Journal of Pathology
J Pathol 2011; **224:** 517–528
Published online in Wiley Online Library
(wileyonlinelibrary.com) DOI: 10.1002/path.2938

ORIGINAL PAPER

Influence of whole arm loss of chromosome 16q on gene expression patterns in oestrogen receptor-positive, invasive breast cancer

Daniela Hungermann,[1#] Hartmut Schmidt,[2#] Rachel Natrajan,[4] Nicola Tidow,[2] Kathrin Poos,[3] Jorge S Reis-Filho,[4] Burkhard Brandt,[5] Horst Buerger[6] and Eberhard Korsching[3]*

[1] *Institute of Pathology, University of Münster, Germany*
[2] *Institute of Clinical Chemistry, University of Münster, Germany*
[3] *Institute of Bioinformatics, University Medical Centre, University of Münster, Germany*
[4] *The Breakthrough Breast Cancer Research Centre, London, UK*
[5] *Institute of Tumour Biology, University of Hamburg, Germany*
[6] *Institute of Pathology, Paderborn, and Cooperative Breast Centre Paderborn/Höxter, Germany, and Institute of Pathology, Utrecht University, The Netherlands*

*Correspondence to: Eberhard Korsching, Institute of Bioinformatics, Niels-Stensen-Strasse 12-14, 48149 Münster, Germany.
e-mail: korschi@uni-muenster.de

[#]These authors contributed equally to this study.

Abstract

A whole chromosome arm loss of 16q belongs to the most frequent and earliest chromosomal alterations in invasive and *in situ* breast cancers of all common subtypes. Besides E-cadherin, several putative tumour suppressor genes residing on 16q in breast cancer have been investigated. However, the significance of these findings has remained unclear. Thus, other mechanisms leading to gene loss of function (eg haploinsufficiency, or distortion of multiple regulative subnetworks) remain to be tested as a hypothesis. To define the effect on gene expression of whole-arm loss of chromosome 16q in invasive breast cancer, we performed global gene expression analysis on a series of 18 genetically extensively characterized invasive ductal breast carcinomas and verified the results by quantitative real-time PCR (qRT–PCR). The distribution of the differential genes across the genome and their expression status was studied. A second approach by qRT–PCR in an independent series of 30 breast carcinomas helped to narrow down the observed effect. Whole-arm chromosome 16q losses, irrespective of other chromosomal changes, are associated with decreased expression of a number of candidate genes located on 16q (eg *CDA08, CGI-128, SNTB2, NQO1, SF3B3, KIAA0174, ATBF1, GABARAPL2, KARS, GCSH, MBTPS1* and *ZDHHC7*) in breast carcinomas with a low degree of genetic instability. qRT–PCR provided evidence to suggest that the expression of these genes was reduced in a gene dosage-dependent manner. The differential expression of the candidate genes according to the chromosomal 16q-status vanished in genetically advanced breast cancer cases and changed ER status. These results corroborate previous reports about the importance of whole-arm loss of chromosome 16q in breast carcinogenesis and give evidence for the first time that haploinsufficiency, in the sense of a gene dosage effect, might be an important contributing factor in the early steps of breast carcinogenesis.

Keywords: breast cancer; chromosome 16q; global gene expression

Received 6 January 2011; Revised 15 April 2011; Accepted 9 May 2011

No conflicts of interest were declared.

Introduction

Physical losses of the whole arm of 16q belong to the most frequent and earliest findings in *in situ* and invasive breast cancer cases of the ductal and lobular subtypes [1]. Loss of chromosome 16q has been repeatedly associated with favourable clinical features, indicating a good prognosis in invasive breast cancer [2–7]. Based on the distribution of chromosome 16q losses within breast cancers cases of varying tumour grade, breast cancer models with multiple parallel pathways have been proposed [8–11]. Whereas well-differentiated breast cancer cases displayed whole 16q losses in a frequency of >60% by comparative genomic hybridization (CGH), this was a rather rare finding overall in poorly-differentiated breast cancers. These differences, however, could not be shown by means of microsatellite analysis with similar frequencies of loss of heterozygosity (LOH) in breast cancers of all histological grades [12].

It has become apparent that the underlying mechanisms of 16q LOH differed substantially in well and poorly-differentiated breast cancers. Physical losses of

the whole arm of 16q were the predominant mechanisms in well-differentiated breast cancers. In contrast, 16q losses in poorly-differentiated breast cancer were accompanied by mitotic recombination [13]. In conclusion, the results allowed the postulation of multiple tumour suppressor genes which might be the target of LOH at chromosome arm 16q. However, the recent research of putative tumour suppressor genes in the sense of the Knudson two-hit model on 16q have not taken into account these findings and was rather focused on tumour suppressor genes in poorly differentiated breast cancers [14–17]. Tumour suppressor genes residing on 16q might therefore differ in well and poorly differentiated breast cancers. We therefore hypothesized that a gene dosage effect as a consequence of the loss of the whole 16q arm may contribute to differential gene expression of a constellation of genes through haploinsufficiency.

Based on our previous results, we performed an integrative analysis of high-throughput microarray-based CGH and global gene expression profiling of 18 invasive breast cancer cases. The results were verified in a second, independent cohort of 30 additional breast cancers, using quantitative real-time PCR (qRT–PCR).

Materials and methods

Ethics

Data were analysed anonymously. Nonetheless, we have performed according to the principles expressed in the Declaration of Helsinki. The study was approved by the Institutional Review Board of the University of Münster. We have acquired tissue samples only with the informed consent of the patients or patients' next of kin, with the understanding by all parties that it may well be used for research. All patients provided written informed consent for the collection of samples and subsequent analysis.

CGH and array CGH

Based on our hypothesis, we have chosen a series of 18 ductal invasive breast carcinomas out of a series of 144 breast cancer cases previously described [18]. To reduce the influence of putative secondary cytogenetic changes other than 16q losses, we have predominantly chosen invasive breast carcinomas with a low number of cytogenetic alterations. All tumours were ER-positive (ER^+) breast cancers with a low to medium degree of cytogenetic instability. Further details are shown in Table 1.

All cases were fully characterized by means of CGH and array CGH, using DNA from fresh frozen tumour samples. Array CGH was performed with the 32 K BAC re-array collection (CHORI) tiling path aCGH platform, which was constructed at the Breakthrough Breast Cancer Research Centre, as described previously [19]. This type of BAC array platform has been shown to be as robust as, and to have comparable resolution to, high-density oligonucleotide arrays [20–22]. DNA labelling, array hybridizations and image acquisition were performed as previously described [23]. aCGH data were preprocessed and analysed using R script (version 2.9.0; http://www.r-project.org/), as previously described [24,25]. After filtering polymorphic BACs, a final dataset with unambiguous mapping information according to the build hg19 of the human genome

Table 1A. Overview of the basic set of ductal invasive grade 1–3 carcinomas used for the primary analysis

Case no.	c1	c2	Grade	ER	PR	Number of alterations	T-status	N-status	Ploidy status	Cytogenetic alteration status
2468			2	+	+	5	1	1	di	$1q^+$, $7q^+$, $^+8$, $16q^-$, -21
2517			2	+	+	2	1	0	tetra	$1q^+$, $16q^-$
2704			2	+	+	3	2	0	an	$1q^+$, $+8$, $16q^-$
3011			2	+	+	5	1	0	di	$2q^+$, $4q^+$, $9q^-$, $16q^-$, $17p^-$
2974			1	+	+	2	1	0	NA	$1q^+$, $16q^-$
2978			1	+	–	4	1	0	NA	$1q^+$, $16q^-$, -17, -21
181			2	+	+	7	2	1	an	$1q^+$, $8p^-$, $14q^+$, $15q^-$, $16q^-$, $17q^+$, -22
2979			2	+	+	6	1	0	di	-6, $8p^-$, $8q^+$, $13q^-$, $16q^-$, -21
3040			2	+	+	9	4	1	di	$1q^+$, $3q^-$, $5p^+$, $9p^-$, $16q^-$, $17q^+$, $18q^-$, $20q^+$, -22
185			2	+	–	5	2	2	di	$1q^+$, $17q^+$, $18q^-$, $20q^+$, -22
2472			2	+	+	2	2	0	di	$8p^-$, $8q^+$
2700			2	+	+	4	1	0	di	$6q^-$, $11q^+$, $+12$, -22
3050			2	+	+	4	2	1	NA	$7q^-$, $8q^+$, $17p^-$, -22
184			3	+	–	5	2	0	an	$1q^+$, $3q^+$, $6p^+$, $8q^+$, $10q^-$
2516			3	+	+	6	4	2	an	$1q^+$, $+6$, $+7$, $9q^-$, -11, $20q^+$
2528			3	+	–	2	2	2	di	$1q^+$, $8q^+$
2529			3	+	–	2	1	0	di	$1q^+$, $6q^-$
3041			3	+	–	1	2	2	an	$9q^+$

Columns 2 and 3 give an overview of the selections used for the differential gene expression analysis of 16q minus versus 16q normal cases: c1, a selection of grade 2 cases with a low number of alterations; c2, the second differential gene expression analysis is based on a mixed grouping comprising cases with both a low and a higher number of alterations. Clinical parameters of the cases and the cytogenetic alteration status are described in the following columns. Ploidy status: di, diploid; an, aneuploid; tetra, tetraploid; NA, not available. Detailed information on these cases has already been published [18]. Black shading, $16q^-$ cases; grey shading, 16q normal cases.

