Characterization of the Genetic Landscape of Feline Oral Squamous Cell Carcinoma

---

A Thesis presented to

the Faculty of the Graduate School

at the University of Missouri – Columbia

In Partial Fulfilment

Of the Requirements for the Degree

Master of Science

By

Alana R. Rodney

Dr. Wes Warren, Thesis Advisor

May 2022

The undersigned, appointed by the Dean of the Graduate School, have examined the thesis entitled

CHARACTERIZATION OF THE GENETIC LANDSCAPE OF FELINE ORAL SQUAMOUS CELL CARCINOMA

Presented by Alana R. Rodney
A candidate for the degree MASTER OF SCIENCE
We hereby certify that in our opinion it is worthy of acceptance

---

Wes Warren, PhD – Thesis Advisor

Leslie A. Lyons, PhD

Owen Skinner, BVSc, DECVS, DACVS-SA, MRCVS

Chris Elsik, PhD

ACKNOWLEDGEMENTS

First, I would like to thank Dr. Warren for his great mentorship over the past several years. Learning from you has been a great honor, and I am excited to take these skills to complete a PhD this fall. I would also like to thank my committee, Dr. Lyons whose mentorship, support, and encouragement has been invaluable, especially during the last months of preparation. I would also like to thank Dr. Owen Skinner for his help through the writing process and consistent constructive and helpful feedback. I very much appreciated Dr. Elsik's calm cool and collected teaching style, as script writing and bioinformatic can be very stressful, and it was great to have a soft reassuring pace to land during frustrating times. Finally, I want to thank my lab mates, friends, and family, because without their support, I would not be where I am today.

TABLE OF CONTENTS

LIST OF TABLES

Table

LIST OF FIGURES

FIGURE

LIST OF ABBREVIATIONS

WGS: Whole Genome Sequencing

WES: Whole Exome Sequencing

FeLV: Feline Leukemia Virus

FIV: Feline Immunotherapy Virus

STS: Soft Tissue Sarcoma

FeSV: Feline Sarcoma Virus

FMC: Feline Mammary Cancer

ISS: Injection Site Sarcoma spelling?

FOSCC: Feline Oral Squamous Cell Carcinoma

HNSCC: Human Head and Neck Squamous Cell Carcinoma

PKD: Polycystic Kidney Disease

ADPKD: Autosomal Dominant Polycystic Kidney Disease

TMB: Tumor Mutational Burden

ABSTRACT

Over 94 million domestic cats are susceptible to cancers and other common and rare diseases. While cancer treatment in cats increasingly mirrors that available in humans, treatment failures are more frequent. There are no FDA or USDA approved cancer drugs for cats and a paucity of cancer treatments beyond surgery, radiotherapy, and cytotoxic chemotherapy indicate an urgent need to define the molecular properties of aggressive feline cancers including common cancers such as soft tissue sarcoma, mammary carcinoma, lymphoma, and feline oral squamous cell carcinoma. Whole exome sequencing (WES) is a proven strategy to study these disease-causing variants in humans and other organisms. To examine the effectiveness of WES in the study of feline cancer and Mendelian diseases, whole exome sequencing was conducted on 41 cats with known and unknown genetic diseases and traits, of which ten cats had matching whole genome sequence (WGS) data available, used to validate WES performance. Within the 41 cats, we identified 31 previously known causal variants and discovered new gene candidate variants, including novel missense variance for polycystic kidney disease and atrichia in the Peterbald cat. The WES data was sufficient to identify novel gene candidate alleles for diseases and traits in a feline model. Feline oral squamous cell carcinoma (FOSCC) is a cancer of the squamous cell lining in the oral cavity and represents up to 80% of all oral cancers in cats, with a low one-year survival rate of <10%. The cancer pathology associated with feline oral squamous cell carcinoma is similar to human head and neck squamous cell carcinoma (HNSCC), which accounts for 90% of oral cancers in humans. FOSCC may present as a potential model to study HNSCC due to spontaneous formation, similar genetic landscape, pathology, and survival rates. We have generated single nucleotide variant calls using GATK-Mutect2 on six cats with FOSCC and have fully annotated and identified driver genes in common with HNSCC. Due to low sample size, a

larger cohort is needed to draw stronger association of disease-causing traits. Our results show some overlap in the genetic landscape of both cancers, with five samples have mutations in p53, a common mutation in HNSCC, and two samples having four genes in common with HNSCC each. Several samples with mutations in p53 and mutations in genes implicated in HNSCC suggests that the domestic cat could be a viable model for HNSCC.

CHAPTER 1

INTRODUCTION

1. Introduction

Today, at least 94 million cats reside in US households that account for 7.4 billion USD in veterinary services[1,2]. While companion animals, hereafter referred to as cats and dogs, share environments, disease predispositions, risk factors and occurrence rates are distinct in differing breeds, let alone species[3]. From a cancer perspective, there is a rather different profile of disease in cats than dogs. A major complication to finding the origins of substantial differences are companion animals are less likely to show symptoms until later stages, and without the ability to articulate symptoms, leading to challenges in identifying the earlier stages of disease. Cats and dogs both have a shorter average life span than humans, and a smaller window to study disease progression presents an opportunity for identification of potential therapies that may be successful in humans[4]. Dogs present with cancers of various types, including a higher incidence than cats but no confirmed explanations why[3,5]. For example, osteosarcoma is found at a rate of 72 to 126 cases reported per 10,000 dogs depending on the breed, compared to 4.9 cases per 100,000 cats[6,7].

The varying history between cat and dog domestication, with dogs having lower rates of heterozygosity due to strict breeding practices, may offer some genetic or epigenetic explanations for these large differences in cancer rates. Dog and cat domestication differ greatly, with dog domestication events estimated around 15,000 to 30,000 years ago depending on the study[8], compared with cat domestication where available evidence estimated the domestication period up to 9,000 years ago[9]. Genetic data shows dog domestication began in middle eastern

and east Asian grey wolves, with two periods of intense selection[10]. The first selection period began approximately 10,000 years ago, where dogs were selected for docility and better interactions with humans leading to ancient breeds, like the chow chow and New Guinea singing dog that are significantly divergent from modern breeds[10]. Modern dog breeds are a result of a second intense selection period in the Victorian era, 1830 - 1900, where dogs began to be bred under the direction of breed clubs[10]. These practices resulted in a wide range of variations with 195 AKC registered breeds, and over 350 breeds known worldwide [11]. Breeds vary in size, skeletal and cranial proportion, and characteristics including sight, smell, and tendencies for herding, swimming, aggression and laziness[8] . This wide divergence of dog breed phenotypic diversity is largely due to the discrete fixation of DNA variants that have a large effect in individual lineages, which are then crossed into other groups, followed by the selection of this trait in the F2 generation[8,10] . This intense artificial selection causes the haplotype where the variants are fixed to rise in frequency that in combination with breeding of related individuals, leads to reduced levels of genome-wide heterozygosity. In dogs, this inbreeding is likely an important contributor to higher rates of specific diseases in some breeds versus others. For example, a study evaluating degenerative myopathy in Collies found a heterozygous carrier rate of 27.6% for a mutation in SOD1, which was implicated in disease progression in the small (less than 100) populations of collies in Japan. This study highlights the risk of inbreeding on the occurrence of degenerative myopathy[12]. The genetic origins of these higher breed-specific incidence rates, however, is understudied and often unknown.

The genetic history of the domestic cat is dramatically different to the dog. Cats were suggested to cohabitate with humans in Cyprus around 9,000 years ago. Domestication is thought to be mutualistic between humans and cats, where cats were first adopted to control rodents eating the

grains stored within agricultural villages near the Mediterranean fertile crescent [9]. Today, domestic cats are considered a subspecies of wildcats due to their morphological and genetic similarities, and they are often visibly indistinguishable from wild cats, especially those cats with a tabby coat[13]. Cats have fewer than 40 true breeds appearing in only the last 150 years, compared to over 400 recognized dog breeds, and most domestic cats are categorized as random bred, leading to lower rates of inbreeding in the general cat population[14-17]. Further differing from the dog, coat color, length, and texture trait aesthetics are the main drivers of breed formation[16-18]. While the establishment of cat breeds began with random-bred cats throughout the world, allowing the outcrossing of many breeds, where in dogs mating of close relatives leads to an increased risk of genetic disorders and diseases[19]. It is all these genetic features of feline genome evolution, i.e., increased rates of heterozygosity and a less stringent genetic bottleneck, that has led to hypotheses that domestic cats can offer a contrasting model of why cat cancer types differ substantially from the dog in many aspects.

Tumors in companion animals are staged based on tumor specific protocols and are classified from stages I to IV as in humans. Staging informs the extent of cancer locally, distally and in the draining lymph nodes to determine the course of treatment. Risk factors for many cancers in companion animals include exposure to tobacco smoke, tinned tuna, flea collars, infection of the feline leukemia virus, and reoccurring inflammation at sites of trauma due to injections and microchipping[20-22]. Many cancers in cats are very aggressive and treatments available include surgery, radiation, chemotherapy, and ultimately palliative care. While various treatments can be effective with early intervention, new therapies are desperately needed to improve prognosis.

Companion animals and humans suffer from many of the same ailments, with over 70 genes [23] shown in cats to contain single and multiple DNA variants in common with humans contributing

to disease causation such as cardiomyopathy and polycystic kidney disease[24,25]. Genomic medicine is currently being practiced in human health where one sequences the genome, coding regions, or select genes of the patient to identify variants that cause or contribute to the morbidity of a disease in order to improve treatment options. This process is rapidly being implemented in cancer therapeutic decisions with druggable gene mutations as the outcome but not in companion animals. Recent whole genome sequencing (WGS) and whole exome sequencing (WES) cat studies have shown many more causative or associated disease variants await discovery[26]. By uniquely identifying a homozygous missense variant in *NPC1,* a gene responsible for causing Niemann-Pick disease type C1, a cat with an undiagnosed neurological disorder can now inform the veterinarians treatment strategy[27]. In non-disease cases, for instance breeds with rare phenotypes like dwarfism, WGS was able to identify a novel structural variant in the *UGDH* gene associated with cats displaying dwarfism[27,28]. With actionable genetic information available, veterinarians can be aware of disease risk and proactively treat patients to prevent severe symptoms. However, the practice of genomic medicine in veterinary clinics is very limited today. In cancer, this is due to the scarcity of WGS or WES studies that ascertain somatic mutations linked to tumor progression. Our knowledge of variants associated with domestic cat cancers should substantially increase, thus, genomic medicine will be feasible for cats in the near future with cancer or unknown aliments[27]. A survey of the literature follows for the most common types of cat cancer with an emphasis on what is known for genetic causality.

From this point forward, a discussion of the most common cat cancer types with contrasts to dog, where appropriate, is presented. While feline cancer is rare, with one study citing cats having an incidence rate of 63 per 100,000 cats per year compared to dog with about 143 per 100,000 dogs per year[29] both compared to humans, which have a 1 in 3 chance in developing cancer[30]. Cancers

such as lymphoma, fibrosarcoma and mammary cancers are commonly seen while tumors of the brain, lung, and liver are much less frequent, but all often result in poor prognosis due to their late stage of detection. The possible risk factors, prognosis, and when known molecular causes for these feline cancers, including lymphoma, fibrosarcoma, oral squamous cell carcinoma, and pulmonary neoplasia are briefly reviewed **(Figure 1.1).**

