Truncating prolactin receptor mutations promote tumor growth in murine estrogen receptor-alpha mammary carcinomas

Obi L. Griffith  
*Washington University School of Medicine in St. Louis*

Szeman Ruby Chan  
*Washington University School of Medicine in St. Louis*

Malachi Griffith  
*Washington University School of Medicine in St. Louis*

Kilannin Krysiak  
*Washington University School of Medicine in St. Louis*

Zachary L. Skidmore  
*Washington University School of Medicine in St. Louis*

See next page for additional authors

Follow this and additional works at: [https://digitalcommons.wustl.edu/open_access_pubs](https://digitalcommons.wustl.edu/open_access_pubs)

Recommended Citation
[https://digitalcommons.wustl.edu/open_access_pubs/5364](https://digitalcommons.wustl.edu/open_access_pubs/5364)
Authors

This open access publication is available at Digital Commons@Becker: https://digitalcommons.wustl.edu/open_access_pubs/5364
Truncating Prolactin Receptor Mutations Promote Tumor Growth in Murine Estrogen Receptor-Alpha Mammary Carcinomas

Graphical Abstract

Highlights
- Mice lacking Stat1 develop mammary tumors that model human ERα+ breast cancer
- DNA sequencing finds a Prlr truncating mutation hotspot in 100% of Stat1−/− tumors
- At least 77.8% of DCIS examined also contain PRLR truncating mutations
- Co-expression of truncated and full-length PRLR promote tumor growth in MEFs

Authors
Obi L. Griffith, Szeman Ruby Chan, Malachi Griffith, ..., Richard K. Wilson, Robert D. Schreiber, Elaine R. Mardis

Correspondence
schreiber@immunology.wustl.edu (R.D.S.), emardis@wustl.edu (E.R.M.)

In Brief
Griffith et al. report on whole genome sequencing of mammary tumors that spontaneously develop in mice lacking Stat1. They find a 100% recurrent hotspot of mutations resulting in truncation of Prlr. Co-expression of truncated and full-length Prlr in mouse embryonic fibroblasts is shown to promote tumor growth.
Truncating Prolactin Receptor Mutations Promote Tumor Growth in Murine Estrogen Receptor-Alpha Mammary Carcinomas

Obi L. Griffith,1,2,3,9 Szeman Ruby Chan,4,8,9 Malachi Griffith,1,3,5 Kilannin Krysiak,1,2 Zachary L. Skidmore,1 Jasreet Hundal,1 Julie A. Allen,6 Cora D. Arthur,1 Daniele Runci,4 Mattia Bugatti,1 Alexander P. Miceli,8 Heather Schmidt,1 Lee Trani,1 Krishna-Latha Kanchi,1 Christopher A. Miller,1,2,7 David E. Larson,1,5 Robert S. Fulton,1,7 William Vermi,4,7 Jaspreet Hundal,1 Julie A. Allen,6 Cora D. Arthur,1 Daniele Runci,4 Mattia Bugatti,1 Alexander P. Miceli,8 Heather Schmidt,1 Lee Trani,1 Krishna-Latha Kanchi,1 Christopher A. Miller,1,2,7 David E. Larson,1,5 Robert S. Fulton,1,7 William Vermi,4,7

1McDonnell Genome Institute, Washington University School of Medicine, 4444 Forest Park Ave., St. Louis, MO 63108, USA
2Department of Medicine, Washington University School of Medicine, 660 S Euclid Ave., St. Louis, MO 63110, USA
3Section of Pathology, Department of Molecular and Translational Medicine, School of Medicine, University of Brescia, Piazza del Mercato, 15, 25121 Brescia, Italy
4Department of Pathology and Immunology, Washington University School of Medicine, 4921 Parkview Pl., St. Louis, MO 63110, USA
5Department of Genetics, Washington University School of Medicine, 660 S Euclid Ave., St. Louis, MO 63110, USA
6Center for Human Immunology and Immunotherapy Programs, Washington University School of Medicine, 425 S Euclid Ave., St. Louis, MO 63110, USA
7Section of Pathology, Department of Molecular and Translational Medicine, School of Medicine, University of Brescia, Piazza del Mercato, 15, 25121 Brescia, Italy
8Present address: Janssen Research & Development, Johnson and Johnson, 1400 McKean Road, Spring House, PA 19477, USA
9Co-first author
10Lead Contact
*Correspondence: schreiber@immunology.wustl.edu (R.D.S.), emardis@wustl.edu (E.R.M.)

http://dx.doi.org/10.1016/j.celrep.2016.08.076

SUMMARY

Estrogen receptor alpha-positive (ERα+) luminal tumors are the most frequent subtype of breast cancer. Stat1−/− mice develop mammary tumors that closely recapitulate the biological characteristics of this cancer subtype. To identify transforming events that contribute to tumorigenesis, we performed whole genome sequencing of Stat1−/− primary mammary tumors and matched normal tissues. This investigation identified somatic truncating mutations affecting the prolactin receptor (PRLR) in all tumor and no normal samples. Targeted sequencing confirmed the presence of these mutations in precancerous lesions, indicating that this is an early event in tumorigenesis. Functional evaluation of these heterozygous mutations in Stat1−/− mouse embryonic fibroblasts showed that co-expression of truncated and wild-type PRLR led to aberrant STAT3 and STAT5 activation downstream of the receptor, cellular transformation in vitro, and tumor formation in vivo. In conclusion, truncating mutations of PRLR promote tumor growth in a model of human ERα+ breast cancer and warrant further investigation.

INTRODUCTION

The transcription factor STAT1 functions as a tumor suppressor in mammary gland epithelial cells (Chan et al., 2012, 2014; Schneckenleithner et al., 2011). Selective loss of STAT1 expression in breast cancer cells is associated with a significant percentage of human estrogen receptor alpha-positive (ERα+) luminal breast tumors (Chan et al., 2012). The lack of STAT1 expression in mice predisposes these animals to mammary adenocarcinoma development (Chan et al., 2012; Schneckenleithner et al., 2011). We demonstrated that spontaneous mammary tumors that develop in Stat1−/− female mice progressed in a manner similar to human ERα+ progesterone receptor-positive invasive ductal carcinoma (Chan et al., 2012). These tumors also displayed transcript expression profiles that clustered more closely with human ERα+ luminal breast cancers than other murine mammary tumor models and thus recapitulated the molecular characteristics of the luminal breast tumor subtype (Chan et al., 2012; Pfefferle et al., 2013).

To further identify the biological consequences of STAT1 loss in ERα+ luminal breast cancer, we set out to uncover genomic event(s) that fully transform the phenotype of mammary gland epithelial cells into cancer cells in the Stat1−/− mouse model. We performed whole genome sequencing of 14 primary Stat1−/− mammary tumors, 5 primary Stat1−/− ovarian hormone-independent tumors, and 3 Stat1−/− tumor-derived cell lines for a total of 22 independent Stat1−/− tumors (Figure 1). We compared genomic variations in tumor samples with those in control samples that consisted of 10 wild-type, 5 tumor-free Stat1−/− mammary glands, and 15 Stat1−/− tails. Our analysis revealed relatively few copy number variation (CNV) events in primary Stat1−/− mammary tumors and a point mutation rate consistent with that observed in human breast cancers. A number of key genes reported in human cohorts were also mutated in the Stat1−/− mammary tumors including Trp53, Brca1, and Mll3 and the Arid family. Strikingly, we identified a truncating mutation hotspot within the prolactin receptor (Prlr), with mutations...
RESULTS

