Introduction

An important mechanism for the activation of oncogenes in human cancers is gene amplification, which results in gene overexpression at both the message and the protein levels (1, 2). Oncogenes, such as ERBB2 at 17q12, CCND1 at 11q13, and C-MYC at 8p24, have previously been identified as amplification targets linked to the development, progression, or metastasis of human cancers, including breast, prostate, lung, and other cancers (2, 3). ERBB2 is the most frequently amplified oncogene in breast cancer, and its overexpression is associated with poor clinical outcomes. The prognostic and predictive values of ERBB2 amplification and overexpression have been used to guide treatment decisions for patients with both lymph node–positive and -negative diseases. More significantly, recognition of the mechanistic roles of ERBB2 in breast cancer has led to the development of ERBB2-targeting drugs such as trastuzumab to treat breast cancer (4–6).

Amplification of 8p11-12 occurs in approximately 15% of human breast cancers, and this region of amplification is significantly associated with disease-specific survival and distant recurrence in breast cancer patients (7–11). Chin et al. performed an analysis of the association of gene amplification and disease-free survival and distant relapse in human breast cancer specimens (12). They identified 23 genes from the 8p11-12 region as being correlated with progression. Recently, our laboratory published results of a detailed analysis of copy number and gene expression in the 8p11-12 region in a panel of breast cancer cell lines and primary human breast cancers (10). We identified 21 genes that are overexpressed when their copy number is increased (10). Furthermore, we directly tested the transforming function of eight 8p11-12 amplified genes in human mammary epithelial cells. From these experiments, we identified several genes, including LSM1, BAG4, and C8orf4 (TC-1), as having the transforming properties in vitro (10, 13, 14). Accumulating evidence suggests that the 8p11-12 amplicon contains multiple candidate oncogenes that could play a role in breast cancer development (7–11).

Recent extensive genomic analyses and siRNA knockdown studies have identified the Wolf-Hirschhorn syndrome candidate 1-like 1 gene (WHSC1L1, also known as NSD3) as one of the major candidate oncogenes of the 8p11-12 amplicon in breast cancer (7–11). WHSC1L1 is the third member of a gene family that includes NSD1 and WHSCI (NSD2; refs. 15, 16). De novo translocation of NSD1 genes causes the childhood overgrowth syndrome, Sotos syndrome, that is associated with elevated risks of cancer, whereas de novo deletion of NSD2/WHSCI causes the Wolf-Hirschhorn syndrome that displays growth retardation (17, 18). WHSC1L1/NSD3, NSD1, and WHSCI/NSD2 show strong sequence similarity and share multiple functional domains (15). WHSC1L1 has two isoforms that are derived from alternative splicing of exons 10, and both WHSC1L1 protein isoforms contain a PWWP domain.
domain. The PWPP domain belongs to the royal superfamily that includes chromodomain, tudor, malignant brain tumor, and plant agent motifs, and these domains exist in multiple histone modifying proteins. The NH2-terminal half of the PWPP domain exhibits a β-barrel structure that resembles a SAND domain, whereas the COOH-terminal portion is made up of a 5-helix bundle. Both the crystal and nuclear magnetic resonance (NMR) solution structures of the superfamily complexes show that the β-barrel structure recognizes and binds the histone lysine pocket (19, 20). A study on PWPP function in the DNA methyltransferase DNMT3B and binds the histone lysine pocket (19, 20). A study on PWPP function in the DNA methyltransferase DNMT3B showed that the PWPP domain binds methylated DNA (21). Recently, Wang et al. showed that a PWPP domain protein binds to histone lysine in vitro and in vivo, and regulates Set9-mediated H4K20 methylation (22). Their results showed that the PWPP domain is a methyl-lysine recognition motif that plays important roles in epigenetic regulation.

In the present study, we systematically investigated the transforming properties of 13 newly identified 8p11-12 candidate oncogenes in vitro. We found that WHSC1L1, DDHD2, and ERLIN2 are the most potently transforming oncogenes we tested from the 8p11-12 region based on the number of altered phenotypes expressed by the cells. Knockdown of WHSC1L1 in 8p11-12–amplified breast cancer cells resulted in profound loss of growth and survival of these cells. Further, we identified several WHSC1L1 target genes, one of which is irioquois homeobox 3 gene (IRX3), a member of the Iroquois homeobox transcription factor family.

