Journal of Veterinary Diagnostic Investigation
2021, Vol. 33(2) 261–278

DOI: 10.1177/1040638720985242
jvdi.sagepub.com

Novel application of single-cell next-generation sequencing for determination of intratumoral heterogeneity of canine osteosarcoma cell lines

Jordan Ayers, Rowan J. Milner,[1] Galaxia Cortés-Hinojosa, Alberto Riva, Sandra Bechtel, Bikash Sahay, Matthew Cascio, Amandine Lejeune, Keijiro Shiomitsu, Carlos Souza, Oscar Hernandez, Marc Salute

Abstract. Osteosarcoma (OSA) is a highly aggressive and metastatic neoplasm of both the canine and human patient and is the leading form of osseous neoplasia in both species worldwide. To gain deeper insight into the heterogeneous and genetically chaotic nature of OSA, we applied single-cell transcriptome (scRNA-seq) analysis to 4 canine OSA cell lines. This novel application of scRNA-seq technology to the canine genome required uploading the CanFam3.1 reference genome into an analysis pipeline (10X Genomics Cell Ranger); this methodology has not been reported previously in the canine species, to our knowledge. The scRNA-seq outputs were validated by comparing them to cDNA expression from reverse-transcription PCR (RT-PCR) and Sanger sequencing bulk analysis of 4 canine OSA cell lines (COS31, DOUG, POS, and HMPOS) for 11 genes implicated in the pathogenesis of canine OSA. The scRNA-seq outputs revealed the significant heterogeneity of gene transcription expression patterns within the cell lines investigated (COS31 and DOUG). The scRNA-seq data showed 10 distinct clusters of similarly shared transcriptomic expression patterns in COS31; 12 clusters were identified in DOUG. In addition, cRNA-seq analysis provided data for integration into the Qiagen Ingenuity Pathway Analysis software for canonical pathway analysis. Of the 81 distinct pathways identified within the clusters, 33 had been implicated in the pathogenesis of OSA, of which 18 had not been reported previously in canine OSA.

Key words: canine; neoplasia; osteosarcoma; single-cell transcriptomics; tumor heterogeneity.

Introduction

Osteosarcoma (OSA) is the most common primary osseous malignancy of the canine patient; OSA is highly aggressive, both locally and systemically.[13,21] Canine OSA (cOSA) is commonly used as a model system for human (pediatric) OSA (hOSA) given strong parallels in biologic behavior and response to treatment.[9,70] OSA seeds the body with micrometastatic lesions (primarily to the lungs) that can avoid detection prior to removal of the primary tumor. These lesions may then develop into metastatic lesions following removal of the primary tumor. The addition of chemotherapy following amputation is necessary to combat this metastatic progression. With chemotherapy and surgery, canine patients have reported median survival times of 9–15.7 mo, with 30–37% of patients alive at 1 y.[57] A sharp drop-off in survival is evident once metastatic lesions are diagnosed, reducing survival times to 2–3 mo.[5,6,71] In hOSA, 68% of patients survive for 5 y, and that survival rate decreases dramatically to < 30% once metastasis is present.[13,14,54] Although the addition of chemotherapy to OSA treatment protocols > 30 y ago improved survival outcomes for cOSA and hOSA, these survival times have sadly remained stagnant.[6,8,53,54]

OSA is a highly heterogeneous and genetically chaotic tumor type.[61] Whole-genomic sequencing paired with whole-exome sequencing in hOSA has revealed distinct classes of DNA mutations, with 14 genes revealed to be the main drivers of 87% of hOSA.[61] Canine OSA is also genetically chaotic and possesses a high mutational load that shares multiple genetic mutations with hOSA.[18,63,67,70] In the genetic assessment of tumor samples using second-generation or next-generation sequencing (NGS) techniques, bulk sequencing techniques are commonly employed. The term "bulk sequencing" refers to the technique by which the target tissue is lysed as a whole, and the DNA (or RNA) is an amalgamation of all of the DNA (or RNA) of those cells present.[4,60]

Departments of Small Animal Clinical Sciences (Ayers, Cortés-Hinojosa, Salute, Bechtel, Lejeune, Shiomitsu, Souza, Hernandez, Milner) and Infectious Diseases and Immunology (Sahay), College of Veterinary Medicine; and Pediatric Hematology-Oncology, Department of Pediatrics, College of Medicine (Cascio) and ICBR Bioinformatics Core (Riva), University of Florida, Gainesville, FL.

[1]Corresponding author: Rowan J. Milner, Department of Small Animal Clinical Sciences, University of Florida, PO Box 100126, Gainesville, FL 32610. milnerr@ufl.edu

When second-generation or NGS is performed, the expression patterns observed reflect the majority of the cells—but the degree of majority is unknown, and less prominent mutations and biomarkers may be hidden or appear less significant. As an adjunct to bulk analysis, single-cell RNA sequencing (scRNA-seq) has been developed. ScRNA-seq is the pooled sequencing of a sample in which the cells are individually barcoded, allowing the transcriptomes to be separated on a cell-by-cell level to determine intra-tumoral heterogeneity of transcribed genes.[76] Thus, scRNA-seq can identify complex and rare cell populations, expose regulatory relationships between genes, and keep track of the development of distinct cell lineages.[24]

Using established OSA cell lines in in-vitro assays is an important first step toward understanding the molecular biology of OSA while controlling for confounding variables associated with in vivo models (e.g., DNA or RNA from normal stromal cells). In our laboratory at the University of Florida (UFL; Gainesville, FL), we have the following cOSA cell lines: COS31, POS, and DOUG, which are derived from primary cOSA tumor sites.[25,68] We also have the highly metastatic HMPOS cell line, which was reported to be derived from the primary POS cell line using a mouse xenograft model.[3] No formal studies have been performed to characterize the transcriptome of these cell lines using NGS. Thus, our overall goal was to characterize the cOSA cell lines (COS31, DOUG, POS, and HMPOS) using scRNA-seq to detect possible tumor heterogeneity and the possible existence of rare populations of cOSA cells. Our first aim, in order to verify the future scRNA-seq results, was to use validated reverse-transcription PCR (RT-PCR) methods and Sanger sequencing to interrogate the cDNA of the 4 cOSA cell lines for oncogenes and tumor suppressor genes known to be implicated in the pathogenesis of cOSA (*PTEN*, *P53*, *P16*, *MYC*, *P21*, *IGF1*, *ERBB2*, *RB1*, *EGFR*, *MET*, and *STAT3*).[14,15,20,21,30,41,44,47,61,66] Our second aim was to use the 10X Genomics single-cell transcriptomics solution (scRNA-seq platform) to characterize and profile gene expression in the DOUG and COS31 cell lines using the canine reference genome CanFam3.1.[22] We then compared scRNA-seq results to the RT-PCR and Sanger results to validate the results for COS31 and DOUG cell lines. In addition, we hoped with the analysis of the scRNA-seq data to identify novel targets for therapy, biomarkers for prognosis and diagnosis, and cell types without the constraints of pre-selecting the targets or masking smaller populations.[37,80]

The innovative nature of scRNA-seq data allowed us to perform pathway analysis on the DOUG and COS31 cell lines. The novelty of our research includes reporting on the previously unknown transcriptomic data of 4 commonly used canine cell lines, the validation of the scRNA-seq pipeline using the 10X Genomics platform for the identification and expression of canine genes and new pathways implicated in cOSA, and the confirmation of significant tumor expression heterogeneity in cOSA.

Materials and methods

Bulk sequencing of OSA cell lines using RT-PCR and Sanger sequencing

Study population. To establish the study population, we cultured 4 cOSA cell lines: COS31,[68] POS,[25] HMPOS,[3] and DOUG. COS31 is a cell line isolated from a primary OSA of the left distal radius of an 8-y-old male St. Bernard dog.[68] POS was established from a primary OSA lesion of the left femur of a 1.5-y-old male mixed-breed dog; HMPOS (highly metastasizing POS) was derived from metastatic lung lesions using the POS cell line in a xenografted mouse model.[3] DOUG was established at UFL from a primary OSA lesion of the right proximal humerus of an 8-y-old castrated male Rottweiler dog. All 4 cell lines were tested for *Mycoplasma* spp. contamination using PCR methods and ELISA (MycoAlert mycoplasma detection kit; Lonza) and were found to be negative. To confirm the cell lines as canine in origin, Sanger sequencing results from all 4 cell lines were compared to the mouse, human, and canine genome; homology was 99% for all cell lines compared to the canine genome.

Cells were seeded into 6-well plates, with triplicates of each cell line plated and extracted. COS31 was maintained in Dulbecco modified Eagle medium (DMEM) supplemented with 10% non–heat-inactivated fetal bovine serum (10% FBS; Cellgro Mediatech). POS and HMPOS were maintained in Roswell Park Memorial Institute medium supplemented with 10% heat-inactivated FBS (10% HI-FBS; MilliporeSigma) and the addition of L-glutamine, sodium pyruvate, vitamin solution, nonessential amino acids, and penicillin–streptomycin. DOUG was maintained in DMEM and 10% HI-FBS. Cells were grown to 75–80% confluence to reduce cellular stress, then washed with Hanks balanced salt solution (BSS) and detached with 0.25% trypsin. Cells were counted on an automated cell counter (Cellometer; Nexelcom Bioscience) with dual fluorescence to determine the live-to-dead ratios and cell concentration.

RNA isolation and cDNA generation

RNA was extracted from the cOSA samples (RNeasy plus mini kit; Qiagen) per the manufacturer's instructions. The final product was eluted with RNase-free water. Purity (via the 260/280 nm ratio of ~2.0) and concentration were determined (NanoDrop 8000 spectrophotometer; Thermo Fisher). cDNA was then generated (QuantiTect reverse transcription kit; Qiagen) per the manufacturer's instructions, and incubated (42°C for 30 min, 95°C for 3 min) to inactivate the reverse transcriptase. Then, 30 µL of DNase-free water was added to the samples, and purity and concentration were determined via spectrophotometry with a 260/280 ratio of ~1.8.

Primer design for RT-PCR

Based on a literature review of clinically and biologically relevant mutations in cOSA and hOSA, 11 genes were chosen for RT-PCR in all cell lines, including *PTEN*, *P53*, *P16*, *MYC*, *P21*, *IGF1*, *ERBB2*, *RB1*, *EGFR*, *MET*, and *STAT3*.[14,15,21,31,41,45,47,48,62,66,70] Primers were designed using BLAST (http://blast.ncbi.nlm.nih.gov/Blast.cgi; Suppl. Table 1). Primers were selected to cover 350–700 bp of the cDNA sequence of the selected gene. Sequences from the CanFam3.1 genome[22] were analyzed using Primer3 software (primer3.org), and 18–20-bp primers were chosen that yielded the most stable products (melting temperature close to 60°C, 40–60% GC content, moderate specificity to bind at the 3'-end). The primers were then run through a virtual PCR software (Amplify 4; University of Wisconsin), and the primers selected were confirmed to have no mispriming to any other section of the sequence and no hairpin formation.

RT-PCR and gel extraction

The cDNA generated underwent RT-PCR with the primers for the gene of interest and SYBR Green RT-PCR master mix (Thermo Fisher), according to the manufacturer's instructions. The reference gene, glyceraldehyde 3-phosphate dehydrogenase (GAPDH), was used as an internal control. The RT-PCR thermocycler protocol (SimpliAmp thermal cycler; Thermo Fisher) utilized was 95°C for 10 min to activate the *Taq* polymerase, followed by 40 cycles of 95°C for 15 s for denaturing, and then 60°C for 1 min to anneal and synthesize. The products were then run in 1% agarose gel at 90V until the product was approximately one-half to three-quarters down the gel (Fig. 1A–D). The gels were cut at the location of the bands, and cDNA was extracted (Gel extraction kit; Qiagen) per the manufacturer's instruction. The quantity and purity of the cDNA were assessed (NanoDrop 8000 spectrophotometer) and submitted for Sanger sequencing (Genewiz).

