Activation of multiple cancer-associated genes at the **ERBB2** amplicon in breast cancer

**P Kauraniemi and A Kallioniemi¹**

Laboratory of Cancer Biology and ¹Laboratory of Cancer Genetics, University of Tampere and Tampere University Hospital, 33014 Tampere, Finland

(Requests for offprints should be addressed to P Kauraniemi; Email: paivikki.kauraniemi@uta.fi)

**Abstract**

During the past decade the role of the **ERBB2** (**neu/HER2**) oncogene as an important predictor of patient outcome and response to various therapies in breast cancer has been clearly established. This association of **ERBB2** aberrations with more aggressive disease and poor clinical outcome, together with the high prevalence of such alterations in breast cancer, has also made **ERBB2** an attractive target for therapy. A specific antibody-based therapy, Herceptin, directed against the extracellular domain of the **ERBB2** receptor tyrosine kinase, was recently developed and several clinical trials have shown the therapeutic efficacy of this drug against **ERBB2**-positive breast cancer. However, a relatively large fraction of patients does not benefit from Herceptin treatment, indicating that other factors beyond **ERBB2** itself must influence therapy response in **ERBB2**-positive tumors. It is well known that amplification of the 17q12-q21 region is the most common mechanism for **ERBB2** activation in breast cancer and that it leads to simultaneous activation of several other genes. These co-amplified and co-activated genes may have an impact on disease progression and the clinical behavior of **ERBB2**-positive tumors and thus represent important targets of research. In this paper we discuss the current knowledge on the structure of the **ERBB2** amplicon, the genes involved, and their possible contribution to breast cancer pathogenesis.

*Endocrine-Related Cancer* (2006) 13 39–49

**Introduction**

The **ERBB2** (**neu/HER2**) oncogene, located at chromosome 17q12, is one of the most intensively studied genes in cancer, especially in breast cancer. **ERBB2** encodes for a 185 kDa transmembrane glycoprotein that belongs to the family of epidermal growth factor receptors (EGFRs). Other members of this family include EGFR (**ERBB1**/**HER1**), **ERBB3** (**HER3**), and **ERBB4** (**HER4**) (Stern 2000). Ligand binding to the extracellular domain of these receptors induces homodimer (e.g. EGFR–EGFR) or heterodimer (e.g. EGFR–**ERBB2**) formation leading to the activation of the intracellular tyrosine kinase domain and subsequent signaling cascade (Rubin & Yarden 2001, Järvinen & Liu 2002). Several ligands capable of activating the ERBB receptors have been identified, two major groups being EGF-like ligands binding to EGFR and neuregulins binding to **ERBB3** and **ERBB4** receptors (Pinkas-Kramarski et al. 1997, Riese & Stern 1998, Hynes et al. 2001). Interestingly, no ligand specific for **ERBB2** has been identified, instead **ERBB2** has been shown to be a preferable interaction partner for all other ERBB receptors (Tzahar et al. 1996, Graus-Porta et al. 1997). **ERBB2**-containing heterodimers are long lived and have a particularly high signaling potency, partly due to a slower rate of ligand dissociation and receptor internalization (Graus-Porta et al. 1997, Pinkas-Kramarski et al. 1997, Worthylake et al. 1999, Brennan et al. 2000), thus leading to enhanced activation of downstream signaling pathways, such as the MAP kinase and phosphatidylinositol-3-kinase (PI3-K) pathways (Yarden & Sliwkowski 2001).

From the clinical point of view, amplification of the **ERBB2** oncogene is one of the most relevant genetic aberrations in breast cancer, occurring in 10–34% of cases. In 1987, Slamon and co-workers demonstrated that **ERBB2** amplification is a significant predictor of
both overall survival and time to relapse in breast cancer patients (Slamon et al. 1987). Since then, several studies assessing the relationship between ERBB2 abnormalities and breast cancer outcome have been published with varying results. A recent meta-analysis, summarizing data from 81 studies with 27,161 patients, revealed that in the great majority (90%) of studies either ERBB2 amplification or protein overexpression did indeed correlate with poor outcome of the patients (Ross et al. 2003). The possible role of ERBB2 in predicting response to therapy has also been evaluated in multiple studies. The overwhelming majority of these studies agree that there is a strong association between ERBB2 abnormalities and resistance to tamoxifen therapy as well as sensitivity to anthracycline treatment (Ross & Fletcher 1999, Cooke et al. 2001, Nunes & Harris 2002, Ross et al. 2003). A specific antibody-based therapy, Herceptin, targeted against the extracellular domain of the ERBB2 receptor, was developed by Carter et al. (1992). Clinical trials evaluating the efficacy and safety of Herceptin as a single agent therapy have provided overall response rates ranging from 11.6 to 26% for patients with metastatic ERBB2-positive breast cancer (Baselga et al. 1996, 1999, Cobleigh et al. 1999, Vogel et al. 2001, 2002). These relatively low response rates indicate that, despite the targeted nature, all patients do not benefit from Herceptin treatment and therefore other factors must influence treatment responses in ERBB2-positive tumors.