Materials and Methods

Genomic array comparative genomic hybridization (CGH)

The isolation and culture of the SUM series of human breast cancer cell lines and MCF10A cells have been described in the Supplementary Materials and Methods (10, 23). Genomic array CGH experiments were performed using the Agilent 4K human genome CGH microarray chip (Agilent Technologies). Agilent’s CGH Analytics software was used to calculate various measurement parameters, including log2 ratio of total integrated Cy-5 and Cy-3 intensities for each probe.

Semi quantitative reverse transcriptase-PCR reactions

Total RNA was prepared from human breast cancer cell lines and the MCF10A cell line by standard methods (10, 24). For reverse transcriptase-PCR (RT-PCR) reactions, RNA was converted into cDNA via a reverse transcription reaction using random hexamer primers. Primers were ordered from Invitrogen, and all the relevant primer sequences are available on request. A GAPDH primer set was used as a control. Semiquantitative RT-PCR was done using the iQSYBR Green Supermix (Bio-Rad).

Lentivirus construction and transduction of cells

The lentiviral expression constructs containing the 13 genes tested in the present experiments, listed in Table 1, were established as previously described (10). Briefly, we first created entry clones from cDNA of SUM-44 cells using the pENTR directional TOPO cloning kit and then performed the LR recombination reaction to transfer the gene into the Gateway destination vector, pLenti6/V5-DEST. Specifically, the pLenti-WHSC1L1 construct was established from the full-length short isoform, which only contained the PWPP domain. The lentivirus for each construct was generated and used to infect immortalized, nontransformed mammary epithelial MCF10A cells. Control infections with pLenti-LacZ virus were performed in parallel. Selection began 48 hours after infection in growth medium with 10 μg/mL blasticidin in the absence of insulin. Upon confluence, selected cells were passaged and serially cultured.

Growth in soft agar and Matrigel

Soft agar assays were performed as previously described (10). For three-dimensional (3D) morphogenesis assays in Matrigel, cells grown in monolayer culture were detached by trypsin/EDTA treatment and seeded in Matrigel (BD Biosciences) precoated 8-well chamber slides. The appropriate volume of medium was added and maintained in culture for 10 to 18 days. Phase-contrast images and immunostained images were photographed with bright-field and confocal microscopy (25).

Lentivirus-mediated short hairpin RNA (shRNA) knockdown of gene expression

We knocked down the expression of the human WHSC1L1 gene in the breast cancer cell lines SUM-44 and SUM-52 and in the MCF10A cell line using the Expression Arrest GIPZ lentiviral shRNAAmir system (OpenBiosystems). Lentivirus was produced by transfecting 293FT cells with the combination of the lentiviral expression plasmid DNA and Trans-Lentiviral packaging mix (OpenBiosystems). For cell infection, viral supernatants were supplemented with 6 μg/mL polybrene and incubated with cells for 24 hours. Cells expressing shRNA were selected with puromycin for 2 to 3 weeks for functional studies (cell proliferation and colony formation assays) and for 4 to 10 days after infection for RNA extraction.

Results

The effect of different 8p11-12 genes on growth factor–independent proliferation

Recently, our group identified 21 candidate oncogenes within the 8p11-12 ampiclon in breast cancer based on statistical analysis of copy number increase and gene overexpression. We tested 8 of the 21 candidate oncogenes for transforming function in vitro and identified three genes, namely, LSM1, BAG4, and C8orf4 (TC-1), that could induce transformed phenotypes (10). In the present report, we expanded our analysis to the remaining 13 candidate oncogenes. Table 1 shows the original 21 gene list, with the 13 genes tested in the present experiments highlighted with an asterisk. Details on the origins and sequence validations of each clone are given in Materials and Methods and in Supplementary Data.
To systematically investigate the transforming properties of 13 8p11-12 candidate breast cancer oncogenes, we transduced MCF10A cells, which are highly growth factor dependent, with individual lentiviral expression vectors for each gene. Growth factor–independent proliferation of MCF10A cells transduced with each candidate gene was investigated. RT-PCR was performed to confirm the expression of the gene using primers specific for the gene and for the vector. Overexpression of WHSC1L1 protein in MCF10A-WHSC1L1 cells was further confirmed by Western blot (Supplementary Fig. S1). As shown in Fig. 1A, MCF10A cells expressing six genes, namely, WHSC1L1, DDHD2, PROSC, BRF2, ASH2L, and ERLIN2, formed expanding colonies in insulin-free medium, and then grew continuously in the absence of insulin-like growth factors. Colony formation assays in MCF10A cells with equalized viral titer of the tested genes indicated that overexpression of WHSC1L1 and DDHD2 resulted in the highest number of insulin-independent colonies. Growth curves of MCF10A cells overexpressing the five genes (WHSC1L1, DDHD2, PROSC, BRF2, and ERLIN2) were performed within two passages of isolation in insulin-free medium. Data in Fig. 1B show that overexpression of WHSC1L1 not only resulted in the largest number of colonies emerging in insulin-free medium, but also gave rise to cells with the most rapid proliferation rate under these conditions. These results extend our previous findings and indicate that a total of nine genes from the 8p11-12 have the ability to induce insulin-like growth factor–independent proliferation when overexpressed in MCF10A cells.