Sequences were then cross-referenced against the canine genome (*Canis lupus familiaris* CanFam 3.1)[22] using Geneious Prime (v.2019) to determine the degree of homology (Table 1). All sequences were run in triplicate for all 4 cell lines. If a gene was not present, the process was repeated in triplicate. All reactions were run with a negative control eluted with RNase-free water.

Characterization of cOSA cell lines using 10X Genomics single-cell transcriptomics solution

Study population. Only COS31 and DOUG were selected for 10X Genomics single-cell analysis given the cost associated with single-cell analysis and NextSeq 550 high-throughput sequencing. DOUG was selected because the Milner Comparative Laboratory (UFL) has histopathology samples from the original tumor, allowing for further investigation. The cells were cultured as described in the bulk sequencing section. After removal of media and rinsing with Hanks BSS, the cells were incubated with trypsin at 37°C for 5 min. The cells were washed twice following centrifugation at 168 rcf for 5 min, and then cells were resuspended and counted. The cells were centrifuged again at 168 rcf for 3 min and the supernatant removed. One mL of a mixture of 1× PBS with bovine serum albumin (0.04% or 400 µg/mL) was added, and second centrifugation (168 rcf × 3 min) was performed before the cells were strained (Flomi cell strainer tip; Bel-Art). The cells were counted, and cell viability was determined using an automated cell counter (Cellometer) using dual field fluorescence and acridine orange–propidium iodide staining. The samples were verified to have >90% of cells alive, with an ideal concentration of 1,200–1,600 cells/µL.

Library preparation and Illumina high-throughput sequencing

The libraries for both COS31 and DOUG were prepared (10X Chromium single-cell transcriptome kit; 10X Genomics) following the manufacturer's instructions. Both samples had a concentration of 1,200 cells/µL (total of ~10,000 cells/sample) with a target minimum of 1,000 cells/µL for recovery. The 10X Genomics platform has an ~65% mean recovery rate of cells per sample. The platform uses a droplet-based system to generate a gel bead-in-emulsion (GEM) droplet, which encapsulates the cells and reagents. The GEMs then undergo reverse transcription to generate barcoded cDNA. All cDNA generated from individual cells shares a common 10X Genomics barcode. The produced barcoded libraries are then run through standard Illumina sequencers (Suppl. Fig. 1). The libraries were run with the NextSeq 550 high-throughput sequencer, which yielded 20,000 read pairs per cell via paired-end single index sequencing in a 28+91 bp configuration, to obtain 400 million reads. The read depth was based on recommendations by 10X Genomics.

Cell Ranger pipeline

Briefly, the Cell Ranger pipeline (v.3.0.0; 10X Genomics) analyzes the NextSeq 550 high-throughput sequencer (in our experiment) data produced from the Chromium single-cell gene expression kit and processes the data generated using Feature Barcode technology. To perform the analysis, we first created a Cell Ranger genome index from the annotated canine genome (CanFam3.1).[22] Raw sequencing reads were demultiplexed by Cell Ranger based on a sample spreadsheet in comma-separated values (CSV) format. The Cell Ranger Count tool then filters and aligns the reads from the file containing the reads for each sample and counts the unique molecular identifiers (UMIs) of the aligned reads to quantify the expression of each gene in each cell. Alignment is performed using STAR software (Spliced

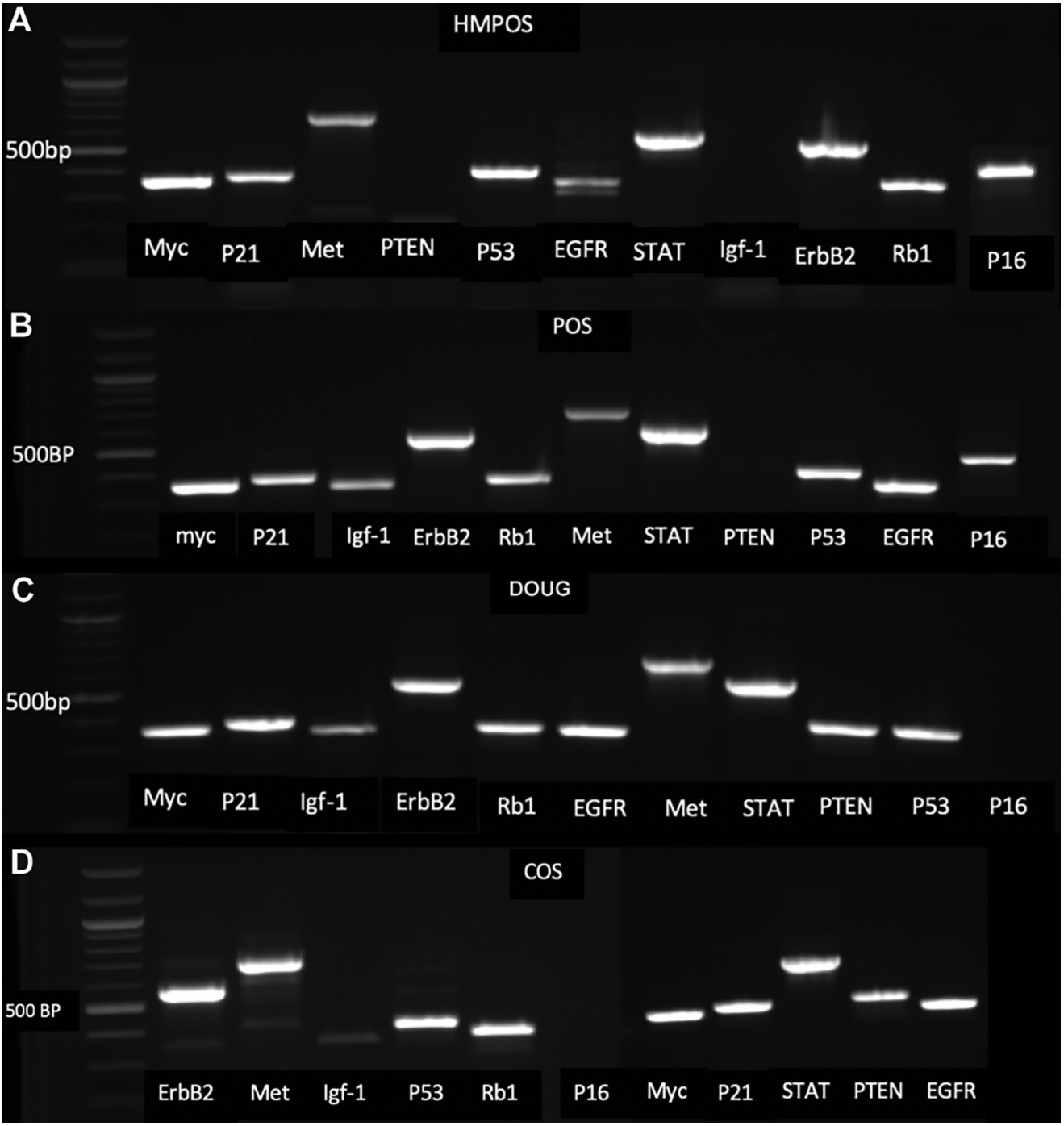

Figure 1. RT-PCR agarose gels from 4 canine osteosarcoma cell lines showing cDNA expression of *MYC*, *ERBB2*, *IGF1*, *P21*, *RB1*, *MET*, *STAT3*, *PTEN*, *P53*, *EGFR*, and *P16* genes. Primers were selected to cover 350–700 bp of the cDNA sequence of the selected gene (Suppl. Table 1). **A.** HMPOS expressed 9 of 11 genes but did not express *PTEN* or *IGF1*. **B.** POS expressed 10 of 11 genes but did not express *PTEN*. **C.** DOUG expressed 10 of 11 genes but did not express *P16*. **D.** COS31 expressed 10 of 11 genes but did not express *P16.* Both DOUG and COS31 cell lines showed lower *IGF1* band intensity relative to other genes, and a lower expression was later confirmed by scRNA-seq (Suppl. Fig. 4B).

Transcripts Alignment to a Reference).[10] STAR software aligns high-throughput long and short RNA-seq data to the canine reference genome (CanFam3.1)[22] using uncompressed suffix arrays. The software can detect canonical junctions, non-canonical splices, and chimeric transcripts, and map full-length RNA sequences.[10]

By using the combination of Chromium cellular barcodes and UMIs, the Cell Ranger Count tool generates a gene-barcode

Table 1. The expression of cDNA sequences of selected genes in 4 canine osteosarcoma cell lines (COS31, DOUG, POS, and HMPOS).

Gene (cDNA)	COS31	DOUG	POS	HMPOS
c*PTEN*	Expressed, 100% sequence match to canine transcriptome	Expressed, 100% sequence match to canine transcriptome	Not expressed	Not expressed
c*P53*	Expressed, 100% sequence match to canine transcriptome	Expressed, 100% sequence match to canine transcriptome	Expressed, 100% sequence match to canine transcriptome	Expressed, 100% sequence match to canine transcriptome
c*P16*	Not expressed	Not expressed	Expressed, 100% sequence match to canine transcriptome	Expressed, 100% sequence match to canine transcriptome
c*MYC*	Expressed, 100% sequence match to canine transcriptome	Expressed, 100% sequence match to canine transcriptome	Expressed, 100% sequence match to canine transcriptome	Expressed, 100% sequence match to canine transcriptome
c*P21*	Expressed, 100% sequence match to canine transcriptome	Expressed, 100% sequence match to canine transcriptome	Expressed, 100% sequence match to canine transcriptome	Expressed, 100% sequence match to canine transcriptome
c*IGF1*	Expressed, 100% sequence match to canine transcriptome	Expressed, 100% sequence match to canine transcriptome	Expressed, 100% sequence match to canine transcriptome	Not expressed
c*ERBB2*	Expressed, 100% sequence match to canine transcriptome	Expressed, 100% sequence match to canine transcriptome	Expressed, 100% sequence match to canine transcriptome	Expressed, 100% sequence match to canine transcriptome
c*RB1*	Expressed, 100% sequence match to canine transcriptome	Expressed, 100% sequence match to canine transcriptome	Expressed, 100% sequence match to canine transcriptome	Expressed, 100% sequence match to canine transcriptome
c*EGFR*	Expressed, 100% sequence match to canine transcriptome	Expressed, 100% sequence match to canine transcriptome	Expressed, 100% sequence match to canine transcriptome	Expressed, 100% sequence match to canine transcriptome
c*MET*	Expressed, 100% sequence match to canine transcriptome	Expressed, 100% sequence match to canine transcriptome	Expressed, 100% sequence match to canine transcriptome	Expressed, 99% sequence match to canine transcriptome, single base substitution of A to G @base 3466, no change in predicted protein
c*STAT3*	Expressed, 100% sequence match to canine transcriptome	Expressed, 100% sequence match to canine transcriptome	Expressed, 100% sequence match to canine transcriptome	Expressed, 100% sequence match to canine transcriptome

The expression of the genes was determined by the presence of gene-specific cDNAs (expected base pair number) on the RT-PCR gel and by comparing the sequence following RT-PCR to the canine genome (CanFam3.1) on BLAST software. COS31 and DOUG did not express *P16*; POS and HMPOS did not express *PTEN*; HMPOS did not express *IGF1*.

matrix that is then analyzed to identify clusters of cells showing similar gene expression patterns. The result is a collection of files containing quality control data, gene expression values, and clustering results, which can then be displayed through the Loupe Cell Browser (a free tool provided by 10X Genomics).