Summary of Somatic Alterations

In a discovery set of 22 Stat1−/− tumor samples, using whole genome sequencing (WGS), we detected over 10,112 single nucleotide variants (SNVs) and 3,331 insertions and deletions (indels) within or near coding regions of known genes. Filtering and manual review reduced the set to 1,770 SNVs and 88 indels (Table S1). The 1,858 mutations occurred in 1,649 genes with 139 recurrently mutated genes (occurring in two or more samples) across all 22 tumors (Table 1 and Table S2). Mutational significance analysis revealed 16 significantly mutated genes (false discovery rate [FDR] < 0.05, likelihood ratio test [LRT] method; Table S3) (Dees et al., 2012). The most recurrently and significantly mutated gene was the prolactin receptor (Prlr), found in 17 out of 22 samples. These Prlr mutations will be discussed in greater detail below. Beyond the Prlr gene, mutations were observed in many of the same key genes and pathways reported previously for human breast and ovarian cancer, including Trp53, DNA repair genes (Brca1, Rad50, Rfc2, Poln, and Polr2a), chromatin modifiers (Arid1a and Arid1b), transcription factors (Zfp335, Zfp523, Zfp119a, Zfp119b), and kinases and phosphatases (Ptprb, Pik3r2, Pik3cd, Mapk7, and Src) (Banerji et al., 2012; Cancer Genome Atlas Network, 2012; Ellis et al., 2012; Jones et al., 2010; Stephens et al., 2012; Wiegand et al., 2010). Other mutations of note were in Ip6k2, which encodes a protein that affects the growth suppressive and apoptotic activities of interferon-beta in ovarian cancers (Morrison et al., 2001); Tiam1, which is a t-lymphoma invasion and metastasis-inducing...
Table 1. Recurrently Mutated Genes with Mutations in More Than Two Tumors from WGS

<table>
<thead>
<tr>
<th>Gene Common Name</th>
<th>Mutations (n)</th>
<th>Samples (n)</th>
<th>HGNC Symbol</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prlr*</td>
<td>17A</td>
<td>17A</td>
<td>PRLR</td>
</tr>
<tr>
<td>Olfr1062*</td>
<td>5</td>
<td>5</td>
<td>N/A</td>
</tr>
<tr>
<td>Mll3</td>
<td>4</td>
<td>4</td>
<td>MLL3</td>
</tr>
<tr>
<td>493243P20Rik*</td>
<td>4</td>
<td>4</td>
<td>WDR87</td>
</tr>
<tr>
<td>Gata3</td>
<td>3</td>
<td>3</td>
<td>GATSL3</td>
</tr>
<tr>
<td>Gnl1*</td>
<td>3</td>
<td>3</td>
<td>GNL1</td>
</tr>
<tr>
<td>Esmrg*</td>
<td>3</td>
<td>3</td>
<td>ESRRG</td>
</tr>
<tr>
<td>Galectin5*</td>
<td>3</td>
<td>3</td>
<td>GALNT5</td>
</tr>
<tr>
<td>bc006779*</td>
<td>3</td>
<td>3</td>
<td>N/A</td>
</tr>
<tr>
<td>Rbbp6</td>
<td>3</td>
<td>3</td>
<td>RBBP6</td>
</tr>
<tr>
<td>Scl39a12*</td>
<td>3</td>
<td>3</td>
<td>SLC39A12</td>
</tr>
<tr>
<td>Zfp335</td>
<td>3</td>
<td>3</td>
<td>ZNF335</td>
</tr>
<tr>
<td>Gm6369</td>
<td>3</td>
<td>3</td>
<td>N/A</td>
</tr>
<tr>
<td>Fbxl7</td>
<td>4</td>
<td>3</td>
<td>FBXL7</td>
</tr>
<tr>
<td>Krt15*</td>
<td>3</td>
<td>3</td>
<td>KRT15</td>
</tr>
<tr>
<td>Tgoln1*</td>
<td>3</td>
<td>3</td>
<td>TGOVL2</td>
</tr>
<tr>
<td>Gm16372</td>
<td>3</td>
<td>3</td>
<td>N/A</td>
</tr>
<tr>
<td>Trps3*</td>
<td>3</td>
<td>3</td>
<td>TP53</td>
</tr>
<tr>
<td>Taar7e*</td>
<td>3</td>
<td>3</td>
<td>N/A</td>
</tr>
<tr>
<td>4930503E14Rik*</td>
<td>3</td>
<td>3</td>
<td>N/A</td>
</tr>
<tr>
<td>ENSMUSG0000077055</td>
<td>3</td>
<td>3</td>
<td>N/A</td>
</tr>
<tr>
<td>Tmem181b-ps</td>
<td>3</td>
<td>3</td>
<td>TMEM181</td>
</tr>
<tr>
<td>Gm10601</td>
<td>3</td>
<td>3</td>
<td>N/A</td>
</tr>
<tr>
<td>A230087F16Rik</td>
<td>3</td>
<td>3</td>
<td>N/A</td>
</tr>
<tr>
<td>Gm11867</td>
<td>3</td>
<td>3</td>
<td>N/A</td>
</tr>
<tr>
<td>Gm16957</td>
<td>3</td>
<td>3</td>
<td>N/A</td>
</tr>
<tr>
<td>Vmn2r90</td>
<td>3</td>
<td>3</td>
<td>N/A</td>
</tr>
</tbody>
</table>

Mutations, total number of mutations identified; samples, number of samples with at least one mutation. Multiple mutations within the same gene per sample are possible. *Prlr mutations were called in 17/22 tumor samples during initial calling from WGS data. Manual review and subsequent validation assays confirmed mutations in 22/22 tumor samples. **Significantly mutated genes (Table S3). See also Table S1, Table S2, and Table S3.

protein; and Esmrg, which encodes the estrogen-related receptor gamma protein. No significant differences were observed in mutation frequencies between ovarian dependent and independent tumors or cell lines, although we are admittedly underpowered to detect such differences.

The numbers of mouse whole genomes sequenced in our study (n = 22) limit the direct comparison of mutation frequencies. Despite this, we compared our cohort to mutations frequently observed in human luminal breast cancers by identifying genes mutated at >5% frequency in The Cancer Genome Atlas (TCGA) luminal A and B cohort (n = 699) (Cancer Genome Atlas Network, 2012). We observed Trps3 and Mll3 mutations at frequencies comparable to the human dataset with 14% versus 16% and 18% versus 8%, respectively (Figure S1). The lack of Pik3ca and Map3k1 mutations is perhaps expected in the context of PRLR truncation (see below) given that their activity is downstream of PRLR and, therefore, activating mutations in these genes may not be required for tumor formation. Gata3 mutations were also not observed in our dataset, although we previously showed upregulation of Gata3 in Stat1−/− tumors, consistent with ERα human breast cancer (Chan et al., 2012).

Summary of Copy Number Variation Results

As a positive control, in each tumor we verified that the Stat1 exon 3–5 deletion was detectable by read-depth-based CNV analysis using CopyCat. Only moderate additional copy number changes were observed in tumor samples (Table S9). Virtually no CNV events were observed in wild-type or Stat1−/− tumor-free mammary glands (Table S9).

Recurrent Truncating Mutations of PRLR

As described above, the most recurrently mutated gene in Stat1−/− mammary tumors was that of prolactin receptor (Prlr) (Table 1). Prlr mutations were not observed in any matched normal tails (0/17), wild-type mammary glands (0/10), or tumor-free Stat1−/− mammary glands (0/5). Our alignment and variant calling pipelines and further manual inspection of WGS data for the Prlr region revealed Prlr mutations in a total of 21/22 tumors (Table 2; Appendix S1). Only the TAC246 tumor sample had no evidence of a Prlr mutation in the initial discovery WGS dataset. All discovery tumor samples were further Sanger sequenced for the Prlr region of interest to validate the observed indels (Table S4; Appendix S2). Sequence traces consistent with the WGS mutations were confirmed for 19 of 21 tumors. Traces for two samples were ambiguous. However, detection of indels from Sanger traces is difficult and this, in our experience, represents a very high indel validation rate. MiSeq sequencing of a formalin-fixed paraffin-embedded (FFPE) sample from the TAC246 tumor identified a Prlr mutation in this sample that was missed in the original discovery set by WGS (giving a sensitivity of 95.5% for the ~30X WGS approach of detecting Prlr mutations). A second mutation was additionally detected by MiSeq data of an FFPE sample from the TAC247 tumor that was also missed by WGS. As a result, 100% of the original discovery samples were found to contain at least one Prlr mutation (Figure S2). Extension sequencing, by the Sanger method, was performed on an additional 10 tumors and 35 non-tumor samples from 10 additional mice (Appendix S3). Non-tumor samples included 10 normal tails, 8 uteri, 7 ovaries, 8 livers, and 2 mammary glands. Prlr mutations were observed in all additional tumors and none of the non-tumor samples (Figure S2).