Table 1B. Overview of the ductal invasive grade 1–3 carcinomas used to study the impact of cases with a high number of alterations and ER-negative status on the outcome of the 16q–effect

Case no.	c3	Grade	ER	PR	Number of alterations	T-status	N-status	Ploidy status	Cytogenetic alteration status
1921		3	+	+	14	2	1	NA	4p , $4q^+$, $7p^+$, $8p^-$, $8q^+$, 11 q^-, 11q13$^+$, $13q^-$, $15q^+$, $16q^-$, $17q^+$, $20q^+$, – 22, Xq^+
2048		2	+	+	11	1	2	NA	$1q^+$, + 2, + 3, + 5, $9p^-$, 11 q^-, 11q13$^+$, $13q^-$, $16q^-$, $17q^+$, $18q^-$
2053		2	+	+	2	2	1	NA	$1q^+$, $16q^-$
2066		2	+	+	3	4	0	NA	$1q^+$, + 5, $16q^-$
2068		2	+	+	5	4	x	NA	$1q^+$, $8p^-$, $8q^+$, $15q^-$, $16q^-$
2071		1	+	+	4	1	1	NA	$1q^+$, $7p^+$, $7q^-$, $16q^-$
2072		1	+	+	3	1	1	NA	$12q^+$, $12q^-$, $16q^-$
2519		3	+	+	9	1	0	an	$1q^+$, $5q^-$, $7p^+$, $8q^+$, $9q^-$, 11 q^-, $^+$11q13, $15q^-$, $16q^-$
2523		3	+	+	15	2	1	NA	$1q^+$, $2q^-$, $3p^-$, + 4, + 5, + 6, $8p^-$, $8q^+$, $^-$9, 11 q^-, $^+$12, $13q^-$, $16q^-$, $17p^-$, $17q^+$
2698		2	+	+	9	1	0	NA	$1q^+$, + 5, + 8, 11 q^-, $16q^-$, $17p^-$, + 18, $20q^+$, – 22
2703		2	+	+	12	2	0	an	$1q^+$, + 5, – 6, – 7, $9q^-$, 11 p^+, 11 q^-, $13q^+$, $15q^+$, $16q^-$, 20q, – 22
2973		2	+	+	16	1	0	an	$1q^+$, $3p^-$, + 4, + 5, – 6, $7p^+$, $8p^-$, $8q^+$, $9q^-$, – 11, $13q^-$, $15q^-$, $16q^-$, $17p^-$, $17q^+$, + 18
3015		2	+	+	4	4	1	NA	$1q^+$, $7p^-$, $8p^-$, $16q^-$
3017		2	+	+	9	2	0	NA	$1q^+$, $3p^-$, $3q^+$, $8p^-$, $9p^-$, – 12, $16q^-$, – 17, – 22
3036		3	+	+	25	1	1	an	$2q^+$, + 3, $4q^-$, $5p^+$, $5q^-$, + 7, $8p^-$, $8q^+$, – 9, + 10, 11 q^-, + 12, $13q^-$, $14q^+$, $14q^-$, $15q^-$, $16q^-$, $17p^-$, $17q^+$, $18q^+$, $19q^+$, $20q^+$, + 21, – 22, Xq^-
3042		3	+	+	10	4	2	NA	$1q^+$, $3p^-$, $3q^+$, $6p^-$, $6q^+$, $6q^-$, $16q^-$, $17p^-$, $17q^+$, – 22
1915		3	–	–	1	3	1	NA	$1q^+$
2467		2	–	–	8	3	2	an	$1q^+$, $2q^-$, $7p^+$, $7q^-$, $9q^-$, $10q^+$, $17p^-$, $17q^+$
2471		2	–	–	4	1	1	NA	$1q^+$, + 3, 6q , – 21
2473		2	+	+	5	2	1	di	$1q^+$, $8p^-$, $^+$11q13, $13q^-$, $20q^+$
2511		3	–	–	12	2	0	di	$1q^+$, $4p^-$, $5p^+$, $5q^-$, $6q^+$, $8p^-$, $8q^+$, 10q , 13q , $17q^+$, 18q , Xq^+
2521		2	–	–	11	2	1	an	$1q^+$, + 5, $8q^+$, 11 q^-, $^+$11q13, $13q^-$, $14q^-$, $15q^-$, $18q^-$, $20q^+$, – 22
2526		3	–	–	12	4	2	di	$1q^+$, $2q^-$, $3q^+$, $6q^-$, $7q^-$, $8p^-$, $8q^+$, $134q^-$, $14q^-$, $17q^+$, $20q^+$, Xq^-
2697		3	–	–	15	2	0	tetra	$1q^+$, $2p^-$, + 3, – 4, – 5, $6p^+$, 6q , $8q^+$, $9p^+$, $10p^+$, $13q^+$, $15q^-$, – 17, + 18, – 22
2968		3	–	–	11	1	0	an	$1q^+$, $2p^+$, $3q^+$, $5p^+$, + 6, $7q^+$, $8p^-$, $9p^+$, $13q^-$, $20q^+$, Xq^-
2969		3	–	–	2	2	0	di	$8p^-$, $8q^+$
2980		2	–	–	14	2	1	an	$1q^+$, $3p^-$, $5p^+$, $5q^-$, $6q^+$, + 7, 11 p^-, $12p^+$, $13q^+$, $14q^-$, $15q^-$, $17p^-$, $18q^-$, Xq^+
3038		3	+	+	15	1	0	an	$1q^+$, $2p^+$, $4q^-$, $5p^+$, $6q^+$, $7p^+$, $8p^-$, $8q^+$, $9p^+$, $13q^+$, $13q^-$, $14q^+$, $15q^-$, $20q^+$, – 22
3041		3	–	–	1	2	2	an	$8q^+$
3044		3	–	–	24	4	1	an	$1q^+$, $2p^+$, $2q^-$, $3p^-$, $3q^+$, – 4, $5p^+$, $5q^-$, + 6, $7p^-$, $7q^+$, $8q^+$, + 9, $10p^+$, $10q^-$, $10q^+$, – 11, $^+$11q13, $12p^+$, $12q^-$, $13q^-$, $14q^-$, $18q^-$, Xq^-

Column 2 gives an overview of the selection used for the qRT–PCR study (c3). Clinical parameters of the cases and the cytogenetic alteration status are described in the following columns. Ploidy status: di, diploid; an, aneuploid; tetra, tetraploid; NA, not available. Black shading, $16q^-$ cases; grey shading, 16q normal cases.

(http://www.ensembl.org) was smoothed using the circular binary segmentation (cbs) algorithm [26]. A categorical analysis was applied to the BACs after classifying them as representing amplification (>0.40), gain (>0.08 and ≤0.4), loss (≤0.08) or no change, according to their cbs-smoothed log2 ratio values. Threshold values were determined and validated as previously described [11,27]. The DNA ploidy status was available for 15 cases. The thresholds in conventional CGH for the definition of losses and gains are described elsewhere [11,27–31]. For the definition of whole chromosome arm losses within array CGH, at least 75% of all probes mapping in the 16q region had to be a smoothed signal ratio ≤0.08.

