Supplementary Methods and Information:

Direct Solution Exome Capture and Sequencing and Mapping:

Exome design

Five (PD2125, PD2126, PD2144, PD2147, PD3441) of the matched clinical sample ccRCC pairs (normal + tumour) enrichment was performed using the Agilent SureSelect Human Exon Kit (Agilent, G3362) corresponding to the exons annotated within the CCDS database. For the remaining 2 clinical sample matched pairs (PD2126, PD3295), a custom in-house design was submitted and baits synthesized and supplied by Agilent (Agilent Technologies Inc, Santa Clara, CA, USA). The custom design included additional exonic regions over those present in CCDS and comprised a total of 288,654 unique exons from 46,275 transcripts of 20,921 Ensembl protein-coding genes, 33,621 transcripts of 13,772 manually annotated protein-coding genes, and 1635 miRNA genes (Coffey et al, manuscript in preparation). Baits for both exomes were provided in a single tube solution format.

Genomic library preparation

Genomic DNA (5ug) was fragmented by Adaptive Focused Acoustics on a Covaris E120 (Covaris Inc, Woburn, MA, USA) for 90 sec with a duty cycle of 20%, intensity of 5 and cycles per burst of 200. The fragmented DNA was purified using a Qiaquick PCR purification column (Qiagen, 28104) and quantified on a Bioanalyser using the Agilent DNA 1000 kit (Agilent, 5067-1504). The resulting DNA ranged in size from ~100-400bp, with a modal fragment size of ~250bp. Genomic libraries were prepared using the Illumina Paired End Sample Prep Kit following the manufacturer’s instructions (Illumina, San Diego CA, USA). Adapter-ligated DNA was purified using AMPure beads (Agencourt BioSciences Corporation, Beverly, MA, USA) following the manufacturer’s protocol, and eluted in 40ul of nuclease-free water. The prepared library was used directly in the subsequent enrichment procedure without prior size-selection or PCR amplification.

Exon enrichment

The genomic library (500ng) was mixed with 7.5ug human C_{0}t1 DNA, lyophilized in a speedvac for 30 min at 45°C and rehydrated in 3.4ul of nuclease-free water. Enrichment of the genomic DNA was performed using the Agilent SureSelect kit with minor modifications to the manufacturer’s protocol. Briefly, the genomic DNA library (3.4ul) was combined with 2.5ul of
Block reagent 1, 2.5ul of Block reagent 2 and 0.6ul of Block reagent 3 and transferred to a well of a microtitre plate. The sample was denatured by incubating the plate on a thermocycler at 95°C for 5 min then snap-cooled on ice. A hybridization mix was prepared comprising 25ul of Hyb reagent 1, 1ul of Hyb reagent 2, 10ul of Hyb reagent 3 and 13ul of Hyb reagent 4. A 13ul aliquout of this mastermix was added to the denatured DNA, and the sample incubated at 95°C for 5 min, then 65°C for 5 min. In a separate microtitre plate, the baits were prepared by combining 5ul of SureSelect capture library with 1ul of nuclease free water and 1ul of RNAse block, and the plate incubated at 65°C for 3 min. The pre-warmed DNA (22ul) was transferred to the pre-warmed bait mix and the solution incubated for 24h at 65°C. Following hybridization, the captured DNA was isolated using streptavidin-coated magnetic Dynabeads, (Invitrogen, 653.05) and washed following the standard Agilent SureSelect protocol. The isolated DNA was purified using a Qiagen MinElute purification column, eluted in 15ul of elution buffer and PCR-amplified for 14 cycles as previously described1.

**Substitution variant calling:**

Mapping of paired-end read data to the human genome (Build 37) was done using BWA². An average of 5 gigabases of uniquely mapping and 3.7 gigabases of uniquely mapping reads on target were obtained per sample, with an average of 74% of all reads mapping on target. Sixty-percent of target bases had 20X or greater coverage and 50 percent had 40X or greater coverage.

CaVEMan (Cancer Variants through Expectation Maximisation), a bespoke Java application using a simple expectation maximisation algorithm implementation³ was used to call single nucleotide substitutions. Through comparison of reads from both tumour and normal with the reference genome, CaVEMan calculates a probability for each possible genotype per base (given tumour and normal copy number). In order to provide more accurate estimates of sequence error rates within the algorithm, thus aid identification of true variants, variables such as base quality, read position, lane, and read orientation are incorporated into the calculations. Once CaVEMan was run, several post processing filters were applied in order to further increase the specificity of somatic mutation calls.

1. At least 1/3 of mutant alleles in tumour reads are of quality >= 25.
2. At least 1 mutant allele in a tumour read must fall in the middle third of the read, unless the tumour read depth is less than 10, when a mutant allele the first third is acceptable.
3. There is no more than 1 high quality (>= 20) mutant allele in a normal read.
Insertion/Deletion variant calling:

A modified version of Pindel\(^4\) was used to call insertions and deletions. By modifying the input file generation process we were able to increase sensitivity and increase confidence in events detected by BWA which was used as the initial mapping tool. The accepted approach for generating input for Pindel is to provide all read pairs where one end is unmapped and the other is confidently mapped to the genome, an anchor read. We found that by including read pairs where both ends map to the genome but allowing for one of the pair to have mismatches, insertions or deletions we could greatly increase coverage over smaller events (in some cases both ends are used as an anchor, creating two input records). The majority of these small events are detected by the BWA mapping algorithm, however, this increases confidence that the events are worth investigating. A second modification to the input generation was included to help identify small events close to large scale deletions or repetitive regions. In regions such as these we would not be able to capture any of the smaller events that can be detected within a single end of a read that is confidently mapped but with some form of mismatch, insertion or deletion. In these cases we generated an artificial anchor co-ordinate so that Pindel can attempt a realignment of these reads. Software that can generate input files of this form can be obtained by contacting the authors.

Once Pindel was run several post processing filters were applied. We considered there to be 2 classes of event in our data, large events > 4 b.p. and small events <= 4 b.p. which are detectable by BWA (non-SW). For many of the filters the mapping depths within the BAM file are used to aid filtering of poor confidence calls.

For both classes the following filters were applied to the raw output:

1. Event must occur in tumour reads
2. >3 tumour reads must support call
3. <5% of calls must occur in wildtype
4. When no wildtype coverage in BAM, Pindel must not call event in wildtype

For small events these filters were applied:

1. Tumour with BAM depth of < 200 reads must have variant call in >=8% of reads
2. Tumour with BAM depth of >= 200 reads must have variant call in >=4% of reads
3. Wildtype BAM must have >5 reads spanning the region
4. Pindel calls in wildtype reads must be <= 5% of the wildtype BAM depth
5. If the tumour BAM depth > wildtype BAM depth, normalise the Pindel wildtype calls against this, discarding if new value is >= 5% reference
6. Apply poly nucleotide tract filter for events with repetitive region > 9 repeats
7. Wildtype BAM depth must be >=8% of tumour BAM depth
8. Tumour BAM bust have <8% BWA reference calls vs BWA variant calls.

Further, for large events no wildtype reads should be called as part of an event by Pindel and exome data results must annotate to coding regions of the genome. Novel germline variants (verified by PCR based capillary sequencing) not previously reported in dbSNP or found in other sequencing screens are given in Supplementary Table 7.

**PBRM1 mutation screening.**

The coding exons of PBRM1 were sequenced via PCR-based capillary sequencing as previously described. Data were analysed semi-automated mutation detection followed by visual inspection of sequencing traces as previously described. The primer sequences for PBRM1 amplification and sequencing are given in Supplemental Table 8.

**Missense mutation analyses**

In order to evaluate the functional effects of the found missense mutations we fixed a scoring system using protein domain alignments from Pfam. The gene PBRM1 contains three kinds of functional domains: six copies of the Bromo domain (Pfam entry PF00439), two copies of the BAH domain (PF01426) and one copy of the HMG-box domain (PF00505). For each domain, we have used the Pfam seed alignment to construct a HMM-profile. In the Pfam full alignments all reported observations of this domain are aligned to this HMM-profile. We have extended these full alignments by the (6/2/1) hits within PBRM1 to fix the coordinate system. We denote the counts of amino acid $a$ in the alignment column $i$ by $n_i(a)$ and compare this observation to a null distribution $p_o(a)$ (overall genomic frequencies of amino acids). Taking the log odds ratio of the amino acid frequencies within the alignment column and the null gives a so called position specific score.
\[ s_i(a) = \log \frac{q_i(a)}{p_0(a)} = \log \frac{n_i(a) + p_0(a)}{(N_i + 1)p_0(a)} \quad (1) \]

where \( N_i \) is the total number of residues in the column. The above construct of the observed distribution uses pseudo-counts\(^8,9\) proportional to \( p_0 \) to account for non-observed residues in the finite sample. The two extreme cases are columns that are highly conserved - where the most prevalent letter receives a large positive score and all others large negative ones - and columns that are highly variable and close to neutral - where all letters receive scores close to zero. For similar conservation based scoring schemes for disease related variation see e.g. the recent review\(^10\) and in the context of cancer mutations\(^11,12\). For a given missense mutation (falling onto alignment column \( i \)), we can now record the score difference between the final and the initial residue:

\[ \Delta s_i = s_i(a_{\text{final}}) - s_i(a_{\text{initial}}) \quad (2) \]

Out of the 9 missense mutations we could score 3 using the Pfam alignments (T232P \( \Delta s = -7.78 \), A597D \( \Delta s = -9.69 \), H1204P \( \Delta s = -2.76 \)). In order to assess if these three somatic mutations differ significantly from random mutations we generated \textit{in silico} all possible point events in PBRM1 (transcript ENST00000337303) that result in a missense mutation which falls onto our scoring system (i.e. mutational opportunity space). From this set we drew 10,000 sets of 3 mutations randomly and evaluated the mean score for each set - the resulting distribution is shown together with the somatic value in Figure 2 in the main paper. Somatic mutations are significantly different from the null set (p-value 0.01). More specifically, the somatic mutation set has a lower mean negative score (i.e. they are predicted to be more deleterious on average) than the null model - thus making them interesting candidates for follow up functional studies.
Confirmation of exon trapping by RT-PCR in mouse pancreatic tumours

Total RNA (1 μg) from tumors with transposon insertions in Pbrm1 was reverse transcribed into single stranded cDNA using Reverse Transcriptase III (Invitrogen) and Random Hexamers (Invitrogen) following the manufacturer protocol. 1 μL of the resulting cDNA was used as a template in a first round of PCR using specific primers corresponding to exon 23 of Pbrm1 (5’-TGGCTGAAGGTTGGTGATTG-3’) and Carp-β-Actin Splice acceptor sequence (5’-TAAATTCCCGCGAATCCATC-3’). The product of this reaction was used as a template in a second round of nested PCR using specific primers corresponding to Pbrm1 exon 24 (5’-TTGAGAAAGTATGGGTCCGAGA-3’) and a second external primer corresponding to Carp-β-Actin Splice acceptor sequence (5’-CATACCGCTACGTTGCTAA-3’). The resulting bands were capillary sequenced.

