AWARD NUMBER: W81XWH-20-1-0294

TITLE: Chromatin Accessibility and the Convergent Oncogenic Pathways of Angiosarcomas

PRINCIPAL INVESTIGATOR: Jong Hyuk Kim

CONTRACTING ORGANIZATION: University of Minnesota, Twin Cities, Minneapolis, MN

REPORT DATE: July 2022

TYPE OF REPORT: Annual

PREPARED FOR: U.S. Army Medical Research and Development Command
Fort Detrick, Maryland 21702-5012

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
July 2022	Annual	01Jul2021-30Jun2022

4. TITLE AND SUBTITLE

Chromatin Accessibility and the Convergent Oncogenic Pathways of Angiosarcomas

5a. CONTRACT NUMBER
W81XWH-20-1-0294

5b. GRANT NUMBER
CA191225

5c. PROGRAM ELEMENT NUMBER

5d. PROJECT NUMBER
00081733

5e. TASK NUMBER

5f. WORK UNIT NUMBER

6. AUTHOR(S)

Jong Hyuk Kim, DVM, Ph.D
E-Mail: jkim@umn.edu; jkim19@ufl.edu

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

University of Minnesota
420 Delaware St., SE
Minneapolis, MN 55455-2070

University of Florida
Grinter Hall
PO Box 115500
Gainesville, FL 32611-5500

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

U.S. Army Medical Research and Development Command
Fort Detrick, Maryland 21702-5012

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT

Approved for Public Release; Distribution Unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

Angiosarcoma is a rare type of soft tissue sarcoma, with a prevalence of fewer than 300 cases in the US annually. Our understanding of the oncogenic mechanisms of aggressive angiosarcomas is rudimentary, and our ultimate goal is to develop appropriate and effective treatment options and protocols for patients with this disease. Angiosarcoma are genomically complex; however, they share a histological morphology that consists of disorganized, malignant vessel-forming cells. Our hypothesis is that chromatin accessibility is necessary to establish the mutational landscape, which consequently activates convergent signaling pathways that contribute to angiosarcoma development. In this report period, we established chromatin accessibility and the transcriptomic landscape in TP53 mutant hemangioblast cells differentiated from human induced pluripotent stem cells. This approach allows us to develop *in vitro* tumor models to define molecular mechanisms that regulate convergent oncogenic pathways in angiosarcomas. We will determine if p53 deficiency in hemangioblasts contributes to angiosarcoma development. This project will impact our understanding of aggressive angiosarcomas and specifically enhance our basic knowledge of how morphologic convergence with genetic chaos arises and contributes to angiosarcoma development. This career development award also supported the PI, Dr. Kim's career goal to develop an independent research program to advance our understanding of aggressive sarcomas. Likewise, Dr. Kim was hired by University of Florida as a tenure-track faculty in this period, and he successfully continues to establish his outstanding research program.

15. SUBJECT TERMS
None listed.

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			USAMRDC
Unclassified	Unclassified	Unclassified	Unclassified	35	19b. TELEPHONE NUMBER *(include area code)*

TABLE OF CONTENTS

1. INTRODUCTION:

Angiosarcoma is a rare type of soft tissue sarcoma, with a prevalence of fewer than 300 cases in the US each year. These tumors are highly aggressive and metastatic, and more than half of the patients with this disease die within the first year of diagnosis. Angiosarcomas are genomically complex; however, they share a histological morphology that consists of disorganized, malignant vessel-forming cells. Our objective is to establish chromatin accessibility and the mutational landscape, which activate convergent signaling pathways contributing to angiosarcoma development. Specifically, we develop tumor models to define molecular mechanisms that regulate convergent oncogenic pathways in angiosarcomas using induced pluripotent stem cells and genome engineering. From this approach, this project tests a new concept that could change the paradigms for addressing the fundamental oncogenic mechanisms of angiosarcomas.

2. KEYWORDS:

Angiosarcoma; rare cancer; induced pluripotent stem cell; CRISPR/Cas9; ATAC-seq; RNA-seq; chromatin accessibility; transcriptomics; hemangioblast; sarcoma modeling; single cell genomics

3. ACCOMPLISHMENTS:

What were the major goals of the project?

The project has three specific aims. In this report for Year-2, we list specific aim 1 and 2 with pertinent major tasks and subtasks below. Details regarding the completion status are provided on section 5, Changes/Problems.

Specific Aim 1: To establish chromatin accessibility and the transcriptomic landscape in angiosarcomas.

Major Task 1: Generation of next generation sequencing data from human angiosarcomas project

Subtask 1: Submit HRPO and ACURO documentation to DoD

- Completed on Jun 3rd and Jun 25th, 2020 at the University of Minnesota; ongoing update on documentation at the University of Florida attributable to PI's transition

Subtask 2: Sample preparation of primary tissue samples for ATAC-Seq and RNA-Seq generation

- Partially completed (50% of completion) and deferred

Subtask 3: Quality control analysis of samples and sequencing run

- Partially completed (50% of completion) and deferred

Subtask 4: Initial bioinformatic analysis of sequenced data

- Partially completed (50% of completion) and ongoing

Subtask 5: Advanced bioinformatic analysis of ATAC-Seq and RNA-Seq

- Partially completed (50% of completion) and ongoing

Subtask 6: Project meetings

- Partially completed (90% of completion) for this period due to PI's transition

Subtask 7: Career development for Dr. Kim (learning ATAC-Seq analysis and developing the application)

- Partially completed (75% of completion) and ongoing

Major Task 2: Establishment of angiosarcoma xenografts and generation of single cell-ATAC-Seq and -RNA-Seq

Subtask 1: Ordering and housing mice in animal facilities
- Deferred

Subtask 2: Culture and expand angiosarcoma (AS5 and ISO-HAS) cells
- Partially completed (50% of completion) and deferred

Subtask 3: Mice xenograft experiment by transplantation of angiosarcoma cells
- Deferred

Subtask 4: Sacrifice mice, generation of histological samples, and harvest xenograft tumors to prepare single cells of tumors
- Deferred

Subtask 5: Generation of single cell-sequencing (scATAC-Seq and scRNA-Seq) data libraries
- Partially completed (40% of completion) and deferred

Subtask 6: Quality control analysis of samples and sequencing run
- Partially completed (40% of completion) and ongoing

Subtask 7: Initial bioinformatic analysis of sequenced data
- Partially completed (50% of completion) and ongoing

Subtask 8: Project meetings
- Partially completed (90% of completion) for this period due to PI's transition

Subtask 9: Career development for Dr. Kim (learning scATAC-Seq and scRNA-Seq)
- Partially completed (50% of completion) and ongoing

Specific Aim 2: To develop in vitro tumor models to define molecular mechanisms that regulate convergent oncogenic pathways in angiosarcomas
Major Task 1: Gene engineering and differentiation in iPSCs

Subtask 1: Generation of iPSCs and preparation of reagents; commercially available iPSC cell lines (iPS12-10 and BYS-0110)
- Partially completed (75% of completion) and deferred

Subtask 2: Engineering gene mutation (*TP53*, *PIK3CA*, *TP53*/*PIK3CA*) in iPSCs
- Partially completed (40% of completion) and deferred

Subtask 3: Differentiation of hemangioblasts from engineered iPSCs
- Partially completed (50% of completion) and deferred

Subtask 4: Functional validation of engineered cells
- Partially completed (50% of completion) and deferred

Subtask 5: Cell line authentication and Mycoplasma screening (iPSCs-derived cells, HUVEC, fibroblasts, AS5, ISO-HAS)
- Partially completed (50% of completion) and deferred
We performed cell line authentication for AS5 and ISO-HAS cells as described above. We have not grown HUVEC and fibroblasts in this period, and we will continue to do cell authentication when cell culture work resumes in the PI's new lab.

Subtask 6: Project meetings
- Partially completed (90% of completion) for this period due to PI's transition

Subtask 7: Career development for Dr. Kim (acquisition of new experimental skills for iPSCs generation and genome engineering)
- Partially completed (50% of completion) and ongoing

Subtask 8: Career development for Dr. Kim (starting development of strategies to secure funding)
- Partially completed (75% of completion) and ongoing

What was accomplished under these goals?
For this report period, we provide major changes and accomplishment of the project regarding PI's transition from the University of Minnesota, an originally awarded institution to the University of Florida on section 5, Changes/Problems.

Specific Aim 1: To establish chromatin accessibility and the transcriptomic landscape in angiosarcomas.

Subtask 3: Quality control analysis of samples and sequencing run
- Partially completed (50% of completion) and deferred
In the previous report, we showed that ATAC-seq and RNA-seq data libraries were generated from commercially available normal endothelial cells and osteoblast (HUVEC, HUAEC, HOB). In this project period, we generated additional ATAC-seq and RNA-seq from iPSC-derived cells performed in Specific Aim 2. We determined quality of the sequencing data and ran bioinformatic algorithms for data analysis.

ATAC-seq data: Human iPSCs (day 0) were differentiated into mesoderm (day 2), hemangioblast (day 5), and endothelial cells (day 8 or 9). A total of 28 samples obtained at each differential stage including replicates were sequenced. In addition, to determine biological consequence of *PIK3CA* mutations, we induced *PIK3CA* mutations in a canine hemangiosarcoma cell line, called DHSA-1426 (as alternative cell line for human angiosarcoma) using CRISPR/Cas9. Two mutant clones (C8 and C35) and one WT clone were obtained by single cell limiting dilution. A total of 12 samples were sequenced for ATAC-seq.

Cell culture and sequencing were done as described in previous report. Each sample was sequenced to a targeted depth of 50 million reads (2 × 50 pair-end) on NovaSeq. The mean quality score was above the cutoff of 30 for all sequencing libraries. The total number of sequencing reads per sample and sample annotation data are provided in **Table 1**.

