Supplementary methods

Generation of mammary specific Nmi-floxed mice: $Nmi^{fl/fl}$ mice were custom created by GenOway (http://www.genoway.com/, France) on a mixed 129SV/C57BL/6 background using 129SV-derived ES cells. PCR amplification of the murine Nmi locus from 129sv genomic DNA was performed to create the targeting vector. Homologous recombination of the targeting vector in ES cells inserted loxp sites flanking exons 8, 9, and the polyadenylation sequence. Recombined clones were selected using a neomycin cassette which was subsequently removed by breeding to a Flp deleter mouse. Integration of the construct into ES cells as well as chimeras was confirmed via PCR and southern blot analysis of genomic DNA using the following primers: 5'AGCAATGGTTATGAGTTCTCCAAGGGTG3',5'TGCAATGAAAGTATTTCAAGTGGTAAGG GG-3'.

Mice were backcrossed onto C57BL/6 and FVB/N backgrounds simultaneously using microsatellite assisted selection (performed by Jackson Labs, Bar Harbor, Maine) for more than 5 generations to ensure greater than 95% desired background before crossing to congenic K14-Cre and MMTV-Neu mice that were utilized for downstream experiments.

Genotyping: DNA was extracted from tail fragments using the REDExtract-N-Amp™ Tissue PCR Kit (Sigma Aldrich, St. Louis, MO), and analyzed with PCR using the following primer sets: Nmi floxed locus (5'AGCAATGGTTATGAGTTCTCCAAGGGTG3' and 5'TGCAATGAAAGTATTTCAAGTGGTAAGGGG3'), K14-cre transgene (5'GCG GTC TGG CAG TAA AAA CTA TC3' and 5'GTG AAA CAG CAT TGC TGT CAC TT3'), and the MMTV- Neu transgene (5'TTT CCT GCA GCA GCC TAC GC3' and 5'CGG AAC CCA

CAT CAG GCC3').

BAT-gal 5' CGTGGCCTGATTCATTCC 3' and 5'ATCCTCTGCATGGTCAGGTC3'

Nmi Antibody Generation: Polyclonal Nmi antibodies were generated by Thermo Scientific Pierce (Rockford, IL). Briefly, two peptides were designed against the N terminal sequence of Nmi corresponding to amino acids 13-28 (DERSAEMDDMRGEQSM) or amino acids 63-80 (ENVPEKKLKLTSVESPKD). Two rabbits were immunized for each of the corresponding KLH-conjugated peptides for the course of a 120 day protocol, and sera was collected at day 98,100 and tested for reactivity and specificity via both western blot as well as IHC. Upon determination of specific antibody production two individual rabbits from each peptide immunization were further extended to 140 and 196 days. Animals were euthanized following the second extension and serum was collected. Upon validation of antibody production from the extended bleeds further purification was done by affinity column and further validated via ELISA. Purified antibodies were stored in BSA at a final concentration of 3% and used for further downstream applications. PA8076 (aa-63:80) was used for western blot and PA7977 (aa-13:28) was used for IHC.

Isolation and enrichment of primary cells: Second and third mammary glands were isolated from mice on the first day of lactation, minced and dissociated in digestion medium (HBSS containing collagenase I (Sigma Aldrich, St. Louis, MO) (1mg/mL) and Pronase (Sigma Aldrich) (0.1mg/mL)) for two hours at 37^0C with shaking. Epithelial organoids were washed in PBS and enriched by pulse centrifugation to 1500rpm at least three times before subsequent assays. To isolate cancer cells, tumors were excised and minced before incubating in digestion medium at 37^0C Primary tumor

cells were grown in DMEM/F12 (phenol red free) media containing 10% FBS, EGF (20ng/mL), hydrocortisone (0.5µg/mL), cholera toxin (100 ng/mL), bovine insulin (10µg/mL), and Pen Strep.

RNA Isolation and Real Time PCR: Total RNA was harvested from mammary epithelial organoids with TRIzol® reagent (Invitrogen, Carlsbad, CA) and purified using the RNEasy kit (Qiagen Inc., Valencia, CA). RNA was quantitated and assessed using spectrophotometry (NanoDrop Lite, Thermo Scientific, Wilmington, DE). Total RNA (1 µg) was transcribed into complementary DNA using the High Capacity cDNA Reverse Transcription kit (Applied Biosystems, Foster City, CA). Quantitative real-time PCR was performed with TaqMan® Fast Advanced Master Mix (Applied Biosystems) in a Stepone Plus (Bio-Rad Laboratories) real- time PCR detection system. TaqMan® Gene Expression Assays (Life Technologies) were used for the following genes: Nmi (Mm01324653_m1), l3-actin (Mm00607939_s1). Data collected was normalized to the endorsed control gene, l3-actin, and expressed as fold change in relative expression using the LiLiC_t method (2^{-LiCt}).

Next-Generation Sequencing on Illumina Platforms.

mRNA-sequencing was performed on the Illumina HiSeq2500 using the latest versions of the sequencing reagents and flow cells. Briefly, the quality of the total RNA was assessed using the Agilent 2100 Bioanalyzer. Total RNA with a RIN of 27.0 was used for the RNA library prep that started with 2 rounds of poly A+ selection to isolate the mRNA. The SureSelect Strand-Specific mRNA library generation kit was used as per the manufacturer's instructions (Agilent, Santa Clara, CA) to generate the sequence ready libraries. The cDNA libraries were quantitated using qPCR in a Roche LightCycler 480 with the Kapa Biosystems kit for library quantitation (Kapa Biosystems, Woburn, MA) prior to cluster generation. mRNA- Sequencing was done on the Illumina

HiSeq2500 with onboard clustering by loading 11pM of the pooled libraries and paired end 50bp sequencing by standard techniques (Illumina, Inc., San Diego CA). Approximately 30 million reads were generated per sample.

Cell lines: HC11 cells were obtained from ATCC (CRL-3062) and maintained in RPMI with 10% FBS, 20ng/ml EGF and 5µg /mL insulin. To differentiate HC11 cells, insulin was removed from the media 24 hours prior to the addition of media containing 100nM dexamethasone (Tocris), insulin (5µg/mL), and ovine prolactin(5µg/mL) (Sigma). Nmi was silenced in HC11 cells using SMARTvector lentiviral mouse Nmi mCMV-TurboRFO shRNA plasmids (Dharmacon, Lafayette, CO). Cells were transfected using Lipofectamine 2000 and selected with puromycin (500ng/mL) prior to FACS sorting to enrich a mixed population for RFP positive cells (BD FACS Aria). Mycoplasma-free cell lines were used in all of our experiments.

Image Analysis

ImageJ was used to quantify whole mounts, and histological sections. NIS elements basic research software was used to calculate the invaded area of the ductal epithelial tree during pubertal development. The immunoscore was derived as the product of percentage of cells at each intensity. (0 (no staining) to 4 (strongest possible intensity of staining) and the corresponding intensity. The products were added to get an immunoscore for the section.