Microarray-based gene expression profiling

RNA was isolated from fresh frozen tumour material (tumour sections), using the ABI PRISM 6100 Nucleic Acid PrepStation in combination with the Tissue RNA Isolation System (both Applied Biosystems, Foster City, CA, USA), as previously described [32]. The tumour sections contained at least 70% of tumour cells. Two to five (depending on the tumour size) 10 μm sections were lysed in nucleic acid purification lysis solution and digested with 10 U proteinase K (both part of the Tissue RNA Isolation System) per ml lysis buffer for 1–2 h at room temperature. Purification of RNA from the lysate was done according to the manufacturer's instructions. Purified RNA

was quantified spectrophotometrically (Biophotometer, Eppendorf, Hamburg, Germany).

Whole-genome microarray expression analysis was performed at GeneSys Laboratories GmbH (Münster, Germany), using kits, reagents and the chemiluminescent microarray analyser 1700 from Applied Biosystems according to the manufacturer's protocols. In brief, 2 μg high-quality total RNA was reverse-transcribed for 2 h at 42 °C in a thermocycler [GeneAmp 9700; Applied Biosystems (now Life Technologies Corp.), Carlsbad, CA, USA], using a T7-oligo (d)T primer and the reverse transcription *in vitro* transcription (RT–IVT) labelling kit. Synthesis of the second strand was performed at 16 °C for 2 h and cDNA was purified using the RT–IVT purification components. During IVT labelling for 9 h at 37 °C, digoxigenin-labelled UTP (Roche Diagnostics, Mannheim, Germany) was incorporated into the cRNA. cRNA was purified using the RT–IVT purification components and the quantity was assessed by measuring the absorbance at 260 and 320 nm. Labelled cRNA (10 μg) was fragmented for 30 min at 60 °C in fragmentation buffer and subsequent addition of fragmentation stop buffer was used to end the reaction.

Human Genome Survey Microarrays (Applied Biosystems) were prehybridized at 55 °C for 1 h (hybridization oven Minitron; Infors AG, Einsbach, Germany). Fragmented labelled cRNA was added to the hybridization mixture and rapidly transferred into each microarray cartridge. During hybridization for 16 h at 55 °C, the microarrays were agitated at 100 rpm.

After hybridization, the microarrays were removed from the cartridge and hybridization washes were carried out in a wash tray on a rocking platform shaker (VWR Rocking Platform Shaker, Darmstadt, Germany) in wash buffer 1 for 5 min, wash buffer 2 for 5 min, and twice in chemiluminescent rinse buffer for 5 min. Binding of anti-digoxigenin-AP, Fab fragments (Roche Diagnostics GmbH, Mannheim, Germany) was performed for 20 min on a rocking platform shaker. Remaining antibody was washed away with three washes of chemiluminescent rinse buffer for 10 min each. During the chemiluminescent reaction, the microarrays were washed with chemiluminescent enhancing rinse buffer for 10 min and incubated with chemiluminescent enhancing solution for 20 min. After an additional wash in chemiluminescent enhancing rinse buffer the chemiluminescent substrate was added onto the microarray. Chemiluminescence was detected using a 1700 chemiluminescent microarray analyser (Applied Biosystems).

Relative expression quantification of differentially expressed genes by real time qRT–PCR

Expression of 10 of the putatively differentially expressed genes was quantified in all 18 specimens described above and 30 further invasive breast cancer cases by qRT–PCR. Total RNA derived from these tumours was reverse-transcribed using oligo-dT18 primer in the First Strand Synthesis Kit [Amersham Pharmacia Biotech (now GE Healthcare Europe GmbH), Munich, Germany] according to the manufacturer's instructions. qRT–PCR analyses were performed on an ABI 7900 HT Sequence Detection System (Applied Biosystems) in triplicate, using the qPCR MasterMix Plus (Eurogentec, Liege, Belgium). The expressions of the target genes were normalized to the reference gene *PPIA* (Cyclophilin A). For the reference, a Pre-Developed TaqMan Assay Control Kit (Applied Biosystems) was used. For expression quantification of the target genes we applied TaqMan Assays on Demand (Applied Biosystems; see Supporting information, Table S1). qRT–PCR set-up was automatically done in a 384-well layout on a Genesis Workstation 150 (TECAN, Männedorf, Switzerland) equipped with a low volume. The conditions for qRT–PCR were: initial denaturation for 10 min at 95 °C, followed by 40 cycles, each consisting of 15 s at 95 °C and 1 min at 60 °C, for primer annealing and extension. Relative expression analyses were done with help of SDS 2.1 software (Applied Biosystems). The expression ratios between the samples were calculated from differences in threshold cycles at which an exponential increase in reporter fluorescence could first be detected (Ct values). Further data analyses was also performed with SDS 2.1 software, using fold change $= 2^{-\Delta\Delta Ct}$ [33].

Statistical analysis of microarray data

Applied Biosystems 1700 AB-Navigator Software Version 1.0.0.3 was used to correct the measured raw signal values in relation to background and intra-array shifts (default parameters recommended by Applied Biosystems). These values were exported from the Applied Biosystems software and imported into the statistics software package S-Plus 6.2 (TIBCO Software, Palo Alto, CA, USA). In this environment, all the different experiments were scaled to the same interquartile range and median [34].

The significant gene expression differences between the defined experimental groups were evaluated using a Shapiro–Wilk test (test on Normal distribution in both groups; $\alpha > 0.05$: Normal distribution, provided as a quality parameter), a variance test ($\alpha > 0.1$: no difference in the variance), and depending on that a *t*-test or the Welch variant of a *t*-test (two-sample, two-sided, $\alpha_{c1} < 0.05/\alpha_{c2} < 0.01$: assumed to be different). The additional filter criteria to call a gene 'differentially expressed' are a fold change value >1.5 and a difference of the group means >10. The stability of the candidates was additionally tested by a sampling procedure [35]. The applied variant is similar to the proven significance analysis of microarrays approach, which controls the false discovery rate [36].

To determine whether the proportions of the found differentially expressed 16q genes in conjunction with their found differentially expressed non-16q genes differed significantly from the proportions of the total

number of 16q genes and the total number of non-16q genes, respectively, we applied a χ^2 test (Pearson). Note that the normal fold-change scale ranges from 0 to ∞. The range 0–1 (decrease) corresponds to 1–∞ (increase). So our positive fold change values (increase) are directly taken from that scale, while the negative fold change values (decrease) are calculated by '1/fold change' to ease the comparability. This applies accordingly to all fold-change values in the text (± fold change values outside the parentheses and 0–∞ values in parentheses).

Visualization of the results of the differential gene expression analysis

To visualize the results of chromosome arm-specific gene expression analysis, we applied a function which maps the group of differentially expressed genes to their genomic location. The S-Plus (TIBCO Software, Palo Alto, CA, USA)-based function was written by the last author and is available on request. The genome is displayed as a panel of ordered metaphase ideograms of the human chromosomes 1–22 and X. The scaling along the chromosomes is linear, going from the physical position 0 at the top of the ideogram to the maximum length of the chromosomal sequence at the bottom. Each of the target genes is placed at its physical position on the respective chromosome, denoted by a tick mark and the gene identification code. This representation does not reflect the complex structure present in the genome and therefore may hide some special neighbourhood effects. Nevertheless, using this layout the overall distribution of the differentially expressed genes across the genome can be studied.

Results

CGH and array CGH

Genetic tumour characteristics defined by conventional, chromosome-based CGH of all 18 tumours of the first cohort are listed in Table 1A (column headed 'Cytogenetic alteration status'). Further analysis by array CGH revealed the genetic alterations shown in Figure 1A for 16q$^-$ and Figure 1B for 16q normal

