bioRxiv preprint doi: https://doi.org/10.1101/272815; this version posted February 27, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

# *NF1* deficiency correlates with estrogen receptor signaling and diminished survival in breast cancer

Patrick S. Dischinger[1], Elizabeth A. Tovar[1], Curt J. Essenburg[1], Zachary B. Madaj[2], Eve Gardner[1], Megan Callaghan[1], Ashley N. Turner[3], Anil K. Challa[3], Tristan Kempston[4], Bryn Eagleson[4], Robert A. Kesterson[3], Roderick, T. Bronson[5], Megan J. Bowman[2], Carrie R. Graveel[1¶], Matthew R. Steensma[1,6,7¶]

[1]Center for Cancer and Cell Biology, Van Andel Research Institute, Grand Rapids, MI, USA
[2]Bioinformatics & Biostatistics Core, Van Andel Research Institute, Grand Rapids, MI, USA
[3]Department of Genetics, The University of Alabama at Birmingham, Birmingham, AL, USA
[4]Vivarium and Transgenics Core, Van Andel Research Institute, Grand Rapids, MI, USA
[5]Rodent Histopathology Core Dana Farber/Harvard Cancer Center, Harvard Medical School, Boston, MA, USA
[6]Helen DeVos Children's Hospital, Spectrum Health System, Grand Rapids, MI, USA
[7]Michigan State University College of Human Medicine, Grand Rapids, MI, USA
[¶] Contributed equally

Corresponding Author:

Carrie R. Graveel, PhD
Van Andel Research Institute
333 Bostwick Ave. NE
Grand Rapids MI, 49503
Matt.Steensma@vai.org
Phone: 616-234-5378

Short Title: NF1 deficiency correlates with poor outcome in breast cancer

bioRxiv preprint doi: https://doi.org/10.1101/272815; this version posted February 27, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

**Abstract**

The key negative regulatory gene of the RAS pathway, *NF1,* is mutated or deleted in numerous cancer types and is associated with increased cancer risk and drug resistance. Even though women with neurofibromatosis (germline *NF1* mutations) have a substantially increased breast cancer risk at a young age and *NF1* is commonly mutated in sporadic breast cancers, we have a limited understanding of the role of *NF1* in breast cancer. Much of our understanding of the mechanisms underlying the functional loss of *NF1* comes from mouse models that do not completely recapitulate the phenotypes of human *NF1*. We utilized CRISPR-Cas9 gene editing to create *Nf1* rat models to evaluate the effect of *Nf1* deficiency on tumorigenesis. The resulting *Nf1* indels induced highly penetrant, aggressive mammary adenocarcinomas that express estrogen receptor and progesterone receptor. We identified distinct *Nf1* isoforms that were altered during tumorigenesis.

To evaluate *NF1* in human breast cancer, we analyzed genomic changes in a breast cancer dataset of 2,000 clinically annotated breast cancers. We found *NF1* shallow deletions in 25% of sporadic breast cancers, which correlated with poor clinical outcome. To identify biological networks impacted by *NF1* deficiency, we constructed gene co-expression networks using weighted gene correlation network analysis (WGCNA) and identified a network connected to *ESR1* (estrogen receptor). Moreover, *NF1*-deficient cancers correlated with established RAS activation signatures. Estrogen-dependence was verified by estrogen-ablation in *Nf1* rats where rapid tumor regression was observed. These results demonstrated the significant role NF1 plays in both NF1-related breast cancer and sporadic breast cancer.

bioRxiv preprint doi: https://doi.org/10.1101/272815; this version posted February 27, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

## Introduction

Deregulated RAS signaling promotes several "hallmarks of cancer", such as sustained proliferation, invasion and metastasis, and angiogenesis.[1] The importance of RAS deregulation in cancer is demonstrated by the fact that *KRAS* is the most commonly mutated oncogene and occurs at a high frequency in lung (17%), pancreatic (57%), and colon (33%) cancers.[2] Even though KRAS is mutated in only 4% of sporadic breast cancers (*HRAS* = 1%; *NRAS*=2%)[2], the RAS/ERK pathway is hyperactivated in approximately 50% of breast cancers.[3-5] This discrepancy suggests that there is another mechanism underlying RAS activation, besides mutation, in human breast cancers. The key negative regulatory gene of the RAS pathway, *NF1,* is mutated or deleted in a wide range of cancers and is increasingly recognized as a significant cancer driver. The *NF1* gene encodes neurofibromin, a GTPase-activating protein that regulates RAS (including HRAS, NRAS, and KRAS) and loss of neurofibromin function results in hyperactivated RAS.[6] Exons 20-27 encode the GAP-related domain (GRD), which is the RAS regulatory domain of human neurofibromin.[7] Mutations or deletions within the GRD result in decreased NF1 functionality and dysregulated RAS signaling driven through the RAS–RAF–MEK–ERK cascade.[8] Neurofibromatosis type 1 (NF1) is caused by germline mutations in the *NF1* gene and is the most common single-gene disorder affecting 1 in 3,000 live births.[9,10] The majority of NF patients develop benign cutaneous neurofibromas and may also develop peripheral nerve tumors, neurocognitive disorders, and bone stigmata (tibial dysplasia, scoliosis, and osteoporosis). Importantly, NF patients have an increased risk of developing several adult cancers including breast, ovarian, liver, lung, bone, thyroid, and gastrointestinal cancers.[11] These diverse clinical manifestations reveal the impact of loss of NF1 function and dysregulated RAS in numerous tissue types.

In addition to germline mutations, *NF1* mutations and deletions commonly occur in sporadic cancers and are associated with increased cancer risk and drug resistance. *NF1* is the

bioRxiv preprint doi: https://doi.org/10.1101/272815; this version posted February 27, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

third most prevalent mutated or deleted gene in glioblastoma,[12] fourth most mutated gene in ovarian cancer,[13] and the second most common mutated tumor suppressor in lung adenocarcinoma.[14] More recently, clinical evidence has mounted demonstrating that women with NF1 have a significantly increased breast cancer risk. A risk analysis of 3,672 NF patients found that women with NF1 have an increased relative risk of developing breast cancer in their younger years compared to the general population (relative risk was 6.5 at 30-39 years, 4.4 at 40-49 years, and 2.6 at 50-59 years).[15] Another key study of 1,404 NF patients identified an unequivocal increased risk for breast cancer with standardized incidence ratios of 11.1 (95% CI, 5.56 to 19.5) for breast cancer in women with NF1 age < 40 years and the overall breast cancer mortality ratio was 5.20 (95% CI, 2.38 to 9.88).[16] This study, in addition to several others, has established the increased breast cancer risk and associated poor outcome in patients with neurofibromatosis.[17] Comprehensive genomic analyses of sporadic breast cancers revealed that NF1 is commonly mutated and may be an important driver in sporadic breast cancer.[18,19] A study using the *Mcm4*$^{Chaos3}$ mouse model of chromosomal instability identified *Nf1* deletions in almost all of the *Mcm4*$^{Chaos3}$ mammary adenocarcinomas.[20] These findings suggest that NF1 is a critical tumor suppressor and potential driver of breast tumorigenesis.

Much of our understanding of the mechanisms underlying the functional loss of NF1 and tumorigenesis come from studies of genetically engineered mouse models.[21-24] These models have been valuable in defining how NF1 loss and deregulated RAS/MAPK signaling promote tumorigenesis. However, mouse models have general limitations, particularly with respect to reproducing human pharmacokinetics and recreating putative interactions between genetic/environmental factors and induced gene deficiency.[25] It is well established that mice with germline heterozygous *Nf1* mutations do not spontaneously express key aspects of the human phenotype and require additional crosses into other germline mutants such as *Tp53*, or cell-specific conditional *Nf1* mutation, to elicit more representative phenotypes.[22,26-30] Prior to the demonstration of CRISPR-Cas9 gene editing, manipulation of the rat genome was far more

bioRxiv preprint doi: https://doi.org/10.1101/272815; this version posted February 27, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

difficult than the mouse. The rat offers significant advantages over the mouse because of its larger size, more representative physiology to human disease, higher degree of cognition and memory, and ease of use in pharmaceutical studies.[31,32] In our study we utilized CRISPR-Cas9 gene editing capabilities to create several rat models of *Nf1* deletions in order to evaluate the effect of *Nf1* deficiency on tumorigenesis. The resulting *Nf1* indels induced highly penetrant, aggressive mammary adenocarcinomas in multiple rat founder lines. Moreover, both male and female *Nf1*-mutant animals developed mammary adenocarcinomas.

To evaluate the impact of *NF1* in sporadic breast cancer we analyzed genomic changes in a large breast cancer dataset composed of more than 2,000 clinically annotated breast cancers. We found that *NF1* shallow deletions are present in 25% of sporadic breast cancers and correlated with poor clinical outcome.[33] To identify biological networks impacted by *NF1* deficiency impacts we constructed gene co-expression networks using weighted gene correlation network analysis (WGCNA) and identified co-expression networks. A module associated with *NF1* shallow deletion contained several genes that are considerably important in both ER+ breast cancers and endocrine resistance, including *ESR1* and *FOXA1*. Unsupervised hierarchical clustering revealed that breast cancers with *NF1* shallow deletions form a distinct cluster that correlate with ER-negative breast cancer and RAS activation. We validated estrogen dependence of *Nf1*-deficient breast cancers by ablating mammary tumors in our *Nf1* rat model by ovarectomy. These results demonstrated the significant role *NF1* plays in both NF1-related breast cancer and sporadic breast cancer. Moreover, this novel *Nf1* rat model is invaluable for interrogating the role of NF1, estrogen-dependent breast cancer, and deregulated RAS signaling in sporadic and inherited breast cancer.

bioRxiv preprint doi: https://doi.org/10.1101/272815; this version posted February 27, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

## Results

Establishing *Nf1* rat models using CRISPR –Cas9 nucleases

To investigate the effects of altered *Nf1* function in rat, we used two unique sgRNAs to target the GRD region in exon 20 and disrupt neurofibromin function (Figure 1A). Two unique CRISPR/sgRNAs were synthesized and co-injected with Cas9 mRNA into one-cell-stage Sprague-Dawley rat embryos, which were then transferred to pseudopregnant females. From two rounds of injections, 19 pups were born. Genotyping was performed by amplifying a 452 bp fragment encompassing the sgRNAs target sites from DNA isolated from tail biopsies, followed by a heteroduplex mobility assay (HMA).[34] PCR amplicons with distinct HMA profiles were cloned into a plasmid vector and plasmids from multiple individual colonies were sequenced to identify the genetic lesions resulting from non-homologous end joining (NHEJ). Based on HMA profiles, indels were identified in 18 of 19 pups (Supplementary Tables 1 and 2). Multiple alleles were present in the majority of the G0 animals and were confirmed by Sanger sequencing, which revealed 34 mutant alleles, including 25 unique mutant alleles (Supplementary Table 2). Of the 34 mutations observed, 25 (73.5%) were frameshift mutations in both the 5' and 3' CRISPR target sites; 14 mutations (41.2%) were deletions ranging from 54-63 bp, spanning both the target sites. Interestingly, all the large deletions were in-frame with the translated protein coding sequence (Figure 1B-C). HMA profiles of multiple G0 animals revealed the presence of additional, smaller indels at the 5' and 3' CRISPR target regions (Figure 1B, i.e. lanes 3, 8, 10, etc.) as have been observed in other studies.[34,35] To verify the presence of the smaller indels and the larger in-frame deletions in each animal, we employed a three-step process of HMA, restriction digest, and Sanger sequencing (Figure 1D-E, Supplementary Table 1-2). The majority of smaller indels 10/19 (52.6%) were detected in the 5' CRISPR target region of G0 animals. Within the 5' CRISPR target region, there is a unique Hpy81 restriction site that is lost in each of the smaller indels. Consequently, PCR amplification and Hpy81 restriction

bioRxiv preprint doi: https://doi.org/10.1101/272815; this version posted February 27, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

digestion was used to confirm the presence of the smaller indels in each animal (Figure 1D). Each of these small indels at the 5' CRISPR target region resulted in premature stop codons in all cases (Figure 1E, Supplementary Table 1). Indels were also observed in the 3' CRISPR target region in 5/19 (26.3%) animals; however these were not predicted to have an effect on protein translation due to the presence of premature stop codons at the 5' CRISPR target region.

To investigate the effects of both the *Nf1* in-frame (referred to as **IF**) and premature stop (referred to as **PS**) indels, we bred four founder ($F_0$) rats (2 male and 2 female) that carried a mix of indels consisting of in-frame deletions and premature stops with wildtype Sprague-Dawley rats (Figure 1 and Supplementary Table 1). From two male founders we generated two unique lines: ***Nf1***$^{IF\text{-}57/+}$ from male founder #2 and ***Nf1***$^{IF\text{-}57/PS\text{-}8}$ derived from male #3. From two female founders, the lines ***Nf1***$^{IF\text{-}63/PS\text{-}11}$ (female #10) and ***Nf1***$^{IF\text{-}54/PS\text{-}11}$ (female #13) were generated. Segregation of the IF and PS alleles occurred in the G1 generations for lines ***Nf1***$^{IF\text{-}57/PS\text{-}8}$ and ***Nf1***$^{IF\text{-}54/PS\text{-}11}$ resulting in ***Nf1***$^{PS\text{-}8/+}$, ***Nf1***$^{PS\text{-}11/+}$, ***Nf1***$^{IF\text{-}57/+}$ lines (male #3). These *Nf1* lines were bred out for the following studies.

Upon performing PCR-HMA and sequence analysis, we observed the presence of more than two *Nf1* alleles in individual animals. In ***Nf1***$^{IF\text{-}63/PS\text{-}11}$ we identified 3 *Nf1* alleles: WT allele, an allele with an 11 bp deletion (premature stop), and an allele with a -63 bp deletion (Supplementary Figure 1, Supplementary Table 1). The presence of more than 2 alleles was confirmed independently by HMA and sequencing analysis using unique primer sets in two separate labs. Furthermore, these 3 alleles were transmitted through the germline in G1 and G2 generations and did not segregate in either generation (Supplementary Figure 1). Even though this finding was unanticipated, this is not unique in that the rat genome is known to have additional allelic copies of other genes.[36,37]

<u>*Nf1* female and male rats develop aggressive mammary adenocarcinomas</u>

bioRxiv preprint doi: https://doi.org/10.1101/272815; this version posted February 27, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

*Nf1* mutant rats (including in-frame and premature stop indels) were aged to determine the effects of *Nf1* alterations on tumor development. In *Nf1* female rats, we observed aggressive, multifocal tumorigenesis in the mammary glands at a young age (6-8 weeks). Mammary tumors were observed in 100% (11/11) of G0 females at the average age of 51 days (Supplementary Table 1). Tumor development progressed rapidly in all but one female (*rNf1*#19), whereas the majority of G0 females had to be euthanized before reaching a mature breeding age. We were able to breed two females before their tumor burden required euthanasia (*rNf1*#10 and #13) and support the pups through fostering. Multiple mammary adenocarcinomas (2 - 8) were observed in each animal and were detected in mammary glands 1-5. The detailed tumor pathology for each animal is described in Supplementary Table 3. Notably, two male founders (*rNf1*#2 and #3) developed mammary tumors at 14-16 months of age. Male #2 developed four unique mammary adenocarcinomas, whereas male #3 developed one mammary adenocarcinoma that had to be euthanized due to a larger fibrous histiocytic sarcoma. This aggressive mammary phenotype was highly penetrant and multifocal mammary tumors were observed in the subsequent G1 and G2 generations from each of the *Nf1* lines, including lines derived from male founders (Supplementary Table 3).

