Table of contents

Section 1: Supplementary Materials and Methods

Cell line validation methods

The following methods were used for cell line validation.

A. Mycoplasma Testing

We used a PCR based method with an internal standard to detect mycoplasma contamination on each cell lines. This protocol is utilized at the Department of Human and Animal Cell Lines (DSMZ), Germany (1).

B. Species PCR

We tested each of the genomic DNA samples by a multiplex PCR reaction that amplifies a unique fragment for each of 6 commonly used species including human, cat, Chinese hamster, rat, dog, and mouse. This PCR reaction is based on the protocol developed by Cooper et al. 2007 (2).

C. STR Analysis

Isolated genomic DNA from each cell line was amplified using the StockMarks™ for Dogs Genotyping Kit, canine (Catalog number: 4307481) from Thermo Fisher Scientific. This kit utilizes multiple fluorescent dye labeled PCR primers to amplify 10 unique microsatellite markers (STRs) simultaneously in a multiplex PCR reaction. The resulting PCR products are analyzed using fragment analysis with size standards. Identified peaks are hand-binned and compared to 65 previously analyzed canine cell lines using an algorithm written in java (Supplementary Table S2) (3).

Putative driver mutation identification

Using the cbioPortal (v1.14.0, http://www.cbioportal.org/) website, driver mutations were identified in cancer genes. We manually queried 232 cancer genes, identified in this study, against all 224 studies with 64,461 samples incorporated in cbioportal (http://bit.ly/2u9UfWt).

Any canine cancer gene variant homologous to known human oncogenic or predicted oncogenic variants was classified as putative cancer driver. The variant driver predictions were based on OncoKB, cancer hotspots, and literature searches.

Functional annotation of level 2 variants

The protein coding genes (level 2) were used to identify the signaling pathways and gene ontology terms that were enriched in each cell line. The gene lists for each cell line were used as input for the DAVID functional annotation tool (4). The three databases used for this study were Gene Ontology Consortium, Kyoto Encyclopedia of Genes and Genome (KEGG), and Pfam (5-7). In order to identify additional relevant genes that may contribute to cancer development and progression, we selected functional categories that have: at least one cancer gene and were in at least six of the 33 cell lines.

Serum starvation analysis

Constitutive activation of MAPK and PI3K/AKT pathways was assessed by examining the effects of serum starvation on the phosphorylation status of ERK1/2 and pan-AKT, respectively. After 24 hours of serum starvation, the cells from all 33 cell lines were harvested and total and phosphorylated protein levels were assessed with western blot analysis. The cell lines were categorized based on ERK or AKT activation: 1. Constitutive – if there was no decrease in levels of proteins or phosphorylation was increased with serum starvation; 2. Activated: if phosphorylation was high in the presence of serum but decreased or absent with serum starvation; and 3. Low Expression/Phosphorylation: if expression or phosphorylation was low relative to other cell lines.

Western blot analysis

Cells were washed with ice-cold PBS and lysed with 100 µl of RIPA buffer (20 mM Tris-HCl, 150 mM NaCl, 1 mM Na2EDTA, 1 mM EGTA, 1% NP-40, and 1% sodium deoxycholate, 50