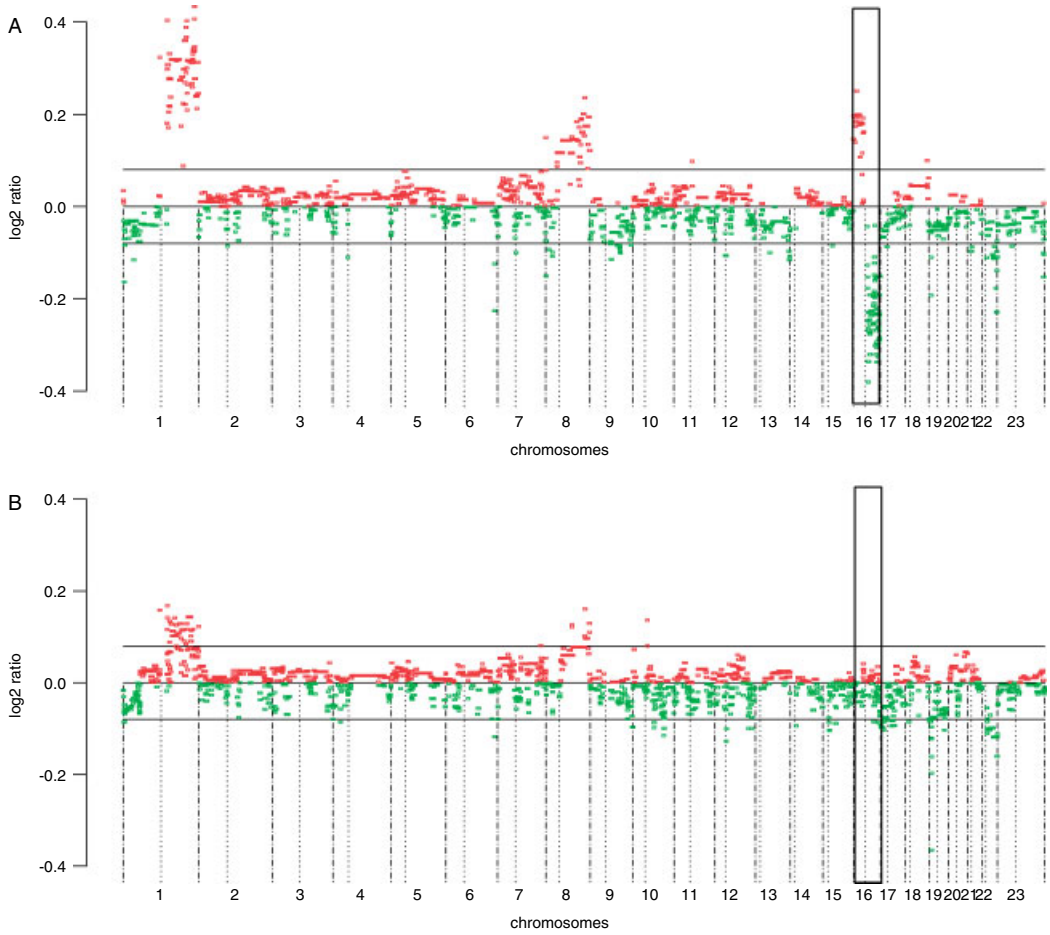

Figure 1. Genomic array CGH profiles. The two panels show median profiles of ductal invasive carcinomas with 16q loss (16q$^-$) and 16q normal state, respectively, from the first cohort (cf Table 1A). The *y* axis denotes log2 values generated by the circular binary segmentation (CBS) algorithm. The *x* axis shows genomic positions and chromosomes in ascending order. Copy number gains or losses, respectively, are denoted by a horizontal line (cut-off ratio >0.08 or ≤0.08; see Materials and methods). (A) Ductal invasive carcinomas holding a 16q loss. These cases harbour on average 4.8 genomic alterations. (B) Ductal invasive carcinomas with a 16q normal state. This group is characterized by on average 3.4 alterations. It can be clearly seen that both selections differ very distinctly at chromosome 16, more pronounced in the q arm (denoted by two black boxes at chromosome 16).

cases. Using the previously defined definition of whole chromosome arm losses [37], all breast cancer cases showing a loss of 16q in conventional CGH revealed an identical loss of 16q in array CGH. Identical results could be shown for all tumours with a normal 16q status in both techniques. Therefore, the definition of whole arm 16q losses was independent of the chosen method. In the following sections, loss of 16q always indicates the loss of the whole 16q arm, verified by both methods.

Microarray-based gene expression profiling

The analysis of the differential gene expression in relation to the chromosomal 16q status was performed to support the assumed asymmetric distribution of gene expression changes across the human genome by 16q alterations. To screen for systematic trends two different approaches differing in 16q status and tumour grade were analysed.

In a first step we compared four ductal invasive grade 2 carcinomas with a chromosomal 16q loss against four ductal invasive grade 2 carcinomas with normal 16q status concerning their panel of differentially expressed genes. All carcinomas displayed five or less genetic alterations per case by means of CGH. Both groups harboured 3.8 alterations/case on average.

An overview of all 86 differentially expressed genes in relation to their chromosomal location is shown in Figure 2A. Twelve differentially expressed genes were located on 16q (*CDA08, CGI-128, SNTB2, NQO1, SF3B3, KIAA0174, ATBF1, GABARAPL2, KARS, GCSH, MBTPS1* and *ZDHHC7*). All these genes displayed a decreased expression level in tumours with 16q loss. In contrast, 26 further differentially expressed genes located on 1q were up-regulated in tumours with 16q losses.

In a second step, we compared all 18 ductal invasive breast cancer cases of all grades with regard to their 16q status. The average number of chromosomal alterations/case in all tumours was 4.1. Nine tumours with 16q loss displayed, on average, 4.8 genetic alterations/case. The average number of genetic alterations per case with normal 16q status was 3.4. Seventy-eight genes were differentially expressed (Figure 2B). Twenty-three of these genes were located on 16q. Again, all genes exhibited a reduced expression in tumours with 16q loss. Only one differentially expressed gene was located on 1q, which means that the 1q effect of the first differential approach was completely diluted by the added cases. The 16q effect, in contrast, remained stable.

A detailed overview of all differentially expressed genes located on 16q in the two approaches is given in the Supporting information (Table S2). This table also includes intersections concerning 1q and further genomic locations. Eight genes located on 16q were differentially expressed in both comparisons (grey-shaded genes in Table 2). To test whether the distribution of the genes across the genome is not generated by chance, we performed a χ^2 test. Indeed, both result sets differed significantly (p_{c1} and $p_{c2} < 0.000001$) from the underlying distribution of all currently known genes in the human genome. This underlines the observed 16q effect. The other differentially expressed genes were distributed randomly over the non-16q and non-1q chromosomal regions.

Verification of microarray results by qRT–PCR

To verify the results described above, we performed real-time qRT–PCR for 10 candidate genes located on 16q in all 18 ductal invasive breast cancers. The candidate selection was taken from the intersection of the differential analyses (*SNTB2, SF3B3, MBTPS1*), candidates solely present in the second (c2) differential analysis (*KIAA1007, AARS, PSMD7, TERF2IP*) and some random candidates (*KIFC3, CFDP1, ANKRD11*) of 16q. This design was chosen to thoroughly verify the 16q effect independently from the microarray results.

The respective, cumulative results of the qRT–PCR are shown in Figure 3A. Clear, significant differences could be observed for carcinomas with loss of 16q in contrast to normal 16q status (Figure 3A, graph at left). The fold change between $16q^-$ and 16q normal is shown in Figure 3A (graph at right). On average, a fold change of—1.6 (0.5–0.7) was observed, indicating a reduction of the respective gene expression of slightly more than 50% in carcinoma with a loss of 16q.

Analysis of the impact of cases with a high number of alterations and ER-negative status on the outcome of the 16q effect by qRT–PCR

In a next step we analysed further 30 ductal invasive breast cancer cases of all grades again based on their chromosome 16q status (cf Table 1B, c3 analysis). Additionally, nearly all tumours with a normal 16q status showed ER-negativity, while all tumours with a 16q normal state exhibit ER-positivity. The average number of genetic alterations per tumour was 9.4, and 9.6 in carcinomas with and without chromosomal 16q losses, respectively.

In Figure 3B only three of the analysed genes still showed a decreased expression status for the tumour group harbouring a 16q loss. Despite the second cohort comprising a higher case number in comparison to the first cohort, the standard deviation (SD) was increased and therefore the results defined no significant fold change for seven of the genes analysed. The fold change values for both groups were within a range of −1.6 and 1.25 (0.64–1.25). The 16q loss effect disappeared in this cohort.

Discussion

The whole arm loss of 16q belongs to the earliest cytogenetic alterations in ER-positive, invasive breast cancer and some of its precursor lesions [38].

