Supporting Information

Laurin et al. 10.1073/pnas.1213050110

SI Materials and Methods

Compendium of Microarray Breast Cancer Datasets. Gene expression datasets, gene annotations and clinical information have been downloaded from http://compbio.dfci.harvard.edu/pubs/sbtpaper/, companion website of the publication of Haibe-Kains et al. (1). To facilitate comparison between datasets, we applied a robust linear scaling to Dedicator of Cytokinesis 1 (*DOCK1*) expression measurements such that the quantiles 2.5% and 97.5% of the expression values were set to −1 and +1, respectively.

Gene Signatures Score. For a specific data set, the signature score was computed for each sample as

$$score = \sum s_i x_i / \sum |s_i|,$$

where x_i is the expression of a gene in the module that is present in the dataset platform, and s_i is either +1 or −1 depending on the sign of the fold change (over- or underexpression) computed during comparison between Mouse Mammary Tumor Virus (MMTV)-NeuNDL2-5-Internal Ribosome Entry Site (IRES)-Cre (NIC) *NIC*$^{+}$*DOCK1* loxP (flx)^{flx} and the control *NIC*$^{+}$*DOCK1*$^{wt/wt}$ tumors. Similarly to scaling of *DOCK1* expression levels, we applied a robust linear scaling such that the quantiles 2.5% and 97.5% of the signature scores were set to −1 and +1, respectively.

Survival Analysis. All survival analyses have been performed with R version 2.15. Distant metastasis-free survival whenever available and relapse-free survival otherwise have been considered for clinical outcome. All survival data were censored at 10-y hazard ratios (HRs), calculated using Cox regression (using the survival package version 2.36-14) with the dataset as stratum indicator, allowing for different baseline hazard functions between cohorts. Survival curves were estimated using the Kaplan–Meier estimator and compared using the two-sided log-rank test.

Cell Culture, Transfection, and Plasmids. HEK293T and the human ductal breast epithelial tumor cell line T47D were routinely cultured in DMEM supplemented with 10% (vol/vol) FBS and a mixture of penicillin and streptomycin (Gibco). Normal Murine Mammary Gland epithelial (NMuMG) cells were grown in DMEM supplemented with 10% (vol/vol) FBS, 2 mM L-glutamine, 20 mM Hepes pH 7.4, 10 μg/mL insulin and penicillin/streptomycin. HEK293T cells were transfected by the calcium phosphate method, whereas Lipofectamine 2000 was used for transfection of NMuMG and T47D. NMuMG cells stably expressing empty vector (EV) or an activated and oncogenic form of rat HER2/Neu (NeuNT) were generated by transfection of empty pcDNA3.1 or pcDNA3.1 rat NeuNT (Neo resistance; kind gift from P. Siegel, McGill University, Montreal). Both cell populations were selected in Geneticin (G418, Invitrogen) and injected in the mammary fat pads of nude mice (Charles River) as described (2). A cyst from the NMuMG-EV and tumors from NMuMG-NeuNT were harvested, dissociated and cells were explanted in culture in DMEM supplemented with 10% (vol/vol) FBS, 2 mM L-glutamine, 20 mM Hepes pH 7.4, 10 μg/mL insulin, 500 μg/mL G418, and penicillin/streptomycin. Down-regulation of *DOCK1* expression in NMuMG-NeuNT cells was achieved as follow: NMuMG-NeuNT explanted cells were independently transfected to express two different *DOCK1*-specific shRNAs (Open Biosystems, Oligo ID: V2MM_203272 for sh1 DOCK1 and V2MM_109441 for sh2 DOCK1 from the pSM2 retroviral plasmid shRNAmir library). A total of 1 μg/mL puromycin was added to the media and clones were isolated and tested for DOCK1 protein knockdown. T47D cells were transfected with 100 nM ON-target smart pool human DOCK1 siRNA (Dharmacon) to achieve transient *DOCK1* down-regulation or 100 nM NON-targeting siRNA (Dharmacon) as control. pCNX2 Flag-DOCK1 plasmid was obtained from M. Matsuda (Kyoto University, Kyoto). pcDNA3.1 Myc-ELMO1 plasmid was described previously (3). pGEX-4T-1-DOCK1$^{1228–1865}$ plasmid used for the production of the GST-DOCK$^{1228–1865}$ recombinant protein, was generated by amplifying DOCK1 by PCR (see Table S2 for primer details) from residue 1228–1865. After ligation of the PCR product into pGEM plasmid (Promega), the insert was digested and transferred into pGEX-4T-1 using BamHI and NotI enzymes.

Immunoprecipitation, Rac GTP Assays, and Western Blot Analyses. For total cell lysate preparation, minced tumor pieces and T47D cells were lysed in Radio ImmunoPrecipitation Assay (RIPA) buffer (50 mM Tris pH 7.5, 0.1% SDS, 0.5% deoxycholic acid, 1% Nonidet P-40, 150 mM NaCl, 5 mM EDTA). For immunoprecipitation with M2-Flag beads (Sigma), HEK293 cells were lysed in Nonidet P-40 buffer (150 mM NaCl, 50 mM Tris pH 7.5, 1% Nonidet P-40). For immunoprecipitation on endogenous DOCK1 with anti-DOCK1 (H4 beads; Santa Cruz Biotechnology), T47D cells were treated with or without Heregulin β1 (HRG) for 15 min, lysed in RIPA buffer, and immediately diluted with 50 mM Hepes pH 7.4, 150 mM NaCl, 0.1% Triton-X100, 10% glycerol (HNTG buffer). For Rac activation assay, T47D cells were serum starved in DMEM for 48 h, treated with DMSO or with 100 μM 4-[3′-(2″-chlorophenyl)-2′-propen-1′-ylidene]-1-phenyl-3,5-pyrazolidinedione (CPYPP) DOCK1 inhibitor (4) for 60 min before treatment with or without 20 ng/mL HRG for 15 min and lysed in the Rac/Cdc42 buffer (20 mM Tris pH 7.4, 150 mM NaCl, 5 mM MgCl2, 0.5% Nonidet P-40, 5 mM β-glycerophosphate, 1 mM DTT). The Rac GTP status was analyzed by precipitation with the purified p21 binding domain of PAK kinase expressed as a GST fusion (GST PAK-PBD) as previously described (3). Equal amounts of proteins lysates or pull downs were separated by SDS/PAGE and Rac was detected by immunoblotting.

GST-Protein Purification and Kinase Assays. GST and GST-DOCK1$^{1228-1865}$ recombinant proteins were prepared as previously described (3). For both hot and cold kinase assays, 2 μg of GST and GST-DOCK1$^{1228–1865}$ were mixed with either 400 ng Human EFG Receptor 2 (HER2) (31166; Active Motif) or 50 ng SRC (31195; Active Motif) recombinant kinases and 40 μL of Kinase Reaction Buffer (KRB) buffer (60 mM Hepes pH 7.5, 3 mM $MgCl_2$, 3 mM $MnCl_2$, 50 μL ATP, 3 μM $NaVO_4$, 1.2 mM DTT). For hot kinase assays, 10 μCi of ATP gamma ^{32}P was added (Perkin-Elmer). Reactions were incubated for 60 min at 30 °C. Reactions were stopped by adding SDS sample buffer and denaturing for 5 min at 95 °C and were loaded on gel. Radioactive products were visualized by autoradiography, whereas the non-radioactive products were detected with anti-pDOCK1^{Y1811}.

Cell Migration Assays. Cell migration assays were performed using 8-μm pores Boyden Chambers (Corning) coated with 25 ng/mL Collagen I (Roche) in 24-well plates as previously described (3). For chemical inhibition of DOCK1, serum starved T47D cells were pretreated with either DMSO or 100 μM CPYPP DOCK1 inhibitor for 60 min before being detached and washed with

DMEM 0.1% BSA. A total of 150,000 cells were seeded in DMEM/0.1% BSA-containing media supplemented with either DMSO or 100 μM CPYPP in the top chamber and cells were allowed to migrate for 5 h toward the bottom chamber containing DMEM/0.1% BSA supplemented or not with 20 ng/mL HRG. For RNAi mediated inhibition of DOCK1, T47D cells were either transfected with a NON-targeting siRNA for control or ON-target smart pool DOCK1 siRNA. The next day cells were serum starved for 24 h before being detached and washed with DMEM 0.1% BSA. A total of 150,000 cells were seeded in DMEM/0.1% BSA-containing media in the top chamber and cells were allowed to migrate for 5 h toward the bottom chamber containing DMEM/0.1% BSA supplemented or not with 20 ng/mL HRG. Cells were then washed twice with PBS and fixed with 4% paraformaldehyde. The top surface of the membrane was wiped with a cotton swab and the membrane was mounted on a microscope slide using SlowFade Gold antifade reagent mounting media (Invitrogen). The average number of migrating cells in 10 independent 20× microscope fields was evaluated and each experiment was performed in duplicate ($n = 7$).

Generation of the *DOCK1* Conditional Mouse Line. Mice with a floxed *DOCK1* allele were generated in collaboration with the Cell Migration Consortium (http://cellmigration.org). The *DOCK1* locus spans 502 kb in C57BL/6 mice and consists of 52 exons (genome browser: genome.ucsc.edu). A conditional targeting vector was designed to delete a 1-kb fragment that included a part of the promoter region, exon 1 that contains the start codon (ATG) and a part of intron 1 upon recombination between two inserted LoxP sites (Fig. S4). The targeting vector was also designed to insert a phosphoglycerate kinase (PGK) promoter fused to-Neo sequence into intron 1 between two short flippase recognition target (FRT) sites. We previously had success abrogating *DOCK1* expression when targeting the same region of the *DOCK1* locus in a nonconditional knockout mouse model (5) and therefore a similar strategy was expected to completely inactivate *DOCK1* upon Cre recombinase expression. Briefly, the targeting vector was designed by amplifying by PCR three genomic arms from a purified C57BL/6 BAC (no. RP230347M11; Children's Hospital Oakland Research Institute) and by cloning them into the PGKNeoF2L2DTA plasmid (provided by P. Soriano, Mount Sinai School of Medicine, New York) (see Table S2 for primer details). A first 3.2-kb arm was cloned into a NotI site, the 1-kb conditional arm was cloned into a SmaI site, and the third 2.7-kb arm was cloned into a HindIII site as depicted in Fig. S4 and all arms were fully sequenced. To obtain recombinant ES cells with the desired homologous recombination in the *DOCK1* locus, 25 μg of XmnI linearized vector was electroporated in C57/B6 embryonic stem cells using standard procedures at Massachusetts Institute of Technology, Cambridge, MA. Following selection with G418 antibiotics, more than 200 surviving clones were expanded and screened by Southern blots to identify recombinant cells. Purified genomic DNA was digested with SpeI and analyzed by Southern blotting with a 3′ probe located in intron 1 in a region residing outside the targeting vector. This genomic probe was amplified by PCR from the C57/B6 BAC DNA and cloned into the TOPO vector. For the wild-type and mutant allele, bands of 16.3 kb and 7.3 kb in sizes, respectively, were expected (Fig. S4 and Table S2).

