High-Resolution Genomic Analysis of the 11q13 Amplicon in Breast Cancers Identifies Synergy with 8p12 Amplification, Involving the mTOR Targets S6K2 and 4EBP1

Elin Karlsson, Marie Ahnström, Josefine Bostner, Gizeh Perez-Tenorio, Birgit Olsson, Anna-Lotta Hallbeck and Olle Stål

Linköping University Post Print

N.B.: When citing this work, cite the original article.

This is the authors’ version of the following article:

Elin Karlsson, Marie Ahnström, Josefine Bostner, Gizeh Perez-Tenorio, Birgit Olsson, Anna-Lotta Hallbeck and Olle Stål, High-Resolution Genomic Analysis of the 11q13 Amplicon in Breast Cancers Identifies Synergy with 8p12 Amplification, Involving the mTOR Targets S6K2 and 4EBP1, 2011, Genes, Chromosomes and Cancer, (50), 10, 775-787.

which has been published in final form at: http://dx.doi.org/10.1002/gcc.20900

Copyright: Wiley-Blackwell


Postprint available at: Linköping University Electronic Press

http://urn.kb.se/resolve?urn=urn:nbn:se:liu:diva-70514
High-resolution genomic analysis of the 11q13 amplicon in breast cancers identifies synergy with 8p12 amplification, involving the mTOR targets S6K2 and 4EBP1

Elin Karlsson, Marie Ahnström Waltersson, Josefine Bostner, Gizeh Pérez-Tenorío, Birgit Olsson, Anna-Lotta Hallbeck and Olle Ståhl

Authors’ affiliations
Department of Clinical and Experimental Medicine, Division of Oncology, Faculty of Health Sciences, Linköping University, Linköping, Sweden

Corresponding author
Elin Karlsson, Department of Clinical and Experimental Medicine, Division of Oncology, Faculty of Health Sciences, Linköping University, SE-581 85 Linköping, Sweden. Phone: +46 010 1038117; Fax: +46 013-127465; e-mail: elin.karlsson@liu.se

This study was supported by grants from the Swedish Cancer Foundation and the Swedish Research Council.

Running title
11q13 amplification in breast cancer
Abstract

The chromosomal region 11q13 is amplified in 15-20% of breast cancers; an event associated with oestrogen receptor (ER) expression but also implicated in resistance to endocrine therapy. Coamplifications of the 11q13 and 8p12 regions are common, suggesting synergy between the amplicons. The aim was to identify candidate oncogenes in the 11q13 region based on recurrent amplification patterns and correlations to mRNA expression levels. Furthermore, the 11q13/8p12 coamplification and its prognostic value, was evaluated at the DNA and the mRNA levels. Affymetrix 250K NspI arrays were used for whole genome screening of DNA copy number changes in 29 breast tumours. To identify amplicon cores at 11q13 and 8p12, Genomic Identification of Significant Targets in Cancer (GISTIC) was applied. The mRNA expression levels of candidate oncogenes in the amplicons (RAD9A, RPS6KB2 (S6K2), CCND1, FGF19, FGF4, FGF3, PAK1, GAB2 (11q13); EIF4EBP1 (4EBP1), PPAPDC1B and FGFR1 (8p12)) were evaluated using real-time PCR. Resulting data revealed three main amplification cores at 11q13. ER expression was associated with the central 11q13 amplification core, encompassing CCND1, whereas 8p12 amplification/gene expression correlated to S6K2 in a proximal 11q13 core. Amplification of 8p12 and high expression of 4EBP1 or FGFR1 was associated with a poor outcome in the group. In conclusion, SNP arrays have enabled mapping of the 11q13 amplicon in breast tumours with high resolution. A proximal 11q13 core including S6K2 was identified as involved in the coamplification/coexpression with 8p12, suggesting synergy between the mTOR targets S6K2 and 4EBP1 in breast cancer development and progression.
Introduction