Loupe Browser

Briefly, the Loupe Browser (v.5.0.0; 10X Genomics) is a desktop application for Windows and Macintosh operating systems that allows easy visualization and analysis of Chromium single-cell 5′ and 3′ gene expression data. The software is used to identify significant genes and cell types, and to explore substructure within cell clusters. The Loupe Browser viewer uses single points that represent cell barcodes. To visualize the data in the Loupe Browser 2D space, Cell Ranger passes PCA-reduced (Principal Components Analysis) data into a t-SNE (t-Stochastic Neighbor Embedding) plot, which uses a nonlinear dimensionality reduction method with modifications by 10X.[43] Cell Ranger determines the vectors that are the most significant, which the Loupe Browser then displays as 2D t-SNE scatter plots (Fig. 2).

The Loupe Browser offers a "Categories" mode allowing the user to label subpopulations of cells in the clustering plot with a specific category. This view offers 2 options: "Significant Feature Comparison" and "Feature Type". The Significant Feature Comparison option works to distinguish a cluster from every other cluster in a dataset (globally distinguish) and to find what distinguishes that cluster from other clusters within its category (locally distinguish; Fig. 2). In the Feature Type expression mode, a specific feature or

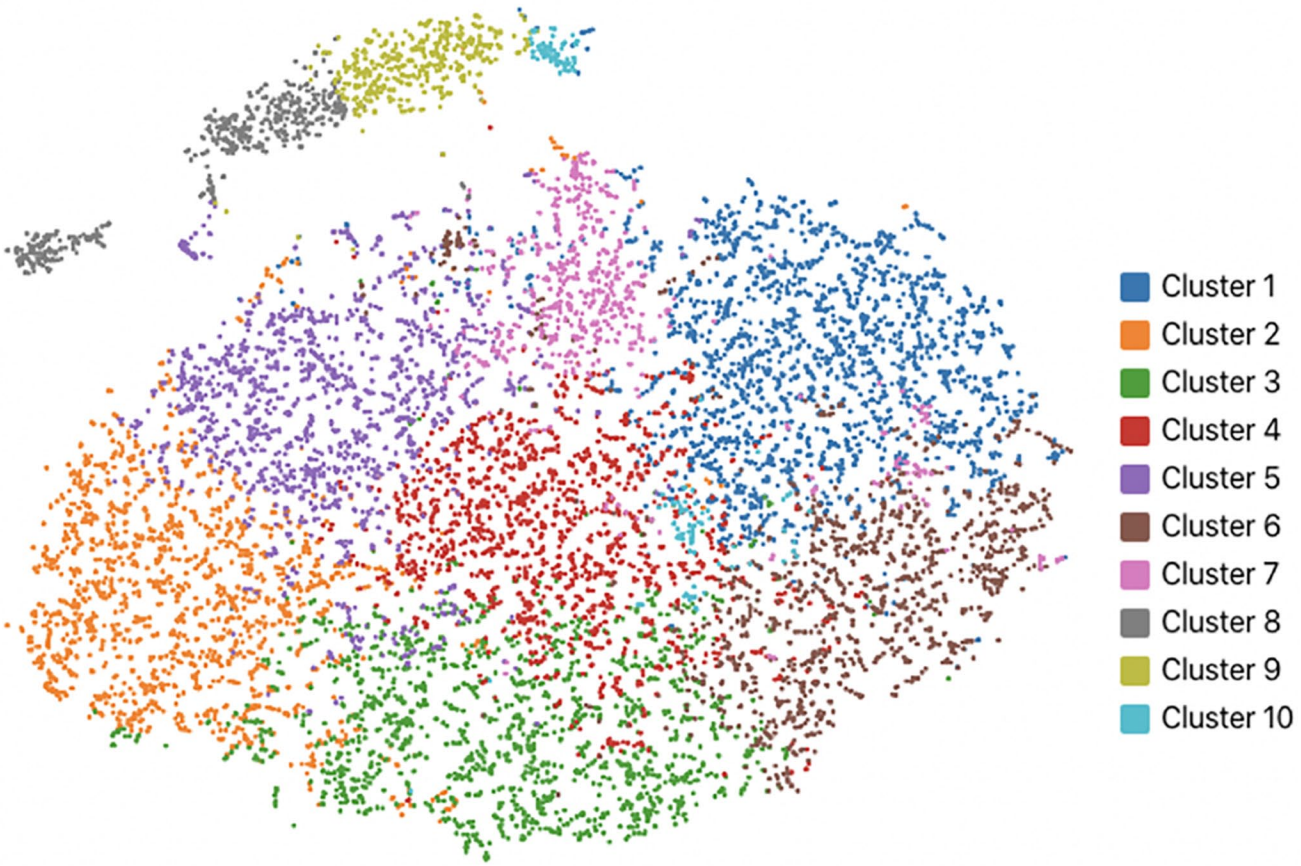

Figure 2. Loupe Browser output displayed as a t-SNE plot for the COS31 canine osteosarcoma cell line, which is based on globally distinguishing gene clusters from 10,319 barcoded cells. Globally distinguishing genes are those genes that are highly expressed within a cluster relative to the entire dataset. Ten clusters were identified: cluster 1 (1,929 cells); 2 (1,527 cells); 3 (1,454 cells); 4 (1,378 cells); 5 (1,355 cells); 6 (1,117 cells); 7 (576 cells); 8 (459 cells); 9 (387 cells); 10 (137 cells) (see online version for colors).

gene can be represented as a heat map of expression values across all represented cell clusters (Fig. 3).

Statistical analysis of single-cell data

Differential gene expression between the cell groups was determined with a negative binomial exact test (sSeq method).[7] Clusters were run through the algorithm and compared against all other cells and within clusters to yield genes that were differentially expressed in the cluster relative to the rest of the sample. Cell Ranger performed batch effect corrections, which allowed the samples to be run in separate wells and combine the datasets from these runs. The algorithm identified similar cell subpopulations between batches and merged them. The algorithm is based on the "Mutual Nearest neighbor" method (i.e., a pair of cells from 2 different batches that contain each other's nearest neighbor); differences between the cells provided an estimate of the batch effect.

The identified cell clusters were then compared to each other to determine the set of marker genes for each cluster (i.e., genes that were highly expressed in a cluster with respect to all other clusters). The output from this process is a list of marker genes for each cluster, with an associated fold change (in $\log_2$ scale) and p value (with Benjamini–Hochberg correction for multiple testing). These statistical values were then exported for further analysis with external software.

Pathway analysis using Ingenuity Pathway Analysis

The Loupe Browser outputs were then formatted for upload into Ingenuity Pathway Analysis (IPA; Qiagen).[28] In the output data, some genes were listed as loci instead of the formal gene name. All loci-listed genes that had a corresponding genetic identifier were manually altered in the gene identifier column. The fold change and p value columns from the Loupe Browser output data were used for analysis. The core analysis selected was the expression analysis based on the fold expression changes for base genes, only with consideration for direct or indirect relationships. Pathway activity

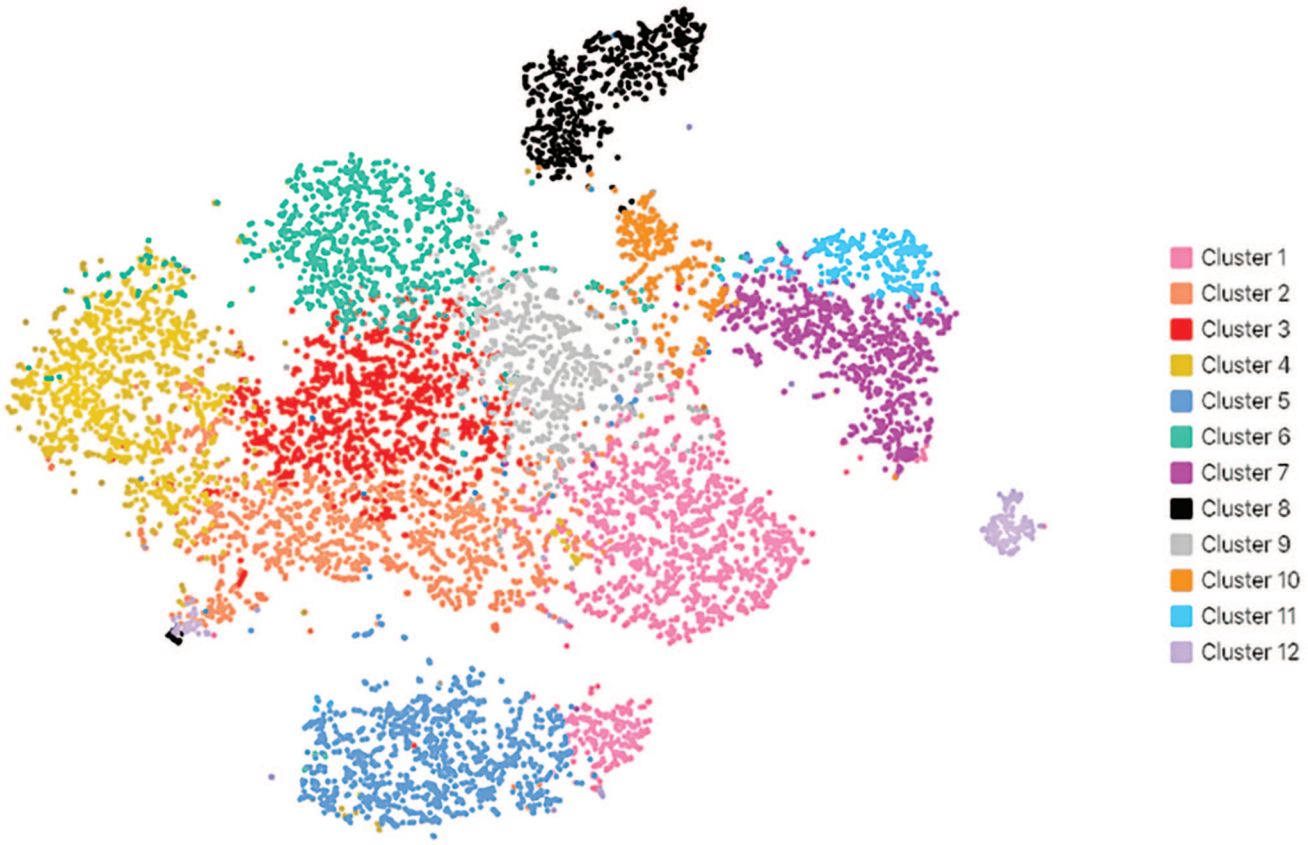

Figure 3. Loupe Browser output displayed as a t-SNE plot for the DOUG canine osteosarcoma cell line which is based on globally distinguishing gene clusters from 9,791 barcoded cells. Globally distinguishing genes are those genes that are highly expressed within a cluster relative to the entire dataset. Twelve clusters were identified: cluster 1 (1,470 cells); 2 (1,202 cells); 3 (1,142 cells); 4 (1,078 cells); 5 (1,018 cells); 6 (977 cells); 7 (788 cells); 8 (691 cells); 9 (648 cells) 10 (294 cells); 11 (281 cells); 12 (202 cells) (see online version for colors).

is indicated by the z score (≥2 = activated; ≤2 = inhibited) and is computed on directional expression. Pathways are organized based on the *p* value of a right-tailed Fisher exact test ($p \leq 0.05$), and the color depth of the pathway reflects the z score. Upstream pathways were also predicted on statistical significance ($p \leq 0.05$) for the clusters.

Results

RT-PCR and Sanger sequencing

RT-PCR and Sanger sequencing revealed differential cDNA expression of the 11 genes analyzed (Table 1, Fig. 1A–D). Three extractions from each cell line were performed and run in triplicate, all with the same results for cDNA expression. The COS31 and DOUG cell lines expressed 10 of 11 genes, with *P16* not being detectable in either cell line. POS expressed 10 of 11 genes but did not express *PTEN*. HMPOS expressed 9 of 11 genes but did not express *PTEN* or *IGF1*. Using validated methods to identify the cDNA for *PTEN*, *P53*, *P16*, *MYC*, *P21*, *IGF1*, *ERBB2*, *RB1*, *EGFR*, *MET*, and *STAT3* in the 4 cOSA cell lines, which was cross-referenced against the canine genome, a cDNA expression profile was created to compare with the scRNA-seq finding.