The most common mechanism for ERBB2 activation in breast cancer is gene amplification. Gene amplification can occur either intrachromosomally, typically as homogeneously staining regions, or extrachromosomally, as cytogenetically visible double minute chromosomes or submicroscopic episomes (Schwab 1998). Several different models, including re-replication, unequal exchange, episome excision, and the breakage-fusion-bridge (BFB) cycle, have been proposed to explain the formation of gene amplification in cancer (Schwab 1998, 1999). The BFB model has been best documented. It is initiated by a double-strand break, typically at chromosomal fragile sites, or telomere erosion and leads to accumulation of multiple copies of a DNA segment organized as inverted head-to-head repeats (Toledo et al. 1992, Coquelle et al. 1997, Hellman et al. 2002). Recently, examination of amplification events at the 17q21 locus revealed a head-to-tail orientation of amplification units, thus indicating that mechanisms other than BFB cycles are involved in the formation of the ERBB2 amplicon (Kuwahara et al. 2004).

In order to search for factors that might influence the clinical behavior of ERBB2-positive breast tumors, especially the response to targeted therapy, one logical approach is look at the ERBB2 locus itself. It is well known that the amplified DNA segment (amplicon) in cancer is often rather large and typically covers multiple genes (Ethier 2003). Such co-amplified genes might have an impact on the phenotype and clinical characteristics of ERBB2-amplified tumors. It is possible that the co-amplified genes contribute to disease progression and clinical behavior, such as response to the Herceptin treatment, and therefore might explain some of the differences currently observed in the clinical management of ERBB2-amplified tumors. Overall, such co-amplified genes represent ideal putative clinical markers that might be used in the future as diagnostic and prognostic markers or as additional targets for therapy. Here we review current knowledge about the ERBB2 amplicon in breast cancer, and the genes involved and how they might contribute to breast cancer pathogenesis. In some cases, the gene nomenclature has changed during the years and different names have been used for the same gene in different publications. For the sake of clarity, we have used gene names given in the original publications but will provide the current official gene symbol in parentheses when applicable. Table 1 provides full gene names and official gene symbols for all known genes located within an approximately 2 Mb region around the ERBB2 oncogene as well as a few more centromeric genes implicated in studies reviewed below.

**Identification of genes co-amplified with ERBB2**

Over the years, multiple genes have been reported to be co-amplified with ERBB2 at the 17q12-q21 chromosomal region in breast cancer. These include previously known oncogenes, such as c-erbA (official gene symbol THRA; van de Vijver et al. 1987), as well as other previously known genes, for example the topoisomerase II α (TOP2A) and retinoic acid receptor α (RARA) genes that were shown to be amplified in a subset of breast tumors with ERBB2 amplification (Keith et al. 1993). Growth factor receptor-bound protein 7 (GRB7) was initially identified through a search for a novel SH2 domain containing proteins (Stein et al. 1994). GRB7 was localized to the ERBB2 locus and subsequently demonstrated to be co-amplified with ERBB2 in breast cancer (Stein et al. 1994). In an effort to identify novel genes involved in
cancer progression, Tomasetto et al. (1995) performed a differential screening using a human metastatic lymph node cDNA library. This approach revealed four novel genes from the ERBB2 region, MLN50 (current official gene symbol LASP1), MLN62, MLN64 (STARD3), and MLN64 (STARD3), that were amplified in breast tumors (Tomasetto et al. 1995, Bie`che et al. 1996). More recently, peroxisome proliferator-activated receptor binding protein (PPARBP), a nuclear receptor coactivator located at

