Supplementary Figure legends

Supplementary Figure S1. Correlation between *TLOC1* suppression and cell proliferation. **(A)** Effects of *TLOC1*-specific shRNAs on *TLOC1* mRNA levels. The effects of expressing two *TLOC1*-specific shRNAs (shTLOC1_1591 and shTLOC1_1267) on *TLOC1* mRNA were assessed in MCF7 cells by quantitative real time PCR. A GFP-targeting shRNA was used as a control. (**B**) Effects of *TLOC1*-specific shRNAs on cell proliferation in the 11 cell lines that harbor 3q26 copy number gain ($p<0.005$ and $p<10^{-6}$, student's t-test).

Supplementary Figure S2. *SKIL* induces Matrigel invasion. *SKIL* induced invasion from the invasion screen shown in Figure 2D was confirmed in a separate experiment. SKIL overexpression induced Matrigel invasion in HMLE-MEKDD cells ($p=0.02$, student's t-test). Suppression of *CDH1* (E-cadherin) was used as positive control. Bars indicate the number of invades cells per vision field ± SD.

Supplementary Figure S3. 3q26 copy number in cancer cell lines. Genomic copy number at 3q26 was determined in six cell lines by quantitative real time PCR on genomic DNA. Human mammary epithelial cells (HMEC) were used as a reference for normal copy number levels of 3q26. MDAMB231 was diploid at 3q26. H28 and COLO320 were found to display moderate amplification levels as compared to HCC95 and T47D. The genomic quantitative real time PCR analysis was performed on genomic DNA with primers specific for 3q26 at the base position 169,940,153. Primers specific for LINE-1 elements were used to normalize for genomic DNA input. Bars indicate the relative copy number as compared to the HMEC cells ± SD.

Supplementary Table S1. Gene suppression induced by shRNAs determined by quantitative real time PCR in MCF7 cells. The relative expression levels of targeted genes were determined after suppression by the corresponding shRNAs as compared to control (suppression of GFP). The levels were determined by quantitative real time PCR from cDNA prepared from infected MCF7 cells. For shRNAs where suppression levels were not readily detected the values were set to 1.

Supplementary Table S2. Gene list derived from a gene expression based comparative marker selection between cell lines with normal or amplified levels of 3q26. The gene list is arranged according to the genes that are the most highly expressed in cell lines with 3q26 amplification, which is denoted by 3q26amp in the "Upregulated in" column. Genes that reside in the minimal common amplified region of 3q26 are marked with bold text. The "Chr" column shows on what chromosomes the genes reside. #N/A designates that information is missing. The score shows the output from the comparative marker selection, which was performed using a T-test metric with GenePattern's comparative marker analysis module.

Supplementary Table S3. Proteins that associate with TLOC1. Proteins that associated with TLOC1 by immunoprecipitation of V5-tagged TLOC1 followed by mass spectrometry analysis are listed in the table. The first and second column describes the amount of identified unique and total peptides per protein. The list is ranked according to the proteins with most unique identified peptides. The peptide coverage for each identified protein is listed in the third column. The suggested function in the fourth column is based on literature and database searches.

Supplementary Table S4. Proteins that have been reported to associate with the *TLOC1 Drosophila melanogaster* ortholog *Trp1*. Four proteins (bold) identified from the TLOC1 mass spectrometry experiments were in common with 16 previously identified proteins found to associate with the TLOC1 *Drosophila* ortholog Trp1. The list is retrieved from the *Drosophila* Protein Interaction Map database (DPiM, https://interfly.med.harvard.edu/). The human and mouse orthologs were added manually.

Supplementary Table S5. Intersection of gene expression signatures that correlate with either high SKIL or low SMAD4 gene expression. A significant overlap in gene expression signatures was found between ones derived from cells with either high SKIL expression or low SMAD4 expression. The gene expression signatures were derived from analysis of data from 634 cell lines in the cancer cell line encyclopedia (all cell lines except ones of hematological origin). High and low expression gene expression of SKIL and SMAD4 was defined as one standard deviation above or below average. Of 634 cell lines, 93 cell lines were identified to have high SKIL expression and 59 cell lines low SMAD4 level. Of these 15 had both high SKIL and low SMAD4 levels. Comparative maker selection analyses were performed in GenePattern between cell lines with high SKIL and the rest or low SMAD4 and the rest. The top 50 comparative markers for each comparison are listed in the table. 13 genes were shared between these two lists and are displayed in the middle column. This overlap is significantly higher than expected by chance ($p<0.001$ by a binomial distribution test).

Supplementary Table S6. Primers used for generating *TLOC1* truncation mutants. Primers used for PCR amplification of *TLOC1* truncation mutants. The forward primer was used in combination with all the other reverse primers to generate full length, short splice and truncation variants. The bold text in the primer sequence denotes the common Gateway BP cloning sequence.

Supplementary Table S7. List of shRNAs used in this study. The shRNA vectors are listed with their corresponding symbol, clone ID and clone name. The shRNAs are arranged according for which experiment they were used and which is marked on the top of each paragraph.

Supplementary Table S8. Primers used for quantitative real time PCR in this study. The sequence for each gene specific forward and reverse primers are listed in the table.