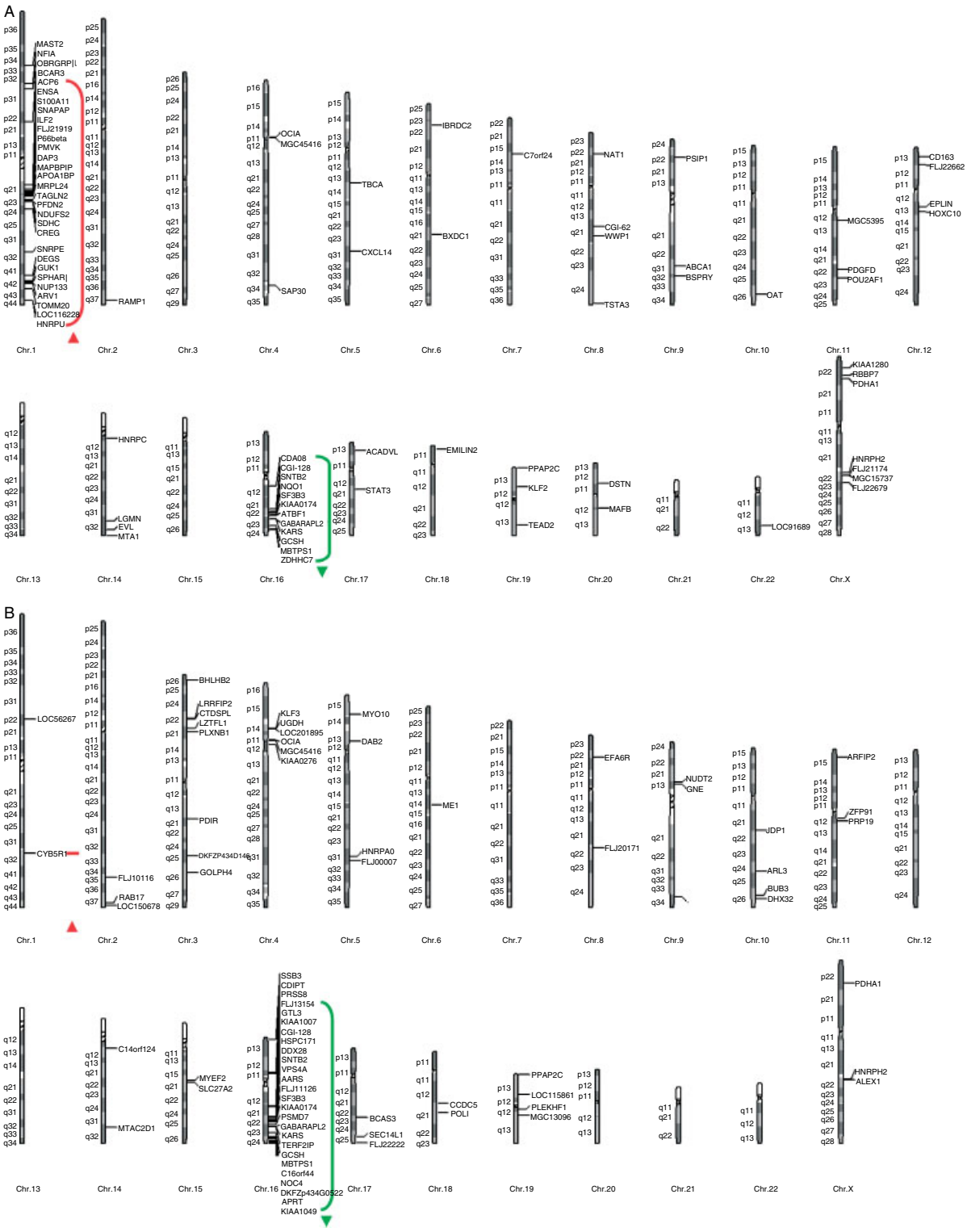

Figure 2. Overview of all differentially expressed genes with regard to their chromosomal location in three different comparisons. The genome is displayed as a panel of ordered metaphase ideograms of the human chromosomes 1–22 and X. To facilitate the orientation in these ideograms, shaded and non-shaded areas with adjacent annotations on the left side point to some prominent cytobands. The scaling along the chromosomes is linear (physical position 0 at the top of each of the ideograms). The differentially expressed genes are mapped to their genomic location on the right side of the ideogram and described by their gene symbol code. Up-regulated genes on 16q and 1q are accompanied by red brackets/lines (red arrow pointing up), while green brackets/lines denote down-regulated genes (green arrow pointing down). The purpose of using this layout is to study the overall distribution of the differentially expressed genes across the genome. (A) The differential genes resulting from the 16q normal versus $16q^-$ analysis with a low number of alterations [c1; cf Table 1, c1 column; average alterations $16q^-$/16q normal, both 3.8, average fold change ($16q^-$: 16q normal):1q: +1.9, 16q: −1.8]. On 16q and 1q a significant accumulation of differentially genes can be seen. (B) The cases of situation (A) complemented by cases with grade 3 and cases harbouring slightly more alterations [c2; cf Table 1, c2 column; average alterations now $16q^-$/16q normal: 4.8/3.4, average fold change ($16q^-$: 16q normal): 1q: +1.6, 16q: −1.6]. Here the 1q effect becomes diluted, while the 16q effect remains stable. All the 16q genes revealed a decreased expression, while 1q genes did show increased expression.

Table 2. List of all differentially expressed genes located on 16q detected by gene chip analysis

Gene name (HGNC)	Physical position start (chr. 16)	GO biological function
Intersection of c1 and c2 analyses		
CGI-128 (FAM96B)	66965959	Protein amino acid binding
SNTB2	69221032	Dystrophin-associated glycoprotein complex, microtubule, cytoskeleton, transport vesicle membrane
SF3B3	70557691	mRNA processing/splicing, via *trans*-esterification reactions, protein complex assembly
KIAA0174	71929304	Cell division, cell cycle
GABARAPL2	75600249	Protein transport, intra-Golgi vesicle-mediated transport, autophagy, positive regulation of ATPase activity
KARS	75661622	Translation, interspecies interaction between organisms, lysyl-tRNA aminoacylation, tRNA processing
GCSH	81115566	Glycine catabolic process
MBTPS1	84087366	Regulation of transcription factor import into nucleus, steroid metabolic process, proteolysis
c1 analysis only		
CDA08 (ITFG1)	47188298	Cell membrane component, T cell immunomodulatory protein
NQO1	69743304	Response to oxidative stress, positive regulation of (neuron) apoptosis, negative regulation of catalytic activity
ATBF1 (ZFHX3)	72816784	Zinc finger homeobox protein 3, regulation of myoblast differentiation, DNA-dependent regulation of transcription, transcription factor activity
ZDHHC7	85007787	Protein palmitoylation, subcellular trafficking of proteins between membrane compartments, modulating protein-protein interactions
c2 analysis only		
GTL3 (C16orf80)	58147492	Transcription factor IIB
KIAA1007 (CNOT1)	58553855	Regulation of transcription, protein binding, transcription complex subunit 1 interaction
HSPC171 (TMEM208)	67261016	Transmembrane protein
AARS	70286290	Translation, tRNA processing, tRNA aminoacylation, protein folding, negative regulation of (neuron) apoptosis
PSMD7	74330673	Anaphase-promoting complex-dependent proteasomal catabolic process, regulation of ubiquitin-protein ligase activity involved in mitotic cell cycle
TERF2IP	75681684	Regulation of telomere maintenance, regulation of transcription
DDX28	68055174	RNA binding, ATP-dependent helicase activity, hydrolase activity, located in the mitochondrion and nucleus
VPS4A	69345279	ESCRT III complex interaction, ATP binding, Rho GTPase binding, endosome transport, function in the cell cycle
KLHL36	84682107	Protein binding
KLHDC4	87741418	Protein binding
APRT	88875877	Purine ribonucleoside salvage pathway, adenine phosphoribosyltransferase activity
FLJ13154 (C16orf57)	58035305	This gene might be part of the RECQL4/SMAD4 pathways
FLJ11126 (DDX19A)	70380732	ATP-dependent RNA helicase involved in mRNA export from the nucleus (by similarity)
NOC4 (COX4NB)	85812230	Not yet characterized
KIAA1049 (TCF25)	89939970	Member of the basic helix-loop-helix (bHLH) family of transcription factors that are important in embryonic development (Steen and Lindholm, 2008); may play a role in cell death control

The exact location and known or suspected functions are listed. A second acronym denotes a synonym and is added as a courtesy for the reader. Grey shading, factors present in both differential gene expression analyses (intersection). The following blocks list genes only present in one differential analysis. *NA*, not available; *EST*, expressed sequence tag, no functional information available. Most factors are expressed during all live phases. They show up in normal and diseased states, eg a broad range of tumour types. They are also ubiquitously expressed in many different cell types (eg epithelial, endothelial, stem cells). To conclude, many of these genes seem to have functions according to basic cellular regulation and maintenance. Their differential pattern therefore points towards basic changes in cellular regulation.

