 30 loci with the maximum of 12 and the average was 6.1. Totally 404 gene abnormalities were found with 166 amplifications and 238 deletions. Gene amplifications includes high level amplifications (cluster pattern) and low level amplifications (dot like pattern). As shown in **Table 1**, gain of MDMx was most frequently observed in this cohort of cases with a frequency of 54.5%. Other common CNAs were amplification of CCNE2, HER2, IGF1R, CKS1a, c-myc and deletion in Chek1, p53, Rb1, CDH1, Chek2, Nek9, all of which were observed in more than 20% of the analyzed cases. CKS1a and CCNE2 as well as MDMx gene showed low level increase of gene copy number.

CNAs in DCIS and invasive carcinoma

We successfully compared copy number of 30 genes in DCIS and the invasive area in all of the tumors. Majority of cases showed identical CNAs though differential CNAs of a few genes were observed in DCIS and invasive carcinoma. 341 events of gene abnormality were reported in DCIS, with average of 5.2 genes showing CNAs for each tumor. We identified differential

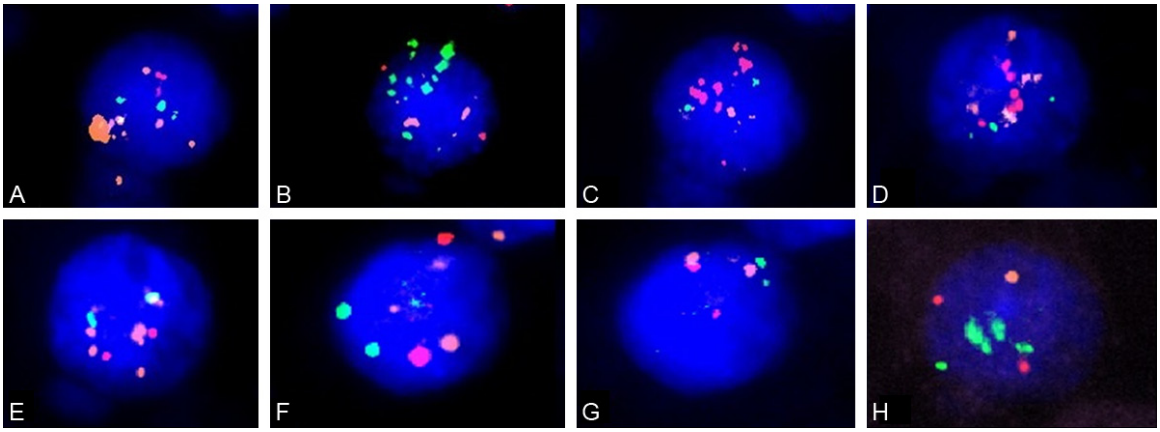

Figure 1. Case 18 showing MDMX, Cyclin D1, c-myc, MDM2, PIK3A, TBX2 amplification, and Chek1, LZST1, MADH4 deletion using QM-FISH. A. CDH1 (green), MDMX (orange), PAX7 (red) and MAPK3 (purple); B. Cyclin D1 (green), Chek1 (orange), Chek2 (red), FHIT (purple); C. HER2 (green), LZST1 (orange), c-myc (red), p53 (purple); D. FGFR1 (green), MAP3K7 (orange), MDM2 (red), JARID-2 (purple); E. CCNE2 (green), Rb1 (orange), p16 (red), PIK3A (purple); F. LATS2 (green), IGF1R (orange), MADH4 (red), TPTE (purple); G. CKS1a (green), MAD4 (red), WATP (purple); H. TBX2 (green), STK6 (orange), Nek9 (red).

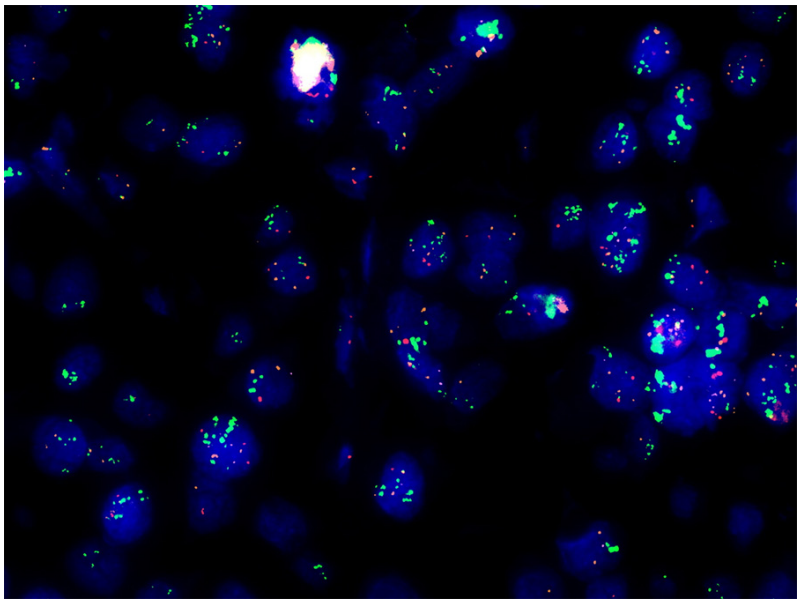

Figure 2. QM-FISH detection showing gene copy number changes of CCND1 and Chek1 (×630).

CNAs of Chek1 (34.8% vs. 41.0% in DCIS and invasive cancer respectively, P=0.591), CCNE2 (28.8% vs. 34.8%, P=0.564), Cyclin D1 (12.1% vs. 18.2%, P=0.467), Chek2 (34.8% vs. 37.9%, P=0.857), CDH1 (30.3% vs. 34.5%, P=0.711), TPTE (16.7% vs. 19.7%, P=0.822), MDMx (48.5% vs. 54.5%, P=0.711), LZST1 (4.6% vs. 7.6%, P=1), MAP3K7 (10.6% vs. 13.6%, P=0.791), PAX7 (9.1% vs. 10.6%, P=1), c-myc (25.8% vs. 27.3%, P=1), Rb1 (36.4% vs. 37. 9%, P=1), LATS2 (13.6% vs. 15.2%, P=1).