The positive ES cell clones were electroporated with a plasmid coding for the flippase to remove the PGK-Neo cassette between the FRT sites. The ES clones were analyzed by Southern blot with the same strategy as above. The FRT-recombined band is at 6.7 kb and the final selected clones were used for injection in 129Sv blastocysts. Chimeras were obtained from the injection and the highest chimeric males were set up for mating with 129Sv mouse strain females to result in the successful germ-line transmission of the conditional *DOCK1* locus. The genotype of these F1 animals was confirmed by Southern blot. From that point, genomic DNA from mice tails was extracted using standard procedures and mice were genotype by PCR as depicted in Fig. S4 and Table S2. *DOCK1* mice were backcrossed for six generations in the FVB/NJ mouse strain genetic background before performing breast cancer studies.

Animal Care and Mouse Strains. Transgenic *NIC* (*MMTV-NeuNDL2-5-IRES-Cre*) FVB/NJ mice were previously described (6). FVB/NJ and athymic nude (nu/nu) mice were obtained from Charles River. Transgenic *MMTV-Cre* mice were previously reported (7). Mice were housed in a mouse-specific pathogen-free (SPF) facility. All mouse experiments were approved by the Institut de Recherches Cliniques de Montréal Animal Care Committee and complied with the Canadian Council of Animal Care rules.

Tumorigenesis Studies. *DOCK1* flox mice were crossed with *NIC* (*MMTV-NeuNDL2-5-IRES-Cre*) transgenic mice. Large cohorts of $NIC^{+}DOCK1^{wt/wt}$, $NIC^{+}DOCK1^{wt/flx}$, and $NIC^{+}DOCK1^{flx/flx}$ female mice were generated. Each female with a genotype of interest was analyzed in detail 5 wk after tumor onset, which was determined by physical palpation. At the time of necropsy, the amount and volume of tumor nodules, the presence of neoplastic lesions and lung metastases were determined exactly as done in ref. 8. Tumor volume was calculated as previously described (9).

Experimental Metastasis Assay. A total of 2×10^5 NMuMG-NeuNT, NMuMG-$NeuNT^{sh1\ DOCK1}$, and NMuMG-$NeuNT^{sh2\ DOCK1}$ cells were injected in the lateral tail vein of 12 athymic nu/nu mice per condition. At 30 d after injection, lungs were harvested and fixed in 4% paraformaldehyde. The total number of metastatic lesions per animal lung was counted under a dissection microscope (Zeiss).

Histology and Immunohistochemistry. Tumors, mammary glands, and lungs were fixed in 4% paraformaldehyde, paraffin embedded, sectioned at 5 μm, and stained with H&E for histological analysis. Lung sections were collected every 50 μm to determined metastases incidence and quantification. For immunohistochemistry analysis, tissue tumor microarray and paraffin sections were deparaffinized in xylene, rehydrated with ethanol gradient, treated using 10 mM sodium citrate buffer (pH 6) for antigen retrieval procedure according to standard method, blocked with 3% H_2O_2, permeabilized with immunohistochemistry (IHC) buffer (0.5% Triton X-100, 0.02% Tween-20/PBS), blocked with IHC buffer with 1% BSA and incubated with primary antibody overnight. (Primary antibodies are described in *Antibodies for Western Blotting and Immunohistochemistry*). Sections were washed three times with IHC buffer and incubated with biotinylated secondary antibody. Sections were washed three times with IHC buffer and incubated with streptavidin-HRP. Sections were washed three times with IHC buffer, one time with PBS, and the reaction was revealed using DAB peroxidase substrate kit (SK-4100, Vector Laboratories). Sections were counterstained with hematoxylin.

Tumor Microarray Staining. The human breast cancer tumor microarray was obtained from the Breast Cancer Genomic Group at the McGill University Health Centre Research Institute. Briefly, slides were subjected to an antigen retrieval procedure (See *Histology and Immunohistochemistry* for a complete description of the method). DOCK1 expression was visualized using DOCK1 (H4) as a primary antibody (1:100; Santa Cruz Biotechnology). Secondary antibody staining was performed as described above. Levels of DOCK1 expression on each breast tumor biopsy was attributed either a negative, weak, moderate, or strong value.

Whole-Mount Analysis of Mammary Glands. To analyze if normal mammary gland development requires *DOCK1*, we intercrossed *MMTV-Cre* and $DOCK1^{flx/flx}$ animals. Mammary glands 4 and 9 from 9, 12,

and 15 wk old *MMTVCre*$^{+}$*DOCK1*$^{wt/wt}$ and *MMTVCre*$^{+}$*DOCK1*$^{flx/flx}$ mice were harvested. Mammary glands were fixed in 4% paraformaldehyde, removed from fat using two changes of acetone, and stained overnight in carmine red solution. Tissues were rehydrated following an ethanol gradient, cleared in xylene, and mounted using Permount (SP15-500, Fisher). The developmental status of the mammary gland was assessed from micrograph of multiple glands per genotype.

Quantification of Mammary Intraepithelial Neoplastic Lesions. Tumor-free mammary glands 4 and 9 from *NIC*$^{+}$*DOCK1*$^{wt/wt}$, *NIC*$^{+}$*DOCK1*$^{wt/flx}$, and *NIC*$^{+}$ *DOCK1*$^{flx/flx}$ female mice were isolated at 5 wk after tumor onset and mounted for whole-mount analysis as described above. The average number of neoplastic lesions from these mammary glands was determined by counting the total amount of mammary intraepithelial neoplastic lesions (MINs) under the microscope.

mRNA Isolation, RNA-Sequencing and Analysis, Q-PCR Validation. The total mRNA from 4 *NIC*$^{+}$*DOCK1*$^{wt/wt}$ and 4 *NIC*$^{+}$*DOCK1*$^{flx/flx}$ average size tumors was extracted using TRIZOL reagent (Invitrogen) and according to recommended procedures. Total RNA was cleaned up using a RNeasy column (Qiagen) and on-column DNase treatment was performed using RNase-Free DNase set (Qiagen). An initial input of 10 μg of total RNA for each sample was used to generate expression libraries, cBot cluster and to performed deep sequencing using Illumina TruSeq RNA Sample Preparation kit, Illumina TruSeq SR cluster kit v2 and Illumina TruSeq SBS Kit V2 50 cycles, respectively, and by following manufacturer procedures. Sequencing was performed at the McGill University and Genome Quebec Innovation Center using the Illumina HiSEq 2000 platform. Alignment of short reads was performed using TopHat version 1.4.1 (10); gene expression values were subsequently estimated using Cufflinks version 1.3.0 (11) Class comparisons were performed using DESeq version 1.8.3 (12). Genes were considered significantly differentially expressed when their nominal *P* value was less than 0.05 and genes presenting an average differential expression between the normalized counts of at least twofold between *NIC*$^{+}$*DOCK1*$^{wt/wt}$ and *NIC*$^{+}$*DOCK1*$^{flx/flx}$ conditions were selected to generate the *DOCK1*-null signature.

To validate the differential gene expression data obtained using RNA-seq, cDNAs were generated from the total RNA isolated from 4 *NIC*$^{+}$*DOCK1*$^{wt/wt}$ and four *NIC*$^{+}$*DOCK1*$^{flx/flx}$ tumors using SuperScript II (Invitrogen) and random primers (Invitrogen) as recommended by the manufacturer. The expression of eight differentially expressed genes was confirmed by RT-qPCR (see Table S2 for gene list and primer information). RT-qPCR was realized using SYBR Green PCR Master Mix (Applied Biosystems). Specificity of the reaction was verified by melt curve analysis for each primer set. TATA box binding protein was used as an internal control. All qPCR reactions were performed for 5 min at 45 °C, for 3 min at 95 °C, followed by 40 cycles of 15 s at 45 °C and 30 s at 60 °C.

Expression levels in *NIC*$^{+}$ tumors of all of the currently described Rho GTPases, Rho Guanine Exchange Factors (RhoGEFs), Rho GTPase Activating Proteins (RhoGAPs), Rho Guanine nucleotide Dissociation Inhibitors (GDIs) and Engulfment and cell motility (ELMOs) in the literature were obtained by performing mRNA deep sequencing. Expression levels were quantified by using the normalized raw count generated by mRNA deep sequencing after deseq analysis. Data were expressed as mean normalized raw count ± SD.

To look at the expression level of down-regulated genes in cell lines, 2.2×10^6 NMuMG-EV, NMuMG-NeuNT, and NMuMG-NeuNT$^{sh1\ DOCK1}$ were plated and allowed to grow for 24 h. The total mRNA from NMuMG-EV, NMuMG-NeuNT, NMuMG-NeuNT$^{sh1\ DOCK1}$ was extracted using TRIZOL reagent (Invitrogen) and according to recommended procedures ($n = 6$ for each cell line). After DNase I treatement (NEB), cDNAs were generated from the total RNA isolated using M-MuLV Reverse Transcriptase (NEB) and random primers (NEB) as recommended by the manufacturer. The expression of seven genes was analyzed by RT-PCR (see Table S2 for gene list and primer information).

