**CCR-14-0330R1:** **SUPPLEMENTARY INFORMATION**

**Concurrent Alterations in *TERT*, *KDM6A*, and the BRCA Pathway in Bladder Cancer**

Michael L. Nickerson1, Garrett M. Dancik2,^, Kate M. Im1, Michael G. Edwards3, Sevilay Turan1, Joseph Brown4, Christina Ruiz-Rodriguez1, Charles Owens2, James C. Costello5, Guangwu Guo6, Shirley X. Tsang6, Yingrui Li6, Quan Zhou6, Zhiming Cai7, Lee E. Moore8, M. Scott Lucia9, Michael Dean1, and Dan Theodorescu2,5,10,\*

**SUPPLEMENTARY METHODS**

**NGS and Variant Analysis and Annotation**

Raw image files were processed using Illumina base calling software 1.7, aligned to the human genome (hg19) reference sequence using BWA ([1](#_ENREF_1)). Duplicate reads were removed by SAMtools ([2](#_ENREF_2)). Single nucleotide variants and short indels were identified using SOAPsnp ([3](#_ENREF_3)) and SAMtools, respectively, and were annotated using ANNOVAR([4](#_ENREF_4)). SNPs with a base quality of at least 20 and a depth of coverage of at least 4 reads were included. Variants in segmental duplications and in >3 samples were removed as likely germline or false positives. Synonymous, intronic and 3’ UTR variants were assessed for potential splicing functions using the Genomic Workbench (CLC Bio, Cambridge, MA)([5](#_ENREF_5)). Primers for selected variants were designed using Primer 3 ([6](#_ENREF_6)) and ExonPrimer (UCSC Genome Browser) and synthesized by Invitrogen (Carlsbad, CA). Variants were annotated using the UCSC Genome Browser, ANNOVAR ([4](#_ENREF_4)), UniProt (The UniProt Consortium, 2010), SIFT ([7](#_ENREF_7)), POLYPHEN ([8](#_ENREF_8)), COSMIC ([9](#_ENREF_9)), and published literature.

**Gene Expression Datasets**

Transcriptome sequencing (RNA-seq) was performed on 43 of the DNA-sequenced tumors with confirmed somatic mutations (BGI cohort([10](#_ENREF_10))). Additional gene expression datasets were obtained for network analysis, mutation signature validation, and evaluation in patients. Cell line datasets include: HEK293 kidney cell lines transfected with siRNA against *TP53* or *lacZ* (negative control, Elkon set([11](#_ENREF_11))); H1299 non-small lung cancer cell lines transfected with mutant KDM6A (Chen set, ArrayExpress database([12](#_ENREF_12)), accession number E-MTAB-84); mammary epithelial cells expressing *HRAS* or green fluorescent protein (GFP control, Bild set([13](#_ENREF_13))). Patient datasets include: 93 primary bladder tumors and 38 normal bladder samples collected at Memorial Sloan-Kettering Cancer Center (MSKCC cohort([14](#_ENREF_14))), 165 primary bladder tumors, 10 normal bladder samples, and 22 recurrent tumors collected at Chungbuk National University Hospital (CNUH cohort([15](#_ENREF_15))), 144 primary bladder tumors and 12 normal bladder samples collected at the University Hospital of Lund, Sweden (Lindgren cohort([16](#_ENREF_16))), 404 primary bladder tumors from patients treated in hospitals in Denmark, Sweden, Spain, France, and England (Dyrskjot cohort([17](#_ENREF_17))), 90 primary lung adenocarcinomas with known *TP53* mutation status (Takeuchi cohort([18](#_ENREF_18))), and 116 primary lung adenocarcinomas with known *TP53* mutation status (Tomida cohort([19](#_ENREF_19))). We used the publicly available processed data for each dataset. Tumors were profiled in duplicate in the Dyrskjot cohort and these duplicates were averaged to produce a single gene expression profile for each tumor. Missing values were imputed in the Tomida cohort using the *impute* package (impute.knnfunction) in *R* with default parameters([20](#_ENREF_20)). The gene expression datasets are summarized in **Supplementary Table S9**. P-values were calculated by the non-parametric Wilcoxon rank-sum test and each p-value adjusted to obtain the false discovery rate ([21](#_ENREF_21)). When multiple microarray probes matched to a single gene (Affymetrix or Illumina annotation), the probe with the highest mean expression was selected for the *KDM6A* signature score ([22](#_ENREF_22)).

***KDM6A* Plasmids, Transfection, and RT-PCR**

Short hairpin RNA (shRNA) sequence 5’-CCGGGATGCAAGTCTATGACCAATTCTCGAGAA-TTGGTCATAGACTTGCATCTTTTTG-3' in pLKO.1-puro plasmid was used for human *KDM6A* depletion (TRCN0000107763, Sigma-Aldrich, St. Louis, MO, USA) in MGHU3 cells. The shRNA control non-targeting plasmid pLKO.1-puro was used for the shRNA control (SHC002, Sigma-Aldrich). Mammalian expression vectors containing *KDM6A* or empty vector (control) were constructed using a modified Gateway Multisite Recombination system ([Life Technologies](http://www.lifetechnologies.com/us/en/home.html), Carlsbad, CA) by the Protein Expression Laboratory (SAIC-Frederick, National Cancer Institute, Frederick, MD, USA). N-terminal FLAG-tagged human *KDM6A* plasmid (11648-X06-515) included a cytomegalovirus promoter for *KDM6A* over-expression in T24T cells. Empty plasmid vector (pEL124-490) was used as a control in *KDM6A* over-expression experiments. Transfection-ready DNA was prepared using the GenElute XP Maxiprep kit (Sigma-Aldrich) and verified by sequencing, restriction digest and agarose gel electrophoresis.

*KDM6A* depletion in MGHU3 cells and *KDM6A* over-expression in T24T cells, two frequently utilized bladder tumor-derived cell lines ([23](#_ENREF_23)), was validated by quantitative RT-PCR using the SYBR Green SuperMix protocol with standard curves on an iQ5 Cycler (Bio-Rad Laboratories, Hercules, CA, USA). Standard curves were generated using mRNA from *KDM6A* T24T over-expressing cells. RNA was harvested from the cell lines using the Qiagen RNEasy Plus kit (Qiagen, Valencia, CA, USA) and converted to cDNA using the IScriptcDNA Synthesis kit (Bio-Rad Laboratories, Hercules, CA, USA). Quantitative RT-PCR reactions were performed to amplify *KDM6A* using the Primetime qPCR primers: forward 5′-TGGAAACGTGCCTTACCTG-3′ and reverse 5′- TGCCGAATGTGAACTCTGAC-3′ (Integrated DNA Technologies, Coralville, IA, USA). *GAPDH* was used as the internal control housekeeping gene for quantitative RT-PCR reactions with primers: forward 5’-TCTTTTGCGTCGCCAGCCGA-3’ and reverse 5’-ACCAGGCGCCCAATACGACC-3’. For determination of *KDM6A* expression, the ΔΔCT method was used. Expression was normalized to the appropriate control cells.

***In Vitro* and *In Vivo* Growth**

Anchorage independent growth was assessed by plating 8 x 103 cells in 1.2 mL media with 0.4% SeaPlaque low melting temperature agarose (Lonza, Rockland, ME, USA) in three times in duplicate wells in 12-well plates. Colonies formed were stained with Nitro-BT (Sigma-Aldrich, St. Louis, MO, USA) and counted using ImageJ software (Rasband, W.S., ImageJ, U. S. National Institutes of Health, Bethesda, Maryland, USA, http://imagej.nih.gov/ij/, 1997-2012). Cell numbers were assessed by plating 1 x 103 cells per well in 96 welled plates in quadruplicate for proliferation studies. Cell numbers was determined by CyQUANT® Assay (Life Technologies Corporation, Carlsbad, CA, USA). Cell migration was determined by plating 2 x 104 cells in quadruplicate to the upper chambers of transwell filters with 8.0 µm pores (Becton Dickinson, Franklin Lakes, NJ, USA) in a 24-well tissue culture plate. The lower chambers contained media with 2% FBS. Plating control assays were done in quadruplicate containing the same media but with no transwell filters. After 18 hours, cells remaining on the upper surface of the filters were removed with cotton swabs and cells on the lower surface were fixed with 100% methanol, stained with crystal violet, and counted using ImageJ software. For *in vivo* assessment, 5-week-old male NCr*nu*/*nu* mice (NCI-Frederick, Frederick, MD, USA) were injected with 2 x 106 cells stably expressing *KDM6A* shRNA, or non-target control shRNA in the right and left flanks (2 sites/flank) of each mouse for subcutaneous tumor growth. Animals were maintained according to University of Colorado IACUC guidelines. Tumors were measured and tumor volumes calculated as described previously ([24](#_ENREF_24))

**SUPPLEMENTARY FIGURE LEGENDS**

**Supplementary Figure S1.**

The ubiquitin C (UBC) network in bladder cancer. The most significant network (Fisher’s Exact Test, P<10-68) constructed by IPA for genes with confirmed somatic mutations in bladder cancer. The number of independently identified mutations is listed above each respective gene, and all of the transcribed proteins from these genes are known to bind the UBC protein (solid lines). Colored genes indicate decreased (green) or increased (red) gene expression in tumors compared to normal bladder samples in at least two out of three patient cohorts (FDR < 5%).

**Supplementary Figure S2.**

Bladder cancer network containing debiquitinating enzymes (DUBs), chromosome remodeling genes and genes associated with VEGF signaling. The second most significant network (Fisher’s Exact Test, P<10-46) constructed by IPA using genes with confirmed somatic mutations in bladder cancer. The number of independently identified mutations is listed above each respective gene, and colored genes indicate decreased (green) or increased (red) gene expression in tumors compared to normal bladder samples in at least two out of three patient cohorts (FDR < 5%).

