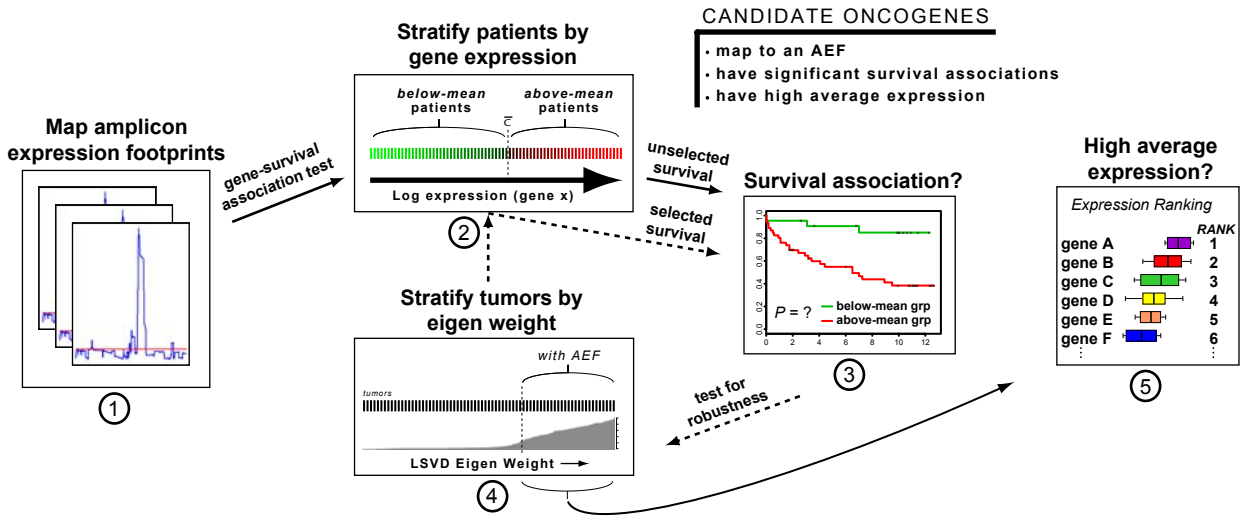
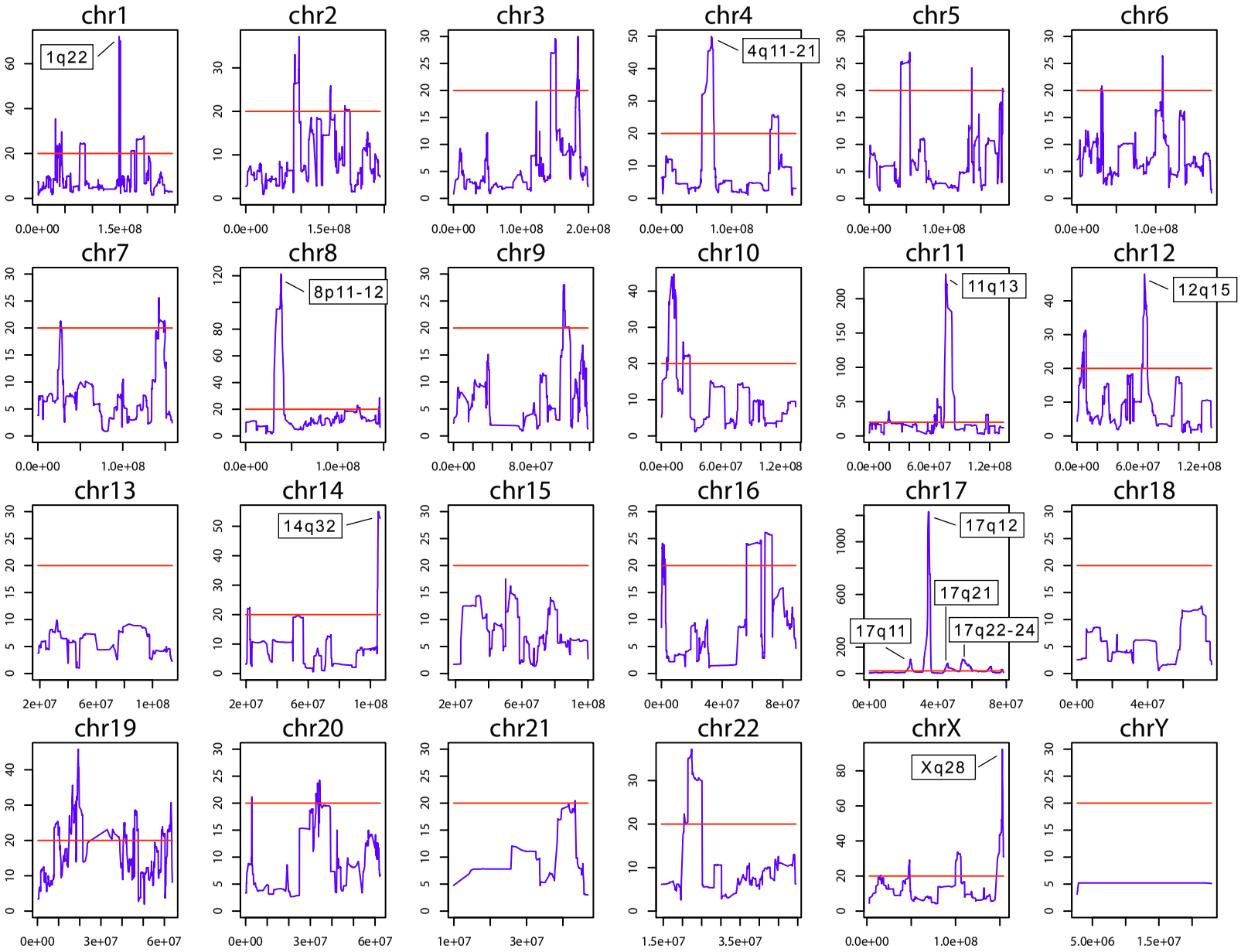
SUPPLEMENTAL FIGURE 1.