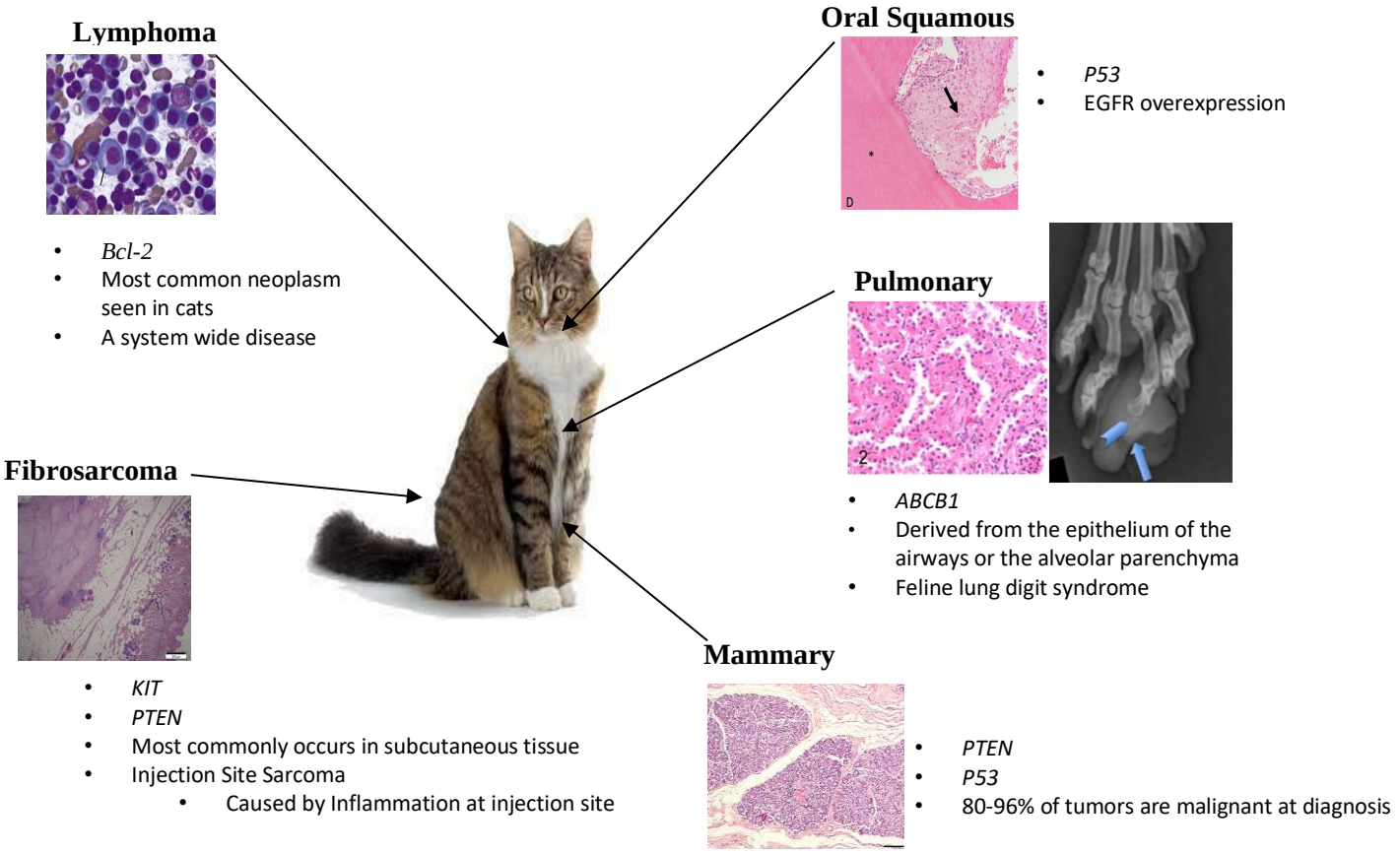


Figure 1.1 Summary of the most common types of cancer in cats with candidate driver genes of each cancer

278

279

280

281

282

283

284

285

286

287

CHAPTER 2

LITERATURE REVIEW

2.1 Feline Cancer Molecular Landscape

While cat cancers are less frequent compared to dogs, there are several that are often seen in cats. The most common by far is lymphoma followed by soft tissue sarcoma, which includes injection site sarcoma. Mammary cancer accounts for 12% of all cats' cancers, followed by oral squamous cell carcinoma which is the most common oral cancer diagnosed in cats, with pulmonary carcinoma diagnosed less often, but often resulting in metastasis[31]. While several reports have advocated for the overlooked potential of studying the only other spontaneous naturally occurring model of cancer outside human, the dog[32,33] lost in this discussion are the interesting genetic features of domestic cat cancer even with its lower rate of incidence. Improved genome references, better gene annotation, and greater WGS sampling of domestic cats for variant ascertainment, have established resources for more transcriptomic and genomic investigations of cat cancer. Addressing this, in turn, will facilitate the finding of shared and novel molecular drivers of cat cancer compared to the human and dog.

2.2 Lymphoma.

A malignancy of lymphocytes, various versions of lymphoma are the most common neoplasms seen in cats[34]. Lymphoma is typically considered a system-wide disease, as these cells travel through the entire body through the lymphatic system. In the cat, lymphoma can arise at any site but generally presentations are alimentary/gastrointestinal, peripheral nodal, and extra nodal, and most frequently is seen in the lymph nodes, spleen and bone marrow[34,35] with B, T, or NK cells forming the neoplastic population[36]. Feline leukemia virus (FeLV) is the most common cause of

lymphoma in cats, however after the introduction of the FeLV vaccine in the 1980's in combination with the loss of infected animals, there has been a decline in FeLV-associated lymphoma. The median age of cats diagnosed with lymphoma has increased to 12 years from 3 to 5 years of age as a result[37]. There is an association between feline immunodeficiency virus (FIV) infection and increased incidence of lymphoma, resulting in a five-fold increase of developing feline lymphoma[38], suggesting that FIV plays an indirect role in tumorigenesis causing chronic dysregulation of the immune system and activation of oncogenic pathways. Siamese and Oriental breeds are reported to be at higher risk, with one study reporting a 1.5:1 male to female ratio and another study finding no association with neutering status[37,34]. Clinically, identification of lymphoma includes a complete blood count (CBC) with differential cell and platelet count, serum biochemistry profile, urinalysis and a retroviral screen[35]. Tumors are graded and then staged to evaluate the extent of the disease. Tumors that present as low grade have strong remission rates with conservative treatment protocols including oral chlorambucil and prednisolone resulting in average survival rates of 1.5 to 3 years[39]. Most feline lymphomas present as medium and high-grade tumors at any anatomical site that require more aggressive multiagent combination therapies. CHOP therapies, consisting of a combination of cyclophosphamide (C), hydroxydaunorubicin (H), oncovin (O) and prednisone (P)[35,20] , are most successful in treating medium and high-grade lymphoma. CHOP is a modified treatment protocol based on human protocols, and only have complete remission rates of 50 to 60%, with an even lower survival rate of less than one year. Radiation treatment has been demonstrated as an alternative effective therapy because lymphatic cells are sensitive to radiation[34]. Lastly, surgery is typically not an adequate treatment for this disease but is most often used to achieve a diagnosis or relieve gastrointestinal obstruction but has no clear effect on survival[21,40].

With the use of genomic medicine approaches, human patient survival times for blood borne cancers have increased and that has already accelerated investigations of all types of lymphoma maladies among companion animals, albeit at a much slower pace in cats. Interestingly, we can perhaps also study heritable risk since there is a lymphoma predisposition observed in the oriental cat breeds [41]. Changes in oncogenic pathways, epigenetic changes and signal transduction alterations found in human are suspected to be at work in the domestic cat but again unproven to date[35]. One cat study saw increased *Bcl-2* expression which controls cellular proliferation and cell cycle apoptosis. In this study, lymphoma cell lines had higher levels of *Bcl-2* compared to normal peripheral mononuclear blood cells which had lower levels of *Bcl-2* expression, suggesting that *Bcl-2* expression may be useful in the differential diagnosis of feline tumors[42].

Since very few studies explore the underlying driver genes of feline lymphoma or their accompanying transcriptomes cat lymphoma outcomes haven't changed much. In the past, gene microarray techniques were utilized to broadly understand the role of chromosomal aberrations and gene expression changes associated with lymphoma[35] but no treatment changes were implemented as a result. The adoption of fast developing genomic methodologies promises a brighter therapeutic outcome for feline lymphoma.

2.3 Soft Tissue Sarcoma

Soft tissue sarcomas (STS) are locally aggressive cancers that arise from mesenchymal tissues and can arise in connective tissues including the muscle, adipose neurovascular, fascial, and fibrous tissues[41]. This type of cancer is relatively common with an incidence rate of 17 per 100,000 cats. Soft tissue sarcoma can arise at any anatomical location, but most commonly occurs in subcutaneous tissue and is the second most prevalent skin tumor in cats[41]. In cats, STS

has three known subtypes: spontaneous formation of fibrosarcoma, feline sarcoma virus and injection site sarcoma (ISS). Fibrosarcomas can be rapidly growing and present as multiple cutaneous or subcutaneous nodules. These nodules are normally locally invasive and can metastasize to the lungs and other sites[43]. The feline sarcoma virus (FeSV) is a recombinant version of the FeLV, where the recombination of the FeLV genome with cellular oncogenes of the infected cat causes uncontrolled cell proliferation[43,44]. Through this recombination, FeSV can acquire one of the multiple oncogenes including *FES, FMS,* or *FGR*[43]. Oncogenes associated with FeSV were discovered by transforming fibroblasts and producing a fibrosarcoma that is multicentric in younger cats[43,44]. Older studies show that only about 2% of the fibrosarcomas found in cats are virally induced and as cases of FeLV have decreased so have cases of FeSV[43-45].

The most common histological type of fibrosarcoma is ISS. While ISS is included as a type of soft tissue sarcoma, they are histologically heterogeneous and have distinctive biological behavior, are more aggressive, and appear in younger cats when compared to non-injection site fibrosarcoma[46,47]. ISS is often thought to be the result of the proliferation of fibroblasts and myofibroblasts at the sites of chronic inflammation caused by a sustained immune response induced by the injected material[4,46,48]. It is also characterized as the proliferation of atypical spindle cells and a variable amount of multinucleated giant cells, which are not generally seen in non-injection site feline fibrosarcomas[4]. In this cancer, cats develop a subcutaneous inflammatory response at the site of injections and the tumor generally develops within the first 3 years of vaccination[4,46]. This inflammation has been associated with the use of vaccines that contain aluminum-based adjuvants, and while the increasing use of non-adjuvant vaccines has caused a reduction of ISS, it remains a concern in feline veterinary medicine[2,49]. Vaccine-

associated sarcomas have been reported to have elevated mitotic rates, with higher-grade classification, less differentiated, and contained cells with more variable size and shape compared to non-vaccine site sarcomas[49]. Other causes of ISS include microchipping, the use of long-acting injectable antibiotics, and injectable glucocorticoids. ISS typically presents as a subcutaneous mass at the site of injection. Cats can also present with decreased appetite, and difficulty breathing if the cancer has metastasized, most frequently to the lungs[47,48]. Staging is an important step in ISS treatment plans since this tumor type is highly aggressive and treatment generally includes tumor resection depending on the stage and grade of the tumor[48,47]. Of note, there are marked institutional variations here with some centers that routinely involve radiation treatment in treatment while others perform radical excision as a locally curative intent therapy without RT (as long as margins are acceptable) [50].

The study of molecular causes associated with ISS have been limited to a few genes and typically rely on the understanding of the analogous injection site sarcoma cancer type in human for the most likely driver gene candidates. A cohort of ISS cats was found to harbor homozygous deletions in *PTEN*, amplification of *KIT*, and an overall DNA copy number imbalance that was correlated to a more aggressive tumor behavior[2]. Another study found that increased expression of *MMP-9*, *MMP-2*, and *TIMP-2* in all ISS tumors investigated was suggestive of tumor aggressiveness[48,51,52]. With the aggressive nature and often unknown specific causes of ISS, early detection and surgery are still the first-line defenses against this cancer type with druggable genes not on the clinical horizon[53-56].

Excitement and perhaps too much anticipation is now placed on immunotherapy opportunities. Yet, one study showed that viruses expressing interleukin-2 are not very effective, reducing the tumor reoccurrence to 28% compared to 52% at 12 months[53]. Combinatorial strategies used to

treat human head and neck cancers (HNSCC) and melanomas, such as targeted type I interferon pathway inhibitors and oncolytic viruses could offer better tumor control in feline sarcomas[57], with one feline study showing reoccurrence rates dropping from 61% to 28%[58]. Inflammation at injection sites causes mutations in many oncogenic pathways[53], potentially leading to injection site sarcomas; however, improved vaccine technology, early detection, and better adjuvant therapy, seem to be reducing ISS incidence.

2.4 Feline Mammary Cancer.

Feline mammary cancer (FMC) is the third most diagnosed feline neoplasm. Approximately 80 - 96% of FMCs are malignant with multiple tumors being common at diagnosis[59]. FMC carries an incidence rate of 24.4 cats per 100,000 female cats, and accounts for 12% of all tumors in cats regardless of sex[22]. A major risk factor for FMC is age, with risk of diagnosis becoming significantly increased at 7 to 9 years[60], and mean age of diagnosis being 10 to 12 years in age. Breed is another major risk factor with studies suggesting short-haired cats have a higher risk. Siamese cats are at twice the risk of developing FMC and are diagnosed at a younger age[61,22]. Another major risk factor is hormonal exposure, with gonadally intact cats having a 7-fold higher risk of developing FMC than spayed cats; a is similarly effect is observed in dogs[60]. Cats exposed to regular progesterone have an increased risk of FMC, so ovariohysterectomy, a tissue source of this steroid, is protective against mammary tumor development[62]. It is a highly aggressive cancer, with locally infiltrative and metastatic features, as the primary tumor metastasizes frequently to regional lymph nodes and lungs[59]. Fortunately during routine veterinary visits FMC based on the palpation of firm nodular masses on the mammary gland[63]. Thoracic radiographs, abdominal ultrasound, and fine-needle aspiration are then often used assess the extent of FMC[63]. Just as in human breast cancer, most feline tumors are classified as

moderately or poorly differentiated[63]. In today's practice, many veterinarians have adopted the Elston and Ellis histological grading system that is considered the “gold standard” grading system for human breast cancer[64]. Feline mammary histological grading is based on cellular differentiation and degree of tubule formation; nuclear pleomorphism; and mitotic frequency[63]. Many factors affect the prognosis of FMC including grade, mitotic count, and disease stage. Tumors smaller than 2 cm have seen a survival a mean of 54 months, tumors 2-3 cm average of 24 months and 3-6cm have an average survival of only 6 months[60,63]. Once again, the first line of FMC treatment is surgical excision, with studies showing bilateral mastectomy improves survival time[65].