Concurrent of gene CNAs

Genetic abnormalities of 30 genes for an individual cell were analyzed. Some gene alterations occurred concurrently in the same cells were found (examples are shown in **Figure 1**). 9 out of 12 Cyclin D1 amplification positive cases showed concurrent Chek1 deletion (**Figure 2**). There was significant association between Cyclin D1 amplification and Chek1 deletion (P=0.011). C-myc amplification was identified in 16 out of 23 cases with CCNE2 amplification (P<0.001). All 5 tumors with LZST1 deletion had c-myc amplification (P=0.002). Other co-occurring CNAs included CKS1a and IGF1R amplification (in 9 cases, P=0.01), c-myc amplification and HER2 amplification (in 8 cases, P=0.025). Inverse correlation between deletion of CDH1 and p16 was found (P=0.001) as well.

Relationship between CNAs of genes and the clinicopathological characteristics

The correlation of CNAs with clinicopathological features was evaluated. Overall, there were

Table 2. Summary of patients and tumor characteristics and CNAs of 30 genes in studied cases with breast cancer

Characteristics	Patients no. (%)	Means of CNAs	*P* value
Age at diagnosis (y)			
≤35 ys	10	5.3	0.134
>35 ys	56	6.3	
Tumor size[1]			
d≤20 mm	22	5.5	0.874
>20 mm	44	6.4	
Grades			
Low	47	5.0	0.008
High	19	6.8	
Nodal Status[2]			
N0	26	5.7	0.745
N+	40	6.5	
ER			
Negative	22	6.2	0.582
Positive	44	6.1	
PR			
Negative	21	6.4	0.378
Positive	45	6.1	
HER2			
Negative	51	5.9	0.855
Positive	15	6.8	
Ki67			
Low	27	5.3	0.033
High	39	6.6	
Tumor Ploidy			
D	22	5.2	0.032
A	44	6.6	

[1]d, diameter; [2]N0, node metastasis negative; N+, node metastasis positive.

more CNAs in aneuploid tumor (A-tumor) than those in diploid tumor (D-tumor) (average 6.6 vs. 5.2, P=0.032). Tumors with high ki67 index had more CNAs than tumors with low ki67 (P=0.033). There were 5.0 CNAs in average in low grade tumors (Elson and Ellis grade 1 and 2) compared with 6.8 in high grade tumors (Elson and Ellis grade 3) (P=0.008) (**Table 2**; **Figure 3**). No significant association was seen between copy number changes and ER status, PR status, HER2, tumor size and node status.

Amplification of HER2, c-myc, CCNE2, CKS1a, Cyclin D1 were more frequent in A-tumor than that in D-tumor (P=0.001, P<0.001, P<0.001, P=0.019, P=0.049 respectively) while D-tumors has more deletion of PAX7 (P=0.036). The incidence of Nek9 deletion significantly increased with grade (P=0.027). Axillary lymph node metastasis positive tumors showed more IGF1R amplification (P=0.036) and MAPK3 and LATS2 deletion (P=0.009 and 0.005 respectively). Cyclin D1 amplification was present more often in ER and PR positive tumors (P=0.049 and 0.045) while more HER2 amplification and WTAP deletion were present in ER and PR negative tumors (P<0.001, P<0.001 and P=0.039, P=0.046). There were more CNAs of HER2, CCNE2 and IGF1R in high ki67 group than those in low ki67 group (P=0.017, P=0.034, P=0.009), more details seen in Supplementary Table 1.

Correlation between CNAs of the genes and prognosis of the patients

The survival analysis based on CNAs of 30 genes was investigated. There were less CNAs in survival group than in death group (5.3 vs 7.7, P<0.001). Survival analysis revealed that the gene deletions of p16 (P<0.001), CDH1 (P=0.004), Chek1 (P=0.025), TPTE (P=0.009), Nek9 (P=0.002), p53 (P=0.001) were significantly associated with a poor prognosis (**Figure 4**).

Discussion

The investigation of CNAs is of high relevance for improving the diagnosis of breast cancer and prognosis of the patients. During the last decades, modern techniques such as cytogenetic analysis, gene sequencing or array based analyses were developed to detect genomic alterations of tumors [9-13]. However, the problem of resolution and lack of commercial availability limited clinical applications of most genomic array technologies. So microarray studies are mostly used to screen genes aberration. CNAs identified by microarray always include a contiguous set of genes. Up to now FISH is still the rapid, most precise means to assess copy numbers of specific sequences in tumor samples due to its simplicity and reliability in evaluating the key biomarkers [7]. Currently, conventional FISH could detect 1 or 2 genes once a time on one section. Weak signals or losing one of the hybridization signals may occur accompanied with high background and consequently results in a relatively low signal-to-noise ratio when multi-genes are detect-