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Alias</th>
<th>Description</th>
<th>Start bp position</th>
</tr>
</thead>
<tbody>
<tr>
<td>TRAF4</td>
<td>MLN62</td>
<td>TNF receptor-associated factor 4</td>
<td>24095173</td>
</tr>
<tr>
<td>TIAF1</td>
<td></td>
<td>TGFβ1-induced anti-apoptotic factor 1</td>
<td>24424665</td>
</tr>
<tr>
<td>MLLT6</td>
<td></td>
<td>Myeloid/lymphoid or mixed-lineage leukemia; translocated to, 6</td>
<td>34115412</td>
</tr>
<tr>
<td>LOC284106</td>
<td></td>
<td>Hypothetical protein LOC284106</td>
<td>34140014</td>
</tr>
<tr>
<td>PCGF2</td>
<td>ZNF144</td>
<td>Polycomb group ring finger 2</td>
<td>34143676</td>
</tr>
<tr>
<td>PSMB3</td>
<td></td>
<td>Proteasome (prosome, macropain) subunit, β type, 3</td>
<td>34162528</td>
</tr>
<tr>
<td>PIP5K2B</td>
<td></td>
<td>Phosphatidylinositol-4-phosphate 5-kinase, type II, β</td>
<td>34177324</td>
</tr>
<tr>
<td>RPL23</td>
<td></td>
<td>Ribosomal protein L23</td>
<td>34259865</td>
</tr>
<tr>
<td>LASP1</td>
<td>MLN50</td>
<td>LIM and SH3 protein 1</td>
<td>34279894</td>
</tr>
<tr>
<td>FLJ43826</td>
<td></td>
<td>FLJ43826 protein</td>
<td>34439685</td>
</tr>
<tr>
<td>PLDXC1</td>
<td></td>
<td>Plexin domain containing</td>
<td>34473083</td>
</tr>
<tr>
<td>ARL12</td>
<td></td>
<td>ADP-ribosylation factor-like 12</td>
<td>34561546</td>
</tr>
<tr>
<td>CACNB1</td>
<td></td>
<td>Calcium channel, voltage-dependent, β1 subunit</td>
<td>34583235</td>
</tr>
<tr>
<td>RPL19</td>
<td></td>
<td>Ribosomal protein L19</td>
<td>34610096</td>
</tr>
<tr>
<td>STAC2</td>
<td></td>
<td>SH3 and cysteine rich domain 2</td>
<td>34620316</td>
</tr>
<tr>
<td>FBXL20</td>
<td></td>
<td>F-box and leucine-rich repeat protein 20</td>
<td>34670367</td>
</tr>
<tr>
<td>PPARBP</td>
<td>TRAP220</td>
<td>PPAR-binding protein</td>
<td>34816380</td>
</tr>
<tr>
<td>CRK7</td>
<td></td>
<td>CDC2-related protein kinase 7</td>
<td>34871818</td>
</tr>
<tr>
<td>NEUROD2</td>
<td></td>
<td>Neurogenic differentiation 2</td>
<td>35014575</td>
</tr>
<tr>
<td>PPP1R1B</td>
<td>DARPP-32</td>
<td>Protein phosphatase 1, regulatory (inhibitor) subunit 1B</td>
<td>35036705</td>
</tr>
<tr>
<td>STARCD3</td>
<td>MLN64</td>
<td>START domain containing 3</td>
<td>35046938</td>
</tr>
<tr>
<td>TCAP</td>
<td></td>
<td>Titin-cap (telethonin)</td>
<td>35075125</td>
</tr>
<tr>
<td>PNMT</td>
<td></td>
<td>Phenylethanolamine N-methyltransferase</td>
<td>35078033</td>
</tr>
<tr>
<td>PERL2</td>
<td>MGC9753</td>
<td>per-like domain containing 1</td>
<td>35089092</td>
</tr>
<tr>
<td>ERBB2</td>
<td></td>
<td>v-erb-b2 erythroblastoid leukemia viral oncogene homolog 2</td>
<td>35109922</td>
</tr>
<tr>
<td>C17orf37</td>
<td>MGC14832</td>
<td>Chromosome 17 open reading frame 37</td>
<td>35138937</td>
</tr>
<tr>
<td>GRB7</td>
<td></td>
<td>Growth factor receptor-bound protein 7</td>
<td>35147744</td>
</tr>
<tr>
<td>ZNINF1A3</td>
<td></td>
<td>Zinc finger protein, subfamily 1A, 3 (Aiolos)</td>
<td>35174724</td>
</tr>
<tr>
<td>ZPBP2</td>
<td></td>
<td>Zona pellucida binding protein 2</td>
<td>35277995</td>
</tr>
<tr>
<td>GSDML</td>
<td></td>
<td>Gasdermin-like</td>
<td>35314376</td>
</tr>
<tr>
<td>ORM1L3</td>
<td></td>
<td>ORM1-like 3</td>
<td>35330822</td>
</tr>
<tr>
<td>GSDM1</td>
<td></td>
<td>Gasdermin 1</td>
<td>35372752</td>
</tr>
<tr>
<td>PSMD3</td>
<td></td>
<td>Proteasome (prosome, macropain) 26S subunit, non-ATPase, 3</td>
<td>35390586</td>
</tr>
<tr>
<td>CSF3</td>
<td></td>
<td>Colony stimulating factor 3 (granulocyte)</td>
<td>35425214</td>
</tr>
<tr>
<td>THRA</td>
<td>c-erbA</td>
<td>Thyroid hormone receptor, α</td>
<td>35472589</td>
</tr>
<tr>
<td>NR1D1</td>
<td></td>
<td>Nuclear receptor subfamily 1, group D, member 1</td>
<td>35502567</td>
</tr>
<tr>
<td>CASC3</td>
<td>MLN51</td>
<td>Cancer susceptibility candidate 3</td>
<td>35550100</td>
</tr>
<tr>
<td>RAPGEFL1</td>
<td></td>
<td>Rap guanine nucleotide exchange factor (GEF)-like 1</td>
<td>35551394</td>
</tr>
<tr>
<td>WIRE</td>
<td></td>
<td>WIRE protein</td>
<td>35666238</td>
</tr>
<tr>
<td>CDC6</td>
<td></td>
<td>CDC6 cell division cycle 6 homolog</td>
<td>35697672</td>
</tr>
<tr>
<td>RARA</td>
<td></td>
<td>Retinoic acid receptor, α</td>
<td>35740896</td>
</tr>
<tr>
<td>GJC1</td>
<td></td>
<td>Gap junction protein, χ 1, 31.9 kDa (connexin 31.9)</td>
<td>35770433</td>
</tr>
<tr>
<td>TOP2A</td>
<td></td>
<td>Topoisomerase (DNA) II α 170kDa</td>
<td>35798321</td>
</tr>
</tbody>
</table>