Uploading canine genome and generation of Cell Ranger outputs

Canine genome CanFam3.1 was chosen because of the need for an annotated genome index for Cell Ranger analysis. The file format of the genome was not consistent among the annotated genes within CanFam3.1 and required conversion of genes from GFF (general feature format) into GTF (gene transfer format). Some genes required manual conversion to the correct gene name to be recognized by the software. Once converted, the data were run against the CanFam3.1 genome.

The Cell Ranger output read 10,319 cells for COS31 and 9,791 cells for DOUG, with 167,518,632 and 125,779,629 mean reads per cell, respectively (Suppl. Figs. 2, 3). In COS31, 92.8% of reads were mapped to the canine genome (85.1% with $p < 0.05$ confidence), and in DOUG, 92.2% were mapped to the genome (85.0% with $p < 0.05$ confidence). The number of reads per cell performed depended on

the concentration of cells per sample and recovery rate, which were similar for both COS31 and DOUG cell lines. A median of 2,797 genes was identified per cell in COS31, and a median of 2,780 genes was identified in DOUG. Overall, 23,564 genes were detected in COS31 and 22,241 in DOUG. The Q30 (the fraction of bases with a Q score of at least 30 for a cell barcode/UMI count) of the reads greatly exceeded the minimum of 70–80% for a clean run in both samples, with a Q30 of 98.1% in COS31 and 98.3% in DOUG (Q30 Bases in Barcode). The saturation of repeated reads was low for both COS31 and DOUG at 8.9% and 6.3%, respectively. Cell Ranger clustering analysis identified 10 clusters for COS31 and 12 for DOUG (Figs. 2, 3). Each cluster represented a group of cells within the cOSA samples, which shared genetic similarities in the over- or under-expression of the identifiable genes.

Loupe Browser output data

Ten clusters from 10,319 barcoded cells were identified from the COS31 cell line and displayed in the t-SNE plot format (Fig. 2, Suppl. Fig. 4). The cluster with the highest number of barcoded cells was cluster 1 with 18.7% of cells, followed by cluster 2 with 14.8% of cells, cluster 3 with 14.1% cells, cluster 4 with 13.4% of cells, cluster 5 with 13.1% of cells, cluster 6 with 10.8% of cells, cluster 7 with 5.6% of cells, cluster 8 with 4.4% of cells, cluster 9 with 3.8% of cells, and cluster 10 with 1.3% of cells. Similarly, the SNE plot for DOUG (Fig. 3, Suppl. Fig. 4) identified 12 clusters from 9,791 barcoded cells. Cluster 1 had 15.0% of barcoded cells, followed by cluster 2 with 12.3% of cells, cluster 3 with 11.7% of cells, cluster 4 with 11.0% of cells, cluster 5 with 10.4% of cells, cluster 6 with 10.0% of cells, cluster 7 with 8.0% of cells, cluster 8 with 7.1% of cells, cluster 9 with 4.0% of cells, cluster 10 with 3.0% of cells, cluster 11 with 2.9% of cells, and cluster 12 with 2.1% of cells.

A hypothesis for the differences in the number and separation distance of clusters (Figs. 2, 3) between cell lines could be that COS31 is a cell line that has been used extensively in in-vitro experiments, whereas DOUG is unique to UFL and thus has a lower passage rate. We estimate the passage rate for DOUG to be 30–40 times from the first plating. In addition, the DOUG cell line has a greater distance between clusters compared to COS31, which may be as a result of DOUG having had fewer passages than COS31. For example, the COS31 cells (Fig. 2) have 3 clusters (8–10) that are distinctly separate from the larger clusters, whereas the DOUG cells (Fig. 3) have clusters, 5, 7, 8, 10–12 that are separated from the larger clusters. Furthermore, these clusters occur in smaller populations of (barcoded) cells, which confirms the ability of the 10X Genomics platform to identify smaller unique cell populations. The t-SNE plots combined with the statistical data confirm the transcriptional heterogeneity of the cells in the COS31 and DOUG OSA cell lines.

The 10 globally distinguishing genes expressed in each cluster from COS31 and DOUG (Tables 2, 3) represent the highest cluster average and the $\log_2$-fold change relative to all other clusters based on the nearest-neighbor algorithm. The genes are ranked in order of cluster average and log-fold changes. However, these data (Tables 2, 3) should be viewed with caution given that the importance of individual genes cannot be fully discerned because the expression pattern is related to the cell line itself and not to a normal cell baseline expression. Single-cell transcriptomic analysis is a group expression pattern comparison and is limited in its ability to discern the importance of individual genes given that it cannot detect point mutations nor reflect mutational status. The over- or under-expression of different genes relative to other clusters from the same cell line is their discriminating character; however, the relevance of the genes to neoplastic progression cannot be widely discerned outside of pathway analysis because some genes are likely performing housekeeping roles or are differentially expressed relative to other clusters but lacking in overall significance to pathogenesis. Therefore, to determine if genes were relevant to cellular pathways and if their differential expression were important to these pathways, we ran pathway analyses with IPA.

Cluster analysis with Ingenuity Pathway Analysis

The IPA software was successfully applied to the Cell Ranger outputs once all loci were labeled correctly. IPA has specifications for gene nomenclature, and the canine genome required significant manual reformatting to correct the nomenclature before the genes could be identified accurately. The COS31 (Suppl. Fig. 5) cell line had 89 canonical pathways identified; the DOUG (Suppl. Fig. 6) cell line had 124 canonical pathways identified, which were either up- or down-regulated within the clusters. Although similar pathways were shared among the 2 cell lines, especially the most significant pathways, there remained significant heterogeneity between pathway expression patterns. For example, some pathways represented in COS31 were either absent or under-expressed in DOUG, and vice versa (Suppl. Figs. 5, 6). Upstream regulator analyses yielded similar results. Within the clusters of both cell lines, the pathways expressed were overall consistent—most of the clusters expressed the same important pathways. However, there was considerable heterogeneity between clusters as pathways were either up-regulated or down-regulated (Suppl. Figs. 5, 6). The variability of these expression patterns reflected the "globally distinguishing" features of each cluster relative to the other clusters within the same tumor and may reflect populations of cells with variable features of malignancy.

The top 5 canonical pathways expressed in the COS31 cells were EIF2 signaling, oxidative phosphorylation, glycolysis I, CDC 42 signaling, and tRNA charging (Suppl. Fig. 5).

Table 2. The top 10 globally distinguishing genes expressed in COS31, representing the genes that have the highest cluster average and log-fold change relative to all other clusters based on the nearest-neighbor algorithm. The genes are in order of cluster average and log-fold changes.

Cluster 1 genes (n = 1929)	*ASPM*	*CENPF*	*UBE2C*	*TOP2A*	*PLK1*	*CCNB1*	*AURKA*	*KIF20A*	*CENPA*	*AURKB*
Cluster avg.	1.90	5.27	8.43	5.69	1.41	2.49	1.29	2.32	1.26	2.49
Log_2-fold change	2.19	2.08	1.98	1.97	1.93	1.80	1.69	1.66	1.65	1.64
Cluster 2 genes (n = 1527)	*OGN*	*C10H2orf40*	*RGCC*	*FGFR2*	*CTGF*	*COL8A1*	*OCIAD2*	*HPCAL1*	*TMEM59*	*DSTN*
Cluster avg.	1.03	1.21	1.43	0.98	4.99	1.04	1.39	2.41	1.60	3.66
Log_2-fold change	1.17	1.17	0.98	0.96	0.80	0.73	0.70	0.69	0.69	0.69
Cluster 3 genes (n = 1454)	*DKK 1.00*	*PTHLH*	*ANKRD1*	*F3*	*INHBA*	*SERPINE1*	*CTGF*	*MCM3*	*BAMBI*	*MYH10*
Cluster avg.	5.06	1.13	2.57	2.16	1.33	9.11	5.59	1.06	1.12	1.86
Log_2-fold change	1.79	1.48	1.43	1.40	1.29	1.28	1.07	0.94	0.85	0.77
Cluster 4 genes (n = 1378)	*PCNA*	*MCM3*	*RRM2*	*PCLAF*	*RFC3*	*DCTPP1*	*SIVA1*	*HMGA1*	*TYMS*	*NMRAL1*
Cluster avg.	2.74	1.07	2.07	1.60	1.27	1.34	2.14	2.62	2.15	3.80
Log_2-fold change	0.98	0.96	0.89	0.86	0.66	0.62	0.58	0.50	0.48	0.45
Cluster 5 genes (n = 1355)	*CDC20*	*FAM46A*	*PTTG1*	*PSMC5*	*CITED2*	*HSP90AA1*	*PRDX1*	*TXNL1*	*PSMC4*	*CAPZA1*
Cluster avg.	2.22	1.72	2.53	2.66	1.58	30.32	25.29	1.73	2.48	1.31
Log_2-fold change	0.75	0.59	0.56	0.43	0.43	0.43	0.41	0.41	0.40	0.39
Cluster 6 genes (n = 1117)	*CDK1*	*CCNA2*	*NDC80*	*SPDL1*	*TUBB4B*	*TOP2A*	*UBE2C*	*CENPE*	*TUBB*	*AURKB*
Cluster avg.	1.66	1.04	0.91	1.06	4.09	3.50	5.14	1.16	4.00	1.67
Log_2-fold change	0.74	0.73	0.69	0.67	0.66	0.65	0.64	0.62	0.61	0.61
Cluster 7 genes (n = 576)	*HSP70*	*ZFANDA*	*CCNB1*	*HSPH1*	*TUBB4B*	*GAS1*	*CDC20*	*DNAJA1*	*PLK1*	*HSPA8*
Cluster avg.	2.63	1.35	2.44	1.73	5.62	1.47	2.62	1.61	1.10	8.75
Log_2-fold change	2.85	1.34	1.24	1.17	1.10	1.00	0.96	0.95	0.94	0.89
Cluster 8 genes (n = 459)	*SLC5A3*	*AHNAK*	*ITGB3*	*SLC7A11*	*TIA1*	*RNF217*	*IGF2R*	*COL12A1*	*MYSM1*	*ADAMTS1*
Cluster avg.	1.05	1.63	1.03	1.01	1.35	1.05	1.80	6.08	1.47	1.01
Log_2-fold change	1.87	1.74	1.73	1.66	1.65	1.64	1.63	1.63	1.61	1.59
Cluster 9 genes (n = 387)	*NRIP1*	*STK17A*	*AHNAK*	*FOXP1*	*RPS28*	*ZNF250*	*UQCRQ*	*MET*	*SCD*	*SLC5A3*
Cluster avg.	1.36	1.41	1.17	1.40	10.91	1.25	1.68	2.81	2.78	0.58
Log_2-fold change	1.61	1.34	1.29	1.12	1.12	1.05	1.04	1.03	1.03	1.01
Cluster 10 genes (n = 137)	*RPS28*	*COPS9*	*NXN*	*ATP5E*	*RPS27*	*CENPW*	*CCDC34*	*EPAS1*	*MRPL36*	*RPS21*
Cluster avg.	8.20	2.75	1.06	1.73	5.63	1.28	0.88	0.98	1.17	2.56
Log_2-fold change	0.71	0.69	0.67	0.63	0.61	0.59	0.58	0.58	0.58	0.57

n = number of barcoded cells. Negative log_2-fold change indicates down-regulation.

Table 3. The top 10 globally distinguishing genes expressed in DOUG, representing the genes that have the highest cluster average and log-fold change relative to all other clusters based on the nearest-neighbor algorithm. The genes are in order of cluster average and log-fold changes.