**SUPPLEMENTARY TABLES**

**Supplementary Table S1. Clinical characteristics of U.S. patients and tumors (N=54).\***

\*Includes 14 tumors for discovery exome sequencing (discovery panel, samples 1-14) and 40 tumors for *BAP1* mutation frequency determination by PCR and Sanger sequencing (validation panel, samples 15-54).

| **Sample****Count** | **Sex** | **Age****(years)** | **Stage****(TNM)** | **Matched****Normal** | **Grade****(WHO)** | **Analysis****Type** |
| --- | --- | --- | --- | --- | --- | --- |
| 1 | M | 56 | T4N0M0 | Yes | 3 | Exome |
| 2 | M | 77 | T3N0M0 | Yes | 3 | Exome |
| 3 | M | 81 | T3N0M0 | Yes | 3 | Exome |
| 4 | M | 71 | T4N0M0 | Yes | 3 | Exome |
| 5 | M | 68 | TaN0M0 | Yes | 1 | Exome |
| 6^ | M | 73 | TaN0M0 | Yes | 2 | Exome |
| 7 | M | 59 | TaN0M0 | Yes | 1 | Exome |
| 8 | M | 66 | TaN0M0 | Yes | 1 | Exome |
| 9 | M | 73 | TaN0M0 | Yes | 1 | Exome |
| 10 | M | 65 | T2N0M0 | Yes | 3 | Exome |
| 11 | M | 65 | T3N0M0 | Yes | 3 | Exome |
| 12 | M | 61 | T4N0M0 | Yes | 3 | Exome |
| 13 | M | 76 | T2N0M0 | Yes | 3 | Exome |
| 14 | M | 85 | T3N0M0 | Yes | 3 | Exome |
| 15 | M | 77 | T3N0M0 | Yes | 3 | Validation |
| 16 | M | 75 | T2N0M0 | Yes | 1 | Validation |
| 17 | M | 55 | T4N3M0 | Yes | 3 | Validation |
| 18 | M | 70 | T2N0M0 | Yes | 3 | Validation |
| 19 | M | 64 | T3N0M0 | Yes | 3 | Validation |
| 20 | F | 73 | T3N2M0 | Yes | 3 | Validation |
| 21 | M | 70 | T4N2M0 | Yes | 2 | Validation |
| 22^ | M | 66 | T3N0M0 | Yes | 3 | Validation |
| 23 | M | 71 | T1NxM0 | Yes | 3 | Validation |
| 24 | F | 63 | T3N0M0 | Yes | 3 | Validation |
| 25 | F | 78 | T3N0M0 | Yes | 3 | Validation |
| 26 | M | 57 | T3N0M0 | Yes | 3 | Validation |
| 27 | M | 50 | T3N2M0 | Yes | 3 | Validation |
| 28 | M | 54 | T2N0M0 | Yes | 3 | Validation |
| 29 | M | 57 | T4N0M0 | Yes | 3 | Validation |
| 30 | M | 80 | T3N1M0 | Yes | 2 | Validation |
| 31 | F | 53 | T3N0M0 | Yes | 3 | Validation |
| 32 | M | 65 | T4N2M0 | Yes | 3 | Validation |
| 33 | M | 73 | T4N0M0 | Yes | 3 | Validation |
| 34 | M | 42 | T4N3M0 | Yes | 3 | Validation |
| 35^ | M | 73 | T4N0M0 | Yes | 3 | Validation |
| 36 | F | 61 | T3N0M0 | Yes | 3 | Validation |
| 37 | M | 63 | T1N0M0 | Yes | 3 | Validation |
| 38 | M | 74 | T4N0M0 | Yes | 3 | Validation |
| 39 | M | 58 | T2N0M0 | Yes | 3 | Validation |
| 40 | M | 58 | T3N1M0 | Yes | 3 | Validation |
| 41 | M | 57 | T3N3M0 | Yes | 3 | Validation |
| 42 | M | 62 | TaN0M0 | No | 2 | Validation |
| 43 | M | 89 | TaN0M0 | Yes | 3 | Validation |
| 44 | M | 46 | T1N0M0 | Yes | 3 | Validation |
| 45 | M | 52 | TaN0M0 | Yes | 3 | Validation |
| 46 | F | 67 | T1N0M0 | Yes | 3 | Validation |
| 47 | M | 71 | T2N0M0 | Yes | 1 | Validation |
| 48^ | M | 77 | T1N0M0 | Yes | 3 | Validation |
| 49 | M | 69 | T1N0M0 | Yes | 2 | Validation |
| 50 | F | 81 | T1N0M0 | Yes | 3 | Validation |
| 51 | F | 78 | TaN0M0 | No | 2 | Validation |
| 52 | M | 75 | TaN0M0 | Yes | 2 | Validation |
| 53 | M | 71 | T2N0M0 | Yes | 3 | Validation |
| 54 | M | 68 | T1N0M0 | Yes | 3 | Validation |

Exome, whole exome sequencing; Validation, PCR and Sanger sequencing; Race, self-identified as Caucasian or ^, indicating African American.

**Supplementary Table S2. Exome sequencing statistics for 14 bladder tumors.**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Sample ID** | **Sample Type** | **Total Reads** | **# reads uniquely mapped to target region** | **Mean fold coverage on target region** | **% of targets covered by at least 4×** | **% of targets covered by at least 10×** |
| 1 | Tumor | 71,373,144 | 59,443,722 | 87.13 | 96.1% | 90.8% |
| 2 | Tumor | 84,940,294 | 69,990,721 | 102.39 | 96.9% | 92.6% |
| 3 | Tumor | 91,685,548 | 76,536,592 | 108.53 | 96.5% | 91.5% |
| 4 | Tumor | 85,529,556 | 69,848,114 | 102.76 | 97.1% | 93.4% |
| 5 | Tumor | 70,640,930 | 59,376,408 | 82.02 | 95.7% | 89.7% |
| 6 | Tumor | 65,128,006 | 54,664,239 | 81.54 | 96.1% | 90.4% |
| 7 | Tumor | 74,805,560 | 62,297,988 | 93.06 | 96.8% | 92.3% |
| 8 | Tumor | 73,655,412 | 60,764,000 | 90.17 | 96.8% | 92.6% |
| 9 | Tumor | 81,917,930 | 66,484,801 | 97.54 | 95.6% | 89.8% |
| 10 | Tumor | 70,783,582 | 59,088,253 | 86.35 | 96.3% | 91.0% |
| 11 | Tumor | 83,159,866 | 67,283,208 | 97.65 | 97.9% | 95.0% |
| 12 | Tumor | 81,742,674 | 67,498,718 | 98.52 | 96.1% | 90.8% |
| 13 | Tumor | 70,238,814 | 58,057,487 | 84.99 | 94.5% | 87.8% |
| 14 | Tumor | 59,467,792 | 49,948,318 | 72.77 | 96.0% | 90.3% |

**Supplementary Table S3**. **Somatic variants confirmed by Sanger sequencing of tumor-normal DNA in 14 exomes and 40 validation tumors.**