Genetic and epigenetic alterations are fundamental characteristics in the development of malignancy. The chromosome band 11q13 is amplified in a significant proportion of cancers, including 15-20% of breast cancers (Borg et al., 1991; Karlseder et al., 1994; Schuuring, 1995; Courjal et al., 1997; Al-Kuraya et al., 2004). Amplification of this region has been reported also in benign and in situ lesions of the breast (Weinstat-Saslow et al., 1995; Simpson et al., 1997; Alle et al., 1998), suggesting that this is an early event, connected to initiation of carcinogenesis, rather than a secondary effect of genomic instability. Several studies have indicated a clinical significance of the 11q13 amplicon, but its complex amplification patterns and the fact that the region is extremely gene dense have obstructed a definite identification of driving oncogenes. To date, CCND1, encoding the cell cycle regulator Cyclin D1, is by far the most extensively studied candidate oncogene in this area in breast cancer. However, it has been suggested that the region harbours several independent amplification cores, indicating additional driving oncogenes (Ormandy et al., 2003). In the first detailed study of 11q13 amplifications in breast cancer, 389 carcinomas were screened for amplification status of ten molecular markers, which identified four independent amplification cores around D11S97 (62 Mb), CCND1 (69.5 Mb), EMS1 (70 Mb) and D11S833E (76-77Mb) (Karlseder et al., 1994). These areas were further confirmed as independent ampilicons by Southern blot, real-time PCR and comparative genomic hybridization (CGH) (Bekri et al., 1997; Courjal et al., 1997; Hui et al., 2002; Rodriguez et al., 2004). In a recent study, 18 breast tumours with 11q13 amplification were screened using array CGH with 101 BAC clones covering the 11q12-14 area (Kwek et al., 2009). The study revealed an even more complex amplification pattern, with several separated areas of amplification. However, minimal areas of amplification could be condensed into five cores (around 69.5 Mb, 70.0 Mb, 76.2 Mb, 76.5 Mb and 77.1 Mb).
11q13 amplification has been correlated to positive oestrogen receptor (ER) status in breast tumours (Letessier et al., 2006; Elsheikh et al., 2008). Due to the availability of antioestrogen therapies, ER positivity is considered as a marker of good prognosis. However, among patients with ER positive tumours, those also harbouring 11q13 amplification seem to constitute a subgroup with higher grade and a relatively poor prognosis (Al-Kuraya et al., 2004). A possible explanation is a connection between the amplicon and endocrine resistance, which has been implicated in both in vitro (Wilcken et al., 1997; Bindels et al., 2002; Hui et al., 2002) and in clinical studies (Kenny et al., 1999; Stendahl et al., 2004; Ahnström et al., 2005; Jirström et al., 2005; Holm et al., 2006; Bostner et al., 2007; Rudas et al., 2008).

Several studies have demonstrated that the 11q13 region is commonly coamplified with regions at 8p11-12 (Cuny et al., 2000; Letessier et al., 2006). The presence of both of these amplifications seems to predict a worse patient outcome than either amplicon alone, suggesting a synergy where potential oncogenes remain to be elucidated (Cuny et al., 2000).

In the present study, whole-genome single nucleotide polymorphism (SNP) arrays were used to perform a comprehensive analysis of copy number alterations in breast tumours harbouring amplification in the 11q13 region and to map this area with high resolution. By using the statistical approach Genomic Identification of Significant Targets In Cancer (GISTIC) (Beroukhim et al., 2007), three main amplicon cores at 11q13 could be identified. The mRNA levels of candidate oncogenes in these hotspots were evaluated using quantitative real-time PCR. Among recurrent genome-wide copy number alterations, amplification in the 8p12 region was most frequently found. This connection was evaluated at DNA and mRNA levels and the results suggest a prognostic significance of 11q13/8p12 coamplification and coexpression. Furthermore, coamplification was found to frequently engage the amplicon cores including the mTOR targets S6K2 and 4EBP1.
Materials and methods

Patients

The study comprised a tumour material of 34 breast tumours with 11q13 amplification, selected from a cohort of 185 breast tumours from postmenopausal patients with stage II breast cancer. 11q13 positivity was defined as four or more gene copies of CCND1 (encoding Cyclin D1) and/or PAK1 (encoding p21-activated kinase 1) and/or RPS6KB2 (S6K2, encoding S6 kinase 2) (Bostner et al., 2007; Perez-Tenorio et al., 2010). From these, 29 were available for DNA extraction and array analysis, and 23 (whereof 21 included in the array analysis) were available for RNA extraction and real-time PCR. All samples were histologically estimated to contain at least 50% malignant cells.

DNA and RNA extraction

Genomic DNA was extracted from fresh-frozen tumour tissue using the Puregene® DNA purification kit (Gentra, Minneapolis, MN, US) according to the manufacturer’s instructions. Briefly, the tissue was digested in proteinase-K solution, 55°C, until totally dissolved after approximately 4-5 days. The cell lysate was treated with RNase, for 15 min, at 37°C. Proteins were precipitated using protein precipitation solution and pelleted by centrifugation (16000*g, 3 min, room temperature). The supernatant was collected and DNA was precipitated with isopropanol and pelleted by centrifugation (16000*g, 1 min, room temperature). Finally, DNA was washed with 70% ethanol, dried in air and dissolved in sterile water for storage at -70°C. DNA concentration was measured using a NanoDrop ND-1000 Spectrophotometer (Nanodrop Technologies, Wilmington, DE, US). DNA quality, referred to as minimum degradation, was assessed by running the samples on a 1% agarose gel for 30-60 minutes, using genomic DNA 103, included in the Affymetrix 250K Nsp kit, as a positive control for intact DNA.
Of the 11q13 positive samples initially included in the study, 23 were available for RNA extraction, whereof 21 were included in the array analysis. Fresh frozen tumour tissue was homogenised with a microdismembrator (B Braun, Melsungen, Germany) and total RNA was isolated using the mirVana™ miRNA Isolation kit (Ambion, Austin, TX, USA), according to manufacturer’s recommendations. Purified RNA was eluted in nuclease-free water and RNasin® Ribonuclease Inhibitor (Promega, WI, USA) was added before storage at -70°C. RNA quantity and quality was assessed with an Agilent 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA, US).