Statistical Analysis. Kaplan–Mayer curves were analyzed using Prism software and *P* value were calculated using log-rank test. For other statistical analysis, *P* values were calculated by using Student *t* test or by using ANOVA test and Bonferroni's posttest and Prism software. Gene ontology analysis was realized using DAVID tool (13).

Antibodies for Western Blotting and Immunohistochemistry. Antibody dilution and catalog numbers for Western blots are shown in brackets. Antibodies from Santa Cruz Biotechnology include: DOCK1 (1:5,000; H-70), phospho-Neu^{Y877} (1:1,000; sc-101695), Neu (1:1,000; C-18), and Myc-9E10 (1:2,000; sc-40). M2-Flag (1:10 000; F3165) and TUBULIN (1:10 000; T5168) antibodies were from Sigma. Rac1 (1:3,000; 17-283) was obtained from Upstate Biotechnologies. Antibodies from Cell Signaling Technology include: phospho-DOCK1^{Y1811} (1:1,000; kind gift from Susan Keezer, Cell Signaling Technology, Boston), phospho-STAT3^{Y705} (1:1,000; 9145S), phospho-STAT3^{S727} (1:1,000; 9134S), STAT3 (1:1,000; 9132S), phospho-STAT1^{Y701} (1:1,000; 9167), STAT1 (1:1,000; 9172), phospho-IKKα/β$^{S76/180}$ (1:1,000; 2697), IKKα (1:1,000; 2682), phospho-AKT^{S473} (1:1,000, 9271S), AKT (1:1,000; 9272), phospho-RS6K$^{S235/236}$ (1:1,000; 2211S), phospho-p130CAS (1:1,000; 4011S), phospho-SRC^{Y416} (1:1,000; 2101S), SRC (1:1,000; 2108); phospho-FAK^{Y397} (1:1,000; 3283S) phospho-c-JUN^{S73} (1:1,000; 3270S).

Antibodies used in immunohistochemistry include: DOCK1 (1:100; C19; Santa Cruz Biotechnology), Cre (1:600; PRB106C; Covance), Ki67 (1:250; 275R-15; Cell Marque), Cleaved-caspase3 (1:200; 9661S; Cell Signaling), pSTAT3^{Y705} (1:100; 9145S; Cell Signaling), STAT3 (1:500; 9132S; Cell Signaling), pSTAT1^{Y701} (1:100; 9167; Cell Signaling), STAT1 (1:500; 9172; Cell Signaling), pc-JUN^{S73} (1:100; 3270S; Cell Signaling).

1. Haibe-Kains B, et al. (2012) A three-gene model to robustly identify breast cancer molecular subtypes. *J Natl Cancer Inst* 104(4):311–325.
2. Northey JJ, et al. (2008) Signaling through ShcA is required for transforming growth factor beta- and Neu/ErbB-2-induced breast cancer cell motility and invasion. *Mol Cell Biol* 28(10):3162–3176.
3. Côté JF, Vuori K (2002) Identification of an evolutionarily conserved superfamily of DOCK180-related proteins with guanine nucleotide exchange activity. *J Cell Sci* 115(Pt 24):4901–4913.
4. Nishikimi A, et al. (2012) Blockade of inflammatory responses by a small-molecule inhibitor of the Rac activator DOCK2. *Chem Biol* 19(4):488–497.
5. Laurin M, et al. (2008) The atypical Rac activator Dock180 (Dock1) regulates myoblast fusion in vivo. *Proc Natl Acad Sci USA* 105(40):15446–15451.
6. Ursini-Siegel J, et al. (2012) The ShcA SH2 domain engages a 14-3-3/PI3′K signaling complex and promotes breast cancer cell survival. *Oncogene* 31(48):5038–5044.
7. Andrechek ER, et al. (2000) Amplification of the neu/erbB-2 oncogene in a mouse model of mammary tumorigenesis. *Proc Natl Acad Sci USA* 97(7):3444–3449.
8. Ranger JJ, Levy DE, Shahalizadeh S, Hallett M, Muller WJ (2009) Identification of a Stat3-dependent transcription regulatory network involved in metastatic progression. *Cancer Res* 69(17):6823–6830.
9. Brantley-Sieders DM, et al. (2008) The receptor tyrosine kinase EphA2 promotes mammary adenocarcinoma tumorigenesis and metastatic progression in mice by amplifying ErbB2 signaling. *J Clin Invest* 118(1):64–78.
10. Trapnell C, Pachter L, Salzberg SL (2009) TopHat: Discovering splice junctions with RNA-Seq. *Bioinformatics* 25(9):1105–1111.
11. Trapnell C, et al. (2012) Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nat Protoc* 7(3):562–578.
12. Anders S, Huber W (2010) Differential expression analysis for sequence count data. *Genome Biol* 11(10):R106.
13. Huang W, Sherman BT, Lempicki RA (2009) Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc* 4(1):44–57.

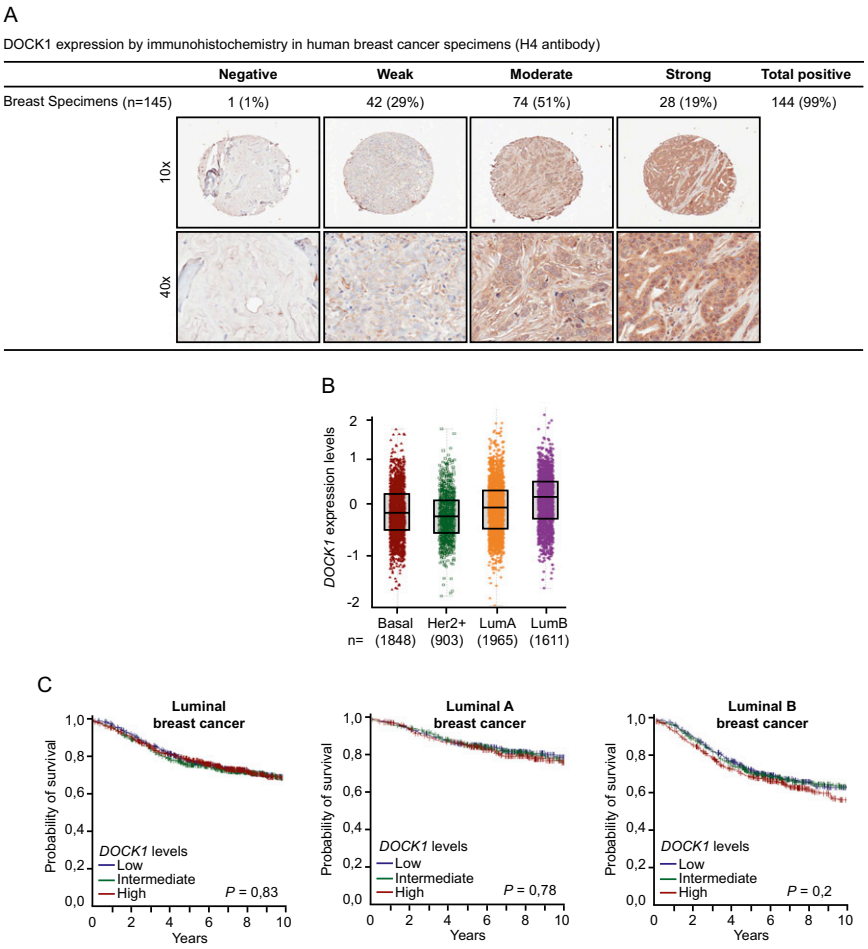

Fig. S1. *DOCK1* is expressed in human breast cancer. (*A*) DOCK1 immunohistochemistry analysis on a breast tumor microarray containing 145 human breast tumor specimens. Levels of DOCK1 expression on each section was attributed either a negative, weak, moderate, or strong value. (*B*) Quantification of the levels of *DOCK1* expression in breast cancer subtypes. (*C*) High levels of *DOCK1* expression is not indicative of prognostic in luminal breast cancer subtypes.