mM NaF, 2.5 mM sodium pyrophosphate, 1 mM beta-glycerophosphate, 1mM Na3VO4, and 1 ug/ml leupeptin) for 5 min on ice. Lysates were collected, sonicated on ice, and centrifuged at 14,000 x g for 10 min at 4 °C. Supernatants were collected and stored at -80 °C until use. Total protein was quantified (Pierce BCA Protein Assay Kit, Catalog number: 23225, Thermo Fisher) and samples were boiled for 5 min at 95 °C in 2X Laemmli Sample Buffer (BioRad) to denature protein. Equal amounts of protein were resolved on 4-20% Criterion TGX Protein Gels (BioRad) and transferred to PVDF membranes. Membranes were blocked for one hour at room temperature in blocking buffer (20 mM Tris HCl, pH 8.0; 150 mM NaCl; 0.1% Tween 20; 5% BSA). Membranes were assessed for NF1, p-ERK1/2, total ERK1/2, p-AKT, total AKT, and tubulin expression using recommended conditions and antibody dilutions. The primary antibodies used in this study were: NF1 (C terminus) #ab17963, P-p44/42 MAPK (Thr202/Tyr204) #4370, p44/42 MAPK #4695, Phospho-AKT (Ser473) #9271, pan-AKT #9272, Cell Signaling, α-tubulin mouse ab #T5168, Sigma Aldrich (1:1000 dilution), and beta actin (ab8227), Abcam (1:5000 dilution). When multiple primary antibodies were tested on the same blot, membranes were stripped twice using Re-Blot Plus Western Blot Strong Antibody Stripping Solution (1x) (Millipore) for 15-25 minutes at room temperature. These blots were blocked twice in blocking buffers for 5 minutes and re-probed with primary antibodies at recommended conditions and dilutions. Following incubation with primary antibodies, blots were washed with TBST (20 mM Tris HCl, pH 8.0; 150 mM NaCl; 0.1% Tween 20) and incubated with secondary antibodies (goat anti-rabbit IgG HRP-conjugated or goat anti-mouse IgG HRP-conjugated,1:10,000 dilution, BioRad) for one hour at room temperature. Membranes were washed with TBST, developed with Clarity Western ECL Blotting Substrate (BioRad), and imaged with a Chemi Doc XES+ System (BioRad).

Command lines used for processing the canine cancer cell line WES data.

Trimming of raw reads - Trimmomatic

```
java -jar -Xmx48G Trimmomatic-0.36/trimmomatic-0.36.jar PE -phred33 \
SampleName_R1.fastq.gz SampleName_R2.fastq.gz \
SampleName_R1_paired.fastq.gz SampleName_R1_unpaired.fastq.gz \
SampleName_R2_paired.fastq.gz SampleName_R2_unpaired.fastq.gz \
Trimmomatic-0.36/adapters/TruSeq3-SE.fa:2:30:10 \
CROP:130 LEADING:3 TRAILING:3 SLIDINGWINDOW:4:20 MINLEN:50
```

Mapping of trimmed reads against CanFam3.1 genome – BWA
Reference genome was downloaded from ftp://ftp.ensembl.org/pub/release-90/fasta/canis_familiaris/dna/ and indexed using BWA index tool

```
bwa mem -M -t 12 \
-R '@RG\tID:ID\tSM:SampleName\tPL:illumina\tLB:lib1\tPU:indexsequence' \
CanFam3 \
SampleName_R1_paired.fastq.gz SampleName_R2_paired.fastq.gz >
```

Sorting SAM file - Picard Tools

```
java -jar picard-tools-1.119/SortSam.jar \
TMP_DIR=tempSampleName \
I=SampleName_BWA.sam \
O=SampleName_BWA_sorted.sam \
SORT_ORDER=coordinate
```

Marking duplicates - Picard Tools

```
java -jar picard-tools-1.119/MarkDuplicates.jar \
TMP_DIR=tempSampleName \
I=SampleName_BWA_sorted.sam \
O=SampleName_dedup.bam \
M=SampleName_marked_dup_metrics.txt
```

Indexing BAM file - Samtools

```
samtools index SampleName_dedup.bam
```

GATK RealignerTargetCreator - GATK

```
java -jar GenomeAnalysisTK.jar \
-T RealignerTargetCreator \
-R Canis_familiaris.CanFam3.1.dna.toplevel.fa \
-I SampleName_dedup.bam \
-o SampleName_target_intervals.list
```

GATK realignment of indels - GATK

```
java -jar GenomeAnalysisTK.jar \
-T IndelRealigner \
-R Canis_familiaris.CanFam3.1.dna.toplevel.fa \
-I SampleName_dedup.bam \
-targetIntervals SampleName_target_intervals.list \
-o SampleName_realigned_reads.bam
```

GATK Recalibration of bases - GATK
The variant file - Canis_familiaris.vcf was downloaded from :-
ftp://ftp.ensembl.org/pub/release-90/variation/vcf/canis_familiaris/

```
java -jar GenomeAnalysisTK.jar \
-T BaseRecalibrator \
-R Canis_familiaris.CanFam3.1.dna.toplevel.fa \
-I SampleName_realigned_reads.bam \
-knownSites Canis_familiaris.vcf \
-o SampleName_recal_data.grp
```