### Table 1. List of the 21 candidate genes of the 8p11-12 region

<table>
<thead>
<tr>
<th>Gene</th>
<th>Description</th>
<th>Tested in the current study.</th>
</tr>
</thead>
<tbody>
<tr>
<td>ZNF703</td>
<td>Zinc finger protein 703</td>
<td></td>
</tr>
<tr>
<td>ERL12*</td>
<td>ER lipid raft associated 2</td>
<td></td>
</tr>
<tr>
<td>PROSC*</td>
<td>Proline synthetase cotranscribed homolog (bacterial)</td>
<td></td>
</tr>
<tr>
<td>BRF2*</td>
<td>BRF2, subunit of RNA polymerase III transcription initiation factor, BRF1-like</td>
<td></td>
</tr>
<tr>
<td>RAB11FIP1</td>
<td>Rab coupling protein = RCP</td>
<td></td>
</tr>
<tr>
<td>EIF4EBP1</td>
<td>Elongation factor 4 binding protein 1</td>
<td></td>
</tr>
<tr>
<td>ASH2L*</td>
<td>Ash2 (absent, small, or homeotic)-like (Drosophila)</td>
<td></td>
</tr>
<tr>
<td>LSM1</td>
<td>LSM1 homolog, U6 small nuclear RNA associated (S. cerevisiae)</td>
<td></td>
</tr>
<tr>
<td>BAG4</td>
<td>BCL2-associated athanogene 4</td>
<td></td>
</tr>
<tr>
<td>DDHD2*</td>
<td>DDHD domain containing 2</td>
<td></td>
</tr>
<tr>
<td>PPAPDC1B*</td>
<td>Phosphatidic acid phosphatase type 2 domain containing 1B</td>
<td></td>
</tr>
<tr>
<td>WHSC1L1*</td>
<td>Wolf-Hirschhorn syndrome candidate 1-like 1</td>
<td></td>
</tr>
<tr>
<td>LETM2*</td>
<td>Leucine zipper-EF-hand containing transmembrane protein 2</td>
<td></td>
</tr>
<tr>
<td>FGFR1</td>
<td>Fibroblast growth factor receptor 1 (fms-related tyrosine kinase 2, Pfeiffer syndrome)</td>
<td></td>
</tr>
<tr>
<td>TACC1</td>
<td>Transforming acidic coiled-coil</td>
<td></td>
</tr>
<tr>
<td>TM2D2*</td>
<td>TM2 domain containing 2</td>
<td></td>
</tr>
<tr>
<td>C8orf4</td>
<td>Chromosome 8 open reading frame 4 = TC-1</td>
<td></td>
</tr>
<tr>
<td>AP3SM2*</td>
<td>Adaptor-related protein complex 3, μ 2 subunit</td>
<td></td>
</tr>
<tr>
<td>POLB*</td>
<td>Polymerase (DNA directed), β</td>
<td></td>
</tr>
<tr>
<td>VDAC3*</td>
<td>Voltage-dependent anion channel 3</td>
<td></td>
</tr>
<tr>
<td>HOOK3*</td>
<td>Hook homolog 3 (Drosophila)</td>
<td></td>
</tr>
</tbody>
</table>

*Tested in the current study.