In order to determine whether Prlr mutations are an early tumor-initiating event, ductal carcinoma in situ (DCIS) components were identified from FFPE blocks of additional Stat1−/− mammary glands. DCIS DNA samples were amplified for the mutated Prlr region and sequenced by MiSeq (Appendix S4). In total, seven of nine (77.8%) DCIS samples showed evidence of truncating Prlr mutations. Immunohistochemical analysis also indicated that activated STAT3 and STAT5 were present in a majority of the atypical cells in DCIS (Figure S3), which was consistent with our previous results for pSTAT3/5 in primary Stat1−/− tumors (Chan et al., 2014). These results suggest that mutations...
Table 2. Summary of Prlr Mutations in Discovery and Extension Sets

<table>
<thead>
<tr>
<th>Start</th>
<th>Stop</th>
<th>Variant</th>
<th>Sample</th>
<th>Variant Effect</th>
<th>AA Change</th>
<th>Set</th>
<th>WGS</th>
<th>PCR Sanger</th>
<th>PCR MiSeq</th>
</tr>
</thead>
<tbody>
<tr>
<td>10258182</td>
<td>10258182</td>
<td>G/0</td>
<td>B3R15 L. tho.</td>
<td>frame_shift_del</td>
<td>p.G330fs</td>
<td>discovery</td>
<td>Y</td>
<td>Y</td>
<td>N/A</td>
</tr>
<tr>
<td>10258180</td>
<td>10258180</td>
<td>C/0</td>
<td>B3R12L1 R. tho.</td>
<td>frame_shift_del</td>
<td>p.P329fs</td>
<td>discovery</td>
<td>Y</td>
<td>Y</td>
<td>N/A</td>
</tr>
<tr>
<td>10258151</td>
<td>10258151</td>
<td>G/0</td>
<td>O VX3L2 R. tho.</td>
<td>frame_shift_del</td>
<td>p.L320fs</td>
<td>discovery</td>
<td>Y</td>
<td>Y</td>
<td>N/A</td>
</tr>
<tr>
<td>10258151</td>
<td>10258151</td>
<td>G/0</td>
<td>O VX6R2 L. cerv.</td>
<td>frame_shift_del</td>
<td>p.L320fs</td>
<td>discovery</td>
<td>Y</td>
<td>Y</td>
<td>N/A</td>
</tr>
<tr>
<td>10258147</td>
<td>10258147</td>
<td>G/0</td>
<td>SSM1</td>
<td>frame_shift_del</td>
<td>p.E318fs</td>
<td>discovery</td>
<td>Y</td>
<td>Y</td>
<td>N/A</td>
</tr>
<tr>
<td>10258139</td>
<td>10258140</td>
<td>0/TGAGGAGCAGC</td>
<td>SSM2</td>
<td>frame_shift_ins</td>
<td>p.E319fs</td>
<td>discovery</td>
<td>Y</td>
<td>Y</td>
<td>N/A</td>
</tr>
<tr>
<td>10258195</td>
<td>10258195</td>
<td>A/0</td>
<td>SSM3</td>
<td>frame_shift_del</td>
<td>p.K334fs</td>
<td>discovery</td>
<td>Y</td>
<td>Y</td>
<td>N/A</td>
</tr>
<tr>
<td>10258180</td>
<td>10258180</td>
<td>C/0</td>
<td>TAC171 R. tho.</td>
<td>frame_shift_del</td>
<td>p.P329fs</td>
<td>discovery</td>
<td>Y</td>
<td>Y</td>
<td>N/A</td>
</tr>
<tr>
<td>10258147</td>
<td>10258147</td>
<td>G/0</td>
<td>TAC183 R. tho.</td>
<td>frame_shift_del</td>
<td>p.E318fs</td>
<td>discovery</td>
<td>Y</td>
<td>Y</td>
<td>N/A</td>
</tr>
<tr>
<td>10258147</td>
<td>10258148</td>
<td>0/GATGGCT</td>
<td>TAC186 L. cerv.</td>
<td>frame_shift_ins</td>
<td>p.E318fs</td>
<td>discovery</td>
<td>Y</td>
<td>Y</td>
<td>N/A</td>
</tr>
<tr>
<td>10258178</td>
<td>10258186</td>
<td>ATCCGGCT/0</td>
<td>TAC246 R. ing.</td>
<td>in_frame_del</td>
<td>p.Y328*</td>
<td>discovery</td>
<td>N</td>
<td>N</td>
<td>Y</td>
</tr>
<tr>
<td>10258147</td>
<td>10258148</td>
<td>G/0</td>
<td>TAC247 L. cerv.</td>
<td>frame_shift_del</td>
<td>p.K334fs</td>
<td>discovery</td>
<td>Y</td>
<td>Y</td>
<td>N/A</td>
</tr>
<tr>
<td>10258147</td>
<td>10258148</td>
<td>GA/0</td>
<td>TAC270 L. tho.</td>
<td>frame_shift_del</td>
<td>p.E318fs</td>
<td>discovery</td>
<td>Y</td>
<td>Y</td>
<td>N/A</td>
</tr>
<tr>
<td>10258147</td>
<td>10258148</td>
<td>GA/0</td>
<td>TAC270 L. ing.</td>
<td>frame_shift_del</td>
<td>p.E318fs</td>
<td>discovery</td>
<td>Y</td>
<td>Y</td>
<td>N/A</td>
</tr>
<tr>
<td>10258179</td>
<td>10258179</td>
<td>T/A</td>
<td>TAC271 R. tho.</td>
<td>nonsense</td>
<td>p.E318fs</td>
<td>discovery</td>
<td>Y</td>
<td>A</td>
<td>N/A</td>
</tr>
<tr>
<td>10258146</td>
<td>10258168</td>
<td>CGAGCGGCTAATGCCATCCATT/0</td>
<td>TAC272 L. tho.</td>
<td>frame_shift_del</td>
<td>p.E318fs</td>
<td>discovery</td>
<td>Y</td>
<td>Y</td>
<td>N/A</td>
</tr>
<tr>
<td>10258195</td>
<td>10258195</td>
<td>A/0</td>
<td>TAC273 L. tho.</td>
<td>frame_shift_del</td>
<td>p.K334fs</td>
<td>discovery</td>
<td>Y</td>
<td>Y</td>
<td>N/A</td>
</tr>
<tr>
<td>10258146</td>
<td>10258149</td>
<td>CGAG/0</td>
<td>TAC273 R. ing.</td>
<td>frame_shift_del</td>
<td>p.E318fs</td>
<td>discovery</td>
<td>Y</td>
<td>Y</td>
<td>N/A</td>
</tr>
<tr>
<td>10258147</td>
<td>10258147</td>
<td>G/T</td>
<td>TAC274 R. ing.</td>
<td>nonsense</td>
<td>p.E318*</td>
<td>discovery</td>
<td>Y</td>
<td>Y</td>
<td>N/A</td>
</tr>
<tr>
<td>10258184</td>
<td>10258184</td>
<td>G/0</td>
<td>O VX13R1R2 L. ing.</td>
<td>frame_shift_del</td>
<td>p.G330fs</td>
<td>extension</td>
<td>N/A</td>
<td>Y</td>
<td>N/A</td>
</tr>
<tr>
<td>10258181</td>
<td>10258181</td>
<td>C/0</td>
<td>TAC297 L. tho.</td>
<td>frame_shift_del</td>
<td>p.P329fs</td>
<td>extension</td>
<td>Y</td>
<td>Y</td>
<td>N/A</td>
</tr>
<tr>
<td>10258148</td>
<td>10258149</td>
<td>AG/0</td>
<td>TAC298 L. tho.</td>
<td>frame_shift_del</td>
<td>p.E318fs</td>
<td>extension</td>
<td>Y</td>
<td>Y</td>
<td>N/A</td>
</tr>
<tr>
<td>10258147</td>
<td>10258147</td>
<td>G/T</td>
<td>TAC299 R. tho.</td>
<td>nonsense</td>
<td>p.E318*</td>
<td>extension</td>
<td>Y</td>
<td>N</td>
<td>A</td>
</tr>
<tr>
<td>10258181</td>
<td>10258181</td>
<td>C/0</td>
<td>TAC300 L. tho.</td>
<td>frame_shift_del</td>
<td>p.P329fs</td>
<td>extension</td>
<td>N/A</td>
<td>Y</td>
<td>N/A</td>
</tr>
<tr>
<td>10258194</td>
<td>10258194</td>
<td>T/0</td>
<td>TAC300 R. ing.</td>
<td>frame_shift_del</td>
<td>p.K334fs</td>
<td>extension</td>
<td>N/A</td>
<td>Y</td>
<td>N/A</td>
</tr>
<tr>
<td>10258151</td>
<td>10258169</td>
<td>GGCTAATGCCATCCCA TTC/0</td>
<td>TAC301 L. tho.</td>
<td>frame_shift_del</td>
<td>p.R319fs</td>
<td>extension</td>
<td>N/A</td>
<td>Y</td>
<td>N/A</td>
</tr>
<tr>
<td>10258142</td>
<td>10258149</td>
<td>AGGACGAG/0</td>
<td>TAC302 R. cerv.</td>
<td>frame_shift_del</td>
<td>p.E316fs</td>
<td>extension</td>
<td>N/A</td>
<td>Y</td>
<td>N/A</td>
</tr>
<tr>
<td>10258184</td>
<td>10258184</td>
<td>G/0</td>
<td>TAC311 L. cerv.</td>
<td>frame_shift_del</td>
<td>p.G330fs</td>
<td>extension</td>
<td>N/A</td>
<td>Y</td>
<td>N/A</td>
</tr>
<tr>
<td>10258147</td>
<td>10258147</td>
<td>G/T</td>
<td>TAC311 L. ing.</td>
<td>nonsense</td>
<td>p.E318*</td>
<td>extension</td>
<td>N/A</td>
<td>Y</td>
<td>N/A</td>
</tr>
<tr>
<td>10258222</td>
<td>10258223</td>
<td>0/A</td>
<td>TAC299 L. tho.</td>
<td>frame_shift_ins</td>
<td>p.D343fs</td>
<td>DCIS</td>
<td>N/A</td>
<td>Y</td>
<td>N/A</td>
</tr>
<tr>
<td>10258151</td>
<td>10258173</td>
<td>GGCTAATGCCATCCCAT TCCAAA/-</td>
<td>TAC312 L. tho.</td>
<td>frame_shift_del</td>
<td>p.L320fs</td>
<td>DCIS</td>
<td>N/A</td>
<td>N/A</td>
<td>Y</td>
</tr>
<tr>
<td>10258177</td>
<td>10258177</td>
<td>T/-</td>
<td>TAC314 L. tho.</td>
<td>frame_shift_del</td>
<td>p.Y328fs</td>
<td>DCIS</td>
<td>N/A</td>
<td>N/A</td>
<td>Y</td>
</tr>
<tr>
<td>10258151</td>
<td>10258173</td>
<td>GGCTAATGCCATCCCAT TCCAAA/-</td>
<td>TAC319 L. ing.</td>
<td>frame_shift_del</td>
<td>p.L320fs</td>
<td>DCIS</td>
<td>N/A</td>
<td>N/A</td>
<td>Y</td>
</tr>
<tr>
<td>10258147</td>
<td>10258147</td>
<td>G/-</td>
<td>TAC319 R. ing.</td>
<td>frame_shift_del</td>
<td>p.E318fs</td>
<td>DCIS</td>
<td>N/A</td>
<td>N/A</td>
<td>Y</td>
</tr>
<tr>
<td>10258154</td>
<td>10258191</td>
<td>TAATGCCATCCATCCAAA GAGATATCCGGGTCAAGGT/-</td>
<td>TAC322 L. tho.</td>
<td>frame_shift_del</td>
<td>p.L320fs</td>
<td>DCIS</td>
<td>N/A</td>
<td>N/A</td>
<td>Y</td>
</tr>
<tr>
<td>10258147</td>
<td>10258147</td>
<td>G/-</td>
<td>TAC323 L. tho.</td>
<td>frame_shift_del</td>
<td>p.E318fs</td>
<td>DCIS</td>
<td>N/A</td>
<td>N/A</td>
<td>Y</td>
</tr>
</tbody>
</table>