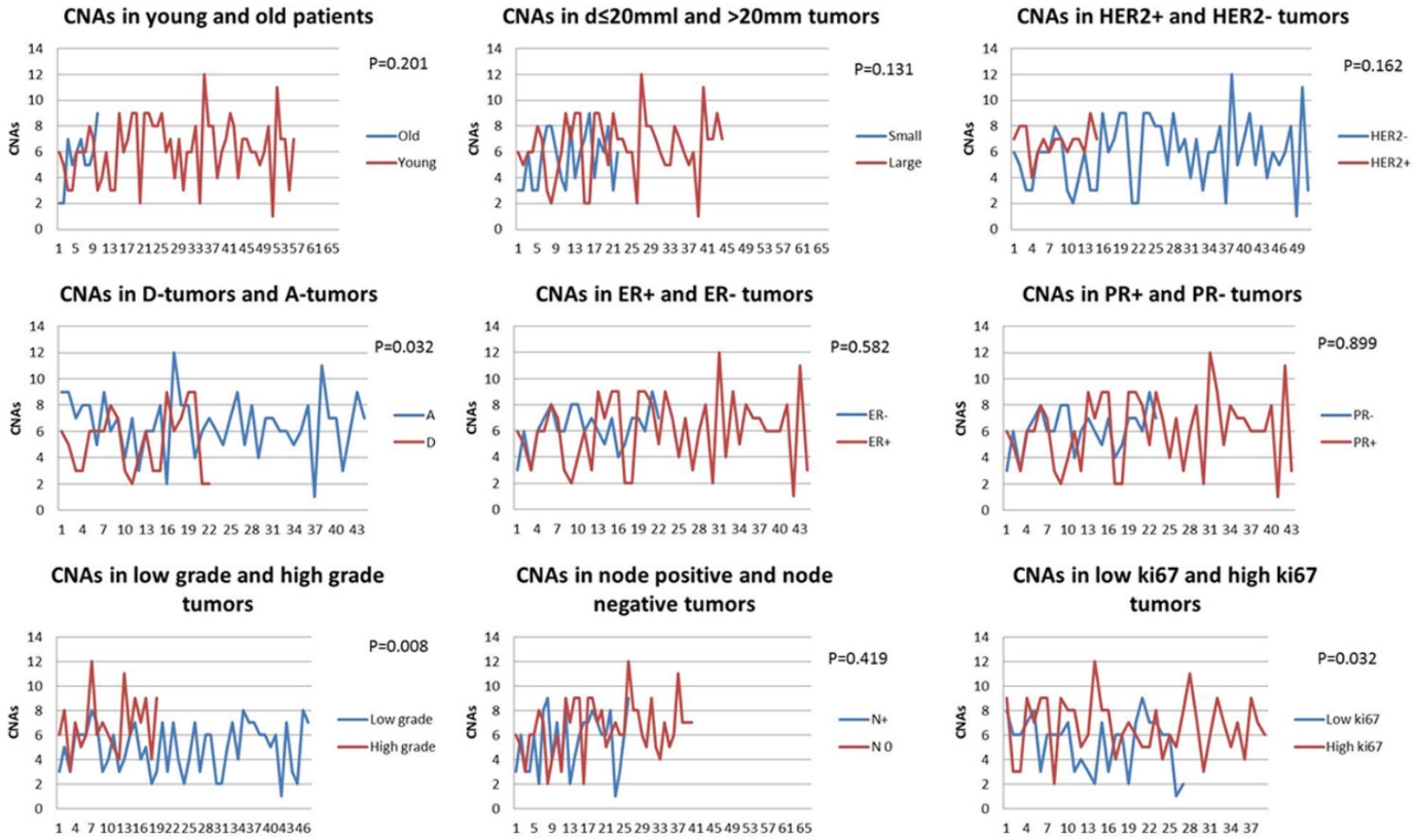

Figure 3. CNAs of 30 genes were more common in A-tumors, high grade tumors and tumors with high ki67 index (N0, node metastasis negative; N+, node metastasis positive).