### Transforming properties of 8p11-12 candidate genes

To assess the expression of other transformed phenotypes of MCF10A cells overexpressing the newly identified candidate oncogenes, we evaluated each of them for their ability to form colonies in soft agar and for altered morphogenesis in Matrigel. Figure 2A shows that after three weeks in culture, MCF10A cells overexpressing WHSC1L1, DDHD2, and ERLIN2 formed colonies in soft agar. MCF10A cells overexpressing WHSC1L1 and DDHD2 had the highest soft agar colony-forming efficiency (Fig. 2B). By contrast, MCF10A cells overexpressing ASH2, BRF2, and PROSC did not form soft agar colonies. We also examined whether these six genes affect the growth or morphology of colonies in 3D Matrigel culture, as aberrant behavior in this environment is frequently associated with transformation and/or tumor progression (26). In 3D basement membrane cultures, the immortalized, nontransformed mammary epithelial cells, MCF10A, formed acinar-like structures consisting of a single cell layer of polarized, growth-arrested mammary epithelial cells surrounding a hollow lumen (Fig. 2A). As shown in Fig. 2A, MCF10A cells overexpressing WHSC1L1 formed strikingly abnormal acini that were enlarged and disorganized, and contained filled lumens. In contrast, MCF10A cells overexpressing DDHD2 formed disorganized, small abnormal acini. MCF10A cells overexpressing ERLIN2 also formed large, highly proliferative colonies, whereas insulin-independent MCF10A cells overexpressing the other three candidate oncogenes formed polarized, growth-arrested acinar structures with hollow lumens.
similar to MCF10A parental cells (data not shown). These experiments show that three of the transforming genes, i.e., PROSC, ASH2L, and BRF2, induced insulin-independent growth and no other altered phenotypes. By contrast, WHSC1L1, DDHD2, and ERLIN2 were the most transforming oncogenes based on the number of altered phenotypes expressed by the cells.

**Amplification of WHSC1L1 isoforms in breast cancer**

In our analysis of the transforming properties of the 8p11 candidate oncogenes, we were surprised by the potency of WHSC1L1 for transforming MCF-10A cells. As described above, WHSC1L1-overexpressing MCF-10A cells exhibited the highest transforming efficiency. The cells had the highest growth rate in insulin-free medium, and the cells grew with high efficiency in soft agar, while forming very abnormal colonies in Matrigel. Because of the extraordinary transforming potency of WHSC1L1, the role of this gene as a driver oncogene in breast cancer cell lines and specimens with the amplicon was examined further.

Expression of the WHSC1L1 gene results in two alternatively spliced variants: a long isoform and a short isoform that are derived from alternative splicing of exon 10. The WHSC1L1 long isoform encodes a 1437-amino-acid protein containing two PWPP domains, two PHD-type zinc finger motifs, a TAN2 domain, an AWS domain, and a SET domain. The short isoform encodes a 645-amino-acid protein containing a PWPP domain only (Supplementary Fig. S2). The data shown in Fig. 3 show that both SUM-44 and SUM-52 cells have amplifications of the full-length gene, but at the protein level expression of the short isoform predominates. The transformation data for WHSC1L1-overexpressing MCF10A cells shown above were obtained using an expression construct coding for the short isoform, and similar results were obtained when we transduced MCF10A cells with a vector coding for full-length WHSC1L1 (data not shown).
Alternative splicing in cancer is an important mechanism for gene regulation and for generating proteomic diversity. Interestingly, we identified one primary breast cancer specimen (10173A) with the 8p11-12 amplicon in which array CGH showed genomic loss of the COOH-terminal region of the WHSC1L1 long isoform but with amplification of exons 1 to 10. We validated this finding in that particular breast cancer specimen by genomic PCR using primers specific for the short isoform exon 10 (S-10) and the long isoform exon 20 (L-20) as shown in Fig. 3C and Supplementary Fig. S2. To further determine whether the WHSC1L1 short isoform protein, which only contains a PWWP domain, is also localized in the nucleus, we generated expression constructs containing the short isoform WHSC1L1 coding sequences fused to the EGFP epitope at the COOH-terminus. The constructs were transfected into MCF10A and HEK293 cells, and localization of the proteins was examined by fluorescence microscopy. The WHSC1L1 short isoform was localized to the nucleus as expected (Supplementary Fig. S3). These results indicate that both WHSC1L1 protein isoforms are localized to the nucleus, and may act as transforming oncoproteins in breast cancer cells bearing the 8p11p12 amplicon.