Cluster 1 genes (n = 1470)	*SERPINF1*	*CENPF*	*AURKB*	*TOP2A*	*CCNB1*	*UBE2C*	*CCNB2*	*PLK1*	*CDCA3*	*PRC1*
Cluster avg.	1.66	0.05	0.06	0.11	0.06	0.07	0.05	0.09	0.09	0.06
Log_2-fold change	2.45	−4.16	−4.00	−3.77	−3.72	−3.68	−3.69	−3.46	−3.41	−3.48
Cluster 2 genes (n = 1202)	*TOP2A*	*AURKA*	*CENPF*	*AURKB*	*UBE2C*	*HMGB2*	*CDK1*	*CKS2*	*MGP*	*SPON2*
Cluster avg.	0.22	0.09	0.15	0.16	0.18	0.44	0.19	1.16	2.65	0.56
Log_2-fold change	−2.73	−2.66	−2.60	−2.46	−2.27	−1.97	−1.96	−1.48	1.55	−1.37
Cluster 3 genes (n = 1142)	*CENPF*	*SPON2*	*NUPR1*	*CCNB1*	*SERPINF1*	*PCNA*	*PLK1*	*LCP1*	*TSC22D1*	*ARL6IP1*
Cluster avg.	0.23	0.38	1.58	0.27	0.20	3.51	0.37	0.32	1.43	0.62
Log_2-fold change	−2.00	−1.95	−1.62	−1.52	−1.44	1.03	−1.31	−1.28	−1.17	−1.22
Cluster 4 genes (n = 1078)	*CCNB1*	*PLK1*	*CCNB2*	*CDC20*	*TACC3*	*CENPA*	*KIF23*	*NUF2*	*KIF20B*	*CENPF*
Cluster avg.	2.11	2.30	1.42	3.40	1.17	1.87	1.18	1.72	1.41	1.77
Log_2-fold change	1.83	1.67	1.53	1.48	1.42	1.42	1.33	1.31	1.29	1.29
Cluster 5 genes (n = 1018)	*ANXA8L1*	*TAGLN*	*MAL*	*CTGF*	*GAP43*	*SPON2*	*CHCHD10*	*DCN*	*EBPL*	*TFPI2*
Cluster avg.	0.51	6.79	0.16	1.06	0.42	2.49	3.43	0.58	0.32	3.91
Log_2-fold change	−2.41	1.48	−1.71	1.10	−1.39	1.04	1.02	0.92	−1.20	0.89
Cluster 6 genes (n = 977)	*TOP2A*	*CENPF*	*CDK1*	*AURKB*	*UBE2C*	*ASPM*	*AURKA*	*PRC1*	*CCNB1*	*PLK1*
Cluster avg.	4.33	2.33	1.88	2.09	2.12	1.67	1.23	1.36	1.52	1.80
Log_2-fold change	2.31	1.98	1.83	1.82	1.79	1.76	1.66	1.46	1.34	1.33
Cluster 7 genes (n = 788)	*RASL11A*	*C10H2orf40*	*MGP*	*GPNMB*	*OCIAD2*	*TUBB3*	*CENPF*	*IGFBP2*	*ASPM*	*CLDN1*
Cluster avg.	3.37	0.21	0.14	0.54	0.28	0.21	0.29	3.48	0.25	0.97
Log_2-fold change	1.53	−2.02	−3.10	−1.74	−1.67	−1.65	−1.62	−1.44	−1.47	1.08
Cluster 8 genes (n = 691)	*GLTP*	*TXNL4A*	*TMX3*	*SERPINB2*	*LMAN1*	*SEC11C*	*ATP8B1*	*TXNL1*	*TCF4*	*MBD2*
Cluster avg.	1.29	2.11	0.53	0.34	1.46	1.82	0.86	0.96	0.53	1.15
Log_2-fold change	2.33	0.52	0.05	−1.13	−0.26	0.56	0.67	−0.12	−0.48	0.64
Cluster 9 genes (n = 294)	*CCNB1*	*PCNA*	*TXNL4A*	*TMX3*	*SERPINB2*	*LMAN1*	*SEC11C*	*ATP8B1*	*TXNL1*	*TCF4*
Cluster avg.	0.26	3.65	1.30	0.56	0.60	1.73	1.26	0.69	0.91	0.82
Log_2-fold change	−1.56	1.02	−0.21	0.16	−0.32	−0.02	0.02	0.36	−0.22	0.14
Cluster 10 genes (n = 281)	*TXNL4A*	*TMX3*	*SERPINB2*	*LMAN1*	*SEC11C*	*ATP8B1*	*TXNL1*	*TCF4*	*MBD2*	*CTGF*
Cluster avg.	1.84	0.40	0.37	1.16	1.65	0.58	1.06	0.62	2.84	0.84
Log_2-fold change	0.33	−0.36	−1.02	−0.61	0.44	0.10	0.02	−0.28	0.25	0.18
Cluster 11 genes (n = 281)	*UBE2C*	*AURKA*	*PLK1*	*AURKB*	*CCNB1*	*HMGB2*	*TK1*	*H1F0*	*CENPF*	*KIF23*
Cluster avg.	3.06	1.65	2.61	2.37	2.12	4.64	1.55	1.43	2.13	1.33
Log_2-fold change	2.11	1.88	1.74	1.72	1.69	1.66	1.58	1.51	1.49	1.42
Cluster 12 genes (n = 202)	*TMX3*	*SERPINB2*	*LMAN1*	*TCF4*	*CTGF*	*BCLAF1*	*GINM1*	*PCMT1*	*EZR*	*ACAT2*
Cluster avg.	1.24	2.68	4.06	5.76	0.66	1.46	1.03	1.41	2.49	4.70
Log_2-fold change	1.28	1.85	0.66	1.71	−0.16	1.39	0.27	1.10	1.69	0.14

n = number of barcoded cells. Negative log_2-fold change indicates down-regulation.

Within the clusters, these pathways were variably expressed. For example, within the EIF2 signaling pathway (based on z scores), clusters 1, 3, 6–8 (5 of 10) show down-regulation, whereas clusters 2, 4, 5, 9, and 10 (5 of 10) showed up-regulation ($p < 0.001$). The cluster with the most robust down-regulation of the EIF2 pathway was cluster 1 (the largest cluster, with 1,929 cells), and the cluster with the most up-regulation of the EIF2 pathway was cluster 10 (the smallest cluster, with 137 cells). Interestingly, on the COS31 t-SNE plot, cluster 10 was 1 of 3 clusters separated from the main cluster grouping (Fig. 2). The most significant disease biofunctional pathway predicted for COS31 was cell death ($p < 0.001$). The major upstream regulatory pathways identified were the *RICTOR*, *MYC/MYCN*, *NFE2L2*, *RABL6* (*RAS* oncogene family–like 6), and *TGFB1.*

In the DOUG cells (Suppl. Fig. 6), the top 5 canonical pathways expressed were (in order of importance) oxidative phosphorylation, cholesterol biosynthesis (I–III), EIF2 signaling, tRNA charging, and death receptor signaling. When comparing the clusters, variation in expression was significant within all 12 clusters. For example, within oxidative phosphorylation, clusters 2, 3, 9, and 10 (4 of 12) were up-regulated, whereas clusters 1, 4, 5, 8, 11, and 12 (6 of 12) were down-regulated. Clusters 6 and 7 (2 of 12) clusters showed no change in baseline expression of the pathway ($p < 0.001$). Unlike COS31, where cell death pathway was predicted, in DOUG, up-regulation of pathways associated with cell proliferation and cell viability was expressed in a higher number of clusters. Nevertheless, cell death in DOUG cells was still expressed in all clusters but was only up-regulated in 6 of 12 clusters. This could indicate that the DOUG cells were likely less stressed and more replicative at the time of extraction relative to COS31. The 5 major upstream regulators for DOUG were *MYC, RICTOR, ATF4, HSF2*, and *CLPP* ($p < 0.001$).

Comparing RT-PCR and Sanger sequencing results with 10X Genomics outputs for DOUG and COS31

We compared the RT-PCR and Sanger sequencing data for COS31 and DOUG to the scRNA-seq data from the 10X Genomics platform; real-time PCR (rtPCR) was not performed to quantify the expression profiles of the genes of interest. The genes reported in the RT-PCR bulk analysis (*PTEN*, *P53*, *P16*, *MYC*, *P21*, *IGF1*, *ERBB2*, *RB1*, *EGFR*, *MET*, and *STAT3*) were examined for distribution in the scRNA-seq clusters (Table 1, Fig. 4, Suppl. Fig. 4). Similar to the RT-PCR cDNA findings, transcription of *P16* (*CDKN2A*) in scRNA-seq clusters was absent in both cell lines (Fig. 4). The remaining genes were distributed diffusely throughout the clusters, albeit at different expression levels (Suppl. Fig. 4). Thus, the RT-PCR data from COS31 and DOUG genes of interest were in concordance with the presence or absence of gene expression from the scRNA-seq data.

Discussion

In the first aim of our study, bulk sequencing of the cOSA cell lines (COS31, POS, HMPOS, and Doug) using RT-PCR and Sanger sequencing identified differences in the presence or absence of key genes (*PTEN*, *P53*, *P16*, *MYC*, *P21*, *IGF1*, *ERBB2*, *RB1*, *EGFR*, *MET*, and *STAT3*). COS31 and DOUG both lacked *P16* expression, POS and HMPOS lacked *PTEN*, and HMPOS lacked *IGF1* expression. Although we report on the cDNA expression patterns of these genes in the 4 cOSA cell lines, we did not sequence the entire genes to detect point mutations, nor did we quantify the RNA expression using rtPCR because this was not our goal at the time. The profile of genes generated allowed for further validation of our novel application of scRNA-seq to the canine genome when comparing the presence or absence of key OSA genes. The RT-PCR expression patterns in COS31and DOUG matched the outputs from the scRNA-seq run with *P16* (*CDKN2A*), being absent in both cell lines. Interestingly, the remaining 10 genes were distributed diffusely throughout the scRNA-seq clusters, albeit at different expression levels.

The characterization of COS31 and DOUG cell lines utilizing scRNA-seq provided a transcriptomic expression profile not achievable previously with bulk analysis. The scRNA-seq data also allowed us to perform detailed canonical pathway analyses, which identified significant transcriptional tumor heterogeneity in both cOSA cell lines. NGS using scRNA-seq is a powerful tool and, to our knowledge, has not been attempted previously with the canine genome, nor has it been applied to OSA samples of any species. This required significant development work by the authors to upload the canine genome (CanFam3.1) into the 10X Genomics Cell Ranger pipeline. The output from the pipeline identified transcriptomes from 20,110 individual cOSA cells from both cell lines, which resulted in the identification of previously unreported pathways in cOSA. A limitation of all scRNA-seq platforms is the relatively low capture efficiency, which is ~65% for the 10X Genomics platform. The cell capture inefficiency can be overcome by increasing the number of viable cells and the purity of the cells. However, a cell number ceiling exists of 10,000 cells per sample; increasing the cell number above the ceiling leads to a higher saturation of repeated reads.[2,42,78] These limitations plus the lack of transcriptomic data were an incentive for us to start with the 4 cOSA cell lines, thereby controlling for viable cell numbers and non-cancerous cell cDNA contamination. Samples were optimized for a high live-to-dead ratio by extracting COS31 and DOUG at 75% confluence for the scRNA-seq runs. Confirmation of this approach was reflected in the quality of the samples for the scRNA-seq runs, which is supported by our quality control data.[42] Nevertheless, some evidence of

cell proliferation and cell death pathway differences existed between cell lines, which requires further investigation.