**SNP arrays and data analysis**

Affymetrix 250K NspI arrays were used to profile the 29 tumour samples for whole genome DNA copy number changes. The assay was performed according to the Affymetrix recommended protocol (Affymetrix, Santa Clara, CA, US). Briefly, 250 ng of total genomic DNA was digested with NspI, ligated to adaptors and PCR amplified using adaptor specific primers. The purified PCR products were fragmented, labelled with biotin and hybridised to the arrays. Affymetrix Scanner 3000 7G was used to scan the arrays and generated data were converted to cell intensity files in GeneChip® Operating Software (GCOS). Cell intensity data were processed by Affymetrix GeneChip® Genotyping Analysis Software (G-TYPE) 4.1. The BRLMM algorithm was applied for data normalisation and generation of genotype calls for each SNP in each sample. Resulting data was utilised for copy number analysis with the Copy Number Analysis Tool (CNAT) 4 algorithm as follows. Briefly, the relative copy number of each SNP was primarily estimated by calculating the log2ratio of the normalised signals; log2(S_{\text{sample}}/\text{mean } S_{\text{ref}}), where 30 female samples from the HapMap project (available at www.Affymetrix.com) were used as references. To reduce the level of noise and to capture important patterns in the data, Gaussian smoothing was performed on the log2ratios, using the
default bandwidth of 100kb. Copy number state (CNstate) for the SNPs were estimated by applying a Hidden Markov Model (HMM), using default parameters in the software, where five hidden states represent the following copy number aberrations; homozygous deletion (CNstate=0), hemizygous deletion (CNstate=1), copy neutral (CNstate=2), single copy gain (CNstate=3), and amplification (CNstate=4). Gene copy number was also expressed as a continuous variable by calculating $2^{(\log_2 \text{ratio} + 1)}$ for each SNP. Resulting data was visualised by CNATviewer and Affymetrix Genotyping Console (GTC) Browser. The NCBI tool MapViewer was applied for mapping of known genes within regions of gains and losses, whereby positions of genes were received from the Human genome build 36, ver3.

Quantitative real-time PCR

RNA was reverse transcribed into cDNA using the high-capacity cDNA reverse transcription kit (Applied Biosystems, Stockholm, Sweden), following the manufacturer’s instructions. For each reaction, 200 ng RNA was added to a final reaction volume of 20 µl. mRNA levels for RAD9A, S6K2, CCND1, FGF19, FGF4, FGF3, PAK1, GAB2, EIF4EBP1 (4EBP1), PPAPDC1B and FGFR1 were measured in the 23 samples as previously described, using quantitative fast real-time PCR. TaqMan® assays were ordered as assay-on-demand with the exception of FGF4 for which primers and probes were designed using Primer Express v1.5a (Applied Biosystems) (Table 1). Real-time PCR was performed using TaqMan® Fast Universal PCR Master Mix (Applied Biosystems) in reactions of 10 µl. TaqMan® assays were diluted according to manufacturer’s instructions, whereas a final concentration of 0.2 µM were used for FGF4 primers and probes. TaqMan® chemistry was performed on an ABI Prism 7900ht (Applied Biosystems) using the default thermal program: 95°C for 20 s, followed by 40 cycles of 95°C for 1 s, and 60°C for 20 s. The expression of FGF3, FGF4 and FGF19 was low in most samples. To increase the sensitivity of these assays; a preamplification step was performed in which the thermal program: 95°C for 60 s, followed
by 15 cycles of 95°C for 1 s, and 60°C for 20 s, was run in a Peltier Thermal Cycler (PTC-200, MJ research) for each sample. Relative expression of each gene was calculated using the standard curve method, with ACTB as an endogenous control (Table 1). Briefly, standard curves were constructed with four-fold serial dilutions of cDNA from the following cell lines: SKBR3 (S6K2, CCND1, PAK1, GAB2, 4EBP1 and FGFR1), HCT-116 (RAD9A and PPAPDC1B) and SW613-S; kindly provided by Olivier Brison, Curie Institute, Paris, France (FGF19, FGF4 and FGF3). Standard curves were included on each plate for genes of interest as well as for the endogenous control. All samples were run in triplicates and the median Ct-values were used to calculate a relative expression value (C) for each gene, based on the respective standard curve. Final mRNA quantitation was performed by calculating the ratio (C gene of interest/C endogenous control) for each sample.

Gene copy number analysis of CCND1, PAK1 and S6K2 by real-time PCR have been described elsewhere (Bostner et al., 2007; Perez-Tenorio et al., 2010)

### Table 1 TaqMan® assays used for mRNA quantification by fast real-time PCR

<table>
<thead>
<tr>
<th>Gene</th>
<th>Assay (Applied Biosystems)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACTB</td>
<td>Part no 4310881E</td>
</tr>
<tr>
<td>RAD9A</td>
<td>Hs00270240_m1</td>
</tr>
<tr>
<td>RPS6KB2</td>
<td>Hs00177689_m1</td>
</tr>
<tr>
<td>CCND1</td>
<td>Hs00765553_m1</td>
</tr>
<tr>
<td>FGF19</td>
<td>Hs00192780_m1</td>
</tr>
<tr>
<td>FGF3</td>
<td>Hs01104462-m1</td>
</tr>
<tr>
<td>FGF4</td>
<td>Forward:5’-AGCAAGGGCAAGCTCTATG</td>
</tr>
<tr>
<td></td>
<td>Reverse:5’-GTGTGGGAGGAGGAGAATCTCCT</td>
</tr>
<tr>
<td></td>
<td>Probe: 5’-FAM-TCGCCCTTCTTCACCAGATGACTGC-TAMRA</td>
</tr>
<tr>
<td>PAK1</td>
<td>Hs00176815_m1</td>
</tr>
<tr>
<td>GAB2</td>
<td>Hs00373045_m1</td>
</tr>
<tr>
<td>EIF4EBP1</td>
<td>Hs00607050_m1</td>
</tr>
<tr>
<td>PPAPDC1B</td>
<td>Hs00261167_m1</td>
</tr>
<tr>
<td>FGFR1</td>
<td>Hs00241111_m1</td>
</tr>
</tbody>
</table>
Evaluation of ER and ERBB2 status