PBRM1 knockdown and functional analyses

Cell lines and Transfections

Cell lines tested including ACHN, 786-O, SN12C, U031 A704 Caki-1 and TK10 were cultured in complete medium supplemented with 10% FBS (v/v) under 37°C and 5% CO2. PBRM1 or scrambled control siRNAs (Santa Cruz, CA) were transfected into renal cell lines using Lipofectamine 2000 (Invitrogen, CA) according to the manufacturer's conditions.

siRNA sequences, which detect all three PBRM1 splice forms corresponding to NM_018165, NM_018313 AND NM_181042, were as follows:

C CCAUAGUUGUAGCUACAAA
C GAAAGCAUCACUUCUUUA
C GCACUCAGCUAUACCACAA
Real-time PCR

Total RNA was extracted from 48 hour post-transfected cells using TriPure (Roche, pIN). cDNA synthesis was carried out by using iScript™ cDNA Synthesis Kit (Bio-Rad, CA). Real-time PCR was performed to determine expression level of PBRM1 and β-actin by SsoFast EvaGreen Supermix using CFX96™ Real-Time PCR Detection System (Bio-Rad, CA). Primers used for amplification were: PBRM1-F (5'-GTGTGATGAACCAAGGAGTGGC-3'); PBRM1-R (5'-GATATGGAGGTGGTGCCTGCTG-3'); β-actin-F (5'-GATCAGCAAGCAGGAGTATGACG-3') and β-actin-R (5'-AAGGGTTGAACGCAACTAAGTCATAG-3'). Relative expression of PBRM1 was normalized with β-actin expression level.

Western blot analysis

Cellular proteins were extracted with phosphate buffered saline (PBS) containing 0.1% (v/v) Triton X-100 (Sigma, LA) in the presence of protease inhibitors. Proteins resolved by SDS-PAGE were electroblotted to a nitrocellulose membrane (Amersham, Buckinghamshire) and the membrane was incubated overnight at 4°C with blocking buffer (PBS containing 5% (w/v) skim milk and 0.05% (v/v) Tween-20). Primary and secondary antibody incubations were done in
blocking buffer. Anti-PBRM1 antibody was purchased from Bethyl Laboratories (TX) and anti-β-actin antibody was from Sigma (LA). The membranes were washed with PBS containing 0.05% (v/v) Tween-20 followed by analysis using the Supersignal Chemiluminescent kit (Pierce, IL) according to the manufacturer’s recommendations.

**Proliferation assay**

After 48 hour transfection, $2 \times 10^3$ cells were plated per well in 96-well plate. Growth of PBRM1 siRNA- and scramble siRNA-transfected cells was determined using the colorimetric 3-(4,5-dimethylthiazol-2-yl)-5-(3-carboxymethoxyphenyl)-(4-sulfophenyl)-2H-tetrazolium assay according to the manufacturer’s protocol (MTS; Promega, WI). The assay was performed in triplicate.

**Migration assay**

After 48 hour transfection, $2.0 \times 10^5$ cells in serum-free medium were seeded into the upper chamber of BioCoat inserts containing filters with 8 μm pores for migration assay (BD Pharmingen, CA). The lower chamber was filled with 10% (v/v) serum-containing medium as attractant. Cells that did not migrate through the filters after 22 hours post-incubation were removed with cotton swabs. Cells that traversed through the filter were fixed and stained by Diff-Quik Solution (Dade Behring, DE). After staining, cells were taken photos.

**Soft Agar Assay**

SN12C cells were cultured in a two-layer agar system to prevent their attachment to the plastic surface. After transfection, cells ($4 \times 10^4$) were trypsinized to single-cell suspensions, resuspended in 0.4% agar (Sigma, LA), and added to a preset 1% bottom agar layer in six-well plates. The top agar cell layers were covered with culture medium. Cells were incubated in 5% CO$_2$ at 37°C for 14 days, and colonies were counted under ×2.5 object. Experiments were performed in triplicate.

**PBRM1 knockdown expression phenotype analyses**

**Gene expression data generation and processing.** RNA was isolated from 786-O, SN12C, and TK10 cells that were either transfected with scrambled siRNA or transfected with PBRM1
targeting siRNA. Single color gene expression data was generated using the HG-U133 Plus 2.0 chipset (Affymetrix, Santa Clara, CA) as described\textsuperscript{13} and deposited in the Gene Expression Omnibus (GEO22316). Gene expression analysis was performed using R/BioConductor version 2.0 software\textsuperscript{14}. Summary expression values were computed using the RMA method as implemented in the \textit{affy} package using updated probe set mappings (hgu133plus2hsentrezgcf version 12) such that a single probe set is associated with each well measured gene\textsuperscript{15,16}.

**Gene expression analysis.** Gene set enrichment analysis was performed using curated gene sets obtained from MSigDB (http://www.broadinstitute.org/gsea/msigdb/) and using additional curated gene sets obtained from the \textit{PGSEA} package. Log-transformed relative expression values derived from comparison of targeted versus scrambled siRNA were computed for each cell line. For each cell line, gene sets that were significantly enriched in up-regulated genes were identified using the mean-rank method with permutation (n=10,000) as implemented in the \textit{limma} package\textsuperscript{17}. Gene sets that were significantly deregulated (P < 0.05) in all three cell lines were identified and sorted based on the lowest average p-value. Individual genes that were deregulated within specific gene sets were identified using a moderated t-statistic and significance values adjusted to control for multiple testing using the Benjamini & Hochberg approach as implemented in the \textit{limma} package.

Gene expression data generated from renal cell carcinoma samples and non-diseased kidney samples were obtained from the Gene Expression Omnibus (GSE17895) as previously described\textsuperscript{5}. The set of samples that displayed the hypoxic phenotype (n=90) were isolated and correlations between PBRM1 expression and other genes computed using Pearson's correlation.

1. Mamanova, L. et al. Target-enrichment strategies for next-generation sequencing. \textit{Nat Meth} 7, 111-118.
2. Li, H. & Durbin, R. Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics* **26**, 589-595 (2010).

3. Do, C.B. & Batzoglou, S. What is the expectation maximization algorithm? *Nat Biotech* **26**, 897-899 (2008).

4. Ye, K., Schulz, M.H., Long, Q., Apweiler, R. & Ning, Z. Pindel: a pattern growth approach to detect break points of large deletions and medium sized insertions from paired-end short reads. *Bioinformatics* **25**, 2865-2871 (2009).

5. Dalgliesh, G.L. et al. Systematic sequencing of renal carcinoma reveals inactivation of histone modifying genes. *Nature* **463**, 360-363 (2010).

6. Finn, R.D. et al. The Pfam protein families database. *Nuc Acids Res* **38**, D211-222 (2010).

7. Eddy, S.R. Profile hidden Markov models. *Bioinformatics* **14**, 755-763 (1998).

8. Durbin, R., Eddy, S., Krogh, A. & Mitchison, G. *Biological sequence analysis: probabilistic models of proteins and nucleic acids*, (Cambridge University Press, Cambridge 1998).

9. Lawrence, C.E. et al. Detecting subtle sequence signals: a Gibbs sampling strategy for multiple alignment. *Science* **262**, 208-214 (1993).

10. Jordan, D.M., Ramensky, V.E. & Sunyaev, S.R. Human allelic variation: perspective from protein function, structure, and evolution. *Current Opinion in Structural Biology* **20**, 342-350.

11. Dixit, A. et al. Sequence and Structure Signatures of Cancer Mutation Hotspots in Protein Kinases. *PLoS ONE* **4**, e7485 (2009).

12. Yue, P. et al. Inferring the functional effects of mutation through clusters of mutations in homologous proteins. *Human Mutation* **31**, 264-271 (2010).

13. Yang, X.J. et al. A molecular classification of papillary renal cell carcinoma. *Cancer Res* **65**, 5628-37 (2005).

14. Gentleman, R.C. et al. Bioconductor: open software development for computational biology and bioinformatics. *Genome Biol* **5**, R80 (2004).

15. Dai, M. et al. Evolving gene/transcript definitions significantly alter the interpretation of GeneChip data. *Nuc Acids Res* **33**, e175 (2005).

16. Irizarry, R.A. et al. Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics* **4**, 249-64 (2003).

17. Smyth, G.K. Linear models and empirical bayes methods for assessing differential expression in microarray experiments. *Stat Appl Genet Mol Biol* **3**, Article3 (2004).
### Supplementary Table 1 - Clinical Samples in exome sequencing

| Sample   | Sex | Age | Grade | Histology   | VHL mutation^ | SETD2 mutation | UTX mutation |
|----------|-----|-----|-------|-------------|----------------|----------------|--------------|
| PD2125a  | M   | 82  | 4     | Clear Cell  |                |                |              |
| PD2126a  | F   | 74  | 1     | Clear Cell  | c.236_241delGCAGTC; p.R79_P81>P | c.1801T>A; p.R601* |              |
| PD2127a  | F   | 59  | 4     | Clear Cell  |                |                |              |
| PD2144a  | F   | 63  | 4     | Clear Cell  | c.525delC; p.Y175* |                |              |
| PD2147a  | F   | 50  | 2     | Clear Cell  |                | c.4161_4162delTG; p.Y1387fs*1 |              |
| PD3295a  | M   | 62  | 4     | Clear Cell  |                |                |              |
| PD3441a  | M   | 69  | 1     | Clear Cell  | c.223_225delATC; p.F76_C77>C |                |              |