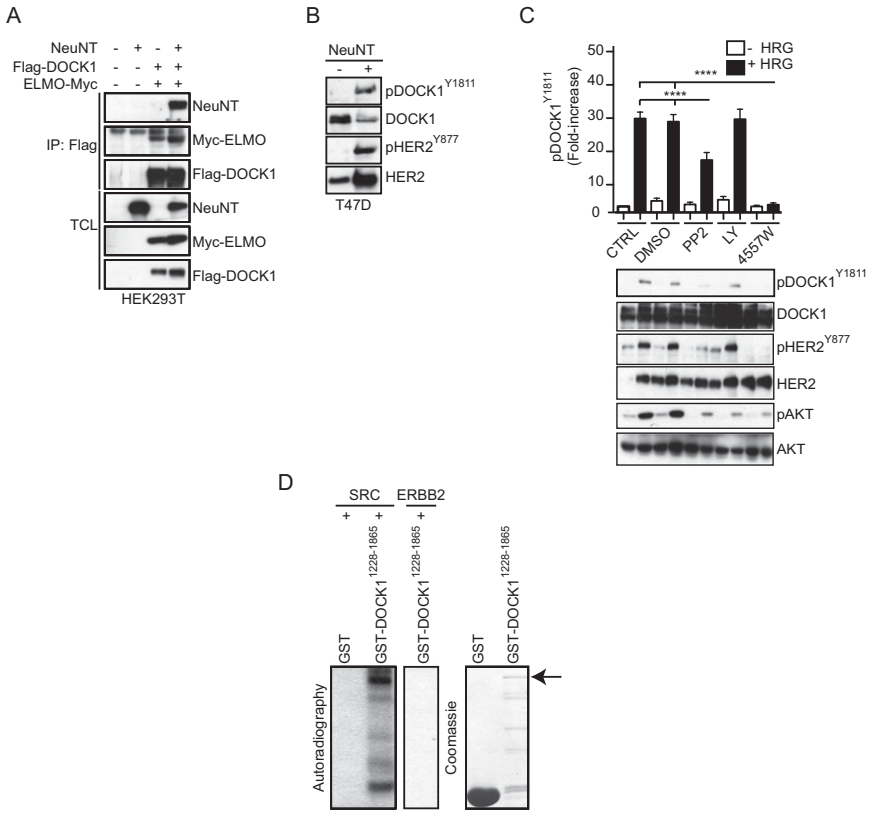

Fig. S2. DOCK1 is phosphorylated by SRC kinase downstream of HER2 activation. (*A*) DOCK1 is found in a complex with oncogenic NeuNT. Flag-DOCK1, Myc-ELMO1, and NeuNT were coexpressed and DOCK1 was immunoprecipitated with anti-FLAG beads ($n = 3$). (*B*) Induction of DOCK1 phosphorylation at Y^{1811} in T47D breast cancer cells upon overexpression of NeuNT ($n = 5$). (*C*) DOCK1 phosphorylation on Y^{1811} is SRC dependent. T47D cells breast cancer cells were treated with 20 ng/mL of HRG for 15 min in the absence or presence of DMSO, 5 μM SRC PP2, 10 μM PI3K LY-294002, or 50 μM EGFR/HER2 4557W inhibitors. Quantification of the pDOCK1^{Y1811} signal is shown (*Upper*) ($n = 7$). (*D*) Active recombinant SRC and HER2 kinases were incubated with recombinant GST or the GST-DOCK1$^{1228\text{-}1865}$ fragment, followed by in vitro kinase assay ($n = 3$). Data were expressed as mean ± SD. One-way ANOVA followed by a Bonferroni test calculated the indicated *P* values; $^{*}P \leq 0.05$, $^{**}P \leq 0.01$, $^{***}P \leq 0.001$, and $^{****}P \leq 0.0001$.

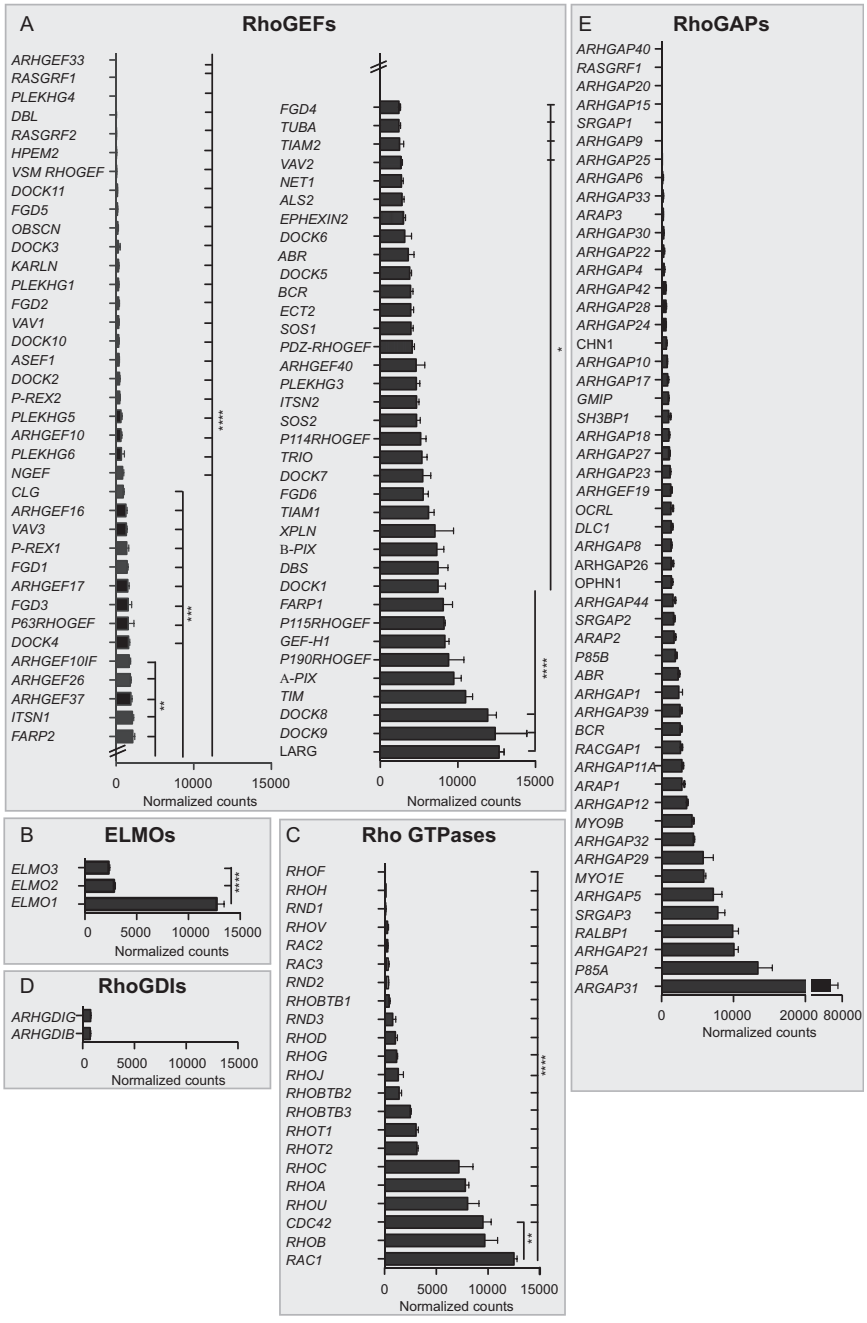

Fig. S3. Portrait of Rho GTPase regulation in Neu-induced tumors. Expression profiling of all RhoGEFs (*A*), ELMOs (*B*), Rho GTPases (*C*), RhoGDIs (*D*), and RhoGAPs (*E*) in Neu(NDL2-5)-induced mouse mammary tumors. (*A–E*) Expression was determined by mRNA deep sequencing. Data were expressed as mean ± SD $n = 4$. One-way ANOVA followed by a Bonferroni test calculated the indicated *P* values; $*P \leq 0.05$, $**P \leq 0.01$, $***P \leq 0.001$, and $****P \leq 0.0001$.

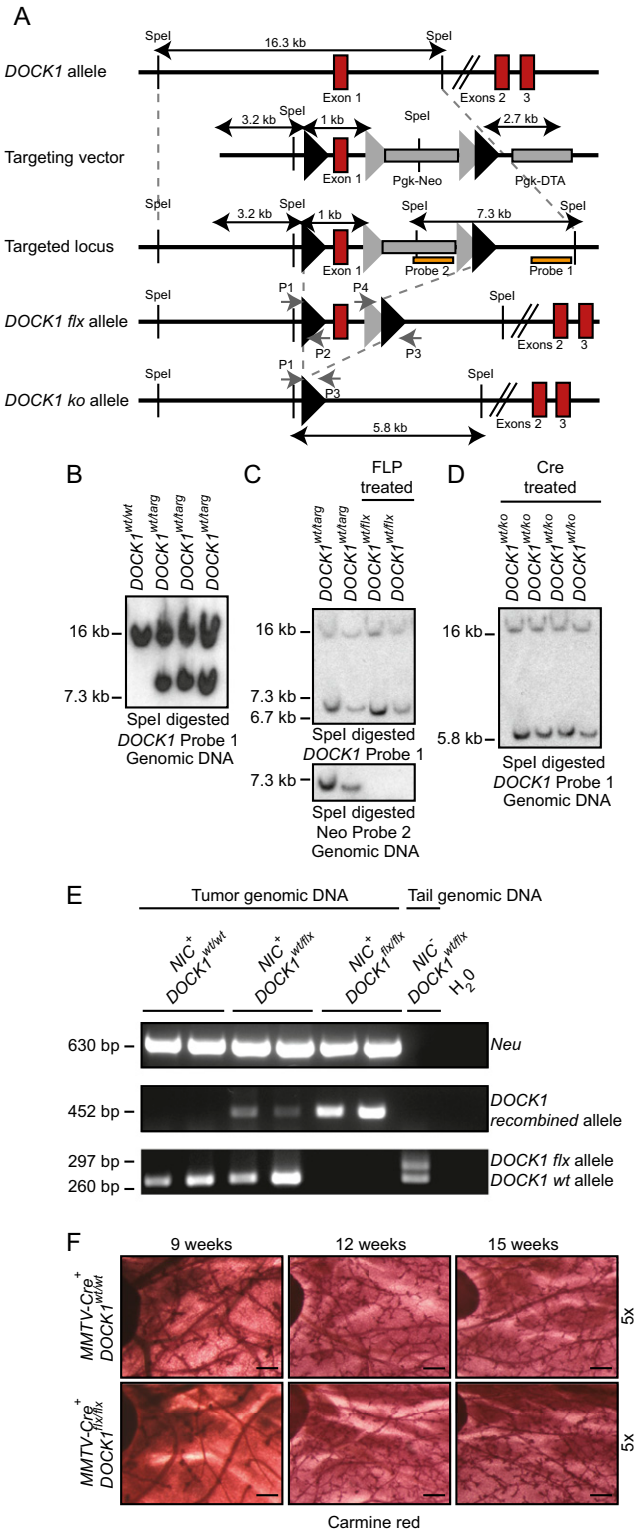

Fig. S4. Generation of a *DOCK1* conditional knockout mouse. (*A*) Partial representation of the *DOCK1* locus, structure of the targeting vector, organization of the rearranged *DOCK1* targeted locus (*targ*) and rearranged flox allele (*flx*) after FLP treatment and knockout allele (*ko*) after *Cre* treatment. The probes used for Southern blot and the expected size of the SpeI fragments are indicated. Locations of PCR primers used for genotyping are also indicated. (*B*) Southern blot analysis using a *DOCK1* genomic probe (probe 1). *Wt* and *targ* alleles produced 16 kb and 7.3 kb bands, respectively. (*C*) Southern blot analysis using DNA isolated from ES cells treated with FLP recombinase confirmed removal of the PGK-neo cassette. Genomic DNA was digested with SpeI and analyzed using *DOCK1* (probe 1) and Neo (probe 2) probes. *WT*, *targ* and *flx* alleles produced 16 kb, 7.3 kb and 6.7 kb bands, respectively. (*D*) Southern blot analysis using a *DOCK1* genomic probe (probe 1) confirming recombination of the *flx* allele after *Cre* recombinase treatment. *Wt* and *ko* alleles generated 16 kb and 5.8 kb bands, respectively. (*E*) PCR analysis using P1, P2, P3, and Neu primers sets on genomic DNA extracted from mammary tumors or tails from mice of the indicated genotype. (*F*) Whole mount inguinal mammary glands from *MMTV-Cre*$^{+}$*DOCK1*$^{wt/wt}$ and *MMTV-Cre*$^{+}$*DOCK1*$^{flx/flx}$ mice at 9, 12, and 15 wk of age stained with carmine red showing normal ductal trees. (Scale bar: 200 μm, 5×.) $n = 3$ for each genotype.