Genes mentioned in the review are shown in bold. Only those aliases mentioned in this review are listed. The common region of amplification defined by Kauraniemi et al. (2003) is indicated by dashed lines. TNF, tumor necrosis factor; TGF, transforming growth factor; PPAR, peroxisome proliferative activated receptor; CDC2, cell division cycle 2; ORM, orosomucoid; WIRE, WASP-(Wiskostf-Aldrich syndrome protein) interacting protein--related protein.
17q12, was shown to be amplified in a subset of breast tumors (Zhu et al. 1999).

Microarray-based expression profiling has been widely utilized in cancer research, for example in classification of tumors and prediction of patient outcome. The foundation for the gene expression profiling in human breast cancer was set by Perou and co-workers (2000) who examined gene expression patterns in 65 breast adenocarcinomas using an 8102 clone cDNA microarray. A characteristic ‘molecular portrait’ of each breast tumor was obtained. Classification of the tumors using hierarchical cluster analysis revealed four distinct subgroups of samples; luminal like, basal like, normal breast like, and the so-called ERBB2-positive tumor cluster. The ERBB2-positive tumor cluster was characterized by high expression of ERBB2 as well as a group of other genes (e.g. GRB7, MLN64, MLN62, and TIAF1), most of which are located in the close vicinity of ERBB2 at 17q12-q21 (Table 2). The common high level expression pattern observed in the ERBB2-positive tumor cluster is thus likely to be caused by co-amplification of these genes. The existence of distinct breast tumor subtypes, including the ERBB2-positive tumor cluster, has subsequently been confirmed in larger series of patients representing different patient cohorts as well as microarray formats (Sorlie et al. 2003, Yu et al. 2004).

The discovery of an ERBB2-positive tumor subtype has been followed by many studies that have further explored the ERBB2-related gene expression patterns in breast cancer (Table 2). Dressman et al. (2003) performed gene expression profiling to identify genes with expression profiles similar to ERBB2. They analyzed a set of 34 primary breast tumors using oligonucleotide microarrays representing approximately 7128 genes. Expression analysis identified a total of 23 genes showing similar expression pattern with ERBB2, six of which (PSMB3, PNMT, MLN64, RPL19, GRB7, and NR1D1) mapped within a 1.5 Mb region at the ERBB2 locus (Dressman et al. 2003). Moreover, copy number analysis of ERBB2 and two (PNMT and MLN64) of the co-expressed genes in 12 primary breast tumors demonstrated a statistically significant association between relative gene copy numbers and gene expression levels, thus confirming that increased expression of this set of genes is a result of gene amplification at the 17q12-q21 region in breast cancer (Dressman et al. 2003).

Bertucci and co-workers (2004) aimed to identify an ERBB2-specific gene expression signature by evaluating gene expression changes in a large series of 213 primary breast tumors and 16 breast cancer cell lines with known ERBB2 status. To this end, they utilized a cDNA microarray containing over 9000 genes (Table 2). Supervised analysis of the expression profiles produced a list of 36 genes and expressed sequence tags (ESTs) that were differentially expressed between ERBB2-positive and -negative tumor samples and breast cancer cell lines (Bertucci et al. 2004). In addition to ERBB2 itself, this so-called ERBB2-specific gene expression signature contained six overexpressed genes (PSMB3, RPL19, PPARBP, PPP1R1B, GRB7, and NR1D1) located within less than 1 Mb on either side of the ERBB2 oncogene, indicating probable co-amplification (Bertucci et al. 2004).