FMC is most likely to be hormone receptor-negative and patients presenting with triple-negative mammary carcinoma, an aggressive form of cancer lacking secretion of progesterone, human epidermal growth factor and progesterone receptors[66], had the lowest survival rates[67,68]. Multiple genes with specific somatic mutations have been associated with FMC. Elevated expression of the oncogene *HER2*, which stimulates downstream activation of the AKT pathway [69] that promotes growth factor-mediated cell growth, proliferation, migration, and survival is highly correlated with malignancy and tumor differentation[69,70]. Loss of the *PTEN* gene, a negative regulator of the AKT pathway, has been shown in 76% of FMC cases and in higher-grade tumors predicting a poor prognosis[70]. Recurrent deletions among B2 and E3 chromosomes detected in higher-grade FMC tumors exhibit suggested these genomic imbalances contribute to poor outcome[71,72]. Copy number gain of *EBR2* was greater than 3-fold in high-grade compared to intermediate-grade mammary gland cat tumors that alludes to the important role this gene plays as part of this chromosome instability correlation with a worse outcome[2,71]. There is evidence that molecular markers cyclin A, *p53, RON, VEGF,* and *COX* can be predictors of outcome for

FMC[63]. Studies have shown *VEGFR-3* highly to moderately expressed in all FMC. All these studies demonstrate much more FMC sample sequencing will be needed to resolve the somatic mutations with the highest occurrence by gene.

As stated earlier for other cancer types immunotherapy is also being explored for FMC treatment. Recent studies suggest immunotherapy strategies targeting programmed cell death 1 (*PD-1*) and programmed cell death ligand 1 (*PD-L1*) genes, FMC may be a new promising line of treatment[73,63]. PD-L1 has been implicated in a long list of human cancers, including breast, lung, and metastatic melanoma, and promotes immune suppression and tumor escape[73,74]. Nascimento *et al* discovered elevated levels of PD-1 and PD-L1 in the serum of cats with FMC, and with a newly FDA approved monoclonal antibody that blocks PD-L1 binding better therapeutic outcomes are hopefully possible for FMC[73].

2.5 Oral Squamous Cell Carcinoma.

Feline oral squamous cell carcinoma (FOSCC) is a cancer of the squamous cell lining in the oral and oropharyngeal cavity of the cat and may spread into deeper tissues[75]. These tumors are commonly found in the gingiva, tongue, and sublingual region[76]. FOSCC is the fourth most common cancer in cats and represents 70 - 80% of all oral cancers and shares many risk factors and molecular markers to human head and neck squamous cell carcinoma (HNSCC). These similarities may present an opportunity to test evolving therapies in HNSCC[4,20] but also offer comparative oncological models of their molecular circuits. These tumors are usually locally invasive, however can invade the local bone tissue[77]. FOSCC uncommonly metastasize to distant locations but one study found 35% of cats had metastasis to local lymph nodes[78]. FOSCC is typically a deadly disease for cats with median survival post-diagnosis only reaching a few months, and a one-year survival rate of less than 10%[79]. Cats presenting with FOSCC are an

471 average age of 12.5 - 13 years, with no sex or breed disposition[76,80]. Signs of this disease include
472 oral pain, difficulty swallowing, excessive salivation, anorexia, and loss of teeth[20]. In cats, risks
473 associated with FOSCC include exposure to tobacco smoke as well as flea collars and feeding
474 some canned foods[81]. Human papillomavirus, a known cause of HNSCC, has been associated
475 with more negative outcomes in humans; however, current studies show FOSCC is more closely
476 related to HPV-negative HNSCC[82]. Molecular markers thus far include abnormal *p53*, and *CK2*
477 expression as well as genes implicated in angiogenesis, e.g. *LOX* and *COX*[83]. *EGFR*
478 overexpression is seen in FOSCC and humans, and has been associated with proliferation and
479 migrations of squamous carcinoma cells to other sites[84,85]. Recent studies show that *CD147*, a
480 surface cluster protein found in HNSCC tumors is associated with poor prognosis[80]. Since
481 FOSCC *CD147* associated expression has been observed perhaps its manipulation in the tumor is
482 a therapeutic target[76]. Treatment of FOSCC is first surgical removal from the oral cavity, if
483 feasible. In contrast to some other cat cancer types, FOSCC surgery is often followed by high
484 recurrence rates with estimates ranging from 15.4 to 38%[86-88]. Other standards of treatment
485 following surgery include, radiation, and chemotherapy; but again poor outcomes are the
486 result[89,90]. Starting with FOSCC cell lines, treatment with actinomycin D, dinaciclib,
487 flavopiridol, and methotrexate shows promise yet this drug mixture has not been tested in
488 clinical settings[77]. While current treatments offer unacceptable FOSCC control, new potential
489 strategies that include EGFR inhibitors, CK2 inhibitors, COX/LOC inhibitors, or methods to
490 reverse hypoxia[83] can hopefully avoid euthanasia that remains the most common veterinarian
491 recommendation in FOSCC cases.

492

2.6 Pulmonary Carcinoma.

Primary pulmonary carcinoma is rare a type of lung cancer in the cat and most commonly is derived from the epithelium of the airways or the alveolar parenchyma[91]. This lung cancer is found in about 2.2 per 100,000 cats, almost all tumors are malignant[92] and the most common type of primary lung tumors are adenocarcinomas (60 - 80%), with sarcomas, squamous cell carcinomas, and adenomas less common[93]. Primary lung neoplasia is less common than metastatic tumors that spread to the lung from other primary cancers, most commonly from the breast [91]. Pulmonary carcinoma can also present as feline lung-digit syndrome, an atypical pattern of metastasis where the primary lung tumor cells travel to the digits of the cat, particularly weight bearing regions. This presentation is associated with a poor prognosis. Metastasis to the digits may be due, in part, to high digital blood flow to compensate for heat loss[93]. Risk factors of pulmonary carcinoma include exposure to secondhand smoke and as in other cat cancers increased age but with no higher prevalence among breeds or sex[94]. A recent study has concluded that in areas with high radon exposure, companion animals have a 2-fold higher incidence of primary pulmonary neoplasia[95]. Signs of this cancer include coughing, weight loss, lethargy, lameness, and respiratory changes[31,94]. Computed tomography and thoracic radiographs are commonly used for staging of these tumors prior to removal[31]. Multiple driver genes have been implicated in human lung cancer studies, but none thus far have been investigated in cats. Oncogenes such as *KRAS*, *EGFR*, *BRAF*, *PIK3CA*, and *MET* have each been shown to harbor somatic mutations in as many as 33% of lung tumors[96]. *P53* or *STK11* mutations in combination with *KRAS* mutations have been linked with low survival rates and have been found to cause early-onset cancer in mice[97,98]. *EGFR* is implicated in 20% of lung cancer in humans and is responsible for the early stages of epidermoid carcinoma development[99].

Tyrosine kinase inhibitors, such as gefitinib and erlotinib have been shown to improve survival rates of human patients with susceptible *EGFR* mutations[100]. Treatment of this cancer includes surgical resection as the first line of defense in companion animals as well. Partial or complete lobectomies are generally performed, where a cuff of normal tissue is also removed to ensure a wide margin is obtained[91]. Chemotherapy is a standard approach in human lung cancer, however very few trials have been conducted to evaluate its efficiency in cats[91]. In humans, patients often develop chemoresistance and thereafter have poor survival rates. A rare feline study of chemotherapy resistance found expression of proteins that are associated with human multi-drug chemotherapy resistance in primary lung tumors, P-glycoprotein *ABCB1*, multi-drug resistant protein, and lung resistance-related protein, all of which regulate the export of chemotherapeutic drugs outside of the cell and could present future targets[101]. While the underlying feline genetic mechanism of pulmonary carcinomas, i.e., driver genes with high recurrence, is unknown, further investigation into the predicted similar somatic mutational landscape from the vast human lung cancer sequencing cases for example *EGFR*, may present as treatment options for cats outside surgery and chemotherapy.

2.7 Conclusions

Cats provide humankind with crucial companionship and should receive the best veterinary care. Nearly 6 million cats will be diagnosed with cancer each year[3], and with very little known about the genetic landscape of cancer, many pets will face euthanasia. Adopting the use of genomic medicine approaches being developed for humans may provide an improved standard of care for the cat with cancer. A critical gap in knowledge that can be filled with larger sampling of cat cancers of all types using WGS or WES methods will find somatic or germline variants implicated in cancers. This new approach to therapy may allow for more accurate and earlier

diagnosis and precise therapy leading to an improved prognosis rate in several diseases, and, importantly, deadly cancers. With many overlapping molecular and environmental causes of cancer in cats and humans, several types of cancer present opportunities for comparative therapeutics, such as mammary and oral squamous cell carcinoma. While current treatment is limited to surgery, chemotherapy, and radiation, new treatment approaches are on the horizon.

CHAPTER 3

A Domestic Cat Whole Exome Sequencing Resource for Trait Discovery

3.1 Introduction

Genomic medicine promises new avenues of disease treatment in veterinary medicine[1]. However, the appropriate resources are not yet readily available for robust implementation in clinical practice[27]. One resource which has been successfully applied to the diagnosis of rare diseases in humans is whole exome sequencing (WES) analysis, a cost-effective method for identifying potentially impactful DNA variants in the coding regions of genes[102]. DNA base changes in the exome can alter amino acids in proteins or disrupt their overall structure, so focusing on these regions offers a more direct and biologically interpretable approach to searching for putative disease variants. In comparison, whole genome sequencing (WGS) captures DNA variants spanning the entire genome. However, as the vast majority of the identified variants are within non-coding regions, much of the variation is difficult to interpret. WES also allows for deeper coverage of target sequences due to lower cost compared to WGS. The present study seeks to develop and validate the use of WES as a viable approach for determining novel disease variants in cats.

Over the last decade, a surge of studies using next generation sequencing (NGS), in particular WES, has led to many novel discoveries of candidate disease-causing variants across species. WES is recognized as an efficient means for genome resequencing and is the primary NGS approach used to help diagnose human patients with rare genetic diseases[103,104]. By selectively sequencing all protein-coding regions to a deeper depth than WGS, WES is a dependable method for finding biallelic exonic variants causative of Mendelian inherited diseases that rarely appear in healthy populations[103,104]. In humans, WES is commonly used to

find genetic causes in a wide range of diseases, even complex neurological conditions such as autism spectrum disorder[105]. Its widespread use has led to the discovery of therapeutic targets for drug development and genetic markers for innovative clinical applications[106]. Tumor WES has been especially successful by cost-effectively providing somatic variant information about a patient's normal and tumor exomes, supporting the identification of recurrent somatic mutations among known oncogenes that may suggest a mechanism of action and targets for potential drug therapies[107]. The significant depth of exome coverage is integral to overcoming diluted somatic variant allele frequencies (VAF) due to tumor clonality and purity issues.

Exome sequencing has also proven successful in non-human species. Mouse WES studies have found strong candidate alleles for models of orofacial clefting, urogenital dysmorphology, and autoimmune hepatitis[108]. In companion animals, the development of dog WES has demonstrated that causative allele discovery for common diseases has great potential[109]. Some examples in dogs include the discovery of a two-base pair deletion in *SGCD* causing muscular dystrophy, and a splice site variant in *INPP5E* which is associated with cystic renal dysplasia[110]. As there are many isolated breeds of domestic dogs, this species is an important genetic resource for cancer studies, for which WES demonstrated dogs have similar oncogene variant patterns to humans[111]. However, many oncogene variants are not equivalent to a WES analysis of human, and canine bladder cancers identified novel mutations in *FAM133*B, *RAB3GAP2*, and *ANKRD52* that are unique to canine bladder cancer, emphasizing the need to understand the biological differences in origin[112].