Coordinates for the Prlr gene (Ensembl: ENSMUST00000124470; v67; mm9) are for chromosome 15. Abbreviations: WGS, whole genome sequencing; Y, mutation observed; N, mutation not observed; N/A, mutation data not available for technology; A, mutation calling ambiguous; L. tho., left thoracic; R. tho., right thoracic; L. cerv., left cervical; R. cerv., right cervical. A mutation for “TAC246 R. ing.” was not observed in WGS or PCR Sanger assay but was observed by the PCR MiSeq method. See also Table S4, Table S5, Table S10, Appendix S1, Appendix S2, Appendix S3, and Appendix S4.

*aOvarian-hormone independent tumor.*
in Prlr and activation of the PRLR pathway are early events during tumorigenesis of Stat1−/− mammary epithelial cells.

The final set of 40 Prlr mutations observed included 32 frame-shift-deletions, 3 frame-shift-insertions, 4 non-sense SNVs, and 1 in-frame deletion introducing a stop-gain (Table 2 and Table S5; Figure 2). All mutations were located within an 85 base pair window (chr15:10258139–10258223) of the mouse reference genome (build mm9). These mutations are predicted to produce a truncated PRLR protein only 317 to 349 amino acids (aa) in length compared to the 608 aa full-length wild-type PRLR (Figure S4). The truncated forms result in loss of most, but not all, of the PRLR cytoplasmic tail. They share the first 285 aa with the known “S1b” short-form with 32 to 64 additional amino acids and total lengths ranging between the “S1c” and “S1a” forms (Bole-Feyos et al., 1998; Pujianto et al., 2010). Examination of WGS data showed that nearly all Prlr mutations in primary Stat1−/− mammary tumors occurred in one of two alleles and all of the primary tumors examined so far displayed constitutive PRLR pathway activation, we hypothesized that heterodimers of full-length (FL) and truncated (T) PRLR may be the cause of constitutive PRLR activation and, thus, the tumorigenic phenotype of the Stat1−/− mammary epithelial cells. Endogenous expression of the FL and T PRLR isoforms was verified in Stat1−/− mammary tumor cell lines harboring these mutations (SSM1, SSM2, and SSM3) by immunoprecipitation and western blotting (Figure S5A). In contrast to the SSM2 and SSM3 tumor cell lines, which are heterozygous for the mutation, the SSM1 tumor cell line was homozygous (Table 2; Figure S5A) and failed to display constitutive PRLR-JAK2-STAT3/5 signaling (Chan et al., 2014). To directly examine the activity of the two PRLR isoforms, FL or T PRLR (aa residues 1 to 317) were expressed either alone or together in non-transformed mammary epithelial cells. Expression of the PRLR isoforms was confirmed by flow cytometry using a PRLR-specific monoclonal antibody (Figure S5B) (Chan et al., 2014). In the absence of exogenous PRL stimulation, phosphorylation of STAT3 and STAT5 was detected in cells expressing both FL and T PRLR, but not in cells expressing either FL PRLR homodimers or T PRLR homodimers alone (Figure S5A). Therefore, co-expression of FL and T PRLR led to phosphorylation and activation of STAT3 and STAT5 in the absence of exogenous PRL stimulation.

To further determine the biological significance of FL and T PRLR heterodimers, Stat1−/− MEFs expressing both FL and T PRLR were analyzed for their ability to grow in an anchorage-independent manner. Cells expressing FL and T PRLR developed significantly more colonies than control MEFS (p = 0.0013), or those expressing FL alone (p = 0.0013) or T alone (p = 0.0015), when plated in soft agar (Figures 3B and S5C).