Table 1. ATAC-seg generation of human iPSC-derived cells and canine hemangiosarcoma cells

Sample	Total Sequences	Batch #	Species	Differentiation day	Mutation	Cell line
Trial_8_D0_WT_1	58,288,658	1	Human	0	WT	iPSC
Trial_8_D0_WT_2	40,464,927	1	Human	0	WT	iPSC
Trial_8_D2_WT_1	53,642,401	1	Human	2	WT	iPSC
Trial_8_D2_WT_2	57,465,292	1	Human	2	WT	iPSC
Trial_8_D5_WT_CD34_1	54,788,152	1	Human	5	WT	iPSC
Trial_8_D5_WT_CD34_2	49,311,025	1	Human	5	WT	iPSC
Trial_9_P53KO_CD34_1	89,162,158	2	Human	8	TP53 KO	iPSC
Trial_9_P53KO_CD34_2	93,166,635	2	Human	8	TP53 KO	iPSC
Trial_9_P53KO_D0_1	63,279,819	2	Human	0	TP53 KO	iPSC
Trial_9_P53KO_D0_2	81,706,902	2	Human	0	TP53 KO	iPSC
Trial_9_P53KO_D2_1	77,319,614	2	Human	2	TP53 KO	iPSC
Trial_9_P53KO_D2_2	70,482,434	2	Human	2	TP53 KO	iPSC
Trial_9_P53KO_D5_1	80,828,950	2	Human	5	TP53 KO	iPSC
Trial_9_P53KO_D5_2	77,687,685	2	Human	5	TP53 KO	iPSC
Trial_9_WT_D0_1	70,911,954	2	Human	0	WT	iPSC
Trial_9_WT_D0_2	64,787,520	2	Human	0	WT	iPSC
Trial_9_WT_D2_1	75,215,029	2	Human	2	WT	iPSC
Trial_9_WT_D2_2	71,411,304	2	Human	2	WT	iPSC
Trial_9_WT_D5_1	70,268,611	2	Human	5	WT	iPSC
Trial_9_WT_D5_2	69,040,891	2	Human	5	WT	iPSC
Trial_10_D0_P53KO_1	62,680,585	3	Human	0	TP53 KO	iPSC
Trial_10_D0_P53KO_2	67,558,368	3	Human	0	TP53 KO	iPSC
Trial_10_D2_P53KO_1	45,170,589	3	Human	2	TP53 KO	iPSC
Trial_10_D2_P53KO_2	57,794,465	3	Human	2	TP53 KO	iPSC
Trial_10_D5_P53KO_1	67,750,399	3	Human	5	TP53 KO	iPSC
Trial_10_D5_P53KO_2	71,381,429	3	Human	5	TP53 KO	iPSC
Trial_10_D9_P53KO_1	64,472,928	3	Human	9	TP53 KO	iPSC
Trial_10_D9_P53KO_2	75,514,332	3	Human	9	TP53 KO	iPSC
DHSA_1426_C35_A1	96,875,829	n/a	Canine	n/a	*PIK3CA* mutation	C35
DHSA_1426_C35_A2	102,309,811	n/a	Canine	n/a	*PIK3CA* mutation	C35
DHSA_1426_C35_B1	82,733,756	n/a	Canine	n/a	*PIK3CA* mutation	C35

DHSA_1426_C35_B2	79,984,855	n/a	Canine	n/a	*PIK3CA* mutation	C35
DHSA_1426_C35_C1	59,220,280	n/a	Canine	n/a	*PIK3CA* mutation	C35
DHSA_1426_C35_C2	60,127,908	n/a	Canine	n/a	*PIK3CA* mutation	C35
DHSA_1426_C73_B1	91,282,731	n/a	Canine	n/a	WT	C73
DHSA_1426_C73_B2	79,673,023	n/a	Canine	n/a	WT	C73
DHSA_1426_C73_C1	38,494,358	n/a	Canine	n/a	WT	C73
DHSA_1426_C73_C2	56,611,070	n/a	Canine	n/a	WT	C73
DHSA_1426_C8_A1	113,473,683	n/a	Canine	n/a	*PIK3CA* mutation	C8
DHSA_1426_C8_A2	90,496,877	n/a	Canine	n/a	*PIK3CA* mutation	C8

RNA-seq data: Human iPSC-derived cells and canine hemangiosarcoma cells were used for RNA-seq data generation in parallel with ATAC-seq data. A total of 28 samples matched to that for ATAC-seq were sequenced for RNA-seq. RNA-seq data (n=6) generated from *PIK3CA* mutant and WT canine hemangiosarcoma cells were also provided. 1 ug of each RNA sample was submitted to the UMGC, and TruSeq stranded mRNA libraries were created and sequenced to a targeted depth of 20 million reads (2 × 150 pair-end) on NovaSeq. The mean quality score was above the cutoff of 30 for all sequencing libraries. The total number of sequencing reads per sample and sample annotation data are provided in **Table 2**.

Table 2. RNA-seq data generation of human iPSC-derived cells and canine hemangiosarcoma cells

Sample	Total Reads	Batch #	Species	Differentiation day	Mutation	Cell line
Trial_8_D0_WT_1	20080053	1	Human	0	WT	iPSC
Trial_8_D0_WT_2	24352776	1	Human	0	WT	iPSC
Trial_8_D2_WT_1	27887402	1	Human	2	WT	iPSC
Trial_8_D2_WT_2	24913590	1	Human	2	WT	iPSC
Trial_8_D5_WT_CD34_1	35111276	1	Human	5	WT	iPSC
Trial_8_D5_WT_CD34_2	21029399	1	Human	5	WT	iPSC
Trial_9_P53KO_CD34_1	17795405	2	Human	8	TP53 KO	iPSC
Trial_9_P53KO_CD34_2	31515107	2	Human	8	TP53 KO	iPSC
Trial_9_P53KO_D0_1	23933395	2	Human	0	TP53 KO	iPSC
Trial_9_P53KO_D0_2	21989156	2	Human	0	TP53 KO	iPSC
Trial_9_P53KO_D2_1	19119977	2	Human	2	TP53 KO	iPSC

Trial_9_P53KO_D2_2	25373460	2	Human	2	TP53 KO	iPSC
Trial_9_P53KO_D5_1	28065884	2	Human	5	TP53 KO	iPSC
Trial_9_P53KO_D5_2	17984490	2	Human	5	TP53 KO	iPSC
Trial_9_WT_D0_1	18369685	2	Human	0	WT	iPSC
Trial_9_WT_D0_2	23250301	2	Human	0	WT	iPSC
Trial_9_WT_D2_1	25491936	2	Human	2	WT	iPSC
Trial_9_WT_D2_2	16350677	2	Human	2	WT	iPSC
Trial_9_WT_D5_1	20649870	2	Human	5	WT	iPSC
Trial_9_WT_D5_2	17959300	2	Human	5	WT	iPSC
Trial_10_D0_P53KO_1	19111930	3	Human	0	TP53 KO	iPSC
Trial_10_D0_P53KO_2	17808400	3	Human	0	TP53 KO	iPSC
Trial_10_D2_P53KO_1	20171737	3	Human	2	TP53 KO	iPSC
Trial_10_D2_P53KO_2	19382984	3	Human	2	TP53 KO	iPSC
Trial_10_D5_P53KO_1	16068929	3	Human	5	TP53 KO	iPSC
Trial_10_D5_P53KO_2	22929275	3	Human	5	TP53 KO	iPSC
Trial_10_D9_P53KO_1	20105420	3	Human	9	TP53 KO	iPSC
Trial_10_D9_P53KO_2	19745893	3	Human	9	TP53 KO	iPSC
DHSA_1426_WT_1	18661467	n/a	Canine	n/a	WT	C73
DHSA_1426_WT_2	20083499	n/a	Canine	n/a	WT	C73
DHSA_1426_C8_1	16203508	n/a	Canine	n/a	*PIK3CA* mutation	C8
DHSA_1426_C8_2	18236695	n/a	Canine	n/a	*PIK3CA* mutation	C8
DHSA_1426_C35_1	20985233	n/a	Canine	n/a	*PIK3CA* mutation	C35
DHSA_1426_C35_2	17395900	n/a	Canine	n/a	*PIK3CA* mutation	C35

Subtask 4: Initial bioinformatic analysis of sequenced data

- Partially completed (50% of completion) and ongoing

The sequenced reads data as FASTQ format were deposited to the MSI server storage and transferred to the University of Florida HiPerGator supercomputer with PI's transition.

ATAC-seq data: In previous report, we showed that initial quality check of ATAC-seq data was done on Galaxy platform. During this period, we performed additional quality control and statistical signal processing of short-read sequencing data, producing alignments and measures of transcription enrichment. Specifically, we used ENCODE ATAC-seq pipeline [Ref #1,2] and esATAC R package [Ref #3, 4, 5].

[1] https://www.encodeproject.org/atac-seq/
[2] https://github.com/ENCODE-DCC/atac-seq-pipeline
[3] https://rdrr.io/bioc/esATAC/
[4] https://github.com/wzthu/esATAC
[5] https://www.bioconductor.org/packages/release/bioc/html/esATAC.html

This approach cross-checks quality control process and peak calling (i.e., regions of transcription enrichment). Both ENCODE ATAC-seq and esATAC pipeline were successfully ran on our FASTQ files, mapped with hg38 human reference genome. Analysis of canine data is in process by building customized canine reference genome as required in each algorithm. Representative screenshots of QC reports generated by those pipelines are provided in **Figure 1** (ENCODE ATAC-seq) and **Figure 2** (esATAC).

QC Report

	general
Report generated at	2021-11-13 12:21:44
Title	Trial_8_D2_WT_1_S33_R1,R2
Description	Trial_8_D2_WT_1_S33_R1,R2
Pipeline version	v2.0.1
Pipeline type	atac
Genome	hg38
Aligner	bowtie2
Sequencing endedness	{'rep1': {'paired_end': True}}
Peak caller	macs2

Alignment quality metrics

SAMstat (raw unfiltered BAM)

	rep1
Total Reads	55772290
Total Reads (QC-failed)	0
Duplicate Reads	0
Duplicate Reads (QC-failed)	0
Mapped Reads	43298352
Mapped Reads (QC-failed)	0
% Mapped Reads	77.60000000000001
Paired Reads	55772290
Paired Reads (QC-failed)	0
Read1	27886145
Read1 (QC-failed)	0
Read2	27886145
Read2 (QC-failed)	0
Properly Paired Reads	38406240
Properly Paired Reads (QC-failed)	0
% Properly Paired Reads	68.89999999999999
With itself	40459926
With itself (QC-failed)	0
Singletons	2838426
Singletons (QC-failed)	0
% Singleton	5.1
Diff. Chroms	324854
Diff. Chroms (QC-failed)	0

Marking duplicates (filtered BAM)

	rep1
Unpaired Reads	0
Paired Reads	15901773
Unmapped Reads	0
Unpaired Duplicate Reads	0
Paired Duplicate Reads	5133521
Paired Optical Duplicate Reads	261868
% Duplicate Reads	32.2827

Filtered out (samtools view -F 1804):

- read unmapped (0x4)
- mate unmapped (0x8, for paired-end)
- not primary alignment (0x100)
- read fails platform/vendor quality checks (0x200)
- read is PCR or optical duplicate (0x400)

Fraction of mitochondrial reads (unfiltered BAM)

	rep1
Rn = Number of Non-mitochondrial Reads	43032753
Rm = Number of Mitochondrial Reads	2423652
Rm/(Rn+Rm) = Frac. of mitochondrial reads	0.053318162753961736