The scRNA-seq identified individual clusters within each cell line with unique expression profiles, demonstrating that COS31 and DOUG OSA cell lines have heterogeneous cell populations with pathways differentially up- and down-regulated based on cluster. The Cell Ranger output revealed 10 cellular clusters in COS31 and 12 cellular clusters in DOUG, as visualized in the 2D t-SNE plots generated by the Loupe Browser. The plots and dataset of the highest cluster average and $\log_2$-fold change reflected the heterogeneous nature of the 2 cell lines. Although many more genes were transcribed, we selected the 10 globally distinguishing genes expressed in clusters from COS31 and DOUG to illustrate the heterogeneity between clusters. Also, DOUG, a recently derived cell line, had more clusters and greater separation between clusters when comparing the spatial nature of cluster patterns of COS31 with DOUG. Although speculative at this time, the differences between the cell lines may be a result of their age; the passage process may apply a selection bias to older cell lines.[72] Our data also served to illustrate the ability of scRNA-seq to identify clusters with low cell numbers in comparison to "dominant clusters"; further investigation of these clusters may result in genotypes associated with "stemness" or tumorigenicity.[50] The t-SNE plots plus the statistical data confirm the transcriptional heterogeneity of the cells in the COS31 and DOUG OSA cell lines.

To further explore canonical gene pathways in the 2 cOSA cell lines, we performed IPA analysis on the Cell Ranger outputs. A constraint of the IPA software was its limit to human and murine species. Thus, any unidentified loci or genes not present in either human or murine species were excluded. The loci excluded from our study were only those that were unidentified in the reference canine genome. Importantly, genes relevant to our study were available for analysis but required manual reformatting into the correct nomenclature to be identified by the software. Clusters largely expressed similar pathways but were highly heterogeneous in expression and were either up- or down-regulated (Tables 4–7, Suppl. Figs. 5, 6). Furthermore, when comparing clusters within cell lines, no single cluster had the same combination of pathway expression. This was also the case when comparing COS31 with DOUG. Of the 81 distinct and significant pathways identified in both COS31 and DOUG, 33 were identified as important to OSA based on our literature review.[14,15,20,21,31,41,44,47,61,66] The remaining pathways were either related to cellular maintenance (reference genes), inflammation, or had not been investigated previously in OSA.

Although the heterogeneity of the pathway expression profiles in the cell lines was complex and requires further investigation, we can comment on the possible significance of the 33 pathways in OSA. Of the 33 pathways important to OSA pathogenesis, 18 have been investigated previously in the cOSA patient (Tables 4, 5). In hOSA and cOSA, a poor prognosis has been associated with alterations in the mitochondrial oxidative phosphorylation, protein kinase A (PKA), integrin pathways, VEGF, polo-like kinase (PLK), and IGF1 pathways.[9,27,33,38,44,45,65,73,75] The CXCR4, aryl hydrocarbon, mTOR, MAPK/ERK, and PII3-AKT pathways have also been demonstrated to play a role in hOSA and cOSA tumorigenesis, with the MAPK/ERK and PII3-AKT pathways investigated as potential therapeutic targets for hOSA and cOSA.[11,19,24,51,55,56,74,82,85,87] The Il-8 and WNT signaling pathways have been implicated in increasing the metastatic potential in canine and human OSA.[12,31,33,56,58]

The cholesterol biosynthesis pathway, which showed significant up-regulation in some clusters in the DOUG cell line, has been reported to be of importance in hOSA metastatic development.[11,23,49,74] Interestingly, a study in cOSA looked at the total cholesterol elevations in cOSA patients versus controls with limb fractures, and found a significant elevation in cholesterol in OSA-bearing patients, indicating that the cholesterol biosynthesis pathway is of probable significance in cOSA patients.[29]

The platelet-derived growth factor (PDGF) pathway was significantly represented in both COS31 and DOUG, and, in prior studies, overexpression was associated with the majority of hOSA and cOSA cases.[46,84] The targeting the PDGF pathway in multi-agent clinical trials in hOSA appears to potentiate OSA cell death—such trials have not been performed in the canine patient to date.[46,84]

Another interesting finding was the significance of the glutathione transferase pathway, which had been investigated previously in COS31, whereby increases in glutathione transferase potentiated resistance to carboplatin cytotoxicity.[69] The specific polymorphisms of the glutathione transferase pathways in hOSA have been associated with an increased risk of OSA development based on evidence from meta-analysis studies of hOSA trials.[81]

Importantly, scRNA-seq identified 15 pathways in both cell lines (Tables 6, 7) that have not been investigated in the cOSA patient or cell lines. The SAPK/JNK pathway, a major apoptotic pathway for eIF (eukaryotic initiation factor 2), and the direct EIF2 pathway mRNA regulating pathway, were significant in both DOUG and COS31 clusters, albeit with significant cluster heterogeneity. Analysis in hOSA tumors has demonstrated that eIF-2α levels are typically lower in cancer cells.[16] With treatment, the increase in phosphorylated eIF-2α was associated both with increased cancer cell death and was synergistic when combined with chemotherapeutic treatments.[83] There were also multiple pathways expressed in COS31 and DOUG that have been implicated in the promotion of metastasis in hOSA, such as the Rho GTPase pathway (encompassing Rac, Rho, and Cdc42 subfamilies), CDK5 pathway, and ephrin receptor pathway.[1,35,64] The suppression of the Rho and Cdc42 pathways has been shown to prevent metastasis and cell growth in culture.[1,35,64] The Rho pathway has also

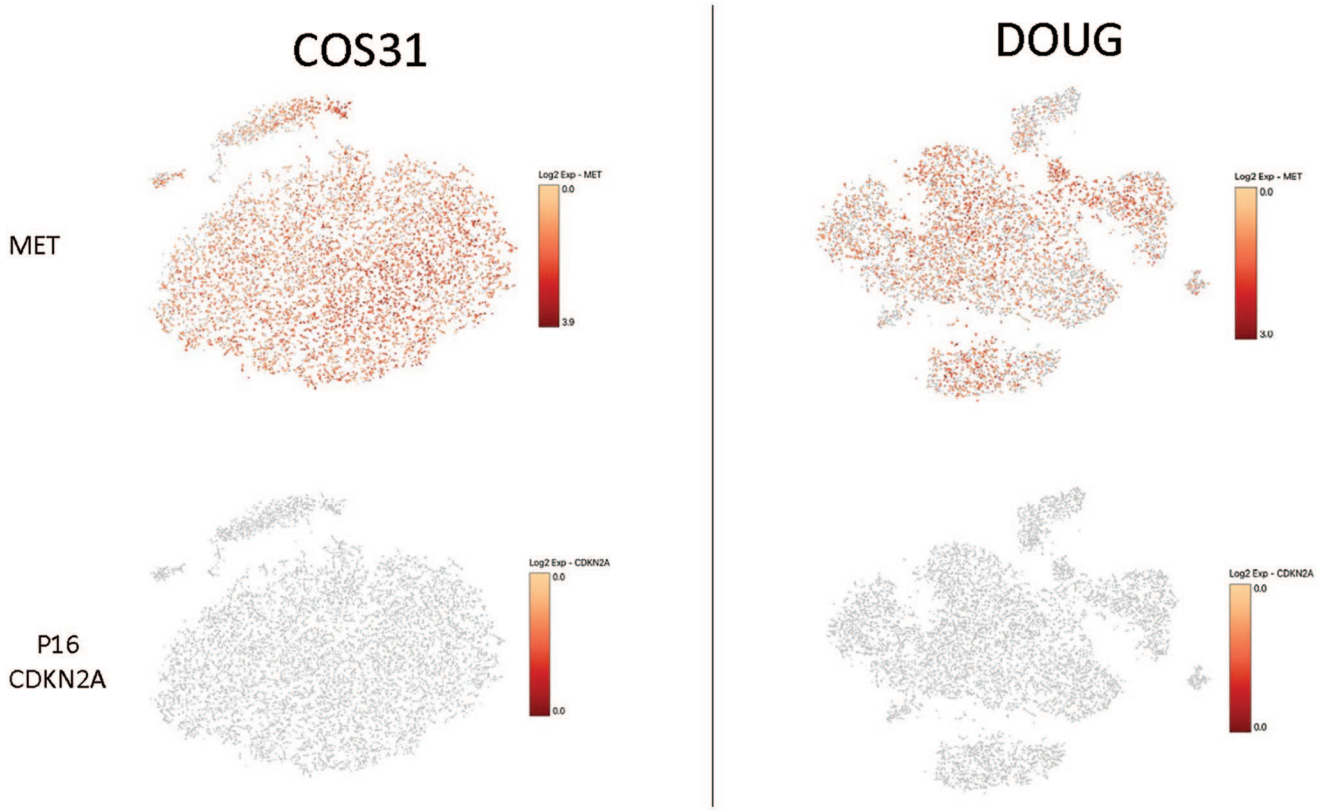

Figure 4. Loupe Browser output as a t-SNE plot for COS31 and DOUG in the active feature search browser. The $\log_2$-expression level is reflected by the color scale and the lower numerical value to the right of the scale. The numerical value reflects if a gene is present or absent; the color scale indicates the intensity of expression in barcoded cells across all clusters. Both COS31 and DOUG lacked *P16* expression (no barcoded cells and 0 $\log_2$-expression) but had intense *MET* expression (positive barcoded cells [color scale] and 0–3.9 [COS31], 0–3.0 [DOUG] $\log_2$-expression) (see online version for colors).

been implicated in alterations in the phosphoinositide phospholipase C (PLC) transduction pathway, and both are being investigated as pathways to block OSA growth.[39,40] EPHA2 (ephrin type-A receptor 2) is being tested as a potential therapeutic target in hOSA.[17] Pathways implicated in chemoresistance (such as the tRNA pathway and the Nrf2 pathway) were also found to be upregulated in multiple clusters in DOUG and COS31.[36,86] For COS31, these pathways were up-regulated in clusters 4–7 representing a population of cells that would likely have been overshadowed by the other more populous clusters. The sirtuin pathway was also significant in both cell lines, and this tumor suppressor pathway has been implicated in the tumorigenesis of hOSA in murine models.[59] Furthermore, sirtuin-1 (SIRT1) was found to impair liver kinase B1 (LKB1) protein in 41% of OSA patients and was also found to be regulated by miR-204.[59] There was also a pathway of potential epigenetic interest, the HOTAIR pathway, which was variably expressed across clusters in both DOUG and COS31. The HOTAIR (long non-coding RNA [lncRNA]) pathway was found to be significantly up-regulated in a study of 900 hOSA samples versus matched controls.[88] The silencing of the pathway in experimental settings has inhibited OSA proliferation, migration, and invasion.[88] Furthermore, methylating agents are currently being investigated to inhibit HOTAIR in hOSA.[32] The androgen receptor pathway was shown to be significant to multiple cellular clusters. The androgen receptor pathway expression has been inversely predictive of survival in hOSA.[34] Interestingly, a large 2019 cOSA study[77] found castrated males and spayed bitches had a greater frequency of OSA compared to intact dogs, indicating that sex hormones may play a role in cOSA. Finally, a specific target for aminobisphosphonate drugs, the mevalonate pathway, was also found to be significant in COS31 and DOUG, and is reported to be implicated in the progression of hOSA and cOSA.[26,52,79]

The weaknesses of our study with regards to scRNA-seq are also those inherent within any NGS study, namely quality and purity of the sample, the integrity of the method, library preparations, level of annotation of the reference genome, and computational challenges.[24] Another weakness of our study was the use of cOSA cell lines versus tissue samples from OSA cases; nonetheless, as a first step, our primary

Table 4. Pathways with previously investigated roles in both hOSA and cOSA in COS31 as expressed by each cluster of cells (10 clusters of cells total, with 10,319 cells accounted for across all clusters) within the IPA output.