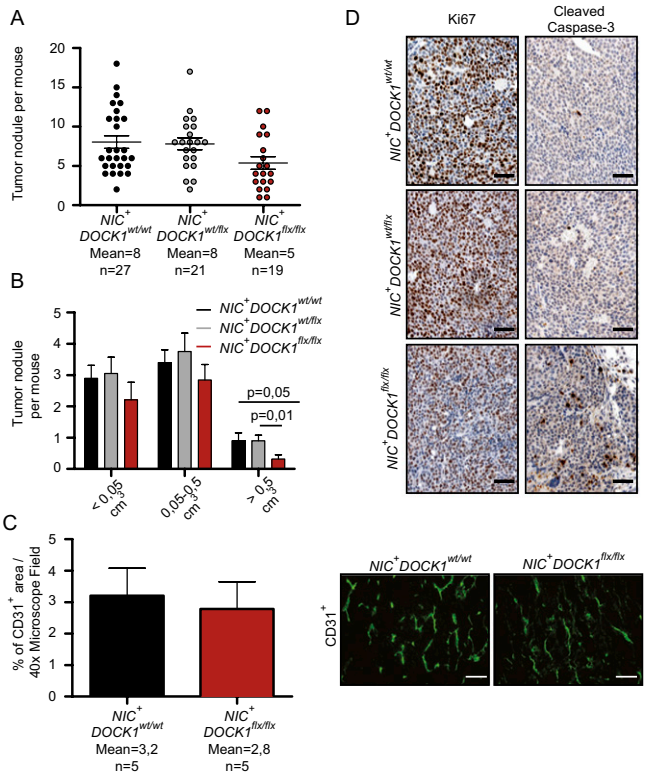

Fig. S5. DOCK1 signaling regulates tumor growth by promoting cell proliferation and blocking apoptosis. (*A*) Quantification of tumor nodules per animal for *NIC*$^{+}$*DOCK1*$^{wt/wt}$, *NIC*$^{+}$*DOCK1*$^{wt/flx}$, and *NIC*$^{+}$*DOCK1*$^{flx/flx}$ mice. (*B*) Size distribution of tumor nodules in all three genotypes. *NIC*$^{+}$*DOCK1*$^{flx/flx}$ mice show a significant reduction in the amount of large tumors. The indicated *P* values were calculated from a two-tailed Student *t* test. (*C*) Average percentage of the area covered by CD31^{+} staining per 40× microscope field (*Left*) and representative immunohistochemistry analysis on frozen tumor sections showing CD31^{+} expression in *NIC*$^{+}$*DOCK1*$^{wt/wt}$ and *NIC*$^{+}$*DOCK1*$^{flx/flx}$ tumors (*Right*). (Scale bar: 50 μm, 40×.) *P* value was calculated using a two-tailed Student *t* test. (*D*) Immunohistochemistry analysis on paraffin embedded tumor sections showing Ki67 (*Left*) and cleaved caspase-3 (*Right*) expression in *NIC*$^{+}$*DOCK1*$^{wt/wt}$, *NIC*$^{+}$*DOCK1*$^{wt/flx}$, and *NIC*$^{+}$*DOCK1*$^{flx/flx}$ tumors. (Scale bar: 50 μm, 40×.)

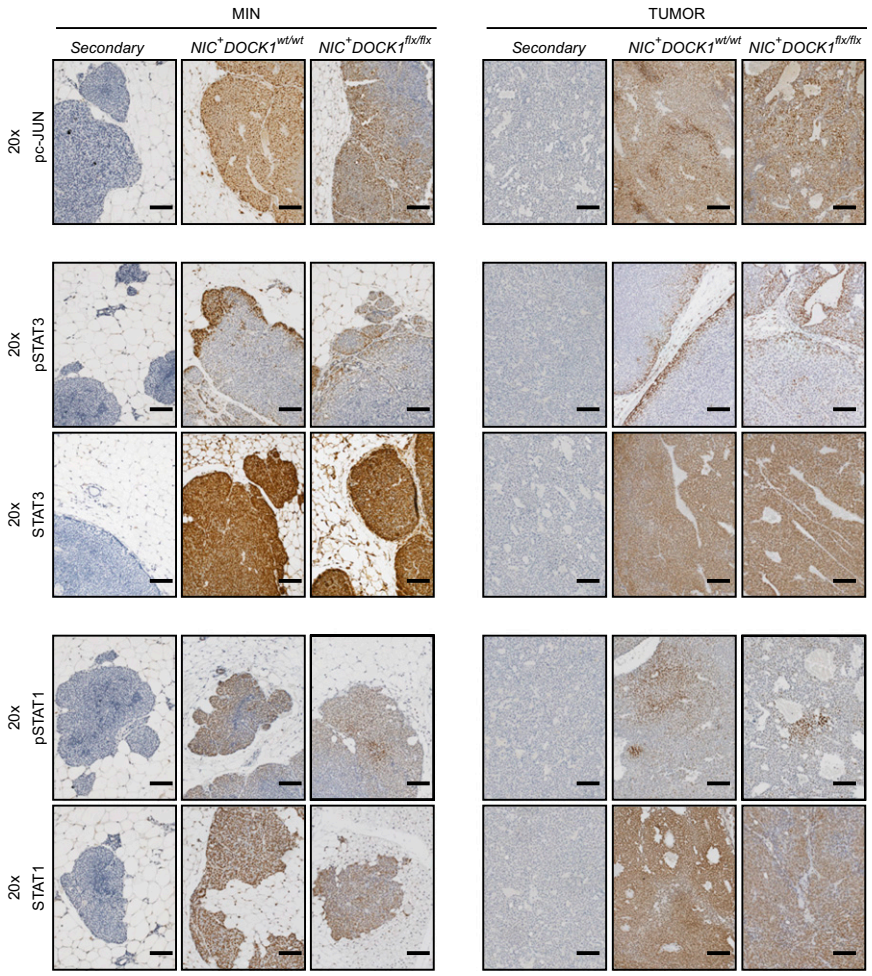

Fig. S6. DOCK1 regulates c-JUN and STAT3 activation in HER2-driven MINs and STAT1 expression levels in HER2-driven tumors. Immunohistochemistry analyses showing pc-JUN, pSTAT3, STAT3, pSTAT1, STAT1 staining in *NIC*$^{+}$*DOCK1*$^{wt/wt}$ and *NIC*$^{+}$*DOCK1*$^{flx/flx}$ mice mammary gland lesions and tumors. (Scale bar: 100 μm, 20×.)

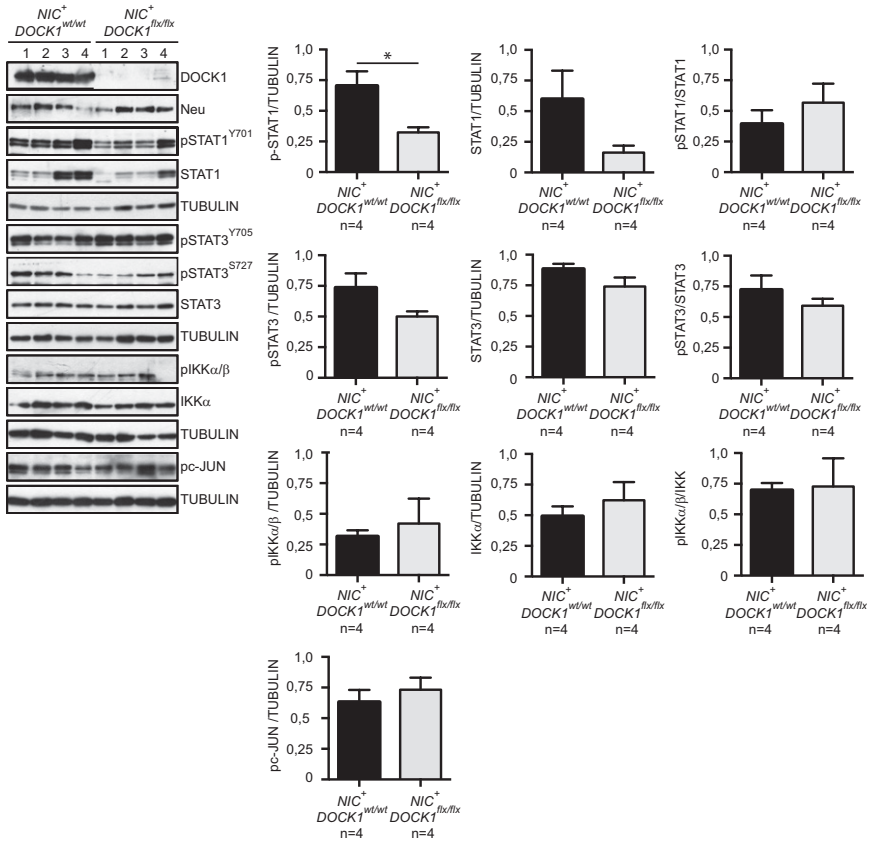

Fig. S7. Reduced levels and activation of STAT1 in HER2-driven tumors in the absence of *DOCK1* expression. Total tumor lysates were immunobloted with the indicated antibodies. Quantification of band intensity was done using Fiji software. *P* value was calculated from a two-tailed Student *t* test.