Histopathologic analysis of the mammary tumors revealed that both *Nf1* in-frame deletions and premature stop indels induced a wide variety of histopathologic mammary tumor types. Mammary tumors with acinar, solid, ductular, and cystic histology were observed in tumors in *Nf1* female rats from each line (Figure 2A). Moreover, *Nf1* animals often developed mammary adenocarcinomas with mixed histology such as acinar and solid (Figure 2A, left panel); cystic, papillary, ductal, and solid (Figure 2A, middle panel); and cystic, acinar, and solid (Figure 2A, right panel). This diverse mammary histology was observed in G0, G1, and G2 animals. In addition, we observed mixed histopathogy in the mammary adenocarcinomas from male *Nf1* rats (Figure 2B). The level of tumor burden in mammary pads of several animals was

bioRxiv preprint doi: https://doi.org/10.1101/272815; this version posted February 27, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

substantial, as can be observed in *rNf1* #413 (Supplementary Table 3). As shown in Supplementary Figure 1, one mammary gland from *Nf1* #413 contained 6 separate mammary tumors (2 additional larger tumors were separated for histopathology). Immunostaining of several tumors revealed that the *Nf1* mammary tumors were positive for estrogen receptor, progesterone receptor, and HER2 receptor and highly proliferative based on Ki67 staining (Figure 2C). The results demonstrate the impact of *Nf1* loss of function on mammary tumor initiation in both male and female hormonal environments.

*Nf1* premature stop indels result in more aggressive tumor progression

To evaluate how in-frame vs. premature *Nf1* indels affect disease burden and survival, we examined the overall survival of G1 heterozygous females. We performed a Kaplan-Meier analysis on overall survival of *Nf1* in-frame (IF) indels (n=35) compared to *Nf1* premature stop (PS) indels (n=24) in three rat lines ($Nf1^{IF\text{-}57/+}$, $Nf1^{IF\text{-}57/PS\text{-}8}$, $Nf1^{IF\text{-}54/PS\text{-}11}$). Survival analysis of tumor onset demonstrated that there was no statistical difference between the time of onset in premature stop compared to in-frame indels (Figure 3A). Conversely, analysis of disease-specific survival ($p<0.06$) and overall survival ($p< 0.0001$) revealed that animals with *Nf1* premature stop indels died due to tumor burden significantly faster than animals harboring *Nf1* in-frame deletions (Figure 3B-C). Notably, all animals in this study died of mammary tumor progression. We compared the overall survival of two $Nf1^{IF\text{-}57}$ lines that were derived from two separate founders (em2 and em3). The divergence in the tumor outcome in the *Nf1* in-frame indels compared to the *Nf1* premature stop indels suggests that the *Nf1* in-frame deletions may result in a partially functional neurofibromin protein and a less aggressive tumor phenotype. Interestingly, we observed that the em2-$Nf1^{IF\text{-}57}$ rats had a significantly shorter survival than em3-$Nf1^{IF\text{-}57}$ rats even though these *Nf1* lines had identical in-frame deletions ($p<0.02$; Figure 3D). The potential reason for this discrepancy in survival is described below. The fact that tumor initiation was comparable between the *Nf1* in-frame and *Nf1* premature stop indels

bioRxiv preprint doi: https://doi.org/10.1101/272815; this version posted February 27, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

suggests that a minimal decrease in NF1 function may be sufficient for mammary tumor initiation, yet a more significant loss of NF1 function (as in the *Nf1* premature stop indels) is required for rapid tumor progression.

Distinct neurofibromin isoform is expressed in mammary tissue

To understand how *Nf1* deficiency via mutation of the GRD domain affects neurofibromin expression and activity, we first examined the *Nf1* isoforms that are expressed in distinct $Nf1^{IF}$ and $Nf1^{PS}$ lines. Rat embryo fibroblasts (REFs) were isolated from day 13.5 rat embryos and mRNA was isolated. RT-PCR was performed using primers in exon 17 and exons 22-23 (Figure 4A). To differentiate RT-PCR products derived from the WT, IF, and PS alleles we performed restriction digests with Hpy81 endonuclease (Figure 4B). As demonstrated in Figure 1D-E, alleles harboring premature stop indels have lost the Hpy81 restriction site at the 5' CRISPR site. RT-PCR of WT REFs resulted in a strong band at 826 bp that created two bands at 390 bp and 368 bp with Hpy81 digestion. RT-PCR and Hpy81 digestion of the $Nf1^{PS\text{-}8}$ fibroblasts resulted in $Nf1^{WT}$ bands and an undigested band at 745 bp. Expression of the WT allele was observed in each of the *Nf1* in-frame and premature stop REF lines. We compared the em2-$Nf1^{IF\text{-}5}$ and em3-$Nf1^{IF\text{-}57}$ lines that had distinct survival curves (Figure 3D). Even though genotyping demonstrated both of these lines have a 57 bp deletion in exon 20, we observed distinct *Nf1* mRNA isoform expression in these lines (Figure 4B, notated with ∗). To determine the difference in the $Nf1^{IF\text{-}57}$ mRNA isoforms, we cloned and sequenced each of the products. Overall, we sequenced seventeen clones where the majority contained the predicted $Nf1^{IF\text{-}57}$ and $Nf1^{WT}$ isoforms, but we identified 3 unique clones harboring exon 21 deletions or extended exon 20 deletions resulting in *Nf1* premature stops. For example, one mRNA isoform contained the -57 bp in-frame deletion in exon 20 and an additional 140 bp deletion in exon 21 (Figure 4C). This exon 21 deletion mRNA was exclusive to the em2-$Nf1^{IF\text{-}57}$ line and resulted in a

bioRxiv preprint doi: https://doi.org/10.1101/272815; this version posted February 27, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

premature stop. These results substantiate the requirement of a functional GRD domain within NF1 to suppress tumor progression.

Several unique isoforms have been characterized in the *Nf1* gene with the main isoform identified at 250 kD; however the majority of these studies have been performed in brain, neuronal, and muscle tissues.[38-40] To our knowledge, neurofibromin expression has not been previously evaluated in the mammary glands of humans or rodents. To understand the effect of *Nf1* deficiency in the mammary gland, we first performed Western blot analysis of neurofibromin in the normal mammary, brain, muscle, spleen, and liver of adult rats and observed distinct isoform expression patterns among the tissue types (Figure 4D). Even though the established 250 kD neurofibromin isoform is highly expressed in brain tissue, its expression was significantly less in the other tissues including the mammary gland. In the mammary gland, the predominant neurofibromin isoform was observed at 125 kD. This 125 kD isoform was also observed in the muscle, spleen, and liver, yet minimally expressed in the brain. In Figure 4D, we compared neurofibromin expression in the brain, normal mammary glands, and mammary adenocarcinomas from $Nf1^{WT}$, $Nf1^{IF}$, and $Nf1^{PS}$ rats. In the brain, the 250 kD isoform was highly and equally expressed in the $Nf1^{WT}$, $Nf1^{IF}$, and $Nf1^{PS}$ animals. In the normal mammary gland there was no difference in expression of the 250 kD isoform between the $Nf^{WT}$, $Nf1^{IF}$, and $Nf1^{PS}$ animals, however there was a substantial decrease in expression of the 125 kD isoform in both the $Nf1^{IF}$ and $Nf1^{PS}$ mammary glands compared to WT mammary glands. In mammary adenocarcinomas, 125 kD isoform was further decreased in both the $Nf1^{IF}$ and $Nf1^{PS}$ tumors and there appeared to be a slight loss of the 250 kD isoform in the $Nf1^{IF}$ mammary tumor. Overall these results demonstrate the distinct neurofibromin isoform tissue preferences that are present and the inverse relationship of the 250 kD vs. 125 kD isoforms in the brain and mammary tissues.

bioRxiv preprint doi: https://doi.org/10.1101/272815; this version posted February 27, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

*NF1* is commonly deleted in human sporadic breast cancer and correlates with ER networks

Even though it is known that neurofibromatosis patients have a significantly increased breast cancer risk and that *NF1* is mutated in sporadic breast cancers, the impact of NF1 deficiency and RAS deregulation in sporadic breast cancer is often overlooked. To interrogate the impact of *NF1* loss in breast cancer we performed an analysis of the METABRIC breast cancer dataset which contained 2051 patients with clinical annotation including CNV and SNP genotypes.[33] In this cohort, there were 43 patients (2.1%) with truncating mutations, including frameshift deletions or insertions, nonsense mutations, and splice site alterations (Figure 5A). Notably, these mutations occurred throughout the *NF1* gene similar to the mutation diversity observed in neurofibromatosis patients. After removing patients with *NF1* SNP mutations, we analyzed *NF1* copy number alterations (CNAs) and identified CNAs in 32.9% of patients, with 24.5% having *NF1* shallow deletions (defined as potential heterozygous deletions). To determine the effect of *NF1* shallow deletions on survival, we conducted a survival analysis using the Cox proportional hazards model with random effects (frailty model). When accounting for ER status and age at diagnosis, a patient with a *NF1* shallow deletion is 1.65X ($p < 1.6e-05$) more likely to die within the first 10 years compared to patients with diploid *NF1* status (Figure 5B).

To identify gene expression networks that correlate with *NF1* copy number status, we utilized weighted gene co-expression network analysis (WGCNA).[41] From the METABRIC dataset, expression data were analyzed for patients with available copy number and gene expression data (n=1427). Since *NF1* is in close proximity to *HER2* on chromosome 17 and is often co-amplified, we removed HER2-amplified patients from this analysis. From the 18,049 genes in the METABRIC dataset, 2,218 were placed in 12 co-expression modules. Modules contained between 30-500 genes, with an average module size of 200 genes. Interestingly, one module (179 genes) associated with *NF1* shallow deletion copy number status contained

bioRxiv preprint doi: https://doi.org/10.1101/272815; this version posted February 27, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

several genes that are considerably important in both ER+ breast cancers and endocrine resistance, including *ESR1* (Figure 5C). Hub genes for this module, or genes with high connectivity, included *FOXA1* and *MLPH*. *FOXA1*, a Forkhead family transcription factor, regulates ER binding and transcriptional activity.[42] *FOXA1* expression correlates with luminal subtype A breast cancer and is a significant predictor of cancer-specific survival in ER+ tumors.[43] *MLPH* expression is also associated with longer survival in breast cancer.[44] Further network visualization of genes within this ER-associated co-expression network revealed additional connections with *AP1*, *AR* (androgen receptor) and *C/EBP*$\alpha$. AP1 is also known as the transcription factor JUN. RAS/ERK signaling is known to regulate JUN at the transcriptional and post-translational (phosphorylation) level and can modify lysine acetylation in JUN DNA binding regions.[45] Recent studies have demonstrated that *AR* is expressed in 77% of breast cancers (88% ER+, 59% HER2+, 32% TNBC)[46] and is involved in endocrine resistance in ER+ breast cancers.[47,48] Although there are studies demonstrating interaction of C/EBP (CCAAT/enhancer binding protein) at ER transcriptional binding sites, the consequence of C/EBP activity in breast cancer is unclear.

Unsupervised clustering of genes from this WGCNA module with METABRIC clinical data indicates a strong association with patient ER status, specifically ER-negative breast cancer (Figure 5D). To understand if shallow deletions of *NF1* correlate with deregulated RAS signaling, we employed expression signatures developed from KRAS-mutant cancers and performed a secondary WGCNA analysis and unsupervised clustering with 788 genes related to ER, RAS and KRAS signaling pathways. This analysis distinguished three modules associated with *NF1* copy number status, with one module associated specifically with *NF1* shallow deletions (Figure 5E). These results validate the presence of RAS activation in *NF1*-related breast cancers. Moreover, our WGCNA analyses revealed novel connections between NF1 deficiency, RAS signaling, and ER signaling.

bioRxiv preprint doi: https://doi.org/10.1101/272815; this version posted February 27, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

*Nf1*-deficient tumors are estrogen-dependent

To examine whether our *Nf1* deficient rat breast cancer model was estrogen dependent, we performed ovariectomies on rats containing at least one tumor greater than 1500 $mm^3$. Ovariectomies were performed on 5 rats with multiple tumors from three distinct Nf1 lines (Figure 6F and Supplementary Figure 3). Upon ovary removal, tumor size diminished at a surprisingly rapid pace (mean reduction per day = 5.33%; CI [4.89, 5.78]; $p < 0.000000001$). Tumor reduction was observed in all of the tumors regardless of Nf1 status (in-frame vs. premature stop indels) and the mean total percent reduction in tumor volume was 90.95%; (CI [78.79, 97.28], $p = 0.0000018$) (Figure 6G and Supplementary Figure 3). These results confirm estrogen-dependency in the rat *Nf1*-deficient breast cancer model and support the biological significance of the NF1-ER networks that were identified by WGCNA.

bioRxiv preprint doi: https://doi.org/10.1101/272815; this version posted February 27, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

## Discussion

The RAS signaling pathway is intricately involved in the initiation and progression of numerous cancer types; however the mechanism by which RAS is deregulated in breast cancer is unclear. Here we present data that illuminates the role that *NF1* plays in breast cancer and a novel model for interrogating the interaction between NF1 and ER in both sporadic and NF1-related breast cancer. Even though NF patients develop a variety of cancers, the high breast cancer risk in women with neurofibromatosis was not been highlighted until recently.[15-17] *NF1* was firsts implicated in sporadic breast cancer by the initial comprehensive genome breast cancer studies where *NF1* mutations were identified in 2-4% of breast cancers.[18,19] Our analysis of the extensive METABRIC dataset verified the *NF1* mutation frequency, but also discovered *NF1* shallow deletions in 25% of sporadic breast cancers. This verifies the *NF1* shallow deletions that have been identified in the TCGA breast cancer dataset.[20] Most importantly, our analysis revealed that *NF1* shallow deletions correlated with poor outcome after the first 10 years. One clinical challenge for oncologists is determining which patients primary breast cancer should be treated aggressively. Our results indicate that *NF1* copy number may be a potential prognostic indicator in breast cancer. In addition to *NF1*, the RasGAP gene *RASAL2* is altered in a sporadic breast cancer. McLaughlin et al. demonstrated that *RASAL2* expression is substantially decreased in luminal B breast cancers and low *RASAL2* expression correlates with metastasis and poor survival.[49] Our WGCNA analysis revealed an unknown network between NF1, ER, and RAS signaling in breast cancer. Together, these studies indicate that loss of RAS suppression may be more important in breast cancer than activation of RAS signaling through mutation or amplification.