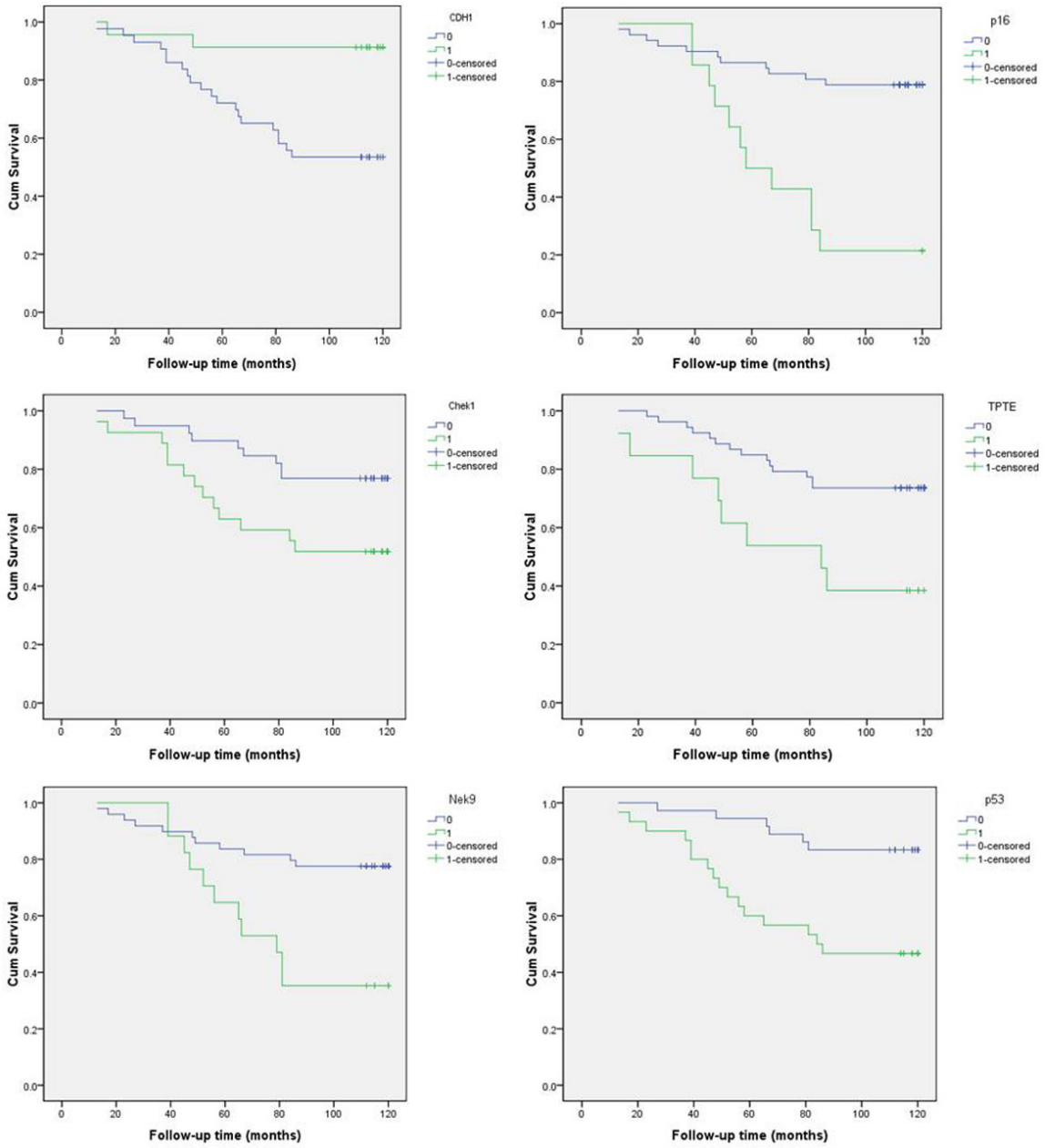

Figure 4. Kaplan-Meier analysis of overall survival and CNAs of genes.

ed. In current study we used three to five different BAC probes with partially overlapping DNA sequences for each gene to enhance resolution. Using QM-FISH copy number values of dozens of genes in individual cells can be measured and high signals/noise ratio with intact nuclear morphology is reached.

Although a flood of CNAs have been reported, they only cover a small proportion of the functional genes and tend to occur at distinct 'hotspots'. Therefore, as for a certain tumor, the leading couples of genetic changes could describe its genetic features. Our study was designed to detect CNAs of 30 genes based on previous gene array study. Some are well known genes (such as HER2 and c-myc) involved in breast cancer, however some are rather new (such as LZST1 and Nek9). To my knowledge there have not been such studies before. The data showed all cases had at least one chromosomal aberration out of 30 loci with an average

6 CNAs. It confirmed that CNA is a hallmark of malignant tumors and indicated it plays an important role in driving breast cancer. MDMx, CCNE2, HER2, IGF1R, CKS1a, c-myc, Chek1, p53, Rb1, CDH1, Chek2, Nek9 were frequently affected chromosomal region in breast cancers. We firstly found Nek9 deleted frequently in breast cancer.

DCIS is a non-obligate pre-invasive lesion for invasive breast cancer. Histologic and epidemiologic studies showed similarities between these two lesions. Several studies suggested that molecular genetic alterations in DCIS of mixed DCIS-invasive cancer strongly resemble that seen in invasive ductal carcinoma [14-16]. However, over the last decade few studies have specifically compared the CNAs between these two lesions, and more related genomic alterations remain unknown. To better understand the mechanisms behind progression to invasive breast cancer, we used QM-FISH to detect specific genetic alterations. We found the degree of genomic alteration of DCIS resembled that of invasive carcinoma, indicating that these alterations may be essential for the early phase of DCIS development. This also suggests that, in most cases, CNA occurs before the acquisition of an invasive phenotype and may not contribute to the development of invasion. Combined DCIS-invasive breast cancer and pure DCIS have been reported to be genetically distinct. Kim et al performed whole-exome sequencing and copy number profiling for six cases of pure DCIS and five pairs of synchronous DCIS and invasive ductal carcinoma and found genomic features of DCIS associated with invasive cancer were closer to invasive cancer than pure DCIS [17]. To get more information, step studies with comparing pure DCIS and invasive carcinoma and with expanded sample size are needed. This raises the possibility of using QM-FISH of a panel of such genes to predict disease progression.

It is well known that breast cancers develop through an accumulation of a variety of genetic alterations [18-20]. It is essential to carry out comprehensive CNAs profiling on tumor specimens to assess the role and impact of coincident CNAs. The study on concurrent genetic changes may provide insight into the underlying mechanisms that drive this divergent biological and clinical behaviour of breast cancer and facilitate the discovery of potential biomarkers with therapeutic value. A great advantage of QM-FISH is that it is relatively easy to be conducted in a clinical setting, thereby gives the chance to study the concurrent of various genes at the same single cell level, which can avoid intra-tumor heterogeneity. Significantly, we uncovered high frequency of co-existing genetic alterations. This is the first report for most of the concurrent gains and deletions in breast cancer by QM-FISH. Meanwhile it indicated a more complex gene expression/genomic variation interaction in breast cancer.

Our study demonstrated more frequent gene copy number gains/losses in A-tumors, and high grade tumors. Moreover, more copy number abnormality of a few genes was observed in A-tumor, tumors in larger size, node positive tumors and high grade tumors, confirming more aggressive behaviour of the tumors with more frequent CNAs. Our data also indicated higher CNAs in death patients. As correlated with low survival, the copy numbers of p16, CDH1, Chek1, TPTE, Nek9 and p53 might be used as prognosis markers for breast cancer.