**Knockdown of WHSC1L1 inhibits cell proliferation in breast cancer cells**

To directly assess the contribution of endogenous WHSC1L1 overexpression on the transformation of human breast cancer, we examined the effects of knockdown of WHSC1L1 in SUM-44 and SUM-52 cells where WHSC1L1 is amplified and overexpressed, in SUM-149 cells that do not have the amplicon, and in the control cell line MCF10A. To perform shRNA knockdown experiments, we obtained eight pGIPZ-WHSC1L1

![Image](https://example.com/image.png)

**Figure 2.** A, top, representative pictures of MCF10A cells that stably overexpress DDHD2 and WHSC1L1 genes and control (Ctrl) cell soft agar colonies. Cells were grown for three weeks in soft agar and stained with the vital dye p-iodonitrotetrazolium violet. Bottom, effects of DDHD2 and WHSC1L1 on mammary acinar morphogenesis. MCF10A-DDHD2, WHSC1L1, and control cells were cultured on a bed of Matrigel as described in Materials and Methods. Representative images of structures with staining for actin with phalloidin conjugated to Alexa Fluor-568 (red), and 4′, 6-diamidino-2-phenylindole (DAPI) as a marker of nuclei (blue). B, soft agar colony-forming efficiency in MCF10A cells stably overexpressing the six genes (WHSC1L1, BRF2, DDHD2, PROSC, ERLIN2, and ASH2L) and control cell soft agar colonies. Data are the mean number of colonies per low power field three weeks after seeding 10^5 cells per well.
shRNA expression constructs from OpenBiosystems. In this vector, TurboGFP and shRNA are part of a bicistronic transcript allowing the visual marking of the shRNA-expressing cells. SUM-44, SUM-52, and control MCF10A cells were infected with these 8-shRNA lentivirus supernatants, pooled or individually, to determine which gave the best knockdown of WHSC1L1. Nonsilencing shRNAmir lentiviral control, at the same titer as WHSC1L1 shRNA, was used in parallel as the negative control. First, the consequence of knockdown of WHSC1L1 on colony formation using all eight shRNAs was evaluated in all three cell lines. WHSC1L1 knockdown suppressed proliferation of SUM-44 and SUM-52 cells, whereas WHSC1L1 shRNAs had no effect on the growth of SUM-149 cells or MCF10A cells (Supplementary Fig. S4). Next, we identified the two most efficient shRNAs with respect to knockdown of WHSC1L1 expression levels in SUM-44 and SUM-52 cells. Quantitative RT-PCR and Western blot data revealed that the WHSC1L1-shRNAs 2 and 6 resulted in decreases in mRNA and protein levels to approximately 20% to 30% of the level seen in the nonsilencing control-infected cells (Fig. 4A). As shown in Fig. 4B and C, WHSC1L1 knockdown with both shRNA constructs slowed cell growth of SUM-44 and SUM-52 cells. The results were most striking for SUM-44 cells in which WHSC1L1 knockdown inhibited cell proliferation by ~90% (Fig. 4C). WHSC1L1 knockdown with shRNA 2 and 6 had an undetectable effect on the cell growth of MCF10A cells (data not shown). Thus, knockdown of WHSC1L1 inhibits cell proliferation in breast cancer cells with WHSC1L1 gene amplification.