Together our CRISPR rat *Nf1* model and our WGCNA analysis of human breast cancer indicate that NF1 is intricately connected to ER networks, including several genes that have been directly linked to endocrine resistance (i.e FOXA1 and AR). The rate of tumor regression in the *Nf1* tumors after estrogen ablation was remarkably rapid and the fact that all of the *rNf1*

bioRxiv preprint doi: https://doi.org/10.1101/272815; this version posted February 27, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

tumors diminished regardless of mutation (PS or IF) suggests that *Nf1*-deficienct tumors are universally estrogen-dependent. Interrogating the mechanisms by which NF1 interacts with ER will involve a careful analysis both transcriptional and signaling interactions. NF1 is a large protein and there is a limited knowledge of NF1 protein-protein interactions besides the RAS interactions at the GRD domain. It is possible that NF1 may directly interact with genes involved identified in the NF1-ER network at a transcriptional level, yet indirect signaling NF1-RAS-ER signaling interactions are also likely since RAS signaling regulates JUN (AP1) at the transcriptional and post-translational level.[45] A recent study that assessed driver genes during the genomic evolution of breast cancer metastasis revealed that *NF1* mutations substantially increase in recurrent breast cancers, but this increase was only observed in ER-positive breast cancers[50]. This findings in addition to 1) the correlation of *NF1* shallow deletions with poor outcome, 2) network connectivity between NF1-ER-FOXA1-AR, and 3) estrogen-dependency in the Nf1 tumors, indicate that NF1 may be a key player in endocrine reistance.

The *Nf1* rat model also revealed some unexpected alternative mRNA and proteins isoforms that underscore our limited understanding of NF1 expression. For example, we observed two distinct neurofibromin isoforms that are differentially expressed in the brain and mammary gland. Both in-frame and premature stop sequence variants resulted in significantly reduced neurofibromin mammary expression, whereas expression of the 250kD isoform was preserved in the brain regardless of *Nf1* status. It is unclear why the wild type allele was unable to effectively compensate for heterozygous loss of NF1 function in the mammary gland. Collectively, these findings suggest that RAS regulation in the breast is distinct from other tissues, and that breast cancer predisposition may be linked to a unique isoform of neurofibromin that is abundantly expressed in mammary tissue.

One issue that vexed us initially was the divergence of survival between two genotype-matched $Nf1^{IF\text{-}57}$ lines. Analysis of the mRNA in em2 $Nf1^{IF\text{-}57}$ and em3 $Nf1^{IF\text{-}57}$ lines identified an

bioRxiv preprint doi: https://doi.org/10.1101/272815; this version posted February 27, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

additional deletion in exon 21 in em2 *Nf1*$^{IF\text{-}57}$. Differential mRNA isoform expression has been observed in breast cancer and may occur through alternate promoter usage, alternate splicing, and alternate 3'UTR usage.[51] Importantly, differential mRNA isoform expression has been identified and associated with distinct breast cancer subtypes. In NF patients, there is often a discrepancy between genotype and phenotype. For example, one NF patient may develop numerous cutaneous and plexiform neurofibromas, where another patient with an identical *NF1* mutation may have few cutaneous neurofibromas and develop breast cancer at a young age. Our data raises the possibility that *NF1* isoform expression may play a role in these clinical inconsistencies.

In summary, we have developed a novel *Nf1* rat model is invaluable for interrogating deregulated RAS signaling and ER networks in sporadic and NF-related breast cancer. Moreover, we identified a frequent mechanism for RAS deregulation through *Nf1* shallow deletion that is present in 25% of sporadic breast cancers. The correlation of *Nf1* shallow deletions with poor prognosis and ER-FOXA1-AR networks indicates that NF1 may be an important prognostic indicator and therapeutic target in endocrine-resistant breast cancers.

bioRxiv preprint doi: https://doi.org/10.1101/272815; this version posted February 27, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

## Methods

### CRISPR/sgRNA design and synthesis, and Cas9 mRNA synthesis

CRISPR targets were identified in exon 20 of the rat *Nf1* gene (NCBI rn5 rat genome) using an online design tool (crispr.mit.edu). A region near the middle of exon 20 was targeted by one guide sequence (5' CRISPR), GCCCTGTCAGTGAACGCAAA (*GGG*), while a downstream region was targeted by a second guide sequence (3' CRISPR), CGGTCCATAAATCTGCTGAC (*AGG*). The rn5 rat genome reference sequence and annotation for CRISPR target design.

### Rat breeding, microinjection into zygotes and blastocyst culture

Sprague-Dawley rats were purchased from Charles River Laboratories. Female rats were superovulated with injection of 20 IU PMSG, followed by injection of 50 IU HCG and immediate mating to CD Sprague Dawley studs 48 hours later. The pronuclei of fertilized rat zygotes were injected with 40 ng/µL Cas9 mRNA and 20 ng/µL each of two sgRNAs. Surviving eggs were surgically transferred to CD Sprague-Dawley pseudopregnant females. All animal experiments were performed in accordance with institutional IACUC protocols.

### Detecting the presence of indels by PCR-HMA and amplicon sequencing

Genomic DNA from rat-tail biopsies was obtained by phenol extractions. DNA samples were directly used as a PCR template to amplify a region flanking the CRISPR target site. PCRs were set up using the following oligonucleotide primers: Ex20-F 5'-TCA ACA TGA CTG GCT TCC TC–3' and Ex20-R 5'-CAT TGG ATA CAG AGC AGG ACT C–3' to obtain a 284 bp fragment. The amplicons were subjected to denaturation-slow renaturation to facilitate formation of heteroduplexes using a thermocycler. These samples were then resolved on polyacrylamide gels (6%) and the resulting mobility profiles used to infer efficiency of CRISPR-Cas9 nuclease activity. PCR using same primers flanking CRISPR target site (Ex20-F and Ex20-R) was performed to amplify regions with indels (284 bp). The amplified products were cloned using the TOPO-TA cloning kit (Invitrogen, Carlsbad, CA). Ten representative colonies were picked from

bioRxiv preprint doi: https://doi.org/10.1101/272815; this version posted February 27, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

each plate and grown in 4mL liquid cultures to isolate plasmid DNA. Plasmid DNA was sequenced using M13F and M13R primers.

Genotyping

Rat DNA was collected from tail clips by phenol extraction (Invitrogen 15593-031). Target regions of the DNA were PCR amplified using the forward primer rNF1-1A 5'– CTT AGG CTG CAG AAA GTC TTC – 3' and reverse primer rNF1-1B 5'– CTT CAC CTG TCC TTG AGA GTC – 3'. PCR products were digested with Hpy8I (Thermo Fisher Scientific). After digestion, DNA fragments were separated by electrophoresis on a 2% agarose gel and stained with ethidium bromide. Premature stop (439 bp) and in-frame (185 bp) mutations were identified by gel imaging.

RT-PCR

RNA was isolated from rat tissue by Trizol extraction (Invitrogen). RNA was treated with DNase I (Thermo Fisher Scientific #18068015). Complementary DNA (cDNA) was preapred by using random primers (Life Technologies 48190-011) and SuperScript™ II Reverse Transcriptase (Invitrogen 18064-022). PCR was performed with forward primer (5' – TCAGTACACAACTTCTTGCC – 3') and reverse primer (5' – GAACACGAACATATCTGACC – 3') to create a 826 bp amplicon. These PCR products were digested with the restriction enzyme Hpy8I as described above to identify mutations. Digested PS amplicons make a 745 bp product and digested wildtype make a 390 bp and 368 bp product. Digested PCR products were ran on a 10% TGX-PAGE gel (BioRad) and stained with ethidium bromide.

Histology

Tumors were fixed and processed following the previous methods [52]. DAKO PT Link was used for antigen retrieval. Immunohistochemistry staining was done with ER (Thermo MA5-13304), HER2 (Thermo MA5-13105), Cytokeratin (DAKO AE1/AE3), and PR (Thermo MA1-411) using

bioRxiv preprint doi: https://doi.org/10.1101/272815; this version posted February 27, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

the DAKO Autostainer Link 48. Ki67 (Spring Bioscience SP6) was stained using the Ventanna Discovery Ultra.

Rat Embryo Fibroblast (REF) Isolation

At embryonic day 13.5 embryos were harvested and isolated independently in petri dishes with PBS. The placenta was removed followed by the yolk sac. Embryos were then dissected by removing the head, organs, and blood vessels. The remaining tissue was transferred to a new petri dish and finely chopped using surgical blades. Tissue was then digested in 0.25% trypsin for 15 minutes at 37C. Cells were passed through a cell strainer to isolate single cells. Cells were plated and grown in DMEM (Gibco 11995-065) supplemented with 10% FBS and 1% penicillin.

Immunoblot Analysis

Rat tissue was homogenized with a hand held pellet pestle in RIPA buffer (20 mM TrisHCL, pH 7.6; 5 mM EDTA; 150 mM NaCl; 0.5% NP-40; 50 mM NaF; 1 mM beta-glycerpphosphate) supplemented with PhosSTOP (Roche) and protease inhibitors (Roche). Samples were resolved by SDS-polyacrylamide gel electrophoresis on a 7.5% TGX-PAGE gel (BioRad) and transferred overnight at 4°C. Immunoblotting was performed using the following antibodies: neurofibromin H-12 (SC-376886) and D (SC-67) from Santa Cruz Technology; McNFn27b #MA1-085 from Thermo Scientific; β-actin (#3700), HSP90 (#4877) and GAPDH (#2118) from Cell Signaling Technology.

Statistical Analysis

For analysis of *Nf1* rat survival, data were plotted using Kaplan-Meier curves and p-values were calculated using a Cox mixed-effects model with a frailty term for litter (https://CRAN.R-project.org/package=coxme). Linear contrasts with Benjamini-Hochberg false discovery rate adjustments were used to test specific hypotheses. All analyses described below were

bioRxiv preprint doi: https://doi.org/10.1101/272815; this version posted February 27, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

performed using R v 3.4.0. (https://cran.r-project.org/) All hypotheses were two-sided with a significance level of 0.05. Ovariectomy analysis was analyzed using a multi-level mixed-effects beta regression with a random intercept for tumor, clustered within rat lines. This model was fit using the R package glmmTMB[53] and the squeeze algorithm.[54]

METABRIC analysis

Survival analysis was conducted in R v(3.4.2) using the packages survival (https://cran.r-project.org/web/packages/survival/index.html), survminer (https://cran.r-project.org/web/packages/survminer/index.html), and coxme. Clinical, copy number and gene expression data from the METABRIC dataset[33] were accessed using the cgdsr (https://cran.r-project.org/web/packages/cgdsr/index.html ) package in R. Patients were selected based on the availability of both copy number and gene expression data, and HER2 negative status. Gene-co expression network analysis was conducted using the WGCNA (v1.61) package in R to identify gene coexpression modules from mRNA expression data (Illumina Human v3 Microarray).[41] Data were normalized as previously described. Unsigned correlations were used with a soft threshold value β of 10 and a treecut value of 0.15, a minimum number of genes in the module set to 30. The β and treecut parameters were chosen after assessing the quality of modules detected, and all other parameters used default settings. Unsupervised clustering of WGCNA results was run using the pheatmap package in R (v1.0.8). Specific gene networks were visualized using MetaCore from Thomson Reuters (v6.33 build 69110). The RAS and KRAS gene lists used in the analysis were provided by the Broad Institute MSigDB database (v6.1) Hallmark gene set collection and BioCarta (c) (2000-2017 BioCarta, all rights reserved).

bioRxiv preprint doi: https://doi.org/10.1101/272815; this version posted February 27, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

**Acknowledgements:** We dedicate this work in honor of Patricia Graveel and all neurofibromatosis patients dealing with breast cancer. The authors would like to thank Dr. Ben Johnson and Jamie Grit for their critical review of this manuscript, Dr. Casey Droscha and Dr. Bart Williams for assistance with CRISPR designs, and also Lisa Turner (VAI Histology) and the VAI Vivarium staff for their strong work on this project. Funding for this research was made possible by the Breast Cancer Research Foundation (BCRF-17-159) and the VARI Faculty Innovation Award.

## Author Contributions

Study conception and design: PSD, CRG, MRS; acquisition and analysis of data: PSD, EAT, EJE, ZM, EEG, MC, ANT, AKC, TK, RTB, MJB, CRG, MRS; provided reagents: BE, RAK; drafting of manuscript: CRG, MRS; critical revision: PSD, EAT, ZM, MJB; CRG and MRS are the guarantors of this study.

## ADDITIONAL INFORMATION

**Supplementary information** is available at the NPJ Breast Cancer website.

**Competing interests:** The author declares that they have no competing financial interests.

bioRxiv preprint doi: https://doi.org/10.1101/272815; this version posted February 27, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

## References


1. Hanahan, D. & Weinberg, R.A. Hallmarks of cancer: the next generation. *Cell* **144**, 646-74 (2011).
2. Pylayeva-Gupta, Y., Grabocka, E. & Bar-Sagi, D. RAS oncogenes: weaving a tumorigenic web. *Nat Rev Cancer* **11**, 761-74 (2011).
3. Sivaraman, V.S., Wang, H., Nuovo, G.J. & Malbon, C.C. Hyperexpression of mitogen-activated protein kinase in human breast cancer. *J Clin Invest* **99**, 1478-83 (1997).
4. von Lintig, F.C. *et al.* Ras activation in human breast cancer. *Breast Cancer Res Treat* **62**, 51-62 (2000).
5. Mueller, H. *et al.* Potential prognostic value of mitogen-activated protein kinase activity for disease-free survival of primary breast cancer patients. *Int J Cancer* **89**, 384-8 (2000).
6. Bollag, G. *et al.* Loss of NF1 results in activation of the Ras signaling pathway and leads to aberrant growth in haematopoietic cells. *Nat Genet* **12**, 144-8 (1996).
7. Upadhyaya, M. *et al.* Mutational and functional analysis of the neurofibromatosis type 1 (NF1) gene. *Hum Genet* **99**, 88-92 (1997).
8. Ratner, N. & Miller, S.J. A RASopathy gene commonly mutated in cancer: the neurofibromatosis type 1 tumour suppressor. *Nat Rev Cancer* **15**, 290-301 (2015).
9. Lammert, M., Friedman, J.M., Kluwe, L. & Mautner, V.F. Prevalence of neurofibromatosis 1 in German children at elementary school enrollment. *Arch Dermatol* **141**, 71-4 (2005).
10. Evans, D.G. *et al.* Birth incidence and prevalence of tumor-prone syndromes: estimates from a UK family genetic register service. *Am J Med Genet A* **152A**, 327-32 (2010).
11. Seminog, O.O. & Goldacre, M.J. Risk of benign tumours of nervous system, and of malignant neoplasms, in people with neurofibromatosis: population-based record-linkage study. *Br J Cancer* **108**, 193-8 (2013).
12. Comprehensive genomic characterization defines human glioblastoma genes and core pathways. *Nature* **455**, 1061-1068 (2008).
13. Integrated genomic analyses of ovarian carcinoma. *Nature* **474**, 609-15 (2011).
14. Ding, L. *et al.* Somatic mutations affect key pathways in lung adenocarcinoma. *Nature* **455**, 1069-75 (2008).
15. Seminog, O.O. & Goldacre, M.J. Age-specific risk of breast cancer in women with neurofibromatosis type 1. *Br J Cancer* **112**, 1546-8 (2015).
16. Uusitalo, E. *et al.* Distinctive Cancer Associations in Patients With Neurofibromatosis Type 1. *J Clin Oncol* **34**, 1978-86 (2016).
17. Howell, S.J., Hockenhull, K., Salih, Z. & Evans, D.G. Increased risk of breast cancer in neurofibromatosis type 1: current insights. *Breast Cancer (Dove Med Press)* **9**, 531-536 (2017).
18. Stephens, P.J. *et al.* The landscape of cancer genes and mutational processes in breast cancer. *Nature* **486**, 400-4 (2012).
19. Comprehensive molecular portraits of human breast tumours. *Nature* **490**, 61-70 (2012).

bioRxiv preprint doi: https://doi.org/10.1101/272815; this version posted February 27, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

20. Wallace, M.D. *et al.* Comparative oncogenomics implicates the neurofibromin 1 gene (NF1) as a breast cancer driver. *Genetics* **192**, 385-96 (2012).