In conclusion, this is the first report of detection of 30 genes using QM-FISH in the same tumor section. Our protocol of high resolution QM-FISH would be useful in detection of CNAs of multigenes and give the chance to get more information on co-occurrence of CNAs. Frequently affected gene aberrations of 30 genes includes amplification of MDMx, CCNE2, HER2, IGF1R, CKS1a, c-myc and deletion in Chek1, p53, Rb1, CDH1, Chek2, Nek9, with a frequency of over 20%. These genes can more precisely reveal the overall abnormalities present in breast cancer patients relative to a single gene. Detecting CNAs of a panel of above recurrent genes will assist accurate prognosis and facilitate individual therapeutic plans. In addition, synchronous DCIS and invasive breast cancer have similarities at genomic level.

Materials and methods

Patients and tumor characteristics

The study was approved by the Ethics Committee of Shandong University (no. MECSDUMS2012052). Sixty-six consecutively cases of non-specified invasive breast carcinoma associated with various percentage of DCIS were collected from 66 women undergoing surgery

Table 3. 30 genes labeled by 4 fluorophores

	Group 1	Group 2	Group 3	Group 4	Group 5	Group 6	Group 7	Group 8
Spectrum Green	CDH1	Cyclin D1	C-erbB2	FGFR1	CCNE2	LATS2	CKS1a	TBX2
Spectrum Orange	MDMX	Chek1	LZST1	MAP3K7	Rb1	IGF1R	MAD4	STK6
Texas Red	PAX7	Chek2	c-myc	MDM2	p16	MADH4	WATP	Nek9
Cy 5 (purple)	MAPK3	FHIT	p53	JARID-2	PIK3A	TPTE		

in Qilu Hospital, Shandong University, Jinan, China, from January to June 2005. None of the patients received either neoadjuvant chemotherapy or radiation therapy prior to operation. From each specimen ten contiguous sections were prepared and used for H&E staining and QM-FISH procedure.

The ages of patients at the time of the diagnosis ranged from 23 to 85 years (mean 46.3 years). The tumor size ranged from 1.0 to 7.8 cm in the greatest dimension (mean 2.9 cm). Detailed patient and disease characteristics are documented in **Table 1**. Patients were followed up from the date of surgery until death or the last observation (median follow-up, 71.4 months, ranging 13-120 months). At the time of the last follow-up, 44 patients (66.7%) were alive, 22 patients (33.3%) were dead from the disease.

Tumor ploidy evaluation

The ploidy of 66 tumors was evaluated by measurement of DNA content using image cytometry on Feulgen stained sections as previously described [21]. DNA histograms were interpreted according to a modified subjective method. The normal control cells were given the value 2c, denoting the normal diploid DNA content, and all tumor-cell DNA values were expressed in relation to that. The histograms were divided into two groups. Cases with a major peak near the 2c region (1.8c-2.2c), and less than 10% cells exceeding 2.5c were denoted D-tumor. DNA profiles with a stem line outside the diploid and tetraploid region and distinctly scattered DNA values exceeding the tetraploid region (3.8c-4.2c) were classified as A-tumor.

Preparation of probes

The UCSC (University of California, Santa Cruz) genome browser database (http://genome.ucsc.edu/) was used to design probes for all genes. The probes were created from bacterial artificial chromosomes (BAC) plasmid DNA (Deutsche Ressourcenzentrum für Genomforschung, RZPD, Berlin) isolated by using a Qiagen® plasmid purification kit. All probes were labeled with Spectrum Green-dUTP, Spectrum Orange-dUTP (Abbott Molecular Inc., Downers Grove, IL), Alexa Fluo 594-5-dUTP (Texas Red, Invitrogen) and Cy5-dUTP (purple, Amersham Life Science) with dUTP by nick translation. Three to five different BAC probes with partially overlapping DNA sequences were chosen for each gene to boost the hybridization signal and enhance resolution. For example, five continuous sequences p53-N1, p53-N2, p53-N, p53-N3, p53-N4 were used as probes for p53 gene. All probes tested on HDF cells in order to obtain probes with as high specificity as possible. 30 genes were divided into eight groups, which were summarized in **Table 3**. Labeled probes were mixed with human Cot1 DNA, tRNA and salmon sperm DNA to compete for repetitive elements. Sodium acetate and ice cold absolute ethanol were added in and the mixture was precipitated at -80°C overnight. Ethanol was removed by centrifugation and the pellet was washed once with 70% ethanol. The pellet was air dried and redissolved in hybridization buffer.

FISH procedure

Formalin-fixed paraffin-embedded tissue sections were removed excess wax followed by dehydration in absolute alcohol. Antigenic recovery was performed through incubating the slides for 1 hour at 80°C in 0.1 M citric buffer (pH 6.0). A 10-minute digestion with pepsin (1 mg/ml in 0.01 M HCL) was performed followed by fixation in 1% formaldehyde. The probes were dissolved in the hybridization mixture. Denaturation of probe and target DNA were performed simultaneously at 90°C for 10 min and each slide was incubated in a moist chamber for hybridization at 47°C overnight. After hybridization, slides were washed in 4×SSPE for 10 min at 37°C and 47°C respectively. Nuclei were mounted and counterstained with

DAPI (4', 6-diamino-2-phenylindole; Vector Laboratories) followed by view in the fluorescence microscope. After photographing, the slides were stored in the cold room for review and next hybridization.

Sequential FISH

The cover slips were removed for hybridization with next probe. The signals were removed by incubation of sections with 70% formamide at 70°C for 3 min, immediately immersed in precooling 70% ethanol and washed. Then, slides were dehydrated, denatured, hybridized and mounted as described above. After mounting for every round, the slides were placed in the microscope. Fields were chosen with filter for DAPI, and the position was recorded for image acquisition. One monochrome image with each of the filter sets for detection of DAPI, Spectrum Green, Spectrum Orange, Alexa 594, and Cy3 was acquired using ×630 objective lens before moving the focal plane to the next position. After re-hybridization the slides were reinserted in the microscope, they were repositioned using the automatic repositioning system of the Delta Vision system. So images were obtained from the same area.

