Supplementary Methods

**DNA copy number analysis**

Tissue was collected from Spectrum Health Hospital of Grand Rapids, MI and the Cooperative Human Tissue Network (CHTN) under the VARI Institutional Review Board. All tissue was snap-frozen in liquid nitrogen immediately after nephrectomy and stored at -80°C. Genomic DNA was extracted using a Jetquick Maxiprep kit (Genomed, St. Louis, MO, USA) per the manufacturer’s instructions. The SNP mapping assay was performed according to Affymetrix’s protocol. SNP intensity data was analyzed using Affymetrix GeneChip Genotyping analysis software (GTYPE) version 4.0. Starting from the CEL files, the R `oligo` package was used to read the raw SNP intensities. The SNP’s A or B allele signal was averaged by its sense and antisense signals respectively. To obtain SNP raw copy numbers, we used a total of 56 normal references either downloaded from Affymetrix (n = 48) or obtained by our own scan on normal kidney tissues (n = 8) to calculate the reference signals for each SNP's A and B alleles. The overall signal for the SNP was the sum of its A and B signals. The raw copy number for a SNP of a target sample was its overall signal subtracted by the averaged overall signals from the normal references (as all signals are in logarithm scale). A divide-and-conquer algorithm was used to partition the raw copy numbers into segments based on the maximum likelihood estimate of the break points. The segmentation algorithm was a two round algorithm. In the first round, the algorithm identifies segments with unusual large or small test scores (e.g., 5 standard deviations away from the mean), will mark the segment as outliers. The reported scores were the output from the second round of segmentation based on the raw copy numbers after removal of the outliers identified from the first round. The raw copy numbers of SNPs within any identified segment will be regarded identical and a (one-sided) t-test score (to test location=0) would replace the raw copy numbers within the identified segment and was reported thereafter. A segment identified in an individual sample would be declared either a gain or a loss if the score
for the segment was larger or smaller than a preselected cutoff value (e.g., 10). The SNP data has been deposited at the Gene Expression Omnibus (GEO 25399).

**PTTG1 and ECT2 expression analysis**

Expression data was generated from the Affymetrix HG-U133 Plus 2.0 platform as previously described (Yang et al., 2006). A Cox Proportional Hazards model using cancer-specific survival (CSS) and the expression level of PTTG1 or ECT2 as a continuous predictor. The significance values (p-values) from the Cox models are reported in the figures and in the text. To visualize the relationship between gene expression levels and CSS, Kaplan-Meier survival analysis was performed. The gene expression values were turned into dichotomous variables (high and low) using an optimal cut procedure. For PTTG1 and ECT2, the expression range was divided into 25 evenly spaced points to create 25 cut points. These cut points were used to create 25 dichotomous variables for PTTG1 and ECT2 where high expression was defined as those samples that had expression values above the cut point and low expression was defined as those samples that had expression values less than or equal to the cut point. A cut point was discarded if either the high expression sample group or the low expression sample group contained less than 25 percent of samples. The cut point that resulted in the lowest p-values for PTTG1 and ECT2 were then used to create the Kaplan-Meier survival curves used in the figures. To identify differentially expressed genes following PTTG1 knockdown, expression profiles from treated and control samples were generated from the Affymetrix HG-U133 Plus 2.0 using RMA analysis and data was subsequently unlogged. Genes with expression values under 20 were considered to be not expressed (NE) in the indicated cell line. Fold-change of each gene was defined to be the ratio of expression intensity of that gene in the treated sample over that in the control sample, if the ratio was positive; otherwise, the fold change was defined to be the negative of inverse value of the ratio.