21. Bajenaru, M.L. *et al.* Astrocyte-specific inactivation of the neurofibromatosis 1 gene (NF1) is insufficient for astrocytoma formation. *Mol Cell Biol* **22**, 5100-13 (2002).

22. Zhu, Y. *et al.* Inactivation of NF1 in CNS causes increased glial progenitor proliferation and optic glioma formation. *Development* **132**, 5577-88 (2005).

23. Cichowski, K. *et al.* Mouse models of tumor development in neurofibromatosis type 1. *Science* **286**, 2172-6 (1999).

24. Zhu, Y. *et al.* Early inactivation of p53 tumor suppressor gene cooperating with NF1 loss induces malignant astrocytoma. *Cancer Cell* **8**, 119-30 (2005).

25. Lin, J.H. Applications and limitations of genetically modified mouse models in drug discovery and development. *Curr Drug Metab* **9**, 419-38 (2008).

26. Brannan, C.I. *et al.* Targeted disruption of the neurofibromatosis type-1 gene leads to developmental abnormalities in heart and various neural crest-derived tissues. *Genes Dev* **8**, 1019-29 (1994).

27. Silva, A.J. *et al.* A mouse model for the learning and memory deficits associated with neurofibromatosis type I. *Nat Genet* **15**, 281-4 (1997).

28. Vogel, K.S. *et al.* Mouse tumor model for neurofibromatosis type 1. *Science* **286**, 2176-9 (1999).

29. Reilly, K.M., Loisel, D.A., Bronson, R.T., McLaughlin, M.E. & Jacks, T. Nf1;Trp53 mutant mice develop glioblastoma with evidence of strain-specific effects. *Nat Genet* **26**, 109-13 (2000).

30. Bajenaru, M.L. *et al.* Optic nerve glioma in mice requires astrocyte Nf1 gene inactivation and Nf1 brain heterozygosity. *Cancer Res* **63**, 8573-7 (2003).

31. Russo, J. Significance of rat mammary tumors for human risk assessment. *Toxicol Pathol* **43**, 145-70 (2015).

32. Aitman, T., Dhillon, P. & Geurts, A.M. A RATional choice for translational research? *Disease Models & Mechanisms* **9**, 1069-1072 (2016).

33. Curtis, C. *et al.* The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups. *Nature* **486**, 346-52 (2012).

34. Challa, A.K. *et al.* Novel Hypomorphic Alleles of the Mouse Tyrosinase Gene Induced by CRISPR-Cas9 Nucleases Cause Non-Albino Pigmentation Phenotypes. *PLoS One* **11**, e0155812 (2016).

35. Lambert, L.J. *et al.* Increased trabecular bone and improved biomechanics in an osteocalcin-null rat model created by CRISPR/Cas9 technology. *Dis Model Mech* **9**, 1169-1179 (2016).

36. Liu, R.Z., Li, X. & Godbout, R. A novel fatty acid-binding protein (FABP) gene resulting from tandem gene duplication in mammals: transcription in rat retina and testis. *Genomics* **92**, 436-45 (2008).

37. Aitman, T.J. *et al.* Copy number polymorphism in Fcgr3 predisposes to glomerulonephritis in rats and humans. *Nature* **439**, 851-5 (2006).

bioRxiv preprint doi: https://doi.org/10.1101/272815; this version posted February 27, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

38. Gutmann, D.H., Geist, R.T., Wright, D.E. & Snider, W.D. Expression of the neurofibromatosis 1 (NF1) isoforms in developing and adult rat tissues. *Cell Growth Differ* **6**, 315-23 (1995).

39. Andersen, L.B. *et al.* A conserved alternative splice in the von Recklinghausen neurofibromatosis (NF1) gene produces two neurofibromin isoforms, both of which have GTPase-activating protein activity. *Mol Cell Biol* **13**, 487-95 (1993).

40. Gutmann, D.H., Geist, R.T., Rose, K. & Wright, D.E. Expression of two new protein isoforms of the neurofibromatosis type 1 gene product, neurofibromin, in muscle tissues. *Dev Dyn* **202**, 302-11 (1995).

41. Langfelder, P. & Horvath, S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics* **9**, 559 (2008).

42. Hurtado, A., Holmes, K.A., Ross-Innes, C.S., Schmidt, D. & Carroll, J.S. FOXA1 is a key determinant of estrogen receptor function and endocrine response. *Nature Genetics* **43**, 27 (2010).

43. Badve, S. *et al.* FOXA1 Expression in Breast Cancer—Correlation with Luminal Subtype A and Survival. *Clinical Cancer Research* **13**, 4415-4421 (2007).

44. Thakkar, A. *et al.* High Expression of Three-Gene Signature Improves Prediction of Relapse-Free Survival in Estrogen Receptor-Positive and Node-Positive Breast Tumors. *Biomarker Insights* **10**, BMI.S30559 (2015).

45. Lopez-Bergami, P., Lau, E. & Ronai, Z.e. Emerging roles of ATF2 and the dynamic AP1 network in cancer. *Nature Reviews Cancer* **10**, 65 (2010).

46. Collins, L.C. *et al.* Androgen receptor expression in breast cancer in relation to molecular phenotype: results from the Nurses' Health Study. *Mod Pathol* **24**, 924-31 (2011).

47. Ciupek, A. *et al.* Androgen receptor promotes tamoxifen agonist activity by activation of EGFR in ERalpha-positive breast cancer. *Breast Cancer Res Treat* **154**, 225-37 (2015).

48. D'Amato, N.C. *et al.* Cooperative Dynamics of AR and ER Activity in Breast Cancer. *Mol Cancer Res* **14**, 1054-1067 (2016).

49. McLaughlin, S.K. *et al.* The RasGAP gene, RASAL2, is a tumor and metastasis suppressor. *Cancer Cell* **24**, 365-78 (2013).

50. Yates, L.R. *et al.* Genomic Evolution of Breast Cancer Metastasis and Relapse. *Cancer Cell* **32**, 169-184 e7 (2017).

51. Stricker, T.P. *et al.* Robust stratification of breast cancer subtypes using differential patterns of transcript isoform expression. *PLoS Genet* **13**, e1006589 (2017).

52. Linklater, E.S. *et al.* Targeting MET and EGFR crosstalk signaling in triple-negative breast cancers. *Oncotarget* **7**, 69903-69915 (2016).

53. Brooks, M.E. *et al.* Modeling Zero-Inflated Count Data With glmmTMB. *bioRxiv* (2017).

54. Smithson, M. & Verkuilen, J. A better lemon squeezer? Maximum-likelihood regression with beta-distributed dependent variables. *Psychol Methods* **11**, 54-71 (2006).

bioRxiv preprint doi: https://doi.org/10.1101/272815; this version posted February 27, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

## Figure Legends

**Figure 1: In-frame and premature indels are present in *Nf1* G0 rats.** A) CRISPR-Cas9 targeting of the *Nf1* GRD domain in exon 20 using two sgRNAs (red). PAM sequences (green) and the Hpy81 site (purple) used for HMA/restriction digest analysis are shown. B) Heteroduplex PCR amplicons from 19 *Nf1* G0 animals. C) DNA and peptide sequence alignment of the in-frame deletions within *Nf1*exon 20. D) Heteroduplex PCR amplicons from 19 *Nf1* G0 animals were digested with Hpy81 to reveal the presence of smaller indels. No change between the –Hpy81 and +Hpy81 indicates the Hpy81 restriction digestion was lost during CRISPR recombination events. E) DNA and peptide sequence alignment of the small indels at the 5' and 3' CRISPR target region.

**Figure 2: *Nf1* deficiency induces $ER^+/PR^+$ mammary tumors.** We observed mammary tumors with diverse mammary histology in F0-G2 animals. A) *Nf1* females and B) *Nf1* males developed mammary adenocarcinomas with mixed histology including acinar, cystic, papillary, ductal, and solid features. All images were taken at 100X magnification. C) Immunostaining of *Nf1* mammary tumors (rNf1 #6) for estrogen receptor, progesterone receptor, HER2, Ki67, and pan-cytokeratin (CK AE1/AE3). All immunostaining images were taken at 200X magnification.

**Figure 3: Survival analysis of $Nf1^{IF}$ vs $Nf1^{PS}$ reveals distinct effects of Nf1 deficiency on tumor onset and survival.** Survival analysis was performed on $Nf1^{IF}$ (n=35) vs $Nf1^{PS}$ (n=24) in the $Nf1^{IF\text{-}57/+}$, $Nf1^{IF\text{-}57/PS\text{-}8}$, $Nf1^{IF\text{-}54/PS\text{-}11}$ lines. Kaplan Meier plots of A) Tumor onset, B) disease-specific survival, and C) overall survival are shown. D) Overall survival analysis of em2 and em3 lines harboring the $Nf1^{IF\text{-}57/+}$ indel showed a discrepancy in survival. All animals in this study died of mammary tumor progression.

bioRxiv preprint doi: https://doi.org/10.1101/272815; this version posted February 27, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

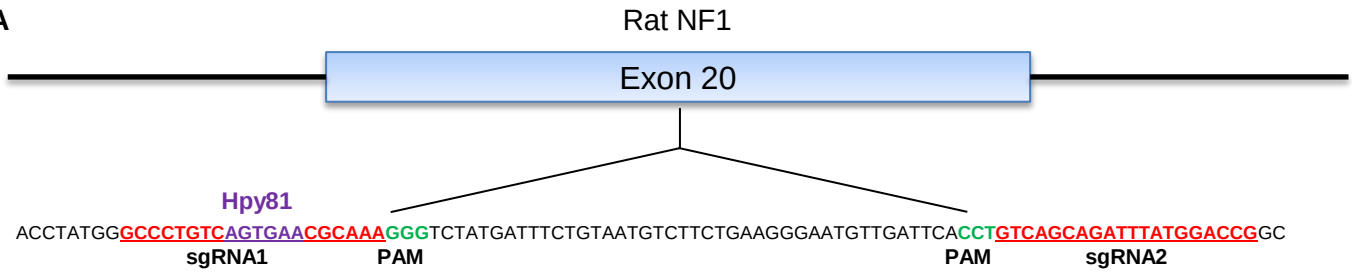
**Figure 4: Analysis of *Nf1*$^{IF}$ and *Nf1*$^{PS}$ mRNA and protein isoforms.** A) Schematic of RT-PCR analysis of *Nf1* exons 17-23. B) RT-PCR products were digested with Hpy81 and separated by PAGE to differentiate *Nf1*$^{IF}$ and *Nf1*$^{PS}$ mRNA. Unique bands were observed in the em2 and em3 *Nf1*$^{IF\text{-}57}$ REF cDNA (noted by *). C) Sequence analysis of em2-*Nf1*$^{IF\text{-}57}$ and em3-*Nf1*$^{IF\text{-}57}$ REF cDNA. Immunoblot analysis showing D) tissue-specific expression of rat neurofibromin isoforms and E) differential expression of the neurofibromin 250 and 125 kDa bands in matched brain, normal mammary tissue, and mammary tumor at both a long (top) and short (bottom) exposure.

**Figure 5: *NF1* shallow deletions are frequently present in sporadic breast cancers and associate with ER and FOX1A.** A) Truncating mutations were detected throughout *NF1* in METABRIC patients (cBioportal) B) Survival analysis of breast cancer patient survival comparing *NF1* diploid and shallow deletion copy number status. C) Network visualization of WGCNA module genes associated with *NF1* shallow deletion copy number status. D) Heatmap of METABRIC gene expression data of WGCNA module genes correlated with *NF1* copy number and ER status. D) Heatmap of METABRIC gene expression data of WGCNA module genes associated with RAS activation signatures.

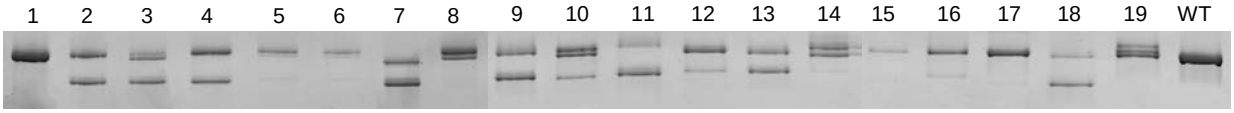
**Figure 6: Estrogen depletion via ovariectomy results in rapid tumor regression in *Nf1*-deficient tumors.** A) Tumor volume was plotted after ovariectomies (Day 0) for individual *Nf1*$^{IF}$ and *Nf1*$^{PS}$ rats. B) Percent tumor change over time of multiple tumors in individual *Nf1*$^{IF}$ and *Nf1*$^{PS}$ rats after ovariectomy (Day 0). C) The percent reduction in tumor size for all tumors from 3 Nf1 rat lines is shown. Distinct colors are used to represent individual rats and the mammary pad location of each tumor is notated. Note that several rats had multiple measurable tumors within an individual mammary pad.