Similar to canines, domestic cats have long been recognized for their potential in modeling human diseases, such as retinal blindness[113,114]. Approximately 150 variants in domestic cats are associated with over 100 genetic traits or diseases, many mimicking human

disease phenotypes[23]. As feline genomic resources continue to advance, more diseases caused by single base variants are being discovered, such as two novel forms of blindness in Persian and Bengal cats[115,116]. However, a feline WES resource has not been described to date for the discovery of novel disease gene candidates. Here we describe the first feline exome resource, a WES analysis of 41 cats, and its use in the discovery of known and novel variants associated with feline phenotypes, healthy and diseased. A comparison of WES and WGS methods was also completed to understand the efficiency, depth of coverage, and sequence specificity, for variant calling from each approach.

3.2 Methods

Exome Design. The annotated exons from the Felis_catus_9.0 reference genome assembly were used as the basis to design the exome capture probes, incorporating the NCBI RefSeq release 92 annotation, containing 19,590 refGene names completed by Roche Sequencing Soultions. The coding sequences (CDS) for the primary chromosomes were extracted and consolidated into a non-overlapping set of features, and repetitive probes were removed totaling 35,724,716 bases divided over 201,683 regions. Of those bases, only 395,115 bp are not covered directly or indirectly. GO functions for removed genes were olfactory genes or unidentifiable. Since Y chromosome genes are not represented in the Felis_catus_9.0 reference, a set of coding sequence features from the Felis catus Y chromosome genomic sequence (NCBI accession KP081775) was used[117]. The cat exome panel was designed by Roche Sequencing Solutions (Madison, USA)[118]. A capture probe dataset was constructed for the full cat genome by tiling variable-length probes, ranging from 50 - 100 bases in length, at a five-base step across all sequences. Each capture probe was evaluated for repetitiveness by constructing a 15-mer histogram from the full genome sequence and then calculating the average 15-mer count across each probe, a sliding

window size of 15 bases across the length of each probe. Any probe with an average 15-mer count greater than 100 was considered to be repetitive and excluded from further characterization. Non-repetitive probes were then scored for uniqueness by aligning each capture probe to the full cat genome using SSAHA v3[119]. A close match to the genome was defined as a match length of 30 bases, allowing up to five insertions/deletions/substitutions. Capture probes were selected for each coding sequence feature by scoring one to four probes in a 20-base window, based on repetitiveness, uniqueness, melting temperature, and sequence composition, and then choosing the best capture probe in that window. The start of the 20 base windows was then moved 40 bases downstream and the process repeated. Selected probes were allowed to start up to 30 bases before the 5' start of each feature and overhang the 3' end by 30 bp. A maximum of five close matches in the genome was allowed when selecting the capture probes.

Samples and DNA Isolation. Cat DNA samples for WES were donated by owners and archived in the Lyons Feline Genetics Laboratory at the University of Missouri, College of Veterinary Medicine in accordance with the University of Missouri Institutional Animal Care and Use Committee protocol study protocols 9056, 9178, and 9642. DNA was isolated from 41 whole blood or tissue cat samples using standard organic methods[120] and verified for quantity and quality by DNA fluorescence assay (Qubit, Thermo Fisher) and ethidium bromide staining after 0.7% agarose gel electrophoresis. Ten cats with existing whole genome sequence (WGS) data were initially tested, followed by 31 novel cats for additional screening.

Sequencing. All WGS cat data used in this study was obtained from Buckley et al, and library preparation was completed by Washington University.[121] Genomic DNA (250 ng) was fragmented on the Covaris LE220 instrument targeting 250 bp inserts. Automated dual indexed libraries were constructed with the KAPA HTP library prep kit (Roche) on the NGS platform

(Perkin Elmer). The libraries were PCR-amplified with KAPA HiFi for 8 cycles. The final libraries were purified with a 1.0x AMPureXP bead cleanup and quantitated on the Caliper GX instrument (Perkin Elmer) and were pooled pre-capture generating a total 5µg library pool. Each library pool was hybridized with a custom NimbleGen probe set (Roche), targeting 35.7 Mb. The libraries were hybridized for 16 - 18 hours at 65°C followed by washing to remove non-specific hybridized library fragments. Enriched library fragments were eluted following isolation with streptavidin-coated magnetic beads and amplified with KAPA HiFi Polymerase prior to sequencing. PCR cycle optimization was performed to prevent over-amplification of the libraries. The concentration of each captured library pool was determined via qPCR utilizing the KAPA library Quantification Kit (Roche) to produce appropriate cluster counts prior to sequencing. The Illumina NovaSeq6000 instrument was used to generate paired-end 2 x 150 bp length sequences to yield an average of 14 Gb of data per 35.7 Mb target exome, producing ~80x exome sequencing depth of coverage. Exome sequencing data are available at the Sequence Read Archive under accession number PRJNA627536.

Variant Discovery. The following tools/packages were applied to WGS and WES samples in accordance with variant processing as previously described[121] 71: BWA-MEM version 0.7.17[122], Picard tools version 2.1.1 (http://broadinstitute.github.io/picard/), Samtools version 1.9[123], and Genome Analysis toolkit version 3.8[124,125,126] by Rueben Buckley. Code used for the variant calling workflow can be found at https://github.com/mu-feline-genome/batch_GATK_workflow. For WES processing, GATK tools were restricted to exons annotated in Ensembl release 99 with an additional 100 bp of flanking sequence[127]. Following processing, samples were genotyped in three separate cohorts. The first cohort consisted of all 41 WES samples. The second and third cohorts were ten matched WES and WGS samples. Variants in all three cohorts were tagged

using the same variant filtering criteria. For SNVs, the filtering criteria were QD < 2.0, FS > 60.0, SOR > 3.0, ReadPosRankSum < -8.0, MQ < 40.0, and MQRankSum < -12.5. For indels, the filtering criteria were QD < 2.0, FS > 200.0, SOR > 10.0, and ReadPosRankSum < -20.0. Although five Y chromosome genes were included in the exome probe set, these genes had not been added to the aligning reference. For WGS/WES comparison, matched WES/WGS samples were annotated using variant effect predictor (VEP)[128]. Variants from both cohorts were independently tagged as to whether they were biallelic, SNVs, or passed filtering criteria. Before analysis, variants flanking the exome primary target regions +/- 2bp were removed . Variant processing and comparisons were performed in the R statistical environment using the vcf R package[129]. Common variants between both platforms were determined as those at the same position with the same reference and alternate alleles. Exclusive variants were determined as those where the position and/or the alleles were specific to a particular platform.

Disease and trait variant detection. Variants for all 41 cats were evaluated using VarSeq software (GoldenHelix, Inc.). SNVs were annotated as having high, moderate, or low impacts on gene function. High impact variations were those that were a protein-truncating variant caused by stop gain or loss and splice-site acceptor or donor mutations[130]. Moderate impacts include missense mutations or in-frame insertions, and lastly, low impact variants are characterized by synonymous base changes, splice region variants, or intron variants. Known variants for diseases and traits were evaluated in each cat.

Polycystic Kidney Disease. A pointed cat of the Siberian breed (a.k.a. Neva Masquerade, a pointed Siberian) was diagnosed with polycystic kidney disease based on signs of renal disease (polydipsia, polyuria) and ultrasonography **(Table 3.1, cat 37)**. DNA was submitted using buccal swabs and a whole blood sample to two different commercial testing laboratories in which both

confirmed the absence of the currently known autosomal dominant polycystic kidney disease in polycystin-1 (PKD1)[131 132]. The dam and a sibling were also reported as having PKD by ultrasonography but were not available for genetic analyses.

Cystinuria. A three-month-old European shorthair kitten from the isle of Korfu, Greece, was presented to the AniCura Small Animal Hospital, Bielefeld, FRG, for heavy straining during urination, and the owner reported the kitten would fall over from time to time (Table 1, cat 17). The kitten had been pretreated with two injections of cephalexine and dexamethasone for suspected cystitis, however, difficulty in urination worsened. Upon hospital admission, the kitten was in good general condition. Abdominal palpation revealed an enlarged urinary bladder. Abdominal X-ray showed over 30 radiolucent urinary stones up to a diameter of half of the width of the last rib. Urinary bladder stones and some urethral stones were removed via cystolithotomy and retrograde flushing of the urethra. Urinary stones were submitted for infraspectroscopic stone analysis. Stone analysis revealed pure cystine stones and a diagnosis of cystinuria was made. Urinary stones reoccurred at six months of age, but the kitten was otherwise healthy.

3.3 Results

Phenotype cohort. WES was performed on 41 individual cats, representing a variety of different diseases and traits, some with known disease alleles **(Table 3.1)**. The 41 cats can be further divided into two separate cohorts: the first is the initial ten cats that had nine known variants for various diseases and aesthetic traits, e.g., coat colors and fur types. These 10 cats also had matched WGS data, which was used to assess the efficacy of WES. The second cohort of 31 represents genetically uncharacterized cats. These cats represented 11 different breeds and include 14 random-bred cats. Groups of cats with similar genetic backgrounds were used to evaluate causes for mediastinal lymphoma, a seizure disorder, eyelid colobomas,

707 hypothyroidism, hypovitaminosis D, blue eyes of Ojos Azules breed, and curly hair coat of the
708 Tennessee Rex. Five cats were reported with cardiac diseases, including hypertrophic
709 cardiomyopathy (HCM). At least seven neurological disorders are represented in the study
710 population, generally representing novel presentations in random-bred cats. Overall, the 41 cats
711 had approximately 31 different unknown disease presentations.

712 **Table 3.1 Description and diseases of 41 cats for WES evaluation.**

No.	Id.	Breed	Sex	Disease / Trait	Gene(s)
1	19725	Lykoi	F	Lykoi	*HR*
2	13230	Mixed Breed	F	Bengal PRA / Bobbed tail	*KIF3B / HES7*
3	14056	Mixed Breed	M	Persian PRA / *Long*	*AIPL1 / FGF5*
4	17994	Mixed Breed	F	Hydrocephalus	*GDF7*
5	19067	Munchkin	F	Dwarfism / Dominant White	*UGDH / KIT*
6	5012	Oriental	M	Lymphoma	*Unknown*
7	20382	Peterbald	M	*Hairless*	*LPAR6**
8	11615	Random Bred	M	*Dominant White*	*KIT*
9	18528	Random Bred	M	*Spotting*	*KIT*
10	20424	Siberian	F	*Long* / Cardiac disease	*FGF5 / Candidate*
11	22550	Bengal	F	Polyneuropathy	*Unknown*
12	20957	Devon Rex	U	Papilloma virus	*Unknown*
13	22752	Devon Rex	M	Neurological disorder	*Unknown*
14-15	21983/ 21464	Ojos Azules	1M:1F	Ojos Azules	*Unknown*
16	20964	Oriental	F	Cardiac disease	*Unknown*
17	22728	Random bred	F	Cystinuria	*SLC3A1**
18	20617	Random Bred	M	Neuronal ceroid lipofuscinosis	*CLN6**
19	20948	Random Bred	M	Cinnamic acid urea	*Unknown*
20	21153	Random Bred	M	Ambulatory paraparesis	*Unknown*
21	22287	Random Bred	F	Myotonia congenita	*Unknown*
22	22397	Random Bred	M	Neurological disorder	*Unknown*
23	22505	Random Bred	M	Cardiac disease	*Unknown*
24	22623	Random Bred	U	Pycnodysostosis	*Candidate*
25	22740	Random Bred	F	Epidemolysis bullosa	*Unknown*
26 – 27	22741/ 22742	Random Bred	1F:1M	Eyelid coloboma	*Unknown*
28	22751	Random Bred	M	Ehlers-Danlos	*Unknown*
29 – 30	22763/ 22764	Random Bred	2F	Hypothyroidism	*Candidate*
31 – 32	22761/ 22762	Savannah	2M	Hypovitaminosis D	*Unknown*
33	21984	Scottish Fold	F	Cardiac disease	*Candidate*
34 – 35	20384/ 20385	Selkirk Rex	1F:1U	Seizures	*Unknown*
36	20953	Siamese	F	Cardiac disease	*Candidate*
37	22622	Siberian	U	PKD	*PKD2**
38	22711	Singapura	F	Hypovitaminosis D	*Candidate*
39 – 40	8641/ 8642	Tennessee Rex	1F:1M	Rexoid hair coat	*Unknown*
41	6623	Oriental	M	Lymphoma	*Unknown*

41		**14 breeds**	**19F:18M:4U**	**~31 diseases & traits**	

A complete description of diseases and traits for entire cohort**.** Candidate genes are potential genes that been identified with less evidence of a causal mutations. U: Unknown sex; F: Female; M: Male *Mutations as tentative causal variants for diseases presented

Sequence coverage and specificity. To assess the performance of this feline exome resource, deep coverage WES data was produced for ten cats with WGS data for comparison. After mapping to Felis_catus_9.0, base quality trimming, and PCR duplicate removal, the average percentage of reads uniquely mapped was 82% **(Table 3.2).** The average sequencing depth was 267x with a range of 76x to 458x **(Supplementary Table 1).** Assessing the depth of coverage, of the 201,683 exonic targets, 98.1% aligned with coverage of >20x. An average of 6.98% of the total reads aligned outside of the targeted regions of the genome (**Supplementary Table 2)**. For the uncharacterized 31 cat exomes, the sequencing depth was adjusted to typical human WES studies; for this group of cats, we estimated the average depth of coverage to be 80x. 96.41% of exonic targets aligned with a coverage of >20x, ranging from 91-98%. An average of 10.41% of total reads aligned off-target is slightly higher when compared to the first 10 higher-coverage cats that can be attributed to lower sequencing depth in the larger cohort. As expected, overall, there is a reduction in mapping at lower depth of coverage; for example, at 40x, 93.5% of targeted bases were covered **(Figure 1)**, conversely, 99% are covered at 2x.