Signal evaluation was carried out using an Axioplan 2 confocal fluorescence microscope (Carl Zeiss AB, Sweden) mounted with a charge-coupled device (CCD) camera Axiocam MRm (Carl Zeiss AB) , coupled with a computer with Axio Vision software (Carl Zeiss AB). Two researchers independently carried out all investigations without knowledge of the clinicopathological data of the patients studied. Only signals in the tumor areas based on both a consecutive section stained by H&E and DAPI morphology were counted and evaluated. At least 100 non-overlapping nuclei per sample were analyzed for each tumor. More than 50% of counted nuclei with 0-1 signal defined those samples containing gene deletion. Gene amplification was defined by the presence of an excess in the number of gene loci (except HER2) over the number of corresponding chromosomes in more than 20% counted tumor cells.

Immunostaining evaluation

The immunostained slides of ER, PR, HER2 and ki67 were reviewed and reevaluated. Tumors were counted as positive for ER and PR if >1% of the nuclei of neoplastic cells showed definitive staining [22]. Ki67 index was scored low if <14% of the nuclei of neoplastic cells were positive, and high if ≥14% of the nuclei of neoplastic cells were positive [23]. HER2 was scored according to ASCO/CAP HER2 clinical practice guideline [24].

Statistical analysis

All statistical analyses were performed using SPSS 16.0 (SPSS Inc. Chicago, IL). The Chi-square and Fisher exact test was used to evaluate the statistical significance between clinicopathological variables and CNAs of genes. Correlations were studied using the Spearman test. Mann-Whitney U test was used to identify the number of CNAs associated with clinicopathological parameters. Survival of patients was analyzed using the Kaplan-Meier and Log-rank test. Statistical significance was set at 0.05 and all *p* value is two-sided.

Acknowledgements

We thank Professor Anders Zetterberg (Cancer Centre, Karolinska Institute, Sweden) for his valuable probes, technical assistance and useful discussions. This study was supported by National Natural Science Foundation of China [81272902 and 30901456 to LL].

Disclosure of conflict of interest

None.

Address correspondence to: Dr. Li Li, Department of Pathology, Shandong University School of Medicine, 44# Wenhua Xi Road, Jinan 250012, Shandong, P. R. China. Fax: +86-0531-82169221; E-mail: lillie6636@sdu.edu.cn

References

[1] Paik S, Shak S, Tang G, Kim C, Baker J, Cronin M, Baehner FL, Walker MG, Watson D, Park T, Hiller W, Fisher ER, Wickerham DL, Bryant J and Wolmark N. A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer. N Engl J Med 2004; 351: 2817-2826.

[2] Nielsen TO, Parker JS, Leung S, Voduc D, Ebbert M, Vickery T, Davies SR, Snider J, Stijleman IJ, Reed J, Cheang MC, Mardis ER, Perou CM, Bernard PS and Ellis MJ. A comparison of PAM50 intrinsic subtyping with immunohistochemistry and clinical prognostic factors in tamoxifen-treated estrogen receptor-positive breast cancer. Clin Cancer Res 2010; 16: 5222-5232.