**Figure 2. Mutational Hotspot Analysis of Prlr**

The diagrams depict the full-length, 608 amino acid residue coding region of Prlr (Ensembl: ENSMUST00000124470; v67) that is encoded by 9,900 base pairs. (A) depicts mutations identified in the original discovery set of tumors by whole genome sequencing and (B) depicts those identified in the extension/validation set (including DCIS samples) by Sanger/MiSeq sequencing. A total of 32 frame-shift deletions, 4 frame-shift insertions, 4 nonsense SNVs, and 1 in-frame deletion introducing a stop codon were identified in Prlr that cluster in 2 hotspots around residues 318 and 330. See also Figure S4, Table S5, Table S10, Appendix S1, Appendix S2, Appendix S3, and Appendix S4.

**Functional Significance of the Truncated Form of PRLR Protein**

Because all but one of the observed Prlr mutations in primary Stat1−/− mammary tumors occurred in one of two alleles and all of the primary tumors examined so far displayed constitutive PRLR pathway activation, we hypothesized that heterodimers of full-length (FL) and truncated (T) PRLR may be the cause of constitutive PRLR activation and, thus, the tumorigenic phenotype of the Stat1−/− mammary epithelial cells. Endogenous expression of the FL and T PRLR isoforms was verified in Stat1−/− mammary tumor cell lines harboring these mutations (SSM1, SSM2, and SSM3) by immunoprecipitation and western blotting (Figure S5A). In contrast to the SSM2 and SSM3 tumor cell lines, which are heterozygous for the mutation, the SSM1 tumor cell line was homozygous (Table 2; Figure S5A) and failed to display constitutive PRLR-JAK2-STAT3/5 signaling (Chan et al., 2014). To directly examine the activity of the two PRLR isoforms, FL or T PRLR (aa residues 1 to 317) were expressed either alone or together in non-transformed Stat1−/− murine embryonic fibroblasts (MEFs). Expression of the PRLR isoforms was confirmed by flow cytometry using a PRLR-specific monoclonal antibody (Figure S5B) (Chan et al., 2014). In the absence of exogenous PRL stimulation, phosphorylation of STAT3 and STAT5 was detected in cells expressing both FL and T PRLR, but not in cells expressing either FL PRLR homodimers or T PRLR homodimers alone (Figure S5A). Therefore, co-expression of FL and T PRLR led to phosphorylation and activation of STAT3 and STAT5 in the absence of exogenous PRL stimulation.

To further determine the biological significance of FL and T PRLR heterodimers, Stat1−/− MEFS expressing both FL and T PRLR were analyzed for their ability to grow in an anchorage-independent manner. Cells expressing FL and T PRLR developed significantly more colonies than control MEFS (p = 0.0013), or those expressing FL alone (p = 0.0013) or T alone (p = 0.0015), when plated in soft agar (Figures 3B and S5C). Stat1−/− MEFS expressing both FL and T PRLR also developed significantly
more tumors in nude mice than MEFs expressing vector alone (p = 4.867E-104), FL PRLR alone (p = 0.0344), or T PRLR alone (p = 3.365E-106) (Figure 3C). Tumor formation occurred more quickly in mice that received FL/T expressing MEFs than FL alone (p = 0.0002), T alone (p = 0.0003), vector alone (p = 0.0002), or MEFs alone (p < 0.0001; Figure S5D). FL/T expressing MEFs formed tumors at a frequency similar to the KRAS expressing positive control (p = 0.238), although at a significantly slower rate (p < 0.0001). Taken together, these results indicate that FL/T PRLR heterodimers promote activation of oncogenic STAT3 and STAT5, anchorage-independent growth, and transformation of non-transformed Stat1−/− MEFs.

**PRLR Mutations and Isoform Usage in Human Breast Cancers**

To assess the prevalence of PRLR mutations in human breast cancers, we examined human breast cancer exome sequence data from 991 patients made publicly available through the TCGA data portal. Using the published MAF file, four mutations in PRLR, including two SNVs and two indels, were identified (Cancer Genome Atlas Network, 2012). One of these mutations, an indel (L360fs), causes a truncating mutation in the human PRLR exon 10 (Ensembl: ENST00000382002; v70_37), analogous to that observed in the mouse Stat1−/− mammary tumors. Manual review of alignment data for this exon identified four additional truncating indels at E313fs (2/47 reads), L315fs (2/44), L360fs (35/42), and K460fs (3/100) from samples TCGA-B6-A0X7, TCGA-A2-A04R, TCGA-AC-A3EH, and TCGA-AR-A5QQ, respectively. L360fs, E313fs, and L315fs were found in luminal subtype breast cancers, whereas K460fs was in a basal breast cancer. The Exome Aggregation Consortium (ExAC) reports only four individuals with rare (allele frequency < 0.00001) germline truncating mutations in PRLR at A597fs, N568fs, S27*, and W180* (Lek et al., 2016).

There are currently eight to ten reported complete protein-coding transcript isoforms for human PRLR according to Ensembl: ENSG00000113494 (v79), University of California, Santa Cruz (UCSC) Genome Browser database: PRLR (GRCh37/ hg19), and UniProt: P16471 (v175) that can be broadly grouped as long, intermediate, and short PRLR isoforms (Bole-Feysot et al., 1998) (Figure S6; Table S6). We investigated the possibility of an increase in the expression of truncated (i.e., short) PRLR relative to full-length (i.e., long) PRLR in human ERα+ luminal breast cancer. We hypothesized that a skewing toward more T PRLR expression could be functionally equivalent to PRLR truncation. The expression ratio of FL to T PRLR was calculated using the TCGA human breast RNA sequencing (RNA-seq) data-sets based on counts for isoform-specific junctions (see Experimental Procedures). We also analyzed the STAT1 expression status of human luminal and basal or Her2 breast cancers and stratified each tumor subtype into STAT1-low and STAT1-high. In a previous report, we demonstrated that STAT1 was
specifically downregulated in tumor cells, but not in stromal cells, in ERα+ luminal breast cancers (Chan et al., 2012). Although the human breast tumor TCGA RNA-seq datasets were generated from whole tumors, we observed an overall significant reduction in the STAT1 expression level in the luminal breast cancer subtype compared to basal and Her2 subtypes \( (p = 2.484 \times 10^{-6}) \), reflecting the selective downregulation of STAT1 in these tumor cells, which was consistent with our previous report (Chan et al., 2012). We then analyzed the FL/T PRLR ratio in STAT1-low and STAT1-high tumors among each subtype. We observed a significant increase in T PRLR expression relative to FL PRLR expression (i.e., lower FL/T ratio) in STAT1-low samples compared to STAT1-high samples among luminal subtype tumors \( (p = 0.0077; \text{Figure } 4) \). No significant difference in FL/T ratio was observed among the basal or Her2 breast tumor group \( (p = 0.723) \). These data indicate that there may be a preferential usage of the truncated PRLR isoform in tumor cells with reduced STAT1 expression among ERα+ luminal breast cancers.

**DISCUSSION**

In this study, we identified recurrent gene mutations that were associated with the tumorigenic landscape of ERα+ Stat1−/− luminal mammary gland tumors. Several of these genes have also been reported as significantly mutated in human breast cancers, underscoring the biological significance of these mutations.
in the pathogenesis of this disease. Our study also revealed a potential mechanism whereby ERα+ luminal breast cancer initiates and progresses. Loss of Stat1 expression in mammary cells favors acquisition of mutations in an 85 base-pair hotspot of exon 10 of the Prir gene (Ensembl: ENSMUST0000124470; v67), resulting in a truncation of the cytoplasmic tail of the prolactin receptor (PRLR). Concurrent expression of full-length and truncated PRLR in the absence of STAT1 promotes phosphorylation and activation of the oncogenic STAT3 and STAT5, anchorage-independent growth of mouse embryonic fibroblasts, and tumor formation in nude mice.