Known variant validation. To further analyze the effectiveness of WES for variant detection, we examined each sample for the presence of known trait-causing variants. The Felis_catus_9.0 Ensembl release 99 gene annotation was used with a selection of exons with +/- 30 bp to match exome capture design and variants were browsed using the VarSeq software (GoldenHelix, Inc). The majority of the previously published 115 trait causing variants in the domestic cat that have been documented as causal for diseases and traits affect either the coding regions or a splice donor/acceptor site[23]. Of these known variants, 44 were identified in our WES cohort. All variants for coat colors and diseases expected to be present in the ten cats were identified,

including the alleles in the loci for Agouti (*ASIP* - a[133]), Brown (*TYRP1* – b[134]), Color (*TYR* – cs[135]),

Table 3.2 Summary of Metrics across both Cohorts

	Average -First 10	Range- First 10	Average- Cohort of 31	Range- Cohort of 31
Depth of Coverage	267x	76-485x	80x	60-108x
% of Bases Covered	99.1%	92.3-100%	96.4%	91-98%
% Reads Aligned	99.9%	99.9-100%	82%	75-85%

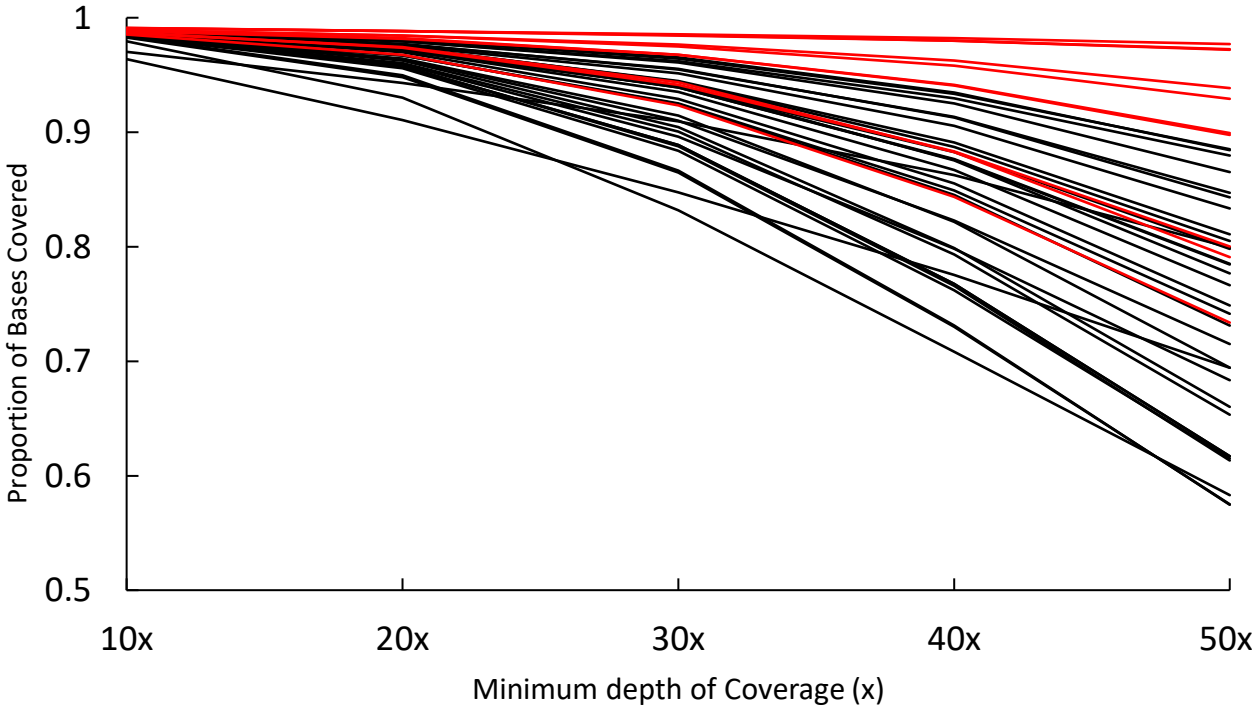


Figure 3.1 The proportion of bases covered with the exome capture probes. The initial 10 samples are colored in red, with the X axis showing the depth of coverage, which is how many times a nucleotide base is covered starting at a depth of 10x and increasing to 50x.

Dense (*MLPH* - d[136]), Longhair (*FGF5* - I[137]), Lykoi (*HR* – hrTN, hrVA[138]), Bengal progressive retinal degeneration (*KIF3B*[139]) and Persian progressive retinal degeneration (*AIPL1*[115]), hydrocephalus (*GDF7*[140]), and others. The cats also had variants known to affect cat blood type as well[141,142]. In accordance with the limitations of our feline exome capture design, neither known structural nor intronic variants were detected. When analyzing discordant reads in a WGS dwarf sample, a deletion and rearrangement indicating a structural variant (SV) was visible in the *UDGH* gene[115], but no read discordance was found in the WES analysis (Supplementary Figure 1). In addition, the *KIT* intron one SV for White and Spotting were not identified[143]. Therefore, the WES approach will fail to identify many complex SVs, an important limitation to consider for future feline trait discovery efforts.

Novel candidate variant discovery. Novel DNA variants were explored as putatively causal for diseases and traits in 33 cats. A novel frameshift mutation in polycystin 2 (*PKD2*[144]), a gene associated with polycystic kidney disease (PKD) was predicted to disrupt protein function in a Siberian cat shown by ultrasound to have PKD. This mutation, a single-base deletion, causes a truncated protein (p.Lys737Asnfs*2). This variant was heterozygous in the affected cat and unique to the exome data and was not identified in the 195-cat cohort of the 99 Lives variant dataset[121]. This variant was also identified in both grandparents on the dam's side of the pedigree, although kidney ultrasound was not available. However, analysis of other Siberian cats with PKD diagnosed by ultrasound failed to identify the c.2211delG variant in *PKD2*, suggesting that this could be a private variant and that other disease-causing PKD variants are yet to be discovered in this breed.

A variant in the lysophosphatidic acid receptor 6 (*LPAR6*) gene associated with the autosomal recessive rexoid (Marsella wave) coat of the Cornish rex breed was detected in a

Peterbald cat, which is a hairless breed[145]. However, the hairless trait is considered autosomal dominant by cat breeders. The annotation predicts a c.249delG causing a p.Phe84Leufs*10; therefore, this Peterbald cat likely is compound heterozygous for two mutations juxtaposed in LPAR6. This variant was heterozygous in the affected cat, unique to the exome data and not identified in the 99 Lives variant dataset.

A known feline disease variant was also re-identified [121]. A solute carrier family 3-member 1 (*SLC3A1*) variant was homozygous in a Greek cat presenting with cystinuria. The c.1342C>T variant, causing a p.Arg448Trp at position A3:66539609 has been previously documented to be associated with this condition[146]. No other cat in the exome dataset had this variant. Many of the variants associated with cat blood group B and its extended haplotype were detected in 11 cats, suggesting five cats as type B, one was confirmed[141]. Variants were detected in *APOBEC3*, which is associated with feline immunodeficiency virus (FIV) infection in cats, and three cats had the allelic combination producing the IRAVP amino acid haplotype that is associated with FIV resistance[147]. Novel findings included two cats that were heterozygous for a porphyria variant in UROS (c.140C>T, c.331G>A)[148,149], one cat which was homozygous for FXII deficiency variant (FXII_1631G>C)[148], and had died as a kitten, and one cat which was heterozygous for a copper metabolism deficiency in *ATP7B*[150]. Additional variants for neuronal ceroid lipofuscinosis, pycnodysostosis, Ehlers-Danlos syndrome, hypothyroidism, and hypovitaminosis D, and several individual-specific variants for hypertrophic cardiomyopathy are under further investigation **(Table 3.1)**.

3.4 Discussion

In humans, WES has flourished over the past few years and is becoming more common in the practice of genomic medicine, especially newborn screening[151]. Well-annotated genomes and

extensive resources, such as for human and mouse, have led to the development of various exome capture products ranging from those with a very limited focus, e.g., oncogene panels, to more extensive designs including 5' and 3' untranslated regions, predicted regulatory elements, and non-coding RNAs. For other mammals, exome capture designs have ranged from 44.6 Mb in pigs[152] to 146.8 Mb in rats[153,154], illustrating the variation in experimental objectives. This is not currently the case for veterinary medicine due to several factors: a dog or cat owner's unwillingness to incur the costs, lower accuracy of available genome references[155], and the uncertainty of treatment options driven by sequence variant data. In companion animals, only the domestic dog has exome capture probes available, which span 53 to 152 Mb with an overlap of 34.5 Mb between the capture designs[156,157]. In this study, a feline exome resource was developed by designing capture probes against the annotated Felis_catus_9.0 genome assembly, a highly contiguous assembly that enabled efficient probe design[155]. The targeted 35.7 Mb accounts for the exons and 30bp of flanking sequences to minimize the loss of detectable splice donor and acceptor variants.

Success in disease variant identification in any species using WES is dependent on multiple factors, including mode of inheritance, sequencing depth, and efficient probe design that covers the regions of interest with high specificity, minimizing the number of off-target reads. Sequence coverage of ≥20x is generally regarded as the standard to efficiently detect heterozygous variants[158]. At this threshold, an acceptable average target coverage of 96.4% was obtained in our study. In our first WES experiment of 10 cats, we achieve maximum exonic coverage of 99% with a mean depth of 267x at aligned bases. However, we have found this high-depth approach is not necessary or cost-efficient for the discovery of feline associated disease variants. The first domestic dog exome design[156], which covered 52.8 Mb distributed over 203,059 regions, had a

range of 87-90% mapped reads at a 102x mean sequencing depth. An updated canine design[156] had 93.5% of the targeted bases (<53 Mb) covered to at least 1X depth of coverage, while in our feline exome design, the on-target reads were nearly 100% at 10x sequencing depth. Whilst absolute dog and cat exome comparisons are difficult due to the differences in annotation, genome assembly accuracy, and design techniques, both of these resources reveal acceptable performance.

The intended application of the cat WES was twofold: the identification of heritable, Mendelian diseases and traits, and somatic mutations in cancer. In this study, the focus was the former and included the assessment of the efficiency of the feline exome design for SNV discovery against ten matched WGS samples. The matched WGS and WES cats had an average of 30x and 267x depth of coverage, respectively, with the vast majority of SNVs and indels in overlapping regions being detected by both platforms. Altogether, these findings suggest the use of this feline exome probe set was extremely consistent with variant discovery from WGS, where 99.4% were uncovered in WGS while only 1.5% were absent from the WES cats. Consistent with large cohort human studies, indel discovery was less consistent (92.5% overlap) with 12.2% of WGS indels absent from WES data owing to the well-known short-read misalignment problem in regions with indels of varying size. Differences in the number of common variants between platforms is due to differential filtering, as common variants were identified prior to when filtering was performed. The percentage of exclusive variants per platform also varied according to variant impact, with high impact variants representing the largest percentage of exclusive variants for their impact class. Since high impact mutations are generally rare due to their impact on normal gene function, their enrichment within platform exclusive variant sets is

845 expected. In the same manner, as low impact variants have no impact on gene function, they are
846 less likely to be identified as platform exclusive within their variant class.