GATK Print reads - GATK

```
java -jar GenomeAnalysisTK.jar \
-T PrintReads \
-R Canis_familiaris.CanFam3.1.dna.toplevel.fa \
-I SampleName_realigned_reads.bam \
-BQSR SampleName_recal_data.grp \
-o SampleName_recal_reads.bam
```

Coverage - GATK
The bed file provides chromosomal intervals used for creating the Agilent Sure Select exome capture

```
java -jar GenomeAnalysisTK.jar \
-T DepthOfCoverage \
-R Canis_familiaris.CanFam3.1.dna.toplevel.fa \
-I SampleName_recal_reads.bam \
-L Dog_Exome.bed \
--outputFormat table \
-o SampleName \
-l DEBUG
```

Extracting raw variants - Freebayes

```
freebayes -f Canis_familiaris.CanFam3.1.dna.toplevel.fa \
--min-alternate-fraction 0.05 \
--min-mapping-quality 50 \
--min-alternate-count 2 \
SampleName_recal_reads.bam > SampleName_freebayes.vcf
```

Annotation of known variants - SnpEff (using version 89 of Ensembl; NCBI 146 built)

```
java -Xmx10g -jar SnpSift.jar annotate \
-id Canis_familiaris.vcf \
SampleName_freebayes.vcf > SampleName_freebayesSS.ann.vcf
```

Adding information on variant type - SnpEff

```
java -Xmx10g -jar SnpSift.jar varType \
SampleName_freebayesSS.ann.vcf >SampleName_freebayesSS-VT.ann.vcf
```

Annotating variants – SnpEff

```
java -Xmx10g -jar snpEff.jar -stats SampleName_snpEff_summary.html \
CanFam3.1.86 SampleName_freebayesSS-VT.ann.vcf > \
SampleName_freebayes.snpEff.ann.vcf
```

Filtering variants with DP>=10 - SnpEff

```
cat SampleName_freebayes.snpEff.ann.vcf | \
java -Xmx4g -jar SnpSift.jar filter " ( DP >= 10 )" \
> SampleName_freebayes.snpEff-filtDP.vcf
```

Filtering variants with QUAL>=20 - SnpEff

```
cat SampleName_freebayes.snpEff-filtDP.vcf | \
java -Xmx4g -jar SnpSift.jar filter " ( QUAL >= 20 )" \
> SampleName_freebayes.snpEff-filtDP-QUAL.vcf
```

Zipping and indexing vcf file - bgzip and tabix

```
bgzip -c SampleName_freebayes.snpEff-filtDP-QUAL.vcf > \
SampleName_freebayes.snpEff-filtDP-QUAL.vcf.gz

tabix -p vcf SampleName_freebayes.snpEff-filtDP-QUAL.vcf.gz
```

Eliminating Germline/Normal variants - vcftools

```
vcf-isec -c -f -d -o SampleName_freebayes.snpEff-filtDP-QUAL.vcf.gz \
Normals.vcf.gz \
| bgzip -c > SampleName_VCFisec.vcf.gz
```

Converting vcf to tab format - bcftools

```
bcftools query -H -f
'%CHROM\t%POS\t%ID\t%REF\t%ALT\t%FILTER[\t%GT\t%DP\t%AD\t%AO\t%
GL]\t%INFO/AF\t%INFO/AC\t%INFO/MQM\t%INFO/VARTYPE\t%INFO/ANN\n'
-u -o SampleName.tab SampleName_VCFisec.vcf.gz
```

SIFT prediction - SIFT4G

```
java -jar -Xmx1024m SIFT4G_Annotator_v2.3.jar -c \
-i SampleName_freebayes_VCFisec-filtered.vcf \
-d SIFT4G_database/CanFam3.1.83/ \
-r SIFT_folder -t
```

Generating lollipop plots – lollipops

```
lollipops -show-motifs -dpi=300 -w=700 -labels -no-patterns -legend
-domain-labels=off -U <Uniprot_ID> <Variations>
```

Section 2. Supplementary Tables.

Supplementary Table S3. Tools and databases used in this study.