# Figure 1

**A**



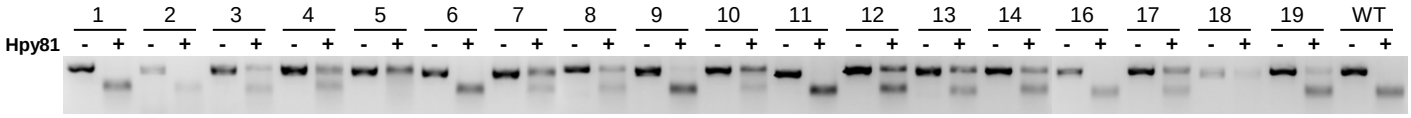

**B**




**C**

| | |
|---|---|
| WT | Y S P P M G P V S E R K G S M I S V M S S E G N V D S P V S R F M D R L L S L M V<br>TACAGCCCACCTATGGGCCCTGTCAGTGAACGCAAAGGGTCTATGATTTCTGTAATGTCTTCTGAAGGGAATGTTGATTCACCTGTCAGCAGATTTATGGACCGGCTTCTGTCCTTGATGGTG |
| rNF1#2 and #3<br>c.2669_2725del<br>19 aa deletion | Y S P P M G P V S E R R F M D R L L S L M V<br>TACAGCCCACCTATGGGCCCTGTCAGTGAACGCA----------GATTTATGGACCGGCTTCTGTCCTTGATGGTG |
| rNF1#10<br>c.2662_2724del<br>21 aa deletion | Y S P P M G P V S R F M D R L L S L M V<br>TACAGCCCACCTATGGGCCCTGTCAGT----------AGATTTATGGACCGGCTTCTGTCCTTGATGGTG |
| rNF1#13<br>c.2669_2722del<br>18 aa deletion | Y S P P M G P V S E R S R F M D R L L S L M V<br>TACAGCCCACCTATGGGCCCTGTCAGTGAACGCA----------GCAGATTTATGGACCGGCTTCTGTCCTTGATGGTG |

**D**




**E**

| | |
|---|---|
| WT | Y S P P M G P V S E R K G S M I S V M S S E G N V D S P V S R F M D R L L S L M V<br>TACAGCCCACCTATGGGCCCTGTCAGTGAACGCAAAGGGTCTATGATTTCTGTAATGTCTTCTGAAGGGAATGTTGATTCACCTGTCAGCAGATTTATGGACCGGCTTCTGTCCTTGATGGTG |
| rNF1#3<br>+1,-8;+4,-10 | Y S P P M G P V I K G L STOP<br>TACAGCCCACCTATGGGCCCTGTCAT-------AAAGGGTCTATGATTTCTGTAATGTCTTCTGAAGGGAATGTTGATTCATGTT----------GATTTATGGACCGGCTTCTGTCCTTGATGGTG |
| rNF1#10<br>-11 | Y S P P M G P V R V Y D F C N V F STOP<br>TACAGCCCACCTATGGGCCCTGTCAG----------GGTCTATGATTTCTGTAATGTCTTCTGAAGGGAATGTTGATTCACCTGTCAGCAGATTTATGGACCGGCTTCTGTCCTTGATGGTG |
| rNF1#13<br>-11;-4 | Y S P P M G P V R V Y D F C N V F STOP<br>TACAGCCCACCTATGGGCCCTGTCAG----------GGTCTATGATTTCTGTAATGTCTTCTGAAGGGAATGTTGATTCACCTGT----AGATTTATGGACCGGCTTCTGTCCTTGATGGTG |

bioRxiv preprint doi: https://doi.org/10.1101/272815; this version posted February 27, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

**Figure 2**

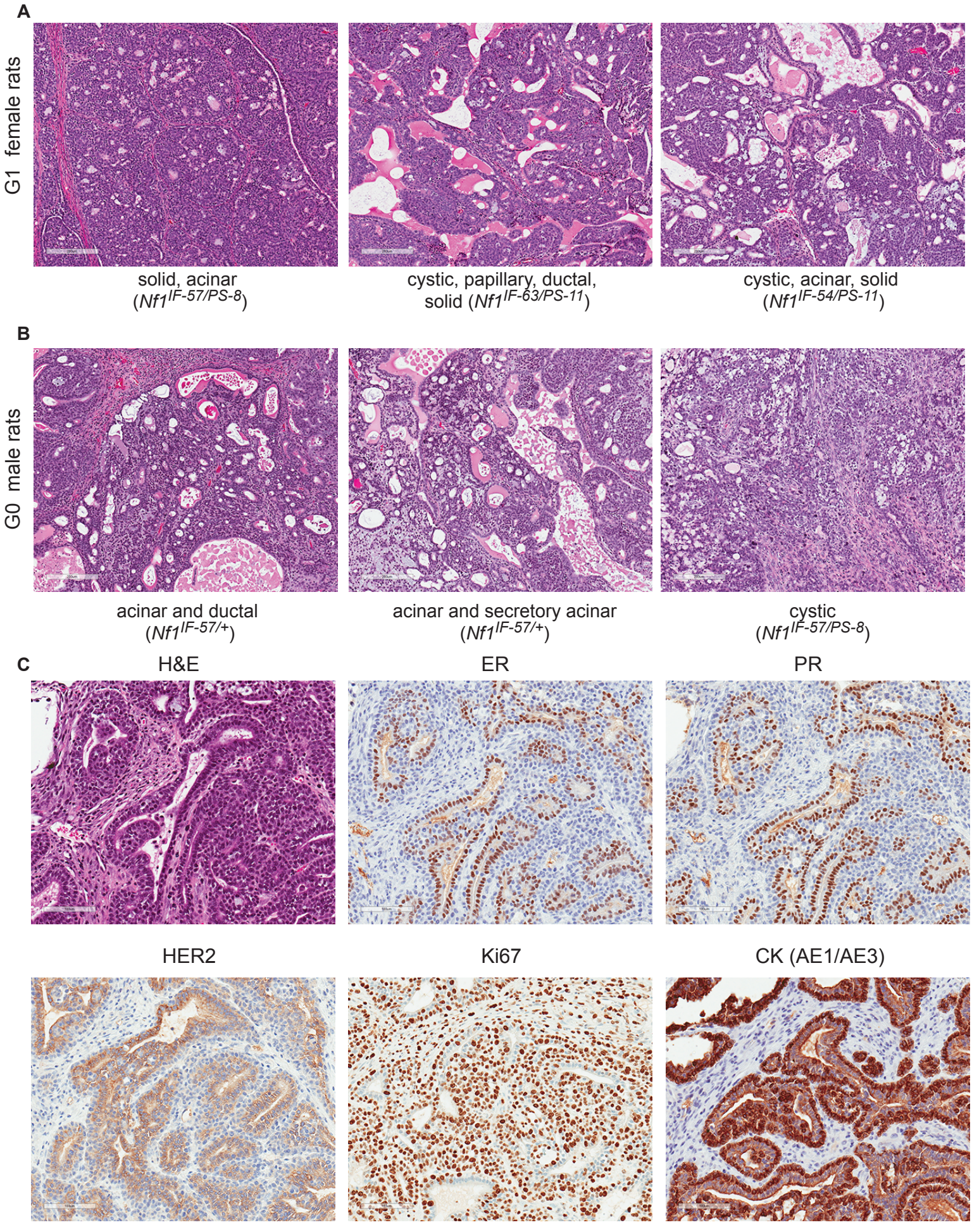
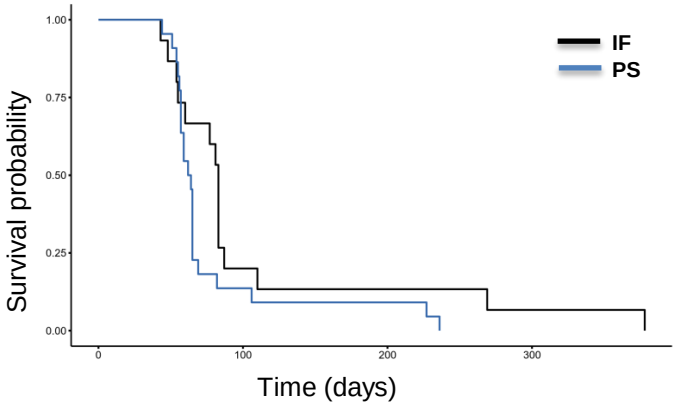
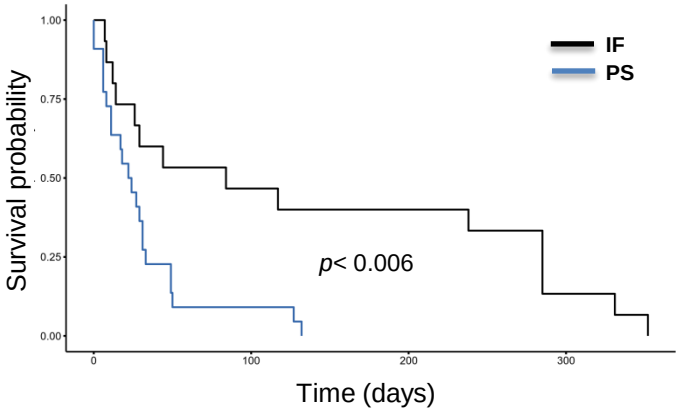


**Figure 3**

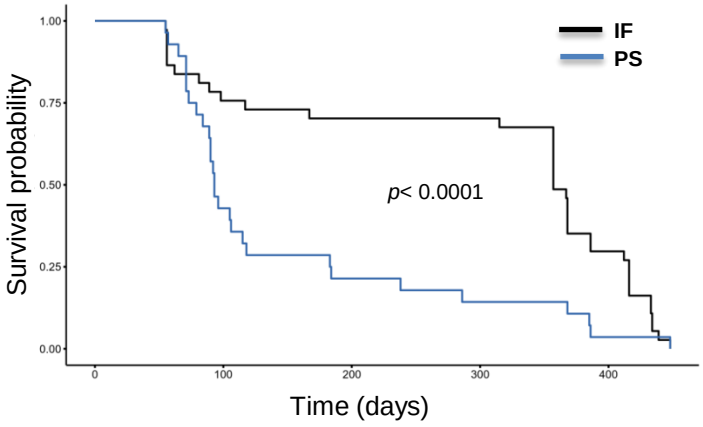
**A Tumor Onset**



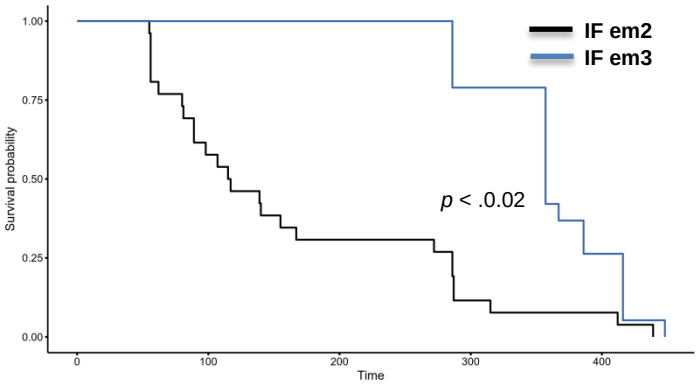

**B Disease-specific survival**




**C Overall Survival**

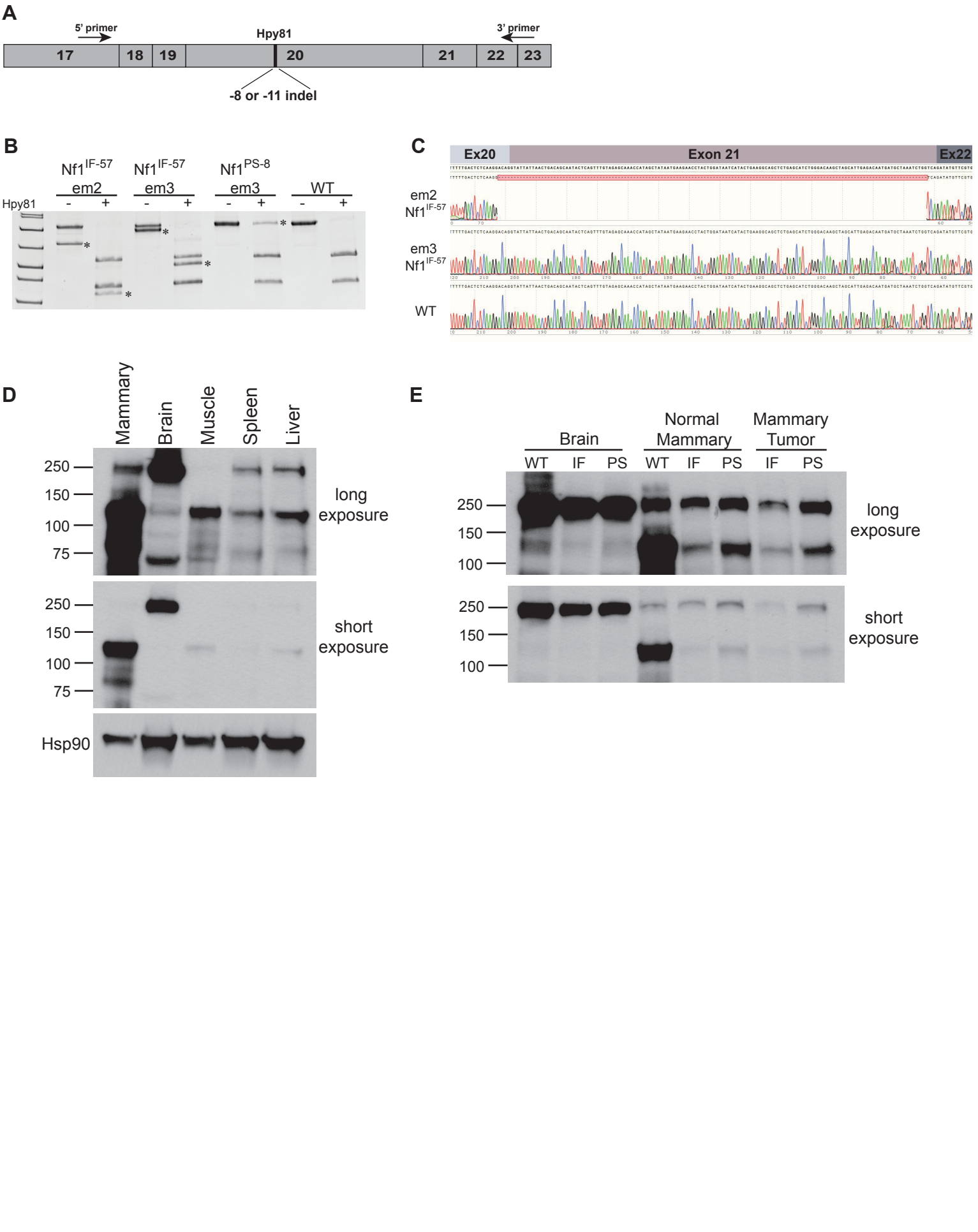



**D Overall Survival *NF1*$^{IF\text{-}57}$ em2 vs em 3**



bioRxiv preprint doi: https://doi.org/10.1101/272815; this version posted February 27, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