TRIAGE steps:	*TRIAGE steps:*
Correlating amplicon genes with distant metastasis-free survival 1. *Amplicon expression footprints (AEFs)* are inferred by genome-wide LSVD; principal eigen peaks (PEPs) that overlap with chromosomal regions of known recurrent amplification are further investigated. 2. Patients are stratified by each amplicon gene into below-mean and above-mean expression groups. 3. For each gene, Cox proportional hazards regression is performed (using distant metastasis-free survival as the endpoint) and the significance of the hazards ratio (p-value) is reported. **Testing for robustness; discriminating oncogenes from passenger genes** 4. Tumors are ranked by eigen weight to estimate the amplicon-containing tumors. Steps 2 & 3 are repeated excluding those patients with amplicon-containing tumors. 5. Genes within the amplicon are ranked by their average expression level in the amplicon-containing tumors.	**RATIONALE** 1. Expression footprints of recurrent genomic amplicons can be detected by LSVD. 2. & 3. Some metastasis-promoting oncogenes that are activated by amplification have expression patterns that correlate with poor DMFS in breast cancer patients, reflecting their functional involvement in metastatic progression. 4. Passenger genes located on survival-associated amplicons may correlate with poor DMFS without functional involvement. Oncogenes activated by multiple cancer mechanisms not limited to amplification should be associated with poor outcome in the absence of the amplicon, whereas passenger genes should not. 5. The oncogenic driver(s) of an amplicon should be relatively highly and consistently expressed in amplicon-containing tumors.

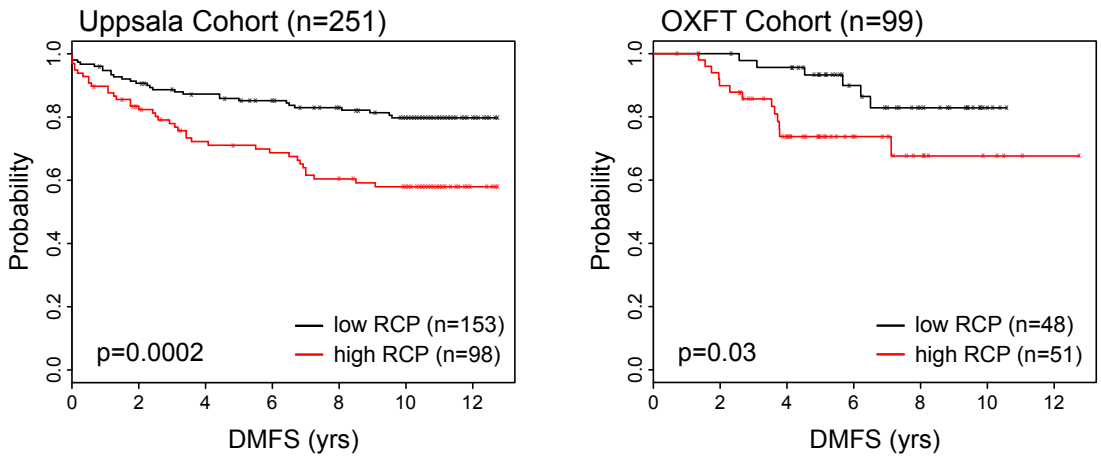
Supplemental Figure 1. Overview of the TRIAGE Methodology. Multi-step schematic of the TRIAGE procedure is shown.

SUPPLEMENTAL FIGURE 2.



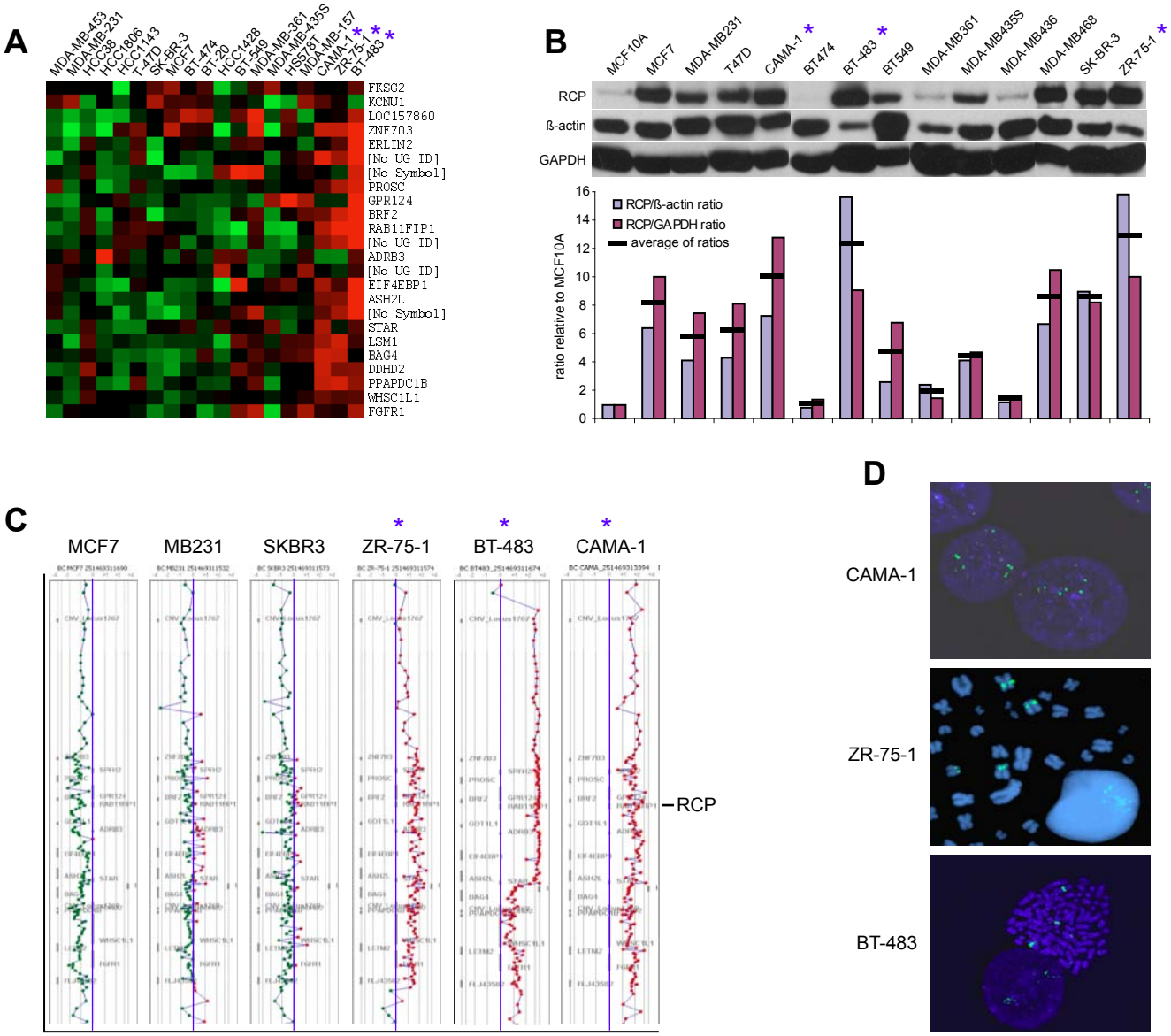
Supplemental Figure 2. LSVD analysis of the breast cancer oncogenome. Eigen values (y-axis) are plotted along the length of each chromosome (x-axis). The genomic locations of the highest-scoring principle eigen