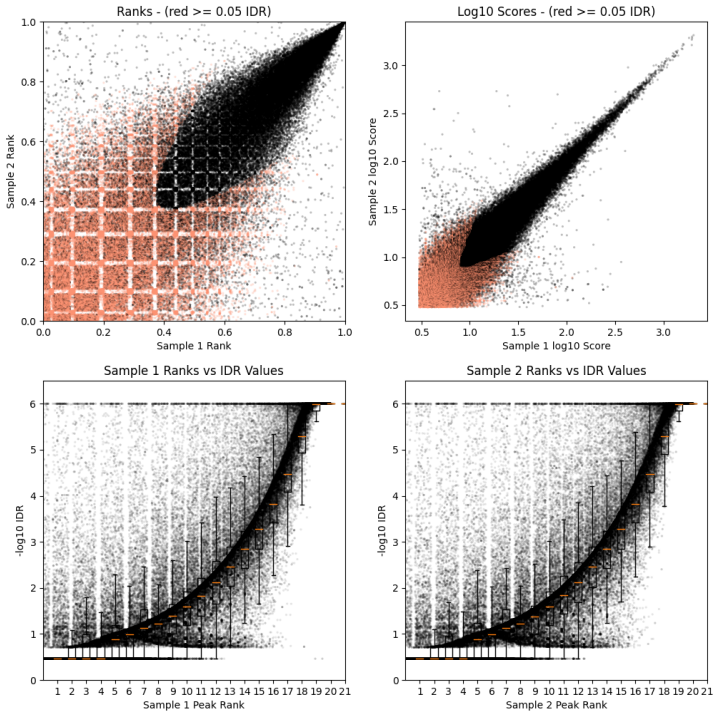
SAMstat (filtered/deduped BAM)

	rep1
Total Reads	21116352
Total Reads (QC-failed)	0
Duplicate Reads	0
Duplicate Reads (QC-failed)	0
Mapped Reads	21116352
Mapped Reads (QC-failed)	0
% Mapped Reads	100.0
Paired Reads	21116352
Paired Reads (QC-failed)	0
Read1	10558176
Read1 (QC-failed)	0
Read2	10558176
Read2 (QC-failed)	0
Properly Paired Reads	21116352
Properly Paired Reads (QC-failed)	0

… (ellipsis)

Replication quality metrics

IDR (Irreproducible Discovery Rate) plots



rep1-pr1_vs_rep1-pr2

Peak calling statistics

Peak region size

	rep1	idr_opt	overlap_opt
Min size	150.0	153.0	150.0
25 percentile	332.0	416.0	360.0
50 percentile (median)	550.0	1066.0	763.0
75 percentile	1411.0	1944.0	1617.0
Max size	6531.0	6531.0	6531.0
Mean	987.3560382890532	1336.8875248215531	1120.7296568546112

…(ellipsis)

Figure 1. Example QC report of ATAC-seq data on representative human samples ran by ENCODE ATAC-seq pipeline.

esATAC Report

Zheng Wei and Wei Zhang

2021-03-27

```
#load("Report.Rdata")
library(esATAC)
loadConfig(file.path("ReportData","SummaryInfo","config.rds"))
suminfo <- readRDS(file.path("ReportData","SummaryInfo","suminfo.rds"))
allsteps <- suminfo[["prevSteps"]]
wholesummary <- suminfo[["wholesummary"]]
filtstat <- suminfo[["filtstat"]]
```

1 Summary Table

Sequence files below are set as inputs of the pipeline.

Mate1.files
<fct>

/Users/jhkim/Desktop/esATAC/WT_D21_OB_R1.fastqsanger.gz

1 row | 1-1 of 2 columns

Summerized infomation on sequence files has been shown showed below. You can see details in later sections

Item	Value	Reference
Sequence Files Type	paired end (PE)	SE / PE
Original total reads	54.3M	
– Reads after adapter removing (ratio)	54.2M (99.75%)	>99%

2 Sequence Statistics

2.1 FastQC

Quality control for the sequence data

Click to Visit Report (WT_D21_OB_R1.FastQC.pdf)

2.2 Remove adapter

The adapter sequence are shown below. For paired end reads, if adapters were not setted, the adapters below are identified by AdapterRemoval.

Item	Value
adapter1	CTGTCTCTTATACACATCTCCGAGCCCACGAGACCGTAAA
adapter2	CTGTCTCTTATACACATCTGACGCTGCCGACGACTCCTTA

The statistic of adapter removing are show below.

Item	Value
Total number of read pairs	54299222
Number of unaligned read pairs	30730252
Number of well aligned read pairs	23568970
Number of discarded mate 1 reads	136490
Number of singleton mate 1 reads	0
Number of discarded mate 2 reads	136490
Number of singleton mate 2 reads	0
Number of reads with adapters[1]	10173500
Number of retained reads	108325464
Number of retained nucleotides	5436193794
Average length of retained reads	50.1839

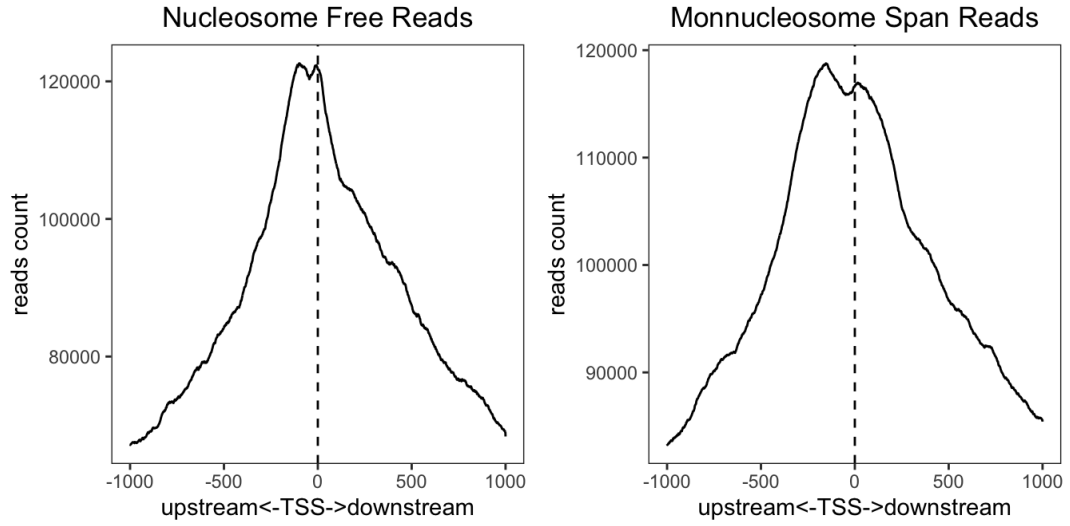
For detail, you can visit Website of AdapterRemoval on Github (https://github.com/MikkelSchubert/adapterremoval).

…(ellipsis)

3.5 TSS enrichment

The nucleosome free reads (<100bp) and monnucleosome span reads (180~247bp) enrichment around transcription starting site (TSS) are shown below.

```
library(ggplot2)
df<-report(allsteps$TSSQCNFR)[["tss"]]
g1<-ggplot(df,aes(pos,counts))+geom_line()+ geom_vline(xintercept = 0, linetype=2)+
xlab("upstream<-TSS->downstream")+ylab("reads count")+theme_bw() + theme(panel.grid
=element_blank()) + labs(title = "Nucleosome Free Reads") + theme(plot.title = elem
ent_text(hjust = 0.5))
df<-report(allsteps$TSSQCneucleosome)[["tss"]]
g2<-ggplot(df,aes(pos,counts))+geom_line()+ geom_vline(xintercept = 0, linetype=2)+
xlab("upstream<-TSS->downstream")+ylab("reads count")+theme_bw() + theme(panel.grid
=element_blank()) + labs(title = "Monnucleosome Span Reads") + theme(plot.title = e
lement_text(hjust = 0.5))
grid.arrange(g1, g2, ncol=2)
```



4 Peak Statistics

4.1 Blacklist ratio

Item	Value
Total peaks	360662
Blacklist regions	19
Ratio	0.00

4.2 DHS ratio

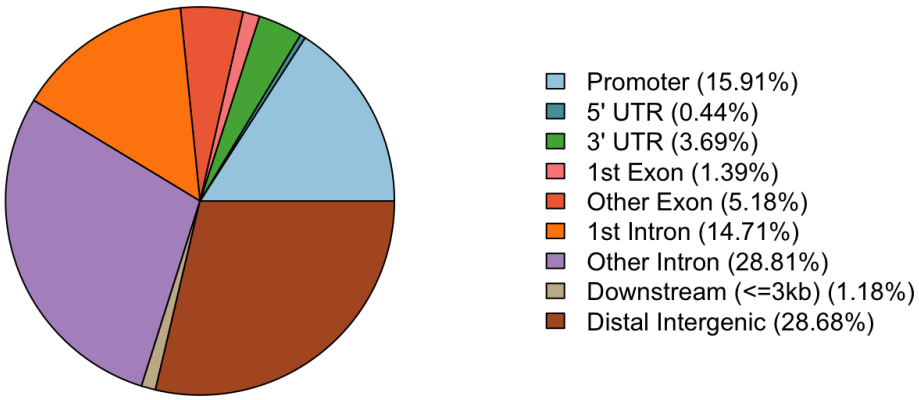
Item	Value
Total peaks	360662

Item	Value
DHS regions	135694
Ratio	0.38

4.3 Fraction of reads in peaks (FRiP)

Item	Value
The number of reads in peak	2725053
The number of total reads	7906889
The number of total peaks	360662
FRiP	0.34

4.4 Peak Annotation



5 Gene Ontology Analysis

Gene ontology analysis for all genes around peak regions.

ID	Description	GeneRatio	pvalue	qvalue
GO:0007409 (GO:0007409)	axonogenesis	455/15765	4.69e-20	2.08e-16
GO:0050769 (GO:0050769)	positive regulation of neurogenesis	456/15765	9.99e-17	2.22e-13

…(ellipsis)

Figure 2. Example QC report of ATAC-seq data on representative human samples ran by esATAC pipeline.

RNA-seq data: All human RNA-seq data (56 original FASTQ files; 28 samples) listed in Table 1 were used for quality check and genome mapping (hg38). The standard CHURP pipeline for RNA-seq at MSI, University of Minnesota was ran on the MSI Linux environment using Python module:

```
% module load python3
% python3 /home/msistaff/public/RNAseq_Tutorial/CHURP-slurm/churp.py bulk_rnaseq -w 72 -f
/home/jhkim/public/Kim6_Project_003/human_ipsc/ -x
/panfs/roc/risdb_new/ensembl/main/homo_sapiens/GRCh38/hisat2/genome -g
/panfs/roc/risdb_new/ensembl/main/homo_sapiens/GRCh38/annotation/Homo_sapiens.GRCh38.100.g
tf.gz -o /scratch.global/jhkim/CHURP_Output -d /scratch.global/jhkim/CHURP_Work
```

Report of the CHURP pipeline run includes QC metrics and mapping alignment as follows:

PURR Summary

2021-09-03, 21:07:54

1 PURR: Pipeline for UMII/RIS RNASeq Analysis

university of minnesota
informatics institute

PURR is the RNASeq analysis pipeline that is housed within CHURP, the **C**ollection of **H**ierarchical **U**MII/**R**IS **P**ipelines. CHURP is developed and maintained by the Research Informatics Solutions (RIS) group at MSI with funding provided by the University of Minnesota Informatics Institute (UMII).

https://www.msi.umn.edu/

https://research.umn.edu/units/umii

2 Key Figures and Summaries

2.1 Experiment Summary

The location of the FASTQ files for this run is `/panfs/roc/groups/13/jhkim/public/Kim6_Project_003/human_ipsc`

This run consists of `28` samples.