| **Gene** | **Sample ID #** | **Genomic** | **Nucleotide****(cDNA)** | **Protein** | **Mutation Type** | **dbSNP137** | **Status** | **RSI** |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| *ABCC12* | 10 | g.chr16:48180307G>C | c.C29G | p.S10X | nonsense | . | Somatic |  |
| *ANK3* | 3 | g.chr10:61835330G>A | c.C5309T | p.T1770M | missense | . | Somatic | 0.4 |
| *ARID1A* | 10 | g.chr1:27092809C>T | c.C2830T | p.Q944X | nonsense | . | Somatic | 0.4 |
| *ARID1A* | 10 | g.chr1:27101683G>C | c.G4314C | p.Q1438H | missense | . | Somatic |  |
| *ARID1A* | 5 | g.chr1:27088714C>T | c.C2323T | p.Q775X | nonsense | . | Somatic | 0.4 |
| *ARID1A* | 4 | g.chr1:27106794insT | c.5754\_5755insT | p.I1918fs | FS ins | . | Somatic | 0.4 |
| *ARID1A* | 24 | g.chr1:27094322delGTTGTATGAG | c.3030\_3039del | p.1009fs | FS del | . | Somatic | 0.5 |
| *ARID1A* | 29 | g.chr1:27101112C>G | c.C4394G | p.S1465C | missense | . | Somatic | 0.4 |
| *ARID1A* | 29 | g.chr1:27094291C>T | c.C2999T | p.S1000F | missense | . | Somatic | 0.3 |
| *ARID1A* | 16 | g.chr1:27105528delAGAGCTCCTTGTAGA | c.5139\_5153del | p.1713\_1717delEELVE | NFS del | . | Somatic | 0.5 |
| *ARID1A* | 16 | g.chr1:27057727C>T | c.C1435T | p.Q479X | nonsense | . | Somatic | 1.0 |
| *ARID1A* | 40 | g.chr1:27101683G>C | c.G4965C | p.Q1655H | missense | . | Somatic | 0.7 |
| *ARID1A* | 54 | g.chr1:27101383T>C | c.T4665C | p.Y1555Y | synonymous | . | Somatic | 1.0 |
| *ARID1A* | 19 | g.chr1:27102066A>T | c.4994-2A>T | . | splicing | . | Somatic | 0.5 |
| *ARID2* | 10 | g.chr12:46231391G>T | c.G1231T | p.D411Y | missense | . | Somatic | 0.4 |
| *ARID3A* | 12 | g.chr19:964909G>C | c.G1027C | p.D343H | missense | . | Somatic | 0.4 |
| *ARID3A* | 7 | g.chr19:966584C>T | c.C1211T | p.A404V | missense | rs145850291 | Somatic | 0.7 |
| *ARID4B* | 2 | g.chr1:235346068C>G | c.G1908C | p.E636D | missense | . | Somatic | 0.4 |
| *ARID5A* | 10 | g.chr2:97217076G>C | c.G811C | p.E271Q | missense | . | Somatic | 0.4 |
| *ASPM* | 12 | g.chr1:197070880G>T | c.C7501A | p.Q2501K | missense | . | Somatic | 0.4 |
| *ATM* | 9 | g.chr11:108175401G>C | c.5497-1G>C | . | splicing | . | Somatic | 0.7 |
| *ATM* | 9 | g.chr11:108200976A>G | c.A7343G | p.D2448G | missense | . | Somatic | 0.4 |
| *ATM* | 10 | g.chr11:108127058G>C | c.G2241C | p.Q747H | missense | . | Somatic | 0.5 |
| *ATM* | 11 | g.chr11:108216590G>C | c.G8539C | p.E2847Q | missense | . | Somatic | 0.5 |
| *ATM* | 12 | g.chr11:108165783C>T | c.C4906T | p.Q1636X | nonsense | . | Somatic | 0.4 |
| *ATM* | 11 | g.chr11:108098576C>G | c.C146G | p.S49C | missense | rs1800054 | Somatic | 0.5 |
| *BAP1* | 1 | g.chr3:52441264T>C | c.A506G | p.H169R | missense | . | Somatic | 0.4 |
| *BAP1* | 9 | g.chr3:52441263G>C | c.C507G | p.H169Q | missense | . | Somatic | 0.7 |
| *BAP1* | 5 | g.chr3:52442575C>T | c.G170A | p.R57Q | missense | . | Somatic | 0.4 |
| *BAP1* | 5 | g.chr3:52443864C>T | c.G31A | p.D11N | missense | . | Somatic | 0.5 |
| *BAP1* | 15 | g.chr3:52440346C>G | c.G821C | p.D236H | missense | . | Somatic | 0.9 |
| *BAP1* | 24 | g.chr3:52436342C>G | c.C2152G | p.R718G | missense | . | Somatic | 0.3 |
| *BAP1* | 31 | g.chr3:52437808G>A | c.G1353T | p.K451K | synonymous | . | Somatic | 0.2 |
| *BAP1* | 42 | g.chr3:52436342C>G | c.C2152G | p.R718G | missense | . | Confirmed | 0.3 |
| *BAP1* | 48 | g.chr3:52440283G>T | c.G769A | p.E257X | nonsense | . | Somatic | 1.0 |
| *BRCA1* | 8 | g.chr17:41243691C>G | c.G3857C | p.S1286T | missense | rs142383077 | Somatic | 0.4 |
| *BRCA1* | 3 | g.chr17:41246458G>C | c.C1090G | p.P364A | missense | . | Somatic | 0.4 |
| *BRCA2* | 11 | g.chr13:32954222C>T | c.C9196T | p.Q3066X | nonsense | rs80359180 | Somatic | 0.4 |
| *BRD1* | 11 | g.chr22:50217330G>T | c.C636A | p.D212E | missense | . | Somatic |  |
| *BRPF1* | 7 | g.chr3:9785273G>C | c.G2305C | p.E769Q | missense | . | Somatic |  |
| *CABIN1* | 3 | g.chr22:24573783-24573806delAAGGTGACCTCAGGGGCTGGGCTG | c.6517\_6519del | p.2173\_2173del | NFS del | . | Somatic |  |
| *CDKN2B* | 13 | g.chr9:22008730C>G | c.G223C | p.G75R | missense | . | Somatic | 0.4 |
| *CHD1* | 3 | g.chr5:98192167G>A | c.C5050T | p.P1684S | missense | rs61749618 | Somatic | 0.4 |
| *CHD1* | 5 | g.chr5:98192167G>A | c.C5050T | p.P1684S | missense | rs61749618 | Somatic | 0.4 |
| *CHD1* | 7 | g.chr5:98236560C>A | c.G814T | p.E272X | nonsense | . | Somatic | 0.7 |
| *CHD1* | 7 | g.chr5:98236575C>T | c.G799A | p.E267K | missense | . | Somatic | 0.4 |
| *CHD1L* | 11 | g.chr1:146758138C>T | c.C2182T | p.Q728X | nonsense | . | Somatic | 0.4 |
| *CHD4* | 11 | g.chr12:6682300G>C | c.C5497G | p.Q1833E | missense | . | Somatic | 0.6 |
| *COL6A3* | 8 | g.chr2:238296788G>C | c.C749G | p.S250X | nonsense | . | Somatic |  |
| *COL6A3* | 12 | g.chr2:238275420C>T | c.G5410A | p.E1804K | missense | . | Somatic |  |
| *CREBBP* | 2 | g.chr16:3832796G>C | c.C1348G | p.L450V | missense | . | Somatic | 0.4 |
| *CREBBP* | 4 | g.chr16:3781465-3781486delCGAAGAAGACCTGCAGGAGAGG | c.4891\_4900del | p.1631\_1634del | FS del | . | Somatic | 0.4 |
| *CSF1R* | 7 | g.chr5:149437090G>C | c.C2198G | p.S733X | nonsense | . | Somatic |  |
| *DHDH* | 11 | g.chr19:49442851G>A | c.G512A | p.G171E | missense | . | Somatic | 0.4 |
| *DNAH8* | 12 | g.chr6:38866115G>C | c.G8371C | p.D2791H | missense | . | Somatic | 0.4 |
| *DNMT1* | 9 | g.chr19:10270719C>G | c.G1016C | p.R339P | missense | . | Somatic | 0.5 |
| *DNMT3A* | 10 | g.chr2:25458613C>A | c.G2560T | p.E854X | nonsense | . | Somatic | 0.7 |
| *DNMT3A* | 11 | g.chr2:25457243G>T | c.C2644A | p.R882S | missense | . | Somatic |  |
| *EP300* | 2 | g.chr22:41548264G>T | c.G3052T | p.E1018X | nonsense | . | Somatic | 0.7 |
| *ERBB3* | 1 | g.chr1:56495001-56495012delAGGAGCCGGAGC | c.3358\_3369del | p.1120\_1123del | NFS del | . | Somatic | 0.6 |
| *ERG* | 12 | g.chr21:39755645C>G | c.G1120C | p.D374H | missense | . | Somatic |  |
| *ERG* | 13 | g.chr21:39755347G>C | c.C1418G | p.S473C | missense | . | Somatic |  |
| *ERG* | 13 | g.chr21:39775589T>C | c.A431G | p.E144G | missense | rs201672509 | Somatic |  |
| *EZH1* | 11 | g.chr17:40870589G>A | c.C814T | p.Q272X | nonsense | . | Somatic | 0.4 |
| *FAT4* | 13 | g.chr4:126337691C>T | c.C6932T | p.S2311L | missense | . | Somatic | 0.4 |
| *FGFR3* | 1 | g.chr4:1803568C>G | c.C746G | p.S249C | missense | rs121913483 | Somatic | 0.4 |
| *FGFR3* | 8 | g.chr4:1803568C>G | c.C746G | p.S249C | missense | rs121913483 | Somatic | 0.4 |
| *FGFR3* | 6 | g.chr4:1806099A>G | c.A1124G | p.Y375C | missense | rs121913485 | Somatic | 0.7 |
| *FLT4* | 7 | g.chr5:180030322C>T | c.G3962A | p.R1321Q | missense | rs79620092 | Somatic |  |
| *GCN1L1* | 1 | g.chr12:120567200G>C | c.C7770G | p.I2590M | missense | . | Somatic | 0.4 |
| *GCN1L1* | 1 | g.chr12:120567248G>C | c.C7722G | p.I2574M | missense | . | Somatic | 0.5 |
| *GCN1L1* | 8 | g.chr12:120597824C>G | c.G2554C | p.D852H | missense | . | Somatic | 0.7 |
| *GCN1L1* | 10 | g.chr12:120582466delG | c.5329delC | p.P1777fs | FS del | . | Somatic | 0.4 |
| *HMCN1* | 10 | g.chr1:185902818G>C | c.G1690C | p.E564Q | missense | . | Somatic | 0.4 |
| *HRAS* | 7 | g.chr11:534289C>T | c.G34A | p.G12S | missense | rs104894229 | Somatic | 1.0 |
| *HTT* | 7 | g.chr4:3132038A>G | c.A1874G | p.Q625R | missense | rs151106561 | Somatic |  |
| *HUWE1* | 5 | g.chrX:53563401C>T | c.G12365A | p.R4122H | missense | . | Somatic |  |
| *HUWE1* | 8 | g.chrX:53589792C>G | c.G7204C | p.D2402H | missense | . | Somatic |  |
| *IFNAR2* | 10 | g.chr21:34621100G>T | c.G481T | p.E161X | nonsense | . | Somatic |  |
| *ITGA10* | 3 | g.chr1:145538740C>T | c.C2851T | p.R951C | missense | rs142228466 | Somatic |  |
| *JMJD1C* | 13 | g.chr10:64967167G>A | c.C3605T | p.S1202L | missense | . | Somatic | 0.4 |
| *JMJD1C* | 13 | g.chr10:64967173G>A | c.C3599T | p.S1200F | missense | . | Somatic | 0.4 |
| *KALRN* | 10 | g.chr3:124379810C>G | c.C6254G | p.S2085X | nonsense | . | Somatic |  |
| *KDM6A* | 9 | g.chrX:44942823T>G | c.T3403G | p.Y1135D | missense | . | Somatic | 0.7 |
| *KDM6A* | 7 | g.chrX:44949073C>T | c.C3634T | p.Q1212X | nonsense | . | Somatic | 0.4 |
| *KDM6A* | 9 | g.chrX:44942823insTAG | c.3403\_3404insTAG | p.Y1135delinsLD | NFS ins | . | Somatic | 0.5 |
| *KDM6A* | 10 | g.chrX:44921990-44921992delCCA | c.1524\_1526del | p.508\_509del | NFS del | . | Somatic | 0.4 |
| *KDM6A* | 14 | g.chrX:44733227-44733235delGGGCAAGGT | c.219\_225del | p.73\_75del | FS del | . | Somatic | 0.4 |
| *KDM6A* | 24 | g.chrX:44870219insT | c.398\_399insT | p.L133fs | frameshift ins | . | Somatic | 0.5 |
| *KDM6A* | 26 | g.chrX:44969465insC | c.4147\_4148insC | p.L1383fs | frameshift ins | . | Somatic | 0.5 |
| *KDM6A* | 16 | g.chrX:44922728delAC | c.1589\_1590del | p.D530fs | frameshift del | . | Somatic | 0.5 |
| *KDM6A* | 42 | g.chrX:44949073C>T | c.C3634T | p.Q1212X | nonsense | . | Confirmed | 1.0 |
| *KDM6A* | 45 | g.chrX:44921885C>T | c.1426-7C>T | . | splicing | . | Somatic | 0.2 |
| *KDM6A* | 48 | g.chrX:44949073C>T | c.C3634T | p.Q1212X | nonsense | . | Somatic | 0.8 |
| *KDM6A* | 51 | g.chrX:44929398C>T | c.C2498T | p.T833I | missense | . | Confirmed | 0.3 |
| *KDM6A* | 22 | g.chrX:44928894insC | c.1994\_1995insC | p.P665fs | FS ins | . | Somatic | 0.5 |
| *LAMA3* | 13 | g.chr18:21437920T>A | c.T4249A | p.C1417S | missense | . | Somatic |  |
| *LRP1B* | 8 | g.chr2:141108467G>A | c.C11791T | p.Q3931X | nonsense | . | Somatic |  |
| *LRP1B* | 8 | g.chr2:141773447C>G | c.G2008C | p.D670H | missense | . | Somatic |  |
| *LRP1B* | 14 | g.chr2:141806752C>T | c.G1592A | p.R531H | missense | . | Somatic |  |
| *LRP2* | 2 | g.chr2:169993931T>A | c.A13591T | p.S4531C | missense | . | Somatic | 0.7 |
| *MDM1* | 7 | g.chr12:68696445G>A | c.C1927T | p.R643X | nonsense | rs14762717 | Somatic |  |
| *MLL* | 12 | g.chr11:118347523C>G | c.C3160G | p.Q1054E | missense | . | Somatic | 0.7 |
| *MLL3* | 6 | g.chr7:151873843G>A | c.C8695T | p.Q2899X | nonsense | . | Somatic |  |
| *MLL5* | 11 | g.chr7:104747026C>G | c.C2654G | p.S885X | nonsense | . | Somatic |  |
| *MLL5* | 11 | g.chr7:104752914C>T | c.C4711T | p.Q1571X | nonsense | . | Somatic | 0.4 |
| *NCOA1* | 3 | g.chr2:24974990G>A | c.G3846A | p.M1282I | missense | . | Somatic |  |
| *NCOR1* | 14 | g.chr17:16004811G>C | c.C2491G | p.P831A | missense | . | Somatic |  |
| *NF1* | 12 | g.chr17:29667550T>G | c.T6886G | p.W2296G | missense | . | Somatic |  |
| *PALB2* | 10 | g.chr16:23641186C>G | c.G2289C | p.L763F | missense | . | Somatic | 0.4 |
| *PALB2* | 10 | g.chr16:23641638G>A | c.C1837T | p.Q613X | nonsense | . | Somatic | 0.4 |
| *PALB2* | 6 | g.chr16:23646264T>A | c.A1603T | p.S535C | missense | . | Somatic | 0.5 |
| *PAX1* | 4 | g.chr20:21689972insC | c.1172\_1173insC | p.P391fs | FS ins |  | Somatic |  |
| *PDZD2* | 2 | g.chr5:32087510G>A | c.G3956A | p.R1319K | missense | . | Somatic |  |
| *RSF1* | 3 | g.chr11:77386155C>T | c.G3488A | p.R1163Q | missense | rs145881006 | Somatic |  |
| *RSF1* | 10 | g.chr11:77388044C>T | c.G3134A | p.G1045E | missense | . | Somatic |  |
| *SETD2* | 10 | g.chr3:47129645A>C | c.T5235G | p.I1745M | missense | . | Somatic | 0.7 |
| *SETDB2* | 7 | g.chr13:50050777C>G | c.C471G | p.I157M | missense | . | Somatic | 0.4 |
| *SIN3A* | 10 | g.chr15:75687030G>C | c.C2268G | p.I756M | missense | . | Somatic |  |
| *STAG2* | 15 | g.chrX:123181311C>T | c.C775T | p.R259X | nonsense | . | Somatic | 0.3 |
| *STAG2* | 16 | g.chrX:123217325A>C | c.A2979C | p.P993P | synonymous | . | Somatic | 0.5 |
| *STAG2* | 41 | g.chrX:123195618insTTT | c.1535-3\_1535-2insTTT | . | splice site | May correspond to rs72250532 | Somatic | 0.4 |
| *STAG2* | 43 | g.chrX:123215311C>T | c.C2857T | p.R953X | nonsense | . | Somatic | 0.9 |
| *STAG2* | 45 | g.chrX:123215311C>T | c.C2857T | p.R953X | nonsense | . | Somatic | 0.9 |
| *STAG2* | 48 | g.chrX:123164976G>T | c.288+1G>T | . | splicing | . | Somatic | 0.7 |
| *STAG2* | 50 | g.chrX:123171381T>G | c.T293G | p.V98G | missense | . | Somatic | 0.5 |
| *STAG2* | 18 | g.chrX:123197031insA | c.1797\_1798insA | p.I599fs | FS ins | . | Somatic | 0.6 |
| *SYNE1* | 1 | g.chr6:152652869C>A | c.G12738T | p.Q4246H | missense | . | Somatic | 0.5 |
| *SYNE2* | 12 | g.chr14:64498043G>C | c.G7189C | p.E2397Q | missense | . | Somatic |  |
| *SYNE2* | 14 | g.chr14:64443292G>A | c.G1140A | p.W380X | nonsense | . | Somatic | 0.4 |
| *TG* | 13 | g.chr8:133980074G>A | c.G5722A | p.E1908K | missense | . | Somatic |  |
| *TP53* | 10 | g.chr17:7578290C>T | c.560-1G>A | . | splicing | . | Somatic | 0.7 |
| *TP53* | 11 | g.chr17:7577517A>G | c.T764C | p.I255T | missense | . | Somatic |  |
| *TP53* | 11 | g.chr17:7577568C>T | c.G713A | p.C238Y | missense | . | Somatic | 0.4 |
| *TP53* | 2 | g.chr17:7578406C>T | c.G524A | p.R175H | missense | rs28934578 | Somatic | 0.7 |
| *TP53* | 3 | g.chr17:7577085C>T | c.G853A | p.E285K | missense | rs112431538 | Somatic | 0.7 |
| *TP53* | 4 | g.chr17:7578550G>A | c.C380T | p.S127F | missense | . | Somatic | 0.9 |
| *TRRAP* | 12 | g.chr7:98503866G>C | c.G1103C | p.R368T | missense | . | Somatic |  |
| *TSC1* | 8 | g.chr9:135804258A>T | c.T2A | p.M1K | missense | . | Somatic | 0.4 |
| *TSC1* | 5 | g.chr9:135797201-135797224delACTTCTTCAAAA | c.492\_510del | p.164\_170del | FS del | . | Somatic | 0.4 |
| *USP38* | 5 | g.chr4:144135290T>A | c.T2161A | p.L721I | missense | . | Somatic |  |