# Figure 4



# Figure 5

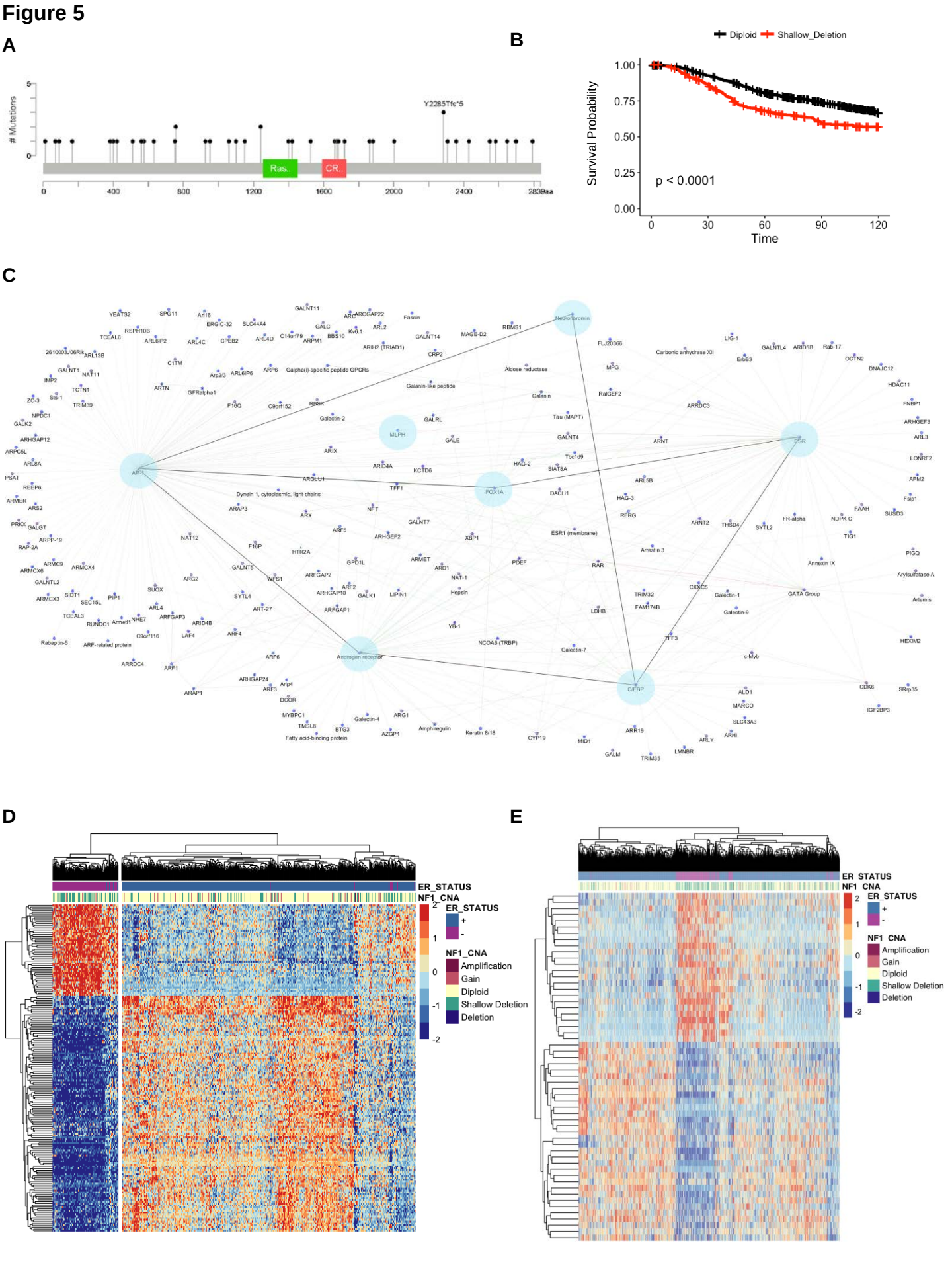
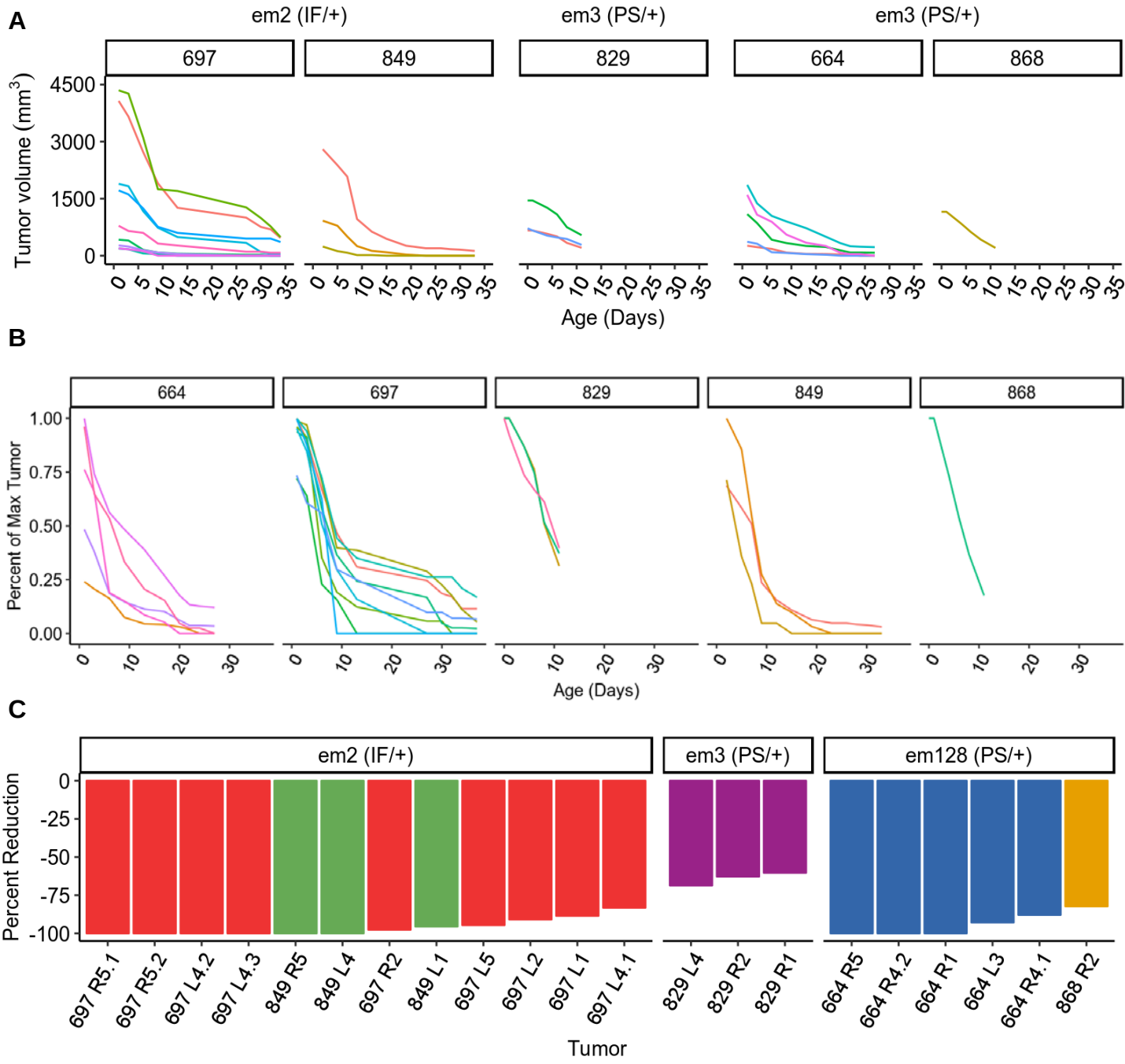
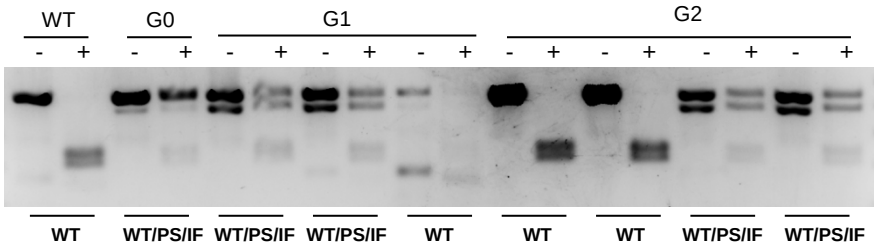

A
# Mutations
Y2285Tfs*5
Ras
CR
0
400
800
1200
1600
2000
2400
2839aa
B
Diploid
Shallow_Deletion
Survival Probability
1.00
0.75
0.50
0.25
0.00
p < 0.0001
0
30
60
90
120
Time
C
Neurofibromin
MLPH
ESR
AP-1
FOXA1
Androgen receptor
C/EBP
D
ER_STATUS
NF1_CNA
ER_STATUS
+
-
NF1_CNA
Amplification
Gain
Diploid
Shallow Deletion
Deletion
2
1
0
-1
-2
E
ER_STATUS
NF1_CNA
ER_STATUS
+
-
NF1_CNA
Amplification
Gain
Diploid
Shallow Deletion
Deletion
2
1
0
-1
-2

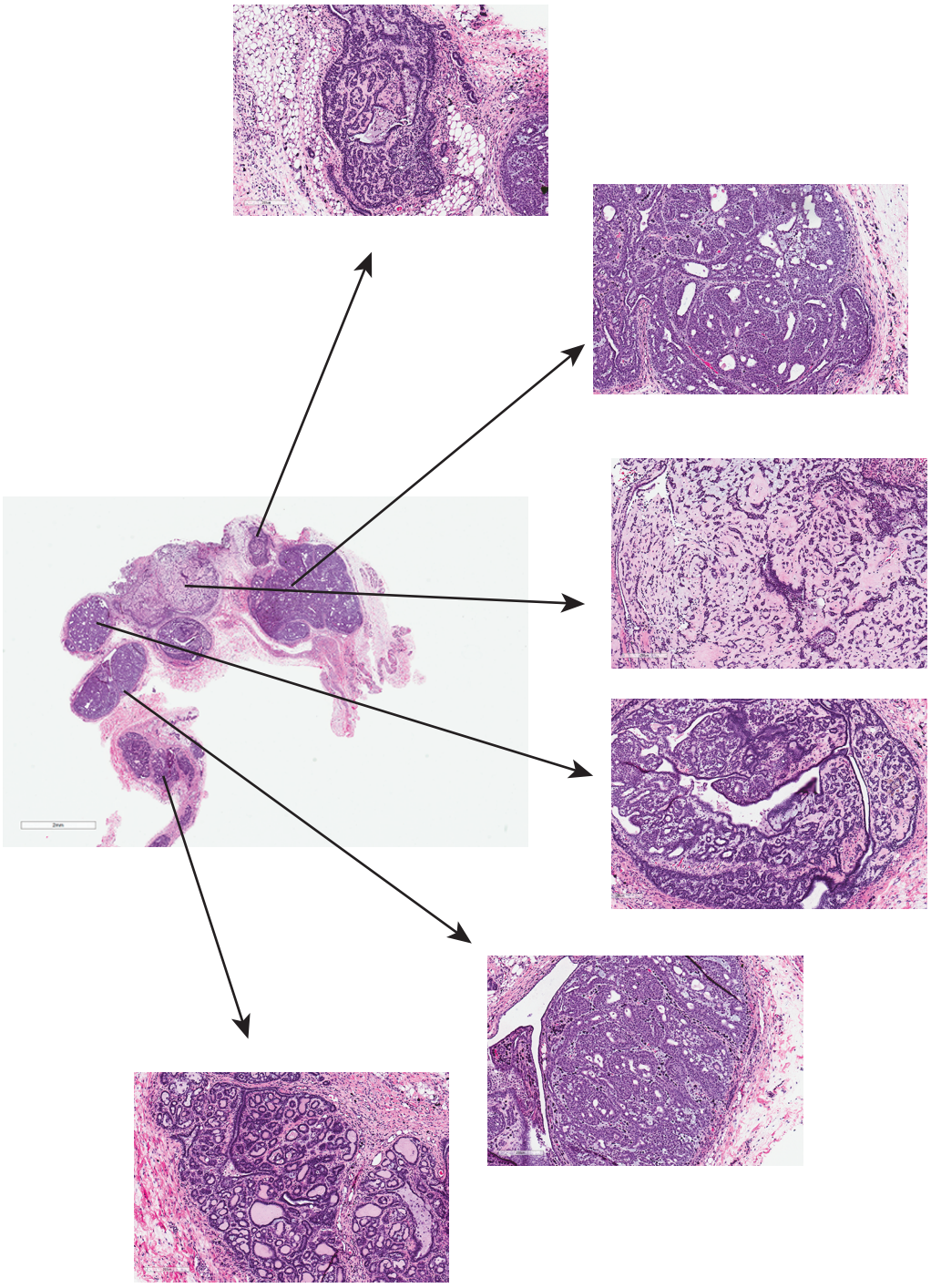
**Figure 6**


A
em2 (IF/+)
em3 (PS/+)
em3 (PS/+)
697
849
829
664
868
Tumor volume (mm3)
4500
3000
1500
0
0 5 10 15 20 25 30 35
Age (Days)
B
664
697
829
849
868
Percent of Max Tumor
1.00
0.75
0.50
0.25
0.00
0 10 20 30
Age (Days)
C
em2 (IF/+)
em3 (PS/+)
em128 (PS/+)
Percent Reduction
0
-25
-50
-75
-100
697 R5.1
697 R5.2
697 L4.2
697 L4.3
849 R5
849 L4
697 R2
849 L1
697 L5
697 L2
697 L1
697 L4.1
829 L4
829 R2
829 R1
664 R5
664 R4.2
664 R1
664 L3
664 R4.1
868 R2
Tumor

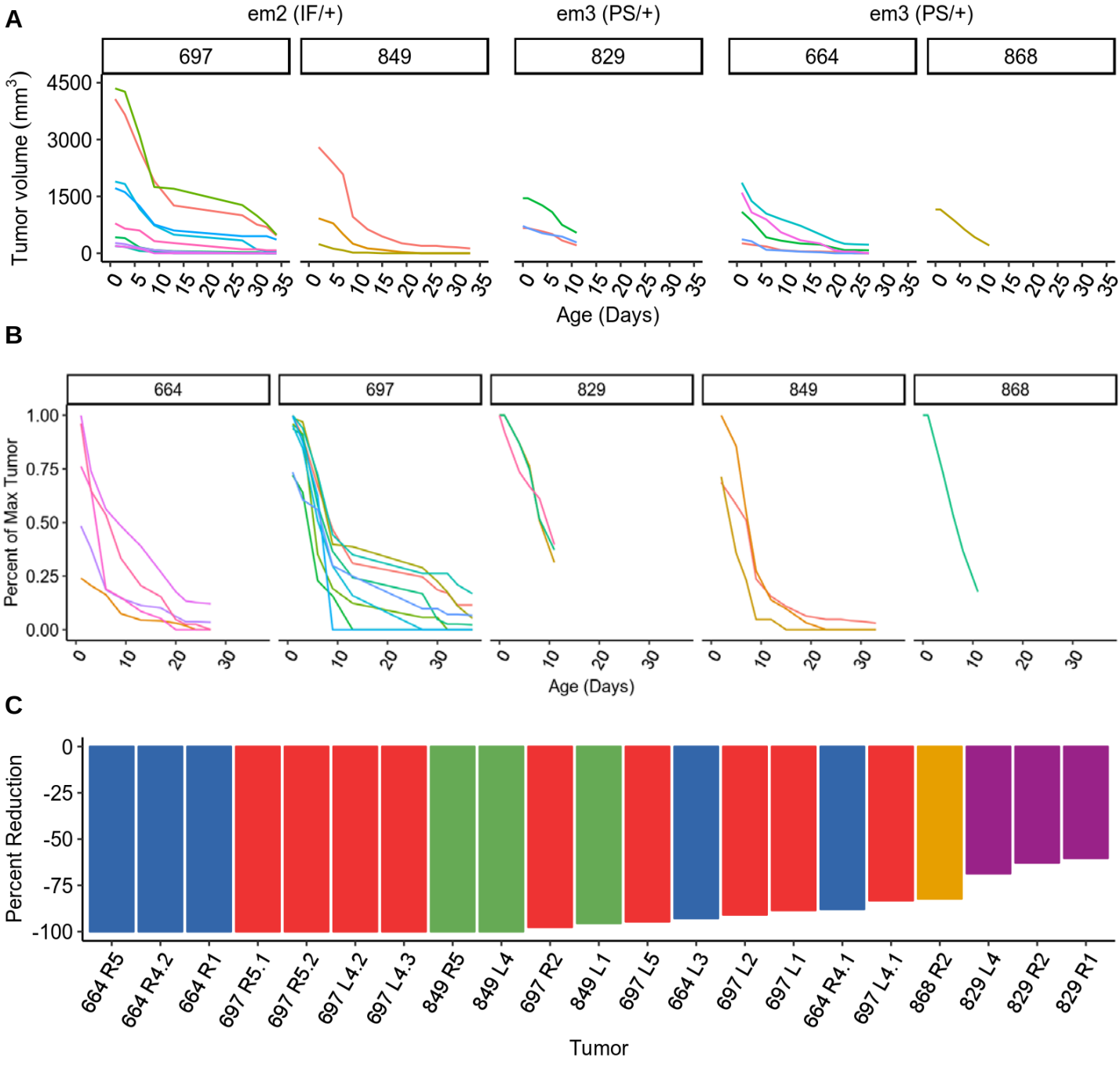



**Supplementary Figure 1: Three Nf1 alleles identified in *Nf1*$^{IF\text{-}63/PS\text{-}11}$ line.** HMA and sequencing identified 3 *NF1* alleles in *Nf1*$^{IF\text{-}63/PS\text{-}11}$ : a WT allele, an allele with an 11 bp deletion (premature stop), and an allele with a -63 bp deletion. The presence of more than two alleles was confirmed using unique primer sets in two separate labs. The three alleles were transmitted through the *Nf1*$^{IF\text{-}63/PS\text{-}11}$ germline in F1 and F2 generations and did not segregate in either generation as shown in the HMA profiles of individual F1 and F2 *Nf1*$^{IF\text{-}63/PS\text{-}11}$ rats.

bioRxiv preprint doi: https://doi.org/10.1101/272815; this version posted February 27, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.