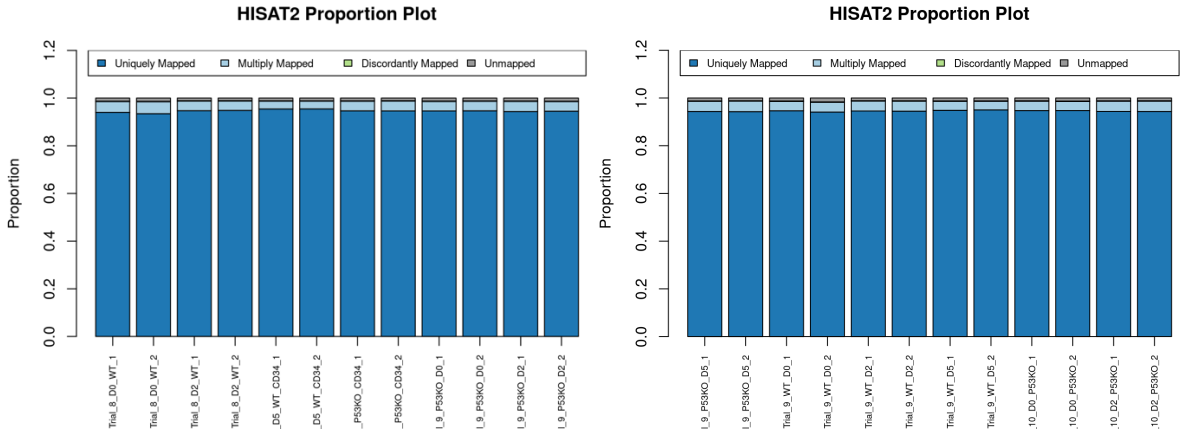
A tabular summary of the samples and reads that were analyzed is shown below:

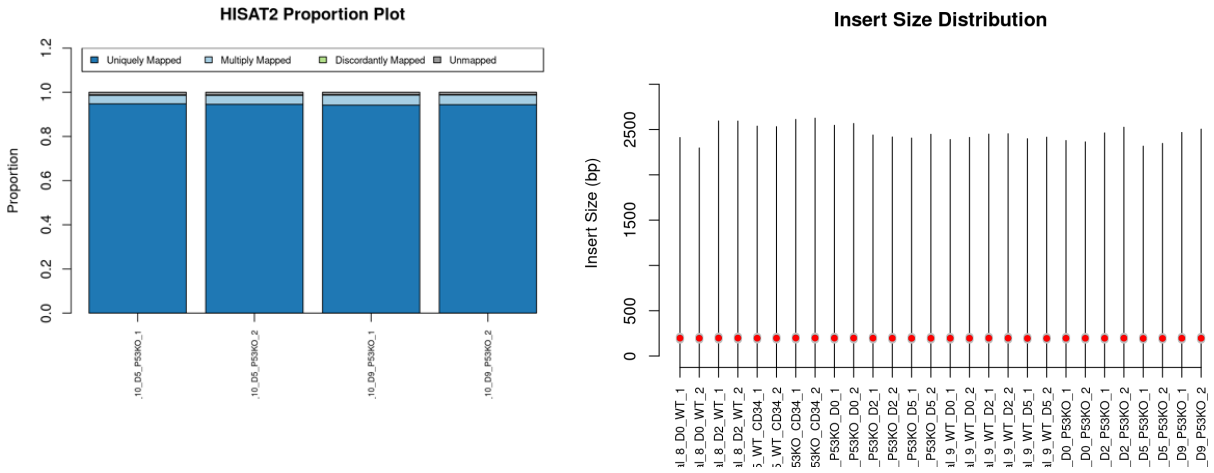
	Sample Name <chr>	**Group** <chr>	**R1** <chr>
1	Trial_10_D0_P53KO_1	NULL	Trial_10_D0_P53KO_1_S37_R1_001.fastq.gz

	Sample Name <chr>	Group <chr>	R1 <chr>
2	Trial_10_D0_P53KO_2	NULL	Trial_10_D0_P53KO_2_S38_R1_001.fastq.gz
3	Trial_10_D2_P53KO_1	NULL	Trial_10_D2_P53KO_1_S39_R1_001.fastq.gz
4	Trial_10_D2_P53KO_2	NULL	Trial_10_D2_P53KO_2_S40_R1_001.fastq.gz
5	Trial_10_D5_P53KO_1	NULL	Trial_10_D5_P53KO_1_S41_R1_001.fastq.gz
6	Trial_10_D5_P53KO_2	NULL	Trial_10_D5_P53KO_2_S42_R1_001.fastq.gz
7	Trial_10_D9_P53KO_1	NULL	Trial_10_D9_P53KO_1_S43_R1_001.fastq.gz
8	Trial_10_D9_P53KO_2	NULL	Trial_10_D9_P53KO_2_S44_R1_001.fastq.gz
9	Trial_8_D0_WT_1	NULL	Trial_8_D0_WT_1_S31_R1_001.fastq.gz
10	Trial_8_D0_WT_2	NULL	Trial_8_D0_WT_2_S32_R1_001.fastq.gz

Next
123
Previous
1-10 of 28 rows | 1-4 of 9 columns

…(ellipsis)





…(ellipsis)

Figure 3. Representative screenshots of RNA-seq analysis report from human samples.

Subtask 5: Advanced bioinformatic analysis of ATAC-Seq and RNA-Seq
- Partially completed (50% of completion) and ongoing

ATAC-seq analysis: We continued to run multiple ATAC-seq pipelines to compare output of different tools and to perform downstream analysis including ATACRseqQC, ChrAccR, and PEPATAC [Ref #6,7,8]. This analysis will allow us to identify open/closed chromatin regions associated with transcriptional regulation after combining ATAC-seq and RNA-seq data.

[6] ATACseqQC R package (https://bioconductor.org/packages/release/bioc/html/ATACseqQC.html)
[7] ChrAccR R package (https://github.com/GreenleafLab/ChrAccR)
[8] PEPATAC (http://pepatac.databio.org/)

RNA-seq analysis: Raw count reads generated by the CHURP pipeline were filtered and normalized using DESeq2 R package as described in previous report. Heatmap and unsupervised clustering data reveal distinct gene expression signatures depending on cell types differentiated from human iPSCs over time (**Figure 4**). Nine groups are classified as provided in **Table 3** for differential expression gene (DEG) analysis. Gene pathway analyses using Gene Ontology Resource (http://geneontology.org/), Gene Set Enrichment Analysis (https://www.gsea-msigdb.org/), and KEGG pathway database (https://www.genome.jp/kegg/) are in process.

Figure 4

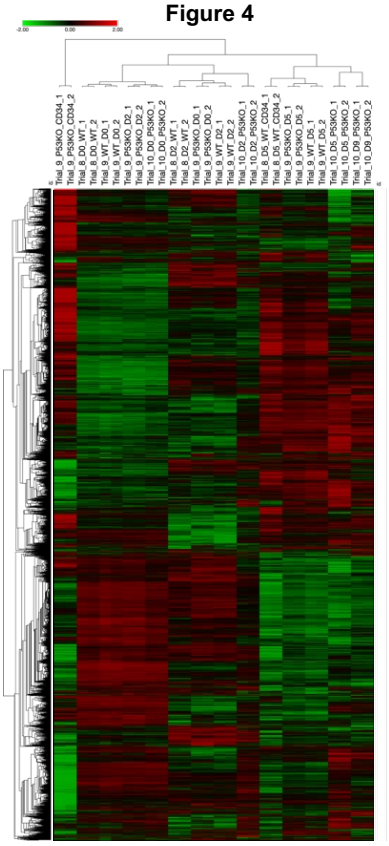


Table 3. Differential expression gene analysis between experiment groups of human iPSC-derived cells

Sample ID	Mutation	Day	Batch	DEG Group	
Trial_8_D0_WT_1	WT	0	1	WT_D0	Group 1
Trial_8_D0_WT_2	WT	0	1	WT_D0	Group 1
Trial_9_WT_D0_1	WT	0	2	WT_D0	Group 1
Trial_9_WT_D0_2	WT	0	2	WT_D0	Group 1
Trial_10_D0_P53KO_1	TP53 KO	0	3	P53_D0	Group 2
Trial_10_D0_P53KO_2	TP53 KO	0	3	P53_D0	Group 2
Trial_9_P53KO_D0_1	TP53 KO	0	2	P53_D0	Group 2
Trial_9_P53KO_D0_2	TP53 KO	0	2	P53_D0	Group 2
Trial_8_D2_WT_1	WT	2	1	WT_D2	Group 3
Trial_8_D2_WT_2	WT	2	1	WT_D2	Group 3
Trial_9_WT_D2_1	WT	2	2	WT_D2	Group 3
Trial_9_WT_D2_2	WT	2	2	WT_D2	Group 3
Trial_10_D2_P53KO_1	TP53 KO	2	3	P53_D2	Group 4
Trial_10_D2_P53KO_2	TP53 KO	2	3	P53_D2	Group 4
Trial_9_P53KO_D2_1	TP53 KO	2	2	P53_D2	Group 4
Trial_9_P53KO_D2_2	TP53 KO	2	2	P53_D2	Group 4
Trial_9_WT_D5_1	WT	5	2	WT_D5	Group 5
Trial_9_WT_D5_2	WT	5	2	WT_D5	Group 5
Trial_10_D5_P53KO_1	TP53 KO	5	3	P53_D5	Group 6
Trial_10_D5_P53KO_2	TP53 KO	5	3	P53_D5	Group 6
Trial_9_P53KO_D5_1	TP53 KO	5	2	P53_D5	Group 6
Trial_9_P53KO_D5_2	TP53 KO	5	2	P53_D5	Group 6
Trial_8_D5_WT_CD34_1	WT	5	1	WT_D5_CD34	Group 7
Trial_8_D5_WT_CD34_2	WT	5	1	WT_D5_CD34	Group 7
Trial_10_D9_P53KO_1	TP53 KO	9	3	P53_D9	Group 8
Trial_10_D9_P53KO_2	TP53 KO	9	3	P53_D9	Group 8
Trial_9_P53KO_CD34_1	TP53 KO	9	2	P53_D9_CD34	Group 9
Trial_9_P53KO_CD34_2	TP53 KO	9	2	P53_D9_CD34	Group 9

Figure 4. Gene expression profiling and unsupervised clustering analysis of RNA-seq data generated from human iPSC-derived cells. (A) Heatmap and hierarchical clustering data reveal distinct gene expression signatures between different cell types differentiated from iPSCs (n=28 as annotated in Table 3) over time. Log2 transformed; average linkage; subtracted row means, divided standard deviation; red bar = upregulated genes, green bar = downregulated genes.