FS, frameshift; NFS, non-FS; UTR, untranslated region; del, deletion; ins, insertion; RSI, relative signal intensity from Sanger sequence chromatograms. Confirmed, variant was confirmed by PCR and Sanger sequencing in tumor DNA but normal DNA was not available. Genomic coordinates HG19 genome build.

**Supplementary Table S4. Germline variants confirmed by PCR and Sanger sequencing of tumor-normal DNA in 14 exomes and 40 validation tumors.**

| **Gene** | **Sample ID #** | **Genomic****(HG19)** | **Nucleotide (cDNA)** | **Protein** | **Mutation Type** | **dbSNP137** | **RSI** |
| --- | --- | --- | --- | --- | --- | --- | --- |
| ***ABCC12*** | 10 | **g.chr16:48130781C>T** | **c.G3071A** | **p.W1024X** | **nonsense** | **rs36102575** |  |
| ***ABCC12*** | 5 | **g.chr16:48174765C>A** | **c.G490T** | **p.G164X** | **nonsense** | **rs141807269** |  |
| *AHNAK* | 10 | g.chr11:62294309C>T | c.G7580A | p.G2527D | missense | rs141117375 |  |
| *AHNAK* | 7 | g.chr11:62298621G>A | c.C3268T | p.P1090S | missense | rs148272375 | 0.7 |
| *AHNAK* | 2 | g.chr11:62294309C>T | c.G7580A | p.G2527D | missense | rs141117375 | 0.5 |
| *AHNAK* | 9 | g.chr11:62303563insCCT | c.8\_9insAGG | p.K3delinsKE | NFS ins | . |  |
| *AHNAK* | 11 | g.chr11:62303563insCCT | c.8\_9insAGG | p.K3delinsKE | NFS ins | . |  |
| ***ALDH1B1*** | 14 | **g.chr9:38396323delG** | **c.578delG** | **p.G193fs** | **FS del** | **.** |  |
| ***ALDH1B1*** | 14 | **g.chr9:38396906delC** | **c.1161delC** | **p.G387fs** | **FS del** | **rs201408956** |  |
| *ANK3* | 10 | g.chr10:62149193C>T | c.G104A | p.R35Q | missense | . |  |
| *ANK3* | 13 | g.chr10:61819543C>A | c.G4981T | p.D1661Y | missense | rs139092048 | 0.7 |
| *ANK3* | 2 | g.chr10:61833684C>T | c.G6955A | p.D2319N | missense | rs140463162 |  |
| *ARID1A* | 4 | g.chr1:27092733G>C | c.G2754C | p.M918I | missense | . | 0.4 |
| *ARID1A* | 4 | g.chr1:27106106G>A | c.G5066A | p.R1689Q | missense | rs41303631 | 0.4 |
| *ARID1A* | 13 | g.chr1:27107263insC | c.\*16\_\*17insC | . | UTR3 | . | 0.7 |
| *ARID1A* | 47 | g.chr1:27087500A>G | c.A2074G | p.I692V | missense | rs146402512 | 0.5 |
| *ARID3B* | 3 | g.chr15:74836434C>T | c.C157T | p.L53F | missense | . |  |
| *ARID4A* | 9 | g.chr14:58830978A>G | c.A2171G | p.N724S | missense | rs2230098 |  |
| *ARID4B* | 5 | g.chr1:235345986A>C | c.T1990G | p.S664A | missense | . |  |
| *ARID5B* | 8 | g.chr10:63661474G>T | c.G6T | p.E2D | missense | . | 0.7 |
| *ASPM* | 6 | g.chr1:197099053T>C | c.A2621G | p.Y874C | missense | . |  |
| *ASPN* | 12 | g.chr9:95228676C>G | c.G565C | p.D189H | missense | rs146775001 |  |
| *ASPN* | 11 | g.chr9:95237026insCTC | c.154\_155insGAG | p.E52delinsEE | NFS ins | . |  |
| *ASPN* | 4 | g.chr9:95237022insTCTCAT | c.158\_159insATGAGA | p.D53delinsDEN | NFS ins | . |  |
| *ATG2B* | 1 | g.chr14:96784164G>A | c.C2908T | p.P970S | missense | rs200385687 |  |
| *ATM* | 14 | g.chr11:108122592C>G | c.C1636G | p.L546V | missense | rs2227924 | 0.7 |
| *AXL* | 13 | g.chr19:41743872G>A | c.G807A | p.M269I | missense | . |  |
| *BAP1* | 45 | g.chr3:52439240A>G | c.1002A>AG | p.L334L | synonymous | rs28997577 | 0.9 |
| *BRCA1* | 13 | g.chr17:41244765C>T | c.G2783A | p.G928D | missense | rs202004680 | 0.4 |
| *BRCA2* | 5 | g.chr13:32906593C>A | c.C978A | p.S326R | missense | rs28897706 | 0.5 |
| *BRCA2* | 6 | g.chr13:32914126C>G | c.C5634G | p.N1878K | missense | rs80358784 | 0.4 |
| ***BRCA2*** | 7 | **g.chr13:32972626A>T** | **c.A9976T** | **p.K3326X** | **nonsense** | **rs11571833** | **1.0** |
| *BRCA2* | 1 | g.chr13:32914236C>T | c.C5744T | p.T1915M | missense | rs4987117 | 0.5 |
| *BRCA2* | 12 | g.chr13:32972930A>C | . | . | UTR3 | . | 0.6 |
| *BRD8* | 5 | g.chr5:137476527C>T | c.G3482A | p.R1161Q | missense | . |  |
| *CABIN1* | 6 | g.chr22:24483514G>A | c.G3373A:p.V1125I | p.V1125 | missense | rs148592192 |  |
| *CHD6* | 1 | g.chr20:40079655G>C | c.C3614G | p.A1205G | missense | rs41278126 |  |
| *CHD6* | 11 | g.chr20:40040870G>A | c.C7165T | p.R2389C | missense | rs61752057 |  |
| *CHD6* | 2 | g.chr20:40040870G>A | c.C7165T | p.R2389C | missense | rs61752057 |  |
| *COL6A3* | 12 | g.chr2:238273001C>T | c.G5909A | p.R1970H | missense | . |  |
| *CREBBP* | 4 | g.chr16:3820723T>C | c.A2614G | p.T872A | missense | rs143247685 |  |
| *CREBBP* | 5 | g.chr16:3820773G>A | c.C2564T | p.S855L | missense | rs142047649 |  |
| ***DHDH*** | 13 | **g.chr19:49447750-49447751delAG** | **c.881\_882del** | **p.294\_294del** | **FS del** | **.** |  |
| *DHDH* | 10 | g.chr19:49442940G>A | c.G601A | p.G201R | missense | rs76719621 | 0.5 |
| ***DHDH*** | 9 | **g.chr19:49442856insG** | **c.517\_518insG** | **p.A173fs** | **FS ins** | **.** | **0.4** |
| *DNAH8* | 9 | g.chr6:38747785C>G | c.C1432G | p.H478D | missense | rs200737379 | 0.4 |
| *DNAH8* | 14 | g.chr6:38704877G>A | c.G146A | p.G49E | missense | . | 0.5 |
| *DNAH8* | 14 | g.chr6:38854646T>C | c.T7688C | p.I2563T | missense | rs142328376 | 0.5 |
| *DNMT1* | 13 | g.chr19:10291473C>T | c.G206A | p.R69H | missense | rs61750053 | 0.5 |
| *DNMT3L* | 2 | g.chr21:45678516C>T | c.G406A | p.A136T | missense | . |  |
| *EP300* | 7 | g.chr22:41546158C>A | c.C2773A | p.P925T | missense | rs148884710 | 0.7 |
| *ERBB3* | 9 | g.chr12:56493456G>A | c.G2864A: | p.R955H | missense | . | 0.7 |
| *FAT4* | 2 | g.chr4:126412106C>G | c.C14129G | p.S4710C | missense | rs147662558 | 0.7 |
| ***FGFBP1*** | 9 | **g.chr4:15938250delC** | **c.6delG** | **p.K2fs** | **FS del** | **.** | **0.7** |
| *FGFR1* | 6 | g.chr8:38287281G>A | c.C277T | p.P93S | missense |  | 0.7 |  |
| *FGFR1* | 4 | g.chr8:38275420C>T | c.G1514A | p.R505H | missense | . | 0.6 |
| *GCN1L1* | 12 | g.chr12:120580413C>G | c.G5727C | p.K1909N | missense | . |  |
| *HIF1A* | 13 | g.chr14:62207816C>T | c.C2003T | p.S668L | missense | . | 0.4 |
| *HMCN1* | 9 | g.chr1:186010161G>A | c.G6197A | p.R2066Q | missense | . |  |
| *HMCN1* | 2 | g.chr1:185891584G>A | c.G974A | p.R325Q | missense | rs202206115 | 0.7 |
| *HMCN1* | 4 | g.chr1:186031015A>G | c.A7345G | p.M2449V | missense | rs199613884 |  |
| *HTT* | 8 | g.chr4:3231645G>A | c.G8141A | p.R2714H | missense | . |  |
| *JMJD1C* | 8 | g.chr10:64974188G>A | c.C1082T | p.T361I | missense | . | 0.4 |
| *JMJD1C* | 6 | g.chr10:65225409G>A | c.C14T | p.T5M | missense | . | 0.4 |
| *JMJD1C* | 5 | g.chr10:64967953delAAACCT | c.3471\_3476del | p.1157\_1159del | NFS del | . | 0.7 |
| *KALRN* | 1 | g.chr3:124045010G>T | c.G1270T | p.A424S | missense | rs138624947 |  |
| *KDM4D* | 5 | g.chr11:94731905C>G | c.C1369G | p.P457A | missense | rs144086807 | 0.4 |
| *KDM4D* | 6 | g.chr11:94731905C>G | c.C1369G | p.P457A | missense | rs144086807 |  |
| ***KDM6A*** | 47 | **g.chrX:44948987G>A** | **c.3548-1G>A** | **.** | **splicing** | **.** | **0.3** |
| *KDM6B* | 13 | g.chr17:7753469G>A | c.G3647A | p.R1216Q | missense | . |  |
| *LAMA3* | 13 | g.chr18:21355857A>G | c.A1375G | p.I459V | missense | rs199968275 |  |
| *LRP1B* | 12 | g.chr2:141093253G>A | c.C12047T | p.P4016L | missense | rs150957163 |  |
| *LRP1B* | 2 | g.chr2:141093245T>G | c.A12055C | p.T4019P | missense | . |  |
| *LRP1B* | 7 | g.chr2:141819751T>C | c.A1105G | p.T369A | missense | rs138996626 |  |
| *MET* | 8 | g.chr7:116411990C>T | c.C2975T | p.T992I | missense | rs56391007 | 0.7 |
| *MET* | 11 | g.chr7:116411923C>T | c.C2908T | p.R970C | missense | rs34589476 | 0.5 |
| *MLL* | 12 | g.chr11:118343378G>A | c.G1504A | p.E502K | missense | rs9332772 | 0.7 |
| *MLL* | 5 | g.chr11:118343378G>A | c.G1504A | p.E502K | missense | rs9332772 | 0.7 |
| *MLL3* | 1 | g.chr7:151860230G>C | c.C10432G | p.Q3478E | missense | rs142835638 | 0.5 |
| *MLL3* | 6 | g.chr7:151877127G>T | c.C7234A | p.P2412T | missense | rs13231116 | 0.4 |
| *MLL3* | 11 | g.chr7:151919751A>G | c.T3340C | p.C1114R | missense | rs200559566 |  |
| *MLL5* | 13 | g.chr7:104747899G>T | c.G2995T | p.G999C | missense | rs117986340 |  |
| *MOS* | 12 | g.chr8:57026507C>A | c.G35T | p.R12L | missense | . |  |
| *NCOR1* | 3 | g.chr17:15973774T>G | c.A4266C | p.L1422F | missense | rs61753150 | 0.4 |
| *NHS* | 14 | g.chrX:17743695G>T | c.G938T | p.G313V | missense | . |  |
| *PALB2* | 12 | g.chr16:23646857A>G | c.T1010C | p.L337S | missense | rs45494092 | 0.7 |
| *PALB2* | 6 | g.chr16:23634293C>T | c.G2993A | p.G998E | missense | rs45551636 | 0.5 |
| *PALB2* | 6 | g.chr16:23641461C>G | c.G2014C | p.E672Q | missense | rs45532440 | 0.7 |
| *PDZD2* | 12 | g.chr5:31799522C>T | c.C167T | p.T56M | missense | rs145138976 |  |
| *PDZD2* | 5 | g.chr5:32074573G>A | c.G3361A | p.V1121M | missense | rs149535005 |  |
| *PDZD2* | 6 | g.chr5:32091049G>A | c.G7495A | p.E2499K | missense | rs146545649 |  |
| ***RB1*** | 3 | **g.chr13:49050853InsA** | **c.2537\_2538insA** | **p.Q846fs** | **FS ins** | **.** | **0.6** |
| *RB1* | 4 | g.chr13:49030485G>A | c.G1960A | p.V654M | missense | . | 0.7 |
| *REL* | 11 | g.chr2:61149189C>T | c.C1379T | p.A460V | missense | . |  |
| *RRP8* | 2 | g.chr11:6623433G>A | c.C112T | p.R38C | missense | rs146487650 |  |
| ***RRP8*** | 13 | **g.chr11:6624695insA** | **c.38\_39insT** | **p.V13fs** | **FS ins** | **.** |  |
| *RSF1* | 14 | g.chr11:77412776C>T | c.G1498A | p.G500S | missense | . |  |
| *SETDB1* | 9 | g.chr1:150922939C>T | c.C1586T | p.P529L | missense | rs143224912 | 0.5 |
| *SIN3A* | 3 | g.chr15:75704005G>A | c.C836T | p.P279L | missense | rs61761938 | 0.5 |
| *SIN3B* | 5 | g.chr19:16980668C>T | c.C2300T | p.P767L | missense | rs117307745 | 0.7 |
| *SPTA1* | 8 | g.chr1:158589105T>C | c.A6437G | p.Q2146R | missense | rs138055271 |  |
| *SPTA1* | 8 | g.chr1:158637728T>C | c.A1958G | p.Y653C | missense | [rs148912436](http://www.ncbi.nlm.nih.gov/SNP/snp_ref.cgi?type=rs&rs=rs148912436) |  |
| *SPTA1* | 8 | g.chr1:158590081A>G | c.T6296C | p.F2099S | missense | . |  |
| *SPTA1* | 4 | g.chr1:158605757C>T | c.G5378A | p.R1793Q | missense | rs200938874 |  |
| *SPTA1* | 12 | g.chr1:158589105T>C | c.A6437G | p.Q2146R | missense | rs138055271 |  |
| *SPTA1* | 12 | g.chr1:158637728T>C | c.A1958G | p.Y653C | missense | rs148912436 |  |
| *STAG2* | 27 | g.chrX:123196819T>G | c.T1706G | p.V569G | missense | . | 0.4 |
| *STAG2* | 44 | g.chrX:123156521A>G | c.A44G | p.Q15R | missense | . | 0.5 |
| *SYNE1* | 11 | g.chr6:152501416C>T | c.G6887A | p.R2296Q | missense | rs138787771 |  |
| *SYNE2* | 8 | g.chr14:64608748A>G | c.A15248G | p.D5083G | missense | rs149617373 |  |
| *SYNE2* | 10 | g.chr14:64634063G>A | c.G16718A | p.R5573Q | missense | rs149227847 | 0.7 |
| *TET1* | 1 | g.chr10:70406520C>T | c.C4034T | p.P1345L | missense | rs150543016 | 0.5 |
| *TET1* | 9 | g.chr10:70406747A>G | c.A4261G | p.T1421A | missense | rs144156611 | 0.7 |
| *TET1* | 14 | g.chr10:70406747A>G | c.A4261G | p.T1421A | missense | rs144156611 | 0.7 |
| *TG* | 1 | g.chr8:134030106A>G | c.A6646G | p.I2216V | missense | rs138690529 |  |
| *TG* | 7 | g. chr8:133899136G>A | c.G1519A | p.A507T | missense | . |  |
| *TRRAP* | 7 | g.chr7:98491422C>T | c.C368T | p.T123M | missense | rs148101267 |  |
| *TSC2* | 3 | g.chr16:2110795G>A | c.G1100A | p.R367Q | missense | rs1800725 | 0.7 |
| *UTY* | 5 | g.chrY:15591171T>A | c.A179T | p.H60L | missense | . | 1.0 |
| *ZFHX3* | 5 | g.chr16:72822033C>T | c.G7400A | p.R2467Q | missense | rs200959815 |  |
| *ZFHX3* | 5 | g.chr16:72828890T>C | c.A4949G | p.Q1650R | missense | rs141564201 |  |