---

### Table 2 Summary of microarray-based studies reporting on genes with expression patterns similar to ERBB2 at the 17q12-q21 region. The genes implicated in each study are listed using the official gene symbols. Genes observed in at least three of the studies are underlined. See text for further details.

<table>
<thead>
<tr>
<th>Reference</th>
<th>No. of samples</th>
<th>Array format</th>
<th>Official gene symbol</th>
</tr>
</thead>
<tbody>
<tr>
<td>Perou et al. 2000</td>
<td>65 primaries</td>
<td>cDNA microarray</td>
<td>GRB7, STRAD3, TIAF1, TRAF4</td>
</tr>
<tr>
<td>Dressman et al. 2003</td>
<td>34 primaries</td>
<td>Oligoarray</td>
<td>GRB7, NR1D1, PNMT, PSMB3, RPL19, STARAD3</td>
</tr>
<tr>
<td>Bertucci et al. 2004</td>
<td>213 primaries, 16 cell lines</td>
<td>cDNA microarray</td>
<td>GRB7, NR1D1, PPARBP, PPTtr1B, PPARBP, GRB7, RPL19, CASC3, CDC6, PNMT, PSMB3, RPL19, STARAD3</td>
</tr>
<tr>
<td>Kauraniemi et al. 2001</td>
<td>7 cell lines</td>
<td>17q12-q21 specific cDNA microarray</td>
<td>GRB7, LASP1, PPARBP, CASC3, CDC6, PNMT, PSMB3, RPL19, STARAD3</td>
</tr>
<tr>
<td>Clark et al. 2002</td>
<td>9 cell lines</td>
<td>cDNA microarray*</td>
<td>GRB7, LASP1, PPARBP, CASC3, CDC6, NR1D1, PNMT, PSMB3, RPL19, STARAD3</td>
</tr>
<tr>
<td>Willis et al. 2003</td>
<td>12 primaries, 2 cell lines</td>
<td>Oligoarray</td>
<td>GRB7, MLLT6, PPARBP</td>
</tr>
<tr>
<td>Orsetti et al. 2004</td>
<td>22 primaries, 30 cell lines</td>
<td>Genomic array</td>
<td>GRB7, LASP1, PPARBP, CASC3, CDC6, NR1D1, PNMT, PSMB3, RPL19, STARAD3</td>
</tr>
</tbody>
</table>

*Constructed from a cDNA library derived from the BT-474 breast cancer cell line.
Several microarray-based studies have been aimed at directly identifying amplification target genes, including those at the ERBB2 region (Table 2). These studies have typically employed a strategy where copy number information, derived from comparative genomic hybridization (CGH) analyses, and expression data have been combined to search for genes that are both amplified and overexpressed. Kauraniemi and coworkers (2001) were the first to utilize this strategy to specifically search for amplification target genes at the ERBB2 locus. They constructed a custom-made chromosome-17-specific cDNA microarray containing 217 genes or ESTs from the 17q12-q21 region and applied it to copy number and expression analysis in seven breast cancer cell lines. These analyses identified a set of 12 co-amplified and overexpressed transcripts, including six known genes (CDC6, ERBB2, GRB7, MLN51, MLN64, and ZNF144 (official symbol PCGF2)) from the 17q12-q21 region (Kauraniemi et al. 2001). In a similar fashion, Clark et al. (2002) constructed a microarray from an 8000 clone cDNA library derived from the BT-474 breast cancer cell line with well-known genomic alterations, including the ERBB2 amplicon. The custom microarray was then utilized for a combination of CGH and expression analysis to search for putative amplification target genes. Using this technique, Clark and co-workers (2002) were able to identify several genes, such as MLN64, MLN50, PPARBP, and GRB7 that have previously been found to be involved in the ERBB2 amplicon. In addition, they revealed amplification and increased expression of PSMD3 and TRAP100 (official gene symbol TRAP4) as well as a novel gene HSPC209. In another study, a combination of a high resolution expression analysis using Affymetrix 40K GeneChip oligonucleotide microarray and chromosomal CGH analysis in 12 breast cancer specimens and two breast cancer cell lines, revealed four genes (ERBB2, GRB7, MLL76, and PPARBP) showing increased expression in combination with gain of the 17q12-q21 region (Willis et al. 2003). Finally, Orsetti et al. (2004) used a genomic array covering chromosome 17 at 0.5 Mb resolution to define regions of copy number change in 30 breast cancer cell lines and 22 primary tumors. They then explored the expression levels using cDNA microarrays and identified seven genes (LASP1, RPL19, PPARBP, PPP1R1B, MLN64, ERBB2, and GRB7) to be both amplified and overexpressed.