Tool/Resource/Database	Version	Reference / Download URL
FASTQC	0.11.5	Andrews 2010 (8)
Trimmomatic	0.36	Bolger et al 2014 (9)
BWA	0.7.15	Li and Durbin 2010 (10)
samtools	1.4.1	Li 2009 (11)
Picard-tools	1.119	http://broadinstitute.github.io/picard/
GenomeAnalysisToolKit	3.7-0-gcfedb67	Van der Auwera et al 2013 (12)
Freebayes	v1.1.0-3-g961e5f3	Garrison and Marth 2012 (13)
SnpEff	4.3p (2017-06-06)	Cingolani et al 2012 (14)
Tabix	1.6-14-g199ab97	Li 2011 (15)
VCFtools	0.1.13	Danecek et al 2011 (16)
bcftools	1.6	https://github.com/samtools/bcftools
SIFT	2.3	Kumar et al 2009 (17)
Lollipops	1.3.2	Jay and Brouwer (18)
R	3.4.4	R Core Team 2018 (19)
ComplexHeatmap	1.17.1	Gu et al 2016 (20)
Canis lupus familiaris genome	Ensembl v89	ftp://ftp.ensembl.org/pub/release-89/fasta/canis_familiaris/dna/Canis_familiaris.CanFam3.1.dna.toplevel.fa.gz
dbSNP vcf file	Ensembl v90; dbSNP built 146	ftp://ftp.ensembl.org/pub/release-89/variation/vcf/canis_familiaris/Canis_familiaris.vcf.gz
COSMIC database	v83; downloaded June 12th 2018	https://cancer.sanger.ac.uk/cosmic/download
Dog SNP database	Downloaded July 10th 2017	ftp:/ai et al 2015/download.big.ac.cn/dogsd/dogsd_total_VCF/DOGSD-vcf.gz ; Bai et al 2015 (21)
David Functional Annotation Tool	6.7	Huang et al. 2009 (4)

Supplementary Table S4. Primers used for validating variants.

Gene Name	Exon	5' Primer	3' Primer
CUX1	17	AGGACGCCAAGTCAGAAGAA	AGCTCAATGGTGTCCACCTC
FAT3	27	TAGACAGCGCAAGTGGCTTA	TAGACAGCGCAAGTGGCTTA
KIT1	11	CATTTGTTCTCTACCTAAGTG	GTTCCCTAAAGTCATTGTTACA
KRAS	1	TCAAATGTTGTTGGGGGAGT	TAAAGCCAATGGAACCCAAG
KRAS	2	TGTCACCTTTGGAGGAATGTC	AAGCCAAGGTTTCAATTCCA
KTM2D	39	AGGGTCCTCACAGACAGGTG	TGCTGCTGTGGAGACAGTAAA
MET1	17	CCCCTGTGCTTTGCTGTTTC	CTAGGCGGTCGCTAAAGAGG
NRAS	2	GCATTGCATTCCCTATGGTT	GCGGTATCCTCATTTCCTGT
PIK3CA	21	ATGCCAGTTTCTGCCTGTCT	CACACTGCTGAACCAGTCAAA
TP53	4	GACCTGACCCTTGACTCTGG	GGACCGAGAGGAGTCTCAAA
TP53	5-6	CGCGCTATGGCCATCTATAA	ATTTCCTTCCACTCGGATGAG
TP53	7	TGGGAAAGACTGAGGCTGAT	GGGGGTACAGTAAGGAAGTGG
TP53	8	TGAGGGTGGCTAGGAGTCAG	TCCCTCCTTCACCTCCTCTT

Supplementary Table S6. Putative driver gene mutations.