^ VHL mutations in PD2126a and PD3441a were not "re-discovered" in exome sequencing due to poor coverage of the highly GC-rich first exon.
| Sample | Chromosome | Position | Gene | Annotated Transcript | WT base | Mut Base | Mutation Type | Protein Annotation | cDNA Annotation |
|--------|------------|----------|------|----------------------|---------|----------|---------------|-------------------|-----------------|
| PD2127a | 16         | 20855285 | AC004381.2 | ENST00000261377     | A       | C        | SYNONYMOUS    | p.L552L           | c.1656A>C       |
| PD2127a | 2          | 179449098 | AC010680.2 | ENST00000356127     | C       | A        | MISSENSE      | p.G19157V         | c.57470G>T      |
| PD2126a | 2          | 51028374 | AC012100.1 | ENST00000261854     | C       | T        | MISSENSE      | p.E286K           | c.856G>A        |
| PD2147a | 7          | 158529750 | AC19084.2  | ENST00000435514     | C       | T        | MISSENSE      | p.V258V           | c.774G>A        |
| PD2147a | 4          | 25677963 | AC092436.2 | ENST00000382051     | G       | C        | SYNONYMOUS    | p.L555L           | c.1665G>C       |
| PD2147a | 17         | 43214437 | ACBD4      | ENST00000321854     | T       | -        | FRAMESHIFT    | p.F116fs*7        | c.347DelT       |
| PD2147a | 5          | 80643677 | ACOT12     | ENST00000307624     | A       | T        | MISSENSE      | p.L190H           | c.569T>A        |
| PD2144a | 2          | 148680563 | ACVR2A     | ENST00000241416     | G       | A        | MISSENSE      | p.A367T           | c.1099G>A       |
| PD2144a | 2          | 9683393  | ADAM17     | ENST00000497134     | G       | C        | NONSENSE      | p.S40*            | c.119G>C        |
| PD2127a | 2          | 100625295 | AFF3       | ENST00000409579     | A       | G        | SYNONYMOUS    | p.S76S            | c.228T>C        |
| PD2147a | 5          | 133220253 | AFF4       | ENST00000378595     | A       | T        | MISSENSE      | p.S757T           | c.2269T>A       |
| PD2147a | 10         | 45498936 | AL353801.1 | ENST00000298299     | C       | T        | SYNONYMOUS    | p.D40D            | c.120C>T        |
| PD2126a | 12         | 45803231 | ANO6       | ENST00000441406     | A       | G        | MISSENSE      | p.M640V           | c.1918A>G       |
| PD2127a | 11         | 55258787 | AP001998.1 | ENST00000314657     | A       | G        | SYNONYMOUS    | p.T23T            | c.699G>A        |
| PD2125a | 19         | 45451775 | APOC2      | ENST00000252490     | C       | T        | SYNONYMOUS    | p.L141L           | c.40C>T         |
| PD2125a | 13         | 111944635 | ARHGEF7    | ENST00000375737     | A       | G        | MISSENSE      | p.T612A           | c.1834A>G       |
| PD2126a | 1          | 27094351 | ARID1A     | ENST00000457599     | G       | A        | MISSENSE      | p.R1020K          | c.3059G>A       |
| PD2127a | 1          | 27106656 | ARID1A     | ENST00000457599     | T       | C        | MISSENSE      | p.L1872P          | c.5615T>C       |
| PD2127a | 10         | 63850639 | ARID5B     | ENST00000297873     | A       | T        | NONSENSE     | p.K473*           | c.1417A>T       |
| PD2147a | 17         | 42249629 | ASB16      | ENST00000293414     | A       | C        | MISSENSE      | p.T173P           | c.517A>C        |
| PD2127a | 1          | 119976966 | ASTN2      | ENST00000373996     | G       | A        | MISSENSE      | p.A229V           | c.686C>T        |
| PD2127a | 10         | 96794874 | ATG2B      | ENST00000359933     | T       | A        | ESSENTIAL_SPLICE | p.---             | c.-A>T          |
| PD2126a | 22         | 46085613 | ATXN10     | ENST00000252934     | C       | A        | SYNONYMOUS    | p.I46L            | c.138C>A        |
| PD3295a | 1          | 171506448 | BAT2D1     | ENST00000367742     | -       | N-FRAME DELETION | p.P782_S787delPNSES | c.2340delGTACCGAAAGTCTGA |
| PD2125a | 6          | 31630173 | BAT4       | ENST00000375896     | G       | T        | MISSENSE      | p.T314N           | c.941C>A        |
| PD2127a | 6          | 38142761 | BTBD9      | ENST00000419706     | T       | A        | STOP_LOST    | p.*S83Y           | c.1749A>A       |
| PD2125a | 10         | 124457732 | C10orf120  | ENST00000329446     | C       | G        | SYNONYMOUS    | p.R175R           | c.525G>C        |
| PD2147a | 15         | 24922328 | C15orf2    | ENST00000329468     | C       | A        | SYNONYMOUS    | p.I438I           | c.1314C>A       |
| PD2125a | 5   | 43388447 | CCL28 | ENST00000361115 | ATAGAAA - | FRAMESHIFT | p.I218fs*379 | c.652delATAGAAA |
|---------|-----|----------|-------|------------------|----------|------------|-------------|----------------|
| PD2127a | 5   | 137686950| COL5A1 | ENST00000371817  | C T      | MISSENSE   | p.P908L     | c.2723C>T      |
| PD2127a | 6   | 158272284| CYTIP  | ENST00000377833  | G A      | MISSENSE   | p.S2020A    | c.6058A>T      |
| PD2127a | 7   | 23018596 | DDX53  | ENST00000327968  | T C      | MISSSE     | p.I64fs*17  | c.190delATgtgag |
| PD2127a | 8   | 43483790 | CCND3B | ENST00000395813  | A C      | MISSENSE   | p.S878R     | c.2634C>G      |
| PD2127a | 9   | 10239283 | FAM173B | ENST00000280330 | A G      | SYNONYMOUS | p.R259R     | c.777G>A       |
| PD2127a | 10  | 10829496 | GPR68  | ENST000003357776 | G A      | SYNONYMOUS | p.L170L     | c.510G>A       |
| PD2127a | 11  | 176272328| FBN2   | ENST00000327968  | A C      | MISSSE     | p.N141T     | c.422A>C       |
| PD2127a | 12  | 42979966 | FAM173B | ENST00000327968 | A C      | MISSSE     | p.I2048V    | c.6142A>G      |
| PD2127a | 13  | 7132566  | COL5A1 | ENST00000332930  | A G      | MISSSE     | p.P908L     | c.2723C>T      |
| PD2127a | 14  | 36900169 | EIF3D  | ENST00000397224  | T A      | MISSSE     | p.P908L     | c.2723C>T      |
| PD2127a | 15  | 53416352 | EIF4B  | ENST00000430205  | C G      | MISSSE     | p.P203R     | c.608C>G       |
| PD2127a | 16  | 23297290 | ENTPD4 | ENST00000417069  | T C      | MISSSE     | p.I333V     | c.997A>G       |
| PD2127a | 17  | 42979966 | FAM173B | ENST00000327968 | A C      | MISSSE     | p.L170L     | c.510G>A       |
| PD2127a | 18  | 127627328| FBN2   | ENST00000327968  | A C      | MISSSE     | p.N141T     | c.422A>C       |
| PD2127a | 19  | 95657584 | DYN1C1 | ENST00000359388  | A T      | MISSSE     | p.I282T     | c.845T>C       |
| PD2127a | 20  | 39376176 | DAB2   | ENST00000339788  | G T      | MISSSE     | p.L506M     | c.1516C>A      |
| PD2127a | 21  | 36900169 | EIF3D  | ENST00000397224  | T A      | MISSSE     | p.P908L     | c.2723C>T      |
| PD2127a | 22  | 53416352 | EIF4B  | ENST00000430205  | C G      | MISSSE     | p.P203R     | c.608C>G       |
| PD2127a | 23  | 23297290 | ENTPD4 | ENST00000417069  | T C      | MISSSE     | p.I333V     | c.997A>G       |
| PD2127a | 24  | 10239283 | FAM173B | ENST00000327968 | A C      | MISSSE     | p.I282T     | c.845T>C       |
| PD2127a | 25  | 42979966 | FAM173B | ENST00000327968 | A C      | MISSSE     | p.I282T     | c.845T>C       |
| PD2127a | 26  | 127627328| FBN2   | ENST00000327968  | A C      | MISSSE     | p.I282T     | c.845T>C       |
| PD2127a | 27  | 42979966 | FAM173B | ENST00000327968 | A C      | MISSSE     | p.I282T     | c.845T>C       |
| PD2127a | 28  | 127627328| FBN2   | ENST00000327968  | A C      | MISSSE     | p.I282T     | c.845T>C       |
| PD2144a | 20 | 23345310 | GZF1 | ENST00000377051 | C | T | MISSENSE | p.A97V | c.290C>T |
|---|---|---|---|---|---|---|---|---|---|
| PD2127a | 7 | 18914186 | HDAC9 | ENST00000401921 | C | A | MISSENSE | p.P880T | c.2638C>A |
| PD2147a | 10 | 96352065 | HELLS | ENST00000441434 | A | G | MISSENSE | p.H568R | c.1703A>G |
| PD3441a | 7 | 92848502 | HEPACAM2 | ENST00000394468 | G | T | MISSENSE | p.F114L | c.342C>A |
| PD3441a | 21 | 38269279 | HLCS | ENST00000399120 | C | T | SYNONYMOUS | p.E444E | c.1332G>A |
| PD2127a | 5 | 149386186 | HMGXB3 | ENST00000261804 | A | C | MISSENSE | p.K245N | c.735A>C |
| PD2126a | 9 | 21077581 | IFNB1 | ENST00000380232 | A | G | SYNONYMOUS | p.S96S | c.288T>C |
| PD2147a | 10 | 96352065 | HELLS | ENST00000441434 | A | G | MISSENSE | p.H568R | c.1703A>G |
| PD3441a | 7 | 92848502 | HEPACAM2 | ENST00000394468 | G | T | MISSENSE | p.F114L | c.342C>A |
| PD3441a | 21 | 38269279 | HLCS | ENST00000399120 | C | T | SYNONYMOUS | p.E444E | c.1332G>A |
| PD2127a | 5 | 149386186 | HMGXB3 | ENST00000261804 | A | C | MISSENSE | p.K245N | c.735A>C |
| PD2126a | 9 | 21077581 | IFNB1 | ENST00000380232 | A | G | SYNONYMOUS | p.S96S | c.288T>C |
| PD2147a | 10 | 96352065 | HELLS | ENST00000441434 | A | G | MISSENSE | p.H568R | c.1703A>G |
| PD3441a | 7 | 92848502 | HEPACAM2 | ENST00000394468 | G | T | MISSENSE | p.F114L | c.342C>A |
| PD3441a | 21 | 38269279 | HLCS | ENST00000399120 | C | T | SYNONYMOUS | p.E444E | c.1332G>A |
| PD2127a | 5 | 149386186 | HMGXB3 | ENST00000261804 | A | C | MISSENSE | p.K245N | c.735A>C |
| PD2126a | 9 | 21077581 | IFNB1 | ENST00000380232 | A | G | SYNONYMOUS | p.S96S | c.288T>C |
| PD2147a | 10 | 96352065 | HELLS | ENST00000441434 | A | G | MISSENSE | p.H568R | c.1703A>G |
| PD3441a | 7 | 92848502 | HEPACAM2 | ENST00000394468 | G | T | MISSENSE | p.F114L | c.342C>A |
| PD3441a | 21 | 38269279 | HLCS | ENST00000399120 | C | T | SYNONYMOUS | p.E444E | c.1332G>A |
| PD2127a | 5 | 149386186 | HMGXB3 | ENST00000261804 | A | C | MISSENSE | p.K245N | c.735A>C |
| PD2126a | 9 | 21077581 | IFNB1 | ENST00000380232 | A | G | SYNONYMOUS | p.S96S | c.288T>C |
| PD2147a | 10 | 96352065 | HELLS | ENST00000441434 | A | G | MISSENSE | p.H568R | c.1703A>G |
| PD3441a | 7 | 92848502 | HEPACAM2 | ENST00000394468 | G | T | MISSENSE | p.F114L | c.342C>A |
| PD3441a | 21 | 38269279 | HLCS | ENST00000399120 | C | T | SYNONYMOUS | p.E444E | c.1332G>A |
| PD2127a | 5 | 149386186 | HMGXB3 | ENST00000261804 | A | C | MISSENSE | p.K245N | c.735A>C |
| PD2126a | 9 | 21077581 | IFNB1 | ENST00000380232 | A | G | SYNONYMOUS | p.S96S | c.288T>C |
| PD2147a | 10 | 96352065 | HELLS | ENST00000441434 | A | G | MISSENSE | p.H568R | c.1703A>G |
| PD3441a | 7 | 92848502 | HEPACAM2 | ENST00000394468 | G | T | MISSENSE | p.F114L | c.342C>A |
| PD3441a | 21 | 38269279 | HLCS | ENST00000399120 | C | T | SYNONYMOUS | p.E444E | c.1332G>A |
| Genbank Accession | CHNS | ENST | NCBI NR ID | Gene | Gene ID | Description | Chromosome | Position | Chromosome Start Position | Chromosome End Position |
|-------------------|------|------|-------------|------|---------|-------------|------------|----------|--------------------------|------------------------|
| PD3295a 13        | 109777493 | MYO16 | ENST00000356711 | T    | A      | MISSENSE     | p.L1168H   | c.3503T>A |                         |                        |
| PD2126a 6         | 76589831 | MYO6  | ENST00000369985 | T    | A      | SYNONYMOUS  | p.P760P    | c.2280T>A |                         |                        |
| PD3295a 8         | 90958403 | NBN   | ENST00000265433 | A    | G      | MISSENSE     | p.Y679H    | c.2035T>C |                         |                        |
| PD3441a 16        | 4519450 | NMRAL1| ENST00000404295 | G    | T      | SYNONYMOUS  | p.S19S     | c.57C>A   |                         |                        |
| PD2126a 16        | 77759415 | NUDT7 | ENST00000268533 | C    | T      | SYNONYMOUS  | p.S41S     | c.123C>T |                         |                        |
| PD2127a X         | 70779127 | OGT   | ENST00000415630 | T    | C      | MISSENSE     | p.L412P    | c.1235T>C |                         |                        |
| PD2144a 19        | 15198065 | OR1I1 | ENST00000209540 | C    | G      | SYNONYMOUS  | p.L63L     | c.189C>G |                         |                        |
| PD2127a 1          | 248309179 | OR2M5 | ENST00000369985 | T    | A      | MISSENSE     | p.H244N    | c.730C>A |                         |                        |
| PD2127a 17        | 3195189 | OR3A1 | ENST00000397187 | C    | T      | SYNONYMOUS  | p.L235L    | c.705G>A |                         |                        |
| PD2127a 11        | 56409040 | OR5AP2| ENST00000302981 | A    | -      | FRAMESHIFT  | p.N26fs*13 | c.76delA |                         |                        |