**Supplementary Figure 2: Multiple tumors are commonly observed in *Nf1* mammary glands.** The level of tumor burden in the mammary pads of several animals was substantial. Shown above is the 4th mammary pad from *rNf1* #413.




**Supplementary Figure 3: Estrogen depletion via ovariectomy results in rapid tumor regression in *Nf1*-deficient tumors.** A) Tumor growth curves after overiectomies were performed (Day 0). B) Percent tumor loss of the largest tumor from each *Nf1* rat. C) The percent reduction in tumor size for all tumors from 3 *Nf1* rat lines is shown. Distinct colors are used to represent individual rats and the mammary pad location of each tumor is notated. Note that several rats had multiple measurable tumors within an individual mammary pad.

bioRxiv preprint doi: https://doi.org/10.1101/272815; this version posted February 27, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

| # | Sex | 5' CRISPR | 3' CRISPR | Tumor Onset (days) | Age of Death (days) | F1 | DNA Indels | Line Name |
|---|---|---|---|---|---|---|---|---|
| 1 | M | wt | wt | | Alive | NA | | |
| 2 | M | -57 (wt) | | NA | 477 | Yes | c.2669_2725del | ***Nf1*$^{IF-57/+}$** |
| 3 | M | +1,-8 bp<br>-57 | +4,-10 | NA | 416 | Yes | c.2659_2660insT,2660_2667del,2715_2716insTGTT,2716-2725del<br>c.2669_2725del | ***Nf1*$^{IF-57/PS-8}$** |
| 4 | F | -61 (wt) | | 53 | 58 | NA | | |
| 5 | F | IF indel (HMA) | | 53 | 63 | NA | | |
| 6 | F | -54 | | 60 | 77 | * | | |
| 7 | M | +1,-4<br>-66 | -20 | | Alive | None | | |
| 8 | F | -11 (wt) | -4 | | 62# | NA | | |
| 9 | F | +8,-1 | -11 | 29 | 49 | NA | | |
| 10 | F | -11<br>-63 (wt) | | 51 | 64 | Yes | c.2661_2671del<br>c.2662_2724del | ***Nf1*$^{IF-63/PS-11}$** |
| 11 | M | -57 | | | 314 | None | | |
| 12 | F | -6<br>-54 | -10 | 47 | 62 | * | | |
| 13 | F | -11<br>-54 | -4 | 54 | 198 | Yes | c.2661_2671del,2721_2724del<br>c.2669_2722del | ***Nf1*$^{IF-54/PS-11}$** |
| 14 | F | -17<br>-54 | | 47 | 68 | NA | | |
| 15 | M | -66 (wt) | | | Alive | NA | | |
| 16 | M | IF indel (HMA) | | | Alive | NA | | |
| 17 | M | -2 (wt)<br>-66 | | | Alive | NA | | |
| 18 | F | -66<br>-11 | 3sub,2sub | 28 | 54 | NA | | |
| 19 | F | | +1 (wt) | 89 | 189 | NA | | |

**Supplementary Table 1: Summary of $G_0$ animals obtained from CRISPR-Cas9 injections.** Summary of indels resulting from CRISPR-Cas9 injections including tumor onset and survival. The most common base pair alterations at the 5' and 3' CRISPR sites are listed (additional indels that were identified in each $G_0$ animal are listed in Supplemental Table 2). Mutant alleles that resulted in large deletions between 5' and 3' CRISPR target regions are listed in the middle of the 5' and 3' CRISPR columns; "wt" refers to the wildtype allele being observed by sequencing. IF indel (HMA) indicates that indels were identified by HMA but not sequence verified. All females were euthanized due to tumor burden except for #8 that was taken for control tissue (notated with #). DNA indel notations for non-founder animals can be found in Supplementary Table 2. * Euthanized when pregnant due to tumor burden. None = bred but no progeny.

bioRxiv preprint doi: https://doi.org/10.1101/272815; this version posted February 27, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

**Supplementary Table 2**

| ID | Sequence (5'-3') 5' CRISPR > 3' CRISPR < | Indels (Clones) | Protein | DNA |
|---|---|---|---|---|
| WT | TACAGCCCACCTATGGGCCCTGTCAGTGAACGCAAAGGGTCTATGATTTCTGTAATGTCTTCTGAAGGGAATGTTGATTCACCTGTCAGCAGATTTATGGACCGGCTTCTGTCCTTGATGGTG | | | |
| | Y S P P M G P V S E R K G S M I S V M S S E G N V D S P V S R F M D R L L S L M V | | | |
| 1 | TACAGCCCACCTATGGGCCCTGTCAGTGAACGCAAAGGGTCTATGATTTCTGTAATGTCTTCTGAAGGGAATGTTGATTCACCTGTCAGCAGATTTATGGACCGGCTTCTGTCCTTGATGGTG | WT (4) | | |
| | Y S P P M G P V S E R K G S M I S V M S S E G N V D S P V S R F M D R L L S L M V | | | |
| 2 | TACAGCCCACCTATGGGCCCTGTCAGTGAACGCA----------GATTTATGGACCGGCTTCTGTCCTTGATGGTG | WT(4);-57(4) | In frame deletion | c.2669 2725del |
| | Y S P P M G P V S E R R F M D R L L S L M V C | 19 aa deletion | | |
| 3 | TACAGCCCACCTATGGGCCCTGTCAT-------AAAGGGTCTATGATTTCTGTAATGTCTTCTGAAGGGAATGTTGATTCATGTT---------GATTTATGGACCGGCTTCTGTCCTTGATG GTG | WT(0);+1,-8,+4,-10(3); | Premature stop | c.2659_2660insT,2660_2667del,2715_2716insTGTT,2716-2725del |
| | Y S P P M G P V I K G L * | 4 aa* + stop | | |
| | TACAGCCCACCTATGGGCCCTGTCAGTGAACGCA----------GATTTATGGACCGGCTTCTGTCCTTGATGGTG | -57(1) | In frame deletion | c.2669_2725del |
| | Y S P P M G P V S E R R F M D R L L S L M V C | 19 aa deletion | | |
| 4 | TACAGCCCACCTATGGGCCCTGTCAGTGAA----------GATTTATGGACCGGCTTCTGTCCTTGATGGTG | WT(2);-61(4) | Premature stop | c.2665_2725del |
| | Y S P P M G P V S E D L W T G F C P * | 8 aa* + stop | | |
| 6 | TACAGCCCACCTATGGGCCCTGTCAGTGAACGCA----------GCAGATTTATGGACCGGCTTCTGTCCTTGATGGTG | WT(0);-54(2) | In frame deletion | c.2669_2722del |
| | Y S P P M G P V S E R S R F M D R L L S L M V | 18 aa deletion | | |
| 7 | TACAGCCCACCTATGGGCCCTGTCAGTGAACT----AGGGTCTATGATTTCTGTAATGTCTTCTGAAGGGAATGTTGATTCACC----------GGCTTCTGTCCTTGATGGTG | WT(0);+1,-4,-20(4) | Premature stop | c.2665_2666insT,2666_2669del |
| | Y S P P M G P V S E L G S M I S V M S S E G N V D S P A S V L D G V * | 23 aa* + stop | | |
| | TACAGCCCACCTATGGGCCCTGTCAG----------ATTTATGGACCGGCTTCTGTCCTTGATGGTG | -66(3) | deletion | c.2661 2726del |
| | Y S P P M G P V R F M D R L L S L M V C N H E | 22 aa deletion | | |
| | TACAGCCCACCTATGGACCCTGTCAG----------ATTTATGGACCGGCTTCTGTCCTTGATGGTG | 1sub,-66(1) | Missense/In frame deletion | c.2651G>A,2661_2726del |
| | Y S P P M D P V R F M D R L L S L M V C N H E | 1 aa* + 22 aa deletion | | |
| | TACAGCCCACCTATGGGCCCTGTCAG----------GGTCTATGATTTCTGTAATGTCTTCTGAAGGGAATGTTGATTCACC----------GGCTTCTGTCCTTGATGGTG | -11,-20(1) | Premature stop | c.2661_2671del,2718 2737del |
| | Y S P P M G P V R V Y D F C N V F * | 9 aa* + stop | | |
| 8 | TACAGCCCACCTATGGGCCCTGTCAG----------GGTCTATGATTTCTGTAATGTCTTCTGAAGGGAATGTTGATTCACCTG----CAGATTTATGGACCGGCTTCTGTTCTTGATGGTG | WT(3) & -11,-4,1sub (1) | Premature stop | c.2661_2671del,2720_2723del,2747C>T |
| | Y S P P M G P V R V Y D F C N V F * | 9 aa* + stop | | |
| | TACAGCCCACCTATGGGCCCTGTCAGTGAACGCAAAGGGTTTATGATTTCTGTAATGTCTTCTGAAGGGAATGTTGATTCACCTGTCAGCAGATTTATGGACCGGCTTCTGTCCTTGATGGTG | 1sub(1) | Missense | c.2675C>T |
| | Y S P P M G P V S E R K G F M I S V M S S E G | 1 aa* | | |
| | TACAGCCCACCTATGGGCCCTGTCAG----------GTCTATGATTTCTGTAATGTCTTCTGAAGGGAATGTTGATTCACCTGTCAGCAGATTTATGGACCGGCTTCTGTCCTTGATGGTG | -12(1) | deletion | c.2661_2672del |
| | Y S P P M G P V R S M I S V M S S E G N V D S | 4 aa deletion | | |
| | TACAGCCCACCTATGGGCCCTGTCAG----------GGTCTATGATTTCTGTAATGTCTTCTGAAGGGAATGTTGATTCACCTG----CAGATTTATGGACCGGCTTCTGTCCTTGATGGTG | -11,-4(1) | Premature stop | c.2661_2671del,2720_2723del |
| | Y S P P M G P V R V Y D F C N V F * | 9 aa* + stop | | |
| | TACAGCCCACCTATGGGCCCTATCAG----------GGTCTATGATTTCTGTAATGTCTTCTGAAGGGAATGTTGATTCACCTG----CAGATTTATGGACCGGCTTCTGTCCTTGATGGTG | 1sub,-11,-4(1) | Premature stop | c.2656G>A,2661_2671del2720_2723del |
| | Y S P P M G P I R V Y D F C N V F * | 10 aa* + stop | | |
| 9 | TACAGCCCACCTATGGGCCCTGTCAGTGAACGGTCTATGT-AAAGGGTCTATGATTTCTGTAATGTCTTCTGAAGGGAATGTTGATTCACCT----------TATGGACCGGCTTCTGTCCT TGATGGTG | WT(0);+8,-1,-11(2) | Premature stop | c.2666_2667insGTCTATGT,2667del,2719_2729del |
| | Y S P P M G P V S E R S M * | 2 aa* + stop | | |
| 10 | TACAGCCCACCTATGGGCCCTGTCAG----------GGTCTATGATTTCTGTAATGTCTTCTGAAGGGAATGTTGATTCACCTGTCAGCAGATTTATGGACCGGCTTCTGTCCTTGATGGTG | WT(1);-11(1) | Premature stop | c.2661_2671del |
| | Y S P P M G P V R V Y D F C N V F * | 9 aa* + stop | | |
| | TACAGCCCACCTATGGGCCCTGTCAGT----------AGATTTATGGACCGGCTTCTGTCCTTGATGGTG | -63(2) | deletion | c.2662 2724del |
| | Y S P P M G P V S R F M D R L L S L M V C N H | 21 aa deletion | | |
| 11 | TACAGCCCACCTATGGGCCCTGTCAGTGAACGCA----------GATTTATGGACCGGCTTCTGTCCTTGATGGTG | WT(0);-57(4) | deletion | c.2669_2725del |
| | Y S P P M G P V S E R R F M D R L L S L M V C | 19 aa deletion | | |
| 12 | TACAGCCCACCTATGGGCCCTGTCAGTGAACGCA----------GCAGATTTATGGACCGGCTTCTGTCCTTGATGGTG | WT(0);-54 (1) | deletion | c.2669_2722del |
| | Y S P P M G P V S E R S R F M D R L L S L M V | 18 aa deletion | | |
| | TACAGCCCACCTATGGGCCCTGTCAGTGAA------GGGTCTATGATTTCTGTAATGTCTTCTGAAGGGAATGTTGATTCACCTGTCAGCAGATTTATGGACCGGCTTCTGTCCTTGATGGTG | -6(1) | deletion | c.2665 2670del |
| | Y S P P M G P V S E G S M I S V M S S E G N V | 2 aa deletion | | |
| | TACAGCCCACCTATGGGCCCTGTCAGTGAACGCCAAAGGGTCTATGATTTCTGTAATGTCTTCTGAAGGGAATGTTGATTCA----------GATTTATGGACCGGCTTCTGTCCTTGATGGTG | +1,-10(2) | Premature stop | c.2667_2688insC,2726_2725del |
| | Y S P P M G P V S E R Q R V Y D F C N V F * | 10 aa* + stop | | |
| 13 | TACAGCCCACCTATGGGCCCTGTCAGTGAACGCA----------GCAGATTTATGGACCGGCTTCTGTCCTTGATGGTG | WT(0);-54(1) | deletion | c.2669_2722del |
| | Y S P P M G P V S E R S R F M D R L L S L M V | 18 aa deletion | | |
| | TACAGCCCACCTATGGGCCCTGTCAG----------GGTCTATGATTTCTGTAATGTCTTCTGAAGGGAATGTTGATTCACCTGT----AGATTTATGGACCGGCTTCTGTCCTTGATGGTG | -11,-4(4) | Premature stop | c.2661_2671del,2721_2724del |
| | Y S P P M G P V R V Y D F C N V F * | 9 aa* + stop | | |
| 14 | TACAGCCCACCTATGG-CCCTGTCAGTGA----------TTTCTGTAATGTCTTCTGAAGGGAATGTTGATTCACCTGTCAGCAGATTTATGGACCGGCTTCTGTCCTTGATGGTG | WT(0);-1,-17(1) | Premature stop | c.2651del,2664_2680del |
| | Y S P P M A L S V I S V M S S E G N D S P V S R F M D R L L S L M V C N H E K V G | > 33 aa* + stop? | | |
| | TACAGCCCACCTATGGGCCCTGTCAGTGA----------TTTCTGTAATGTCTTCTGAAGGGAATGTTGATTCACCTGTCAGCAGATTTATGGACCGGCTTCTGTCCTTGATGGTG | -17(1) | Premature stop | c.2664_2680del |
| | Y S P P M G P V S D F C N V F * | 7 aa* + stop | | |
| | TACAGCCCACCTATGGGCCCTGTCAGTGAACGCA----------GCAGATTTATGGACCGGCTTCTGTCCTTGATGGTG | -54(1) | deletion | c.2669 2722del |
| | Y S P P M G P V S E R S R F M D R L L S L M V | 18 aa deletion | | |
| 15 | TACAGCCCACCTATGGGCCCTGTCAG----------ATTTATGGACCGGCTTCTGTCCTTGATGGTG | WT(3);-66(2) | deletion | c.2661_2726del |
| | Y S P P M G P V R F M D R L L S L M V C N H E | 22 aa deletion | | |
| 16 | TACAGCCCACCTATGGGCCCTGTCAGTGAACGCAAAGGGTCTATGATTTCTGTAATGTCTTCTGAAGGGAATGTTGATTCACCTGTCAGCAGATTTATGGACCGGCTTCTGTCCTTGATGGTG | WT(3) | | |
| | Y S P P M G P V S E R K G S M I S V M S S E G N V D S P V S R F M D R L L S L M V | | | |
| 17 | TACAGCCCACCTATGGGCCCTGTCAGTGAAC--AAAGGGTCTATGATTTCTGTAATGTCTTCTGAAGGGAATGTTGATTCACCTGTCAGCAGATTTATGGACCGGCTTCTGTCCTTGATGGTG | WT(1);-2 (1) | Premature stop | c.2666_2667del |
| | Y S P P M G P V S E Q R V Y D F C N V F * | 10 aa* + stop | | |
| | TACAGCCCACCTATGGGCCCTGTCAG----------ATTTATGGACCGGCTTCTGTCCTTGATGGTG | -66 (1) | deletion | c.2661 2726del |
| | Y S P P M G P V R F M D R L L S L M V C N H E | 22 aa deletion | | |
| | TACGGCCCACCTATGGGCCCTGTCAGTGAA---AAAGGGTCTATGATTTCTGTAATGTCTTCTGAAGGGAATGTTGATTCACCTGTCAGCAGATTTATGGACCGGCTTCTGTCCTTGATGGTG | 1sub,-3(1) | Premature stop? | c.2638A>G,2665_2667del |
| | Y G P P M G P V S E K G S M I S V M S S E G N | < 33 aa* + stop? | | |
| | TACAGCCCACCTATGGGCCCTGTCAGTGAA---AAAGGGTCTATGATTTCTGTAATGTCTTCTGAAGGGAATGTTGATTCACCTGTCAGCAGATTTATGGACCAGCTTCTGTCCTTGATGGTG | -3,1sub(1) | In frame deletion/missense | c.2665_2667del,2738G>A |
| | Y S P P M G P V S E K G S M I S V M S S E G N V D S P V S R F M D Q L L S L M V C | 3 aa deletion + 1 aa* | | |
| 18 | TACAGCCCACCTATGGGCCCTGTCAG----------ATTTATGGACCGGCTTCTGTCCTTGATGGTG | WT(0);-66 (3) | deletion | c.2661_2726del |
| | Y S P P M G P V R F M D R L L S L M V C N H E | 22 aa deletion | | |
| | TACAGCCCACCTATGGGCCCTGTCAG----------GGTCTATGATTTCTGTAATGTCTTCTGAAGGGAATGTTGATTCACCTGAATGTTGATTTATGGACCGGCTTCTGTCCTTGATGGTG | -11,3sub,2sub (1) | Premature stop | c.2661_2671del,2720T>A,2721C>A,2722A>T,2724C>T,2725A>T |
| | Y S P P M G P V R V Y D F C N V F * | 9 aa* + stop | | |
| 19 | TACAGCCCACCTATGGGCCCTGTCAGTGAACGCAAAGGGTCTATGATTTCTGTAATGTCTTCTGAAGGGAATGTTGATTCACCTGTCAAGCAGATTTATGGACCGGCTTCTGTCCTTGATGGTG | WT(1);+1(2) | Premature stop | c.2722_2723insA |
| | Y S P P M G P V S E R K G S M I S V M S S E G N V D S P V K Q I Y G P A S V L D G V * | 12 aa* + stop | | |