	Cluster									
Pathway	1	2	3	4	5	6	7	8	9	10
Oxidative phosphorylation	DR	DR	DR	UR	UR	DR	DR	DR	DR	UR
Integrin signaling	DR	UR	UR	DR	DR	DR	NE	NE	NE	DR
Cholesterol biosynthesis	UR	UR	DR	UR	DR	DR	UR	DR	DR	DR
Actin cytoskeleton	DR	UR	UR	NE	NE	DR	UR	DR	DR	DR
CXCR4 signaling	DR	UR	UR	UR	NE	DR	DR	DR	DR	DR
IL8 signaling	DR	UR	UR	DR	DR	DR	NE	DR	NE	DR
Miotic PLK pathways	UR	DR	DR	DR	DR	UR	UR	DR	DR	DR
mTOR pathway	DR	UR	UR	DR	NE	DR	NE	DR	DR	DR
IGF1 signaling pathway	DR	DR	UR	UR	NE	UR	UR	DR	DR	DR
VEGF signaling	DR	NE	UR	DR	DR	UR	UR	NE	DR	DR
PDGF signaling	DR	DR	UR	UR	DR	UR	UR	DR	DR	DR
Glutathione pathways	NE	UR	NE	DR	DR	DR	DR	DR	DR	UR
Apoptosis signaling	UR	UR	DR	DR	NE	DR	DR	NE	DR	DR
MAPK pathway	DR	UR	UR	DR	DR	DR	DR	DR	DR	DR
PI3K/AKT pathway	DR	DR	UR	UR	UR	DR	UR	DR	DR	DR
Wnt-β-catenin pathway	UR	DR	DR	DR	UR	DR	NE	UR	DR	DR
Aryl-HC receptor	UR	NE	UR	NE	DR	DR	NE	UR	UR	NE
PKA pathway	NE	DR	DR	DR	UR	DR	UR	DR	DR	DR

DR (blue) = down-regulation of the pathway; NE (gray) = no change in baseline expression of the pathway; UR (orange) = up-regulation of the pathway (see online version for colors).

Table 5. Pathways with previously investigated roles in both hOSA and cOSA in DOUG as expressed by each cluster of cells (12 clusters of cells total, with 9,791 cells accounted for across all clusters) within the IPA output.

	Cluster											
Pathway	1	2	3	4	5	6	7	8	9	10	11	12
Oxidative phosphorylation	DR	UR	UR	DR	DR	NE	DR	DR	UR	UR	DR	DR
Integrin signaling	UR	UR	DR	DR	UR	DR	UR	UR	DR	DR	UR	DR
Cholesterol biosynthesis	DR	UR	UR	UR	DR	UR	DR	DR	DR	DR	DR	UR
Actin cytoskeleton	DR	NE	DR	DR	UR	DR	UR	UR	DR	DR	UR	DR
CXCR4 signaling	DR	UR	DR	NE	UR	DR	NE	DR	DR	DR	NE	DR
IL8 signaling	NE	UR	DR	NE	DR	DR	DR	UR	DR	DR	NE	DR
Miotic PLK pathways	DR	DR	DR	UR	UR	UR	DR	DR	NE	DR	UR	DR
mTOR pathway	NE	UR	DR	UR	UR	DR	UR	DR	DR	UR	UR	DR
IGF1 signaling pathway	DR	DR	DR	UR	DR	DR	UR	UR	DR	UR	UR	DR
VEGF signaling	DR	UR	NE	UR	UR	UR	NE	UR	DR	NE	UR	DR
PDGF signaling	NE	NE	DR	UR	UR	DR	UR	UR	DR	UR	UR	DR
Glutathione pathways	UR	UR	DR	DR	NE	DR	DR	DR	UR	NE	DR	NE
Apoptosis signaling	DR	NE	UR	UR	UR	NE	DR	DR	DR	DR	DR	UR
MAPK pathway	UR	NE	DR	NE	DR	DR	NE	UR	DR	NE	UR	DR
PI3K/AKT pathway	DR	UR	UR	UR	UR	DR	UR	UR	DR	DR	UR	DR
Wnt-β-catenin pathway	DR	UR	UR	UR	NE	UR	DR	DR	DR	DR	DR	DR
Aryl-HC receptor	DR	DR	UR	NE	NE	NE	DR	UR	UR	DR	DR	DR
PKA pathway	DR	NE	UR	UR	UR	DR	DR	NE	DR	DR	NE	DR

DR (blue) = down-regulation of the pathway; NE (gray) = no change in baseline expression of the pathway; UR (orange) = up-regulation of the pathway (see online version for colors).

Table 6. Pathways with previously investigated roles in hOSA, which have not been investigated previously in cOSA, which were detected in the IPA output for COS31.

	Cluster									
Pathway	1	2	3	4	5	6	7	8	9	10
eIF2 signaling	DR	UR	DR	UR	UR	DR	DR	DR	UR	UR
Glycolysis/gluconeogenesis	DR	UR	DR	UR	UR	DR	UR	DR	DR	DR
Rho signaling	DR	UR	UR	NE	UR	DR	NE	DR	DR	DR
tRNA charging	DR	DR	UR	DR	DR	UR	DR	DR	DR	DR
Nrf2-med. oxidative stress	DR	DR	DR	UR	UR	UR	UR	DR	DR	DR
Ephrin receptor signal	DR	UR	UR	NE	DR	DR	NE	DR	DR	DR
Mevalonate Signaling	UR	NE	DR	UR	DR	DR	UR	DR	DR	DR
Phospholipase C signaling	DR	UR	UR	NE	NE	DR	DR	DR	DR	DR
Nuclear excision repair	DR	DR	NE	UR	DR	NE	NE	NE	UR	UR
SAPK/JNK signaling	DR	UR	NE	UR	UR	DR	UR	DR	DR	DR
HOTAIR reg pathway	DR	NE	UR	DR	DR	DR	DR	UR	UR	DR
Androgen signaling	NE	DR	UR	NE	UR	DR	UR	DR	DR	DR
Calpain protease	DR	UR	UR	NE	NE	NE	NE	DR	DR	DR
CDK5 signaling	DR	UR	UR	DR	DR	DR	DR	NE	NE	DR
P2Y purinergic receptor sig	DR	DR	UR	DR	NE	NE	NE	DR	NE	DR

DR (blue) = down-regulation of the pathway; NE (gray) = no change in baseline expression of the pathway; UR (orange) = up-regulation of the pathway (see online version for colors).

Table 7. Pathways with previously investigated roles in hOSA, which had not previously been investigated in cOSA, which were detected in the IPA output for DOUG.

	Cluster											
Pathway	1	2	3	4	5	6	7	8	9	10	11	12
eIF2 signaling	DR	DR	NE	DR	NE	DR	NE	NE	NE	UR	DR	DR
Glycolysis/gluconeogenesis	UR	UR	UR	UR	NE	NE	DR	UR	DR	NE	DR	DR
Rho signaling	DR	UR	NE	DR	UR	DR	UR	NE	DR	DR	UR	DR
tRNA charging	UR	DR	DR	DR	UR	DR	UR	DR	UR	DR	UR	UR
Nrf2-med. oxidative stress	DR	NE	UR	UR	DR	UR	DR	DR	NE	DR	DR	DR
Ephrin receptor signal	DR	UR	DR	NE	UR	DR	UR	UR	DR	DR	UR	DR
Mevalonate signaling	DR	UR	UR	UR	NE	UR	NE	DR	DR	DR	NE	DR
Phospholipase C signaling	UR	UR	DR	DR	NE	DR	NE	UR	DR	NE	NE	DR
Nuclear excision repair	DR	NE	UR	NE	DR	UR	UR	UR	NE	UR	NE	DR
SAPK/JNK signaling	DR	UR	UR	UR	UR	DR	UR	NE	DR	NE	UR	DR
HOTAIR reg pathway	DR	NE	DR	UR	NE	DR	DR	UR	NE	DR	DR	UR
Androgen signaling	NE	DR	UR	UR	UR	DR	UR	DR	NE	NE	UR	DR
Calpain protease	UR	UR	DR	NE	UR	DR	UR	UR	DR	NE	UR	NE
CDK5 signaling	UR	NE	UR	UR	DR	NE	NE	DR	DR	DR	UR	DR
P2Y purinergic receptor sig	NE	DR	DR	UR	DR	DR	UR	UR	DR	UR	UR	DR

DR (blue) = down-regulation of the pathway; NE (gray) = no change in baseline expression of the pathway; UR (orange) = up-regulation of the pathway (see online version for colors).

motivation for using the OSA cell lines for RT-PCR and scRNA-seq revolved around the ability to get clean samples in vitro versus in vivo, and to exclude any normal cell DNA/RNA contamination. In addition, little was known about the transcriptome of these commonly used cOSA cell lines.

Acknowledgments

We thank the staff of the Interdisciplinary Center for Biotechnology Research (ICBR), Gene Expression & Genotyping Section, University of Florida, especially the Scientific Director, Yanping Zhang. Also, we are most grateful for the advice

received from the 10X Genomics support staff, and specifically Fred Souret and Egon Ranghini.

Declaration of conflicting interests

The authors declared no potential conflicts of interest with respect to research, authorship, and/or publication of this article

Funding

Our research was funded by a Faculty Research Award, College of Veterinary Medicine, University of Florida; the Gonzmart Family Foundation; and the Milner Comparative Oncology Laboratory.

ORCID iDs

Rowan J. Milner https://orcid.org/0000-0001-6379-2058
Galaxia Cortés-Hinojosa https://orcid.org/0000-0002-6160-9093
Sandra Bechtel https://orcid.org/0000-0002-7773-5959

Supplementary material

Supplementary material for this document is available online.

References

1. Bao H, et al. Potential mechanisms underlying CDK5 related osteosarcoma progression. Expert Opin Ther Targets 2017;21:455–460.
2. Baran-Gale J, et al. Experimental design for single-cell RNA sequencing. Brief Funct Genomics 2018;17:233–239.
3. Barroga EF, et al. Establishment and characterization of the growth and pulmonary metastasis of a highly lung metastasizing cell line from canine osteosarcoma in nude mice. J Vet Med Sci 1999;61:361–367.
4. Baslan T, Hicks J. Unravelling biology and shifting paradigms in cancer with single-cell sequencing. Nature Rev Cancer 2017;17:557–569.
5. Bergman PJ, et al. Amputation and carboplatin for treatment of dogs with osteosarcoma: 48 cases (1991 to 1993). J Vet Intern Med 1996;10:76–81.
6. Boerman I, et al. Prognostic factors in canine appendicular osteosarcoma—a meta-analysis. BMC Vet Res 2012;8:56.
7. Ching T, et al. Power analysis and sample size estimation for RNA-Seq differential expression. RNA 2014;20:1684–1696.
8. Cho Y, et al. Long-term survivals of stage IIB osteosarcoma: a 20-year experience in a single institution. Clin Orthop Surg 2011;3:48–54.
9. Davis LE, et al. Integration of genomic, transcriptomic and functional profiles of aggressive osteosarcomas across multiple species. Oncotarget 2017;8:76241–76256.
10. Dobin A, et al. STAR: ultrafast universal RNA-seq aligner. Bioinformatics 2013;29:15–21.
11. Fan TM, et al. Investigating CXCR4 expression in canine appendicular osteosarcoma. J Vet Intern Med 2008;22: 602–608.
12. Fang F, et al. Targeting the Wnt/β-catenin pathway in human osteosarcoma cells. Oncotarget 2018;9:36780–36792.
13. Fenger JM, et al. Canine osteosarcoma: a naturally occurring disease to inform pediatric oncology. ILAR J 2014;55: 69–85.
14. Fossey SL, et al. Characterization of STAT3 activation and expression in canine and human osteosarcoma. BMC Cancer 2009;9:81.
15. Gamberi G, et al. C-myc and c-fos in human osteosarcoma: prognostic value of mRNA and protein expression. Oncology 1998;55:556–563.
16. Gao W, et al. Knockdown of EIF3C promotes human U-2OS cells apoptosis through increased CASP3/7 and Chk1/2 by upregulating SAPK/JNK. Onco Targets Ther 2019;12:1225–1235.
17. Garcia-Monclús S, et al. EphA2 receptor is a key player in the metastatic onset of Ewing sarcoma. Int J Cancer 2018;143:1188–1201.
18. Gardner HL, et al. Canine osteosarcoma genome sequencing identifies recurrent mutations in DMD and the histone methyltransferase gene SETD2. Commun Biol 2019;2:1–13.
19. Giantin M, et al. Expression of the aryl hydrocarbon receptor pathway and cyclooxygenase-2 in dog tumors. Res Vet Sci 2013;94:90–99.
20. Gong T, et al. Expression of NF-κB and PTEN in osteosarcoma and its clinical significance. Oncol Lett 2017;14: 6744–6748.
21. Guijarro MV, et al. Animal models in osteosarcoma. Front Oncol 2014;4:189.
22. Hoeppner MP, et al. An improved canine genome and a comprehensive catalogue of coding genes and non-coding transcripts. PLoS One 2014;9:e91172.
23. Hu K, et al. mTOR signaling in osteosarcoma: oncogenesis and therapeutic aspects (review). Oncol Rep 2016;36: 1219–1225.
24. Hwang B, et al. Single-cell RNA sequencing technologies and bioinformatics pipelines. Exp Mol Med 2018;50:1–14.
25. Kadosawa T, et al. Establishment and characterization of a new cell line from a canine osteosarcoma. J Vet Med Sci 1994;56:1167–1169.
26. Kamel WA, et al. Simvastatin-induced apoptosis in osteosarcoma cells: a key role of RhoA-AMPK/p38 MAPK signaling in antitumor activity. Mol Cancer Ther 2017;16:182–192.
27. Kelleher FC, et al. Targeting the centrosome and polo-like kinase 4 in osteosarcoma. Carcinogenesis 2019;40: 493–499.
28. Krämer A, et al. Causal analysis approaches in Ingenuity Pathway Analysis. Bioinformatics 2014;30:523–530.
29. Leeper H, et al. Preliminary evaluation of serum total cholesterol concentrations in dogs with osteosarcoma. J Small Anim Pract 2017;58:562–569.
30. Levine RA, Fleischli MA. Inactivation of p53 and retinoblastoma family pathways in canine osteosarcoma cell lines. Vet Pathol 2000;37:54–61.
31. Li C, et al. The canonical Wnt-beta-catenin pathway in development and chemotherapy of osteosarcoma. Front Biosci (Landmark Ed) 2013;18:1384–1391.
32. Li X, et al. A novel interplay between HOTAIR and DNA methylation in osteosarcoma cells indicates a new therapeutic strategy. J Cancer Res Clin Oncol 2017;143:2189–2200.
33. Li Y, et al. The possible role of insulin-like growth factor-1 in osteosarcoma. Current Prob Cancer 2019;43:228–235.
34. Liao Y, et al. Androgen receptor is a potential novel prognostic marker and oncogenic target in osteosarcoma with dependence on CDK11. Sci Rep 2017;7:43941.