Gene Name	Cell Lines (Amino Acid Change)	Alternate Allele Frequency	SIFT Score	Functional Category
ARID1A	Tyler2 (A901fs); Tyler1 (A901fs)	0.49; 0.38	NA	CBO
ASXL1	OSW (G618fs); Jones (G618fs)	0.45; 0.41	NA	CBO
ATRX	Tyler1 (C1810*); Tyler2 (C1810*)	1; 0.99	NA	CBO
BRAF	Vogel (V588E); Tyler1 (V588E); Tyler2 (V588E); Bliley (V588E)	0.3; 0.41; 0.38; 0.43	0.002	KA; GA
CCND3	CLBL1 (R358fs)	0.51	NA	KA
CDKN1A	Parks (D33fs)	0.83	NA	KA
CIC	1771 (P1551fs)	0.68	NA	CBO
EIF1AX	Parks (G6S)	0.72	0.098	RB
EP300	SB (N1570fs)	0.29	NA	CBO
ERBB2	CTAC (V659E)	1	0.011	KA
FBXW7	OSW (R694Q)	0.49	0.006	KA
KDM6A	OSA8 (T986fs); CLL1390 (V1131fs)	1; 0.97	NA	CBO
KIT	17CM98 (V558_E560del); BRMCT (L575P)	0.6; 0.47	NA; 0	KA
KMT2C	OSW (G976fs); CMT12 (R1201fs)	0.41; 0.92	NA	CBO
KMT2D	OSW (Q5395fs)	0.99	NA	CBO
KRAS	D-17 (E63K); Angus (G12D); CML-10C2 (G12V)	0.19; 0.63; 0.76	0.002; 0; 0.001	KA; GA
MED12	Vogel (L1349fs); McKinley (L1938fs); Parks (P22S)	0.41; 0.97; 0.48	NA; 0	CBO
NF1	OSW (N58fs); STSA-1 (M877fs); CML-6M (P1599fs); Nike (A2176fs)	0.99; 0.48; 0.44; 1	NA	KA
NRAS	Jones (Q61R); Moresco (Q61H)	0.46; 0.53	0.048; 0.002	KA; GA
PIK3CA	SB (E453K)	1	0.184	KA
PPM1D	McKinley (E472*)	0.49	NA	CBO
PTEN	OSW (K245fs); Jones (D246fs)	1; 0.43	NA	KA
PTPN11	DH82 (G503V)	0.99	0.008	KA
RAC1	Moresco (P29T)	0.35	0.01	KA; GA
SETD2	BRMCT (R526fs)	0.99	NA	CBO
SPEN	OSW (A8fs)	0.36	NA	RB
SRGAP3	Gracie (V850fs)	0.45	NA	GA
TET2	CLL1390 (Q965*)	0.99	NA	CBO
TP53	BRMCT (A193D, R263C); SB (L184F, A150fs); Vogel (V162L); Abrams (C267F); HMPOS (R164H); CLBL1 (R257H; A127V); OS2.4 (I222T); OSW (R164H); CMT12 (V132E); C2 (V162A); CLL1390 (D60fs); CTAC (T114M)	0.48, 0.46; 0.51, 0.48; 0.75; 0.51; 1; 0.46, 0.45; 1; 0.56; 1; 1; 0.63; 1	0.003, 0; 0.096, NA; 0.004; 0.001; 0.001; 0.003, 0.037; 0.018;; 0.001; 0.001; 0.002; NA; 0	CBO

CBO - Chromatin binding/organization; KA – regulation of kinase activity; GA – GTPase binding and activity; RB – RNA binding

Supplementary Table S8. Mutated MAPK genes mapped to KEGG pathway (cfa:04010).
The KEGG Gene Name corresponds to the gene nomenclature used in the KEGG pathway map (Supplementary Figure S4).