IRX3 is a novel target gene of WHSC1L1

Because WHSC1L1 encodes a PWWP domain nuclear protein that has histone methyl transferase activity, it has been postulated that it can promote malignant transformation by altering the histone code and hence expression of specific target genes. To identify genes that may be altered in their expression by overexpression of the short isoform of WHSC1L1, we performed expression profiling of MCF10A cells, MCF10A-WHSC1L1 cells, and SUM-44 cells. To identify genes most likely to be regulated by overexpression of WHSC1L1 and relevant to human breast cancer, we determined which genes are differentially expressed in MCF10A-WHSC1L1 cells relative to parental MCF10A cells, and then determined which of those genes are also differentially expressed in SUM-44 cells compared with MCF10A cells. This orthogonal analysis resulted in the identification of 148 genes.
differentially expressed in both SUM-44 cells and MCF10A-WHSC1L1 cells, relative to MCF10A cells (Supplementary Table). Of the 148 differentially expressed genes, 36 are coordinately upregulated in MCF10A-WHSC1L1 cells and SUM-44 cells. Figure 5A shows the four upregulated genes (TBL1X, IRX3, RAG1AP1, and RAPGEF3), and two downregulated genes (TFBI and SFRP1) in both SUM-44 cells and MCF10A-WHSC1L1 cells, relative to MCF10A cells. To validate these array-based observations, we examined expression of these genes by quantitative RT-PCR in SUM-44 and MCF-10 cells, and in SUM-44 cells following knockdown of WHSC1L1 (Fig. 5B). Figure 5 shows that IRX3, RAG1AP1, RAPGEF3, and TBL1X are significantly overexpressed at the mRNA level in SUM-44 cells compared with MCF10A cells. Furthermore, knockdown of WHSC1L1 in SUM-44 cells using the shRNA constructs described previously resulted in significant downregulation of these four putative target genes. These results support the array-based analysis and
indicate that \textit{WHSC1L1} regulates the expression of these target genes. Similarly, examination of one downregulated gene, \textit{TGFBI}, also confirmed reduced expression in SUM-44 compared with MCF10A cells, and the expression of this gene was increased in SUM-44 cells bearing the \textit{WHSC1L1} shRNA constructs.

Because \textit{WHSC1L1} regulates the expression of \textit{IRX3} in SUM-44 cells and MCF10A-\textit{WHSC1L1} cells, we examined the genomic state of this potentially important target gene. Interestingly, we found that SUM-44 cells have an amplification at the \textit{IRX3} locus of chromosome 16q12. In addition, SUM-225 cells, which also have the 8p11-12 amplicon, have...
an overlapping region of copy number increase in chromosome 16 (Supplementary Fig. S5). We performed fluorescence in situ hybridization analysis using an IRX3-specific probe prepared from BAC clone RP11-1061L23 and confirmed the presence of an independent IRX3 amplification in SUM-44 and SUM-225 cells (Supplementary Fig. S6). Thus, these findings indicate that in SUM-44 cells, overexpression of an amplified oncogene on chromosome 8p11-12 drives the expression of another amplified gene on a different chromosome. This genetic interaction explains the very high level of expression of IRX3 in SUM-44 cells compared with MCF10A-WHSC1L1 cells.

With respect to downregulated genes, the finding that overexpression of WHSC1L1 resulted in downregulation of the negative regulator of WNT signaling, SFRP1, in MCF10A cells is intriguing. We have recently shown that although the SFRP1 gene is part of the 8p11-12 amplicon and is increased in copy number in SUM-44 cells, it is highly methylated and not expressed in these cells (27). Another downregulated gene is TGFBI, which encodes a secreted protein induced by transforming growth factor-β (Fig. 5). Recent studies with TGFBI-null mice showed that TGFBI loss promotes cell proliferation and predisposes mice to spontaneous tumor development (28). Thus, PWWP-protein WHSC1L1 may regulate a subset of genes involved in various functional pathways in breast cancer.

Discussion

The 8p11-12 amplicon has been the subject of a number of studies using high-resolution genomic analysis of copy number and gene expression in human breast cancer (7–11, 29). Our first studies in this area showed that the 8p11-12 amplicon has a complex genomic structure and the size of the amplicon is variable in three human breast cancer lines: SUM-44, SUM-52, and SUM-225 (9, 30). In that work, we showed that FGFR1 was only one of several candidate oncogenes in the amplicon, and we provided evidence that FGFR1 is not the driving oncogene in every breast cancer with the 8p11 amplicon (30). In addition, our correlative evidence suggested that other genes in the region, including LSM-1, C8orf4 (TC-1), RAB11FIP1, WHSC1L1, and ERLIN2 were good candidate oncogenes based on their overexpression associated with gene amplification (10). Our findings are consistent with those of other laboratories. Gelsi-Boyer et al. performed a comprehensive study combining genomic, expression, and chromosome break analyses of the 8p11-12 region in 37 human breast cancer lines and 134 primary breast cancer specimens (8). They identified four overlapping amplicon cores at 8p11-12 and 14 candidate oncogenes that are significantly overexpressed in relation to amplification. In subsequent work, Bernard-Pierrot et al. carried out BAC-array CGH on 21 human breast cancer lines and 152 ductal breast carcinomas and identified five genes (LSM1, BAG4, DDHD2, PPAPDC1B, and WHSC1L1) within the 8p11-12 amplified region as consistently overexpressed due to an increased gene copy number. Finally, Chin et al. published an analysis of the association of 8p11-12 gene amplification and disease-free survival and distant relapse in human breast cancer specimens and identified 23 genes from the 8p11-12 region as being correlated with progression, all of which have been named already (12). Thus, several groups have performed extensive analyses of the 8p11-12 genomic region in human breast cancer and there is substantial agreement on the candidate oncogenes present in this region. The candidate oncogenes consistently identified by all groups include FGFR1, WHSC1L1, RAB11FIP1, LSM1, BAG4, and ERLIN2.