SUPPLEMENTAL FIGURE 3.



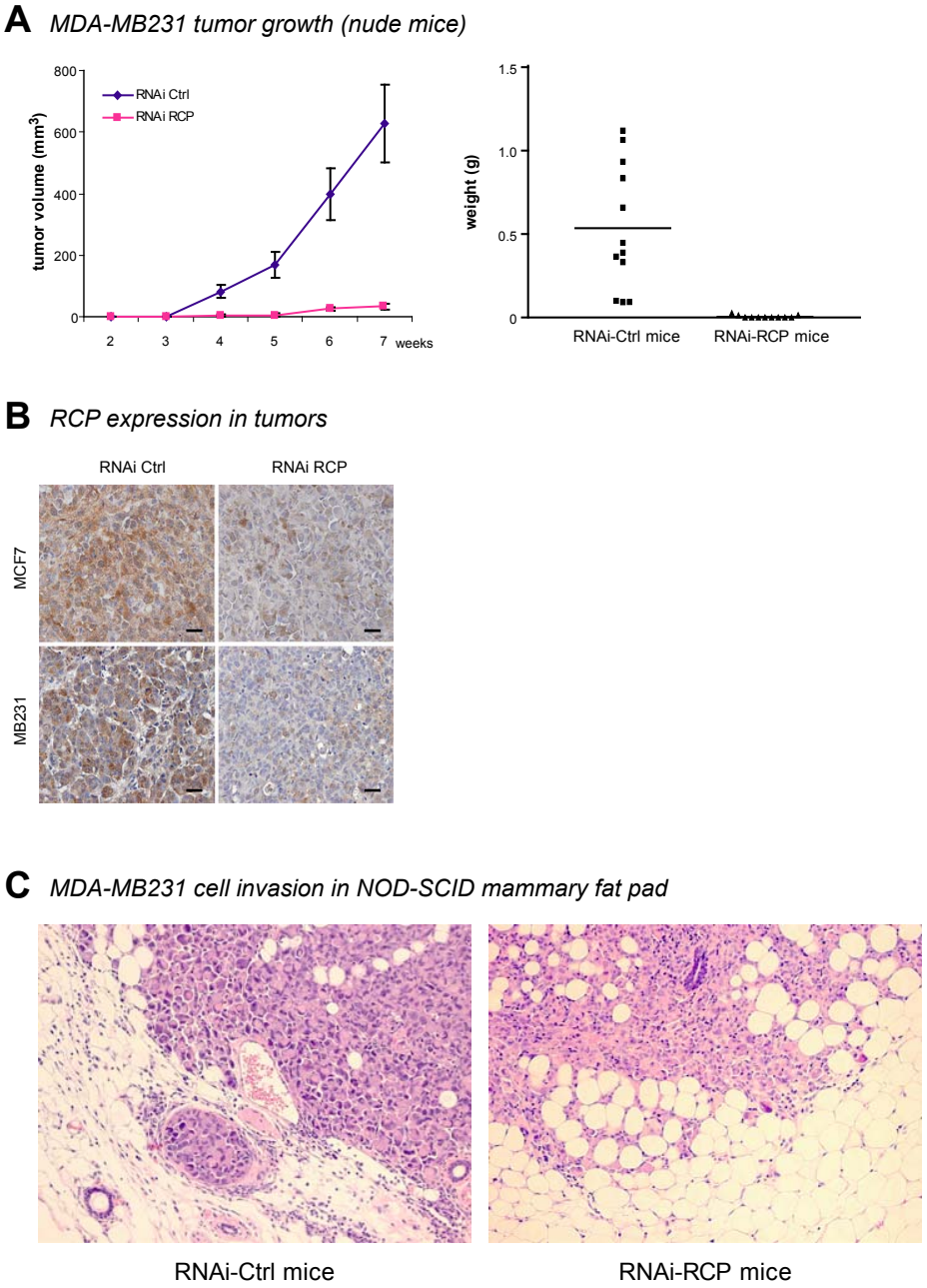
Supplemental Figure 3. RAB11FIP1/RCP expression is prognostic of metastatic recurrence. Breast cancer patients of the Uppsala (Miller, et. al., 2005) and OXFT (Loi, et. al., 2007) cohorts were ranked according to RCP expression levels (Affymetrix probe set 219681_s_at). Distant metastasis-free survival of patients (DMFS) with below-mean expression was compared to that of patients with above-mean expression by Kaplan-Meier (KM) analysis. The likelihood ratio test p-value reflects the significance of the hazard ratio.