The postulation of multiple parallel morphological and genetic pathways in *in situ* and invasive breast cancer is mainly based on the distribution of whole arm 16q losses in breast cancer cases of different grades [8–10,18]. This is further furnished by the fact that the underlying mechanisms of 16q losses differ significantly between these different breast cancer subtypes [13]. Following this line of evidence, it is worth speculating that in consequence different tumour suppressor genes might be involved in different breast cancer evolution lines. Interestingly, in recent years a number of putative tumour suppressor genes located on chromosome 16q have been described, although a definite candidate gene could not be defined [14–17,39].

The definition of differentially expressed genes according to the chromosomal status, especially in relation to chromosomal losses, is a difficult task in the early stages of breast cancer. Mostly breast cancer is a genetically rather far advanced disease with a multitude of genetic alterations, including chromosomal amplifications. It has been speculated that <20% of the global gene expression within breast cancer is a result of a gene dosage effect due to unbalanced chromosomal alterations [40]. Various cell biological reports have dealt with the influence of chromosome transfers into cancer cells and its associated differential gene expression, or with candidate genes associated with chromosomal amplifications [41]. More

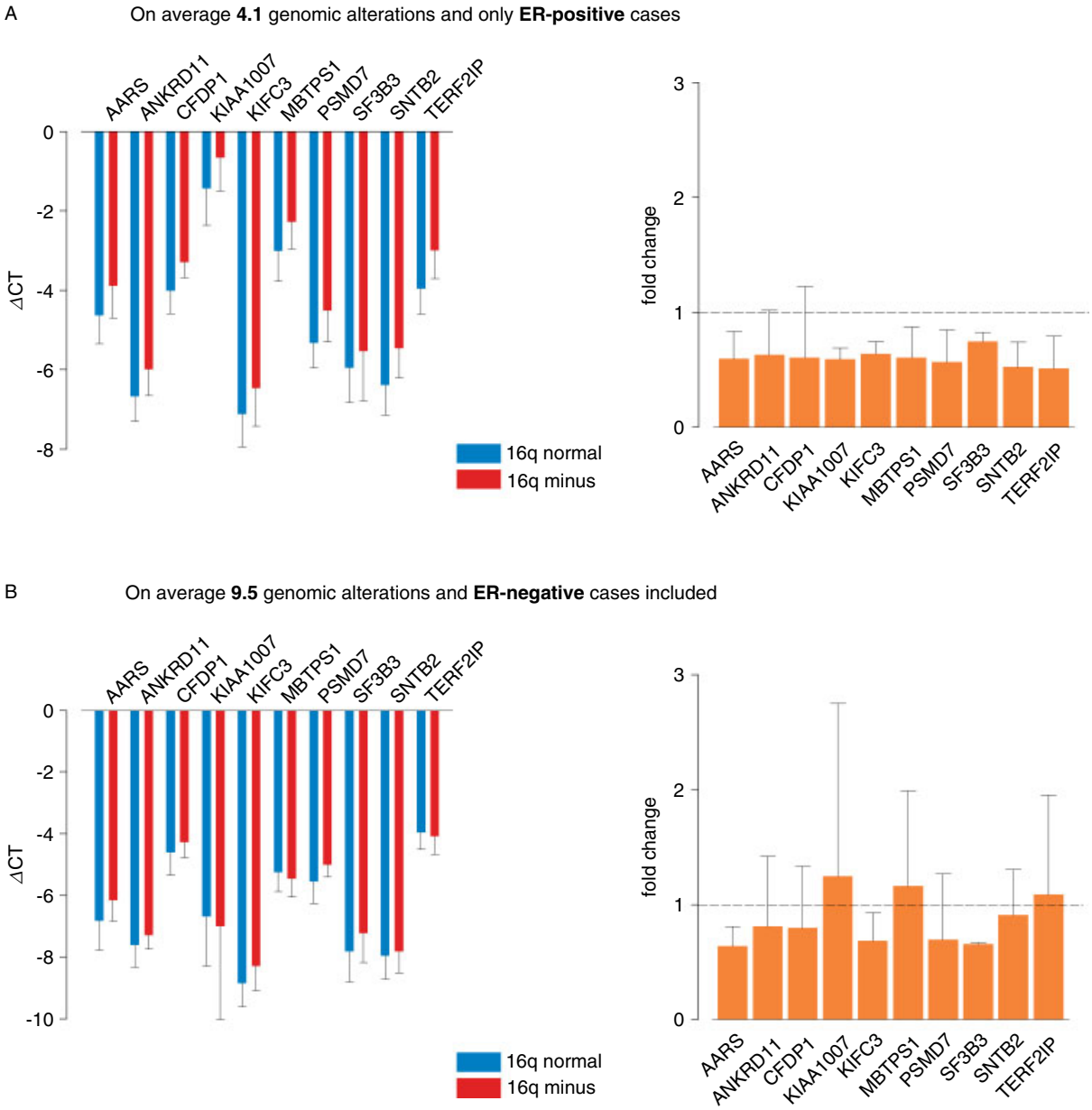

Figure 3. qRT–PCR results of (A) verification of the results of the gene expression microarray study and (B) analysis of these genes in second 16q$^-$/ER$^+$ versus 16q normal mostly ER$^-$ scenario. The left-hand graphs in both panels show the ΔΔCt values for two groups of interest: 16q normal and 16q$^-$ (blue/red). Both graphs show a smaller expression of the genes of interest concerning the used reference gene. The right-hand graphs in both panels present the fold change between the 16q normal and 16q$^-$ groups. Here the dashed line by the *y* axis value 1 denotes no fold change at all (cf fold change definition in Materials and methods). The 10 genes presented are selected from the differentially expressed 16q genes of the microarray study and some random picks on 16q. (A) Verification of the microarray results based on the c2 cases of the primary cohort (cf Table 1A, c2; on average 4.1 alterations). The left-hand graph shows that the difference between the 16q normal and 16q$^-$ (blue/red) groups is unidirectional in all cases. The right-hand graph illustrates this decrease in gene expression in the 16q$^-$ group (fold change on average −1.6, which is slightly more than the loss alone would imply). (B) This analysis is based on a complete independent set of breast cancer cases (cf Table 1B) exhibiting higher alterations (on average 9.5). The comparison is based on 16q$^-$/ER$^+$ and 16q normal/(mainly) ER$^-$ cases. The results show that the differences between the 16q normal and 16q$^-$ group are much smaller (left-hand graph). Many cases show only a weak decrease (fold change on average −1.2) but some cases even show an increase in differential expression (3/10 with a fold change on average of +1.2, right-hand graph). The SD is high and only three cases show a significant fold change. In this study the 16q effect is nearly gone.

recent studies demonstrated that the expression of genes located on 16q was associated with survival and also differed in different breast cancer subgroups [42,43]. In detail, the expression of genes located on 16q was significantly decreased in luminal A, ER-positive tumours. In consequence, the authors postulated that a stratification of breast cancer subgroups into 16q low-expressing and high-expressing tumours could be lead to a sharper distinction of breast cancers with a good and a poor prognosis.

We, in contrast have chosen an approach enabling the definition of differentially expressed genes located on 16q as early events in breast carcinogenesis. In a first step we have investigated ER-positive, ductal

invasive breast carcinomas with less than five genetic alterations, distinguished by the chromosomal 16q status. The concentration on these tumours led to a reduction of secondary effects of a differential expression due to other, non-16q-related events. Our results of this approach showed indeed that chromosomal 16q losses in invasive breast cancer are associated with a specific differential gene expression with a significantly high number of genes located on 16q (14/29% of all differentially regulated genes in the c1/c2 approaches). Interestingly, in accordance with the *a priori* hypothesis, all these genes indeed revealed a significant downregulation in tumours with whole arm chromosomal 16q losses. This was further verified by qRT–PCR.

In contrast to that, the inclusion of genetically more advanced, unstable breast cancer cases in this study tremendously changed this image. Whereas qRT–PCR verified the differential expression of the 16q-specific candidate genes in genetically early breast cancer cases (Figure 3A), indicating a loss of expression of slightly more than approximately 50%, this could not be reproduced in an additional series of 30 breast cancer cases of all grades (Figure 3B). The 16q effect vanished. Since the difference of these two cohorts was the higher degree of genetic instability and the ER status, the interpretation of these findings allows two major conclusions. Based on a classical progression model of breast cancer, 16q losses might represent first tumour initiation events in low-grade tumours, which effects are later diluted or substituted by secondary progression events. Alternatively, it is also possible to speculate that the ER expression status might have a higher impact on gene expression than a single chromosomal alteration.