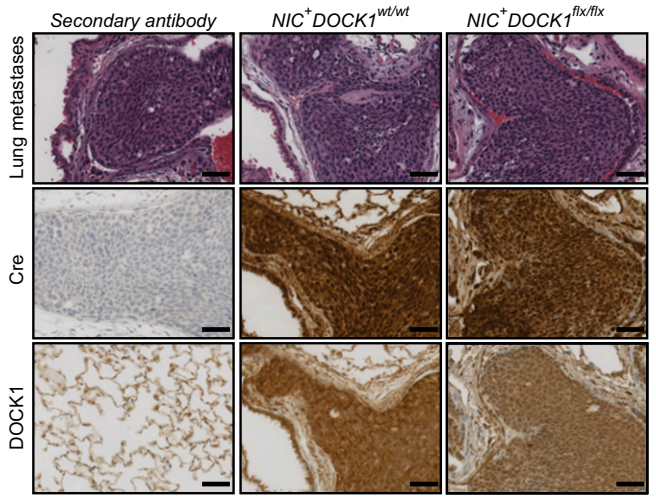

Fig. S8. DOCK1 is expressed in lung metastasis. Immunohistochemistry analysis on paraffin embedded lung sections showing H&E-staining (*Upper*), Cre (*Middle*), and DOCK1 (*Lower*) expression in $NIC^{+}DOCK1^{wt/wt}$ and $NIC^{+}DOCK1^{flx/flx}$ metastasis. (Scale bar: 50 μm, 40×.)

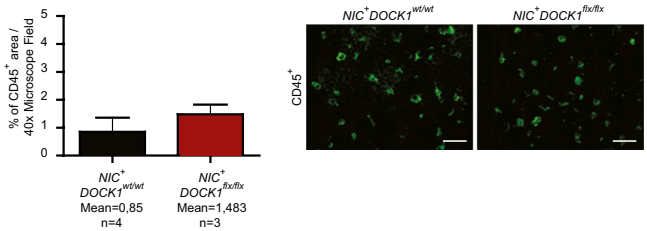

Fig. S9. Deletion of *DOCK1* expression in NIC^{+} tumor does not affect white blood cell recruitment. Average percentage of the area covered by CD45^{+} staining per 40× microscope field (*Left*) and representative immunohistochemistry analysis on frozen tumor sections showing CD45^{+} expression in *NIC^{+}DOCK1$^{wt/wt}$* and *NIC^{+}DOCK1$^{flx/flx}$* tumors (*Right*). (Scale bar: 50 μm, 40×.) *P* value was calculated from a two-tailed Student *t* test.

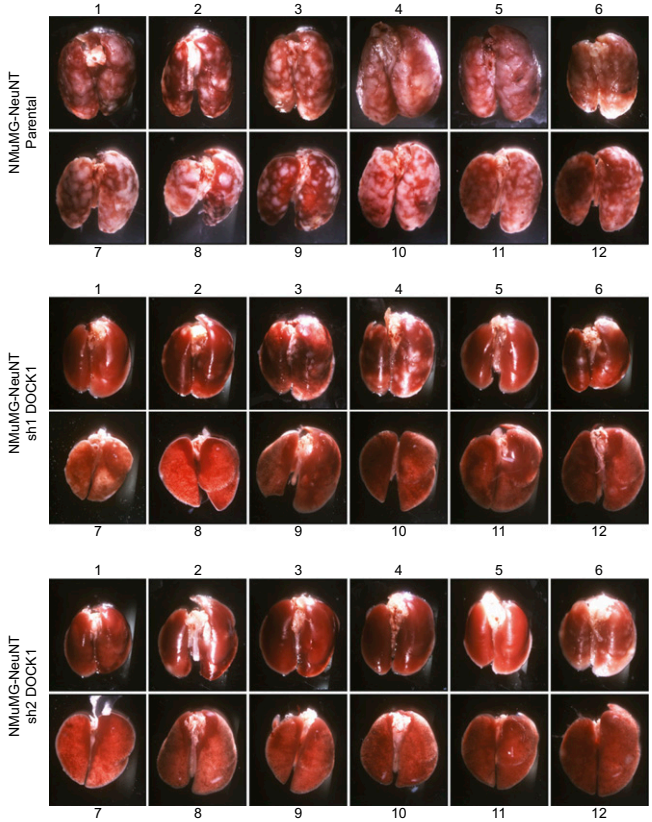

Fig. S10. DOCK1 is essential for HER2-mediated lung metastasis in experimental metastasis assay. Pictures of collected lungs from mice injected with NMuMG-NeuNT, NMuMG-NeuNT$^{sh1\ DOCK1}$, and NMuMG-NeuNT$^{sh2\ DOCK1}$ cells.

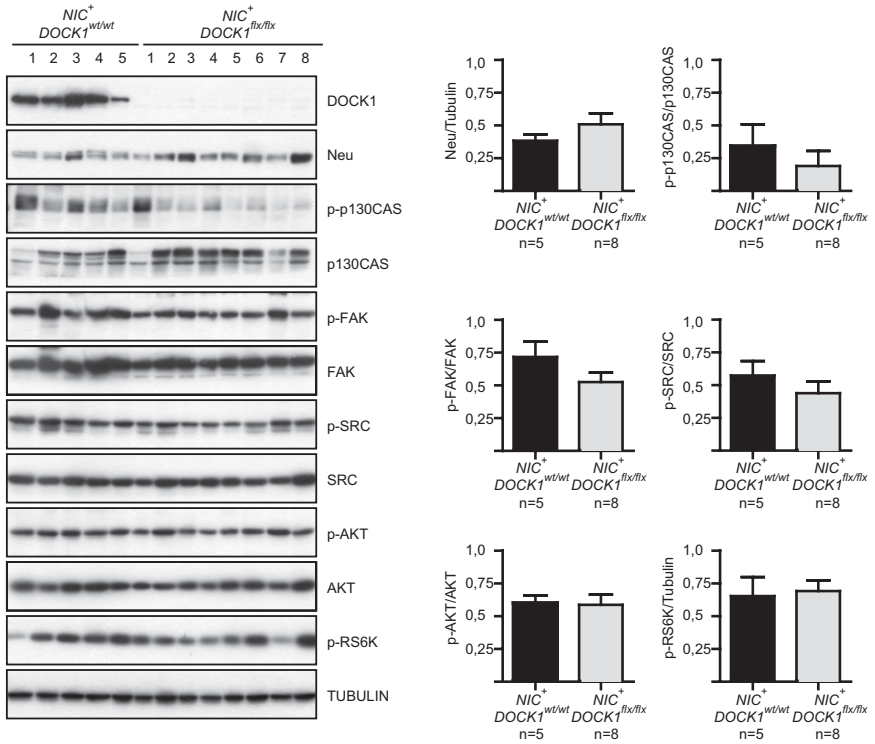

Fig. S11. Components upstream of DOCK1 in the integrin-signaling pathways are activated correctly in *DOCK1*-null tumors. Total tumor cell lysates were immunoblotted with the indicated antibodies. Quantification of band intensity was done using Fiji software. *P* value was calculated from a two-tailed Student *t* test.

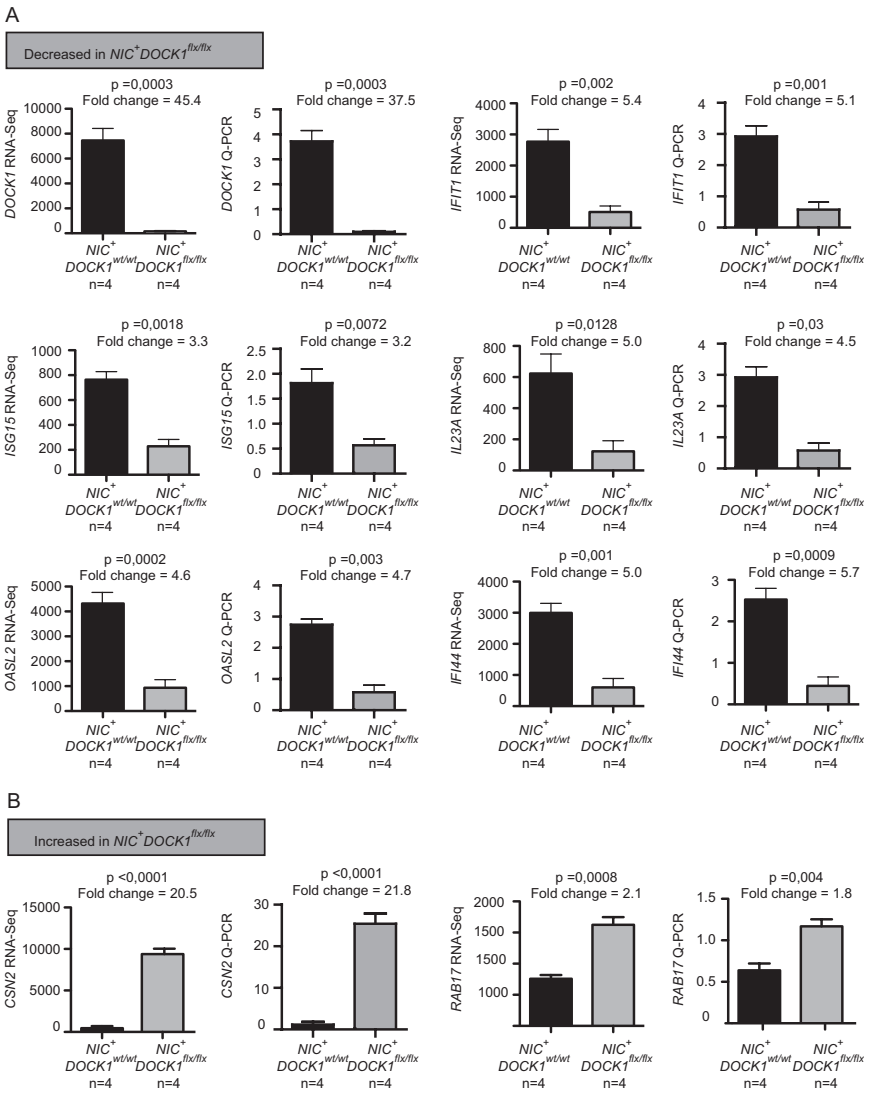

Fig. S12. Validation of the differential expression data obtained by RNA-sequencing using Q-PCR analysis. (*A*) Q-PCR validation of the expression levels of a panel of genes identified as being decreased in *NIC*$^{+}$*DOCK1*$^{flx/flx}$ tumors by mRNA deep sequencing. *Left* graph represents deep sequencing analysis and *Right* graph represents Q-PCR analysis. (*B*) Q-PCR validation of the expression levels of a panel of genes identified as being increased in *NIC*$^{+}$*DOCK1*$^{flx/flx}$ tumors by mRNA deep sequencing. *Left* graphs represent deep sequencing analysis and *Right* graphs represent Q-PCR analysis. The indicated *P* values were calculated from a two-tailed Student *t* test.