Gene Name	KEGG Gene Name	Cell Line	Cancer Type	Protein Variant
BRAF	RafB	Tyler1	Transitional Cell Carcinoma	V588E
BRAF	RafB	Bliley	Transitional Cell Carcinoma	V588E
BRAF	RafB	Tyler2	Transitional Cell Carcinoma	V588E
BRAF	RafB	Vogel	Osteosarcoma	V588E
CACNA1A	CANC	OSW	Leukemia / Lymphoma	A715V
CACNA1C	CANC	CLL1390	Leukemia / Lymphoma	R1374S
CACNA1C	CANC	Tyler2	Transitional Cell Carcinoma	M2114I
CACNA1D	CANC	Jones	Melanoma	R1531G
CACNA1E	CANC	Tyler2	Transitional Cell Carcinoma	P109S
CACNA1E	CANC	Tyler1	Transitional Cell Carcinoma	P109S
CACNA1F	CANC	Parks	Melanoma	P1492L
CACNA1H	CANC	OSW	Leukemia / Lymphoma	R271Q
CACNA2D2	CANC	D-17	Osteosarcoma	E451G
CACNA2D3	CANC	Abrams	Osteosarcoma	L598fs
CDC25B	Cdc25B	Kinsey	Transitional Cell Carcinoma	P30L
CHUK	IKK	D-17	Osteosarcoma	L37P
DAXX	DAXX	OS2-4	Osteosarcoma	R229fs
DUSP16	MKP	CLL1390	Leukemia / Lymphoma	G385A
DUSP5	MKP	D-17	Osteosarcoma	T321M
DUSP9	MKP	Parks	Melanoma	R227C
EGFR	RTK	SB	Hemangiosarcoma	R1039W
FGF17	GF	CTAC	Thyroid Carcinoma	R208C
FOS	c-fos	Kinsey	Transitional Cell Carcinoma	K250fs
KRAS	Ras	CML-10C2	Melanoma	G12V
KRAS	Ras	Angus	Transitional Cell Carcinoma	G12D
KRAS	Ras	D-17	Osteosarcoma	E63K
MAP3K4	MEKK4	BRMCT	Mast Cell Tumor	P32_P33dup
MAP3K4	MEKK4	D-17	Osteosarcoma	P45_E46dup
MAP3K4	MEKK4	McKinley	Osteosarcoma	P45_E46dup
MAP3K5	ASK1	C2	Mast Cell Tumor	V1121_V1122insVLF
MAP3K6	ASK2	Kinsey	Transitional Cell Carcinoma	R531W
MAP3K6	ASK2	1771	Leukemia / Lymphoma	V833fs
MAPK1	ERK	CLL1390	Leukemia / Lymphoma	M1_A2insA
MAPT	Tau	Kinsey	Transitional Cell Carcinoma	S609N
MECOM	Evil	Moresco	Osteosarcoma	K561N
NF1	NF1	OSW	Leukemia / Lymphoma	N58fs
NF1	NF1	STSA-1	Soft Tissue Sarcoma	M877fs
NF1	NF1	Nike	Histiocytic Sarcoma	A2176fs
NF1	NF1	CML-6M	Melanoma	P1599fs
NFATC1	NFAT-2	Vogel	Osteosarcoma	T910M
NFKB1	NF**κ**B	Bliley	Transitional Cell Carcinoma	V21fs
NR4A1	Nur77	HMPOS	Osteosarcoma	R441H
NRAS	Ras	Jones	Melanoma	Q61R
NRAS	Ras	Moresco	Osteosarcoma	Q61H
PLA2G4E	cPLA2	Tyler1	Transitional Cell Carcinoma	I15V
PLA2G4E	cPLA2	Tyler2	Transitional Cell Carcinoma	I15V
PPP3CB	PPP3C	Parks	Melanoma	L140S
RAC1	Cdc42/Rac	Moresco	Osteosarcoma	P29T
RASGRF2	RasGRF	17CM98	Melanoma	T499A
RPS6KA2	RSK2	OSW	Leukemia / Lymphoma	G102fs
TAOK1	TAO	Parks	Melanoma	R562*

Supplementary Table S9. Mutated PI3K-AKT genes mapped to KEGG pathway (cfa:04151). The KEGG Gene Name corresponds to the gene nomenclature used in the KEGG pathway map (Supplementary Figure S5).