| PD3441a 6         | 29323566 | OR5V1 | ENST00000377151 | A    | G      | MISSENSE     | p.L136P    | c.407T>C |                         |                        |
| PD2147a 1         | 158687308 | OR6K3 | ENST00000368146 | C    | T      | MISSENSE     | p.V216M    | c.646G>A |                         |                        |
| PD2126a 11        | 57798938 | OR6Q1 | ENST00000302622 | T    | A      | MISSENSE     | p.F172I    | c.514T>A |                         |                        |
| PD3441a 3         | 125266296 | OSBPL11| ENST00000393455 | T    | C      | MISSENSE     | p.S218G    | c.652A>G |                         |                        |
| PD2147a 20        | 49366431 | PAR6B | ENST00000371610 | C    | T      | SYNONYMOUS  | p.D175D    | c.525T>C |                         |                        |
| PD3441a 3         | 52613158 | PBRM1 | ENST00000394830 | C    | A      | MISSENSE     | p.E1124*   | c.3370C>A |                         |                        |
| PD2126a 3         | 52649441 | PBRM1 | ENST00000337303 | -    | T      | FRAMESHIFT  | p.K621fs*9 | c.1862insT |                        |                        |
| PD2127a 3         | 52678768 | PBRM1 | ENST00000337303 | T    | -      | FRAMESHIFT  | p.K284fs*16 | c.851delA |                         |                        |
| PD3295a 3         | 52613132 | PBRM1 | ENST00000337303 | GC   | -      | FRAMESHIFT  | p.W1157fs*23 | c.3471_3472delGC |                        |                        |
| PD2126a 12        | 96692721 | PCTK2 | ENST00000261211 | G    | A      | MISSENSE     | p.T214I    | c.641C>T |                         |                        |
| PD2126a 10        | 95422894 | PDE6C | ENST00000371447 | T    | G      | MISSENSE     | p.I826S    | c.2477T>G |                         |                        |
| PD2127a 10        | 75675107 | PLAU  | ENST00000446342 | A    | G      | MISSENSE     | p.T340A    | c.1018A>G |                         |                        |
| PD2144a 17        | 37263687 | PLXDC1| ENST00000444435 | G    | A      | MISSENSE     | p.D12V     | c.35C>T  |                         |                        |
| PD2127a 11        | 74329770 | POLD3 | ENST00000263681 | G    | T      | MISSENSE     | p.G194V    | c.581G>T |                         |                        |
| PD2127a 11        | 74329772 | POLD3 | ENST00000263681 | A    | T      | MISSENSE     | p.M195L    | c.583A>T |                         |                        |
| PD2127a 2         | 46313388 | PRKCE | ENST00000306156 | A    | -      | FRAMESHIFT  | p.D493fs*33 | c.1479delIC |                        |                        |
| PD2125a 12        | 50027318 | PRPF40B| ENST00000261897 | G    | A      | MISSENSE     | p.D162N    | c.484G>A |                         |                        |
| PD3295a 14        | 73673169 | PSEN1 | ENST00000394164 | A    | G      | MISSENSE     | p.Y311C    | c.932A>G |                         |                        |
| PD2144a 21        | 30342888 | RN160 | ENST00000361371 | C    | T      | SYNONYMOUS  | p.T387T    | c.1161G>A |                         |                        |
| PD2147a 22        | 39710171 | RPL3  | ENST00000401609 | G    | T      | MISSENSE     | p.L246M    | c.736C>A |                         |                        |
| PD2127a 3         | 38768438 | SCN10A| ENST00000449082 | G    | A      | MISSENSE     | p.R916W    | c.2746C>T |                         |                        |
| PD2125a 7         | 83634829 | SEMA3A| ENST00000265362 | G    | A      | MISSENSE     | p.P396S    | c.1186C>T |                         |                        |
| PD2126a 3         | 47164325 | SETD2 | ENST00000409792 | T    | A      | NONSENSE    | p.R601*    | c.1801T>A |                         |                        |
| Gene    | Ensembl ID | Chromosome | Position | Transcript ID | Coding Region | Change | Mutation Type | Protein Change | Reference 1 | Reference 2 |
|---------|------------|------------|----------|---------------|---------------|--------|---------------|----------------|-------------|-------------|
| SHPRH   | ENST00000275233 | 6         | 146243853 | PD3295a       | A             | T      | MISSENSE     | p.V1222D       | c.3665T>A   |             |
| SIGLEC8 | ENST00000440804 | 19        | 51958879  | PD3295a       | C             | G      | MISSENSE     | p.V282L        | c.844G>C    |             |
| SLC44A3 | ENST00000446120 | 1         | 95330381  | PD2125a       | G             | T      | MISSENSE     | p.G405W        | c.1213G>T   |             |
| SLC4A4  | ENST00000340595 | 4         | 72420907  | PD2127a       | A             | G      | SYNONYMOUS   | p.A871A        | c.2613A>G   |             |
| SLC6A16 | ENST00000454748 | 19        | 49813078  | PD2127a       | C             | T      | MISSENSE     | p.E236K        | c.706G>A    |             |
| SIGLEC8 | ENST00000374299 | X         | 70147450  | PD2127a       | A             | T      | MISSENSE     | p.M356K        | c.1067T>A   |             |
| SLC8A1  | ENST00000408028 | 2         | 40656752  | PD2125a       | C             | G      | MISSENSE     | p.G405W        | c.1213G>T   |             |
| SLC6A16 | ENST00000454748 | X         | 70147450  | PD2127a       | C             | T      | MISSENSE     | p.E236K        | c.706G>A    |             |
| SLC7A3  | ENST00000374299 | X         | 70147450  | PD2127a       | A             | T      | MISSENSE     | p.M356K        | c.1067T>A   |             |
| SMYD4   | ENST00000305513 | 17        | 1690187   | PD2127a       | C             | G      | MISSENSE     | p.G405W        | c.1213G>T   |             |
| SPAM1   | ENST00000413927 | 7         | 123593637 | PD3441a       | A             | C      | MISSENSE     | p.K5Q          | c.13A>C     |             |
| STOX1   | ENST00000399162 | 19        | 49398324  | PD2127a       | T             | C      | MISSENSE     | p.S119P        | c.35T>C     |             |
| TCHH    | ENST00000368804 | 1         | 152081494 | PD2144a       | C             | G      | MISSENSE     | p.R1400P       | c.4199G>C   |             |
| TCL6    | ENST00000357168 | 14        | 96136875  | PD2147a       | T             | C      | MISSENSE     | p.S119P        | c.35T>C     |             |
| TCTN    | ENST00000441824 | 20        | 58011759  | PD2127a       | C             | T      | SYNONYMOUS   | p.G68G         | c.204C>T    |             |
| TMEM97a | ENST00000443106 | 3         | 111782388 | PD2125a       | T             | A      | SYNONYMOUS   | p.S476S        | c.1428T>A   |             |
| TNFRSF13B | ENST00000261652 | 17        | 16855793  | PD2127a       | G             | T      | MISSENSE     | p.H56N         | c.166C>A    |             |
| TOX     | ENST00000361421 | 8         | 59750765  | PD2127a       | C             | T      | MISSENSE     | p.A267T        | c.799G>A    |             |
| TSHR    | ENST00000298171 | 14        | 81606140  | PD2147a       | T             | G      | SYNONYMOUS   | p.L270L        | c.810T>G    |             |
| TULP2   | ENST0000021399  | 19        | 49398324  | PD2125a       | C             | T      | MISSENSE     | p.V149I        | c.445G>A    |             |
| UFP3B   | ENST00000276201 | X         | 118979153 | PD2147a       | tactgt        | -      | SPLICE       | p.---          | c.---       |             |
| USH2A   | ENST00000366942 | 1         | 216497008 | PD3295a       | G             | A      | MISSENSE     | p.T453I        | c.1358C>T   |             |
| UTX     | ENST00000377967 | 4         | 44969479  | PD2147a       | TG            | -      | FRAMESHIFT  | p.Y1387fs*1    | c.4161_4162delTG |             |
| VHL     | ENST00000256474 | 3         | 101915134 | PD2144a       | G             | -      | FRAMESHIFT  | p.Y175*        | c.525delC   |             |
| WDR52   | ENST00000393845 | 3         | 113060702 | PD3295a       | G             | T      | MISSENSE     | p.T257N        | c.770C>A    |             |
| WNT1    | ENST00000293549 | 12        | 49373404  | PD3441a       | G             | T      | SYNONYMOUS   | p.L86L         | c.258G>T    |             |
| YWHAB   | ENST00000353703 | 20        | 43530469  | PD2127a       | G             | A      | MISSENSE     | p.V99I         | c.295G>A    |             |
| ZNF442  | ENST00000420150 | 19        | 12460630  | PD2127a       | T             | A      | MISSENSE     | p.H590L        | c.1769A>T   |             |
| ZNF507  | ENST00000355898 | 19        | 32845488  | PD295a        | C             | T      | SYNONYMOUS   | p.S584S        | c.1752C>T   |             |
| Sample   | Sex | Age | Grade | Histology   |
|----------|-----|-----|-------|-------------|
| PD1580a  | F   | 61  | 3     | clear cell  |
| PD1582a  | F   | 47  | 2     | clear cell  |
| PD1590a  | M   | 64  | 2     | clear cell  |
| PD1593a  | M   | 59  | 4     | clear cell  |
| PD1753a  | M   | 67  | 3     | clear cell  |
| PD1754a  | F   | 42  | 3     | clear cell  |
| PD1759a  | M   | 76  | 3     | clear cell  |
| PD1767a  | F   | 80  | 2     | clear cell  |
| PD1769a  | M   | 43  | 3     | clear cell  |
| PD2125a  | M   | 82  | 4     | clear cell  |
| PD2126a  | F   | 74  | 1     | clear cell  |
| PD2127a  | F   | 59  | 4     | clear cell  |
| PD2129a  | M   | 75  | 4     | clear cell  |
| PD2130a  | F   | 67  | 2     | clear cell  |
| PD2131a  | F   | 63  | 4     | clear cell  |
| PD2133a  | M   | 52  | 3     | clear cell  |
| PD2134a  | F   | 83  | 3     | clear cell  |
| PD2135a  | M   | 73  | 4     | clear cell  |
| PD2136a  | M   | 67  | 3     | clear cell  |
| PD2138a  | M   | 74  | 4     | clear cell  |
| PD2139a  | F   | 62  | 2     | clear cell  |
| PD2140a  | M   | 48  | 3     | clear cell  |
| PD2142a  | M   | 49  | 3     | clear cell  |
| PD2144a  | F   | 63  | 4     | clear cell  |
| PD2145a  | M   | 63  | 2     | clear cell  |
| PD2146a  | M   | 64  | 3     | clear cell  |
| PD2147a  | F   | 50  | 2     | clear cell  |
| PD2148a  | F   | 77  | 2     | clear cell  |
| PD2149a  | M   | 66  | 4     | clear cell  |
| PD2154a  | F   | 64  | 4     | clear cell  |
| PD2155a  | M   | 50  | 3     | clear cell  |
| PD2157a  | M   | 32  | 2     | clear cell  |
| PD2160a  | M   | 73  | 3     | clear cell  |
| PD2161a  | F   | 42  | 2     | clear cell  |
| PD2163a  | M   | 54  | 3     | clear cell  |
| PD2167a  | M   | 59  | 2     | clear cell  |
| PD2168a  | M   | 53  | 4     | clear cell  |
| PD2170a  | M   | 60  | 4     | clear cell  |
| PD2172a  | M   | 67  | 3     | clear cell  |
| PD2173a  | F   | 83  | 1     | clear cell  |
| PD2174a  | F   | 65  | 2     | clear cell  |
| PD2177a  | M   | 49  | 2     | clear cell  |
| PD2180a  | F   | 45  | 2     | clear cell  |
| PD2181a  | F   | 71  | 2     | clear cell  |
| PD2183a  | F   | 66  | 1     | clear cell  |
| PD2185a  | F   | 71  | 1     | clear cell  |
| PD2186a  | M   | 50  | 2     | clear cell  |
| PD2187a  | F   | 49  | 2     | clear cell  |
| PD2190a  | F   | 61  | 2     | clear cell  |
| Patient ID | Sex | Age | Stage | Histology                  |
|------------|-----|-----|-------|----------------------------|
| PD2191a    | M   | 68  | 3     | clear cell                 |
| PD2192a    | M   | 78  | 3     | clear cell                 |
| PD2193a    | M   | 61  | 3     | clear cell                 |
| PD2194a    | M   | 74  | 3     | clear cell                 |
| PD2198a    | M   | 70  | 3     | clear cell                 |
| PD2199a    | M   | 58  | 2     | clear cell                 |
| PD2203a    | F   | 60  | 2     | clear cell                 |
| PD2207a    | M   | 66  | 2     | clear cell                 |
| PD2208a    | F   | 65  | 2     | clear cell                 |
| PD2209a    | F   | 69  | 3     | clear cell                 |
| PD2213a    | M   | 60  | 4     | clear cell                 |
| PD2217a    | M   | 44  | 4     | clear cell                 |
| PD2219a    | M   | 47  | 2     | clear cell                 |
| PD2222a    | F   | 54  | N/D   | clear cell                 |
| PD3284a    | M   | 56  | 3     | clear cell                 |
| PD3285a    | F   | 80  | 3     | clear cell                 |
| PD3286a    | F   | 71  | 4     | clear cell                 |
| PD3287a    | M   | 61  | 2     | clear cell                 |
| PD3290a    | F   | 58  | 2     | clear cell                 |