847 Previously characterized and unknown germline or somatic variants of clinical
848 significance, the former often not identifiable without the parents, were investigated to confirm if
849 each were identical or unique to genes associated with each disease or phenotype in prior studies.
850 Known variants were first confirmed to validate the accuracy of the cat exome design for the
851 following aesthetic traits: Agouti, Brown, Dense, Gloves, Dilution, Extension, Long, Lykoi, and
852 hairless coat types[23]. In addition, disease variants were found in genes earlier shown to be
853 candidate alleles in hydrocephalus[159], hypertrophic cardiomyopathy[160], and progressive retinal
854 atrophy[115]. These results importantly validate our design is capable of detecting variants with
855 prior trait association. Nonetheless, a primary study objective was to find new potential causal
856 variants in our small mixed disease and trait cohort of 31 domestic cats. This cohort was
857 searched to find novel candidate variants for three diseases and traits; feline autosomal dominant
858 polycystic kidney disease (ADPKD), atrichia, hypotricha. ADPKD is a common inherited
859 autosomal dominant disease affecting about 6% of the world's cats[132] and is characterized by
860 fluid-filled cysts that form in the bilateral kidneys that often leads to renal failure[161]. Many of the
861 features of feline ADPKD are similar to human ADPKD and recent studies demonstrated the
862 utility of the cat model[113,162]. The c. 10063C>A mutation in exon 29 of *PKD1* was the only
863 known causative allele for feline ADPKD[132], however, for human ADPKD, variants are found
864 throughout PKD1. A variant in polycystin 2 (*PKD2*), c.2211delG at position B1:134992553,
865 causes a p.Lys737Asnfs*2 and was identified in a Siberian cat from Europe, indicating
866 additional alleles may be segregating for ADPKD in cats.

867 Domestic cats have various forms of atrichia and hypotricha, which even though each is
868 characterized by baldness or loss of hair coat, are not considered diseased cats since breeders
869 have selected upon these observed traits to develop new breeds. Only two breeds are recognized
870 as completely hairless, the Sphynx and Donskoy. Donskoy cats are a breed of Russian cats in
871 which loss of hair is determined by a semi-dominant allele[163]. Peterbald cats were bred in Russia
872 in 1994 as a product of a Donskoy and an Oriental Shorthair cross, and are often born with no
873 hair, or lose their hair over time[164]. Cornish Rex, a hypotrichia breed, that is characterized by a
874 curly coat, is caused by a homozygous deletion mutation in *LPAR6*[165]. The Peterbald cat had an
875 *LPAR6* 4 base pair deletion that is in juxtaposition to a compound heterozygote for the Cornish
876 rex deletion variant. Both variants result in premature stop codons a few amino acids
877 downstream of the variant site. Other disease-associated variants were re-identified, such as
878 cystinuria variants, in which the cat was homozygous and affected. Determination of allele
879 frequencies through the 99 lives project[155] improved the identification of cats that were
880 heterozygous for variants associated with recessive diseases, such as, porphyria[148], Factor XII
881 deficiency[166], and copper metabolism[150]. The inclusion of 99 Lives WGS data was central to
882 establishing the likelihood of variants being causal for diseases and further cross-species
883 explorations of variant frequencies promises to better define variants of uncertain significance[167].

884 Clinical use of sequence variant information in companion animals is in the very early
885 stages, which hampers the ability of veterinarians to rapidly diagnose some diseases without
886 standard or unclear phenotypic determinants. In the future, it could be used to adapt treatments
887 to the specific animal and disease type[168]. Many diagnosed rare diseases have a poor prognosis,
888 with some less than 90 days; thus, cost-effective sequencing approaches may help discover
889 alternate and more effective treatments. The Undiagnosed Diseases Program of the National

Institutes of Health routinely uses WES for this purpose of finding treatments where none exist, suggesting veterinary medicine could benefit in the same manner[169]. We confirm here, as other studies have shown, that WES is cost-effective, data process-efficient (by requiring less computing time), and easier to use than WGS for inferring a variant's biological relevance[170]. As in the dog, a first step is offered toward the use of feline WES for robust disease variant detection, including the validation of previously identified causal alleles and the discovery of novel candidate variants that we suggest are of interest for further experimental scrutiny[171]. We have developed domestic cat-specific WES, and importantly, based on our findings, validated its use for the evaluation of potential disease variants for the future practice of feline genomic medicine.

CHAPTER 4

A Genetic Profile of Feline Oral Squamous Cell Carcinoma

4.1 Introduction

Feline oral squamous cell carcinoma (FOSCC) is the fourth most common cancer, and the most commonly found oral tumor in cats[172], with a one-year survival rate of less than 10% [76].This cancer affects squamous cell lining in the oral and oropharyngeal cavity; gingiva, tongue, and sublingual regions[172] and rarely metastasizes to distant locations; however, local bone invasion and lymph nodes can be affected. FOSCC is often diagnosed at stages too late to intervene with symptoms including oral pain successfully clinically, difficulty swallowing, loss of teeth, and anorexia. Early studies have shown that the use of flea collars, feeding with mostly canned foods were found to increase developing FOSCC 5-fold compared to those who did not have canned foods or tuna[20]. FOSCC has a limited number of treatment options, with the best typically being surgical resection of the tumor from the oral cavity, if feasible, followed by radiation and/or chemotherapy if concerns arise about risk of recurrence or metastasis based on histopathology. However, due to the high reoccurrence rate, these therapies often have poor outcomes and often the recommendation is euthanasia may end up being the best option[77,173]. FOSCC presents with similar risk factors and molecular mechanisms as huaman head and neck cancer (HNSCC) and may present an opportunity as a model for therapy [174].

HNSCC is the sixth most common cancer found among humans worldwide, with 550,000 new cases per year and also has a low 5-year survival rate of less than 50% with 275,000 deaths per year[175,176]. Like FOSCC, if HNSCC is diagnosed in early stages survival rates are much higher at 82%, compared to 26% if the tumor is distantly metastasized[76,177,178]. Known risk

factors for HNSCC include exposure to tobacco smoke, alcohol, and infection with HPV[177-179]. Little is understood about molecular mechanism similarities in cats and humans with only candidate gene approaches in the cat. Both humans and cats show the perturbed function of *p53*, causing issues in a cells metabolism, cycle arrest, and apoptosis[180]. FOSCC and HNSCC are both similar in disease progression and biologic behavior with both cancers being locally invasive with metastasis to regional lymph nodes[174]. Overexpression of *EGFR* is found in 69 to 100% of FOSCC and 90% of HHNSCC, causing cell cycle progression, uncontrolled proliferation, and invasion through the activation of intracellular tyrosine kinase and is associated with poor prognosis in human[84,85,174,181]. Similarities such as low survival rates, similar genetic, and morphological profiles in feline may offer a viable model to study HNSCC.

Naturally occurring animal models of cancer are becoming more integral to a better understanding of tumor evolution and progression such as HNSCC compared to rodent models[174,182]. Murine models, for example, lack important factors that contribute to spontaneous tumor formation and follow-up adaptive immune responses. Understanding the mutually exclusive genetic environment of FOSCC will allow us to determine how informative a comparative model model is to study which molecular canidates are involved in a tumors formation in both FOSCC and HNSCC. With many uncertainties in the genetics of FOSCC, we aimed to characterize the mutational and transcriptional profile of FOSCC thus gaining insight into its molecular pathogenies when comparing to HNSCC. To accomplish this, WES and RNAseq was performed. FOSCC tumor tissue and matching blood samples were used for WES, and RNA-seq was generated on FOSCC tumor tissue and oral cavity samples from healthy cats. Following these experiments, we searched among candidate genes for monotherapy matches that

offer a retrospective view of missed mutation druggable mutations for treating feline patients with FOSCC.

4.2 Methods

Clinical samples. Cryopreserved tissues corresponding to 6 FOSCC tumor samples and 6 matching whole blood samples, as well as 3 normal oral mucosal samples collected from healthy animals, were used (**Table 4. 1**). Tumor samples had been collected during standard-of-care surgical procedures and stored by the Cornell Veterinary Biobank until retrieved for analysis. Sample collection was performed in accordance with a protocol (#2005-0151) approved by Cornell University's Institutional Animal Care and Use Committee. Accordingly, informed consent to authorize the use of tissue samples and clinical data for research purpose was obtained from cat owners prior to sample collection, and undue harm was never inflicted to client-owned cats for the purposes of this study; all methods were performed in accordance with the relevant guidelines and regulations. The diagnosis of FOSCC was validated using routine hematoxylin and eosin-stained samples archived by the Anatomic Pathology Section at Cornell University's College of Veterinary Medicine by a board-certified veterinary pathologist (ADM); tumors were diagnosed following previously described criteria[183] while blinded to molecular assays.

FOSCC histology. Histological assessment of tissues was done using routine hematoxylin and eosin-stained samples archived by the Anatomic Pathology Section at Cornell University's College of Veterinary Medicine by a board-certified veterinary pathologist (ADM); tumors were diagnosed following previously described criteria[183] while blinded to molecular assays **(Figure 4.1)**.

967

Table 4.1 Cohort Characteristics

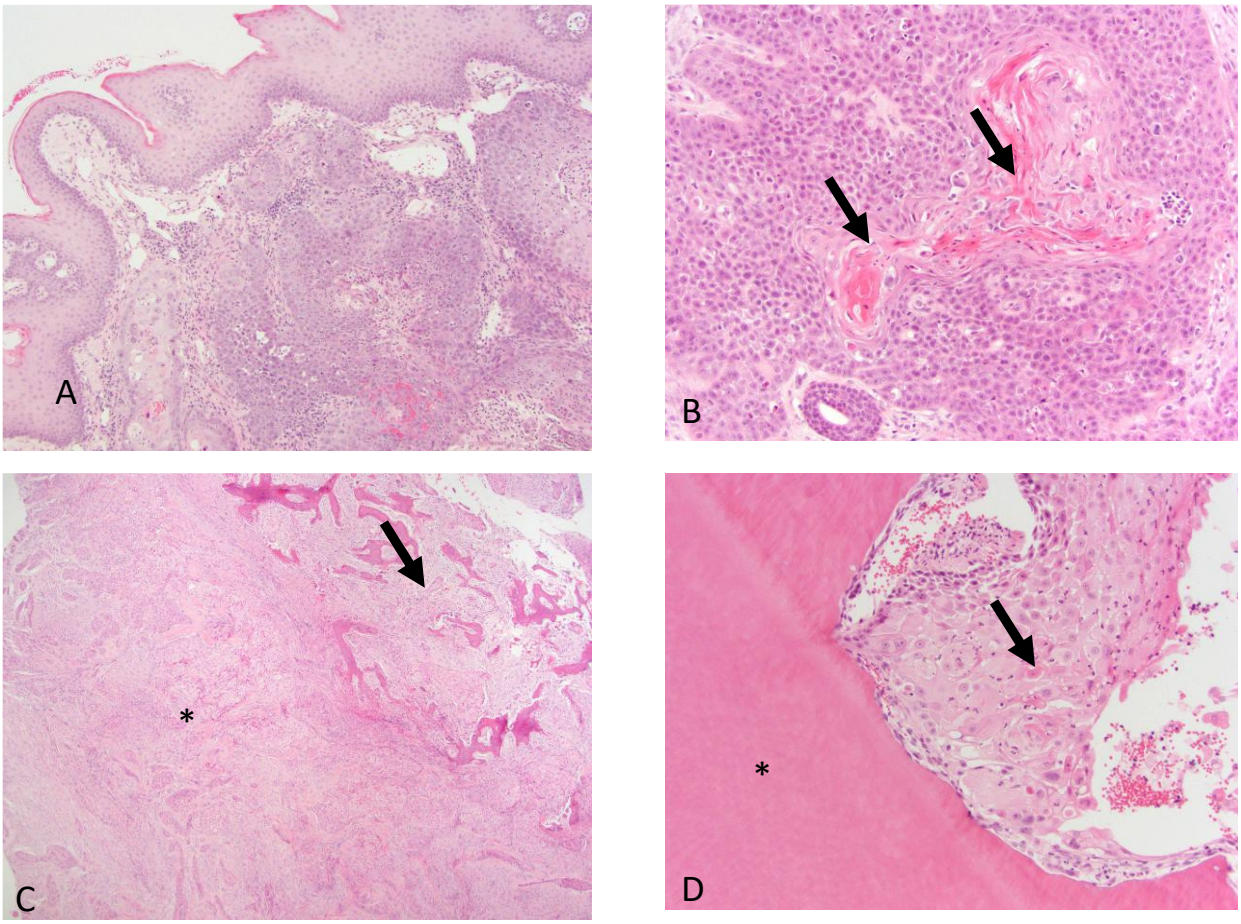
968

Sample	Location Collected	Sex	Age
23263	Tongue	MC	Adult
9895	Mandible/Gingiva	FS	Adult
26903	Mandible	unknown	Adult
7741	Oral Cavity	MC	Adult
24147	Oral Cavity/Gingiva	MC	Adult
605591	Unknown	Unknown	Adult