FS, frameshift; NFS, non-FS; UTR, untranslated region; del, deletion; ins, insertion; RSI, relative signal intensity from

tumor Sanger sequence chromatograms; alterations in bold indicate deleterious alleles; genomic coordinates, HG19

genome build.

**Supplementary Table S5. *BAP1* alterations are associated with papillary histologic morphology in bladder tumors.**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Sample ID #** | **Tumor Type** | ***BAP1* Status** | **Histologic****Morphology** | **Muscle Invasive** | **Grade****(WHO II)** |
| 1 | Urothelial | p.H169R | non-pap | NA | high |
| 3 | Urothelial |  | non-pap | yes | high |
| 4 | Urothelial |  | non-pap | NA | high |
| 5 | Urothelial | p.R57Q, p.D11N | pap | no | low |
| 9 | Inverted Urothelial | p.H169Q | pap | no | low |
| 15 | Urothelial | p.D236H | non-pap | yes | high |
| 24 | Urothelial | p.R718G | non-pap | yes | high |
| 33 | Urothelial |  | non-pap | yes | high |
| 34 | Urothelial |  | non-pap | yes | high |
| 35 | Urothelial |  | non-pap | yes | well |
| 39 | Urothelial | wild type | pap | yes | high |
| 40 | Urothelial |  | non-pap | yes | high |
| 42 | Urothelial | p.R718G | pap | no | low |
| 45 | Urothelial | LOH at rs28997577 | pap | no | high |
| 48 | Urothelial | p.E257X | pap | no | low |
| 54 | Urothelial |  | non-pap | NA | high |

Pap, papillary features were observed in an H&E section of the tumor; black shading, tumors with a *BAP1* mutation and histologic papillary features; grey box, a tumor with wild type *BAP1* and papillary features; NA, not available.