<u>Subtask 6: Project meetings</u>

- Partially completed (90% of completion) for this period due to PI's transition

<u>Subtask 7: Career development for Dr. Kim (learning ATAC-Seq analysis and developing the application)</u>

- Partially completed (75% of completion) and ongoing

During this project, Dr. Kim, PI advanced his ability to run multiple bioinformatic algorithms for ATAC-seq analysis. He learned computational skills by analyzing different types of the analytic pipelines including Galaxy platform, multiple R packages, Linux use, and Python programming. These experiences led him to receive a competitive job offer from the University of Florida, and he was hired through the UF Artificial Intelligence (AI) Initiative. Dr. Kim will continue to develop comprehensive analytic approaches for multi-omics data including ATAC-seq and RNA-seq data by integrating AI and machine learning applications.

Major Task 2: Establishment of angiosarcoma xenografts and generation of single cell-ATAC-Seq and -RNA-Seq

Subtask 1: Ordering and housing mice in animal facilities
- Deferred

Subtask 2: Culture and expand angiosarcoma (AS5 and ISO-HAS) cells
- Partially completed (50% of completion) and deferred
In previous report, we showed that AS5 angiosarcoma cell line was authenticated with STR profiling by IDEXX Laboratories and expanded cells were stored in liquid nitrogen tanks. ISO-HAS cells were obtained from original research group in Japan. Cell authentication of ISO-HAS did not pass the standard criteria showing potential cell contamination, and all cells have been discarded from the inventory. As alternatives, we used two canine hemangiosarcoma cell lines, DHSA-1426 and COSB that have been known to develop malignant vascular tumor comparable with human angiosarcoma in immunodeficient mice. Those cells were authenticated by the same method, and we will use them for tumor xenograft experiments in the next project period.

Subtask 3: Mice xenograft experiment by transplantation of angiosarcoma cells
- Deferred

Subtask 4: Sacrifice mice, generation of histological samples, and harvest xenograft tumors to prepare single cells of tumors
- Deferred

Subtask 5: Generation of single cell-sequencing (scATAC-Seq and scRNA-Seq) data libraries
- Partially completed (40% of completion) and deferred
Single cell-sequencing data generation was originally planned to be done from mouse xenograft experiment of human angiosarcoma cells. The animal experiment was deferred due to the result of cell authentication and PI's transition in this period. Alternatively, we generated scRNA-seq data from canine hemangiosarcoma cells. Specifically, *PIK3CA* mutant cells (C35) and WT cells (C73) established from canine DHSA-1426 cells as described above (Table 1 and 2) were used for scRNA-seq data generation.

Cells were grown in normal culture condition, and cell viability was check for initial capture process. Cell capture and the 3' gene expression libraries met 10x Genomics specifications to move forward with the library preparation. Single cell sequencing libraries were created for targeting 10,000 cells per sample and 50,000 reads per cell (1b billion total reads; 2 x 150 pair-end). 10x Genomics Cell Ranger (v6.1.2) pipeline including `cellranger count` was ran to measure sequenced read counts and gene expression.

Subtask 6: Quality control analysis of samples and sequencing run
- Partially completed (40% of completion) and ongoing
Single cell sequencing results showed that estimated number of cells sequenced was 14,227, mean reads per cells were 84,777, median number of genes expressed per cells was 3,134 in *PIK3CA* WT canine hemangiosarcoma cells (C73; **Figure 5**). In *PIK3CA* mutant canine hemangiosarcoma cells (C35; **Figure 6**), estimated number of cells sequenced was 17,107, mean reads per cell were 90,618, and 3,353 genes were expressed per cell. Detailed sequencing metrics

including mapping results are described in Figure 5 and 6. t-distributed stochastic neighbor embedding (t-SNE) method was used to visualize single cell clusters within the sequenced cells. Individual t-SNE plots presented distinct patterns of single cell clusters between WT and mutant cells (**Figure 7**).

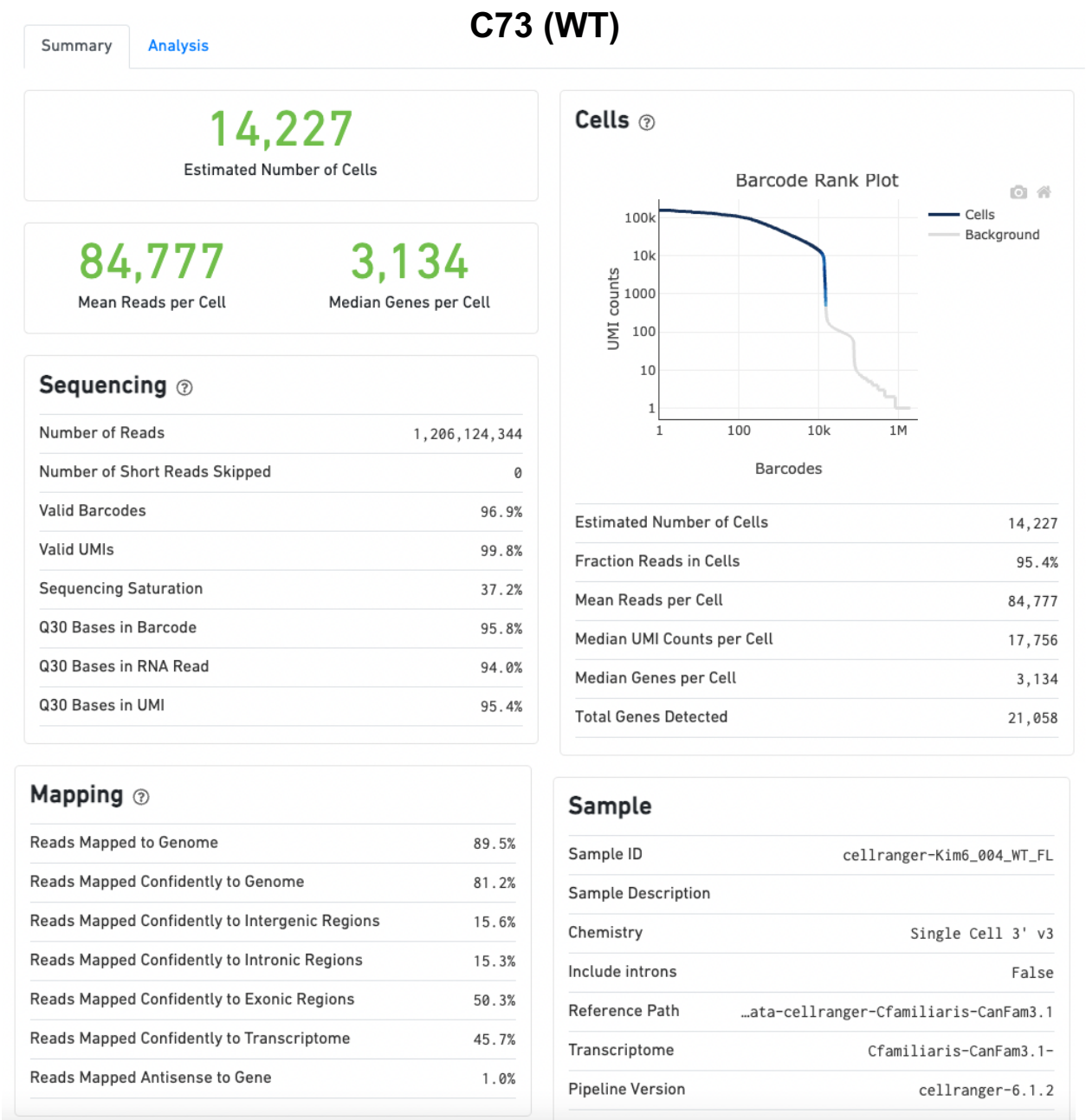


Figure 5. Single cell RNA-seq data of PIK3CA WT canine hemangiosarcoma cells (C73).

C35 (MT)

Summary | Analysis

17,107
Estimated Number of Cells

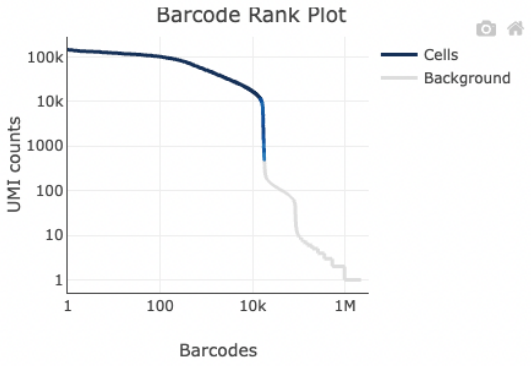
90,618
Mean Reads per Cell

3,353
Median Genes per Cell

Sequencing

Number of Reads	1,550,195,331
Number of Short Reads Skipped	0
Valid Barcodes	97.3%
Valid UMIs	99.8%
Sequencing Saturation	42.5%
Q30 Bases in Barcode	95.8%
Q30 Bases in RNA Read	94.0%
Q30 Bases in UMI	95.3%

Cells



Estimated Number of Cells	17,107
Fraction Reads in Cells	96.3%
Mean Reads per Cell	90,618
Median UMI Counts per Cell	18,845
Median Genes per Cell	3,353
Total Genes Detected	21,986

Mapping

Reads Mapped to Genome	90.4%
Reads Mapped Confidently to Genome	82.8%
Reads Mapped Confidently to Intergenic Regions	16.0%
Reads Mapped Confidently to Intronic Regions	14.5%
Reads Mapped Confidently to Exonic Regions	52.3%
Reads Mapped Confidently to Transcriptome	47.4%
Reads Mapped Antisense to Gene	1.0%

Sample

Sample ID	cellranger-Kim6_004_MT_FL
Sample Description	
Chemistry	Single Cell 3' v3
Include introns	False
Reference Path	…ata-cellranger-Cfamiliaris-CanFam3.1
Transcriptome	Cfamiliaris-CanFam3.1-
Pipeline Version	cellranger-6.1.2

Figure 6. Single cell RNA-seq data of PIK3CA mutant canine hemangiosarcoma cells (C35).