[3] DePinho RA and Polyak K. Cancer chromosomes in crisis. Nat Genet 2004; 36: 932-934.
[4] Balmain A, Gray J and Ponder B. The genetics and genomics of cancer. Nat Genet 2003; 33 Suppl: 238-244.
[5] Chen GK, Chang X, Curtis C and Wang K. Precise inference of copy number alterations in tumor samples from SNP arrays. Bioinformatics 2013; 29: 2964-2970.
[6] Curtis C, Shah SP, Chin SF, Turashvili G, Rueda OM, Dunning MJ, Speed D, Lynch AG, Samarajiwa S, Yuan Y, Graf S, Ha G, Haffari G, Bashashati A, Russell R, McKinney S, Langerod A, Green A, Provenzano E, Wishart G, Pinder S, Watson P, Markowetz F, Murphy L, Ellis I, Purushotham A, Borresen-Dale AL, Brenton JD, Tavare S, Caldas C and Aparicio S. The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups. Nature 2012; 486: 346-352.
[7] Hu L, Ru K, Zhang L, Huang Y, Zhu X, Liu H, Zetterberg A, Cheng T and Miao W. Fluorescence in situ hybridization (FISH): an increasingly demanded tool for biomarker research and personalized medicine. Biomark Res 2014; 2: 3.
[8] Hicks J, Krasnitz A, Lakshmi B, Navin NE, Riggs M, Leibu E, Esposito D, Alexander J, Troge J, Grubor V, Yoon S, Wigler M, Ye K, Borresen-Dale AL, Naume B, Schlicting E, Norton L, Hagerstrom T, Skoog L, Auer G, Maner S, Lundin P and Zetterberg A. Novel patterns of genome rearrangement and their association with survival in breast cancer. Genome Res 2006; 16: 1465-1479.
[9] Ho CC, Mun KS and Naidu R. SNP array technology: an array of hope in breast cancer research. Malays J Pathol 2013; 35: 33-43.
[10] Ueno T, Emi M, Sato H, Ito N, Muta M, Kuroi K and Toi M. Genome-wide copy number analysis in primary breast cancer. Expert Opin Ther Targets 2012; 16 Suppl 1: S31-35.
[11] Kaur H, Mao S, Shah S, Gorski DH, Krawetz SA, Sloane BF and Mattingly RR. Next-generation sequencing: a powerful tool for the discovery of molecular markers in breast ductal carcinoma in situ. Expert Rev Mol Diagn 2013; 13: 151-165.
[12] van der Vegt B, de Bock GH, Hollema H and Wesseling J. Microarray methods to identify factors determining breast cancer progression: potentials, limitations, and challenges. Crit Rev Oncol Hematol 2009; 70: 1-11.
[13] Climent J, Garcia JL, Mao JH, Arsuaga J and Perez-Losada J. Characterization of breast cancer by array comparative genomic hybridization. Biochem Cell Biol 2007; 85: 497-508.
[14] Mardekian SK, Bombonati A and Palazzo JP. Ductal carcinoma in situ of the breast: the importance of morphologic and molecular interactions. Hum Pathol 2016; 49: 114-123.
[15] Rane SU, Mirza H, Grigoriadis A and Pinder SE. Selection and evolution in the genomic landscape of copy number alterations in ductal carcinoma in situ (DCIS) and its progression to invasive carcinoma of ductal/no special type: a meta-analysis. Breast Cancer Res Treat 2015; 153: 101-121.
[16] Heaphy CM, Bisoffi M, Joste NE, Baumgartner KB, Baumgartner RN and Griffith JK. Genomic instability demonstrates similarity between DCIS and invasive carcinomas. Breast Cancer Res Treat 2009; 117: 17-24.
[17] Kim SY, Jung SH, Kim MS, Baek IP, Lee SH, Kim TM, Chung YJ and Lee SH. Genomic differences between pure ductal carcinoma in situ and synchronous ductal carcinoma in situ with invasive breast cancer. Oncotarget 2015; 6: 7597-7607.
[18] Simpson PT, Reis-Filho JS, Gale T and Lakhani SR. Molecular evolution of breast cancer. J Pathol 2005; 205: 248-254.
[19] Allred DC, Mohsin SK and Fuqua SA. Histological and biological evolution of human premalignant breast disease. Endocr Relat Cancer 2001; 8: 47-61.
[20] Lengauer C, Kinzler KW and Vogelstein B. Genetic instabilities in human cancers. Nature 1998; 396: 643-649.
[21] Blegen H, Einhorn N, Sjovall K, Roschke A, Ghadimi BM, McShane LM, Nilsson B, Shah K, Ried T and Auer G. Prognostic significance of cell cycle proteins and genomic instability in borderline, early and advanced stage ovarian carcinomas. Int J Gynecol Cancer 2000; 10: 477-487.
[22] Hammond ME, Hayes DF, Wolff AC, Mangu PB and Temin S. American society of clinical oncology/college of american pathologists guideline recommendations for immunohistochemical testing of estrogen and progesterone receptors in breast cancer. J Oncol Pract 2010; 6: 195-197.
[23] Goldhirsch A, Winer EP, Coates AS, Gelber RD, Piccart-Gebhart M, Thurlimann B and Senn HJ. Personalizing the treatment of women with early breast cancer: highlights of the St Gallen International Expert Consensus on the Primary Therapy of Early Breast Cancer 2013. Ann Oncol 2013; 24: 2206-2223.
[24] Wolff AC, Hammond ME, Hicks DG, Dowsett M, McShane LM, Allison KH, Allred DC, Bartlett JM, Bilous M, Fitzgibbons P, Hanna W, Jenkins RB, Mangu PB, Paik S, Perez EA, Press MF,

Spears PA, Vance GH, Viale G, Hayes DF; American Society of Clinical Oncology; College of American Pathologists. Recommendations for human epidermal growth factor receptor 2 testing in breast cancer: American Society of Clinical Oncology/College of American Pathologists clinical practice guideline update. Arch Pathol Lab Med 2014; 138: 241-256.

Supplementary Table 1. Relation between genes number alterations and clinicopathological characteristics

Characteristics	Patients no.	CDH1	*p* value	MDMX	*p* value	PAX7	*p* value	MAPK3	*p* value	cyclinD	*p* value	Chek1	*p* value
Age at diagnosis (y)													
≤35 ys	10	4	0.730	5	1.000	0	0.583	0	1.000	0	0.187	4	1.000
35 ys	56	19		31		7		2		12		23	
Tumor size[1]													
d≤20 mm	22	8	1.000	15	0.189	0	0.086	0	0.549	6	0.193	8	0.791
>20 mm	44	15		21		7		2		6		19	
Grades													
Low	47	17	0.783	29	0.101	7	0.179	2	1.000	10	0.484	21	0.412
High	19	6		7		0		0		2		6	
Nodal Status[2]													
N0	26	14	0.016**	17	0.208	2	0.695	0	0.515	6	0.517	9	0. 451
N+	40	9		19		5		2		6		18	
ER													
Negative	22	4	0.057	11	0.612	0	0.086	0	0.549	1	0.049*	8	0.791
Positive	44	19		25		7		2		11		19	
PR													
Negative	23	4	0.034*	12	0.801	0	0.086	0	0.539	1	0.045*	8	0.601
Positive	43	19		24		7		2		11		19	
Ki67													
Low	27	12	0.198	15	1.000	5	0.113	1	1.000	5	1.000	10	0.621
High	39	11		21		2		1		7		17	
Tumor Ploidy													
D	22	10	0.274	11	0.612	5	0.036**	2	0.108	1	0.049*	11	0.304
A	44	13		25		2		0		11		16	