| PD3292a    | M   | 76  | 3     | papillary                  |
| PD3293a    | M   | 80  | 3     | clear cell                 |
| PD3294a    | F   | 60  | 3     | clear cell                 |
| PD3295a    | M   | 62  | 4     | clear cell                 |
| PD3296a    | M   | 72  | 3     | clear cell                 |
| PD3306a    | F   | 80  | 2-3   | Clear Cell w/ minor granular component |
| PD3307a    | F   | 64  | 4     | clear cell                 |
| PD3308a    | M   | 43  | 3     | clear cell                 |
| PD3309a    | M   | 67  | 4     | clear cell                 |
| PD3312a    | F   | 81  | 3     | clear cell                 |
| PD3313a    | F   | 48  | 2     | clear cell                 |
| PD3314a    | F   | 58  | 3     | clear cell                 |
| PD3316a    | M   | 44  | 3     | clear cell                 |
| PD3317a    | M   | 62  | 2     | clear cell                 |
| PD3318a    | M   | 74  | 2     | papillary                  |
| PD3324a    | M   | 51  | 3     | clear cell                 |
| PD3332a    | M   | 65  | 1-2   | Papillary (Chromophil)     |
| PD3333a    | M   | 48  | 2     | papillary                  |
| PD3336a    | F   | 74  | 2     | clear cell                 |
| PD3337a    | F   | 71  | 2     | clear cell                 |
| PD3340a    | M   | 49  | 2     | clear cell                 |
| PD3342a    | M   | 65  | 3     | clear cell                 |
| PD3343a    | M   | 69  | 2     | papillary 1                |
| PD3348a    | M   | 57  | 2     | clear cell                 |
| PD3349a    | F   | 56  | 3     | clear cell                 |
| PD3350a    | M   | 51  | 2     | clear cell                 |
| PD3351a    | M   | 65  | 3     | clear cell                 |
| PD3355a    | F   | 59  | 2     | clear cell                 |
| PD3363a    | F   | 72  | 2     | clear cell                 |
| PD3364a    | M   | 70  | 3     | clear cell                 |
| ID   | Gender | Age | Tumor Type | Description               |
|------|--------|-----|------------|---------------------------|
| PD3365a | F     | 54  | 3          | chromophobe               |
| PD3368a | M     | 54  | 3          | clear cell                |
| PD3371a | F     | 56  | 2          | clear cell                |
| PD3372a | F     | 67  | 4          | clear cell/Sarcomatoid    |
| PD3375a | F     | 57  | 2          | clear cell                |
| PD3376a | F     | 82  | 2          | clear cell                |
| PD3378a | F     | 52  | 3          | clear cell                |
| PD3379a | M     | 62  | 2          | clear cell                |
| PD3381a | M     | 66  | 3          | clear cell                |
| PD3382a | M     | 56  | 2          | clear cell                |
| PD3385a | M     | 71  | 4          | clear cell                |
| PD3388a | F     | 73  | 4          | clear cell/sarcomatoid    |
| PD3389a | M     | 48  | 3-4        | clear cell                |
| PD3390a | F     | 67  | 4          | clear cell                |
| PD3391a | M     | 54  | 2          | clear cell                |
| PD3392a | M     | 50  | 2          | clear cell                |
| PD3393a | F     | 54  | 3          | clear cell                |
| PD3394a | M     | 61  | 4          | clear cell                |
| PD3395a | M     | 74  |            | Mucinous Tubular and Spindle Cell Carcinoma |
| PD3397a |        | 47  | 1 (focally 2-3) | clear cell |
| PD3399a | F     | 58  | 1-2        | clear cell                |
| PD3400a | M     | 51  | 3          | clear cell                |
| PD3402a | M     | 53  | 2          | clear cell                |
| PD3403a | M     | 63  | 3          | papillary                 |
| PD3404a | F     | 55  | 2          | clear cell                |
| PD3405a | M     | 55  | 4          | clear cell                |
| PD3408a | F     | 78  | 2          | clear cell                |
| PD3409a | M     | 69  | 2          | clear cell                |
| PD3410a | F     | 54  | 4          | clear cell                |
| PD3411a | F     | 58  | 3          | clear cell                |
| PD3413a | F     | 51  | 3          | clear cell                |
| PD3420a | M     | 58  | 2          | clear cell                |
| PD3421a | M     | 65  | 3          | clear cell                |
| PD3422a | M     | 68  | 1-2        | clear cell                |
| PD3423a | M     | 61  | 2          | papillary 1               |
| PD3424a | M     | 66  | 3          | clear cell                |
| PD3425a | M     | 48  | 3          | clear cell                |
| PD3426a | M     | 51  | 2          | chromophobe               |
| PD3427a | M     | 50  | 2          | clear cell                |
| PD3436a | M     | 38  | 2          | clear cell                |
| PD3437a | F     | 68  | 2          | clear cell                |
| PD3438a | M     | 51  | 3          | clear cell                |
| PD3439a | M     | 59  | 3          | clear cell                |
| PD3440a | M     | 70  | 2          | clear cell                |
| PD3441a | M     | 69  | 1          | clear cell                |
| PD3442a | M     | 73  | 2          | Papillary w/ focal clear cell |
| PD3443a | F     | 48  | 4          | clear cell                |
| PD3446a | M     | 72  | 3          | NOS                       |
| PD3449a | M     | 69  | 4          | Clear Cell                |
| Patient ID | Gender | Age | Tumor Type          |
|------------|--------|-----|---------------------|
| PD3452a    | M      | 68  | Clear Cell          |
| PD3453a    | F      | 64  | clear cell          |
| PD3454a    | M      | 70  | clear cell          |
| PD3455a    | F      | 42  | clear cell/Sarcomatoid |
| PD3456a    | M      | 36  | clear cell          |
| PD3457a    | M      | 39  | clear cell          |
| PD3458a    | F      | 69  | chromophobe         |
| PD3459a    | F      | 52  | 3 (focal areas of 4) clear cell |
| PD3467a    | F      | 72  | clear cell          |
| PD3468a    | M      | 41  | #N/A                |
| PD3469a    | F      | 85  | clear cell          |
| PD3470a    | F      | 58  | clear cell          |
| PD3471a    | F      | 57  | clear cell          |
| PD3472a    | M      | 66  | clear cell          |
| PD3473a    | F      | 73  | clear cell          |
| PD3474a    | M      | 62  | papillary 1         |
| PD3475a    | M      | 69  | papillary           |
| PD3476a    | F      | 46  | clear cell          |
| PD3479a    | M      | 76  | oncocytoma          |
| PD3481a    | F      | 80  | clear cell          |
| PD3483a    | M      | 67  | clear cell          |
| PD3484a    | M      | 72  | clear cell          |
| PD3485a    | M      | 69  | clear cell          |
| PD3486a    | M      | 70  | papillary           |
| PD3487a    | M      | 44  | clear cell          |
| PD3488a    | F      | 69  | clear cell          |
| PD3489a    | F      | 58  | clear cell          |
| PD3490a    | F      | 59  | clear cell          |
| PD3491a    | M      | 87  | clear cell          |
| PD3492a    | M      | 42  | clear cell          |
| PD3493a    | F      | 78  | clear cell          |
| PD3494a    | M      | 42  | clear cell          |
| PD3495a    | M      | 71  | clear cell          |
| PD3497a    | M      | 63  | clear cell          |
| PD3499a    | M      | 55  | clear cell          |
| PD3500a    | M      | 49  | clear cell          |
| PD3501a    | F      | 83  | clear cell          |
| PD3502a    | M      | 84  | chromophobe         |
| PD3503a    | M      | 60  | clear cell          |
| PD3504a    | F      | 37  | clear cell          |
| PD3505a    | M      | 54  | clear cell          |
| PD3506a    | F      | 62  | clear cell          |
| PD3507a    | M      | 57  | clear cell          |
| PD3508a    | F      | 71  | clear cell          |
| PD3509a    | M      | 58  | clear cell          |
| PD3510a    | F      | 89  | clear cell          |
| PD3511a    | F      | 58  | clear cell          |
| PD3512a    | M      | 83  | clear cell          |
| PD3513a    | M      | 69  | clear cell          |
| PD3514a    | F      | 43  | clear cell          |
| PatientID | Gender | Age  | Stage  | Histology     |
|-----------|--------|------|--------|---------------|
| PD3515a   | M      | 80   | 2      | clear cell    |
| PD3516a   | F      | 68   | 3      | clear cell    |
| PD3518a   | M      | 74   | 2      | chromophobe   |
| PD3519a   | F      | 85   | 3      | NOS           |
| PD3520a   | M      | 64   | 2-3    | clear cell    |
| PD3521a   | M      | 54   | 3      | clear cell    |
| PD3522a   | M      | 64   | 3      | clear cell    |
| PD3523a   | F      | 47   | 3      | clear cell    |
| PD3524a   | M      | 66   | 3      | clear cell    |
| PD3525a   | F      | 74   | 1      | clear cell    |
| PD3526a   | M      | 59   | 2      | papillary     |
| PD3528a   | F      | 50   | 2      | clear cell    |
| PD3529a   | M      | 56   | 2      | clear cell    |
| PD3530a   | M      | 69   | 3      | clear cell    |
| PD3532a   | M      | 41   | 2      | clear cell    |
| PD3534a   | M      | 55   | 2      | clear cell    |
| PD3536a   | F      | 71   | 3      | Clear Cell    |
| PD3538a   | M      | 73   | 3      | clear cell    |
| PD3539a   | M      | 78   | 3      | NOS           |
| PD3540a   | F      | 52   | 2      | clear cell    |
| PD3541a   | M      | 61   | 2      | clear cell    |
| PD3542a   | F      | 59   | 2      | clear cell    |
| PD3543a   | M      | 55   | 2      | clear cell    |
| PD3544a   | F      | 50   | 3      | papillary     |
| PD3545a   | M      | 54   | 2      | chromophobe   |
| PD3546a   | F      | 57   | 2      | clear cell    |
| PD3547a   | F      | 63   | 2      | papillary     |
| PD3548a   | M      | 66   | 2      | clear cell    |
| PD3550a   | M      | 44   | 2      | clear cell    |
| PD3552a   | M      | 54   | 2      | clear cell    |
| PD3554a   | F      | 55   | 3      | clear cell    |
| PD3555a   | F      | 66   | 1      | clear cell    |
| PD3556a   | M      | 69   | 2      | clear cell    |
| PD3557a   | F      | 49   | 1      | clear cell    |
| PD3558a   | M      | 55   | 2      | clear cell    |
| PD3559a   | F      | 74   | 2      | clear cell    |
| PD3560a   | F      | 78   | 2      | clear cell    |
| PD3561a   | M      | 72   | 3-4    | clear cell    |
| PD3562a   | F      | 73   | 1      | oncocyntoma   |
| PD3563a   | M      | 57   | 3      | clear cell    |
| PD3564a   | F      | 68   | 3      | clear cell    |
| PD3565a   | M      | 65   | 3      | clear cell    |
| PD3566a   | M      | 44   | 3      | clear cell    |
| PD3567a   | M      | 75   | high   | Papillary Urothelial Carcinoma |
| PD3568a   | F      | 52   | N/D    | Urothelial Carcinoma |
| PD3569a   | F      | N/D  | 3      | clear cell    |
| PD3570a   | M      | 51   | 2      | papillary     |
| PD3571a   | M      | 72   | 3      | papillary 2   |
| PD3572a   | F      | 46   | N/D    | oncocyntoma   |
| PD3573a   | M      | 65   | 2      | clear cell    |
| Patient ID | Gender | Age | Stage | Pathological Diagnosis       |
|------------|--------|-----|-------|------------------------------|
| PD3574a    | M      | 38  | 3     | clear cell                   |
| PD3575a    | F      | 62  | N/D   | inflammatory myofibroblastic tumor |
| PD3576a    | M      | 69  | 2     | papillary                   |
| PD3577a    | F      | 53  | 3     | chromophobe                  |
| PD3578a    | M      | 67  | 2-3   | clear cell                   |
| PD3581a    | M      | 66  | 3     | clear cell                   |
| PD3582a    | M      | 56  | 2     | clear cell                   |
| PD3587a    | M      | N/D | 2     | clear cell                   |
| PD3588a    | F      | 77  | 1     | clear cell                   |
| PD3589a    | M      | 91  | 3     | clear cell                   |
| PD3590a    | M      | 43  | 3     | papillary 2                  |
| PD3591a    | M      | 79  | high  | Urothelial Carcinoma         |
| PD3592a    | F      | 49  | 2     | clear cell                   |
| PD3594a    | M      | 44  | 2     | clear cell                   |
| PD3596a    | M      | 54  | 3     | Clear Cell                   |
| PD3597a    | M      | 60  | 2     | clear cell                   |
| PD3598a    | M      | 70  | 3     | clear cell                   |
**Supplementary Table 4 - PBRM1 somatic mutations**