There are now several studies in the literature reporting experimental analysis of the transforming function of the candidate oncogenes from the 8p11-12 region. We reported that BAG4, LSM1, and C8orf4 (TC-1) are transforming when overexpressed in MCF10A cells (10, 13, 14). In the present report, we provide evidence that three additional genes, namely, WHSC1L1, ERLIN2, and DDHD2, are transforming based on their ability to induce growth factor–independent proliferation, anchorage-independent growth, and altered morphogenesis in Matrigel cultures. As reported in this article, we find WHSC1L1 to be the most potently transforming of all the 8p11 oncogenes we have tested. Our results are consistent with those reported earlier by Bernard-Pierrot et al. who performed RNAi experiments to knock down the expression of candidate genes in two cell lines (CAM-1 and ZR-75-1) with 8p11-12 amplification (31). Their results suggest that PPAPDC1B and WHSC1L1 are two driving oncogenes from this amplicon. Knockdown of WHSC1L1 was found to inhibit the proliferation of ZR-75-1 and CAM-1 cells, but had no effect on MCF-7 cells that lack the 8p11-12 amplicon. Further, inhibition of WHSC1L1 increased the number of apoptotic cells as assessed by terminal deoxynucleotidyl transferase–mediated dUTP nick end labeling assay in ZR-75-1 and CAM-1 cells (31). Recently, Zhang and colleagues applied a novel algorithm, termed TRIAGE (triangulating oncogenes through clinico-genomic intersects), to a collection of microarray expression profiles of primary human breast cancers in an effort to identify candidate genes in amplicons that could contribute to patient outcome (32). They identified RAP11FIP1 and also identified WHSC1L1 as being strongly associated with breast cancer subtype and outcome. They selected RAP11FIP1 for further transfection and knockdown studies and found that RAP11FIP1 is not sufficient to transform naive cells. However, overexpression of RAP11FIP1 in breast cancer cell lines caused decreased growth factor dependence, increased survival under anoikis conditions, and increased motility and invasion. Furthermore, RAP11FIP1 overexpression caused an epithelial-mesenchymal transition in vitro and increased tumor growth in vivo (32). In other studies, Luscher-Firzlaff et al. reported that ASH2L encodes the trithorax protein and cooperates with H-Ras to transform primary rat embryo fibroblasts (33).

Based on findings from several laboratories, WHSC1L1 is clearly emerging as an important transforming gene within the 8p11-12 amplicon in breast and other cancer types. WHSC1L1 is involved in a chromosomal translocation, t(8;11)(p11.2;p15), in acute myeloid leukemia (34). Recent published database of the Affymetrix 250K Sty array in a collection of 244 copy-number profiles of breast cancer samples showed that WHSC1L1 amplification occurred in ~15% samples (Supplementary Fig. S7; ref. 35). With GISTIC (Genomic
Identification of significant Targets in Cancer), a CGH analysis program, WHSC1L1 was identified at the peak of amplification in lung and esophageal squamous cell carcinoma (36). siRNA-mediated knockdown of WHSC1L1 resulted in a 50% reduction in the number of soft agar colonies in a lung cancer cell line (H1703) with WHSC1L1 gene amplification and overexpression (37). Furthermore, deep sequencing of a primary human breast cancer identified a deletion within the WHSC1L1 gene (38). In this study, we identified one primary breast cancer specimen with the 8p11-12 amplicon in which genomic analysis showed a loss of the COOH-terminal region of the WHSC1L1 long isoform but amplification of exons 1 to 10 that coded for the short isoform. At the protein level, WHSC1L1 exists as two isoforms in breast cancer cells with 8p11-12 amplification. Alternative splicing of WHSC1L1 in breast cancer cells can be regulated at different steps of the spliceosome assembly by different splicing factors, and by many different mechanisms that rely on cis–acting elements (39). Future investigations are required to more precisely address the role and mechanism of action of WHSC1L1 isoforms in breast cancer.