Characteristics	Patients no. (%)	FHIT	*p* value	Chek2	*p* value	IGF1R	*p* value	LZST1	*p* value	c-myc	*p* value	P53	*p* value
Age at diagnosis (y)													
≤35 ys	10	0	0.190	1	0.076	2	1.000	0	1.000	2	0.715	6	0.492
>35 ys	56	13		24		14		5		16		24	
Tumor size[1]													
d≤20 mm	22	4	1.000	7	0.593	6	0.764	0	0.160	4	0.380	12	0.310
>20 mm	44	9		18		10		5		14		18	
Grades													
Low	47	10	0.742	18	1.000	10	0.526	4	1.000	14	0.554	19	0.276
High	19	3		7		6		1		4		11	
Nodal Status[2]													
N0	26	4	0.543	6	0.069	2	0.036*	2	1.000	6	0.585	15	0.133
N+	40	9		19		14		3		12		15	
ER													
Negative	22	4	1.000	11	0.184	7	0.367	0	0.160	7	0.571	10	1.000
Positive	44	9		14		9		5		11		20	
PR													
Negative	23	4	1.000	11	0.289	8	0.227	0	0.154	7	0.774	10	1.000
Positive	43	9		14		8		5		11		20	
Ki67													
Low	27	3	0.211	8	0.307	2	0.009*	1	0.641	5	0.263	12	1.000
High	39	10		17		14		4		13		18	
Tumor Ploidy													
D	22	4	1.000	10	0.426	2	0.066	0	0.160	0	0.000*	8	0.432
A	44	9		15		14		5		18		22	

Characteristics	Patients no. (%)	CCNE2	*p* value	Rb1	*p* value	P16	*p* value	PIK3A	*p* value	FGFR1	*p* value	MAPK3	*p* value
Age at diagnosis (y)													
≤35 ys	10	2	0.474	4	1.000	3	0.431	0	1.000	1	1.000	2	0.616
>35 ys	56	21		21		11		5		10		7	

Tumor size[1]													
d≤20 mm	22	6	0.421	8	1.000	5	1.000	2	1.000	4	1.000	2	0.706
>20 mm	44	17		17		9		3		7		7	
Grades													
Low	47	17	0.783	16	0.403	9	0.523	4	1.000	8	1.000	9	0.050
High	19	6		9		5		1		3		0	
Nodal Status[2]													
N0	26	6	0.122	10	1.000	4	0.539	1	0.641	3	0.505	0	0.009*
N+	40	17		15		10		4		8		9	
ER													
Negative	22	11	0.100	6	0.284	6	0.524	2	1.000	2	0.312	2	0.706
Positive	44	12		19		8		3		9		7	
PR													
Negative	23	11	0.174	6	0.188	6	0.536	2	1.000	2	0.304	2	0.478
Positive	43	12		19		8		3		9		7	
Ki67													
Low	27	5	0.034*	11	0.798	5	0.765	2	1.000	6	0.336	4	1.000
High	39	18		14		9		3		5		5	
Tumor Ploidy													
D	22	0	0.000*	10	0.426	4	0.759	2	1.000	2	0.312	4	0.467
A	44	23		15		10		3		9		5	

Characteristics	Patients no. (%)	MDM2	*p* value	JA-RID-2	*p* value	LATS2	*p* value	MADH4	*p* value	TPTE	*p* value	CKS1a	*p* value
Age at diagnosis (y)													
≤35 ys	10	0	1.000	0	1.000	0	0.338	3	0.169	2	1.000	4	0.456
>35 ys	56	2		2		10		7		11		15	
Tumor size[1]													
d≤20 mm	22	0	0.549	0	0.549	2	0.476	0	0.024*	4	1.000	5	0.568
>20 mm	44	2		2		8		10		9		14	
Grades													
Low	47	2	1.000	2	1.000	5	0.136	5	0.136	10	0.742	15	0.550
High	19	0		0		5		5		3		4	
Nodal Status[2]													
N0	26	0	0.515	0	0.515	0	0.005	2	0.293	4	0.543	9	0.419
N+	40	2		2		10		8		9		10	
ER													
Negative	22	0	0.549	0	0.549	2	0.476	6	0.072	6	0.331	7	0.776
Positive	44	2		2		8		4		7		12	
PR													
Negative	23	0	0.539	0	0.539	2	0.474	6	0.085	6	0.351	8	0.569
Positive	43	2		2		8		4		7		11	
Ki67													
Low	27	1	1.000	1	1.000	2	0.180	3	0.508	7	0.353	8	1.000
High	39	1		1		8		7		6		11	
Tumor Ploidy													
D	22	1	1.000	2	0.108	5	0.281	3	1.000	4	1.000	2	0.019*
A	44	1		0		5		7		9		17	

Characteristics	Patients no. (%)	MAD4	*p* value	WATP	*p* value	TBX2	*p* value	STK6	*p* value	NEK9	*p* value	CerbB2	*p* value
Age at diagnosis (y)													
≤35 ys	10	0	1.000	2	0.162	0	1.000	0	1.000	2	1.000	5	0.040**
>35 ys	56	3		3		3		4		15		10	
Tumor size[1]													
d≤20 mm	22	0	0.545	1	0.658	1	1.000	2	0.596	4	0.384	5	1.000
>20 mm	44	3		4		2		2		13		10	
Grades													
Low	47	2	1.000	3	0.621	3	0.551	4	0.316	8	0.027*	8	0.108
High	19	1		2		0		0		9		7	

Nodal Status[2]												
N0	26	0	0.273	2	1.000	1	1.000	2	0.644	4	0.155	7	0.558
N+	40	3		3		2		2		13		8	
ER													
Negative	22	1	1.000	4	0.039**	1	1.000	0	0.292	7	0.552	11	0.000**
Positive	44	2		1		2		4		10		4	
PR													
Negative	23	1	1.000	4	0.046**	1	1.000	0	0.289	7	0.564	12	0.000**
Positive	43	2		1		2		4		10		3	
Ki67													
Low	27	1	1.000	1	0.641	1	1.000	3	0.297	5	0.391	2	0.017*
High	39	2		4		2		1		12		13	
Tumor Ploidy													
D	22	2	0.256	0	0.160	2	0.256	2	0.596	6	1.000	0	0.001*
A	44	1		5		1		2		11		15	

[1]d, diameter; [2]N0, node metastasis negative; N+, node metastasis positive. *significant correlation; **significant negative correlation.