Supplementary Methods:

Sanger sequencing:

Target genes in individual patient samples were PCR amplified using standard techniques and sequenced using conventional Sanger methods, yielding 96•6% of all trimmed reads with an average quality score of 20 or more. All traces were manually reviewed by AAH, using Mutation Surveyor (SoftGenetics, State College, PA). All variants were validated by repeat PCR amplification and Sanger re-sequencing of unamplified diagnostic DNA with the exception of known VHL point mutations from the COSMIC database or VHL frameshift mutations. All mutations not previously reported to be either somatic or germline (dbsnp) were analyzed in matched normal kidney DNA, to determine somatic status. Mutations were further categorized, including frameshift (FS), in-frame deletions (IFD), nonsense (NS), splice site (SS) – defined as essential splices (within first 2 base-pairs of coding region), and missense (MS). Missense mutations were further analyzed for the likelihood of functional impact based on evolutionary conservation of the affected amino acid in protein homologs (www.mutationassessor.org). [1]

IMPACT assay:

Target-specific probes were designed to capture all protein-coding exons of genes of interest for hybrid selection (Agilent SureSelect) as previously described.[2] This list included commonly implicated oncogenes, tumor suppressor genes, and components of pathways deemed actionable by current targeted therapies (see previous publications for complete target list). Barcoded sequence libraries for were prepared using 500 nanograms of input tumor and matched normal DNA according to the manufacturer's instructions (Illumina TruSeq). Libraries were pooled at equimolar concentrations (100 ng per library) and input to a single exon capture reaction as previously described.[3]To prevent off-target hybridization we spiked in a pool of blocker

oligonucleotides complementary to the full sequences of all barcoded adaptors (to a final total concentration of 10 micromolar). Hybridized DNA was sequenced on a single lane of an Illumina HiSeq 2000 flow cell to generate paired-end 75-bp reads. Data were demultiplexed using CASAVA, and reads were aligned to the reference human genome (hg19) using the Burrows-Wheeler Alignment tool.[4] Local realignment and quality score recalibration were performed using the Genome Analysis Toolkit (GATK) according to GATK best practices.[5]

1. Reva B, Antipin Y, Sander C: Predicting the functional impact of protein mutations: application to cancer genomics. Nucleic Acids Res 39:e118, 2011.

2. Gnirke A, Melnikov A, Maguire J, et al. Solution hybrid selection with ultra-long oligonucleotides for massively parallel targeted sequencing. Nature biotechnology 2009;27:182-9.

3. Wagle N, Berger MF, Davis MJ, et al. High-throughput detection of actionable genomic alterations in clinical tumor samples by targeted, massively parallel sequencing. Cancer discovery 2012;2:82-93.

4. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 2009;25:1754-60.

5. McKenna A, Hanna M, Banks E, et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res 2010;20:1297-303.