ER concentration was assessed in clinical routine by isoelectric focusing or enzyme immunoassay (EIA) at the time of diagnosis (Wrange et al., 1978), and these results are highly comparable with present-day immunohistochemistry data (Khoshnoud et al., 2010). ERBB2 protein expression was quantified using flow cytometry (Stål et al., 1994), whereas ERBB2 gene copy number was determined by real-time PCR (Gunnarsson et al., 2003).

Statistical analyses

To give a comprehensive view of the 11q13 amplification patterns across all samples and identify regions of significant aberrations, the principles of a statistical approach referred to as Genomic Identification of Significant Targets in Cancer (GISTIC) was applied (Beroukhim et al., 2007). GISTIC methodology is based on the assumption that the frequencies of occurrence, and to a greater extent the amplitude of the aberrations, are corner stones in the development of functionally important genomic changes. G-scores were calculated for each SNP across all samples for both gains and deletions according to:

\[ G^{\text{gain}} = F^{\text{gain}} \times C^{\text{gain}} \]
\[ G^{\text{del}} = F^{\text{del}} \times C^{\text{del}} \]

where \( F \) represents the frequency of an aberration and \( C \) its average amplitude. The threshold for gain was set to a CNstate \( \geq 3 \) and deletion CNstate \(<2\). G-scores are plotted against chromosomal position and resulting peaks of maximal amplitude are indicative of regions harbouring driving oncogenes.

Spearman’s Rank Order correlation was applied for assessing the correlation between gene copy number and mRNA expression level as well as between \( FGFR1 \) mRNA and ERBB2 protein expression. Multiple linear regression analysis was used for evaluating the
correlations between the different amplification cores at 11q13, and ER expression respectively $FGFR1$ gene copy number. ER expression values were calculated as log10 (x+1) in order to achieve a distribution close to normal. The Kaplan-Meier product limit method was used to estimate the cumulative survival probability in a patient group. Differences in survival between groups were evaluated with a log-rank test. All statistical analyses were performed with Statistica 9.1 (Statsoft Inc, Tulsa, OK, US). The criterion for statistical significance was $P<0.05$.

Results

*Genome-wide copy number profiles of 11q13 amplified breast tumours*

DNA copy number alterations were analysed on a genome wide scale in 29 breast tumours with 11q13 amplification. Using the definitions described by Hicks and colleagues (Hicks et al., 2006), 25/29 samples (86%) belonged to a copy number profile referred to as “firestorm”, that is one or a few regions of clustered, narrow peaks of high level amplification, with the remaining genome containing mainly simple alterations with duplications or deletions comprising entire chromosomes or chromosome arms. In 2/29 (7%) samples, no focal amplifications were evident, but some regions, including 11q13, demonstrated segments of duplications or deletions covering entire chromosomes or arms. These samples can be considered as having a simplex amplification pattern. Two samples (7%) exhibited a complex sawtooth pattern, with several narrow amplifications and deletions throughout the genome. Recurrent genome-wide copy number alterations are reported in Table 2. Copy number alterations identified in >50% of 11q13 amplified tumours included gain of 1q, 8p, 8q, loss of 16q and the distal part of 11q. The combination of 1q gain and 16q loss occurred in 12/29 (41.4%) of the tumours. In addition to 11q13, recurrent focal amplicons occurred at 4q13 ($EREG, AREG, RCHY$), 8p12, 12q14 ($MDM2$) and 17q12 ($ERBB2$).
Table 2 Frequencies of recurrent copy number alterations among 29 breast tumours with 11q13 amplification (-deletion, +gain, ++amplification).

<table>
<thead>
<tr>
<th>Copy number alteration</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>1p-</td>
<td>48.3% (14/29)</td>
</tr>
<tr>
<td>1q+</td>
<td>69.0% (20/29)</td>
</tr>
<tr>
<td>3p12-21-</td>
<td>34.5% (10/29)</td>
</tr>
<tr>
<td>4q13+</td>
<td>13.8% (4/29)</td>
</tr>
<tr>
<td>6q-</td>
<td>37.9% (11/29)</td>
</tr>
<tr>
<td>8p12++</td>
<td>41.4% (12/29)*</td>
</tr>
<tr>
<td>8q+</td>
<td>51.7% (15/29)</td>
</tr>
<tr>
<td>9p-</td>
<td>48.3% (14/29)</td>
</tr>
<tr>
<td>10q23-</td>
<td>17.2% (5/29)</td>
</tr>
<tr>
<td>11q-</td>
<td>96.6% (28/29)</td>
</tr>
<tr>
<td>12q14++</td>
<td>7.0% (2/29)**</td>
</tr>
<tr>
<td>13-</td>
<td>34.5% (10/29)</td>
</tr>
<tr>
<td>16p+</td>
<td>48.3% (14/29)</td>
</tr>
<tr>
<td>16q-</td>
<td>62.1% (18/29)</td>
</tr>
<tr>
<td>17q12-21++</td>
<td>17.2% (5/29)**</td>
</tr>
<tr>
<td>17p-</td>
<td>37.9% (11/29)</td>
</tr>
<tr>
<td>18p-</td>
<td>37.9% (11/29)</td>
</tr>
<tr>
<td>20q+</td>
<td>37.9% (11/29)</td>
</tr>
<tr>
<td>22q-</td>
<td>37.9% (11/29)</td>
</tr>
</tbody>
</table>