| Sample  | Chr | Position | WT allele | Mut Allele | Annotated Trascript | Protein annotation | cDNA annotation | Type           |
|---------|-----|----------|-----------|------------|---------------------|-------------------|----------------|----------------|
| PD1580a | 3   | 52662980 | T         | -          | ENST00000337303    | p.N458fs*17       | c.1373delTA    | INDEL          |
| PD1590a | 3   | 52712590 | AT        | -          | ENST00000337303    | p.Y54fs*1         | c.162_163delTA | INDEL          |
| PD1754a | 3   | 52712583 | GAT       | -          | ENST00000337303    | p.I57del          | c.169_171delATC | INFRAME DEL    |
| PD1759a | 3   | 52620608 | T         | A          | ENST00000337303    | p.K1074*          | c.3220A>T      | NONSENSE       |
| PD1767a | 3   | 52702550 | GCTGG     | A          | ENST00000337303    | p.Q117fs*56       | c.348_352>T    | INDEL          |
| PD2127a | 3   | 52643913 | T         | A          | ENST00000337303    | p.K661N           | c.1983A>T      | MISSENSE       |
| PD2129a | 3   | 52643561 | G         | A          | ENST00000337303    | p.Q779*           | c.2335C>T      | NONSENSE       |
| PD2130a | 3   | 52696193 | C         | -          | ENST00000337303    | p.D162fs*12       | c.484delG      | INDEL          |
| PD2131a | 3   | 52620701 | TTAAGTA   | -          | ENST00000337303    | p.Y1043fs*9       | c.3127_3134delTACTTTAA | INDEL |
| PD2135a | 3   | 52597493 | G         | A          | ENST00000337303    | p.Q1298*          | c.3892C>T      | NONSENSE       |
| PD2140a | 3   | 52637682 | T         | -          | ENST00000337303    | p.E878fs*37       | c.2634delA     | INDEL          |
| PD2145a | 3   | 52651277 | C         | T          | ENST00000337303    | p.?               | Exon 14 +1 G>A | ESSENTIAL_SPLICE |
| PD2146a | 3   | 52677884 | T         | -          | ENST00000337303    | p.I279fs*4        | c.835delA      | INDEL          |
| PD2154a | 3   | 52712548 | AG        | -          | ENST00000337303    | p.C69fs*1         | c.204_205delCT | INDEL          |
| PD2155a | 3   | 52668646 | T         | A          | ENST00000337303    | p.K425*           | c.1273A>T      | NONSENSE       |
| PD2163a | 3   | 52678763 | C         | A          | ENST00000337303    | p.E286*           | c.856G>T       | NONSENSE       |
| PD2170a | 3   | 52595987 | G         | C          | ENST00000337303    | p.?               | Exon 25 -3 C>G | INTRONIC       |
| PD2172a | 3   | 52584514 | G         | C          | ENST00000337303    | p.S1500*          | c.4499C>G      | NONSENSE       |
| PD2174a | 3   | 52651306 | G         | T          | ENST00000337303    | p.A597D           | c.1790C>A      | MISSENSE       |
| PD2181a | 3   | 52597493 | G         | A          | ENST00000337303    | p.Q1298*          | c.3892C>T      | NONSENSE       |
| PD2183a | 3   | 52621485 | ATGTTTC   | -          | ENST00000337303    | p.E1003fs*9       | c.3007_3013delGAAACAT | INDEL |
| PD2186a | 3   | 52610623 | TTCTTTTGTAGAACA | - | ENST00000337303 | p.M1209_E1214delMFYKKE | c.3625_3642delATGTTCTACAAAAA | INFRAME DEL |
| PD2186a | 3   | 52610637 | T         | G          | ENST00000337303    | p.H1204P          | c.3611A>C      | MISSENSE       |
| PD2190a | 3   | 52643742 | TGG       | -          | ENST00000337303    | p.Y718_Q719>*     | c.2154_2156delCCA | INDEL |
| PD2192a | 3   | 52651294 | TCAT      | AT         | ENST00000337303    | p.N601fs*8        | c.1802_1805delTA | INDEL |
| PD2193a | 3   | 52696194 | A         | -          | ENST00000337303    | p.D161fs*13       | c.483delT      | INDEL          |
| PD2194a | 3   | 52643755 | A         | -          | ENST00000337303    | p.M714fs*17       | c.2141delT     | INDEL          |
| PD2199a | 3   | 52595782 | C         | A          | ENST00000337303    | p.?               | Exon 25 +1 G>T | ESSENTIAL_SPLICE |
| PD2203a | 3   | 52663050 | T         | -          | ENST00000337303    | p.T435fs*3        | c.1303delA     | INDEL          |
| PD2207a | 3   | 52661375 | T         | -          | ENST00000337303    | p.E486fs*14       | c.1455delA     | INDEL          |
| PD2208a | 3   | 52668757 | G         | A          | ENST00000337303    | p.Q388*           | c.1162C>T      | NONSENSE       |
| PD2209a | 3   | 52702591 | T         | -          | ENST00000337303    | p.M103fs*10       | c.307delA      | INDEL          |
| Sample ID | Gene ID | Reference | cDNA Change | protein Change | Type |
|-----------|---------|-----------|-------------|----------------|------|
| PD2217a  | ENST00000337303 | p.I279fs*4 | c.835delA | INDEL |
| PD2219a  | ENST00000337303 | p.D589fs*2 | c.1764delA | INDEL |
| PD2222a  | ENST00000337303 | p.? | Exon 6-5 T>G | INTRONIC |
| PD2222a  | ENST00000337303 | p.S652* | c.1955C>G | NONSENSE |
| PD3284a  | ENST00000337303 | p.G646fs*4 | c.1937_1955delGCATTTCTCCTAAAA | AATC | INDEL |
| PD3290a  | ENST00000337303 | p.S652* | c.1955C>G | NONSENSE |
| PD3312a  | ENST00000337303 | p.? | Exon 6 -1 G>A | ESSENTIAL_SPLICE |
| PD3313a  | ENST00000337303 | p.? | Exon 22 -1 del(TTATATTTTCTCCCCAAG) | ESSENTIAL_SPLICE |
| PD3314a  | ENST00000337303 | p.K708fs*14 | c.2124_2125delAA | INDEL |
| PD3317a  | ENST00000337303 | p.R540S | c.1620A>C | MISSENSE |
| PD3336a  | ENST00000337303 | p.? | Exon 15 -1 G>C | ESSENTIAL_SPLICE |
| PD3340a  | ENST00000337303 | p.K644fs*9 | c.1930_1939delAAGAGTGGCA | INDEL |
| PD3349a  | ENST00000337303 | p.? | c.232_236delCGAAG | INDEL |
| PD3355a  | ENST00000337303 | p.? | c.2873_2879delACCATGT | INDEL |
| PD3363a  | ENST00000337303 | p.? | c.3892C>T | NONSENSE |
| PD3371a  | ENST00000337303 | p.? | c.1186delT | INDEL |
| PD3375a  | ENST00000337303 | p.R1185* | c.3553C>T | NONSENSE |
| PD3382a  | ENST00000337303 | p.Q1298* | c.3892C>T | NONSENSE |
| PD3385a  | ENST00000337303 | p.? | c.2240_2246delAAAGATGGCA | INDEL |
| PD3391a  | ENST00000337303 | p.? | c.363_364delCA | INDEL |
| PD3400a  | ENST00000337303 | p.? | c.3695delA | INDEL |
| PD3411a  | ENST00000337303 | p.? | c.307delA | INDEL |
| PD3413a  | ENST00000337303 | p.? | c.363_364delCA | INDEL |
| PD3422a  | ENST00000337303 | p.? | c.3695delA | INDEL |
| PD3438a | 3 | 52582134 | - | A | ENST00000337303 | p.L1565fs*>19 | c.4694_4695insT | INDEL |
| PD3457a | 3 | 52623136 | T | - | ENST00000337303 | p.N972fs*42 | c.2915delA | INDEL |
| PD3467a | 3 | 52649441 | A | T | ENST00000337303 | p.L617* | c.1850T>A | NONSENSE |
| PD3469a | 3 | 52620607 | TT | - | ENST00000337303 | p.K1074fs*32 | c.3221_3222delAA | INDEL |
| PD3470a | 3 | 52649469 | A | - | ENST00000337303 | p.Y608fs*34 | c.1822delT | INDEL |
| PD3472a | 3 | 52637691 | A | - | ENST00000337303 | p.R876fs*39 | c.2625delT | INDEL |
| PD3476a | 3 | 52696289 | TATTCAGGAGAATC | - | ENST00000337303 | p.D130fs*1 | c.388_401delGATTCTCCTGAATA | INDEL |
| PD3487a | 3 | 52620611 | T | - | ENST00000337303 | p.I1073fs*86 | c.3217delA | INDEL |
| PD3490a | 3 | 52712586 | - | T | ENST00000337303 | p.T56fs*6 | c.166_167insA | INDEL |
| PD3492a | 3 | 52677318 | G | T | ENST00000337303 | p.K621E | c.1861A>G | NONSENSE |
| PD3501a | 3 | 52649430 | T | C | ENST00000337303 | p.K621E | c.1861A>G | MISSENSE |
| PD3506a | 3 | 52712614 | C | - | ENST00000337303 | p.? | c.1318C>T | NONSENSE |
| PD3511a | 3 | 52623085 | TTAC | - | ENST00000337303 | p.? | c.1318C>T | NONSENSE |
| PD3524a | 3 | 52663053 | T | A | ENST00000337303 | p.? | c.1318C>T | NONSENSE |
| PD3529a | 3 | 52685778 | T | G | ENST00000337303 | p.T232P | c.694A>C | MISSENSE |
| PD3536a | 3 | 52584541 | G | - | ENST00000337303 | p.P1491fs*14 | c.4472delC | INDEL |
| PD3538a | 3 | 52610715 | C | A | ENST00000337303 | p.? | c.1318C>T | NONSENSE |
| PD3540a | 3 | 52643510 | C | - | ENST00000337303 | p.E796fs*9 | c.2386delG | INDEL |
| PD3541a | 3 | 52613210 | T | - | ENST00000337303 | p.E1132fs*27 | c.3393delA | INDEL |
| PD3543a | 3 | 52582239 | T | - | ENST00000337303 | p.D1530fs*17 | c.4589delA | INDEL |
| PD3548a | 3 | 52682460 | T | C | ENST00000337303 | p.? | c.1318C>T | NONSENSE |
| PD3550a | 3 | 52663035 | G | A | ENST00000337303 | p.Q440* | c.1318C>T | NONSENSE |
| PD3554a | 3 | 52597340 | G | - | ENST00000337303 | p.L1349fs*35 | c.4045delC | INDEL |
| PD3555a | 3 | 52702645 | T | - | ENST00000337303 | p.Y85fs*2 | c.253_254insT | INDEL |
| PD3556a | 3 | 52696272 | GTTTGCAAGCGGCT | - | ENST00000337303 | p.K135fs*11 | c.405_418delAGCCGCTTGGAAAC | INDEL |
| PD3559a | 3 | 52643974 | TTTTTCATTTTTAGGAAATGCCAATCCTCTCTCCTA | - | ENST00000337303 | p.? | c.405_418delAGCCGCTTGGAAAC | INDEL |
| PD3563a | 3 | 52589125 | T | - | ENST00000337303 | p.G1273fs*2 | c.3816delA | INDEL |
| PD3573a | 3 | 52712587 | A | - | ENST00000337303 | p.N55fs*40 | c.165delT | INDEL |
| PD3582a | 3 | 52668716 | C | - | ENST00000337303 | p.Q402fs*2 | c.1203delG | INDEL |
| PD3587a | 3 | 52651439 | T | A | ENST00000337303 | p.K553* | c.1657A>T | NONSENSE |
| PD3588a | 3 | 52613142 | TCAT | - | ENST00000337303 | p.N1154fs*4 | c.3461_3464delATGA | INDEL |
| PD3596a | 3 | 52610766 | T | - | ENST00000337303 | p.E1214fs*4 | c.3639delA | NONSENSE |
| Sample     | Tissue      | Histology                             | Zygosity   | Cff | Position | WT allele | Mut Allele | Annotated Transcript | cDNA Annotation | Protein Annotation | Type   |
|------------|-------------|---------------------------------------|------------|-----|----------|-----------|------------|----------------------|----------------|--------------------|--------|
| NCI-H1793  | Lung        | Adenocarcinoma                        | Heterozygous| 3   | 52692313 | T         | A          | ENST00000337303     | c.547A>T       | p.R183*            | Nonsense |
| OCI-314    | Ovary       | Serous micro papillary carcinoma       | Heterozygous| 3   | 52676277 | G         | A          | ENST00000337303     | c.892C>T       | p.R296*            | Nonsense |
| PANC-10-05 | Pancreas    | Ductal carcinoma                      | Heterozygous| 3   | 52651496 | G         | A          | ENST00000337303     | c.1600C>T      | p.R534*            | Nonsense |
| ESS-1      | Endometrium | Carcinosarcoma-malignant mesodermal mixed tumour | Heterozygous| 3   | 52643768 | G         | A          | ENST00000337303     | c.2126C>T      | p.R710*            | Nonsense |
| HCC2998    | Large intestine, colon | Adenocarcinoma                  | Heterozygous| 3   | 52637678 | G         | A          | ENST00000337303     | c.2126C>T      | p.R710*            | Nonsense |
| ACHN       | Kidney      | Renal cell carcinoma                   | Heterozygous| 3   | 52620674 | G         | A          | ENST00000337303     | c.3154C>T      | p.R11052*         | Nonsense |
| NCI-H226   | Lung        | Squamous cell carcinoma                | Heterozygous| 3   | 52584766 | G         | A          | ENST00000337303     | c.435C>T       | p.Q1452*           | Nonsense |
| TALL-1     | Haematopoietic and lymphoid tissue | Lymphoid leukenoma, acute lymphoblastic T cell leukaemia | Heterozygous| 3   | 52584629 | G         | A          | ENST00000337303     | c.4384C>T      | p.Q1462*           | Nonsense |
| CW-2       | Carcinoma   | Carcinoma                             | Heterozygous| 3   | 52678784 | T         | -          | ENST00000337303     | c.835delA      | p.I279fs*4        | INDEL   |
| NCI-SNU-1  | Stomach     | Carcinoma                             | Heterozygous| 3   | 52678784 | T         | -          | ENST00000337303     | c.835delA      | p.I279fs*4        | INDEL   |
| A704       | Kidney      | Renal cell carcinoma                   | Heterozygous| 3   | 52651383 | -         | AA         | ENST00000337303     | c.1713_1714insTT| p.E572fs*16      | INDEL   |
| NCI-H2196  | Lung        | Small cell carcinoma                   | Heterozygous| 3   | 52637580 | T         | -          | ENST00000337303     | c.2736delA     | p.E913fs*2        | INDEL   |
| TGBC24TKB  | Biliary tract, bile duct | Carcinoma                  | Heterozygous| 3   | 52613114 | -         | A          | ENST00000337303     | c.3489_3490insT| p.V1164fs*17      | INDEL   |
| SUP-T1     | Haematopoietic and lymphoid tissue | Lymphoid leukenoma, acute lymphoblastic T cell leukaemia | Heterozygous| 3   | 52597340 | G         | -          | ENST00000337303     | c.4045delC     | p.L1349fs*35      | INDEL   |
| HCC2998    | Large intestine, colon | Adenocarcinoma                  | Heterozygous| 3   | 52582255 | A         | C          | ENST00000337303     | Exon 28 -4 T>G  | p.?                | INTRONIC |
| NCI-H378   | Lung        | Small cell carcinoma                   | Heterozygous| 3   | 52623082 | C         | T          | ENST00000337303     | Exon 17 +4 G>A  | p.?                | INTRONIC |
| NCI-H650   | Lung        | Bronchioloalveolar adenocarcinoma      | Heterozygous| 3   | 52712607 | C         | G          | ENST00000337303     | c.145G>C       | p.V49L             | MISSENSE |
| 8-MG-BA    | Central nervous system, frontal lobe | Glioma, astrocytoma Grade IV, glioblastoma multiforme | Heterozygous| 3   | 52712586 | T         | C          | ENST00000337303     | c.166A>G       | p.T566A           | MISSENSE |
| SW1417     | Large intestine, colon | Adenocarcinoma                  | Heterozygous| 3   | 52712556 | T         | C          | ENST00000337303     | c.196A>G       | p.R666G           | MISSENSE |
| 647-V      | Urinary tract, bladder | Transitional cell carcinoma             | Heterozygous| 3   | 52702630 | G         | C          | ENST00000337303     | c.268C>G       | p.Q90E             | MISSENSE |
| A4-Fuk     | Skin        | Malignant melanoma                    | Heterozygous| 3   | 52696246 | T         | A          | ENST00000337303     | c.431A>T       | p.Y114F           | MISSENSE |
| A4-Fuk     | Skin        | Malignant melanoma                    | Heterozygous| 3   | 52696198 | T         | G          | ENST00000337303     | c.479A>C       | p.E160A           | MISSENSE |
| SNG-M      | Endometrium | Adenocarcinoma                        | Heterozygous| 3   | 52692256 | G         | A          | ENST00000337303     | c.604C>T       | p.H202C           | MISSENSE |
| CCRF-CEM   | Haematopoietic and lymphoid tissue | Haematopoietic leukenoma, acute lymphoblastic leukaemia | Heterozygous| 3   | 52692244 | C         | T          | ENST00000337303     | c.616G>A       | p.E206K           | MISSENSE |
| CTV-1      | Haematopoietic and lymphoid tissue | Haematopoietic leukenoma, acute myeloid leukaemia, M5 | Heterozygous| 3   | 52680795 | T         | C          | ENST00000337303     | c.677A>G       | p.E226G           | MISSENSE |
| MDA-MB-231 | Breast      | Carcinoma                             | Heterozygous| 3   | 52685790 | T         | C          | ENST00000337303     | c.682A>G       | p.G228V           | MISSENSE |
| OS-RC-2    | Kidney      | Renal cell carcinoma                   | Heterozygous| 3   | 52685774 | A         | G          | ENST00000337303     | c.696T>C       | p.R233T           | MISSENSE |
| OVCA-R-5   | Ovary       | Carcinoma                             | Heterozygous| 3   | 52682407 | C         | T          | ENST00000337303     | c.766G>A       | p.A256T           | MISSENSE |
| IGR-1      | Skin        | Malignant melanoma                    | Heterozygous| 3   | 52676038 | C         | G          | ENST00000337303     | c.1019G>C      | p.G340A           | MISSENSE |
| A388       | NS          | Carcinoma                             | Heterozygous| 3   | 52643864 | G         | A          | ENST00000337303     | c.2032C>T      | p.T678C           | MISSENSE |
| LC4-1      | Haematopoietic and lymphoid tissue | Haematopoietic leukenoma, acute lymphoblastic leukaemia | Heterozygous| 3   | 52637638 | T         | C          | ENST00000337303     | c.2678A>G      | p.Y993C           | MISSENSE |
| TCGA ID | Tissue Type | Tumor Type | Genomic Change | Gene Symbol | Coding Change | Mutation Type |
|---------|-------------|------------|----------------|-------------|--------------|--------------|
| SBC-1   | Lung        | Small cell carcinoma | Homozygous | 3 | 52637638 T C | ENST00000337303 c.2678A>G | p.Y893C | MISSENSE |
| SNU-449 | Liver       | Hepatocellular carcinoma | Heterozygous | 3 | 52637638 T C | ENST00000337303 c.2678A>G | p.Y893C | MISSENSE |
| NCI-H446 | Lung        | Small cell carcinoma | Heterozygous | 3 | 52637633 T A | ENST00000337303 c.2663A>T | p.Y893S | MISSENSE |
| ZR-75-30 | Breast     | Ductal carcinoma | Heterozygous | 3 | 52637552 C G | ENST00000337303 c.2764G>C | p.E922Q | MISSENSE |
| HCC2998 | Liver       | Small cell carcinoma | Heterozygous | 3 | 52637543 T G | ENST00000337303 c.2773A>C | p.K925Q | MISSENSE |
| NCI-SNU-1 | Stomach   | Carcinoma | Heterozygous | 3 | 52589082 C T | ENST00000337303 c.3859G>C | p.E1287Q | MISSENSE |
| MDA-MB-415 | Breast   | Carcinoma | Homozygous | 3 | 52589082 C T | ENST00000337303 c.3859G>C | p.E1287Q | MISSENSE |
| RKO      | Large intestine, colon | Carcinoma | Heterozygous | 3 | 52637638 T C | ENST00000337303 c.3859G>C | p.E1287Q | MISSENSE |
| HCE-T    | Upper aerodigestive tract, sinonasal and nasal cavity, sinus | Squamous cell carcinoma | Heterozygous | 3 | 52637638 T C | ENST00000337303 c.3859G>C | p.E1287Q | MISSENSE |
Supplemental Table X. Deregulated Gene Sets in PBRM1 Knockdown Cellines