SUPPLEMENTAL FIGURE 4.



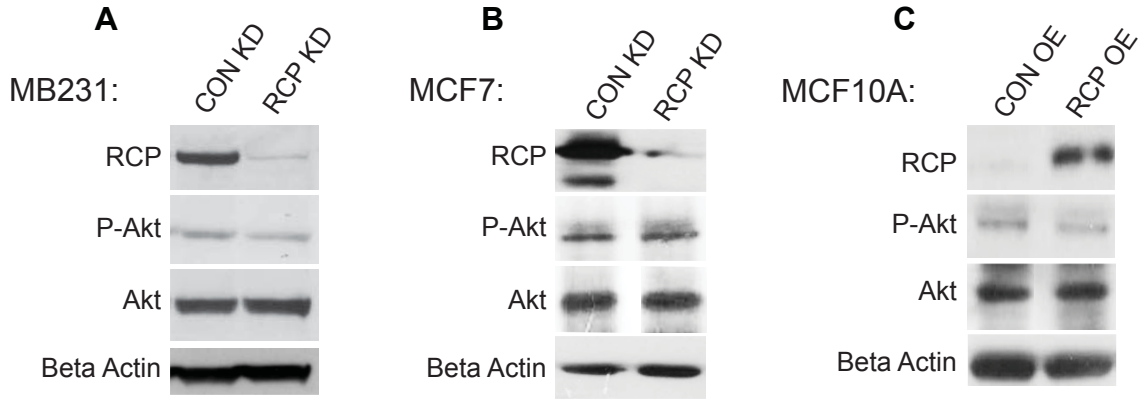
Supplemental Figure 4. Molecular analysis of RCP amplification and expression in breast cancer cell lines. (A) Expression profiles of 19 breast cancer cell lines (GEO accession number GSE3156; Bild et al., 2006) were analyzed by LSVD at the 8p11-12 region and are shown ranked (left to right) by ascending absolute value of tumor eigen weight. (B) Protein levels of RCP are shown by Western blot (upper panel) and plotted as ratios (from densitometric measurements) with β-actin and GAPDH (lower panel) for a panel of breast cancer cell lines. (C) Copy number analysis by array-CGH. 8p11-12 genomic profiles are shown for 6 cell lines; red points to the far right of the center axis indicate copy number gain. The genomic location of RCP is indicated to the right. (D) FISH analysis using a BAC probe spanning the RCP gene locus (BAC clone ID: RP11 933I10).

SUPPLEMENTAL FIGURE 5.



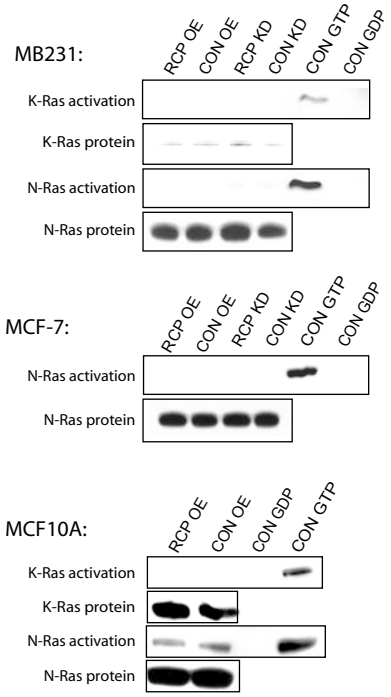
Supplemental Figure 5. RCP in tumor xenografts. (A) Effect of RCP inhibition on MB231 tumor growth in nude mice. Left panel, mean tumor volume plotted as a function of time (mean +/- s.e.m.). Right panel, tumor weight plotted at 5 weeks; mean weight indicated by solid line. (B) Immunohistochemical staining for RCP in tumor samples. (Scale bars, 100 μm.) (C) Invasion of MB231-derived tumor epithelium into adjacent mammary tissue. Representative sections of tumor-fat pad margins in RNAi-Ctrl and RNAi-RCP NOD-SCID mice are shown.

SUPPLEMENTAL FIGURE 6.