**Supplementary Table S6. *TERT* promoter alterations in 54 bladder tumors.**

| **Sample ID** | **Genomic** | **Nucleotide (cDNA)** | **Status** | **RSI** | **dbSNP137/Reference** |
| --- | --- | --- | --- | --- | --- |
| 1 | g.chr5:1295250G>A | c.-146C>T | somatic | 0.6 | Horn et al., 2013; Huang et al., 2013 |
| 1 | g.chr5:1295373C>T | c.-269G>A | somatic | 0.3 | rs35226131 |
| 3 | g.chr5:1295228G>A | c.-124C>T | somatic | 0.6 | Horn et al., 2013; Huang et al., 2013 |
| 4 | g.chr5:1295228G>A | c.-124C>T | somatic | **1.0** | Horn et al., 2013; Huang et al., 2013 |
| 4 | g.chr5:1295349A>G | c.-245T>C | germline | **1.0** | rs2853669, Horn et al., 2013 |
| 5 | g.chr5:1295228G>A | c.-124C>T | somatic | 0.4 | Horn et al., 2013; Huang et al., 2013 |
| 5 | g.chr5:1295349A>G | c.-245T>C | somatic | 0.6 | rs2853669, Horn et al., 2013 |
| 6 | g.chr5:1295228G>A | c.-124C>T | somatic | 0.6 | Horn et al., 2013; Huang et al., 2013 |
| 6 | g.chr5:1295349A>G | c.-245T>C | somatic | 0.5 | rs2853669, Horn et al., 2013 |
| 7 | g.chr5:1295228G>A | c.-124C>T | somatic | 0.4 | Horn et al., 2013; Huang et al., 2013 |
| 7 | g.chr5:1295332T>C | c.-228A>G | germline | 0.3 | **Novel** |
| 7 | g.chr5:1295349A>G | c.-245T>C | germline | 0.5 | rs2853669, Horn et al., 2013 |
| 8 | g.chr5:1295228G>A | c.-124C>T | somatic | 0.6 | Horn et al., 2013; Huang et al., 2013 |
| 8 | g.chr5:1295349A>G | c.-245T>C | germline | 0.6 | rs2853669, Horn et al., 2013 |
| 9 | g.chr5:1295349A>G | c.-245T>C | germline | **1.0** | rs2853669, Horn et al., 2013 |
| 10 | g.chr5:1295228G>A | c.-124C>T | somatic | 0.5 | Horn et al., 2013; Huang et al., 2013 |
| 10 | g.chr5:1295255G>A | c.-151C>T | germline | 0.3 | **Novel** |
| 10 | g.chr5:1295327C>T | c.-223G>A | germline | 0.6 | **Novel** |
| 10 | g.chr5:1295349A>G | c.-245T>C | somatic | 0.6 | rs2853669, Horn et al., 2013 |
| 10 | g.chr5:1295410C>T | c.-306G>A | germline | 0.5 | **Novel** |
| 12 | g.chr5:1295250G>A | c.-146C>T | somatic | 0.6 | Horn et al., 2013; Huang et al., 2013 |
| 13 | g.chr5:1295316G>C | c.-212C>G | somatic | 0.6 | **Novel** |
| 13 | g.chr5:1295228G>A | c.-124C>T | somatic | 0.4 | Horn et al., 2013; Huang et al., 2013 |
| 13 | g.chr5:1295349A>G | c.-245T>C | germline | 0.4 | rs2853669, Horn et al. 2013 |
| 14 | g.chr5:1295349A>G | c.-245T>C | somatic | **1.0** | rs2853669, Horn et al. 2013 |
| 15 | g.chr5:1295228G>A | c.-124C>T | somatic | 0.5 | Horn et al., 2013; Huang et al., 2013 |
| 15 | g.chr5:1295349A>G | c.-245T>C | germline | 0.4 | rs2853669, Horn et al., 2013 |
| 16 | g.chr5:1295349A>G | c.-245T>C | germline | 0.3 | rs2853669, Horn et al., 2013 |
| 17 | g.chr5:1295215G>A | c.-111C>T | somatic | 0.4 | **Novel** |
| 17 | g.chr5:1295216G>A | c.-112C>T | somatic | 0.4 | rs35733142 |
| 17 | g.chr5:1295217G>A | c.-113C>T | somatic | 0.4 | **Novel** |
| 17 | g.chr5:1295228G>A | c.-124C>T | somatic | 0.4 | Horn et al., 2013; Huang et al., 2013 |
| 19 | g.chr5:1295228G>A | c.-124C>T | somatic | 0.6 | Horn et al., 2013; Huang et al., 2013 |
| 20 | g.chr5:1295349A>G | c.-245T>C | germline | 0.4 | Horn et al. 2013, rs2853669 |
| 21 | g.chr5:1295438C>A | c.-334G>T | germline | 0.5 | **Novel** |
| 22 | g.chr5:1295228G>A | c.-124C>T | somatic | 0.6 | Horn et al., 2013; Huang et al., 2013 |
| 22 | g.chr5:1295237G>C | c.-133C>G | somatic | 0.6 | **Novel** |
| 22 | g.chr5:1295349A>G | c.-245T>C | somatic | 0.6 | Horn et al. 2013,rs2853669 |
| 24 | g.chr5:1295228G>A | c.-124C>T | somatic | 0.4 | Horn et al., 2013; Huang et al., 2013 |
| 25 | g.chr5:1295242G>A | c.-138C>T | somatic | 0.4 | Horn et al., 2013 |
| 25 | g.chr5:1295243G>A | c.-139C>T | somatic | 0.6 | rs35550267 |
| 26 | g.chr5:1295270G>A | c.-166C>T | germline | 0.3 | **Novel** |
| 26 | g.chr5:1295349A>G | c.-245T>C | germline | 0.6 | rs2853669, Horn et al. 2013 |
| 27 | g.chr5:1295228G>A | c.-124C>T | somatic | 0.6 | Horn et al., 2013; Huang et al., 2013 |
| 27 | g.chr5:1295349A>G | c.-245T>C | germline | 0.4 | rs2853669, Horn et al. 2013 |
| 28 | g.chr5:1295217G>A | c.-113C>T | germline | 0.5 | **Novel** |
| 28 | g.chr5:1295228G>A | c.-124C>T | somatic | 0.6 | Horn et al., 2013; Huang et al., 2013 |
| 28 | g.chr5:1295230G>A | c.-126C>T | germline | 0.5 | **Novel** |
| 28 | g.chr5:1295260G>A | c.-156C>T | germline | 0.6 | **Novel** |
| 28 | g.chr5:1295305G>A | c.-201C>T | germline | 0.6 | **Novel** |
| 28 | g.chr5:1295349A>G | c.-245T>C | somatic | 0.4 | rs2853669, Horn et al. 2013 |
| 28 | g.chr5:1295388G>A | c.-284C>T | germline | 0.5 | **Novel** |
| 28 | g.chr5:1295433G>A | c.-329C>T | germline | 0.6 | **Novel** |
| 29 | g.chr5:1295143G>C | c.-39C>G | germline | 0.4 | **Novel** |
| 30 | g.chr5:1295228G>A | c.-124C>T | somatic | 0.4 | Horn et al., 2013; Huang et al., 2013 |
| 31 | g.chr5:1295228G>A | c.-124C>T | somatic | 0.4 | Horn et al., 2013; Huang et al., 2013 |
| 31 | g.chr5:1295349A>G | c.-245T>C | germline | 0.6 | rs2853669, Horn et al. 2013 |
| 32 | g.chr5:1295250G>A | c.-146C>T | somatic | 0.6 | Horn et al., 2013; Huang et al., 2013 |
| 32 | g.chr5:1295349A>G | c.-245T>C | germline | **1.0** | rs2853669, Horn et al. 2013 |
| 33 | g.chr5:1295291G>A | c.-187C>T | germline | 0.3 | **Novel** |
| 33 | g.chr5:1295349A>G | c.-245T>C | germline | 0.5 | rs2853669, Horn et al. 2013 |
| 33 | g.chr5:1295369G>C | c.-265C>G | germline | 0.4 | **Novel** |
| 34 | g.chr5:1295214T>A | c.-110A>T | germline | 0.4 | **Novel** |
| 34 | g.chr5:1295228G>A | c.-124C>T | somatic | 0.4 | Horn et al., 2013; Huang et al., 2013 |
| 34 | g.chr5:1295349A>G | c.-245T>C | somatic | 0.4 | rs2853669, Horn et al. 2013 |
| 35 | g.chr5:1295349A>G | c.-245T>C | somatic | 0.5 | rs2853669, Horn et al. 2013 |
| 36 | g.chr5:1295250G>A | c.-146C>T | somatic | 0.6 | Horn et al., 2013; Huang et al., 2013 |
| 37 | g.chr5:1295316G>A | c.-212C>T | germline | 0.6 | **Novel** |
| 37 | g.chr5:1295382C>T | c.-278G>A | germline | 0.5 | **Novel** |
| 38 | g.chr5:1295349A>G | c.-245T>C | somatic | 0.4 | rs2853669, Horn et al. 2013 |
| 39 | g.chr5:1295228G>A | c.-124C>T | somatic | 0.4 | Horn et al., 2013; Huang et al., 2013 |
| 39 | g.chr5:1295349A>G | c.-245T>C | germline | 0.6 | rs2853669, Horn et al. 2013 |
| 40 | g.chr5:1295228G>A | c.-124C>T | somatic | 0.4 | Horn et al., 2013; Huang et al., 2013 |
| 40 | g.chr5:1295349A>G | c.-245T>C | germline | 0.6 | rs2853669, Horn et al. 2013 |
| 41 | g.chr5:1295228G>A | c.-124C>T | somatic | 0.5 | Horn et al., 2013; Huang et al., 2013 |
| 41 | g.chr5:1295349A>G | c.-245T>C | germline | 0.4 | rs2853669, Horn et al. 2013 |
| 42 | g.chr5:1295228G>A | c.-124C>T | somatic | 0.5 | Horn et al., 2013; Huang et al., 2013 |
| 43 | g.chr5:1295349A>G | c.-245T>C | germline | 0.4 | rs2853669, Horn et al. 2013 |
| 44 | g.chr5:1295228G>A | c.-124C>T | somatic | 0.4 | Horn et al., 2013; Huang et al., 2013 |
| 44 | g.chr5:1295349A>G | c.-245T>C | germline | 0.4 | rs2853669, Horn et al. 2013 |
| 45 | g.chr5:1295349A>G | c.-245T>C | germline | 0.3 | rs2853669, Horn et al. 2013 |
| 47 | g.chr5:1295228G>A | c.-124C>T | somatic | 0.4 | Horn et al., 2013; Huang et al., 2013 |
| 47 | g.chr5:1295349A>G | c.-245T>C | germline | 0.6 | rs2853669, Horn et al. 2013 |
| 48 | g.chr5:1295250G>A | c.-146C>T | somatic | 0.6 | Horn et al., 2013; Huang et al., 2013 |
| 48 | g.chr5:1295349A>G | c.-245T>C | germline | **1.0** | rs2853669, Horn et al. 2013 |
| 49 | g.chr5:1295228G>A | c.-124C>T | somatic | 0.4 | Horn et al., 2013; Huang et al., 2013 |
| 49 | g.chr5:1295349A>G | c.-245T>C | germline | 0.4 | rs2853669, Horn et al. 2013 |
| 51 | g.chr5:1295228G>A | c.-124C>T | somatic | 0.4 | Horn et al., 2013; Huang et al., 2013 |
| 51 | g.chr5:1295349A>G | c.-245T>C | germline | 0.4 | rs2853669, Horn et al. 2013 |
| 52 | g.chr5:1295250G>A | c.-146C>T | somatic | 0.6 | Horn et al., 2013; Huang et al., 2013 |
| 53 | g.chr5:1295228G>A | c.-124C>T | somatic | 0.6 | Horn et al., 2013; Huang et al., 2013 |
| 53 | g.chr5:1295250G>A | c.-146C>T | somatic | 0.6 | Horn et al., 2013; Huang et al., 2013 |
| 54 | g.chr5:1295360T>A | c.-256A>T | germline | 0.4 | **Novel** |

Bold RSI indicates amplification of the mutant allele.

**Supplementary Table S7. Gene mutation correlations in 54 bladder tumors.**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | ***ARID1A*** | ***BAP1*** | ***STAG2*** | ***TERT* promoter** |
|  |  **N Y** |  **N Y** |  **N Y** |  **N Y** |
| ***KDM6A*** |  |  |  |  |  |  |  |  |
| **N** | 37 | 5 | 39 | 3 | 37 | 5 | 11 | 29 |
| **Y** | 9 | 3 | 8 | 4 | 10 | 2 | 4 | 8 |
|  | p=0.27 | **p=0.017** | p=0.67 | p=0.70 |
| ***ARID1A*** |   |   |   |   |  |   |  |  |
| **N** |   |   | 41 | 5 | 39 | 7 | 13 | 31 |
| **Y** |   |   | 6 | 2 | 8 | 0 | 2 | 6 |
|  |   |   | p=0.28 | p=0.25 | p=0.80 |
| ***BAP1*** |   |   |   |   |  |   |  |  |
| **N** |   |   |   |   | 42 | 5 | 14 | 31 |
| **Y** |   |   |   |   | 5 | 2 | 1 | 6 |
|  |   |   |   |   | p=0.20 | p=0.37 |
| ***STAG2*** |   |   |   |   |   |   |  |  |
| **N** |   |   |   |   |   |   | 12 | 34 |
| **Y** |   |   |   |   |   |   | 3 | 3 |
|  |   |   |   |   |   |   | p=0.23 |

Correlations in somatic mutation status for pairs of genes sequenced in 54 tumors. N, no mutation; Y, mutation positive; black shaded, Pearson’s correlation p-value <0.05.