However, the interpretation of these gets more confusing in regard to the differentially expressed genes in ER-positive breast cancers located on 1q. 16q losses may be a result of an unbalanced chromosomal translocation, t(1;16), leading to a whole-arm loss of 16q and a whole-arm gain of 1q [2]. Alternatively, 16q losses may appear as a simple chromosomal loss without any further chromosomal alteration. Our presented tumour series indicate that in genetically rather stable, ER-positive tumours the loss of 16q is associated with a gain of 1q, with the respective decreased or increased expression of the respective candidate genes located on 16q and 1q. Even though this might be speculative, an interaction between these genes can not be excluded. However, our gene expression results indicate that the low-grade, ER-positive pathway can be subdivided into tumours with whole arm 16q losses, with or without 1q gains.

The biological function of the genes found in the differential analyses (Table 2) and their role in breast carcinogenesis is largely unknown. One can only speculate about the biological importance of these genes. As far as there is additional information available on the candidate genes, many functions are related to cell cycle regulation, proliferation, exo-respective endocytosis and apoptosis. CGI-128 (synonyms: FAM96B, MIP18) plays a role in chromosome segregation during mitosis. An abnormality here can contribute to the pathogenesis of the genomic structure [44]. SNTB2 belongs to the molecular family of syntrophins, where the family member SNTA1 is discussed as a diagnostic or prognostic marker in oesophageal and breast cancers [45]. SNTB2 is known to be a Cdk5 substrate that controls the motility of secretory granules [46]. SF3B3 (spliceosome-associated subunit, 130 kDa) function effects the regulation of vascular endothelial growth factor (*VEGF*) mRNA [47]. *VEGF* is a regulatory element in tumour angiogenesis and mRNA processing, such as alternative splicing. KIAA0174 (synonym: hIST1) [48] is again part of the cytokinesis. GABA(A) receptor-associated protein-like 2 (GABARAPL2) seems to be active at certain developmental stages [49]. The family member L1 plays a role in conjunction with Atg4D in autophagosome formation, a process linked to staurosporine-induced cell death [50]. MBTPS1 (membrane-bound transcription factor site-1 protease) is discussed by Brandl *et al* [51] to be linked via ATF6a to ER stress response programmes, which in case of malfunction might induce inflammation. TERF2IP and other genes residing on 16q have been recently sequenced in a series of breast cancer cases without any evidence of mutations [52]. This overview on the affected core functionalities–foremost the second to last–show how 16q alterations might initiate a cellular state to acquire further defects and promote tumour progression. However, the diversity of biological functions associated with the candidate genes again underlines the often observed multi-causal nature of cancer phenotypes. Several simultaneously deregulated cellular subnetworks build the stress scenario to initiate tumour pathogenesis [53].

Although these results require verification and mutations of the respective candidate genes have to be excluded, the old hypothesis of a mere gene dosage effect mediating the loss of chromosome 16q should be discussed more intensely and these results might also be regarded as hints to discuss the haploinsufficiency theory in more detail. The effect of haploinsufficiency is associated with the loss of one allele of a specific gene in a tumour cell, whereas the other allele maintains only $<50\%$ of the expression of this gene, leading to a decreased overall expression (gene dosage) in the tumour cell. Dosage sensitivity has been implicated in tumourigenesis, especially for cell-cycle regulatory genes such as p53 and p27, but also for other genes [54]. However, a constant feature of haploinsufficient genes is that tumours generated via this mechanism are of later onset and less severity. In addition, haploinsufficiency has been associated with an early stage of the respective disease. For some genes also a pathway-specific haploinsufficiency effect has

been described. The parallels between these observations and the findings in breast cancer are compelling.

In conclusion, whole-arm loss of 16q has a strong impact on the expression of genes located on 16q in ER-positive breast cancer in a gene dosage-dependent manner, allowing a new round in the hunt for 16q-specific tumour suppressor genes in invasive breast cancer and its precursor lesions.

Acknowledgment

This study was funded in part by the Federal German Ministry of Research (Grant No. 01GM0869 TranSaR-Net) and in part by innovative medical research of the University of Münster Medical School (Grant Nos HU210810 and SC220705).

Author contributions

DH, HS, RN, JSRF, BB, HB, EK performed experiments and evaluations and contributed to the final manuscript; KP performed evaluations; and NT performed experiments.

References

1. Tirkkonen M, Tanner M, Karhu R, *et al*. Molecular cytogenetics of primary breast cancer by CGH. *Genes Chromosomes Cancer* 1998; **21**: 177–184.
2. Tsuda H, Takarabe T, Fukutomi T, *et al*. der(16)t(1;16)/der(1;16) in breast cancer detected by fluorescence *in situ* hybridization is an indicator of better patient prognosis. *Genes Chromosomes Cancer* 1999; **24**: 72–77.
3. Hislop RG, Pratt N, Stocks SC, *et al*. Karyotypic aberrations of chromosomes 16 and 17 are related to survival in patients with breast cancer. *Br J Surg* 2002; **89**: 1581–1586.
4. Richard F, Pacyna GM, Schl, *et al*. Patterns of chromosomal imbalances in invasive breast cancer. *Int J Cancer* 2000; **89**: 305–310.
5. Adeyinka A, Baldetorp B, Mertens F, *et al*. Comparative cytogenetic and DNA flow cytometric analysis of 242 primary breast carcinomas. *Cancer Genet Cytogenet* 2003; **147**: 62–67.
6. Adeyinka A, Mertens F, Idvall I, *et al*. Different patterns of chromosomal imbalances in metastasising and non-metastasising primary breast carcinomas. *Int J Cancer* 1999; **84**: 370–375.
7. Rennstam K, Ahlstedt-Soini M, Baldetorp B, *et al*. Patterns of chromosomal imbalances defines subgroups of breast cancer with distinct clinical features and prognosis. A study of 305 tumors by comparative genomic hybridization. *Cancer Res* 2003; **63**: 8861–8868.
8. Buerger H, Otterbach F, Simon R, *et al*. Comparative genomic hybridization of ductal carcinoma *in situ* of the breast–evidence of multiple genetic pathways. *J Pathol* 1999; **187**: 396–402.
9. Vos CB, ter HN, Rosenberg C, *et al*. Genetic alterations on chromosome 16 and 17 are important features of ductal carcinoma *in situ* of the breast and are associated with histologic type. *Br J Cancer* 2000; **81**: 1410–1418.
10. Roylance R, Gorman P, Harris W, *et al*. Comparative genomic hybridization of breast tumours stratified by histological grade reveals new insights into the biological progression of breast cancer. *Cancer Res* 1999; **59**: 1433–1436.
11. Natrajan R, Lambros MB, Geyer FC, *et al*. Loss of 16q in high grade breast cancer is associated with estrogen receptor status: evidence for progression in tumors with a luminal phenotype? *Genes Chromosomes Cancer* 2009; **48**: 351–365.
12. O'Connell P, Pekkel V, Fuqua SA, *et al*. Analysis of loss of heterozygosity in 399 premalignant breast lesions at 15 genetic loci. *J Natl Cancer Inst* 1998; **90**: 697–703.
13. Cleton-Jansen AM, Buerger H, Haar NN, *et al*. Different mechanisms of chromosome 16 loss of heterozygosity in well versus poorly differentiated ductal breast cancer. *Genes Chromosomes Cancer* 2004; **41**: 109–116.
14. Cleton-Jansen AM, Callen DF, Seshadri R, *et al*. Loss of heterozygosity mapping at chromosome arm 16q in 712 breast tumors reveals factors that influence delineation of candidate regions. *Cancer Res* 2001; **61**: 1171–1177.
15. Kochetkova M, McKenzie OL, Bais AJ, *et al*. *CBFA2T3* (MTG16) is a putative breast tumor suppressor gene from the breast cancer loss of heterozygosity region at 16q24.3. *Cancer Res* 2002; **62**: 4599–4604.
16. Powell JA, Gardner AE, Bais AJ, *et al*. Sequencing, transcript identification, and quantitative gene expression profiling in the breast cancer loss of heterozygosity region 16q24.3 reveal three potential tumor-suppressor genes. *Genomics* 2002; **80**: 303–310.
17. Kumar R, Neilsen PM, Crawford J, *et al*. *FBXO31* is the chromosome 16q24.3 senescence gene, a candidate breast tumor suppressor, and a component of an SCF complex. *Cancer Res* 2005; **65**: 11304–11313.
18. Buerger H, Mommers E, Littmann R, *et al*. Ductal invasive G2 and G3 carcinomas of the breast are the end stages of at least two different lines of genetic evolution. *J Pathol* 2001; **194**: 165–170.
19. Lacroix-Triki M, Suarez PH, Mackay A, *et al*. Mucinous carcinoma of the breast is genomically distinct from invasive ductal carcinomas of no special type. *J Pathol* 2010; **222**: 282–298.
20. Coe BP, Ylstra B, Carvalho B, *et al*. Resolving the resolution of array CGH. *Genomics* 2007; **89**: 647–653.
21. Gunnarsson R, Staaf J, Jansson M, *et al*. Screening for copy-number alterations and loss of heterozygosity in chronic lymphocytic leukemia–a comparative study of four differently designed, high resolution microarray platforms. *Genes Chromosomes Cancer* 2008; **47**: 697–711.
22. Tan DS, Lambros MB, Natrajan R, *et al*. Getting it right: designing microarray (and not 'microawry') comparative genomic hybridization studies for cancer research. *Lab Invest* 2007; **87**: 737–754.
23. Geyer FC, Weigelt B, Natrajan R, *et al*. Molecular analysis reveals a genetic basis for the phenotypic diversity of metaplastic breast carcinomas. *J Pathol* 2010; **220**: 562–573.
24. Mackay A, Tamber N, Fenwick K, *et al*. A high-resolution integrated analysis of genetic and expression profiles of breast cancer cell lines. *Breast Cancer Res Treat* 2009; **118**: 481–498.
25. Natrajan R, Weigelt B, Mackay A, *et al*. An integrative genomic and transcriptomic analysis reveals molecular pathways and networks regulated by copy number aberrations in basal-like, HER2 and luminal cancers. *Breast Cancer Res Treat* 2010; **121**: 575–589.
26. Olshen AB, Venkatraman ES, Lucito R, *et al*. Circular binary segmentation for the analysis of array-based DNA copy number data. *Biostatistics* 2004; **5**: 557–572.
27. Natrajan R, Lambros MB, Rodriguez-Pinilla SM, *et al*. Tiling path genomic profiling of grade 3 invasive ductal breast cancers. *Clin Cancer Res* 2009; **15**: 2711–2722.
28. Korsching E, Packeisen J, Agelopoulos K, *et al*. Cytogenetic alterations and cytokeratin expression patterns in breast cancer: integrating a new model of breast differentiation into cytogenetic pathways of breast carcinogenesis. *Lab Invest* 2002; **82**: 1525–1533.