*) 8p12 gain: 55.2% (16/29)
**) 12q14 gain: 31.0% (9/29)
***) 17q12-21 gain: 41.4% (12/29)
When considering the 11q13 region in detail, four different patterns of amplification could be perceived. In 19/29 (66%) of cases, two or more independent amplicons were evident in concert with a deletion of the remaining q-arm. One continuous amplicon harbouring specific peaks was present in 7/29 (24%) of cases, whereas one narrow, specific peak or one continuous amplicon with no specific peaks could be seen in 1/29 (3%) and 2/29 (7%) of the cases, respectively. In 28/29 (97%) of the tumours, amplification at 11q was followed by deletions at the same chromosomal arm. To identify regions of significant aberrations at 11q13, the GISTIC methodology (Beroukhim et al., 2007) was applied as described below. The resulting chart (Fig 1a) revealed three main amplification cores with central peaks at 66.9Mb, 69.1Mb and 77.0Mb respectively. Altogether, the main amplicon covers the range 65-82Mb that corresponds to cytoband 11q13-q14.1. GISTIC peaks are indicative of regions harbouring driving oncogenes, and were therefore used when candidate genes for expression analyses were selected. Consequently, mRNA expression of RAD9A, S6K2, CCND1, FGF19, FGF4, FGF3, PAK1 and GAB2, was quantified using real-time PCR. The distribution of copy number amplifications/gains in the 11q13-14 region in the 29 samples is presented in Figure 2. To identify possible oncogene candidates, DNA copy number was correlated to relative gene expression using Spearman’s Rank Order correlation (Table 3). A significant correlation between gene copy number and mRNA expression levels could be seen for RAD9A, S6K2, CCND1, PAK1 and GAB2. mRNA expression of FGF19, FGF4 and FGF3 could be detected in 16/23, 8/23 and 13/22 of the samples, respectively. No significant correlation between gene copy number and mRNA expression was seen for these genes.
Figure 1 Comprehensive view of the 11q13 amplicon (a) and the 8p12 amplicon (b) and identification of the most significant amplifications, using the principles of GISTIC methodology (Beroukhim et al., 2007). Positions of candidate oncogenes (previously suggested and/or included in the mRNA expression analysis) were received from the human genome build 36, ver3.
Figure 2 The distribution of copy number amplifications (CNstate ≥ 4, black bars) and gains (CNstate ≥ 3, grey bars) in the 11q13-14 region across the 29 samples.

The CCND1 core correlates to ER expression

Several studies have shown that 11q13 amplification is an event preferably seen in ER positive breast cancers. Consequently, the correlation between ER expression and the three main amplicons at 11q13 according to Figure 1 a, represented by S6K2, CCND1 and PAK1 gene copy numbers was evaluated. In a multiple regression analysis, only the core encompassing CCND1 was significantly correlated to ER expression levels (S6K2: β=0.15, P=0.42; CCND1: β=0.50 P=0.019; PAK1: β=-0.16, P=0.40). The positive correlation between ER expression and CCND1 copy number could also be seen using Spearman’s Rank
order correlation (Spearman’s R=0.48; P=0.009). There was no significant connection between ER levels and mRNA expression of any of the 11q13 located genes included (data not shown).

Table 3 Spearman’s Rank Order correlation estimating the relationship between mRNA expression and DNA copy number for genes centred in each amplification core at 11q13. DNA copy number was assessed by Affymetrix 250K SNParrays and for S6K2, CCND1 and PAK1 also with quantitative real-time PCR as previously described (Bostner et al., 2007; Perez-Tenorio et al., 2010). Real-time PCR copy number data for S6K2, CCND1 and PAK1 were also used for the neighbouring genes in respective core (1-3).