35. Liu JL, et al. MiR-144 inhibits tumor growth and metastasis in osteosarcoma via dual-suppressing RhoA/ROCK1 signaling pathway. Mol Pharmacol 2019;95:451–461.
36. Liu Q, Wang K. The induction of ferroptosis by impairing STAT3/Nrf2/GPx4 signaling enhances the sensitivity of osteosarcoma cells to cisplatin. Cell Biol Int 2019;43:1245–1256.
37. Liu S, Trapnell C. Single-cell transcriptome sequencing: recent advances and remaining challenges. F1000Res 2016;5. doi:10.12688/f1000research.7223.1.
38. Liu Y, et al. High expression levels of Cyr61 and VEGF are associated with poor prognosis in osteosarcoma. Pathol Res Pract 2017;213:895–899.
39. Lo Vasco VR, et al. Expression of phosphoinositide-specific phospholipase C enzymes in human osteosarcoma cell lines. J Cell Commun Signal 2013;7:141–150.
40. Lo Vasco VR, et al. U-73122 reduces the cell growth in cultured MG-63 osteosarcoma cell line involving phosphoinositide-specific phospholipases C. SpringerPlus 2016;5:156.
41. Loukopoulos P, et al. Clinical and pathologic relevance of p53 index in canine osseous tumors. Vet Pathol 2003;40: 237–248.
42. Luecken MD, Theis FJ. Current best practices in single-cell RNA-seq analysis: a tutorial. Mol Syst Biol 2019;15:e8746.
43. Maaten L van der. Accelerating t-SNE using tree-based algorithms. J Machine Learning Res 2014;15:3221–3245.
44. MacEwen EG, et al. C-Met tyrosine kinase receptor expression and function in human and canine osteosarcoma cells. Clin Exp Metastasis 2003;20:421–430.
45. MacEwen EG, et al. IGF-1 receptor contributes to the malignant phenotype in human and canine osteosarcoma. J Cell Biochem 2004;92:77–91.
46. Maniscalco L, et al. PDGFs and PDGFRs in canine osteosarcoma: new targets for innovative therapeutic strategies in comparative oncology. Vet J 2013;195:41–47.
47. Mardanpour K, et al. Coexistence of her2, ki67, and p53 in osteosarcoma: a strong prognostic factor. North Am J Med Sci 2016;8:210.
48. Mason NJ, et al. Immunotherapy with a HER2-targeting *Listeria* induces HER2-specific immunity and demonstrates potential therapeutic effects in a phase I trial in canine osteosarcoma. Clin Cancer Res 2016;22:4380–4390.
49. Mauchle U, et al. Identification of anti-proliferative kinase inhibitors as potential therapeutic agents to treat canine osteosarcoma. Vet J 2015;205:281–287.
50. Maurizi G, et al. Sox2 is required for tumor development and cancer cell proliferation in osteosarcoma. Oncogene 2018;37:4626–4632.
51. Miki Y, et al. Roles of aryl hydrocarbon receptor in aromatase-dependent cell proliferation in human osteoblasts. Int J Mol Sci 2017;18:2159.
52. Milner RJ, et al. Bisphosphonates and cancer. J Vet Intern Med 2004;18:597–604.
53. Mirabello L, et al. International osteosarcoma incidence patterns in children and adolescents, middle ages, and elderly persons. Int J Cancer 2009;125:229–234.
54. Misaghi A, et al. Osteosarcoma: a comprehensive review. SICOT J 2018;4:12.
55. Moriarity BS, et al. A Sleeping Beauty forward genetic screen identifies new genes and pathways driving osteosarcoma development and metastasis. Nat Genet 2015;47:615–624.
56. Paoloni M, et al. Canine tumor cross-species genomics uncovers targets linked to osteosarcoma progression. BMC Genomics 2009;10:625.
57. Phillips B, et al. Use of single-agent carboplatin as adjuvant or neoadjuvant therapy in conjunction with amputation for appendicular osteosarcoma in dogs. J Am Anim Hosp Assoc 2009;45:33–38.
58. Piskun CM, Stein TJ. β-Catenin transcriptional activity is minimal in canine osteosarcoma and its targeted inhibition results in minimal changes to cell line behaviour. Vet Comp Oncol 2016;14:e4–e16.
59. Presneau N, et al. Post-translational regulation contributes to the loss of LKB1 expression through SIRT1 deacetylase in osteosarcomas. Br J Cancer 2017;117:398–408.
60. Ren X, et al. Understanding tumor ecosystems by single-cell sequencing: promises and limitations. Genome Biol 2018;19:211.
61. Rickel K, et al. Molecular genetics of osteosarcoma. Bone 2017;102:69–79.
62. Roy J, et al. Comparative proteomic investigation of metastatic and non-metastatic osteosarcoma cells of human and canine origin. PLoS One 2017;12:e0183930.
63. Sakthikumar S, et al. SETD2 is recurrently mutated in whole-exome sequenced canine osteosarcoma. Cancer Res 2018;78:3421–3431.
64. Satterfield L, et al. miR-130b directly targets ARHGAP1 to drive activation of a metastatic CDC42-PAK1-AP1 positive feedback loop in Ewing sarcoma. Int J Cancer 2017;141:2062–2075.
65. Selvarajah GT, et al. Gene expression profiling of canine osteosarcoma reveals genes associated with short and long survival times. Mol Cancer 2009;8:72.
66. Selvarajah GT, et al. Expression of epidermal growth factor receptor in canine osteosarcoma: association with clinicopathological parameters and prognosis. Vet J 2012;193: 412–419.
67. Selvarajah GT, et al. Reference gene validation for gene expression normalization in canine osteosarcoma: a geNorm algorithm approach. BMC Vet Res 2017;13:354.
68. Shoieb AM, et al. An in vivo/in vitro experimental model system for the study of human osteosarcoma: canine osteosarcoma cells (COS31) which retain osteoblastic and metastatic properties in nude mice. In Vivo 1998;12:463–472.
69. Shoieb AM, et al. In vitro reversal of glutathione-S-transferase-mediated resistance in canine osteosarcoma (COS31) cells. In Vivo 1998;12:455–462.
70. Simpson S, et al. Comparative review of human and canine osteosarcoma: morphology, epidemiology, prognosis, treatment and genetics. Acta Vet Scand 2017;59:71.
71. Skorupski KA, et al. Carboplatin versus alternating carboplatin and doxorubicin for the adjuvant treatment of canine appendicular osteosarcoma: a randomized, phase III trial. Vet Comp Oncol 2016;14:81–87.
72. Solomon DA, et al. Sample type bias in the analysis of cancer genomes. Cancer Res 2009;69:5630–5633.
73. Sottnik JL, et al. Induction of VEGF by tepoxalin does not lead to increased tumour growth in a canine osteosarcoma xenograft. Vet Comp Oncol 2011;9:118–130.
74. Tang Y, et al. CXCR3 from chemokine receptor family correlates with immune infiltration and predicts poor survival in osteosarcoma. Biosci Rep 2019;39:BSR20192134.

75. Thamm DH, et al. Serum vascular endothelial growth factor concentrations and postsurgical outcome in dogs with osteosarcoma. Vet Comp Oncol 2008;6:126–132.
76. Tong M, et al. New insights from the widening homogeneity perspective to target intratumor heterogeneity. Cancer Commun (Lond) 2018;38:17.
77. Tuohy JL, et al. Demographic characteristics, site and phylogenetic distribution of dogs with appendicular osteosarcoma: 744 dogs (2000–2015). PLoS One 2019;14:e0223243.
78. Valihrach L, et al. Platforms for single-cell collection and analysis. Int J Mol Sci 2018;19:807.
79. Wang S, et al. Valproic acid combined with zoledronate enhance γδ T cell-mediated cytotoxicity against osteosarcoma cells via the accumulation of mevalonate pathway intermediates. Front Immunol 2018;9:377.
80. Wang Y, et al. Clonal evolution in breast cancer revealed by single nucleus genome sequencing. Nature 2014;512:155–160.
81. Wang Z, et al. The association of glutathione S-transferase polymorphisms in patients with osteosarcoma: evidence from a meta-analysis. Eur J Cancer Care 2015;24:417–424.
82. Wu J, et al. Clinical significance of the phosphorylation of MAPK and protein expression of cyclin D1 in human osteosarcoma tissues. Mol Med Reports 2017;15:2303–2307.
83. Xian M, et al. Bortezomib sensitizes human osteosarcoma cells to adriamycin-induced apoptosis through ROS-dependent activation of p-eIF2α/ATF4/CHOP axis. Int J Cancer 2017;141:1029–1041.
84. Xu J, et al. PDGF/PDGFR effects in osteosarcoma and the "add-on" strategy. Clin Sarcoma Res 2018;8:15.
85. Yang S-C, et al. Exposure to 2,3,7,8-tetrachlorodibenzo-p-dioxin increases the activation of aryl hydrocarbon receptor and is associated with the aggressiveness of osteosarcoma MG-63 osteoblast-like cells. Oncol Lett 2018;16:3849–3857.
86. Yang X, et al. Identification of genes associated with methotrexate resistance in methotrexate-resistant osteosarcoma cell lines. J Orthop Surg Res 2015;10:136.
87. Zhang J, et al. PI3K/Akt signaling in osteosarcoma. Clin Chim Acta 2015;444:182–192.
88. Zhou Q, et al. Genetic variants of lncRNA HOTAIR contribute to the risk of osteosarcoma. Oncotarget 2016;7:19928–19934.