29. Korsching E, Packeisen J, Helms MW, *et al*. Deciphering a subgroup of breast carcinomas with putative progression of grade during carcinogenesis revealed by comparative genomic hybridisation (CGH) and immunohistochemistry. *Br J Cancer* 2004; **90**: 1422–1428.
30. Korsching E, Packeisen J, Liedtke C, *et al*. The origin of vimentin expression in invasive breast cancer: epithelial-mesenchymal transition, myoepithelial histogenesis or histogenesis from progenitor cells with bilinear differentiation potential? *J Pathol* 2005; **206**: 451–457.
31. Marchio C, Lambros MB, Gugliotta P, *et al*. Does chromosome 17 centromere copy number predict polysomy in breast cancer? A fluorescence *in situ* hybridization and microarray-based CGH analysis. *J Pathol* 2009; **219**: 16–24.
32. Agelopoulos K, Kersting C, Korsching E, *et al*. Egfr amplification specific gene expression in phyllodes tumours of the breast. *Cell Oncol* 2007; **29**: 443–451.
33. 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**: 402–408.
34. Bolstad BM, Irizarry RA, Astrand M, *et al*. A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. *Bioinformatics* 2003; **19**: 185–193.
35. Knowles HJ, Cleton-Jansen AM, Korsching E, *et al*. Hypoxia-inducible factor regulates osteoclast-mediated bone resorption: role of angiopoietin-like 4. *FASEB J* 2010; **24**: 4648–4659.
36. Tusher VG, Tibshirani R, Chu G. Significance analysis of microarrays applied to the ionizing radiation response. *Proc Natl Acad Sci USA* 2001; **98**: 5116–5121.
37. Korsching E, Agelopolous K, Schmidt H, *et al*. Improvements in the analysis strategy make single nucleotide polymorphism analysis a powerful tool in the detection and characterization of amplified chromosomal regions in human tumors. *Pathobiology* 2006; **73**: 18–25.
38. Bombonati A, Sgroi DC. The molecular pathology of breast cancer progression. *J Pathol* 2011; **223**: 307–317.
39. Rakha EA, Green AR, Powe DG, *et al*. Chromosome 16 tumor-suppressor genes in breast cancer. *Genes Chromosomes Cancer* 2006; **45**: 527–535.
40. 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.
41. Klebig C, Seitz S, Korsching E, *et al*. Profile of differentially expressed genes after transfer of chromosome 17 into the breast cancer cell line CAL51. *Genes Chromosomes Cancer* 2005; **44**: 233–246.
42. Nordgard SH, Johansen FE, Alnaes GI, *et al*. Genome-wide analysis identifies 16q deletion associated with survival, molecular subtypes, mRNA expression, and germline haplotypes in breast cancer patients. *Genes Chromosomes Cancer* 2008; **47**: 680–696.
43. Wennmalm K, Calza S, Ploner A, *et al*. Gene expression in 16q is associated with survival and differs between Sorlie breast cancer subtypes. *Genes Chromosomes Cancer* 2007; **46**: 87–97.
44. Ito S, Tan LJ, Andoh D, *et al*. MMXD, a TFIIH-independent XPD–MMS19 protein complex involved in chromosome segregation. *Mol Cell* 2010; **39**: 632–640.
45. Bhat HF, Baba RA, Bashir M, *et al*. α-1-Syntrophin protein is differentially expressed in human cancers. *Biomarkers* 2011; **16**: 31–36.
46. Schubert S, Knoch KP, Ouwendijk J, *et al*. β2-Syntrophin is a Cdk5 substrate that restrains the motility of insulin secretory granules. *PLoS One* 2010; **5**: e12929-.
47. Furumai R, Uchida K, Komi Y, *et al*. Spliceostatin A blocks angiogenesis by inhibiting global gene expression including VEGF. *Cancer Sci* 2010; **101**: 2483–2489.
48. Agromayor M, Carlton JG, Phelan JP, *et al*. Essential role of hIST1 in cytokinesis. *Mol Biol Cell* 2009; **20**: 1374–1387.
49. Liang K, Lin Y, Zhang Y, *et al*. Developmental expression of *Amphioxus* GABAA receptor-associated protein-like 2 gene. *Dev Genes Evol* 2004; **214**: 339–341.
50. Betin VM, Lane JD. Caspase cleavage of Atg4D stimulates GABARAP-L1 processing and triggers mitochondrial targeting and apoptosis. *J Cell Sci* 2009; **122**: 2554–2566.
51. Brandl K, Rutschmann S, Li X, *et al*. Enhanced sensitivity to DSS colitis caused by a hypomorphic Mbtps1 mutation disrupting the ATF6-driven unfolded protein response. *Proc Natl Acad Sci USA* 2009; **106**: 3300–3305.
52. van Wezel T, Lombaerts M, van Roon EH, *et al*. Expression analysis of candidate breast tumour suppressor genes on chromosome 16q. *Breast Cancer Res* 2005; **7**: R998–1004.
53. Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. *Cell* 2011; **144**: 646–674.
54. Santarosa M, Ashworth A. Haploinsufficiency for tumour suppressor genes: when you don't need to go all the way. *Biochim Biophys Acta* 2004; **1654**: 105–122.

SUPPORTING INFORMATION ON THE INTERNET

The following supporting information may be found in the online version of this article:

Table S1. Quantification of differentially expressed genes by qRT–PCR

Table S2. (A) Intersections of differential gene expression results of the c1 and c2 analyses, differentiated according to chromosomal region. (B) Intersections—complete gene list. (C) Differential candidate genes on 16q based on the c1 and c2 analyses.