| Gene Set Descriptiona,b | Significance of enrichment |
|-------------------------|---------------------------|
|                         | 786-O | SN12C | TK10 |
| CHROMOSOME_INSTABILITY - PMID: 16921376 | 0.0001 | 0.0001 | 0.0001 |
| ZHAN_MM_CD138_PR_VS_REST | 0.0002 | 0.0004 | 0.0001 |
| IDX_TSA_UP_CLUSTERS3 | 0.0001 | 0.0022 | 0.0001 |
| SERUM_FIBROBLAST_CELLCYCLE | 0.0001 | 0.0301 | 0.0002 |
| DOX_RESIST_GASTRIC_UP | 0.0001 | 0.0332 | 0.0003 |
| P21_P53_ANY_DN | 0.0054 | 0.0032 | 0.0001 |
| CROONQUIST_IL6_STARVE_UP | 0.0058 | 0.0078 | 0.0001 |
| CROONQUIST_IL6_RAS_DN | 0.0025 | 0.0322 | 0.0001 |
| DNA_REPLICATION_REACTOME | 0.0060 | 0.0418 | 0.0002 |
| ADIP_DIFF_CLUSTER4 | 0.0038 | 0.0083 | 0.0016 |
| GAY_YY1_DN | 0.0011 | 0.0297 | 0.0028 |
| HSA00240_PYRIMIDINE_METABOLISM | 0.0015 | 0.0038 | 0.0189 |
| LEE_TCELLS3_UP | 0.0377 | 0.0161 | 0.0002 |
| OLDAGE_DN | 0.0001 | 0.0287 | 0.0460 |
| PYRIMIDINE_METABOLISM | 0.0012 | 0.0153 | 0.0196 |
| IGF_VS_PDGF_DN | 0.0016 | 0.0267 | 0.0295 |
| GOLDRATH_CELLCYCLE | 0.0087 | 0.0262 | 0.0096 |
| P21_ANY_DN | 0.0160 | 0.0173 | 0.0098 |
| P21_P53_MIDDLE_DN | 0.0043 | 0.0249 | 0.0412 |
| GERY_CEBP_TARGETS | 0.0085 | 0.0190 | 0.0382 |