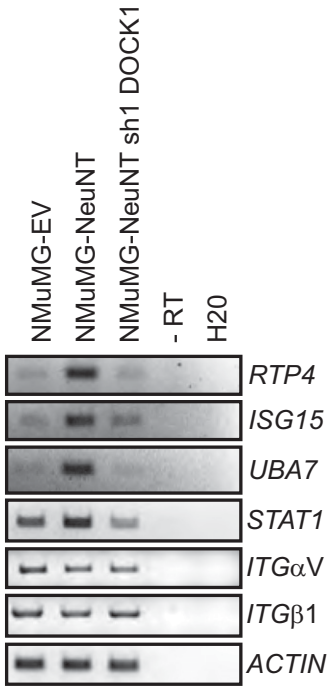

Fig. S13. Oncogenic HER2 elevates IFN-response gene expressions in culture. NMuMG-EV, NMuMG-NeuNT, and NMuMG-NeuNT$^{sh1\ DOCK1}$ were grown for 24 h after plating and analyzed for the expression of selected IFN response genes by RT-PCR.

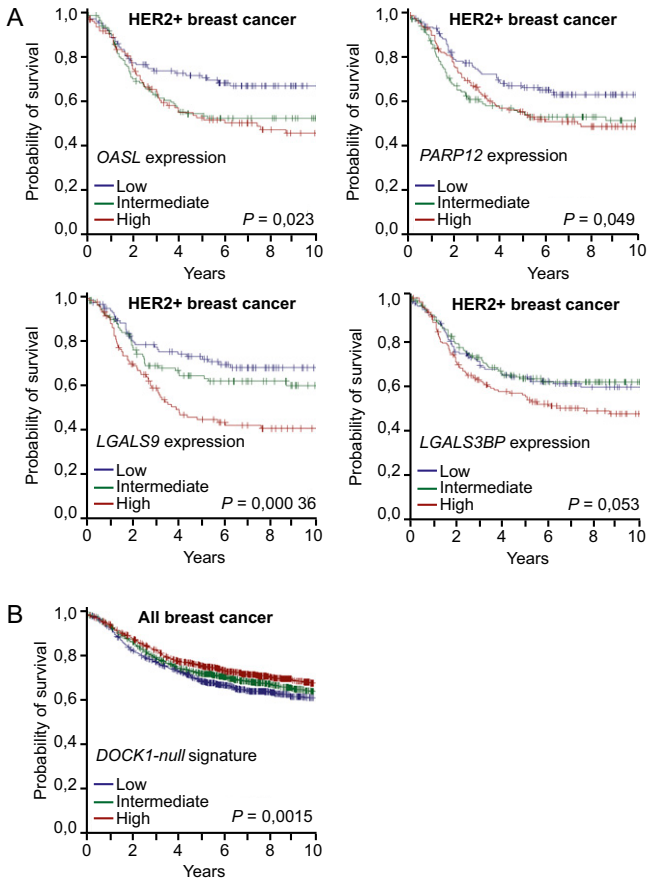

Fig. S14. Expression levels of some *DOCK1*$^{flx/flx}$-signature genes are predictive of disease-free survival in HER2^{+} breast cancer. (*A*) High expression levels of *PARP12*, *OASL*, *LGAlS9*, or *LGALS3BP,* that are part of the genes down-regulated in the *NIC*$^{+}$*DOCK1*$^{flx/flx}$ signature, are associated with poor disease-free survival in HER2^{+} breast cancer patients. (*B*) The *DOCK1*$^{flx/flx}$-associated gene signature predicts outcome for all breast cancers.

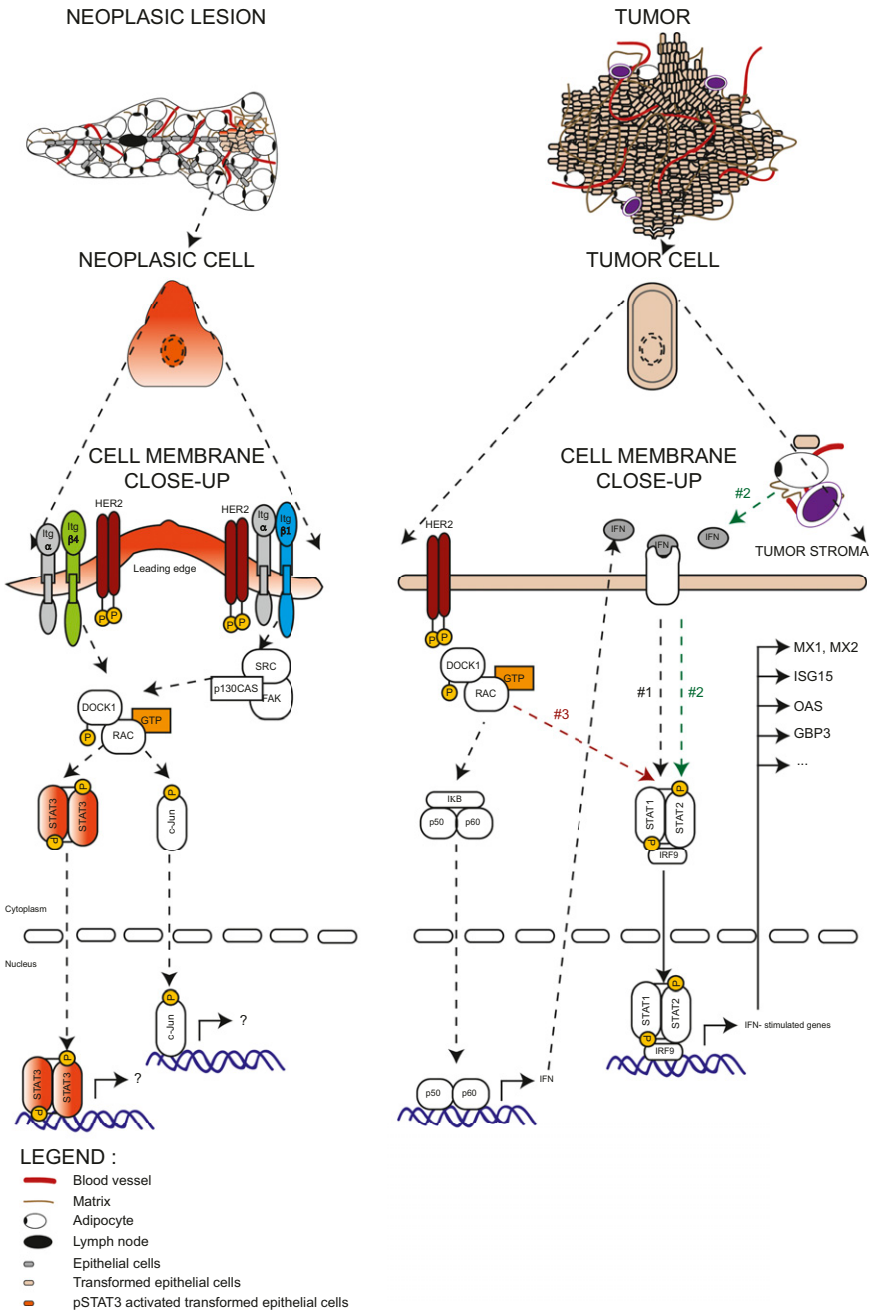

Fig. S15. Model of signaling downstream of DOCK1 in *NIC*$^{+}$ MINs and tumors. See *Discussion* for a description of the model.