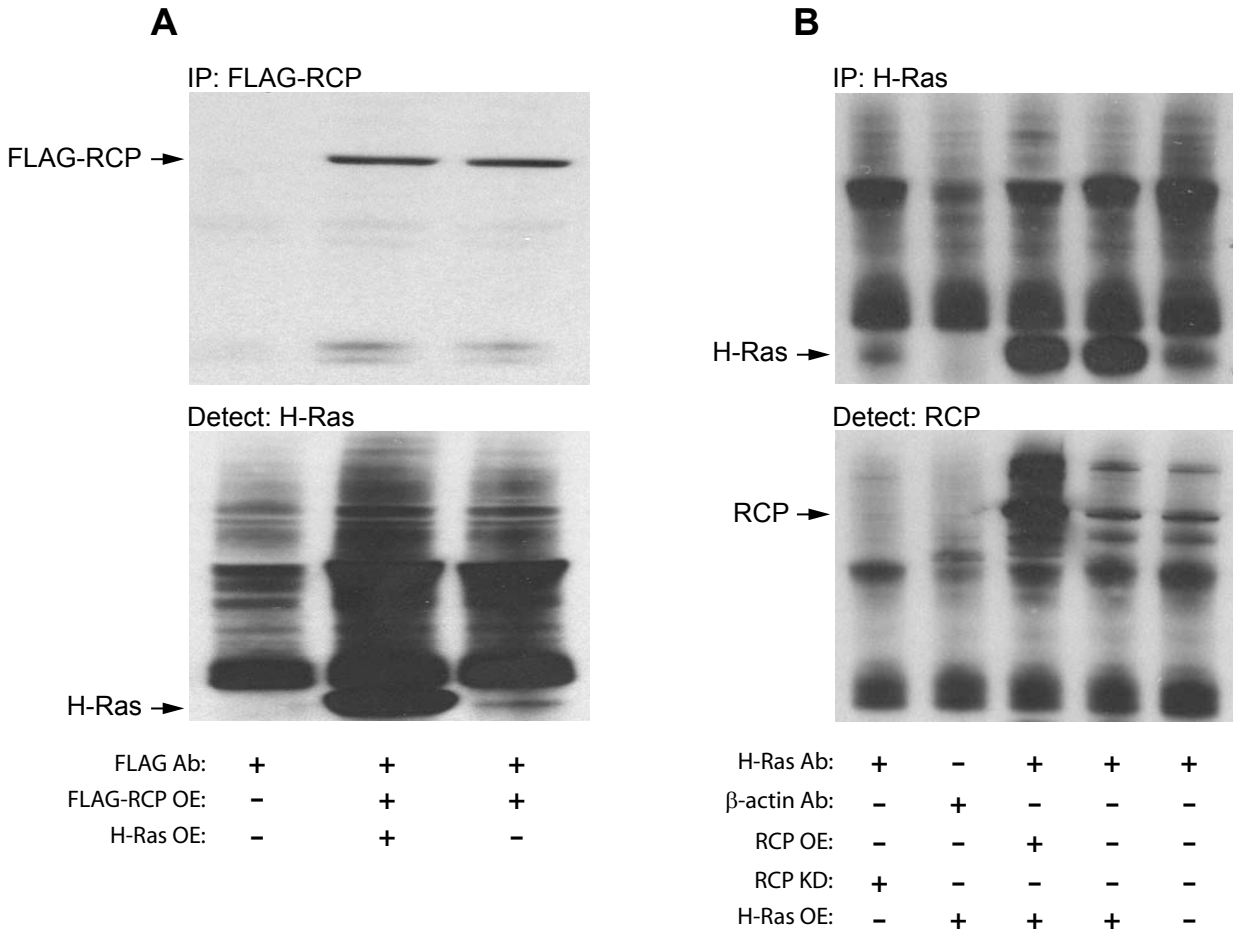
Supplemental Figure 6. AKT phosphorylation is independent of RCP. Akt phosphorylation status (P-Akt), total Akt protein and RCP protein levels in (A) MB231 cells, (B) MCF7 cells and (C) MCF10A cells are shown by Western Blot. Abbreviations: RCP knock down (RCP KD), RCP overexpression (RCP OE) and controls (CON KD, CON OE).

SUPPLEMENTAL FIGURE 7.



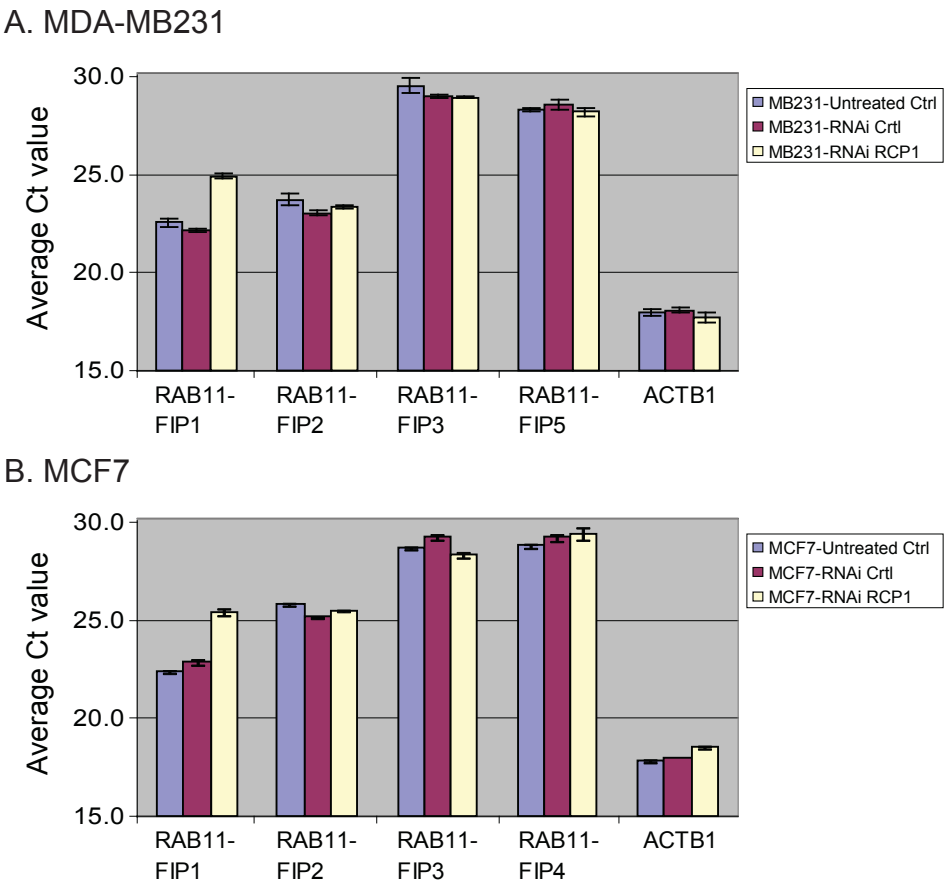
Supplemental Figure 7. N-RAS and K-RAS activation status in cell lines. Shown is the activation status and total protein levels of N-RAS and K-RAS in MB231, MCF7 and MCF10A cells in the context of RCP overexpression, RCP knock down and controls.

SUPPLEMENTAL FIGURE 8.