<table>
<thead>
<tr>
<th>Gene (11q13)</th>
<th>Array copy number data</th>
<th>Real-time PCR copy number data</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Spearman R</td>
<td>p-value</td>
</tr>
<tr>
<td><strong>Core 1</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RAD9A</td>
<td>0.47</td>
<td>0.030</td>
</tr>
<tr>
<td>S6K2</td>
<td>0.38</td>
<td>0.093</td>
</tr>
<tr>
<td><strong>Core 2</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CCND1</td>
<td>0.61</td>
<td>0.003</td>
</tr>
<tr>
<td>FGF19</td>
<td>0.08</td>
<td>0.73</td>
</tr>
<tr>
<td>FGF4</td>
<td>-0.16</td>
<td>0.47</td>
</tr>
<tr>
<td>FGF3</td>
<td>0.18</td>
<td>0.44</td>
</tr>
<tr>
<td><strong>Core 3</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PAK1</td>
<td>0.69</td>
<td>0.0005</td>
</tr>
<tr>
<td>GAB2</td>
<td>0.50</td>
<td>0.022</td>
</tr>
</tbody>
</table>

a) Copy number data for S6K2  

b) Copy number data for CCND1  

c) Copy number data for PAK1
Prognostic significance of 11q13/8p12 coamplification and coexpression

In the present study, amplicons in the 8p12 region were found in 55% (16/29) of the samples of which 41% (12/29) were detected as ≥ 4 DNA copies. Patterns of amplification were similar to those seen at 11q13, including recurrent loss of the distal part of the p-arm. When applying GISTIC on the 8p12 region, a main peak centred on 38.4Mb was evident (Fig 1 b). In a next step, the 29 samples were divided into two groups based on 8p12 amplification status and GISTIC was applied on the 11q13 region in the two cohorts separately (Fig 3). Resulting charts reveal that the most proximal amplicon on 11q13 was exclusively found in the 8p12 positive group (Fig 3 a) compared to the 8p12 negative group (Fig 3 b). In addition, the distal core is most distinctive in the 8p12 negative group. Consequently, the correlation between the 8p12 amplicon (here represented by the FGFRI gene) and the three main amplicons at 11q13 according to Figure 1 a, represented by the CCND1 (69.1Mb), PAK1 (76.7Mb) and S6K2 (66.9Mb) genes were investigated. Using multiple linear regression, the proximal amplification core centred on 66.9Mb (S6K2) correlated significantly to the 8p12 amplicon, whereas no significant correlation could be seen for the other cores (S6K2: β=0.54, P=0.004; CCND1: β=-0.057, P=0.77; PAK1: β=-0.13, P=0.47).

The impact of amplification on mRNA expression was evaluated for the genes FGFRI, PPAPDC1B and 4EBP1, located in the main peak at 8p12 (Fig 1 b). Gene copy number estimated from array data significantly correlated to mRNA expression levels for all these genes (Spearman’s Rank Order Correlation FGFRI: R=0.73; P=0.00015, 4EBP1: R=0.53; P=0.013, and PPAPDC1B: R=0.51; P=0.018). When the connection between 8p12 and the proximal amplicon at 11q13 was investigated at the mRNA level, there was a significant correlation between S6K2 mRNA levels and the three genes at 8p12 (Table 4). RAD9A mRNA expression was not significantly associated with expression levels of any of the
Figure 3 GISTIC view of the 11q13 amplicon in the 29 tumours divided into two groups based on 8p12 status: 8p12+ ≥ 3 DNA copies (a), 8p12- < 3 DNA copies (b).
Table 4 Spearman’s Rank Order correlation estimating the relationship between mRNA expression levels for genes centred in each amplification core at 11q13 versus 8p12 (significant p-values in bold)

<table>
<thead>
<tr>
<th>8p12</th>
<th>4EBP1</th>
<th>PPAPDC1B</th>
<th>FGFR1</th>
</tr>
</thead>
<tbody>
<tr>
<td>11q13</td>
<td>Spearman R</td>
<td>p-value</td>
<td>Spearman R</td>
</tr>
<tr>
<td>RAD9A</td>
<td>0.34</td>
<td>0.11</td>
<td>0.36</td>
</tr>
<tr>
<td>S6K2</td>
<td><strong>0.42</strong></td>
<td><strong>0.044</strong></td>
<td><strong>0.52</strong></td>
</tr>
<tr>
<td>CCND1</td>
<td>0.34</td>
<td>0.11</td>
<td>0.27</td>
</tr>
<tr>
<td>FGF19</td>
<td>0.13</td>
<td>0.55</td>
<td>0.33</td>
</tr>
<tr>
<td>FGF4</td>
<td>0.0029</td>
<td>0.99</td>
<td>0.026</td>
</tr>
<tr>
<td>FGF3</td>
<td>0.33</td>
<td>0.13</td>
<td>-0.062</td>
</tr>
<tr>
<td>PAK1</td>
<td>0.027</td>
<td>0.90</td>
<td>0.34</td>
</tr>
<tr>
<td>GAB2</td>
<td>0.30</td>
<td>0.17</td>
<td><strong>0.43</strong></td>
</tr>
</tbody>
</table>

included genes at 8p12, though a positive trend could be seen. mRNA levels correlated significantly between FGFR1 at 8p12 and FGF3, FGF4 and FGF19 at the central 11q13 core. There was also a significant relationship between mRNA levels for PPAPDC1B at 8p12 and GAB2 at the proximal 11q13 core.