Gene Name	KEGG Gene Name	Cell Line	Cancer Type	Protein Variant
PRKAA1	AMPK	C2	Mast Cell Tumor	V39F
MYB	C-Myb	CLL1390	Leukemia / Lymphoma	N68I
RAC1	Cdc42/Rac	Moresco	Osteosarcoma	P29T
THEM4	CTMP	Nike	Histiocytic Sarcoma	S171L
CCND3	Cyclin	C2	Mast Cell Tumor	A66_S67del
CCND3	Cyclin	CLBL1	Leukemia / Lymphoma	R358fs
CSF3R	CytokineR	Parks	Melanoma	L825F
COL1A2	ECM	Nike	Histiocytic Sarcoma	E1230Q
COL4A2	ECM	Tyler2	Transitional Cell Carcinoma	G755E
COL4A2	ECM	Tyler1	Transitional Cell Carcinoma	G755E
COL4A4	ECM	OS2.4	Osteosarcoma	P1204S
COL4A4	ECM	Nike	Histiocytic Sarcoma	G918R
COL6A3	ECM	CLBL1	Leukemia / Lymphoma	G2293*
COL6A5	ECM	Angus	Transitional Cell Carcinoma	R1207fs
COL6A5	ECM	CLL1390	Leukemia / Lymphoma	I847F
LAMA1	ECM	McKinley	Osteosarcoma	D2753H
LAMA4	ECM	Parks	Melanoma	S495L
LAMC2	ECM	HMPOS	Osteosarcoma	S23P
LAMC2	ECM	Nike	Histiocytic Sarcoma	I955T
LAMC2	ECM	Nike	Histiocytic Sarcoma	N958T
RELN	ECM	Moresco	Osteosarcoma	L972H
THBS2	ECM	OSW	Leukemia / Lymphoma	R482C
TNC	ECM	Tyler2	Transitional Cell Carcinoma	T1394M
TNC	ECM	1771	Leukemia / Lymphoma	G2179R
TNC	ECM	Tyler1	Transitional Cell Carcinoma	T1394M
VWF	ECM	CMT12	Mammary Carcinoma	S182_P184del
VWF	ECM	Parks	Melanoma	W377*
EIF4B	eIF4B	Parks	Melanoma	R182C
MAPK1	ERK	CLL1390	Leukemia / Lymphoma	M1_A2insA
GNG3	GBG	Nike	Histiocytic Sarcoma	F66L
ANGPT4	GF	OSW	Leukemia / Lymphoma	R494C
FGF17	GF	CTAC	Thyroid Carcinoma	R208C
CHUK	IKK	D-17	Osteosarcoma	L37P
ITGA1	ITGA	D-17	Osteosarcoma	A597T
ITGA6	ITGA	Nike	Histiocytic Sarcoma	I859S
ITGA9	ITGA	OSW	Leukemia / Lymphoma	T582fs
ITGAV	ITGA	DEN-HSA	Hemangiosarcoma	M956_P957dup
ITGB7	ITGB	Tyler1	Transitional Cell Carcinoma	R247L
ITGB7	ITGB	Nike	Histiocytic Sarcoma	L133fs
ITGB7	ITGB	Nike	Histiocytic Sarcoma	A132fs
ITGB7	ITGB	C2	Mast Cell Tumor	A132fs
ITGB7	ITGB	Parks	Melanoma	L133fs
ITGB7	ITGB	Parks	Melanoma	A132fs
ITGB7	ITGB	Tyler2	Transitional Cell Carcinoma	R247L
ITGB7	ITGB	C2	Mast Cell Tumor	L133fs
JAK3	JAK	Tyler2	Transitional Cell Carcinoma	W993*
MTOR	mTOR	1771	Leukemia / Lymphoma	R32Q
NFKB1	NFKB	Bliley	Transitional Cell Carcinoma	V21fs
NR4A1	NUR77	HMPOS	Osteosarcoma	R441H

Gene Name	KEGG Gene Name	Cell Line	Cancer Type	Protein Variant
CDKN1A	p21	Parks	Melanoma	D33fs
CDKN1B	p27; p27/Kip1	CLL1390	Leukemia / Lymphoma	Ter199fs
PCK2	PEPCK	Parks	Melanoma	R276W
PHLPP1	PHLPP	HMPOS	Osteosarcoma	A157_P158insP
PHLPP1	PHLPP	CTAC	Thyroid Carcinoma	G971R
PHLPP1	PHLPP	STSA-1	Soft Tissue Sarcoma	P84_A85del
PKN1	PKN	Nike	Histiocytic Sarcoma	C1145Y
PPP2R2B	PP2A	Parks	Melanoma	D215V
PPP2R3A	PP2A	Kinsey	Transitional Cell Carcinoma	R981H
PTEN	PTEN	OSW	Leukemia / Lymphoma	K245fs
PTEN	PTEN	Jones	Melanoma	D246fs
KRAS	Ras	Angus	Transitional Cell Carcinoma	G12D
KRAS	Ras	CML-10C2	Melanoma	G12V
KRAS	Ras	D-17	Osteosarcoma	E63K
NRAS	Ras	Moresco	Osteosarcoma	Q61H
NRAS	Ras	Jones	Melanoma	Q61R
EGFR	RTK	SB	Hemangiosarcoma	R1039W
KIT	RTK	BRMCT	Mast Cell Tumor	L575P
KIT	RTK	17CM98	Melanoma	V558_E560del
MET	RTK	HMPOS	Osteosarcoma	R1188Q
SGK1	SKG	OSW	Leukemia / Lymphoma	F467fs
TLR4	TLR2/4	C2	Mast Cell Tumor	Y38C
EPHA2	RTK	SB	Hemangiosarcoma	G275S
EPHA2	RTK	DEN-HSA	Hemangiosarcoma	D77V
COL6A2	ECM	Vogel	Osteosarcoma	S56F