A finding of particular interest from our study is that IRX3, a member of the homeobox gene family, and TBLIX are target genes of WHSC1L1. Interestingly, IRX3 is also amplified in SUM-44 cells and in SUM-225 cells. This is of interest because in embryonic stem cells IRX3 and TBLIX are linked in a gene expression network that regulates WNT signaling (40). In addition, we have previously shown that in breast cancers with the 8p11-12 amplicon, SFRP1, a negative regulator of WNT signaling, is silenced by promoter methylation, despite being present on the 8p11 amplicon and increased in copy number (27). These results suggest that overexpression of WHSC1L1 and the silencing of SFRP1 result in potent activation of a transcriptional network linked to WNT signaling and expression of stem cell phenotypes.

FGFR1 has long been considered an important candidate breast cancer oncogene from the 8p11-12 region. However, we have consistently failed to find evidence for a direct role of FGFR1 in transformation in mammary epithelial cells. Recently, Turner et al. provided evidence for a functional role of FGFR1 in 8p11-12–amplified breast cancers (41). Many of the results reported by Turner et al. are consistent with our previously published negative results. However, they did show that overexpression of FGFR1 increases the sensitivity and responsiveness of cells to fibroblast growth factor ligands, which influences the response of the cells to 4-OH tamoxifen. These results suggest that FGFR1 overexpression can play a role in endocrine therapy resistance, which may explain the consistent presence of FGFR1 in the amplicon.

In the past, focal amplicons found in cancer specimens were considered to harbor a single driving oncogene that was responsible for the maintenance of the amplicon in the tumor, with the ERBB2 oncogene in the 17q12 amplicon being a prime example. In some cases, amplicons have been thought to harbor more than one driver oncogene that act independently, such as the CCND1 and EMS1 genes present in the 11q12 amplicon. It is possible that the 8p11-p12 amplicon does not follow such a simple paradigm. Indeed, we have proposed that the 8p11-12 amplicon, rather than having a single driving oncogene, can act as an oncogenic unit consisting of multiple interacting transforming genes. This hypothesis is based on the consistent coexpression of several candidate oncogenes with transforming function when the amplicon is present in breast cancers. Within this oncogenic unit are two genes that can regulate the histone code (WHSC1L1, ASH2L), two genes that regulate RNA metabolism (LSM1, BBR2), a receptor tyrosine kinase (FGFR1), a gene that regulates the endoplasmic reticulum stress pathway (ERLIN2), and a gene that influences receptor trafficking (RAR11FIP1). Although it remains possible that each of these genes act independently, and function as driver oncogenes in different tumors with the same amplicon, the possibility that the genes cooperate in mediating neoplastic transformation must now be considered.

Despite the significant and exciting progress in the understanding of the 8p11-12 genomic amplification in breast cancer, we are still in the early stages of functional studies for each 8p11-12 candidate oncogene and its role in breast cancer development. Understanding how the genes in this region influence fundamental cancer processes such as progression, metastasis, and drug resistance will provide potential new avenues for therapeutic development.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Acknowledgments

We thank Michele L. Dziubinski and Katie L. Streicher for technical assistance on the cell culture and three-dimensional morphogenesis assays. The expression profiling was performed by the Genomics Core of the Karmanos Cancer Institute that is supported by grant number P30-CA022453-29. The array CGH work was facilitated by the Microarray and Bioinformatics Core Facility of the Wayne State University Environmental Health Sciences Center, NIEHS P30 ES06869.

Grant Support

Department of Defense Breast Cancer Program (BC083945) to Z-Q. Yang and a grant from the National Institutes of Health (ROI CA100724) to S.P. Ethier.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked advertisement in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received 03/22/2010; revised 08/27/2010; accepted 08/27/2010; published OnlineFirst 10/12/2010.

References