MC=male castrated

FS=female spayed

969



970 **Figure 4.1 Pathology of Oral Squamous Carcinoma**

A. Underlying a moderately hyperplastic gingival epithelium are ribbons, cords, and trabeculae of neoplastic squamous epithelial cells. B. Neoplastic squamous epithelial cells surround and produce brightly eosinophilic keratin (arrows). C. Neoplastic squamous epithelial cells that are enmeshed in abundant scirrhous response (asterisk) are associated with marked bony invasion and remodeling (arrow). D. Neoplastic squamous epithelial cells with dyskeratosis (arrow) invade the dentin layer of a tooth (asterisk). Images provided by……

Image info

A. 79666; 40x mag

B. 16236; 100x mag

C. 76610; 200x mag

D. 89317; 200x mag

Somatic Variant Calling, Annotation, and Filtering. Raw sequence reads were mapped to Felis_Catus_9.0 reference using Burrows-Wheeler Aigner (BWA) v0.7.17. Sam files were sorted and converted to bam and merged using Piccard tools v2.18.9. PCR and optical duplicates were marked using Piccard tools v. 2.19.9. These files were then processed through the Genome Analysis Toolkit (GATK) v.4.0.1 for base quality rescore calibration. Mutect2 was used to identify somatic single nucleotide variants (SNVs) and insertions and deletions (indels) and filtered through the standard MuTect2 filters, such as *t_lod_fstar*, filters out variants with insufficient evidence of presence in tumor sample and panel*_of_normals*, which filters out variants present in at least two samples in the panel of normal (**Supplemental Figure 4.1**). VCF files were further filtered for missing data and minor allele frequencies less than .1% using VCFtools. SNVs were then annotated using Variant effect Predictor (VEP) v. 101.0[184]. The validity of called variants was manually confirmed using IGV by randomly selecting 20 variants per sample and confirming the presence of that variant in the raw BAM files.

Measuring tumor mutational burden (TMB). TMB has been reported to help classify whether a cancer type is more amenable to immunotherapy due to the presentation of multiple antigen targets[185]. For each VCF file, Felis_Catus_9.0 was set as the reference and Ensembl (release 102) genes was used as the definition. TMB was calculated using the number of non-synonymous SNVs, somatic mutations altering the amino acid sequence, found per megabase in the coding regions. We used an exome size of 35MB that was calculated based on the size of coding regions in the cat genome.

Driver Database Version 3/OncoKB*:* Driver Database version 3 was accessed online on 3/15/21(http://driverdb.tms.cmu.edu.tw/cancer). Cancer driver genes database was selected, and then the HNSCC data set was selected that included 263 genes. In addition, OncoKB was

accessed on the same day, and 10 HNSCC druggable genes were obtained. All gene sets were used for further investigation of their presence in the FOSCC samples.

4.3 Results

FOSCC cohort characteristics. In this study, we analyzed six FOSCC samples collected by the Cornell Veterinary Biobank between 6/19/14 and 9/19/19. Only 6 samples were available at the time of analysis. Samples included three castrated males, one spayed female, and two samples of unknown sex (**Table 4.1**). Samples were collected from various locations in the head and neck region, including the mandible, gingiva, and tongue. Due to varied anatomical location and sex, we were unable to account for differences in sex or spatial context for tumor genetic changes.

SNV annotation. In the six samples, we found 1,057 synonymous and nonsynonymous variants with a mean of 176 variants per FOSCC (**Figure 4.2 and 4.3**). After variant visualization using IGV, we found a false call rate of 5% across all 6 samples for 20 genes (**Supplemental Figure 4.2**). Among all somatic SNVs, 56 and 731 were nonsense and missense, respectively. Only one gene, *TP53*, a commonly mutated gene in many cancer types, including head and neck cancer[180] showed multiple occurrences of SNVs with 83% showing somatic mutations as differing positions (**Figure 4.4**). Four samples had mutations in *TP53*, with three being missense and one a frameshift.

TMB has emerged as a biomarker for human patient stratification toward immunotherapy. However, the prognostic value of TMB across cancer types is uncertain. Using our small cohort we generated a preliminary TMB estimate of FOSCC to compare to earlier estimates in HNSCC. Using non-synonymous somatic mutations, we calculated the TMB for each feline tumor (**Supplemental Table 4.1**). The mean TMB for all six samples was 3.7 with a range of 1.4 - 8.5.

We were not able to determine if survival was associated with TMB score due to all cats being euthanized soon after the time of tumor resection.

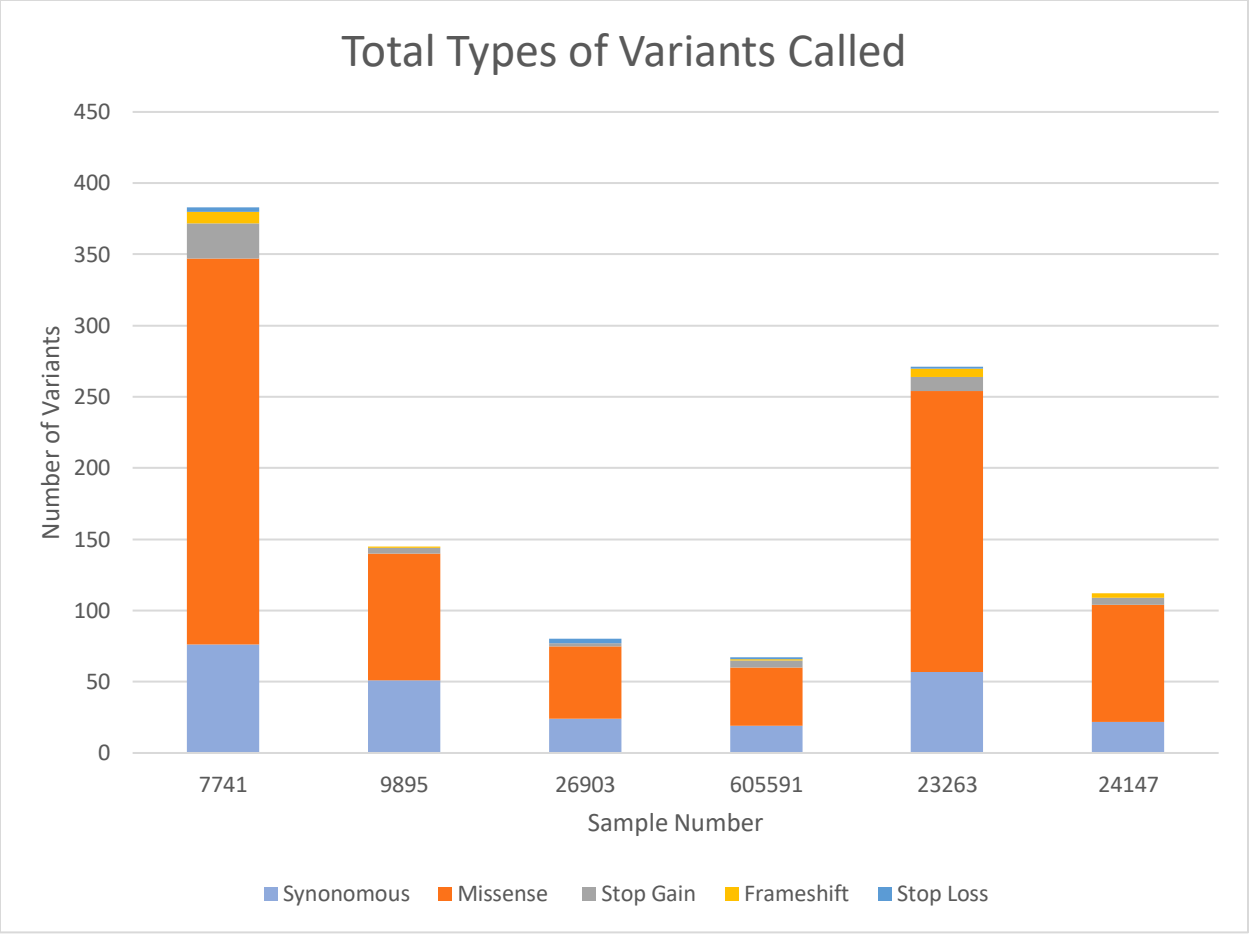


Figure 4.2 All variants called, including synonymous variants, using GATK-Mutect2.

1023

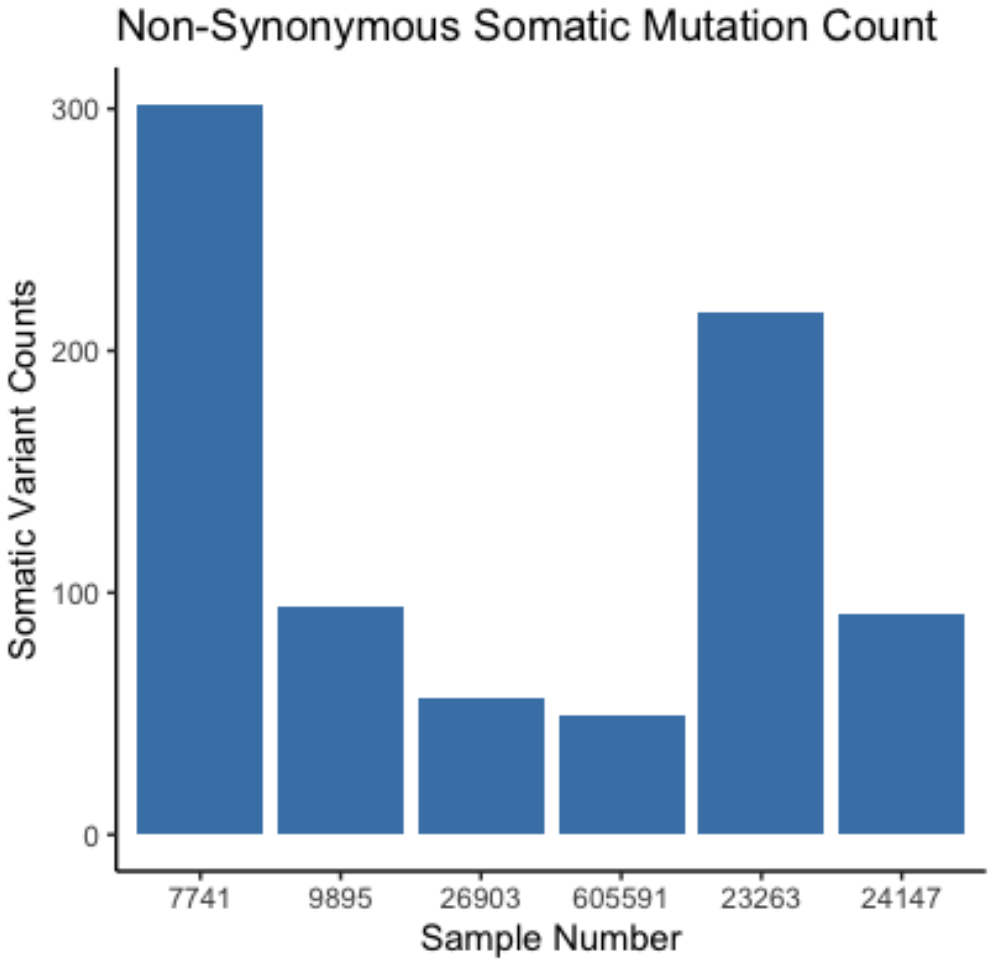


Figure 4.3 Number of Non-Synonymous somatic variants per sample. Non-Synonymous somatic variants called across all samples. Average of 176 variants over six samples including 56 nonsense high impact mutations and 731 missense mutations.

1026

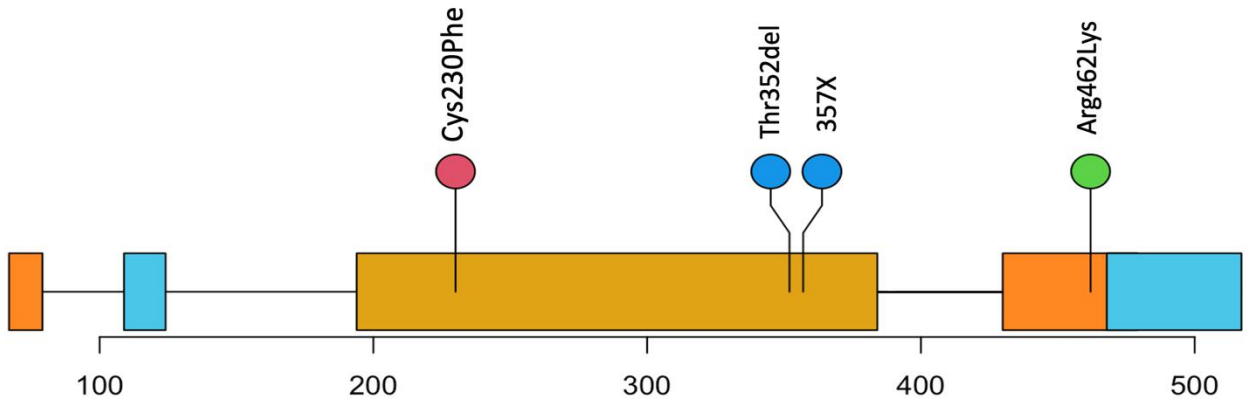


Figure 4.4 Location of the four TP53 mutations found in five of our FOSCC samples.