PRLR is a transmembrane homodimeric receptor with an extracellular region that binds prolactin (PRL). It functions as a cytokine receptor and activates second messenger cascades, including the JAK2-STAT3/5, JAK-RUSH, RAS-RAF-MAPK, and PI3K pathways (Aksamitiene et al., 2011; Helmer et al., 2010; Rui et al., 1994). Over 75% of human ERα+ breast cancers display persistent PRLR-JAK2-STAT3/5 signaling (Chan et al., 2014). Activation of PRLR-JAK2-STAT3/5 signaling has been implicated in the upregulation of steroid hormone receptor expression and malignant progression of breast cancer (Chan et al., 2014; Fiorillo et al., 2013). There is also support for an association between PRLR allelic variations and breast cancer risk (Bogorad et al., 2008; Lee et al., 2007; Mong et al., 2011; Vlachavicek et al., 2006). In lobular neoplasia, amplification of PRLR may also be important for pathogenesis and progression (Thanh et al., 2011). Additionally, mouse mammary cancer models support a role for PRLR signaling in tumor progression. Elevated production of PRL ligand, driven by the Neo promoter, causes development of carcinomas in mice resembling human luminal breast carcinomas (Arendt et al., 2011). We have also shown that loss of STAT1 expression results in unopposed PRLR signaling, promotes expansion of mammary luminal progenitor cells, and leads to development of ductal carcinoma in situ (DCIS) and, finally, invasive mammary carcinomas (Chan et al., 2014). Lack of PRLR signaling has the opposite effect, whereby PRLR-deficiency delays tumor onset in the C3(1) SV40T model of mammary cancer (Oakes et al., 2007). Similarly, pharmacological inhibition of JAK2 (BMS-911543) not only abrogates mammary tumor formation but also causes regression of established Stat1+/− mammary tumors, demonstrating that constitutive activation of the PRLR-JAK2-STAT3/5 pathway promotes tumor progression and maintenance (Chan et al., 2014). Therefore, there is strong evidence supporting the involvement of PRLR signaling in the pathogenesis of ERα+ breast cancer.

The mechanism by which the PRLR pathway is activated during the development of ERα+ breast cancer is, however, less clear. Although elevated serum PRL levels have been associated with increased risk of developing ERα+ breast cancer (Twooger et al., 2015), it has been difficult to definitively show a causal relationship between PRLR pathway activation and breast cancer progression in those individuals who exhibited high PRL levels prior to diagnosis. In a recent study by Twooger and colleagues, half of the patients with ERα+ breast cancer did not show high plasma PRL levels when blood samples were collected <10 years prior to diagnosis (Twooger et al., 2013), suggesting that the association of PRLR signaling and development of ERα+ breast cancer may be more complex than simply elevated PRL production. In our previous report, we showed that excess production of PRL ligand was not the cause of constitutive activation of PRLR-JAK2-STAT3/5 signaling in Stat1−/− mammary tumors (Chan et al., 2014). In this study, we sought to identify a mechanism whereby constitutive oncogenic PRLR signaling was established and maintained in the absence of aberrant overproduction of PRL. We showed that heterodimers consisting of full-length and truncated PRLR activate JAK2-STAT3/5 in the absence of exogenous PRL stimulation. PRLR truncation and STAT3/5 activation were also observed in DCIS in Stat1−/− mammary glands, indicating that persistent PRLR signaling mediated by PRLR mutations was an early tumorigenic event. These results suggest that FL/T PRLR heterodimers could contribute to ERα+ breast cancer development in patients with normal PRL levels due to the intrinsic ability of the heterodimers to signal without ligand stimulation. Because this observation was made in mice and cell lines without STAT1 expression, it is likely that the ability of the heterodimers to confer a tumorigenic effect would be correlated with the loss of the negative regulator that normally controls PRLR signaling. We previously showed that approximately half of ERα+ luminal breast cancers displayed selective downregulation of STAT1 expression in tumor cells (Chan et al., 2012). Given that the Stat1−/SOC51 pathway negatively regulates PRLR signaling (Chan et al., 2014), it would be interesting to determine whether breast cancer cells in patients with normal PRL levels have down-modulated STAT1 and/or SOC51 expression. Our study also indicated that truncated PRLR heterodimers alone failed to transform Stat1−/− MEFs or promote tumor formation in mice (Figure 3). This is consistent with the finding that truncated PRLR lacks the STAT5 binding sites at residues 496 and 597 (Figures 2 and S4) and that STAT5 activation is necessary for transformation. Full-length heterodimers also failed to transform Stat1−/− MEFs in vitro and showed significantly reduced oncogenic activity in vivo. Therefore, it seems that co-expression of FL and T PRLR is required to collaborate with STAT1 loss and promote tumor progression.

One might expect that PRL-mediated tumor induction in the case of the NLR-PRL transgenic model (Arendt et al., 2011) and constitutive PRLR pathway activation in the case of Stat1−/− tumors (Chan et al., 2014) would lead to similar pathological outcomes. However, the tumors developed in Stat1−/− mice were mechanistically different from the tumors developed in NLR-PRL transgenic mice reported by Arendt and colleagues (Arendt et al., 2011; Chan et al., 2014). Arendt et al. (2011) found that PRL-induced mouse mammary carcinomas were heterogeneous with respect to histology, ER/PR expression, and signaling cascades (e.g., STAT5 signaling) and were insensitive to ER-mediated signaling. Our model is less heterogeneous, independent of over-production of PRL ligand, and sensitive to ER-mediated signaling. In addition, although JAK2 is required for initiation of the NRL-PRL tumors, it is not essential for tumor maintenance. In contrast, Stat1−/− tumors require JAK2 activation for both initiation and progression. Although the tumor-initiating cell population in the Stat1−/− tumors is the luminal progenitor subtype (Chan et al., 2014), it is not clear from which specific cell compartment the NLR-PRL tumors are derived. Given the distinct mode of tumorigenesis, the potential difference in targeted tumor-initiating cell populations, and the presence of STAT1 in the NLR-PRL model, it is likely that these differences...
might explain the different endocrine sensitivity and biological outcome of the two models.

We observed 4 PRLR truncating mutations in human breast cancer TCGA datasets. To our knowledge, only 2 exome sequencing studies of DCIS have been published, with 11 and 9 cases, respectively, and neither reports any PRLR mutations (Banerji et al., 2012; Kim et al., 2015). It is possible that PRLR truncation may occur as other types of genomic alterations, such as larger scale deletions, gene fusions, or translocations that are not easily detected from TCGA exome or RNA-seq data. More detailed structural analyses of breast cancer genomes will be required to further explore this possibility. Changes in the expression of naturally occurring alternative transcript isoforms may also provide a mechanism of PRLR deregulation. The human S1a and S1b truncated PRLR isoforms are generated by alternative splicing. Their roles in the development of ER+ breast cancer remain to be clarified. Dufau and colleagues reported an increase in mRNA expression of the full-length PRLR over the short S1a and S1b PRLR isoforms in human breast cancers (Meng et al., 2004). Our analyses compared all of the known short isoforms (S1a, S1b, Δ4 S1b, ΔS4-Δ7/11, Δ7/11, ΔS1) to the full-length isoform and observed a significant increase in expression of the short isoforms over full-length PRLR in STAT1-low luminal, but not basal breast cancer. Because tumor subtypes of the datasets used in the Dufau study were not classified, it was difficult to directly compare their results to our current study.

The biological significance of the PRLR short isoforms has been controversial. Some studies indicate that the PRLR short isoforms act as dominant negatives and block signaling from full-length PRLR (Hu et al., 2001). However, expression of the PRLR short isoform alone is sufficient to rescue the mammary gland differentiation defect in Prlr−/− mice (Binart et al., 2003), indicating that PRLR short isoform is not a dominant negative and FL PRLR and PRLR short isoform heterodimers can indeed transduce signals. None of these previous studies examined the ability of the FL PRLR and PRLR short isoform heterodimers to transform normal cells. We demonstrate the ability of these heterodimers to transform mouse embryonic fibroblasts and promote tumor formation in nude mice. This model also recapitulated the pSTAT3/5 activation we observed in primary tumors and DCIS samples. However, these experiments were performed using MEFs rather than the primary cell of origin, and future studies should assess the tumorigenic property of PRLR heterodimers in Stat1+/− mammary epithelial cells (MECs) with endogenous levels of JAK2. We speculate that the truncated PRLR short form is able to prolong signaling in the absence of PRL ligand because of its increased half-life on the cell surface. Phosphorylation of Ser349 on PRLR recruits the beta-transducin repeats-containing protein (b-TrCP) ubiquitin-protein ligase (Li et al., 2006). This interaction is important for ubiquitin-dependent degradation of the PRLR to terminate signaling. The truncated PRLR described in our current study lacks this critical serine residue, suggesting that it may be insensitive to b-TrCP-mediated degradation. Consistent with this hypothesis, phosphorylation on Ser349 is diminished in human breast cancer cell lines, leading to an increase in PRLR expression levels (Li et al., 2006). In addition, although FL PRLR and the truncated S1b isoform have similar binding affinity to growth hormone, the level of specific binding by S1b is significantly higher than that of FL PRLR on COS-1 or HEK293 cells transfected with either isoform (Hu et al., 2001; Trott et al., 2003). These results suggest that the cell surface expression of S1b PRLR is elevated compared to that of FL PRLR. Therefore, it is conceivably that increased FL/T PRLR heterodimer expression on the cell surface mediates receptor and JAK2 clustering such that autophosphorylation and activation is possible without PRL ligand engagement.