Supplemental Figure 8. RCP/H-RAS co-immunoprecipitation experiments. Constructs expressing RCP, FLAG-RCP, and wildtype H-RAS were co-transfected in various combinations into MCF7 cells. (A) Immuno-precipitated (IP) FLAG-RCP pulls down exogenously expressed H-RAS (lane 2) and endogenous H-RAS (lane 3). (B) Immuno-precipitated H-RAS (exogenous) pulls down exogenously expressed RCP (lane 3) and endogenous RCP (lane 4). Immuno-precipitated endogenous H-RAS pulls down endogenous RCP (lane 5). Ab=antibody; OE=overexpressed; KD=RNAi knock down.

SUPPLEMENTAL FIGURE 9.



Supplemental Figure 9. RCP knock down attenuates transcript levels of RAB11FIP1/RCP but not other RAB11FIP family members. (A) Real-time PCR Ct values are shown for parental MDA-MB231 cells (Untreated Ctrl) or lentivirus-infected MDA-MB231 cells stably expressing scrambled shRNA sequence (RNAi Ctrl) or shRNA targeted to RCP (RNAi RCP1). **(B)** Real-time PCR Ct values are shown for parental MCF7 cells (Untreated Ctrl) or MCF7 cells transiently transfected with scrambled RNAi sequence (RNAi Ctrl) or RNAi targeted to RCP (RNAi RCP1). Transfections were conducted in duplicate, and each transfection analyzed in triplicate by PCR (for 6 PCR measurements per condition; error bars computed from mean +/- s.d.). At 48 hrs post-transfection, RNA was extracted from cells using TRIzol reagent (Invitrogen), followed by purification using the RNeasy mini kit (Qiagen). Two ug total RNA was reverse transcribed using the SuperScriptIII RNA Amplification System (Invitrogen) according to the manufacturer's protocol, and 100 ng cDNA was used as PCR template. Real-time PCR was performed using the ABI 7900HT Fast Real Time PCR System (Applied Biosystems) in 384 well plate format. Pre-constructed primers and probes were purchase from Applied Biosystems' inventoried TaqMan assays for RAB11FIP-1, -2, -3, -4, -5 and beta-actin (Assay IDs: HS00368787_ml, Hs00208593_ml, Hs00608512_ml, Hs00400200_ml, Hs00392033_ml, and Hs00357333_gl (beta-actin), respectively). All RAB11FIP family member transcripts were detected with the exception of RAB11FIP4 (in MDA-MB231 cells) and RAB11FIP5 (in MCF7 cells) which could not be detected (ie, Ct values >35) across a range of starting cDNA amounts (10 ng to 300 ng; data not shown). RAB11FIP1/RCP was the only transcript observed to be consistently and significantly decreased (a gain of ~2.5 Ct increments) when transfected with RNAi RCP1.

SUPPLEMENTAL METHODS (Zhang, et. al)

Local Singular Value Decomposition (LSVD) to identify amplicon expression footprints (AEFs). LSVD analysis of gene expression profiles from a cohort of tumor samples is a key first step in identifying *amplicon expression footprints* (AEFs), from which oncogenes contributing to metastasis may be discovered by the TRIAGE methodology (Triangulating Oncogenes through Clinico-Genomic Intersects). LSVD identifies AEFs based on the hypothesis that a recurrent amplicon will manifest as the coordinated overexpression of genomically-localized genes in a subset of tumor samples. LSVD is built around *singular value decomposition* (SVD) [1] and predicts AEFs as peaks in the coordinated local overexpression plot which is a *local principal eigen value (LPEV) plot* obtained by the local application of SVD on a chromosome-by-chromosome basis. LSVD is achieved by deriving local expression matrices for each genomic position on the genome, transforming it to a binary connectivity matrix and applying SVD on the binary matrix. The principal eigen value obtained upon such application is a measure of the coordinated overexpression of genes at that locus [2,3]. Eigen value "peaks" that exceed a certain basal threshold in the plot of LPEV vs. genomic location (LPEV plot) demarcate the AEFs. The detailed description of the LSVD procedure is provided below.

Let $\boldsymbol{E_c}$ be the gene expression matrix of chromosome C_c; $\boldsymbol{E_{cij}}$ is the $\log_2$ of the expression of gene at location l_i in tumor T_j, where $i = 1, 2, \ldots, N_c$ and $j = 1, 2, \ldots, M$. *AEFs* on C_c are predicted by analyzing $\boldsymbol{E_c}$ using LSVD which has the following sequence of steps: (1) Transforming $\boldsymbol{E_c}$ to binary connectivity matrix $\boldsymbol{A_c}$; (2) Deriving Chromosome Localized Matrices, $\boldsymbol{Ap_c}$; (3) SVD on $\boldsymbol{Ap_c}$; and, (4) Identifying AEFs.

1. Transforming E_c to binary connectivity matrix A_c: $\boldsymbol{E_c}$ is transformed to a binary connectivity matrix $\boldsymbol{A_c}$ through a boolean descretization rule as follows:

$$A_{cij} = 1 \text{ if } E_{cij} > \mu_i + x*v_i$$

$$A_{cij} = 0 \text{ if } E_{cij} \leq \mu_i + x*v_i$$

where μ_i and v_i are the median and adjusted Median Absolute Deviation (aMAD) of $E_{ci} = \{E_{ci1}, E_{ci2}, \ldots, E_{ciM}\}$. aMAD is 1.4826 times the MAD. x is 2.5 for the first pass of coarse map and the 2 for the finer map obtained in the second pass.

2. Deriving Localized Matrices, Ap_c: A local matrix $\boldsymbol{Ap_c}$ at position Lp_c is derived from $\boldsymbol{A_c}$ using genes at positions from p-w to p+w on chromosome C_c where w is the window size, a predefined parameter. w is 50 for the first pass coarse map and 10 for the second pass finer map.