Table S1. RNA-Seq analysis overview

Gene symbol	Norm count 1 *NIC*$^{+}$ *DOCK1*$^{wt/wt}$	Norm count 2 *NIC*$^{+}$ *DOCK1*$^{wt/wt}$	Norm count 3 *NIC*$^{+}$ *DOCK1*$^{wt/wt}$	Norm count 4 *NIC*$^{+}$ *DOCK1*$^{wt/wt}$	Norm count 1 *NIC*$^{+}$ *DOCK1*$^{flx/flx}$	Norm count 2 *NIC*$^{+}$ *DOCK1*$^{flx/flx}$	Norm count 3 *NIC*$^{+}$ *DOCK1*$^{flx/flx}$	Norm count 4 *NIC*$^{+}$ *DOCK1*$^{flx/flx}$	Average norm count *NIC*$^{+}$ *DOCK1*$^{wt/wt}$	Average norm count *NIC*$^{+}$ *DOCK1*$^{flx/flx}$	Deseq *P* value	Fold change wt/flx
CSN2	247	70	1,209	300	10,532	9,004	7,720	10,199	456.5	9,363.75	3.5E-36	0.048731553
CSN1S1	457	33	750	347	4,752	3,686	3,692	3,454	396.75	3,896	1.5E-22	0.101813445
OASL2	4,068	4,298	5,127	3,781	403	1,742	1,200	391	4,318.5	934	1.1E-11	4.623546826
IFI44	3,405	2,795	3,565	2,229	131	1,360	738	189	2,998.5	604.5	2E-10	4.962269191
IFIT1	2,456	2,645	3,887	2,106	154	988	651	255	2,773.5	512	0.000000004	5.41889991
XAF1	1,211	1,642	1,864	1,136	142	721	482	189	1,463.25	383.5	0.00000013	3.813194181
MX2	593	925	891	544	56	327	259	116	738.25	189.5	0.00000036	3.896017074
MX1	722	675	623	367	146	237	190	64	596.75	159.25	0.00000059	3.74508899
GBP3	2,707	3,899	4,835	2,344	387	1,141	965	532	3,446.25	756.25	0.00000064	4.556733509
OAS2	758	851	617	505	151	397	155	74	682.75	194.25	0.00000064	3.513721873
ISG15	695	981	799	583	124	370	264	160	764.5	229.5	0.0000017	3.328794939
DDX60	1,151	1,323	2,182	1,751	138	769	509	231	1,601.75	411.75	0.0000024	3.89061979
RTP4	4,760	4,799	5,816	4,963	538	2,861	2,599	1,626	5,084.5	1,906	0.0000087	2.668446339
UBA7	1,754	2,125	1,588	1,676	356	1,145	970	311	1,785.75	695.5	0.000036	2.568631618
PSMB8	2,685	3,891	3,636	3,309	822	2,527	1,539	897	3,380.25	1,446.25	0.000088	2.337554497
PARP12	2,766	2,791	3,492	2,408	883	1,557	1,450	1,025	2,864.25	1,228.75	0.00012	2.331082396
ZBP1	1,857	1,481	1,217	1,161	354	941	787	281	1,429	590.75	0.00016	2.418311352
LGALS3BP	7,523	9,746	15,019	10,662	2,267	6,625	4,464	1,469	10,737.5	3,706.25	0.00017	2.897884309
IFI47	1,250	791	1,265	1,287	328	885	418	157	1,148.25	447	0.0002	2.568631618
BST2	2,236	3,252	5,074	2,750	580	1,505	1,256	875	3,328	1,054	0.00025	3.157975547
RPL34	2,383	2,279	2,114	156	188	177	210	135	1,733	177.5	0.00025	9.767572101
USP18	1,521	2,391	4,670	1,807	79	730	538	156	2,597.25	375.75	0.00042	6.91629785
GBP7	2,495	2,955	3,367	2,616	877	1,953	1,359	1,135	2,858.25	1,331	0.00045	2.148008943
LGALS9	4,649	4,991	2,911	2,326	1,112	2,102	1,606	828	3,719.25	1,412	0.00064	2.635359903
IRGM1	5,708	5,627	6,721	7,052	2,235	4,410	3,538	2,124	6,277	3,076.75	0.0008	2.040609318
CMPK2	1,094	1,128	2,151	1,250	248	853	531	261	1,405.75	473.25	0.0008	2.971105841
PARP14	4,775	4,091	6,104	3,970	1,144	3,668	2,588	1,281	4,735	2,170.25	0.0009	2.182527754
HERC6	1,505	1,939	1,781	2,074	517	1,444	1,046	517	1,824.75	881	0.0011	2.070529848
OAS1B	423	582	484	472	68	380	236	147	490.25	207.75	0.0012	2.355445579
ESR1	98	250	62	742	1,088	1,250	1,907	588	288	1,208.25	0.0015	0.238489888
DHX58	1,988	2,249	2,233	1,481	367	1,022	1,012	1,513	1,987.75	978.5	0.0016	2.030732202
OAS1A	966	1,257	2,073	1,331	90	887	642	608	1,406.75	556.75	0.0029	2.526254524
IFIH1	3,599	3,137	5,633	4,002	644	3,298	1,967	1,269	4,092.75	1,794.5	0.0034	2.279946545
LALBA	566	276	9,705	14,849	8,291	70,245	24,157	34,113	6,349	34,201.5	0.022	0.185694115
PGF	9,238	9,637	7,894	2,471	3,876	753	4,362	3,888	7,310	3,219.75	0.024	2.270484204
DSH1C	1,982	3,076	4,165	893	425	1,216	1,207	892	2,529	935	0.026	2.705696462
RSAD2	1,065	1,532	3,939	690	125	457	610	129	1,806.5	330.25	0.03	5.467949808
SLFN8	359	406	905	392	236	317	261	88	515.5	225.5	0.034	2.286276671
FBLN5	681	2,391	281	1,504	12,829	6,218	4,944	882	1,214.25	6,218.25	0.036	0.19533197
SLFN2	1,410	942	2,538	1,416	477	1,256	915	253	1,576.5	725.25	0.038	2.173469725
IFI27I2A	289	335	1,177	632	143	372	126	132	608.25	193.25	0.04	3.142690067
OBSL1	1,635	1,058	418	1,295	276	306	785	679	1,101.5	511.5	0.041	2.15546629
MMP12	738	340	412	198	2,205	202	3,354	1,013	422	1,693.5	0.043	0.249307813
SMOC1	920	1,494	33	597	5,591	816	1,350	4,818	761	3,143.75	0.047	0.242154547
STAT1	3,789	4,318	11,040	7,050	1,744	5,077	3,169	1,708	6,549.25	2,924.5	0.047	2.239225777

Table S2. Primers for different procedures

Procedure	Forward	Reverse
To amplify 3.2 kb 5′ first arm	5′-TAG CGG CCG CCT AAC AGC CCA GAT CTC TTT GGG-3′	5′-TAG CGG CCG CCT AAC AGC CCA GAT CTC TTT GGG-3′
To amplify 1 kb conditional arm	5′-TAG ATA TCT CAG ACC CTG GCA AAA TGG GTG-3′	5′-TAG ATA TCG AAT GAG GAG CAC GGT GGA CC-3′
To amplify 2.7 kb 3′ third arm	5′ -TAA AGC TTC TGG CTT CAT ACA GAG GTC TAC T-3′	5′-GCT CTG GTT TTC TAC TAC TGC CA-3′
To amplify *DOCK1* genomic probe	5′-CCC ATC ACG TTC CAC CTT CTG TTT-3′	5′-CAG ACT CAG ATC CTC GAA CAG AAA GC-3′
To screen 5′ loxP site and to genotype *DOCK1* WT and Flx allele	P1 5′-TCA GCA GGC CCA GTT CCT ACT-3′	P2 5′-GCA GAG CTA GGA GTT CAT CGT AGT TC-3′
To screen 3′ loxP site	P4 5′-CAC CAA GCG AGA GGA GAA GTA CG-3	P3 5′-CCT ATC TAC AAC CCT TCA TTC CCA AGG-3′
To genotype *DOCK1* recombined allele	P1 5′-TCA GCA GGC CCA GTT CCT ACT-3	P3 5′-CCT ATC TAC AAC CCT TCA TTC CCA AGG-3′
To genotype *Cre* transgene	5′-GCT TCT GTC CGT TTG CCG-3	5′-ACT GTG TCC AGA CCA GGC-3′
To genotype Neu transgene	5′-TTC CGG AAC CCA CAT CAG GCC-3	5′-GTT TCC TGC AGC AGC CTA CGC-3′
To generate DOCK1$^{1228–1865}$	5′-AGG ATC CAA GGA GTG TGA TAA CTA CAC CGA AGC G-3′	5′-ACT CGA GTC ACT GCA CGA TCC CGG AG-3′
Q-PCR validation *ISG15* and RT-PCR	5′-TGA CTA ACT CCA TGA CGG TGT CAG-3′	5′-GAC CCA GAC TGG AAA GGG TAA G-3′
Q-PCR validation *OASL2*	5′-GAT GGA TAT CCT CCC AGC TTA CG-3′	5′-TTG GTG AGA AGT CAC CAG GGT AG-3′
Q-PCR validation *IFIT1*	5′-GGA GAA CAT GTT GAA GCA GAA GCA-3′	5′-CTG CAA GGC CCT GTT TAG AAG-3′
Q-PCR validation *IL23A*	5′-TCT CGG AAT CTC TGC ATG CTA G-3	5′-CTT GTG GGT CAC AAC CAT CTT CAC-3′
Q-PCR validation *DOCK1*	5′-GTC CAT GCT CCT GAA TGG CAT T-3′	5′-CAG GTC CTT GAG CTT CTC AAT CTG-3′
Q-PCR validation *CSN2*	5′-CCT TGC TCT TGC AAG AGA GAC-3′	5′-TGA ACT TTA GCC TGG AGC ACA TC-3′
Q-PCR validation *RAB17*	5′-TGC GCT CCT GGT TTA TGA CAT CA-3′	5′-GAT CCG TTT TGT TGC CGA CCA-3′
Q-PCR VALIDATION *IFI44*	5′-CCA CAC TCC TGA CAG ATA CCA-3′	5′-TGT CCT TCA GCA GTF GGT CAT-3′
RT-PCR *GBP3*	5′-CTA CAA CAG CAT GAG CAC CAT CAA CC-3′	5′-CTT CAG CTC CAG AGC AAA ATC TCG AAC-3′
RT-PCR *UBA7*	5′-CAT GGG ATC CTG ATG ATG CAG AGA CT-3′	5′-ATA ATG GCT GCC ATG GGG CTT AAG-3′
RT-PCR *STAT1*	5′-ACA GCT GGA CGA CCA GTA CA-3′	5′-TCC TGG GCC TGA TTA AAT CTT TGG G-3′
RT-PCR *RTP4*	5′-AAG TGG AGC CTG CAT TTG GA-3′	5′-CAT GTG GCA CAA GAT CAT CAC CTG AG-3′
RT-PCR *HERC6*	5′-CAT ATG TCT ACA CCA CTG GTC AGG TG-3′	5′- GAC TTC AAT ATC CAC AAG GTC CTC TGC-3′
RT-PCR *ITGβ1*	5′-GTC TTG GAA CGG ATT TGA TG-3′	5′-ACA AGG TGA GCA ATT GAA GG-3′
RT-PCR *ITGαv*	5′-ATG TTC ACA CTT TGG GCT GT-3′	5′-AGA GTT TCC TTC GCC ATT CT-3′
RT-PCR *β-ACTIN*	5′-TGA TGG TGG GAA TGG GTC AGA-3′	5′-TCC ATG TCG TCC CAG TTG GTA A-3′