Anti-tumor agents targeting the PRLR pathway are being investigated for patients with ER+ breast cancer. For example, LFA102, an anti-PRLR antibody, blocks PRLR pathway activation by either inhibiting PRLR dimerization or locking the PRLR dimer in an inactive conformation without affecting PRL ligand binding (Damiano et al., 2013). Because truncated PRLR expression is preferentially increased in ER+ breast cancer and FL/T PRLR heterodimers display constitutive activation, as shown in our current study, it would be of interest to examine whether LFA102 is able to block heterodimerization of full-length and truncated PRLR in future studies. Unfortunately, LFA102 failed to show anti-tumor efficacy in a recent phase 1 clinical trial for metastatic breast cancer (Agarwal et al., 2016). Direct inhibition of JAK2 using small molecule inhibitors will also be worthy of investigation in breast cancers with PRLR activation. Future studies should also aim to clarify the biological outcome of signaling crosstalk between the FL/T PRLR heterodimers and the estrogen receptor pathway, because combination therapy targeting both pathways may be beneficial. In summary, our findings provide a mechanism whereby ER+ luminal breast cancer is initiated and maintained and pose hypotheses of translational and clinical significance in the treatment of this most common human breast cancer subtype.

**EXPERIMENTAL PROCEDURES**

**Mice**

Stat1−/− mammary gland adenocarcinomas have been previously characterized in our laboratory (Chan et al., 2012, 2014; Shankaran et al., 2001). Wild-type (WT) 129S6/SvEv and Stat1−/− mice were purchased from Taconic Farms. Stat1−/− mammary tumors of approximately 10 mm in diameter were harvested from 10–18-month-old retired breeders. Tumor-free Stat1−/− mice about 8 months of age were used as tumor-free controls. To obtain ovarian hormone-independent mammary tumors, ovaries were surgically removed from primary tumor-bearing mice as previously reported (Chan et al., 2014). If tumors did not respond to estrogen-deprivation and grew progressively, tumors were harvested. Tails from both tumor-free and tumor-bearing Stat1−/− mice were also harvested as normal controls. All animal experiments were carried out according to the guidelines of the American Association for Laboratory Animal Science under a protocol approved by the Animal Studies Committees and performed in Association for Assessment and Accreditation of Laboratory Animal Care International (AAALAC)-accredited specific pathogen-free facilities at Washington University School of Medicine in St. Louis.

**Cell Cultures**

The SSIM1, SSIM2, and SSIM3 Stat1−/− mammary tumor cell lines were cultured as previously described (Chan et al., 2012).

**Sample Acquisition**

Genomic DNAs were purified from tumor-free mammary glands, whole tumors, or tails using QIAGEN DNeasy Blood & Tissue kit according to the manufacturer’s instructions (QIAGEN). Fifty-two samples were whole genome sequenced for discovery purposes, and an additional 54 samples were...
sequenced using the Sanger protocol for targeted validation and extension of the Prlr findings (Figures 1 and S2). The discovery set included 14 primary Stat1+/− mammary tumors, 5 ovarian hormone (OH)-independent tumors, and the 3 cell lines (SSM1, SSM2, and SSM3) for a total of 22 tumors from 20 individual mice (2 mice had 2 tumors each). An additional 10 WT and 5 tumor-free Stat1+/− mammary glands from 13 additional mice were also sequenced as controls (2 mice had both tumor-bearing and tumor-free mammary glands). From the Stat1+/− mice, 15 tails were sequenced as matched normal samples for 12 primary tumors, 2 OH-independent tumors, and 5 tumor-free mammary glands. For 2 primary tumors, 3 OH-independent tumors, 3 cell lines, and 10 wild-type mammary glands without matched tails, a pooled sample of 2 unmatched normal tails was used as reference for somatic variant calling (see Table S7 for extensive details of all samples).

**Whole Genome Sequencing**

The yield and integrity of native genomic DNA was verified by a PicoGreen assay to determine mass (Invitrogen). Small insert dual indexed Illumina paired end libraries were constructed with the KAPA LTP sample prep kits according to the manufacturer’s recommendations (KAPA Biosystems) with a few exceptions (Supplemental Experimental Procedures). Each genome was loaded on a HiSeq2000 version 3 flow cell according to the manufacturer’s recommendations (Illumina). 2 × 101 bp read pairs were generated for each sample, yielding an average of 37.1 × coverage for the tumor genomes and 27.2 × coverage sequence for the normal genomes (Table S8).

**Reference Alignment and Somatic Variant Detection**

The Genome Modeling System, an integrated analysis information management system, was used for preliminary analysis of sequence data as previously described (Supplemental Experimental Procedures) (Griffith et al., 2013). Alignment was performed against the mouse reference genome (mm9), and variants were annotated with our custom annotator against Ensembl (version 67). Further filtering of SNVs and indels was performed to exclude (1) random, mitochondrial (MT), and Y contig events; (2) germline events defined as greater than 5% variant allele frequency (VAF) in the normal sample; (3) events with greater than 500 reads at the site; (4) variants from the Sanger Mouse Genomes Project (v2); and (5) variants that appear in 2 or more of the 10 WGS datasets from mammary tissue samples obtained from wild-type mice. These represent likely systematic artifacts of our alignment and calling pipelines. Further manual review of all SNVs, indels, and CNVs was performed in the Integrative Genome Viewer (IGV) to eliminate false positives arising from likely read mapping artifacts. Finally, analysis was performed on all reviewed and somatic variants (Table S1) to identify the recurrent (Table S2) and significantly mutated genes (Table S3) to identify the recurrent (Table S2) and significantly mutated genes were determined using the Mutational Significance in Cancer (MuSiC) pipeline (version 0.4), including non-coding mutations in the background mutation rate calculations (Dees et al., 2012). Reviewed CNV events are also summarized in Table S9. Genomic visualizations (Figure S1A) were created with the GenVisR Bioconductor package (Skidmore et al., 2016).

**Validation and Extension Sequencing of Prlr by Sanger and MiSeq**

The region in which truncating Prlr mutations were observed in the discovery set (chr15:10258139–10258195; mm9), two sets of primers were designed to encompass this region with approximately 50 or 100 bp additional flanking sequence on each side, respectively (Table S10; Supplemental Experimental Procedures). Primers were tagged (p1k/m13 reverse) and ordered from Integrated DNA Technologies (IDT). Amplification was performed in a Bio-Rad thermocycler (Bio-Rad Laboratories). A Lonza flash gel was run to confirm product. Sequencing reactions were performed and loaded on a 3730 DNA analyzer (Life Technologies). Bases were called from sequence trace files using phred and then assembled against a reference scaffold of the amplicon sequence using phrap (Ewing and Green, 1998; Ewing et al., 1998). Sequence variants were identified by manual review of assemblies and sequence traces in Consed (Gordon et al., 1998). We performed Sanger sequencing as described above on the original 22 tumor samples to validate Prlr variants that were called from WGS data and to extend the Prlr findings to 10 additional tumors and 35 non-tumor samples (Figures 1 and S2). For two tumors in the original discovery set, additional FFPE samples were obtained and sequenced on a MiSeq illumina instrument using products of the same PCR protocol described above. Finally we sequenced an additional 9 FFPE samples by MiSeq, to evaluate the presence of PRLR mutations in DCIS tissues.