3. Singular Value Decomposition (SVD): The application of SVD on $\boldsymbol{Ap_c}$ is a crucial step to predict AEFs. The SVD of Ap_c decomposes $\boldsymbol{Ap_c}$ into a product of three matrices $\boldsymbol{Up_c}$, $\boldsymbol{\Sigma p_c}$ and $\boldsymbol{Vp_c}$ i.e. $\boldsymbol{Ap_c} = \boldsymbol{Up_c} \times \boldsymbol{\Sigma p_c} \times \boldsymbol{Vp_c}^T$. $\boldsymbol{Up_c}$ and $\boldsymbol{Vp_c}$ are of $(2w+1)$x$(2w+1)$ and MxM matrices respectively while $\boldsymbol{\Sigma p_c}$ is $(2w+1)$xM diagnol matrix. The column vectors of $\boldsymbol{Up_c}$ and $\boldsymbol{Vp_c}$ are the eigen vectors of $\boldsymbol{Ap_c Ap_c}^T$ and $\boldsymbol{Ap_c}^T\boldsymbol{Ap_c}$ respectively while the diagonal elements of $\boldsymbol{\Sigma p_c}$ are the respective eigen values. The highest eigen value is called the principal eigen value (denoted by λp_c) and the respective eigen vectors are called the principal eigen vectors. *Eigen weight of tumor* T_j at position Lp_c is denoted by Tp_{cj} and it defined as the absolute of the jth component of the principal eigen vector of the matrix $\boldsymbol{Ap_c}^T\boldsymbol{Ap_c}$. Similarly, *eigen weight of a gene* G_i at position Lp_c, denoted by Gp_{cj} is the absolute of the jth component of the principal eigen vector of the matrix $\boldsymbol{Ap_c Ap_c}^T$.

To improve the contrast between true AEF signal and background, in our LSVD implementation, we use SVD on $(\boldsymbol{Ap_c Ap_c}^T)^3$ instead of on $\boldsymbol{Ap_c}$ to find λp_c. The eigen weights of tumors and genes are obtained, following the above definitions, from the principal eigen vectors of $(\boldsymbol{Ap_c}^T\boldsymbol{Ap_c})^4$ and $(\boldsymbol{Ap_c Ap_c}^T)^4$ respectively. Now, λp_c is the ratio of the fourth root of the principal eigen value of $(\boldsymbol{Ap_c Ap_c}^T)^4$ to M^4; the eigen weights of genes and tumors are the fourth root of the respective components of the principal eigen vectors of $(\boldsymbol{Ap_c Ap_c}^T)^4/M^4$ and $(\boldsymbol{Ap_c}^T\boldsymbol{Ap_c})^4/M^4$ respectively.

4. Inferring AEFs: Plotting λp_c vs. Lp_c gives the plot of LPEVs, this is called *LPEV plot*. The peaks in this plot show the AEFs. Higher the value of λp_c at the peak, higher the confidence of AEF at around Lp_c. Finally, an AEF is reported at a peak with range that covers a contiguous region with $\lambda p_c > T_\lambda = 20$ and centered around the peak.

References:

1. Strang G (1998). "Introduction to Linear Algebra". Section 6.7. 3rd ed., Wellesley-Cambridge Press. ISBN 0-9614088-5-5.
2. Kluger Y., Basri R., Chang JT, and Gerstein M., "Spectral Biclustering of Microarray Data: Coclustering Genes and Conditions", Genome Research, 13:703-716, 2003.
3. Kleinberg JM, "Authoritative Sources in a Hyperlinked Environment", Jl of the ACM, 46(5):604-632, 1999.

Cell lines and transient transfection. Human mammary and breast cancer cell lines including MCF10A, MCF7 and MDA-MB231 were obtained from American Type Culture Collection (ATCC) and maintained at 37°C with 5% CO¬2 with growth medium recommended by ATCC. Full-length RCP was ligated into the BamH1 and EcoR1 site of pcDNA3.1. Two RCP-specific RNAi constructs used in this study were designed as previously described (59) and manufactured by Ambion (RNAi RCP1: CGAUAAGCAAGAAGGAGUU) and Qiagen (RNAi RCP2: GGAAGGACUUUCCUU
UCUU). Transient plasmid and RNAi transfection on MCF7 was carried out using Lipofectamine 2000 (Invitrogen) according to the manufacturer's protocol. The specificity of RNAi RCP1 inhibitory effects (with respect to RCP family members RAB11FIP-2, -3, -4, and -

5) was confirmed by real-time PCR in MDA-MB231 and MCF7 cell lines (Supplemental Figure 9).

Generation of stable cell lines by lentiviral infection. For the generation of cell lines stably overexpressing shRNA, oligonucleotides encoding the target sequence (forward: 5'-TCGATAAGCAAGAAGGAGTTTTCAAgAgAAACTCCTTCTTGCTTATCG TTTTTTC-3', and reverse: 5'-TCGAGAAAAAACGATAAGCAAGAAGGAGTTTCT CTTGAAAACTCCTTCTTGCTTATCGA-3') were annealed and cloned into the LentiLox pLL3.7 vector. For control shRNA, a non-targeting "scramble" sequence was cloned into pLL3.7. (forward: 5'-TGAACGGCATCAAGGTGAAC ttcaagagaGTTCACCTTGATGCCGTTCTTTTTTC-3', and reverse: 5'-TCGAGAAAA AAGAACGGCATCAAGGTGAACTCTCTTGAAGTTCACCTTGATGCCGTTCA-3'). For overexpression of RCP by lentivirus, RCP was first cloned into the Invitrogen gateway entry vector pENTR3C and further recombined into the lentivirus vector, pLenti6/V5-DEST (Invitrogen) by LR recombination reaction according to the manufacturer's protocol. For lentivirus production, pLL3.7 or pLenti6/V5-DEST was co-transfected with packaging vectors into 293FT cells and the supernatant was harvested after 48 hours. Virus was concentrated by ultracentrifugation for 2 hours at 25,000 rpm in a Beckman SW28 rotor and resuspended in phosphate-buffered saline. Titers were determined by infecting NIH/3T3 cells with a serial dilution of the concentrated virus. For a typical preparation, the titer was approximately $1\text{-}5\times10^8$ particles for pLL3.7 and $4\text{-}10\times10^7$ particles for pLenti6/V5-DEST vector. For infection of MCF10A, MCF7 and MB231 cells, 2×10^5 cells were incubated in suspension with 1x107 particles and 8ug/ml polybrene for 3 hours at 37°C. The cells were then re-plated and cultured as previously described. To obtain a pure population of RCP knock-down stable cells, GFP-positive cells were sorted after 4 days of culture. For generating RCP overexpressing stable lines, cells were selected with blastacidin for 2 weeks after infection.