Pathways were obtained from the MSigDB with the exception of the CHROMOSOME_INSTABILITY gene set that was obtained from the BioConductor PGSEA package.

Comparisons between PBRM1 targeting and scrambled siRNA.
| Sample   | Chromosome | Position | WT base | Mut Base | Gene       | Annotated Transcript | Type               | Protein Annotation | cDNA Annotation |
|----------|------------|----------|---------|----------|------------|----------------------|--------------------|--------------------|------------------|
| PD2125a  | 1          | 14507757 | G       | A        | PDE4DIP    | ENST00000369359      | MISSENSE           | p.P30S            | c.88C>T          |
| PD2125a  | 4          | 19090439 | T       | C        | AF146191.1 | ENST00000248151      | SILENT             | p.A196A           | c.588A>G         |
| PD2126a  | 11         | 76954792 | G       | T        | GDPD4      | ENST00000376217      | MISSENSE           | p.N396K           | c.1188C>A        |
| PD2127a  | 2          | 10847927 | A       | G        | RGPD4      | ENST00000408999      | SILENT             | p.P781P           | c.2343A>G        |
| PD2127a  | 6          | 31238942 | G       | T        | HLA-C      | ENST00000423188      | SILENT             | p.A176E           | c.527C>A         |
| PD2127a  | 11         | 76954792 | G       | T        | GDPD4      | ENST00000376217      | MISSENSE           | p.N396K           | c.1188C>A        |
| PD2127a  | 12         | 66679    | T       | C        | WASH4P     | ENST00000326592      | SILENT             | p.P260P           | c.780A>G         |
| PD2144a  | 1          | 16973996 | A       | G        | RP5-1182A14.1 | ENST00000334429    | MISSENSE           | p.I97V            | c.289A>G         |
| PD2144a  | 9          | 33385235 | T       | G        | AQP7       | ENST00000379507      | MISSENSE           | p.Y265S           | c.794A>C         |
| PD2147a  | 4          | 79792166 | C       | G        | POTED      | ENST00000299443      | MISSENSE           | p.E172Q           | c.514G>C         |
| PD2147a  | 5          | 14056385 | T       | C        | PCDHB16    | ENST00000361016      | SILENT             | p.T573T           | c.1719T>C        |
| PD2147a  | 9          | 33386167 | G       | C        | AQP7       | ENST0000047660       | MISSENSE           | p.Q13E            | c.37C>G          |
| PD2147a  | 9          | 67968287 | G       | T        | ANKRD20A3  | ENST00000377477      | MISSENSE           | p.A616S           | c.1846G>T        |
| PD2147a  | 9          | 67968295 | G       | T        | ANKRD20A3  | ENST00000377477      | SILENT             | p.S618S           | c.1854C>T        |
| PD2147a  | 11         | 66444273 | T       | A        | RBM4B      | ENST00000310046      | MISSENSE           | p.E93V            | c.278A>T         |
| PD2147a  | 17         | 45219620 | G       | A        | CDC27      | ENST00000066544      | SILENT             | p.A451A           | c.1353C>T        |
| PD3295a  | 10         | 51225281 | C       | G        | AGAP8      | ENST00000420018      | MISSENSE           | p.R100C           | c.298C>T         |
| PD3295a  | 17         | 14095383 | T       | A        | COX10      | ENST00000261643      | MISSENSE           | p.L258H           | c.773T>A         |
| PD3441a  | 17         | 16068340 | C       | T        | NCCR1      | ENST00000417028      | MISSENSE           | p.E50K            | c.148G>A         |
| STS          | Forward Primer        | Reverse Primer         |
|--------------|-----------------------|------------------------|
| stCE03-616895 | AAACAAGGAAGTCCAGGGC   | AAAAAATGGAGATGGCCTTGC  |
| stCE03-616896 | TTGGAAGCGGAGTTTGGGA  | GGCACAGGTGTCAGACGAT    |
| stCE03-616897 | TTTTCTGCTGCTCTTACCT   | GTTTCAAGCAGACTTGGTTAG  |
| stCE03-616898 | CCCCCTAGTACTGAGTGCTG  | ATCTTCCTTCGCTGTTCCAA  |
| stCE03-616899 | CCCAAAATGTGACCTTCTG   | AAGAGATTTCAATTGGTCTTACC|
| stCE03-616900 | AAGATATTTCCATGATTTTA  | AAAAAAGCACAATACCTACAG  |
| stCE03-616901 | CCATAGCAGCAGAAGGTAGGC | AACATGCAGAAAGACTCCAAA  |
| stCE03-616902 | GAATACTGCTGAGAATTTTGT | TGGAGATTGACTTTAAAGTTGTC |
| stCE03-616903 | AAGTACGTCTCAAGGTCCAGA | TAAAAATCATGAGATGTCAGTC |
| stCE03-616904 | GGTCTGTTTTGAATTAGCCTCA | CAACATCCTCCTTTGAACTTATT |
| stCE03-616905 | ATGTCTGTAGTATATTGAATT | TATATCGAAAGATCGTACCA |
| stCE03-616906 | TTAAGCTTGTTGTTAAAGGAGCA | AAAAAGCTTCCACTACAGCTTGA |
| stCE03-616907 | AAACATCTCGTGTTGTTGTTG | GTTTGTTTTGCTAAGGTTT |
| stCE03-616908 | CCTCTTACATGCTGTTTGGTC | CAGGAAAATACGAAATCTCTTT |
| stCE03-616909 | TGCCTTTGAGTTGCTGTTGTC | AAGCTTCCCTAAACCTACCTACTCC |
| stCE03-616910 | ATTGTCCTGAGTTGCTTGAAC | TATATGCAATCCCTACCTCTTC |
| stCE03-616911 | GTAAGATGTATTTTGGAAGCTTGT | AGACATTTCTCTCAACCTACCTACTC |
| stCE03-616912 | GGAACGTTTATCTTTATAATGTA | GGAACGTTTCTGTTCATTGACTGC |
| stCE03-616913 | CAACTTGCCGAAAAGATTCAGTCTTCA | TTCCATCTCATGGCTGACTC |
| stCE03-616914 | GTGTTCCTGGCTTCTGAAAAA | TTGAAATGTTAGATTATTTGAGG |
| stCE03-616915 | TTTGGAATGTAGGTTTATAATGATC | GGTTCAGGGTTCTGTTGAAGGC |
| stCE03-616916 | CCAAACTTGCTGATTTTTGGA | TATATGCAATCCCTACCTCTTC |
| stCE03-616917 | CCTTAAATCTTCCAGCAGATGTT | CTATAGTACCCCTCTCAGC |
| stCE03-616918 | GGTAAAACCATACAAAAGAAGGA | GGAACGTTTCTGTTCATTGACTGC |
| stCE03-616919 | GGAACAATGTTATTTGGAAGATC | GGAACGTTTCTGTTCATTGACTGC |
| stCE03-616920 | GTGTTCCTGGCTTCTGAAAAA | TTGAAATGTTAGATTATTTGAGG |
| stCE03-616921 | CAGGACTTTTGTAAAACCTTGTGGAAGGAGGAC | GGTTCAGGGTTCTGTTGAAGGC |
| stCE03-616922 | CAGGACTTTTGTAAAACCTTGTGGAAGGAGGAC | GGTTCAGGGTTCTGTTGAAGGC |
| stCE03-616923 | GTACAGGCTTCTCTGGTCGTCTC | GGTTCAGGGTTCTGTTGAAGGC |
| stCE03-616924 | GTATAGTACCCCTCTCAGC | CAGGACTTTTGTAAAACCTTGTGGAAGGAGGAC |
| stCE03-616925 | GTAAGATGTATTTTGGAAGCTTGT | AGACATTTCTCTAACCTACCTACTC |
| stCE03-616926 | GTATAGTACCCCTCTCAGC | CAGGACTTTTGTAAAACCTTGTGGAAGGAGGAC |
| stCE03-616927 | GTAAGATGTATTTTGGAAGCTTGT | AGACATTTCTCTAACCTACCTACTC |
| stCE03-616928 | GTATAGTACCCCTCTCAGC | CAGGACTTTTGTAAAACCTTGTGGAAGGAGGAC |