Among the 11q13 positive tumours, FGFR1 (8p12) and ERBB2 amplification (17q21) were mutually exclusive events, giving a significant inverse correlation between the number of gene copies (Spearman’s Rank R -0.58; P=0.001). Of the 8p12 located genes, PPAPDC1B mRNA levels were shown negatively correlated to ERBB2 protein (Spearman’s R -0.54, P=0.0078), whereas a trend was shown for an inverse correlation between FGFR1 mRNA and ERBB2 protein (Spearman’s R -0.35, P=0.099). For 4EBP1 mRNA, no connection to ERBB2 protein levels could be detected (Spearman’s R -0.18; P=0.40).
In univariate analysis, 8p12 gain/amplification was significantly associated with an increased risk of distant recurrence in the cohort of patients with 11q13 positive tumours (Fig 4 a). No prognostic significance could be seen for a high PPAPDC1B mRNA expression (Fig 4 b), but a high expression of FGFR1 (Fig 4 c) as well as 4EBP1 mRNA (Fig 4 d) predicted a worse outcome in the patient group. 8p12 gain/amplification and 4EBP1 mRNA expression remained independent prognostic factors when adjusting for tumour size, nodal status and treatments (8p12 P=0.002; PPAPDC1B P=0.57; FGFR1 mRNA P=0.091; 4EBP1 mRNA P=0.0009).

**Figure 4** Distant recurrence-free survival among 11q13 positive patients in relation to; 8p12 gain/amplification (8p12+: ≥ 3 DNA copies, 8p12-: ≤ 2 DNA copies) (a), high mRNA expression of PPAPDC1B (b), FGFR1 (c) and 4EBP1 (d) (the limit for high mRNA levels was defined by the median for each gene).
Discussion

In the present study, SNP arrays were used to evaluate the patterns of 11q13 amplification, as well as genome wide copy number alterations, in 29 breast tumours. The array technology allows studies of copy number alterations with a high resolution compared to conventional cytogenetic and PCR methods. Using 250K NspI arrays, the 11q12-14 (60-90Mb) area including approximately 30 million base pairs, is covered by 2400 SNPs, with an average of 12 kb between each SNP. Previous studies mapping the 11q13 amplicon in breast cancer using FISH, Southern blot, real-time PCR, chromosome CGH and array CGH have suggested four to five independent amplification cores in the area (Karlseder et al., 1994; Bekri et al., 1997; Courjal et al., 1997; Hui et al., 1997; Rodriguez et al., 2004; Kwek et al., 2009). In the present study, several different amplification cores could be detected in the 11q area. The application of GISTIC methodology, allowed summarising DNA copy number changes at 11q13 across all samples, confirming the 63-82Mb region as relevant. This also revealed three main areas of significant amplification. The location of the most proximal peak identified (66.9 Mb) differs from earlier studies (Karlseder et al., 1994) (61-63 Mb). It was recently shown that S6K2, located in this region, is rarely amplified but commonly displays low level gain in breast tumours (Perez-Tenorio et al., 2010) which could be the reason why this region is identified as relevant by GISTIC, but not with other strategies. Although 11q13 amplifications are heterogeneous, a recurrent feature is deletions at the distal part of 11q, which in turn has been connected to adverse prognosis and a malignant phenotype (Launonen et al., 1998; Laake et al., 1999; Gentile et al., 2001). Studies in squamous cell carcinoma cell lines have suggested that 11q13 amplifications are formed through breakage-fusion-bridge cycles, involving an initiating step where distal 11q is lost (Shuster et al., 2000), which may explain the strong coexistence of 11q13 amplifications and loss of distal 11q in breast cancer.
The present study showed a significant correlation between gene copy number and mRNA levels of all selected genes in the amplification cores, with the exception of \textit{FGF3}, \textit{FGF4} and \textit{FGF19}. Regarding \textit{S6K2}, this connection did not reach statistical significance with copy number data from SNP arrays, whereas a strong correlation was found when applying real-time PCR copy number data. In this case, the resolution of the 250K SNP arrays on the \textit{S6K2} area is not perfect and the closest marker is located approximately 35 kb from the gene, why real-time PCR data is probably more reliable.