**Supplementary Table S8. Top scoring networks in BC.** Ingenuity Pathway Analysis was used to generate interaction networks from the mutated genes confirmed by Sanger sequencing in bladder cancer, based on any direct or indirect relationships known in the scientific literature. Networks are ranked by their score, which is the - log of the probability of obtaining the given relationships by random chance (Fisher’s Exact Test). The number and identity of the mutated genes used in each network is given in the first two columns respectively, with networks limited to 35 molecules each to keep them a usable size. A list of the top over-represented biological functions associated with the respective network molecules is given in the last column.

|  |  |  |  |
| --- | --- | --- | --- |
| **# of mutated genes** | **Gene symbols** | **Score****(log(p-value))** | **Top functions over-represented in network** |
| 35 | ANO5, ATPAF1, AVL9 (includes EG:23080), C20orf72, C8orf76, CEP135, CHD6, CNTNAP3, CSMD3, DDRGK1, DHX35, DIP2C, DISP1, GMIP, GON4L, HOOK3, KATNAL1, KBTBD4, KCTD7, LARP4B, METTL10, MTBP, OSGIN2, PSD4, PWP2, SCRN3, SLC17A5, SLC44A2, SNX14, SPAG1, TDRD1, UBC, VPS13B, XPO4, ZZEF1 | 68 | Hereditary Disorder, Skeletal and Muscular Disorders, Cell Signaling |
| 28 | ANPEP, ARID2, ARID1A, ARID4A, ARID4B, BAP1, CHD4, DUB, HOXB8 (includes EG:15416), KDM5B, KIF15, CHD3, MITF, MKI67, MLL, MTA, NCOR1, NfkB1-RelA, NR1D2, p300-CBP, PARP, PPM1G, PRDM10, SALL4, SIN3A, SMARCC1, USP21, USP26, USP31, USP34, USP36, USP38, USP48, Vegf, WHSC1L1 | 46 | Cellular Compromise, Gene Expression, Cellular Development |
| 27 | ACTB, ASPM, BRD1, BRPF1, CA9, CHD1, DDX1, DNMT1, DNMT3A, FAT2, GART, Hdac, Histone h3, Histone h4, JMJD1C, KDM3A, MLL5, N-cor, NCOA1, NKD1, PADI4, Pias, PRMT5, PTPRU, Rnr, RPS11, SMARCA4, SRRM2, SYNE2, TP53 (includes EG:22059), Troponin t, TRPV5, VitaminD3-VDR-RXR, WDR12, ZNF295 | 42 | DNA Replication, Recombination, and Repair, Gene Expression, Cellular Development |
| 23 | ABCC12, ANK3, Ap2 alpha, AP2B1, ARID3A, ATP5C1, atypical protein kinase C, c-Src, CD1C, EF-1 alpha, EGFR, Egfr dimer, EIF2C4, HUWE1, JINK1/2, LAMA4, MAGED1, MET, MKK3/6, MYH6, Pak, Par6, PARD3, PARD6B, Plexin B, PLXNB1, PLXNB2, PTPRZ1, Smad, SMURF1, SOS2, Sos, UBR4, UBR5, UBR1 (includes EG:197131) | 35 | Cancer, Cellular Response to Therapeutics, Respiratory Disease |
| 23 | Akt, ATG2B, ATM, ATM/ATR, ATR, Basc, BRCA1, BRCA2, BRCA1-BRCA2-FANCD2-FANCN-RAD51, BRIP1, CAPN5, CLSPN, Collagen Type VI, DCAF6, E4F1, ESPL1, I kappa b kinase, mediator, MLH1, Mre11, MTORC2, NFkB (family), NIPBL, PALB2, POLD1, RFC1, SEMA6D, SMC1-SA2, SMC1A, SPECC1, STAG2, TAOK3, TRRAP, Tsc1-Tsc2, TTC3 | 35 | DNA Replication, Recombination, and Repair, Cell Cycle, Cellular Assembly and Organization |
| 20 | ALOX12, Ampa Receptor, amylase, ANK2, Beta adaptin, CaMKII, CSPG4, ERK1/2, GRI, GRIA2, GRIA4, GRIN1, GRIN2B, Integrin alpha 3 beta 1, KALRN, KDM6A, LAMA1, LAMA3, LAMB3, LRP1B, MEP1A, mGluR, MLL3, Ncx, PLCB1, PP1 protein complex group, Pp2b, PPP1R9A, PRKD3, PTPase, Rap, RELN, Secretase gamma, SYNGAP1, TIF2-NCOA1-p300-PCAF-CBP | 29 | Cell-To-Cell Signaling and Interaction, Nervous System Development and Function, Organ Morphology |
| 19 | ABCA4, ACAD10, AKNA, ATAD2, ATP, ATP8A1, ATP8B2, ATP8B3, BTAF1, BTF3L4, C10orf137, CATSPER2, CATSPERG, CD97, CNTRL, CREB3L2, DIDO1, DNAJC22, ENPP7, GPR158, GPR179, GPR89A/GPR89B, GPR89C, GPRC5A, HSPA4L, LEPRE1, MSH5, N4BP2, PCYT1B, PHKA1, RSF1 (includes EG:233532), SMPD4, STX5, TLN2, UBC | 27 | Lipid Metabolism, Molecular Transport, Small Molecule Biochemistry |
| 19 | 14-3-3, ADCY2, AKAP13, AMPK, ATF6, BRAF, CAK, Cdk, CUL5, Cyclin A, Cyclin D, FGFBP1, GLI3, GTPASE, Histone H1, HNRPDL, IQGAP1, Mapk, MTOR, MTORC1, MYCBP2, NF1 (includes EG:18015), NUP62, p70 S6k, PCK1 (includes EG:18534), PEPCK, Pka catalytic subunit, PRKAA, PRKAC, PTPRN2, SESN2, TNPO1, TPO, TSC2, TSH | 25 | Nervous System Development and Function, Organ Morphology, Cancer |
| 18 | Ahr-aryl hydrocarbon-Arnt, ALDH1B1, ARHGAP28, ATP2A3, Coup-Tf, ELF3, ETS, GC-GCR dimer, GSR, HMG CoA synthase, Igh (family), LAMP2, LNX1, MADD, MAGEB18, MED1 (includes EG:19014), NCOA7, NCOA, NFkB (complex), NRG (family), Nuclear factor 1, PNKD, POM121, Rar, Rxr, SERCA, SH3RF1, Smad2/3, SOAT1, SWI-SNF, SYT2, T3-TR-RXR, thyroid hormone receptor, TRIM23, TRIM32 | 25 | Cardiovascular Disease, Hereditary Disorder, Cellular Development |
| 18 | 20s proteasome, ACAT2, ACIN1, Actin, Alpha tubulin, ATP8A1, BAIAP2, caspase, creatine kinase, Ctbp, Cytochrome c, cytochrome-c oxidase, ERG, ESRRA, GALM, HECTD1, Hsp27, Hsp70, Hsp90, HSP, HTT, Ifn gamma, IFT57, KAT8, Nos, OPCML, P38 MAPK, PFK, PFKL, REV1 (includes EG:316344), SASH1, thymidine kinase, TMPRSS2, TPR, ZMYND8 | 24 | Cellular Assembly and Organization, Cellular Function and Maintenance, Developmental Disorder |
| 17 | AIRE, ARID5A, C/ebp, Cdc2, CUBN, CUX1, Cyclin A/Cdk2, Cyclin B, Cyclin E, CYP19, CYP1A1 (includes EG:13076), Dgk, DGKH, E2f, ENO2, G-Actin, Gm-csf, LRP2, MCM3, MCM3AP, P-TEFb, PCIF1, POLR2A, Ras, Rb, RNA polymerase II, Rsk, SETD2, TFIIA, TFIIF, TFIIH, TG, WEE1, YAP1 (includes EG:10413), ZNF83 | 23 | Lipid Metabolism, Molecular Transport, Small Molecule Biochemistry |
| 17 | Alp, Ap1, ATPase, CD3, CDH5, CDKN1A, CHD1L, CHRD, Ck2, Creb, CREBBP, CRNKL1, CYP, Focal adhesion kinase, GCLM, GLDC, HMCN1, HSP90AA1, IL1, LDL, NELL1, NMDA Receptor, Notch, P110, PDGF BB, Pdgfr, PRDM16, Ras homolog, RXRA, SP1, SP110, SRC (family), STAT, SUPT16H, SYNE1 | 23 | Cell Cycle, Cell Death and Survival, Amino Acid Metabolism |
| 18 | alcohol group acceptor phosphotransferase, ATP2A2, CABIN1, Calcineurin A, Calcineurin protein(s), calpain, CANX, CLASRP, CNTNAP2, DTX1, DYRK1A, EPB41L3, Fcer1, GNA13, GRK4, IgG2b, Lh, MAP2K1/2, MAP3K1, MAP3K5, MEF2, NFAT (complex), Nfat (family), NOTCH2, NPR2, PAK3, PI3K (family), Pkc(s), PLA2, PP2A, RYR1, Sapk, SLC12A2, Spectrin, SPTA1 | 21 | Cell Morphology, Hematological System Development and Function, Humoral Immune Response |
| 16 | AChR, AHNAK, AMBRA1, CSF1R, DARS2, Dynamin, Egfr-Erbb2, EGFR/PDGFR/IGFR, ENaC, Endophilin, EPHA8, Fgf, FGFR3, Fgfr, FLT4, Gap, growth factor receptor, Laminin1, Lpa receptor, LRP, MERTK, NCK, PDGFRB, PI3K (complex), PI3K p85, PIK3CA, PIK3R4, PLC gamma, PLCG2 (includes EG:234779), SLC26A4, SMPD3, SYK/ZAP, TSC1, UVRAG, VAV | 21 | Cancer, Reproductive System Disease, Endocrine System Disorders |
| 16 | Adaptor protein 1, Alpha Actinin, Alpha catenin, ANTXR1, CK1, COL22A1, COL3A1, COL5A3, COL6A3, collagen, Collagen type I, Collagen type III, Collagen type IV, Collagen(s), Fc gamma receptor, FLNB, Growth hormone, HSPG2, IL-1R, IL18RAP, Integrin, ITGA5, ITGA10, ITGB8, Jnk, KLF12, Laminin, Ppp2c, Raf, TAOK2, Tgf beta, VANGL2, WNT2, Wnt, WNT10B | 21 | Connective Tissue Disorders, Dermatological Diseases and Conditions, Gastrointestinal Disease |
| 15 | beta-estradiol, C1QC, C4orf17, CARS, CCDC80, CCL18, CDK6, CHD3, CHRNG, CNN2, DSCAML1, ESPL1, FMN2, HSD17B6, IFI30, ITIH5, KIAA1199, mir-23, mir-24, mir-103, mir-145, OXCT1, PDZD2, PLAGL2, PSEN1, RAD21, SDK1, SEMA3C, STAT4, TCEB3B, TGFB1, TNF, ZMYM3, ZNF80, ZNF174 | 19 | Lipid Metabolism, Small Molecule Biochemistry, Cellular Development |
| 14 | BCR (complex), CACNA1S, Calmodulin, CASP8, Caspase 3/7, CD8, CD86, Cofilin, DMD, DOCK8, EMILIN2, F Actin, GOT, HEXA, Ige, IgG1, Igg3, IgG, IgG2a, Igm, IL12 (complex), Immunoglobulin, Interferon alpha, KRAS, Mek, MHC Class I (complex), MHC CLASS I (family), MPO, NFE2L3, OGDH, PKD1L1, Proinsulin, Rock, SACS, SIGLEC1 | 17 | Cancer, Cellular Growth and Proliferation, Developmental Disorder |
| 14 | ADM, AGT, ALS2CR11, BCAS3, CD97, CELSR1, CES2, CHD2, FAT4, FFAR3, FGL1, GALNT13, GAS2 (includes EG:14453), GCN1L1, GPR182, GPX2 (includes EG:14776), HNF1A, IL6, KRT33A, mir-19, mir-26, MIR17HG, MYC, PGLYRP4, RNF113A, RTN2 (includes EG:20167), RYK (includes EG:140585), SAMD9L, SETBP1, SOX3, SUMO1, tretinoin, VIM, XAF1, ZNF536 | 17 | Cancer, Cellular Development, Reproductive System Development and Function |
| 13 | Cbp/p300, EP400, ERBB, ERCC2, ERK, Esr1-Esr1-estrogen-estrogen, Gcn5l, Hat, Hd-neuronal intranuclear inclusions, Hdac1/2, HISTONE, Holo RNA polymerase II, Ifn, IFN alpha/beta, IFN Beta, IFN TYPE 1, IFNAR2, Ikk (family), IRF2, IRF, IRF1 (includes EG:16362), JAK, Oas, RASGRF2, Smad1/5/8, STAM, STAT5a/b, TAF1 (includes EG:270627), TAF4B, TIP60, TRAF3IP1, TRAK1, TRIM25, Ube3, ZFHX3 | 15 | Infectious Disease, Gene Expression, Cancer |
| 13 | ABCA8, APOH, ASB15, BCKDHA, BEST2, C2, C16orf53, CES2, CNN1 (includes EG:1264), EID1, FOXA1, HDL-cholesterol, HNF4A, KDM8, KIAA0141, MIS18BP1, MLL3, MMAB, NCOA6, NR0B2, NR1I2, OAZ2, ONECUT1, ORM1, ORM2 (human), PCBD1, PIPOX, PPFIBP1, SETDB2, TAOK3, TGS1, TMOD2, UTY, ZMYND10, ZNF502 | 15 | Embryonic Development, Tissue Morphology, Metabolic Disease |
| 13 | ACP2, APP, BAI1, BAI3, BDNF, C17orf80, CCDC111, cyclic AMP, DNAH8, DNAH9, DNAH17, DYNLT1, FFAR3, FIGLA, FRMPD1, GNL3L, GPR6, GPR87, GSG1, KLC4, LAMP5, mir-322, miR-16-5p (and other miRNAs w/seed AGCAGCA), NAE1, NRXN3, PRAM1, RPRM, Serpina3g (includes others), SH3BGRL2, SLC6A6, SNCB, TMEM43, TP53 (includes EG:22059), TSPAN12, voltage-gated calcium channel | 14 | Cell Death and Survival, Cellular Assembly and Organization, Nervous System Development and Function |
| 12 | ADCY8, ADCY, Clathrin, Collagen Alpha1, EGF, ERBB3, estrogen receptor, EZH1, G protein, G protein alphai, G protein beta gamma, G-protein beta, GHRHR (includes EG:14602), Gpcr, GRM8, KDM5A, LOXL2, LRP1 (includes EG:16971), NADPH oxidase, p160, p85 (pik3r), Pdgf (complex), Pik3r, Pka, PLC, Pld, pyruvate kinase, Rap1, Sfk, Shc, SLC12A3, TNS3, Trk Receptor, TRPM6, tyrosine kinase | 12 | Hematological Disease, Hereditary Disorder, Cell Signaling |
| 10 | 26s Proteasome, Adaptor protein 2, Beta Arrestin, C15orf55, Cg, EP300, ERBB2, FRRS1, FSH, Gsk3, hemoglobin, HRAS, Ikb, IKK (complex), IL12 (family), Insulin, IREB2, KIAA0430, mir-30, Mmp, MMP23B, MUC13, OXCT1, PORCN, PRELID1, Pro-inflammatory Cytokine, Rac, Rap1GAP, RB1, RPL39 (includes EG:100361661), TCR, Tnf, Tubulin, Ubiquitin, WSB2 | 11 | Cell Cycle, Cellular Development, Connective Tissue Development and Function |
| 1 | DUSP12, MBD6 | 2 | Post-Translational Modification, Cell Death and Survival, Dermatological Diseases and Conditions |
| 1 | CAPN2, LGSN | 2 | Cancer, Cellular Movement, Tumor Morphology |