The microarray-based studies reviewed above highlight a set of genes that are co-amplified with ERBB2 and show similar expression patterns (Table 2). Some genes (e.g. GRB7 and MLN64) are consistently involved across all of these studies whereas others are only described in individual reports. It would be tempting to emphasize the importance of the genes that are implicated most frequently. However, it has to be noted that the number of samples analyzed in each study was rather small and therefore the genetic composition of each individual tumor had a major impact on the results. In addition, the microarray formats and, more importantly, their gene contents vary considerably from one study to another. The methods used to analyze the microarray data have also been variable. It is therefore likely that differences in study designs are responsible for the majority of differences observed. In any case, combined results from the gene expression profiling studies strongly suggest that amplification of the 17q12-q21 region in breast cancer does indeed lead to concomitant activation of several genes located adjacent to the ERBB2 oncogene.

Structure of the ERBB2 amplicon in breast cancer

In an attempt to fully characterize the molecular consequences of the 17q12-q21 amplification, a few studies have evaluated the structure and composition of the amplicon centered around the ERBB2 oncogene. These mapping efforts have been able to define a rather small core region of amplification and have also provided plenty of new information on the genes that are activated through amplification at 17q12-q21 in breast cancer (Fig. 1).

In a study by Luoh (2002), an over 300 kb commonly amplified core segment of the 17q12-q21 amplicon was identified in different breast cancer cell lines by Southern blot analysis using the TRAP220 (PPARBP) and TRAP100 (THRAP4) genes as probes. This region was shown to include altogether 13 amplified genes. Expression analysis by multiplex RT-PCR performed in six breast cancer cell lines showed a near-perfect correlation between amplification and overexpression with 11 (CRK7, PPARBP, PNMT, TCAP, MGC9753 (PERLD1), MLN64, ERBB2, GRB7, ZNFN1A3, PSMD3, and TRAP100 (THRAP4)) of these genes (Luoh 2002). The only exceptions were PPARBP, which displayed similar expression levels across all cell lines studied, and the NEUROD2 gene, which only showed low level expression in a single cell line (Luoh 2002).

Kauraniemi and co-workers (2001) applied fluorescence in situ hybridization using a panel of large insert size genomic clones to characterize the structure of the 17q12-q21 amplicon in breast cancer cell lines. Copy number analysis across the amplified region in 16 breast
cancer cell lines provided detailed information on the structure of the amplicon and allowed definition of a minimal common region of amplification that was restricted to a single bacterial artificial chromosome (BAC) clone (Kauraniemi et al. 2001). In a follow-up study, similar evaluation of the amplicon structure was carried out in a large set of 330 primary breast tumors using the tissue microarray technology (Kauraniemi et al. 2003). Results from this study confirmed the presence of a common minimal region of amplification centered around a single BAC clone. This 280 kb minimal region of amplification is located within the region implicated by Luoh (2002) and its presence was further validated by a recent CGH microarray study that revealed an almost identical minimal common region of amplification in a breast cancer mouse model (Hodgson et al. 2005). The 280 kb region contains ten transcript, eight representing known genes (ERBB2, GRB7, MLN64, PPP1R1B, PNMT, NEUROD2, TCAP, and ZNFN1A3) and two hypothetical proteins (MGC9753 (PERLD1) and MGC14832 (C17orf37)) (Kauraniemi et al. 2003)(Table 1 and Fig. 1). Expression analysis by quantitative real-time RT-PCR demonstrated a statistically significant correlation between amplification and increased expression, thus making them attractive amplification target genes. However, it is quite possible that this correlation is simply due to a bystander effect, i.e. that these genes happen to be located next to the ERBB2 oncogene, and that the increased expression presents no benefit for the cancer cells. An additional insight provided by the amplicon mapping studies reveals that most of the previously reported co-amplified and co-expressed genes, such as THRA, RARA, TOP2A, MLN62 (TRAF4), MLN50 (LASP1), PSMB3, and RPL19, are located outside the minimal common region of amplification (Table 1). However, it has to be noted that in most tumors the amplicon does indeed span beyond the minimal common region of involvement and therefore genes located outside this region might still have some relevance to breast cancer pathogenesis. Overall, it will be extremely important to study the functional consequences of elevated expression to further clarify the possible role of the co-amplified genes in breast cancer progression. We review below what is currently known about the function of the most promising 17q12-q21 amplification candidate target genes, i.e. those located within the minimal region of amplification or its immediate vicinity.