A connection between ER positivity and 11q13 amplification in breast cancer has been indicated in several studies (Letessier et al., 2006; Elsheikh et al., 2008) and the present study suggests that a core centred on \textit{CCND1} is the most relevant in this context. The expression of \textit{CCND1}, encoding the cell cycle regulator Cyclin D1 is highly regulated by ER (Doisneau-Sixou et al., 2003) and in turn, Cyclin D1 can act as an ER activator and induce its transcriptional function by direct binding to the receptor (Fu et al., 2004). However, several genes located at 11q13 are possibly connected to ER pathways. \textit{PAK1} gene amplification (Bostner et al., 2007) as well as high protein expression (Holm et al., 2006; Bostner et al., 2010) has been linked to insensitivity to endocrine treatment among breast cancer patients. Also \textit{GAB2} expression is suggested to be oestrogen regulated and amplification and overexpression of \textit{GAB2} has been reported in breast cancer cell lines and specimens (Daly et al., 2002; Bentires-Alj et al., 2006; Bocanegra et al., 2010). \textit{S6K2} in the proximal amplification core, encodes an mTOR regulated serine/threonine kinase, initially identified as a homologue to p70S6 kinase (S6K1) (Gout et al., 1998; Saitoh et al., 1998). In a recent study, gene copy gain and nuclear protein expression of \textit{S6K2} were shown to predict tamoxifen response (Perez-Tenorio et al., 2010). Altogether, several genes at 11q13 are potentially connected to ER regulation and signalling.
It is a well known fact that the 11q13 region is commonly coamplified with regions at 8p11-12 in breast cancer and present data suggest that the proximal amplicon at 11q13, harbouring \textit{S6K2}, is the most relevant in this context. Here, three candidate oncogenes in the defined 8p12 core region were selected for expression analysis based on previous reports (Armengol et al., 2007; Bernard-Pierrot et al., 2008; Turner et al., 2010). Expression of \textit{S6K2} correlated significantly to expression of both 4\textit{EBP1} and \textit{PPAPDC1B}. It has been demonstrated that coamplification of 11q13 and 8p12 predicts a worse patient outcome than either amplicon alone (Cuny et al., 2000), indicating that genes in the amplicons can promote tumourigenesis in synergy. In accordance with this, although a small group of patients, the present study shows a prognostic relevance of 8p12 amplification and high expression of the 8p12 located genes \textit{FGFR1} and 4\textit{EBP1} among the patients with 11q13 amplified breast tumours. In breast (Rojo et al., 2007) and ovarian cancer (Castellvi et al., 2006) high expression of phosphorylated 4EBP1 has been correlated to higher grade and poor prognosis. In agreement with several other studies (Ray et al., 2004; Yang et al., 2006; Bernard-Pierrot et al., 2008), the present investigation shows a high correlation between 4\textit{EBP1} amplification and increased mRNA expression in breast cancer. Of note, the correlation between \textit{S6K2} and 4\textit{EBP1} expression has also been observed at the protein level (unpublished data). 4EBP1 is generally considered as a tumour suppressor, due to the ability of unphosphorylated protein to inhibit the translational machinery, regulated by mTOR. However, a recent study has shown that MYC decreases the sensitivity to mTOR inhibition and suppresses autophagy by upregulating 4EBP1 expression (Balakumaran et al., 2010). Phosphorylated 4EBP1 has in turn been suggested to be capable of stimulating mTORC1 signalling through binding to the complex (Wang et al., 2006). We recently identified \textit{S6K2} as a novel candidate oncogene in the 11q13 amplicon, and increased gene copy number was related to a poor outcome among breast cancer patients (Perez-Tenorio et al., 2010). In addition, a functional role of \textit{S6K2}
protein in promoting breast cancer cell survival through activation of AKT have been suggested (Sridharan et al., 2011). The results from the present study further emphasize the relevance of \(S6K2\) as an interesting candidate oncogene, together with \(4EBP1\).

\(FGFRI\) was the first reported oncogene in the 8p12 region, and its amplification and overexpression have been related to poor prognosis and endocrine resistance in breast cancer (Elbauomy Elsheikh et al., 2007; Turner et al., 2010). In the present study \(FGFRI\) mRNA expression was found correlated to expression of the ligands encoded by \(FGF3, 4\) and \(19\) located at the central 11q13 core. \(FGF3\) and \(FGF4\) were the first suggested potential oncogenes mapped to the 11q13 area, discovered as a result of their role in virus induced mouse mammary tumours. However, the genes were discarded as amplicon drivers in human breast cancer, as they seemed to be transcriptionally inactive (Dickson et al., 1995). In the present study, mRNA expression of \(FGF3, 4\) and \(19\) could be detected in a substantial proportion of breast tumours and although no significant connection to gene copy number could be seen, tumours with detectable mRNA levels commonly also showed amplification in the area. A role for the FGFs in the 11q13 amplicon may therefore not be excluded. In the cohort of 11q13 positive tumours, 8p12 and \(ERBB2\) amplification were mutually exclusive events. This is in line with recent data, showing 11q13/8p12 coamplifications to be rare among \(ERBB2\) positive tumours, but common in the \(ERBB2\) negative subgroup, suggesting divergences in cellular pathways driving tumour progression (Staaf et al., 2010). Similar findings were reported recently, when 359 breast cancers where classified into six subtypes based on DNA copy-number aberrations, showing that 11q13/8p12 coamplified cases where enriched in a so called “Amplifier subtype”, which in turn did not include any tumours from the HER2 gene-expression subtype (Jönsson et al., 2010).
In conclusion, SNP arrays have enabled mapping of the 11q13 amplicon in breast tumours with high resolution, revealing several physically independent areas of amplification. With the application of GISTIC methodology, three main core regions could be identified. ER expression is associated with a central core, encompassing $CCND1$. A proximal 11q13 core including $S6K2$ is for the first time identified as involved in the coamplification with 8p12, supported by data at both DNA and mRNA levels. In addition, high 4EBP1 mRNA expression seems to predict a poor prognosis among patients with 11q13 amplified breast tumours, suggesting synergy between the mTOR targets S6K2 and 4EBP1 in breast cancer development and progression.

**Conflict of interest**

The authors declare no financial conflict of interest related to this work

**Acknowledgements**

This study was supported by grants from the Swedish Cancer Foundation and the Swedish Research Council.

**References**


Hui R, Campbell DH, Lee CS, McCaul K, Horsfall DJ, Musgrove EA, Daly RJ, Seshadri R, Sutherland RL. 1997. EMS1 amplification can occur independently of CCND1 or INT-2 amplification at 11q13 and may identify different phenotypes in primary breast cancer. Oncogene 15:1617-1623.