**Review of PRLR Sequence Data from TCGA Human Breast Cancers**

The mutated region of mouse Prlr (chr15:10258139–10258195; mm9) and the flanking 50 base pairs on each side were aligned to the human reference genome using BLAST to identify the homologous region of PRLR in the human genome (chr5:35066045–35066101; hg19). This region in human PRLR was then extended to include the entirety of the affected “long” exon of PRLR (Figure S5) as well as the upstream–most exon-intron boundary and exon to identify a target region of interest (chr5:35066191–35066837; hg19) for manual review in human sequence data using IGV. Manual review focused on identifying truncating mutations or deletions of the long exon. For a mutation to be considered credible, at least a 3% tumor VAF was required. A total of 991 tumor and normal pairs of exome sequence data were investigated from the breast TCGA project, including 501 luminal A, 198 luminal B, 171 basal, 77 Her2, 31 normal-like, and 13 of unknown subtype (Cancer Genome Atlas Network, 2012).

**Exome Analysis of TCGA Human Breast Cancers**

Genes with somatic mutations for mouse data (Table S1, n = 22) and human luminal breast cancers from TCGA (n = 699) were compared (Figure S1). Within the comparison, mutation types were restricted to nonsynonymous coding and RNA mutations to ensure results were directly comparable. Using bioinformatics tools, Homo sapiens and Mus musculus ensemble IDs were annotated with orthologs from the other species. Genes without an ortholog with a one to one mapping were excluded from the analysis. Instances in which a sample had more than one mutation in the same gene were treated as having a single mutation.

**RNA-Seq Analysis of TCGA Human Breast Cancers**

Breast cancer RNA-seq data level 3 data corresponding to 10/10/2013 from the firehose pipeline (http://gdac.broadinstitute.org) were obtained from the TCGA data portal. A total of 775 breast tumor samples were represented with 376 luminal A, 181 luminal B, 131 basal, 65 Her2+, and 22 normal-like subtypes. Read counts per kb per million (RPKM) values were calculated as number of mapped reads / (transcript length in bp / 1,000) / (total reads / 1,000,000) and used to define STAT1 and PRLR gene expression levels. Tumor samples (excluding normal-like) were divided into tertiles to define low, intermediate, and high STAT1 expression groups (Figure 4). Junction fragments per million (JPM) values were defined as the raw counts for each junction, divided by the sum of all junctions, and multiplied by one million. PRLR full-length and truncated (short) isoform expression were estimated from the JPM values of FL and T isoform-specific junctions as follows (Figure S6):

\[
\text{Full-length(FL)} = \text{Junction 9B} - (\text{Junction 3A} + \text{Junction 10A})
\]

\[
\text{Truncated(T)} = \text{Junction 10A} + \text{Junction 9A} + \text{Junction 7A} + \text{Junction 3A}
\]

The ratio of PRLR isoform expression was defined as the log2 value of FL expression divided by T expression. As illustrated in Figure S6, the short human PRL isoforms are generated by alternative splicing and will be defined as “truncated” for clarification purposes. It should also be noted that human FL and intermediate isoform expression were grouped as “full-length” because it is unachievable to extract FL expression alone using junction data from TCGA.

**Immunoprecipitation and Western Blotting Analysis**

SSM1, SSM2, and SSM3 cells were lysed using complete RIPA buffer, and PRLR was immunoprecipitated using anti-mPRLR (clone 5A12) as previously described (Chan et al., 2014). Membrane was blotted with biotinylated anti-mPRLR and streptavidin-anti-hamster-IR800 and scanned using the Li-Cor Odyssey detection system.

**Immunohistochemistry**

Immunohistochemical analyses of STAT3 and STAT5 on luminal DCIS lesions were performed as described in a previous study (Chan et al., 2014).
Expression of Full-Length and Truncated Prlr

Full-length or truncated Prlr were cloned into pLVX-Het-1 or pLVX-Het-2, respectively (Clontech). The shortest predicted variant of mutated Prlr (residues 1 to 317) was used as the T Prlr isofrom, as shown in Figure S4. Stat1−/− murine embryonic fibroblasts were transduced with lentivirus expressing either FL or T Prlr alone or both FL and T Prlr together. In cells expressing both FL and T Prlr, FL Prlr was transduced first. PRLR-positive cells were sorted by flow cytometry using the anti-PRLR Ab clone 5A12, and sorted cells were subsequently transduced with T Prlr. Because MEFs did not express sufficient JAK2 to mediate signaling, Stat1−/− MEFs were also transduced with mouse JAK2:ires:GFP and sorted for GFP-positive cells.

Flow Cytometry

Cell surface expression of PRLR in Stat1−/− MEFs was confirmed using a biotinylated monoclonal antibody against murine PRLR (clone 5A12) (Chan et al., 2014) and streptavidin-PE (SA-PE, ebioscience). To examine basal activation of STAT3 and STAT5, MEFs were serum-deprived in 0.05% FBS for 16 hr before analysis. Cells were fixed and permeabilized according to manufacturer’s protocol (BD Biosciences). Tyrosine phosphorylation of STAT3 and STAT5a/b was detected using antibodies specific for the phosphorylated forms of each STAT (Cell Signaling). Monoclonal rabbit immunoglobulin G (IgG) was used as isotype control for gating.

Soft Agar Assay

Single cell suspensions of 20,000 or 50,000 Stat1−/− MEFs expressing FL, truncated PRLR, or both were mixed in 0.3% noble agar in DMEM and plated on top of 0.6% noble agar. Each condition was plated in triplicate in 960 dishes. The number of colonies in each dish was counted after 3 weeks using a Scan 100 colony counter.

Analysis of Tumorigenicity in Nude Mice

Female NCR nude mice (Taconic) were implanted with 1 × 10^6 of immortalized Stat1−/− MEFs expressing JAK2 alone, FL PRLR/JAK2, T PRLR/JAK2, FL/T PRLR/JAK2, or KRAS in 100 μL vehicle. Tumor diameter was measured twice weekly. Animals were censored when progressively growing palpable tumors of at least 2 mm were detected. On day 63, all remaining mice were sacrificed and evaluated for evidence of tumors prior to considering animals to be tumor-free.

Statistical Analyses

Wilcoxon rank sum test with continuity correction was performed to test for an association between PRLR isoform ratio and STAT expression group and sorted cells. Unpaired t test was used to determine the statistical significance between control and experimental groups in the soft agar assay (Figure 3B). All tests were two-sided and a p value ≤ 0.05 was considered significant. Fisher’s exact test was used to compare tumor formation in each group of nude mice implanted with Stat1−/− MEFs to the FL/T expressing group (Figure 3C). Kaplan-Meier survival analysis was performed using log-rank test. All statistical analysis was performed in R or GraphPad Prism.

ACKNOWLEDGMENTS

We thank the McDonnell Genome Institute’s Production group for sequence data production; Analysis Pipeline group for developing the automated sequence analysis pipelines; LIMS group for developing tools and software to manage samples and sequencing; and Systems group for providing the IT infrastructure and HPC solutions required for sequencing and analysis. We would also like to thank Nathan Dees and Scott Smith for their help with adapting the analysis pipelines for the mouse genome during the early phase of this project. O.L.G. was supported by the National Cancer Institute (NIH NCI K22 CA188163). M.G. was supported by the National Human Genome Research Institute (NIH NHGRI K09 HG007940). This work was funded by grants to R.K.W. from the National Human Genome Research Institute (NIH NHGRI U54 HG003079) and R.D.S from the National Cancer Institute (NIH NCI R01 CA043059) and Bristol-Myers Squibb.

REFERENCES


AUTHOR CONTRIBUTIONS


ACCESSION NUMBERS

The accession number for the reference-aligned whole genome sequence data and sample details for 52 tumor and non-tumor mouse tissues have been submitted to NCBI databases (SRA: SRP061941, BioProject: PRJNA248457).

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, six figures, ten tables, and four appendices and can be found with this article online at http://dx.doi.org/10.1016/j.celrep.2016.08.076.