Candidate target genes at the 17q12-q21 amplicon

The amplicon mapping studies reviewed above pinpoint a rather small (280 kb) minimal common region of amplification at the ERBB2 locus (Kauraniemi et al. 2001, 2003, Luoh 2002). Despite its small size, the amplicon was demonstrated to contain a number of genes whose expression was consistently elevated by amplification. In addition to ERBB2 itself, GRB7, MLN64, MGC9753 (PERLD1), and MGC14832 (C17orf37) displayed the most tight correlation between amplification and increased expression, thus making them attractive amplification target genes. However, it is quite possible that this correlation is simply due to a bystander effect, i.e. that these genes happen to be located next to the ERBB2 oncogene, and that the increased expression presents no benefit for the cancer cells. An additional insight provided by the amplicon mapping studies reveals that most of the previously reported co-amplified and co-expressed genes, such as THRA, RARA, TOP2A, MLN62 (TRAF4), MLN50 (LASP1), PSMB3, and RPL19, are located outside the minimal common region of amplification (Table 1). However, it has to be noted that in most tumors the amplicon does indeed span beyond the minimal common region of involvement and therefore genes located outside this region might still have some relevance to breast cancer pathogenesis. Overall, it will be extremely important to study the functional consequences of elevated expression to further clarify the possible role of the co-amplified genes in breast cancer progression. We review below what is currently known about the function of the most promising 17q12-q21 amplification candidate target genes, i.e. those located within the minimal region of amplification or its immediate vicinity.

Growth factor receptor-bound protein 7 (GRB7) encodes for an adaptor-type signaling protein that binds to a variety of cell surface receptor tyrosine kinases via its SH2 domain (Pero et al. 2002, Shen & Guan 2004). Interaction between GRB7 and
multiple cell surface receptors is known to mediate signal transduction to diverse downstream signaling pathways (Shen & Guan 2004). Several studies have demonstrated that GRB7 plays a key role in cell migration through its association with focal adhesion kinase, phosphoinositides, ephrin receptor EphB1, and calmodulin (Han et al. 2000, 2002, Shen et al. 2002, Li et al. 2005), thus implying a likely role in tumor progression. For example, GRB7 expression has been associated with an invasive phenotype and metastatic progression of tumor cells in esophageal carcinoma (Tanaka et al. 1997, 2000). In addition, GRB7 has also been shown to bind ERBB2 with high affinity, indicating functional co-operation between these two proteins and suggesting that this synergistic action may contribute to tumor progression (Stein et al. 1994). Taken together, GRB7 has been suggested as an attractive target for therapeutic intervention in cancer (Pero et al. 2003, Li et al. 2005).

STARD3 steroidogenic acute regulatory protein ((START) domain containing 3), also known as MLN64/CAB1, is a transmembrane protein containing a START domain which is found in a wide variety of proteins involved in lipid transport and metabolism, signal transduction, and transcriptional regulation (Ponting & Aravind 1999). STARD3 shows significant homology with START, a key player in steroid hormone biosynthesis due to its ability to facilitate cholesterol import into the mitochondria (Watari et al. 1997, Moog-Lutz et al. 1997). STARD3 has also been shown to bind cholesterol through its START domain and to enhance steroidogenesis (Watari et al. 1997, Tsujishita & Hurley 2000). It is therefore possible that STARD3 overexpression in cancer cells increases steroid hormone production, thereby promoting the growth of hormone-responsive tumors such as breast cancer (Akiyama et al. 1997, Alpy et al. 2001). It was recently shown that STARD3 and ERBB2 genes are likely to be co-regulated at transcriptional level, since the promoters of both genes are positively regulated by Sp1 transcription factors (Alpy et al. 2003). Co-localization of such non-homologous genes with common transcriptional control has recently been reported and might implicate similar functions or involvement in common biological processes (Cohen et al. 2000, Boutanaev et al. 2002).

Other putative target genes at the 17q12-q21 amplicon include PPP1R1B, and PARBP genes. PPP1R1B (protein phosphatase 1, regulatory (inhibitor) subunit 1B; also known as DARPP-32) has an important role in regulating the efficacy of dopaminergic neurotransmission in the brain (Fienberg et al. 1998). Depending upon which particular amino acid residue is phosphorylated, PPP1R1B can function either as a protein kinase A or as a protein phosphatase-1 inhibitor (Bibb et al. 1999). Over-expression of DARPP-32 and its truncated isoform t-DARPP-32 has been shown to occur in gastric carcinoma (El-Rifai et al. 2002). In a recent study, Beckler and co-workers (2003) showed that DARPP-32 and t-DARPP-32 are also frequently overexpressed in breast, prostate, colon, and stomach carcinomas, suggesting that these proteins may have an important role in tumorigenesis. PPARBP (also known as TRAP220) is a nuclear co-activator that interacts with several receptors, e.g. vitamin D receptor, retinoid acid receptor α, peroxisome proliferator-activated receptor α and γ, thyroid hormone receptor, and estrogen receptor α, in a ligand-dependent manner, activating the transcription of the genes involved, e.g. in cell growth, differentiation, and neoplastic conversion (Yuan et al. 1998, Zhu et al. 1999, Misra et al. 2002). PPARBP has also been shown to regulate p53-dependent apoptosis through mouse double minute 2 homolog (MDM2) activation (Frade et al. 2002). Thus, based on its function, inappropriate activation of PPARBP in cancer might contribute to tumor progression.