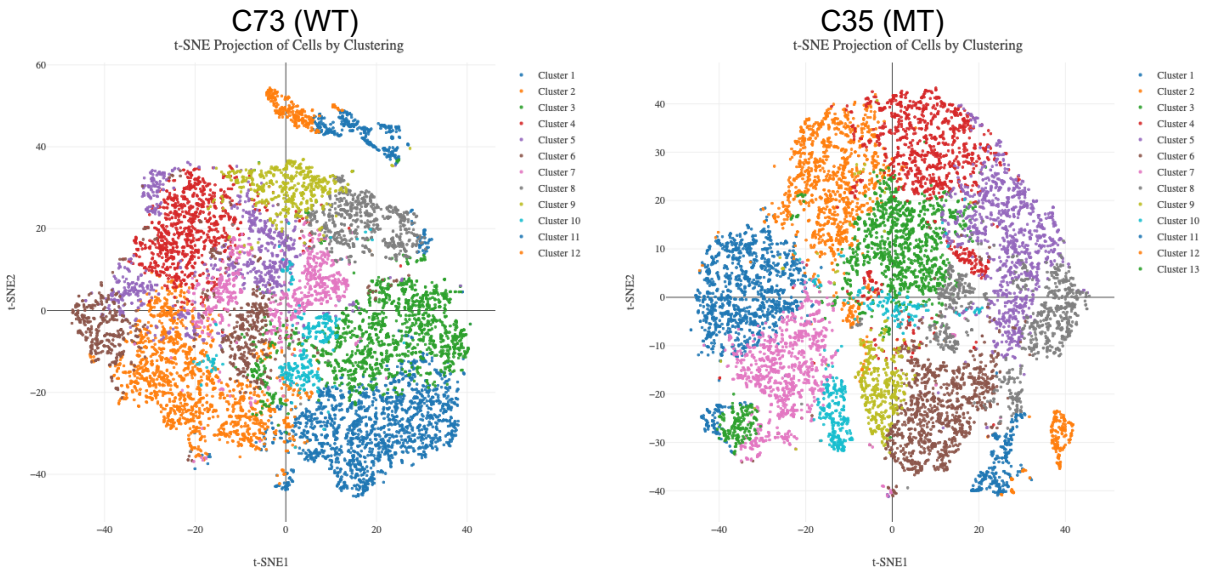


Figure 7. t-SNE plots depict single cell clusters of PIK3CA WT (left) and mutant (right) canine hemangiosarcoma cells.

Subtask 7: Initial bioinformatic analysis of sequenced data

- Partially completed (50% of completion) and ongoing

After completion of quality control process and initial analysis for mapping and read count measurement, we ran `cellranger aggr` tool to aggregate outputs from two samples (C73 and C35), normalizing them to the same sequencing depth and recomputing the gene expression data (i.e., feature-barcode matrices) on the combined data. Then, t-SNE analysis showed distinct gene expression patterns between WT and mutant cells at single cell level with a few of cell clusters overlapped between two datasets (**Figure 8**).

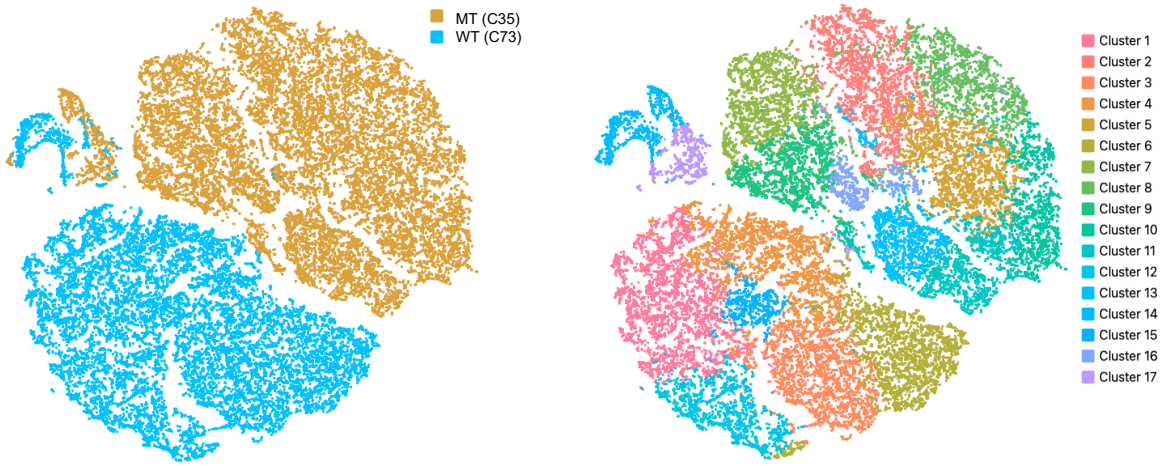


Figure 8. t-SNE plots present single cell clusters of PIK3CA WT and mutant canine hemangiosarcoma cells on combined data. Left plot shows two main clusters by mutation status and right plot present multiple cell clusters by single cell gene expression.

Further, we looked at expression of representative genes individually in the combined data (**Figure 9**). We found that *PECAM1* gene encoding endothelial cell marker, CD31 protein was dominantly expressed in mutant cell clusters, while no change in expression of *CDH2* gene, representative cell adhesion molecule between WT and mutant cells was noticeable. Intriguingly, more cells enriching pro-inflammatory cytokine genes such as *IL6* and CXCL8 (or *IL8)* were found in mutant cell clusters than WT cells. Further downstream analysis is currently being performed.

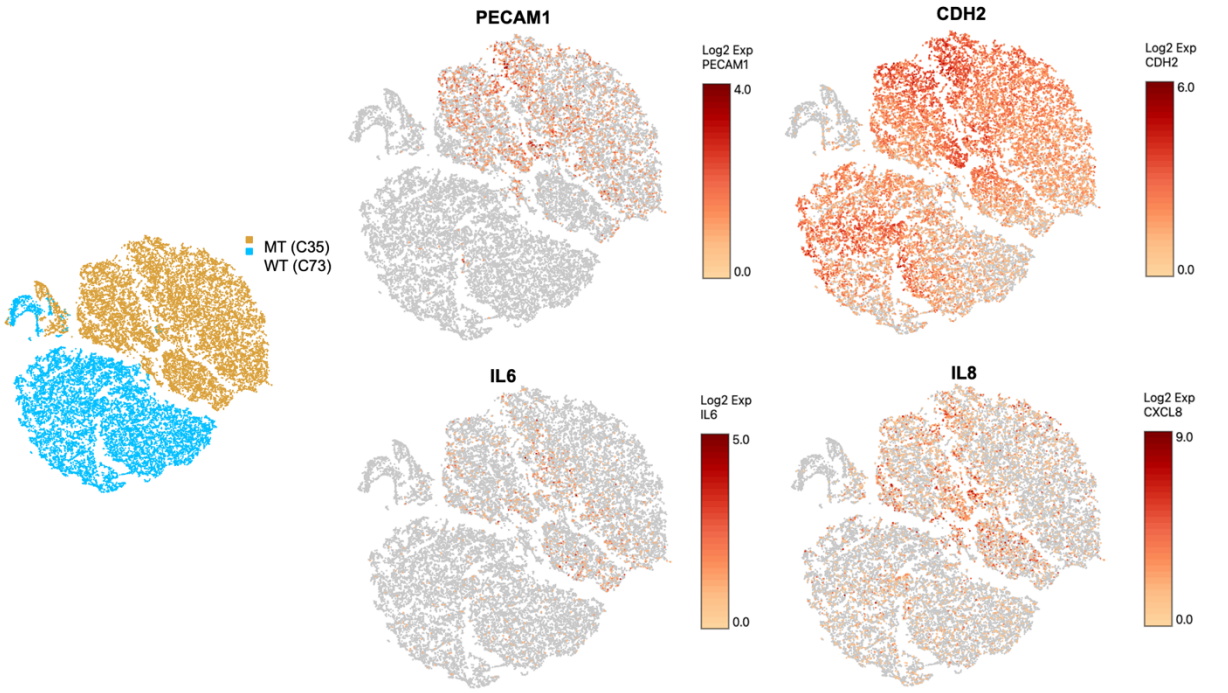


Figure 9. Expression of representative genes associated with endothelial and immune functions. Expression of *PECAM1*, *CDH2*, *IL6*, CXCL8 (or *IL8*) gene is depicted on t-SNE plots.

Subtask 8: Project meetings

- Partially completed (90% of completion) for this period due to PI's transition

Subtask 9: Career development for Dr. Kim (learning scATAC-Seq and scRNA-Seq)

- Partially completed (50% of completion) and ongoing

Dr. Kim coordinated a series of work for scRNA-seq data generation in this period by communicating with people for technical support of 10x Genomics and UMGC sequencing core. He learned overall process including single cell capture, barcoding, quality control, and sequencing libraries creation. He acquired programming skills to run pipelines for single cell read counts, gene expression value, and aggregation of multiple datasets. He visualized the single cell data t-SNE and UMAP data using Loupe Browser as shown above. He continues to run bioinformatics algorithms for downstream analysis including R packages such as seurat, monocle, and SC3 R packages as well as comparative bioinformatics [Ref #9].

[9] Dimitrov D, Türei D, Garrido-Rodriguez M, et al. Comparison of methods and resources for cell-cell communication inference from single-cell RNA-Seq data. Nat Commun. 2022 Jun 9;13(1):3224. doi: 10.1038/s41467-022-30755-0.

Specific Aim 2: To develop in vitro tumor models to define molecular mechanisms that regulate convergent oncogenic pathways in angiosarcomas

Major Task 1: Gene engineering and differentiation in iPSCs

Subtask 1: Generation of iPSCs and preparation of reagents; commercially available iPSC cell lines (iPS12-10 and BYS-0110)
- Partially completed (75% of completion)
We previously reported that iPSCs (iPS12-10 cell line) were used for optimization of iPSC maintenance and differentiation. Further optimization of protocols for BYS-0110 cell line is deferred during PI's transition period.

Subtask 2: Engineering gene mutation (*TP53, PIK3CA, TP53/PIK3CA*) in iPSCs
- Partially completed (40% of completion) and deferred
In previous report, we showed that *TP53* mutation was induced in iPSCs. Induction of additional mutations was deferred during PI's transition. Instead, we generated ATAC-seq and RNA-seq data from undifferentiated iPSC (day 0), mesoderm (day 2), and hemangioblast (day 3) as described above.

Subtask 3: Differentiation of hemangioblasts from engineered iPSCs
- Partially completed (50% of completion) and deferred
We optimized protocols for differentiation of cells from iPSCs. In this period, we focused on data analysis of chromatin accessibility, transcriptomic profile, and flow cytometric analysis generated from distinct differential stages of cells as shown above.

Subtask 4: Functional validation of engineered cells
- Partially completed (50% of completion) and deferred
We previously validated functional characteristics of cells differentiated from WT and *TP53* KO iPSCs. Functional validation of cells engineered with other mutations and from other iPSC line is deferred during PI's transition period.