**Supplementary Table S9. Microarray gene expression datasets used for network and gene expression analyses.**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Cohort** | **Platform** | **N** | **Description** | **Accession no\*** | **Ref** |
| MSKCC | Affymetrix HG-U133A | 129 | 93 primary bladder tumors and 38 normal bladder samples | - | ([25](#_ENREF_25)) |
| CNUH | Illumina Human-6 v2 Expression BeadChip | 197 | 165 primary bladder tumors, 10 normal bladder samples, and 22 recurrent tumors | GSE13507 | ([26](#_ENREF_26)) |
| Lindgren | Swegene Human 27K RAP UniGene188 Array | 156 | 144 primary bladder tumors and 12 normal bladder samples | GSE19915 | ([16](#_ENREF_16)) |

\*GSE datasets are available from the Gene Expression Omnibus ([27](#_ENREF_27)).

**SUPPLEMENTARY REFERENCES**

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

2. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The Sequence Alignment/Map format and SAMtools. Bioinformatics. 2009;25:2078-9.

3. Li R, Li Y, Fang X, Yang H, Wang J, Kristiansen K. SNP detection for massively parallel whole-genome resequencing. Genome Res. 2009;19:1124-32.

4. Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. Nucleic Acids Res. 2010;38:e164.

5. Shirley BC, Mucaki EJ, Whitehead T, Costea PI, Akan P, Rogan PK. Interpretation, stratification and evidence for sequence variants affecting mRNA splicing in complete human genome sequences. Genomics Proteomics Bioinformatics. 2013;11:77-85.

6. Rozen S, Skaletsky H. Primer3 on the WWW for general users and for biologist programmers. Methods Mol Biol. 2000;132:365-86.

7. Kumar P, Henikoff S, Ng PC. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. Nat Protoc. 2009;4:1073-81.

8. Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, et al. A method and server for predicting damaging missense mutations. Nat Methods. 2010;7:248-9.

9. Forbes SA, Bhamra G, Bamford S, Dawson E, Kok C, Clements J, et al. The Catalogue of Somatic Mutations in Cancer (COSMIC). Current Protocols in Human Genetics. 2008;April:Unit 10.1.

10. Guo G, Sun X, Chen C, Wu S, Huang P, Li Z, et al. Whole-genome and whole-exome sequencing of bladder cancer identifies frequent alterations in genes involved in sister chromatid cohesion and segregation. Nat Genet. 2013;45:1459-63..

11. Elkon R, Rashi-Elkeles S, Lerenthal Y, Linhart C, Tenne T, Amariglio N, et al. Dissection of a DNA-damage-induced transcriptional network using a combination of microarrays, RNA interference and computational promoter analysis. Genome Biol. 2005;6:R43.

12. Parkinson H, Sarkans U, Kolesnikov N, Abeygunawardena N, Burdett T, Dylag M, et al. ArrayExpress update-an archive of microarray and high-throughput sequencing-based functional genomics experiments. Nucleic Acids Res. 2011;39:D1002-D4.

13. Bild AH, Yao G, Chang JT, Wang Q, Potti A, Chasse D, et al. Oncogenic pathway signatures in human cancers as a guide to targeted therapies. Nature. 2006;439:353-7.

14. Sanchez-Carbayo M, Socci ND, Lozano J, Saint F, Cordon-Cardo C. Defining molecular profiles of poor outcome in patients with invasive bladder cancer using oligonucleotide microarrays. J Clin Oncol. 2006;24:778-89.

15. Kim WJ, Kim EJ, Kim SK, Kim YJ, Ha YS, Jeong P, et al. Predictive value of progression-related gene classifier in primary non-muscle invasive bladder cancer. Molecular Cancer. 2010;9:3.

16. Lindgren D, Frigyesi A, Gudjonsson S, Sjodahl G, Hallden C, Chebil G, et al. Combined gene expression and genomic profiling define two intrinsic molecular subtypes of urothelial carcinoma and gene signatures for molecular grading and outcome. Cancer Res. 2010;70:3463-72.

17. Dyrskjot L, Zieger K, Real FX, Malats N, Carrato A, Hurst C, et al. Gene expression signatures predict outcome in non-muscle-invasive bladder carcinoma: a multicenter validation study. Clin Cancer Res. 2007;13:3545-51.

18. Takeuchi T, Tomida S, Yatabe Y, Kosaka T, Osada H, Yanagisawa K, et al. Expression profile-defined classification of lung adenocarcinoma shows close relationship with underlying major genetic changes and clinicopathologic behaviors. J Clin Oncol. 2006;24:1679-88.

19. Tomida S, Takeuchi T, Shimada Y, Arima C, Matsuo K, Mitsudomi T, et al. Relapse-related molecular signature in lung adenocarcinomas identifies patients with dismal prognosis. J Clin Oncol. 2009;27:2793-9.

20. Hastie T, Ribshirani R, Balasubramanian N, Chu G. impute: Imputation for microarray data. 2010.

21. Benjamini Y, Hochberg Y. Controlling the False Discovery Rate - a Practical and Powerful Approach to Multiple Testing. J Roy Stat Soc B Met. 1995;57:289-300.

22. Miller JA, Cai C, Langfelder P, Geschwind DH, Kurian SM, Salomon DR, et al. Strategies for aggregating gene expression data: the collapseRows R function. BMC Bioinformatics. 2011;12:322.

23. Titus B, Frierson HF, Jr., Conaway M, Ching K, Guise T, Chirgwin J, et al. Endothelin axis is a target of the lung metastasis suppressor gene RhoGDI2. Cancer Res. 2005;65:7320-7.

24. Overdevest JB, Thomas S, Kristiansen G, Hansel DE, Smith SC, Theodorescu D. CD24 offers a therapeutic target for control of bladder cancer metastasis based on a requirement for lung colonization. Cancer Res. 2011;71:3802-11.

25. Sanchez-Carbayo M, Socci ND, Lozano J, Saint F, Cordon-Cardo C. Defining molecular profiles of poor outcome in patients with invasive bladder cancer using oligonucleotide microarrays. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 2006;24:778-89.

26. Kim WJ, Kim EJ, Kim SK, Kim YJ, Ha YS, Jeong P, et al. Predictive value of progression-related gene classifier in primary non-muscle invasive bladder cancer. Molecular Cancer. 2010;9.

27. Barrett T, Troup DB, Wilhite SE, Ledoux P, Evangelista C, Kim IF, et al. NCBI GEO: archive for functional genomics data sets-10 years on. Nucleic Acids Res. 2011;39:D1005-D10.