Western blot and immunohistochemical analysis. Protein lysates were prepared using RIPA buffer. The proteins were separated by SDS-PAGE and transferred to Hybond-P PVDF membrane (GE Healthcare). Antibodies to RCP (Genway, Cat#: 15-288-21574A), β-actin (Sigma, Cat#: A5441), phospho-Akt (Cell Signaling Technology, Cat#: 9271), Akt (Cell Signaling Technology, Cat#: 9272), phospho-Erk (Cell Signaling Technology, Cat#: 9101), and Erk (Cell Signaling Technology, Cat#: 9102) were used to probe the membrane, and antibody-protein complex was detected by HRP-conjugated antibodies and ECL (Amersham Biosciences). For ERK/AKT activation experiments, 0.5×10^6 cells were plated in 6-well plates in DMEM containing 10% serum for 48 hrs. Cells were then starved in 0.5% serum/DMEM for 2 hrs, then the culture medium was replaced with the original medium for 1 hour to initiate ERK/AKT activation. IHC was performed on MaxArray Human Normal Tissue Microarray Slides (Invitrogen, Cat#: 75-4013, Lot #381086A) and MaxArray Human Breast Carcinoma Tissue Microarray Slides (Invitrogen, Cat#: 75-4043, Lot #473803A) according to the manufacturer's instructions, but with the following modifications. Blocking was carried out with diluted horse serum (Vectastain ABC kit, Vector Laboratories), and slides were incubated with anti-RCP (Genway) at 1:50 dilution at 4^oC overnight. Images of stained tissues were acquired with Nikon Eclipse 90i automated upright microscope under 40× objective.

Cell proliferation and colony formation. Cells were plated at a density of 5000 cells/well in 96-well plates, and cell proliferation was measured in quadruplicate (ie, 4 wells per condition) using WST-1 (Roche) according to the manufacturer's protocol. For the MAPK inhibition assays, cells were treated with MEK inhibitor U0126 (Promega) overnight in medium with 0.5% serum before being re-plated into 96-well plates. To test the effect of RCP on anchorage-independent colony formation, cells were suspended in 250 L of 0.3% agar (Sigma) dissolved in complete medium containing 25% FBS, and plated in quadruplicate (ie, 4 wells per condition) in 24-well plates pre-coated with 500 L of 0.6% agar base. Colony forming efficiency was examined 21 days or more after plating by staining with Iodonitrotetrazolium chloride (Sigma). Colonies of size >50 m were counted using Leica QWin software.

In vitro invasion and migration. Transwell migration and invasion assays were performed using Falcon FluoroBlok 24-Multiwell inserts (BD Biosciences) with 8µm pores. For invasion assays, the inserts were coated with 20µg Matrigel (BD Biosciences) in 80µL serum-free growth medium. For both assays, 5000 cells in 200 L serum-free growth medium were loaded into each transwell insert with 750 L complete growth medium with 10% fetal bovine serum in the lower chamber. At 24 hours, cells that had migrated or invaded through the pores of the inserts were fixed with 3.7% formaldehyde, stained with 2.5 g/mL Hoechst 33342 (Invitrogen) for 15 minutes, washed with PBS and counted using the Target Activation Bioapplication on an ArrayScan VTI (Cellomics). Field size was $1mm^2$. For invasion assays, experiments were performed with 4- to 5-fold replication, with 10-16 fields scanned per experiment. For migration assays, experiments were performed with 3-fold replication, with 10-16 fields scanned per experiment.

Immunofluorescence microscopy. Cell cultures were fixed in 4% paraformaldehyde and permeabilized with 0.25% Triton X-100, followed by blocking with 1% goat serum in PBS. Cells were stained with anti E-Cadherin (BD Pharmingen, CN: 610404) and fibronectin antibody (BD Pharmingen, CN: 610077) at 1:500 dilution, followed by the appropriate secondary antibodies detecting mouse and rabbit IgG conjugated with Alexa Fluor 594 or 488 (Molecular Probes), respectively. F-actin was labeled with Texas Red-X phalloidin (Molecular Probes). DAPI was used to stain the nucleus. Images were captured with a confocal microscope (LSM 510 META, Zeiss).

Cell cycle analysis. MCF10A cells were starved in 1% serum for 48 hours. Cells were then fixed in 75% ethanol, treated with RNase A (0.25 mg/ml), stained with propidium iodide (10 g/ml) and analyzed on a LDR2 flow cytometer (BD Biosciences). Sorted cells was analyzed using FACS DiVa software (BD Biosciences).

Co-immunoprecipitation assays. MCF10A, MCF7 and MB231 cells were co-transfected with pCMV-H-RAS (Clontech), pcDNA3.1-RCP, pcDNA3.1-FLAG-RCP, and empty vectors, alone or in various combinations (see Supplemental Figure 5). At 48 hours post-transfection, cells were lysed in buffer (50mM Tris pH 7.4, 150mM 1Mm EDTA, 1% Triton X-100, sodium orthovanadate and protease inhibitor mix [Roche Applied Science]) and clarified for particulate-free lysates. Briefly, 60uL Protein G beads (Sigma Aldrich) were incubated with 3ug of antibody and 1ml lysate overnight at 4^oC. After washing beads 5x in buffer (50mM Tris7.4, 150mM NaCl,

1mM EDTA, and 0.1% TritonX-100), samples were subjected to SDS-PAGE and Western blot using H-RAS, FLAG and RCP antibodies as previously described.