Subtask 5: Cell line authentication and Mycoplasma screening (iPSCs-derived cells, HUVEC, fibroblasts, AS5, ISO-HAS)
- Partially completed (50% of completion) and deferred
We performed cell line authentication for AS5 and ISO-HAS cells as described above. We have not grown HUVEC and fibroblasts in this period, and we will continue to do cell authentication when cell culture work resumes in the PI's new lab.

Subtask 6: Project meetings
- Partially completed (90% of completion) for this period due to PI's transition

Subtask 7: Career development for Dr. Kim (acquisition of new experimental skills for iPSCs generation and genome engineering)

- Partially completed (50% of completion) and deferred
In previous report, Dr. Kim learned lab experiments of iPSC maintenance, hemangioblast differentiation, and CRISPR/Cas9-based genome engineering. He continues collaboration with Drs. Webber and Moriarity (Co-Is for this project), and he will advance the skills when his lab resumes in the University of Florida, regarding his transition.

Subtask 8: Career development for Dr. Kim (starting development of strategies to secure funding)
- Partially completed (75% of completion) and ongoing
Dr. Kim continued to develop strategies to secure funding during the project Year-2. In previous project Year-1, Dr. Kim developed a new collaborative project using iPSCs and genomic analysis with Dr. Daeha Joung, faculty in physics at the Virginia Commonwealth University. Their invited DoD full application (submitted to FY21 PRCRP) received good scores and recommended for funding as an alternate (i.e., "Recommended as an alternate"). It was stated that there are currently no additional funds available for this fiscal year of the program. Since this project received fundable scores, Drs. Kim and Joung applied for FY22 PRCRP Idea Development Award and their pre-application was invited for a full application. They are currently preparing the full grant application. Dr. Kim developed another project to develop experimental strategies for targeting PI3K pathway to recreate the tumor immune niche in angiosarcoma by developing iPSC-based cancer models. He submitted a grant application for the project as PI to FY22 DoD RCRP this year. Importantly, during this project year, Dr. Kim has moved to the University of Florida as a tenure track Assistant Professor through the UF AI Initiative. He is currently developing an independent research program for comparative oncology in his new institute with the UF support.

What opportunities for training and professional development has the project provided?
This DoD Career Development Award provided great training opportunities for Dr. Kim to advance his research program by obtaining new biotechniques and advancing his skills in stem cell biology, computational applications, and bioengineering. It also allowed him to present his research work at internationally known, the largest cancer research conference, 2022 AACR Annual Meeting. He attended the meeting in person and presented two posters related to this project. By expanding his scientific network, he initiated new collaborative work with Drs. Michael Wagner and Eleanor Chen (sarcoma oncologist and pathologist) in University of Washington Medicine. Recently, they (Dr. Kim as PI) submitted their new project to FY22 DoD RCRP. From these activities, Dr. Kim became an outstanding candidate for the UF AI Initiative faculty hire and received a competitive offer for his transition to a tenure-track faculty position.

How were the results disseminated to communities of interest?
Some results of the project were disseminated through presentations at 2022 AACR Annual Meeting held in New Orleans, NO on Apr 8 - 13, 2022: "Unveiling chromatin accessibility landscape and convergent oncogenic pathway in angiosarcoma models using induced pluripotent stem cells" and "Targeting PI3K pathway to reprogram the tumor immune niche in angiosarcoma."

What do you plan to do during the next reporting period to accomplish the goals?
In this project period, major change was PI's transition to the University of Florida. It was a great achievement in the PI's career development, and Dr. Kim is currently setting up his new lab to

establish comparative oncology program at the University of Florida College of Veterinary Medicine. His primary research focus remains the same in oncology and sarcoma biology, and he will resume research activities and tasks proposed for this DoD award. The University of Florida has established a high-end supercomputing infrastructure, called HiPerGator and they offer computational and bioinformatics support and training courses for researchers. Dr. Kim will use these research resources to accomplish the goals proposed for this project by completing tasks and following the milestones as addressed in Year-2 and -3 in the next period. Dr. Jaime Modiano as Career Guide are mentoring Dr. Kim remotely, and Drs. Webber and Moriarity (Co-Is) continue to work with PI in a collaborative manner.

4. IMPACT:

What was the impact on the development of the principal discipline(s) of the project?

In first project year, we reported that we have developed experimental strategies and optimized protocols for the tumor modeling of angiosarcoma using iPSCs. Human angiosarcomas are rare and they are heterogeneous, with a lack of understanding in genomic and epigenomic program. The project included goals to generate next-generation sequencing data for chromatin accessibility and transcriptomic profile. It required multiple types of tasks for computational programming and bioinformatics, and during this project year, the PI, Dr. Kim learned the skills and achieved the goals for this project as described above. The sequencing data have advanced our understanding of genomic and epigenomic landscape of angiosarcoma.

What was the impact on other disciplines?

New bioinformatics algorithms have been continuously developed that can be implemented for a variety of sequencing data. These methods are useful for studying rare diseases such as angiosarcoma and for better understanding of the pathogenesis and molecular mechanisms. However, the complex datasets must be interpretable in a biological context of cancer; in particular, there is less information in angiosarcoma as rare disease than other common types of cancers. Thus, it needs rigorous and sophisticated approaches with cross validation between multiple methods, when bioinformatic tools are applied. In this period, Dr. Kim has run multiple computational pipelines allowing him to learn those skills. Our ongoing work will develop new approaches to find biologically meaningful information using artificial intelligence and machine learning tools. It could help us unveil complex molecular features of angiosarcoma including genomic and epigenomic landscape as well as single cell gene expression. Furthermore, our computational approach with iPSC differentiation will impact on modeling strategies for other types of diseases.

What was the impact on technology transfer?

- Nothing to Report

What was the impact on society beyond science and technology?

- Nothing to Report

5. CHANGES/PROBLEMS:

Changes in approach and reasons for change

In the project Year-1, we had major changes due to the COVID-19 pandemic. Although we have made significant progress in the project Year-2, the revised activities in first year led to subsequent impact on tasks during this period and alternative experiments were performed as reported above. In addition, the PI, Dr. Kim has moved his primary organization from the University of Minnesota (UMN) to the University of Florida (UF): his appointment was terminated at the UMN on May 13th, 2022; the new appointment started at the UF on May 23rd, 2022. He started conversation with the DoD grant official on grant transfer process immediately after he received a new job offer from the UF in February 2022. He worked with grant officials from all three institutes, UMN, DoD, and UF for the grant transfer, which is in the middle of the process. Dr. Kim's transition and the grant transfer process changed and limited his research activities during this period, while his success on the transition to a more secure position (from non-tenure track to tenure track faculty) represents a great value of the DoD Career Development Award.

Actual or anticipated problems or delays and actions or plans to resolve them

As described above, ISO-HAS cells obtained from original research group failed to pass cell authentication screening. We exclude all the cells stored in our inventory, and as alternatives we used two canine hemangiosarcoma cell lines, DHSA-1426 and COSB. These cell lines have been comparable with human angiosarcoma cells showing malignant cellular characteristics with capacity of vascular formation *in vitro* as well as tumorigenicity in immunodeficient mice [Ref #10-13]. These canine cells were authenticated by the same method, and used for generation of ATAC-seq, RNA-seq, and single cell RNA-seq data as well as bioinformatics analysis in this report period. We were also able to induce *PIK3CA* H1047R mutations in these cells by CRISPR/Cas9 and will use them for mouse xenograft to evaluate their tumorigenicity and establish genomic and molecular landscape in the next project year. In addition, we deferred tasks including genetic engineering and hemangioblast differentiation in iPSCs due to the PI's transition. Despite the impact on timeline and delay due to the transition and grant transfer process, the University of Florida offers great opportunities for Dr. Kim's career development to develop outstanding independent research program. We believe that the PI's transition will continue to make considerable progress towards the stated goals and task of the project in the next phase.

[10] Kim JH, Graef AJ, Dickerson EB, et al., Pathobiology of Hemangiosarcoma in Dogs: Research Advances and Future Perspectives. Vet Sci. 2015 Nov 6;2(4):388-405. doi: 10.3390/vetsci2040388.

[11] Kim JH, Frantz AM, Anderson KL, et al., Interleukin-8 promotes canine hemangiosarcoma growth by regulating the tumor microenvironment. Exp Cell Res. 2014 Apr 15;323(1):155-164. doi: 10.1016/j.yexcr.2014.02.020.

[12] Gorden BH, Kim JH, Sarver AL, et al., Identification of three molecular and functional subtypes in canine hemangiosarcoma through gene expression profiling and progenitor cell characterization. Am J Pathol. 2014 Apr;184(4):985-995. doi: 10.1016/j.ajpath.2013.12.025.

[13] Tamburini BA, Phang TL, Fosmire SP, et al., Gene expression profiling identifies inflammation and angiogenesis as distinguishing features of canine hemangiosarcoma. BMC Cancer. 2010 Nov 9;10:619. doi: 10.1186/1471-2407-10-619.

Changes that had a significant impact on expenditures
- PI's transition changed timeline to perform experiments in his new place. There is no impact on significant budget amount.

Significant changes in use or care of human subjects, vertebrate animals, biohazards, and/or select agents
- Mouse experiments were postponed by PI's transition. Updated IBC and IACUC protocols are in process at PI's new institution, University of Florida. There is no significant change in use or care of animals; instead, we changed timeline to perform the animal experiment due to PI's transition.

Significant changes in use or care of human subjects
- Nothing to Report

Significant changes in use or care of vertebrate animals
- Nothing to Report

Significant changes in use of biohazards and/or select agents
- Nothing to Report

6. PRODUCTS:

- **Publications, conference papers, and presentations**
  - Sophia Wenthe, Kelsie Becklin, Brett Napiwocki, Emma Kozurek, Branden Moriarity, Beau R. Webber, Jong Hyuk Kim. *Unveiling chromatin accessibility landscape and convergent oncogenic pathway in angiosarcoma models using induced pluripotent stem cells*. 2022 AACR Annual Meeting, New Orleans, NO, Apr 8 -13, 2022.
  - Emma Kozurek, Zhiyan Silvia Liu, Hai Dang Nguyen, Jong Hyuk Kim. *Targeting PI3K pathway to reprogram the tumor immune niche in angiosarcoma*. 2022 AACR Annual Meeting, New Orleans, NO, Apr 8 -13, 2022.

- **Journal publications.**
- Nothing to Report

- **Books or other non-periodical, one-time publications.**
- Nothing to Report

- **Other publications, conference papers and presentations**.
- Nothing to Report

- **Website(s) or other Internet site(s)**
- Nothing to Report

- **Technologies or techniques**
- Nothing to Report

- **Inventions, patent applications, and/or licenses**

- Nothing to Report

- **Other Products**

- Nothing to Report

7. PARTICIPANTS & OTHER COLLABORATING ORGANIZATIONS

What individuals have worked on the project?