References

1. CC, Drexler HG. Detection of Mycoplasma contamination in cell cultures. Curr Protoc Mol Biol **2014**;106:28.4.1-14 doi 10.1002/0471142727.mb2804s106.
2. Cooper JK, Sykes G, King S, Cottrill K, Ivanova NV, Hanner R*, et al.* Species identification in cell culture: a two-pronged molecular approach. In Vitro Cell Dev Biol Anim **2007**;43(10):344-51 doi 10.1007/s11626-007-9060-2.
3. O'Donoghue LE, Rivest JP, Duval DL. Polymerase chain reaction-based species verification and microsatellite analysis for canine cell line validation. J Vet Diagn Invest **2011**;23(4):780-5 doi 10.1177/1040638711408064.
4. Huang da W, Sherman BT, Zheng X, Yang J, Imamichi T, Stephens R*, et al.* Extracting biological meaning from large gene lists with DAVID. Current protocols in bioinformatics **2009**;Chapter 13:Unit 13.1 doi 10.1002/0471250953.bi1311s27.
5. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM*, et al.* Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. Nature genetics **2000**;25(1):25-9 doi 10.1038/75556.
6. Kanehisa M, Goto S. KEGG: kyoto encyclopedia of genes and genomes. Nucleic acids research **2000**;28(1):27-30.
7. Finn RD, Bateman A, Clements J, Coggill P, Eberhardt RY, Eddy SR*, et al.* Pfam: the protein families database. Nucleic acids research. Volume 422014. p D222-30.
8. Andrews S. FastQC A Quality Control tool for High Throughput Sequence Data (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/). 2010.
9. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics (Oxford, England) **2014**;30(15):2114-20 doi 10.1093/bioinformatics/btu170.
10. Li H, Durbin R. Fast and accurate long-read alignment with Burrows-Wheeler transform. Bioinformatics (Oxford, England) **2010**;26(5):589-95 doi 10.1093/bioinformatics/btp698.
11. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N*, et al.* The Sequence Alignment/Map format and SAMtools. Bioinformatics (Oxford, England) **2009**;25(16):2078-9 doi 10.1093/bioinformatics/btp352.
12. Van der Auwera GA, Carneiro MO, Hartl C, Poplin R, Del Angel G, Levy-Moonshine A*, et al.* From FastQ data to high confidence variant calls: the Genome Analysis Toolkit best practices pipeline. Current protocols in bioinformatics **2013**;43:11.0.1-33 doi 10.1002/0471250953.bi1110s43.
13. Garrison E, Marth G. Haplotype-based variant detection from short-read sequencing. **2012**.
14. Cingolani P, Platts A, Wang le L, Coon M, Nguyen T, Wang L*, et al.* A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3. Fly **2012**;6(2):80-92 doi 10.4161/fly.19695.
15. Li H. Tabix: fast retrieval of sequence features from generic TAB-delimited files. Bioinformatics (Oxford, England) **2011**;27(5):718-9 doi 10.1093/bioinformatics/btq671.
16. Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA*, et al.* The variant call format and VCFtools. Bioinformatics (Oxford, England) **2011**;27(15):2156-8 doi 10.1093/bioinformatics/btr330.

17. Kumar P, Henikoff S, Ng PC. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. Nature protocols **2009**;4(7):1073-81 doi 10.1038/nprot.2009.86.
18. Jay JJ, Brouwer C. Lollipops in the Clinic: Information Dense Mutation Plots for Precision Medicine. PloS one **2016**;11(8):e0160519 doi 10.1371/journal.pone.0160519.
19. Team RC. 2018 R: The R Project for Statistical Computing. <https://www.r-project.org/>.
20. Gu Z, Eils R, Schlesner M. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. Bioinformatics (Oxford, England) **2016**;32(18):2847-9 doi 10.1093/bioinformatics/btw313.
21. Bai B, Zhao WM, Tang BX, Wang YQ, Wang L, Zhang Z*, et al.* DoGSD: the dog and wolf genome SNP database. Nucleic acids research **2015**;43(Database issue):D777-83 doi 10.1093/nar/gku1174.