**Supplementary Table 2: Detailed sequence analysis of *Nf1* $G_0$ rats.** 34 mutant alleles, including 25 unique mutant alleles, were identified by Sanger sequencing. Sequences for all alleles identified in G0 animals are shown.

bioRxiv preprint doi: https://doi.org/10.1101/272815; this version posted February 27, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

**Supplementary Table 3: Histopathology of *rNf1* animals**

| Animal | Sex | Genotype | Age (days) | Tumor/Slide No | Tissue/Mammary Pad Description | Pathology |
|---|---|---|---|---|---|---|
| **G0/F0 Animals** | | | | | | |
| 2 | M | | 477 | 1 | R1 Mammary Fat Pad | mammary adenocarcinoma - acinar and mucinous |
| | | | | 2 | L4 Mammary Fat Pad | mammary pad - galactorrhea |
| | | | | 3 | R2 Mammary Fat Pad | mammary adenocarcinoma - acinar and ductular |
| | | | | 4 | L1 Mammary Fat Pad | two mammary adenocarcinomas 1-acinar and ductular 2-secretory acinar |
| | | | | 5 | L4 Mammary Fat Pad | mammary pad - galactorrhea |
| | | | | 6 | R1 Mammary Fat Pad | mammary pad - galactorrhea |
| 3 | M | | 416 | 1 | L1 Mammary Fat Pad | mammary adenoma - cystic |
| | | | | 2 | Right flank | fibrous histiocytic sarcoma |
| | | | | 3 | L1 Mammary Fat Pad | mammary adenocarcinoma with two small intraductal tumors |
| | | | | 4 | L4 Mammary Fat Pad | mammary adenoma - cystic |
| 4 | F | | 58 | 1 | R1 Mammary Fat Pad | mammary adenocarcinoma - acinar |
| | | | | 2 | L4 Mammary Fat Pad | mammary adenocarcinoma - acinar and solid |
| | | | | 3 | R4 Mammary Fat Pad | mammary adenocarcinoma - acinar and solid |
| | | | | 4 | L5 Mammary Fat Pad | mammary adenocarcinoma - acinar and solid |
| | | | | 5 | L4 Mammary Fat Pad | mammary adenocarcinoma - acinar and solid |
| 6 | F | | 77 | 1 | R2 Mammary Fat Pad | mammary adenocarcinoma - acinar |
| | | | | 2 | R5 Mammary Fat Pad | mammary adenocarcinoma - acinar |
| | | | | 3 | L2 Mammary Fat Pad | mammary adenocarcinoma - acinar |
| | | | | 4 | L4 Mammary Fat Pad | mammary hyperplasia, mammary adenocarcinoma - acinar and ductular |
| | | | | 5 | R5 Mammary Fat Pad | mammary hyperplasia; mammary adenocarcinoma - acinar, ductular and cystic |
| | | | | 6 | L5 Mammary Fat Pad | mild mammay hyperplasia; mammary adenocarcinoma - cystic and papillary |
| | | | | 8 | R4 Mammary Fat Pad | mammary adenocarcinoma - ductular |
| | | | | N1 | L4-5 Mammary Pad | moderately severe acinar hyperplasia |
| 10 | F | | 64 | 1 | R4 Mammary Fat Pad | mammary adenocarcinoma - acinar and solid |
| | | | | 2 | R5 Mammary Fat Pad | lactating mammary gland with small acinar adenocarcinoma |
| | | | | 3 | R5 Mammary Fat Pad | lactating mammary with adenocarcinoma - acinar and cystic |
| | | | | 4 | L4 Mammary Fat Pad | mammary adenocarcinoma - acinar and solid |
| | | | | 5 | L5 Mammary Fat Pad | mammary adenocarcinoma - acinar and solid |
| | | | | 6 | L5 Mammary Fat Pad | mammary adenocarcinoma - acinar |
| 13 | F | | 198 | 1 | R4 Mammary Fat Pad | mammary adenocarcinoma - acinar and solid |
| | | | | 2 | L5 Mammary Fat Pad | 2 mammary adenocarcinomas - acinar and ductular 2) acinar |
| | | | | 3 | L3 Mammary Fat Pad | mammary adenocarcinoma - acinar and ductular |
| | | | | 4 | Right Brachial Lymph Node | large lymph node with hyperplasia |
| | | | | 5 | Left Brachial Lymph Node | large lymph node with hyperplasia |
| | | | | 6 | Aortic Lymph Node | lymph node |
| | | | | 7 | L4 Mammary Fat Pad | lymph node |
| **rNF1-2 F1 animals** | | | | | | |
| **Folder** | **Sex** | **Genotype** | **Age (days)** | **Tumor/Slide No** | **Tissue/Mammary Pad Description** | **Pathology** |
| 303 | F | | 62 | 1 | R4 Mammary Fat Pad | mammary adenocarcinoma - acinar and solid |
| | | | | 2 | R3 Mammary Fat Pad | mammary adenocarcinoma - solid, ductular with stroma |
| | | | | 7 | L4 Mammary Fat Pad | slight acinar hyperplasia |
| | | | | N1 | L4 normal mammary tissue | mild acinar hyperplasia |
| 304 | F | | 89 | 1 | R4 Mammary Fat Pad | mammary adenocarcinoma - acinar and cystic |
| | | | | 2 | L5 Mammary Fat Pad | mammary adenocarcinoma - solid, acinar and cystic |
| | | | | 3 | L2 Mammary Fat Pad | mammary adenocarcinoma - solid and cystic |
| | | | | 4 | L1 Mammary Fat Pad | mammary adenocarcinoma - solid and cystic |
| 366 | F | | 98 | 1 | L4 Mammary Fat Pad - 1 | mammary adenocarcinoma - acinar and ductular |
| | | | | 2 | L4 Mammary Fat Pad - 2 | mammary adenocarcinoma - acinar and ductular |
| | | | | 3 | R2 Mammary Fat Pad | mammary adenocarcinoma - acinar and ductular |
| | | | | 4 | R4 Mammary Fat Pad -1 | mammary adenocarcinoma - acinar and ductular |
| | | | | 5 | R4 Mammary Fat Pad - 2 | mammary adenocarcinoma - cystic and ductular |
| | | | | 6 | L2 Mammary Fat Pad | mammary adenoma |
| **rNF1-3 F1 animals** | | | | | | |
| **Folder** | **Sex** | **Genotype** | **Age (days)** | **Tumor/Slide No** | **Tissue/Mammary Pad Description** | **Pathology** |
| 357 | F | | 84 | 1 | R5 Mammary Fat Pad | mammary adenocarcinoma - mostly solid with some invasion |
| | | | | 2 | R1 Mammary Fat Pad | mammary adenocarcinoma - acinar |
| | | | | 3 | L4 Mammary Fat Pad - 2 | mammary adenocarcinoma - mostly solid with hemorrhage |
| | | | | 4 | L4 Mammary Fat Pad - 1 | mammary adenocarcinoma - solid |
| | | | | 9 | L2 Mammary Fat Pad | mammary adenocarcinoma - acinar and solid |
| | | | | 10 | R4 Mammary Fat Pad | mammary adenocarcinoma - acinar and solid |
| | | | | N1 | R3/R4 Normal Mammary Tissue | small mammary intraductal solid adenocarcinoma |
| 413 | F | | 105 | 1 | L2 Mammary Fat Pad | mammary adenocarcinoma - acinar |
| | | | | 2 | L4 Mammary Fat Pad - 1 | mammary adenocarcinoma - acinar and ductular |
| | | | | 3 | L3 Mammary Fat Pad | mammary adenocarcinoma - acinar and solid |
| | | | | 4 | L4 Mammary Fat Pad - 2 | mammary adenocarcinoma - acinar and solid |
| | | | | 5 | L4 Mammary Fat Pad - 3 | 6 separate mammary tumors, 4 acinar, one adenoma and one with edema |
| | | | | 6 | L5 Mammary Fat Pad | solid, acinar, ductular and cystic |
| | | | | 7 | R2 Mammary Fat Pad | 5 small mammary acinar tumors in an abscess. |
| | | | | 8 | R5 Mammary Fat Pad | mammary adenocarcinoma - ductular, solid and acinar |
| | | | | 9 | R4 Mammary Fat Pad | mammary adenocarcioma - acinar, ductular, cystic |

bioRxiv preprint doi: https://doi.org/10.1101/272815; this version posted February 27, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

**rNF1-10 F1 animals**

| Folder | Sex | Genotype | Age (days) | Tumor/Slide No | Tissue/Mammary Pad Description | Pathology |
|---|---|---|---|---|---|---|
| 246 | F | | 96 | 1 | R2 Mammary Fat Pad | mammary adenocarcinoma - ductular and solid |
| | | | | 2 | L5 Mammary Fat Pad - 1 | mammary adenocarcinoma - cystic, papillary, solid, and ductal |
| | | | | 5 | L5 Mammary Fat Pad -2 | mammary adenocarcinoma - ductular, acinar and cystic |
| | | | | N1 | R4 Normal Mammary Tissue | mild acinar hyperplasia |
| 252 | F | | 183 | 1 | R4 Mammary Fat Pad | normal mammary gland |
| | | | | 2 | L5 Mammary Fat Pad | mammary adenocarcinoma - acinar |
| | | | | 3 | L4 Mammary Fat Pad | focal hyperplasia and mammary adenoma |
| | | | | 7 | L2 Mammary Fat Pad | mammary adenocarcinoma - solid |
| | | | | N1 | L2 Mammary Fat Pad | normal mammary |

**rNF1-13 F1 animals**

| Folder | Sex | Genotype | Age (days) | Tumor/Slide No | Tissue/Mammary Pad Description | Pathology |
|---|---|---|---|---|---|---|
| 262 | F | | 92 | 1 | R2 Mammary Fat Pad | mammary adenocarcinoma - solid, ductular with stroma |
| | | | | N1 | Normal mammary tissue | normal mammary tissue |
| 391 | F | | 89 | 1 | R4 Mammary Fat Pad - 1 | mammary adenocarcinoma - ductular |
| | | | | 2 | R4 Mammary Fat Pad - 2 | mammary adenocarcinoma - ductular and solid |
| | | | | 3 | R4 Mammary Fat Pad - 3 | mammary adenocarcinoma - ductular |
| | | | | 4 | L4 Mammary Fat Pad | 3 small mammary adenocarinomas - acinar |
| | | | | 5 | L2 Mammary Fat Pad | mammary adenocarcinoma - acinar and solid |
| | | | | N1 | R2 Normal Mammary | normal mammary gland |
| 397 | F | | 118 | 1 | L2 Mammary Fat Pad | mammary adenocarcinoma - solid, acinar and cystic |
| | | | | 2 | L4 Mammary Fat Pad - 2 | mammary adenocarcinoma - acinar and solid |
| | | | | 3 | L5 Mammary Fat Pad | mammary adenocarcinoma - solid |
| | | | | 4 | R5 Mammary Fat Pad | mammary adenocarcinoma - solid and acinar |
| | | | | 5 | L4 Mammary Fat Pad - 1 | mammary adenocarcinoma - ductular |
| | | | | N1 | R4 Normal Mammary | mildly hyperplastic mammary gland |