In addition to known genes, the minimal common region of amplification defined by Kauraniemi and co-workers (2003) contained two uncharacterized hypothetical proteins (MGC9753 and MGC14832) whose expression was statistically highly correlated with amplification in primary breast tumors. The hypothetical protein MGC14832 (currently known as C17orf37, other names ORB3/XTP4) does not show sequence similarity to any other known genes or proteins, leaving the function of this gene/protein currently unknown. On the other hand, MGC9753 (currently known as per1-like domain containing 1 (PERLD1; other names CAB2/PERLD1) encodes for a seven-transmembrane receptor with an extracellular N-terminal six-cysteine domain (Katoh & Katoh 2003). PERLD1 was shown to represent the human homologue of the yeast COS16 gene involved in the regulation of Mn$^{2+}$ homeostasis (Nezu et al. 2002). It is possible that overexpression of PERLD1 alters the intercellular Mn$^{2+}$ levels and thereby impairs the function of Mn$^{2+}$-dependent enzymes, such as those involved in the repair of double-strand breaks, leading to increased genomic instability (Nezu et al. 2002).

**Summary**

Amplification and overexpression of the ERBB2 oncogene is one of the most relevant genetic aberrations...
in breast cancer and has clinical utility both as a prognostic marker and as a predictive factor. \textit{ERBB2} aberrations have been clearly linked to poor patient outcome and resistance to hormonal therapy as well as sensitivity to anthracycline-based chemotherapy (Ross \& Fletcher 1999, Cooke \textit{et al.} 2001, Nunes \& Harris 2002, Ross \textit{et al.} 2003). The use of targeted therapy against the \textit{ERBB2} oncogene has further increased the clinical importance of this gene. However, there is still much to learn about \textit{ERBB2} and its function in breast cancer. It is apparent that there are other genes that influence the clinical behavior of \textit{ERBB2}-amplified tumors and one of the most obvious candidates are genes that are co-amplified with \textit{ERBB2}.

The studies reviewed above reveal a multitude of genes that are co-localized and co-amplified with \textit{ERBB2} and show similar expression profiles in breast cancer. A minimal common region of amplification spanning only about 280 kb and containing a small set of genes with tightly associated copy number and expression patterns has been defined. Although it cannot be assumed that all of these genes have an important and independent role in cancer pathogenesis, some of them might indeed actively contribute to cancer development. It is also possible that the co-ordinated effect of several overexpressed genes provides a growth advantage for the cancer cells. It has also to be noted that several genes, located outside the minimal region of amplification, showed a consistent pattern between amplification and expression. Since in most of the tumors the amplicon extends beyond the minimal region, the role of these genes might also be important and they cannot be ruled out simply based on their location. In order to fully reveal the contribution of all the possible candidate target genes of the 17q12-q21 amplification, studies exploring the functional consequences of overexpression of these genes, both individually and in combinations, are required.

\section*{Acknowledgements}

This work was partly supported by the Academy of Finland, the Pirkanmaa Cultural Foundation, and the Maud Kuistila Foundation. The authors declare that there is no conflict of interest that would prejudice the impartiality of this scientific work.

\section*{References}


Graus-Porta D, Beerli RR, Daly JM & Hynes NE 1997 ErbB-2, the preferred heterodimerization partner of all ErbB receptors, is a mediator of lateral signaling. *EMBO Journal* **16** 1647–1655.

Han DC, Shen TL & Guan JL 2000 Role of Grb7 targeting to focal contacts and its phosphorylation by focal adhesion kinase in regulation of cell migration. *Journal of Biological Chemistry* **275** 28911–28917.


Luoh SW 2002 Amplification and expression of genes from the 17q11 approximately q12 amplicon in breast cancer cells. *Cancer Genetics and Cytogenetics* **136** 43–47.


Vogel C, Cobleigh MA, Tripathy D, Guthell JC, Harris LN, Fehrenbacher L, Slamon DJ, Murphy M, Novotny WF, Burchmore M et al. 2001 First-line, single-agent


Yuan CX, Ito M, Fondell JD, Fu ZY & Roeder RG 1998 The TRAP220 component of a thyroid hormone receptor-associated protein (TRAP) coactivator complex interacts directly with nuclear receptors in a ligand-dependent fashion. *PNAS* **95** 7939–7944.