There is no significant change from original SOW.

Name:	*Jong Hyuk Kim*
Project Role:	*Principal Investigator*
Researcher Identifier (e.g. ORCID ID):	*0000-0002-1645-0036 (ORCID ID)*
Nearest person month worked:	*7.8*
Contribution to Project:	*Dr. Kim has performed work in the area of bioinforamtics and iPSC-derived hemangioblast differentiationg.*
Funding Support:	*DoD Career Development Award*

Has there been a change in the active other support of the PD/PI(s) or senior/key personnel since the last reporting period?

Completed grants and grant status changed during period of project Year-2 are listed below.

Completed Grants (07/01/2021 – 06/30/2022)

a) Title of project: A Novel Approach Combining Oncolytic Virotherapy and Dual Immune Checkpoint Blockade for Metastatic Osteosarcoma

b) Funding agency: Department of Defense

c) Goals of the project: *The goal of this project is to tilt the immune landscape of metastatic osteosarcoma from immunosuppression to immunoreactivity*

d) Specific aims/tasks:

Specific Aim 1: Optimize the dose and schedule for combination VSV-IFNβ-NIS and dual blockade of myeloid and T-cell exhaustion immune checkpoints using P-DIC in orthotopic, syngeneic (immunocompetent) mouse models of metastatic osteosarcoma.

Specific Aim 2: Define the immunological effects and anti-tumor mechanisms of action of VSV-IFNβ-NIS and dual blockade of myeloid and T-cell exhaustion immune checkpoints using P-DIC in orthotopic mouse models of metastatic osteosarcoma.

Specific Aim 3: Characterize the safety, efficacy, and immunomodulatory effects of combination therapy using VSV-IFNβ-NIS and P-DIC in dogs with naturally occurring metastatic osteosarcoma.

e) Estimated start and end date: 07/01/2020 - **06/30/22**

f) Level (%) of effort: 5% effort (0.6 person-calendar months)

g) Point of contact at the funding agency:

Emilee Senkevitch, PhD, PMP

Science Officer
Email: emilee.r.senkevitch.ctr@mail.mil
Phone:

h) Overlap: None

a) Title of project: Novel Immunotherapy for Sarcomas using Combination Oncolytic VSV and IL-18 Superkine
b) Funding agency: The V Foundation
c) Goals of the project: *The goal is to define safety and efficacy of immunotherapy that combines an oncolytic virus to initiate an immune response with an IL-18 analog to expand and sustain the response*
d) Specific aims/tasks:

Specific Aim 1: To establish an optimized protocol for combination VSV and DR-18 therapy in an orthotopic model of mouse osteosarcoma, and to define a safe and biologically active dose profile for DR-18 monotherapy in dogs.

Specific Aim 2: To characterize the safety, efficacy, and immunomodulatory effects of combination therapy using VSV and DR-18 in dogs with spontaneous osteosarcoma.

e) Estimated start and end date: 02/01/19 - **01/31/22**
f) Level (%) of effort: 2% effort (0.24 person-calendar months)
g) Point of contact at the funding agency:
Carole Wegner, PhD, HCLD
Senior VP, Research and Grants Administration
Email: cwegner@jimmyv.org
Phone:

h) Overlap: None

<u>No-Cost Extension request (07/01/2021 – 06/30/2022)</u>

a) Title of project: Reprogramming the Tumor Immune Niche in Canine Hemangiosarcoma
b) Funding agency: AKC Canine Health Foundation
c) Goals of the project: *The goal of this project is to determine if the molecular programs which create the tumor niche are reversible in canine hemangiosarcoma, and that PI3K/AKT/mTOR pathways regulate the expression of inflammatory cytokines that support niche conditioning.*
d) Specific aims/tasks:

Specific Aim 1: To determine if activation of PI3K signaling in hemangiosarcoma cells supports expansion and differentiation of hematopoietic progenitors.

Specific Aim 2: To examine if PI3K/AKT/mTOR pathways regulate the expression of inflammatory cytokines in hemangiosarcoma cells.

e) Estimated start and end date: 07/01/20 - 06/30/22 **(NCE to 03/31/2023 due to PI's transition)**
f) Level (%) of effort: 1% effort; cost-shared (0.12 person-calendar months)
g) Point of contact at the funding agency:
Andrea R. Fiumefreddo, MS
Director of Programs & Operations
Email: andrea.fiumefreddo@akcchf.org

Phone: 919-334-4022

h) Overlap: None

What other organizations were involved as partners?

- Nothing to Report

8. SPECIAL REPORTING REQUIREMENTS

COLLABORATIVE AWARDS:

- Not applicable

QUAD CHARTS:

- Not applicable

9. APPENDICES:

Abstracts presented at 2022 AACR Annual Meeting, New Orleans, NO, Apr 8 -13, 2022.

Unveiling chromatin accessibility landscape and convergent oncogenic pathway in angiosarcoma models using induced pluripotent stem cells

Short title: Modeling angiosarcoma using induced pluripotent stem cells

Sophia Wenthe[1], Kelsie Becklin[1], Brett Napiwocki[2], Emma Kozurek[3,4,5], Branden Moriarity[2,3], Beau R. Webber[2,3], Jong Hyuk Kim[3,4,5]

[1] Comparative and Molecular Biosciences (CMB) Graduate Program, University of Minnesota, St. Paul, MN, USA
[2] Pediatrics, Medical School, University of Minnesota, Minneapolis, MN, USA
[3] Masonic Cancer Center, University of Minnesota, Minneapolis, MN, USA
[4] Veterinary Clinical Sciences, College of Veterinary Medicine, University of Minnesota, St. Paul, MN, USA
[5] Animal Cancer Care and Research Program, College of Veterinary Medicine, University of Minnesota, St. Paul, MN, USA

Section: TB01-10 Other animal and cell models of cancer

Angiosarcoma is a rare soft tissue sarcoma that forms malignant vessels. Angiosarcomas are aggressive and highly metastatic, resulting in a poor prognosis. Half of patients present with metastatic or unresectable disease with a median overall survival of less than 6 months. Recurrent somatic mutations in *TP53* and genes involved in PI3K/AKT/mTOR pathway such as *PIK3CA* and *PIK3R1* are identified in angiosarcomas. However, angiosarcomas are genomically complex, and the oncogenic mechanisms are virtually unknown. Due to its rarity, establishment of experimental tumor models is an unmet need for angiosarcoma research. In this study, we used human induced pluripotent stem cells (iPSCs) to develop a novel, reliable model for angiosarcoma recapitulating the genomic complexity and the tumor immune landscape. Specifically, we induced *TP53* mutations in human iPSCs using CRISPR/Cas9 with validation of

p53 deficiency by gene sequencing and Western blotting. We then established protocols to differentiate genetically engineered iPSCs to mesoderm and subsequently bi-potential hemangioblasts, which are considered the putative cell-of-origin of angiosarcoma. We found that isogenic wild-type (WT) and *TP53* mutant iPSCs were capable of generating hemangioblasts, with no significant differences in morphology or growth patterns observed between WT and mutant cells. Putative iPSC-derived hemangioblasts were CD34+ by flow cytometry and contained primitive endothelial cells with the capacity to form tube-like structures in Matrigel. RNA-seq data libraries were generated to profile global gene expression in non-differentiated cells, mesodermal precursors, hemangioblasts, and endothelial cells derived from WT and p53 mutant iPSCs during differentiation over time (day 0, 2, 5, 8). Principal component analysis revealed that gene expression patterns were altered between WT and p53 mutant cells during differentiation, representing distinct gene signatures unique to each cell type. Intriguingly, transcriptomic alteration of p53 mutant iPSC-derived cells was more variable than that of WT. Our data also showed that p53 mutation induced dysregulation of genes associated with chromosome maintenance, extracellular matrix organization, hemostasis, and receptor tyrosine kinase signaling in iPSC-derived hemangioblasts when compared isogenic WT controls.
Our data highlights the role that mutant p53 plays in the induction of genomic instability and transcriptional programs that regulate hemogenic and endothelial function during differentiation. Additionally, ATAC-seq data were generated from WT and p53 mutant iPSC-derived cells to determine chromatin accessibility dynamics and identify key transcription factors that activate convergent vascular tumorigenic pathways. We are currently generating iPSC harboring co-mutations in *TP53* and *PIK3CA* in order to determine the phenotype and tumorigenic capacity of engineered iPSC-derived hemangioblasts in xenograft models.

Targeting PI3K Pathway to Reprogram the Tumor Immune Niche in Angiosarcoma

Short title: PI3K pathway in angiosarcoma

Emma Kozurek[1,2,3], Zhiyan Silvia Liu,[4] Hai Dang Nguyen,[3,4] Jong Hyuk Kim[1,2,3]

[1]*Veterinary Clinical Sciences, College of Veterinary Medicine, University of Minnesota*
[2]*Animal Cancer Care and Research Program, College of Veterinary Medicine, University of Minnesota*
[3]*Masonic Cancer Center, University of Minnesota*
[4]*Department of Pharmacology, Medical School, University of Minnesota, Minneapolis, MN*

Section: ET06-06 PI3K/AKT inhibitors

Angiosarcoma is an aggressive, albeit rare, cancer in humans. Angiosarcomas are vascular malignancies that can occur anywhere in the body, and their metastatic propensity is high. The vascular sarcomas consist of disorganized, malignant vessel-forming cells. Recurrent mutations in the *PIK3CA* gene have been identified in angiosarcomas, and activation of the PI3K pathway appears to establish angiogenic programs. Naturally occurring hemangiosarcoma in dogs shares clinical and pathological features with human angiosarcoma. Our previous work has revealed that canine hemangiosarcomas and human angiosarcomas share mutational and transcriptional signatures that activate convergent PI3K signaling pathways. Appropriately powered canine

studies take a comparative genomics approach, informing our research into human angiosarcomas. The cause of most sporadic angiosarcomas is unknown. Moreover, no therapeutic targets have been identified to improve outcomes. This study hypothesizes that the activation of *PIK3CA* mutations contributes to the molecular programs that modulate the immune niche in angiosarcoma. We first induced *PIK3CA* hotspot mutations (H1047R) in canine hemangiosarcoma cell line DHSA-1426 using CRISPR/Cas9. Our ongoing work is to determine if *PIK3CA* mutant tumor cells have cell-autonomous capacity to govern hematopoietic progenitors and immune cells, potentially establishing the tumor immune niche. It also includes establishing single-cell gene expression profiles to identify distinct cell clusters between *PIK3CA* mutant and WT cells, along with chromatin accessibility landscape using ATAC-seq. Tumorigenic capacity and immune regulatory mechanisms of *PIK3CA* mutations will be determined in angiosarcoma xenograft models.