FOSCC comparison with HNSCC. We compared the somatic mutations of FOSCC to genes shown have highly recurrent mutations in HNSCC to assess gene model similarities. Genes harboring FOSCC non-synonymous SNVs were matched to the same HNSCC genes in the Driver Database 3[186] and OncoKB [187] databases The *TP53* gene has the most frequent somatic mutations in HNSCC, and was also the most recurrent in FOSCC, with four of the samples having a missense mutation, and sample 24147 containing a compound heterozygous mutation, specifically a missense and splice region variant **(Figure 4.4)**. A missense mutation in *KAT2B* was present in two samples, with several other missense mutations present in only one gene per sample **(Table 4.2)**. Samples 9895 and 23263 had the most genes implicated in HNSCC with four genes in common each **(Table 4.2).** This data indicates some evidence for overlap in mutational background between HNSCC and FOSCC but is very preliminary at this stage.

4.4 Discussion

In this first study of the somatic mutations and gene expression variation present in FOSCC, we describe the similarities and differences when comparing to HNSCC. FOSCC and HNSCC have similarly low survival rates and morphological profiles that suggests a FOSCC model may offer some comparative insight into HNSCC therapeutic strategies or mechanism of action[186]. The scope of FOSCC somatic mutations, however, is not known which we partially address by exploring comparable molecular features between species, i.e., somatic mutation and gene expression patterns. We found several genes **(Table 4.2)** overlapping with HHNSCC, for example the surprising overlap with *TP53,* the most recurrent in our FOSCC samples. *TP53* is the most mutated in all cancer types [188,189], but also has the highest prevalence of driver mutations in HNSCC (41%)[23,2].

1049

1050

Sample	Gene	Annotation	Position
7741	TP53	Frameshift/ Splice region variant	E1:2541686
9895	TP53	Missense	E1:17726778
	MED12L	Missense	C2:115279831
	PXYLP1	Missense	C2:124530451
26903	TP53	In-Frame Deletion	E1:2541698
	KAT2B	Missense	C2:139167628
23263	KAT2B	Missense	C2:139217714
	ARID1A	Missense	C1:20180520
	KMT2D	Missense	B4: 78023698
	TP53	Frameshift	E1:2541322

Table 4.2 Variant annotations for variants found in FOSCC in common with known HNSCC.

Given *TP53* is a tumor suppression gene, and certain somatic mutations are associated with lower survival outcomes in patients with HNSCC[189]. With *TP53* missense variants in three of the samples, an insertion in one, and a compound heterozygous mutation in another it is reasonable to predict these cats experienced rapid tumor progression yet missing clinical follow up prevented us from drawing this conclusion **(Figure 4.4)**. Coincidently, *TP53* mutations have been implicated in other FOSCC studies, with mutations in this gene found in 24 - 69% of cancers[190-192]. In HNSCC, *TP53* mutations are found in 70% of all cases, with variation in *TP53* being a predictive marker for immunotherapy in those with metastatic HNSCC[180].

As estimates of the TMB have been used to predict positive patient response to immune checkpoint inhibitor therapy in some cancer types, e.g. non-small cell lung cancer and melanoma [193,194], we sought to compare FOSCC to HNSCC for this metric. Recent studies on HNSCC have found that mutations in *TP53* are associated with high TMB and low overall survival rates, and high TMB patients responded well to immunotherapy[195,196]. In our FOSCC samples the average TMB was 3.7 with a range of 1.4 - 8.5. Cancer studies that calculated HNSCC TMB as high (>5.0) or low (<5.0), show higher TMB is associated with poor prognosis[197,198]. Our first TMB estimates in FOSCC fall within the observed HNSCC range suggesting some benefits could be gained for immunotherapy outcomes. More FOSCC sample sequencing to obtain better estimates of TMB is needed as well as the future availability of immune checkpoint inhibitors that could collectively be used to substantially improve outcomes for this very lethal cancer.

Other FOSCC recurrent or single gene mutations of interest were *KAT2B, ARID1A, MED12L, HOXB3,* and *PXYLP1* each with interesting features for comparative inference (Table 3). Two samples had a missense variant in *KAT2B* (Table 3). *KAT2B* is a gene that codes for an enzyme that functions as a histone acetyltransferase (HATs) and mutations in this gene have

been implicated in many diseases including cancers[199]; however, it has not been evaluated in FOSCC. *KAT2B* is part of a family of lysine transferases that are responsible for the acetylation of genes that targets a broad range of proteins and can function as tumor suppressors and oncogenes[200]. *KAT2B* is also responsible for inhibiting cell cycle progression and counteracting mitogenic activity. HNSCC cell line studies have shown universal loss of *KAT2B*[175], and a study using HNSCC tumors found significantly lower expressions of *KAT2B* compared to the normal tissue[201]. *KMT2D* was identified in only one sample but has similar epigenetic properties to *KAT2B*. Studies completed by The Cancer Genome Atlas have shown mutations often occurring in *KMT2D*, keeping chromatin in an open state, thereby promoting gene expression[202,203]. Mutations in both *KAT2B* and *KMT2D* possibly induce epigenetic changes in HHNSCC and FOSCC which alternative to immunotherapy could and may open an avenue to study epigenetic drug control of both human and feline oral squamous cell carcinoma[200,203].

A missense mutation in *ARID1A* was also found in one of our samples (**Table 4.2**). *ARID1A* is a gene that is often found to be deleted in many human cancers; however, to our knowledge the effect of this gene in feline cancers is unknown. *ARID1A* functions as a tumor suppressor and tumor stemness repressor by disrupting the perturbed function of p53 or PTEN pathways [204,205]. The upregulation of miR-31 is known to have oncogenic properties in human head and neck cancer and studies show that elevated expression of this miRNA causes reduced expression of *ARID1A*, and patients with low expression of *ARID1A* are found to have the worst survival rates[205,206].

Several other genes were found to overlap in HHNSCC and FOSCC. *MED12L*, *HOXB3*, and *PXYLP1* were also identified as single nucleotide variants in one sample each (**Table 4.4**); however, these genes are understudied in both HHNSCC and FOSCC. Mediator Complex

Subunit 12L (*MED12L*) works by activating the kinase activity of CDK8 which regulates the growth and division of cells[207]. Studies report significant differentiation of the expression of *MED12L* in many cancers including head and neck cancer; however, an altered *MED12* complex is altered in 3.05% of HNSCC patients[208,209]. *HOX* genes regulate a wide range of cell activity including proliferation and migration. HNSCC studies have shown am overall elevation in all HOX genes, including *HOXB3*[210-212]. There was no overlap in actionable genes from OncoKB related to head and neck cancer. We believe this is due to the low sample size, and further studies need to be conducted.

In FOSCC, we have identified the somatic mutations landscape by exome sequencing. For the best-known cancer driver gene, *TP53*, we observe mutations that despite their presence in FOSCC, are not the same variants as those observed in human head and neck cancer. But several other genes also overlap between the two types of cancer in our small cohort suggesting the use of similar genes that initiate tumorigenesis and perhaps future comparative models of treatment. This small exploratory study demonstrated the ability to call variants unique to feline oral squamous cell carcinoma tumors, identified common genes between human and feline oral squamous carcinoma. This study presents a starting point to study FOSCC in a larger cohort of feline oral squamous carcinoma patients and draw more similarities between human and feline oral cancer. With the further development of this technology, we may be able to diagnose this cancer at earlier stages using genomic methods, as well as possibly develop immunotherapy treatments for both cats and humans.

References

1 Buckley, R. M. & Lyons, L. A. Precision/Genomic Medicine for Domestic Cats. *Vet Clin North Am Small Anim Pract* **50**, 983-990, doi:10.1016/j.cvsm.2020.05.005 (2020).

2 Thomas, R. Cytogenomics of Feline Cancers: Advances and Opportunities. *Vet Sci* **2**, 246-258, doi:10.3390/vetsci2030246 (2015).

3 MacVean, D. W., Monlux, A. W., Anderson, P. S., Jr., Silberg, S. L. & Roszel, J. F. Frequency of canine and feline tumors in a defined population. *Vet Pathol* **15**, 700-715, doi:10.1177/030098587801500602 (1978).

4 Cannon, C. M. Cats, Cancer and Comparative Oncology. *Vet Sci* **2**, 111-126, doi:10.3390/vetsci2030111 (2015).

5 Baioni, E. *et al.* Estimating canine cancer incidence: findings from a population-based tumour registry in northwestern Italy. *BMC Vet Res* **13**, 203, doi:10.1186/s12917-017-1126-0 (2017).

6 Poirier, V. J. & Steffen, F. in *Feline Orthopedic Surgery and Musculoskeletal Disease* (eds P. M. Montavon, K. Voss, & S. J. Langley-Hobbs) 97-104 (W.B. Saunders, 2009).

7 Poon, A. C., Matsuyama, A. & Mutsaers, A. J. Recent and current clinical trials in canine appendicular osteosarcoma. *Can Vet J* **61**, 301-308 (2020).

8 Wayne, R. K. & vonHoldt, B. M. Evolutionary genomics of dog domestication. *Mamm Genome* **23**, 3-18, doi:10.1007/s00335-011-9386-7 (2012).

9 Hu, Y. *et al.* Earliest evidence for commensal processes of cat domestication. *Proc Natl Acad Sci U S A* **111**, 116-120, doi:10.1073/pnas.1311439110 (2014).

10 Vonholdt, B. M. *et al.* Genome-wide SNP and haplotype analyses reveal a rich history underlying dog domestication. *Nature* **464**, 898-902, doi:10.1038/nature08837 (2010).

11 *The American Kennel Club*, <https://www.akc.org/?utm_source=google&utm_medium=cpc&utm_campaign=20200413_akc-rev_responsivesearchad1&utm_content=CoreBrand_AKC&utm_term=akc&gclid=Cj0KCQjwr-SSBhC9ARIsANhzu17hOLHgTqkO2zsvYJB-uuMT7dhyRm0r6CTT0NilLAjZioZ8wCbsAF8aAhsAEALw_wcB> (2022).

12 Kohyama, M. *et al.* Degenerative myelopathy in the Collie breed: a retrospective immunohistochemical analysis of superoxide dismutase 1 in an affected Rough Collie, and a molecular epidemiological survey of the SOD1: c.118G>A mutation in Japan. *J Vet Med Sci* **79**, 375-379, doi:10.1292/jvms.16-0391 (2017).

13 Driscoll, C., Yamaguchi, N., O'Brien, S. J. & Macdonald, D. W. A suite of genetic markers useful in assessing wildcat (Felis silvestris ssp.)-domestic cat (Felis silvestris catus) admixture. *J Hered* **102 Suppl 1**, S87-90, doi:10.1093/jhered/esr047 (2011).

14 Lipinski, M. J. *et al.* The ascent of cat breeds: genetic evaluations of breeds and worldwide random-bred populations. *Genomics* **91**, 12-21, doi:10.1016/j.ygeno.2007.10.009 (2008).

15 Leroy, G. Genetic diversity, inbreeding and breeding practices in dogs: results from pedigree analyses. *Vet J* **189**, 177-182, doi:10.1016/j.tvjl.2011.06.016 (2011).

16 Kurushima, J. D. *et al.* Variation of cats under domestication: genetic assignment of domestic cats to breeds and worldwide random-bred populations. *Anim Genet* **44**, 311-324, doi:10.1111/age.12008 (2013).
17 Montague, M. J. *et al.* Comparative analysis of the domestic cat genome reveals genetic signatures underlying feline biology and domestication. *Proc Natl Acad Sci U S A* **111**, 17230-17235, doi:10.1073/pnas.1410083111 (2014).
18 Morris, D. *Cat Watching*. (Ebury PR, 1985).
19 Matsumoto, Y. *et al.* Genetic relationships and inbreeding levels among geographically distant populations of Felis catus from Japan and the United States. *Genomics* **113**, 104-110, doi:10.1016/j.ygeno.2020.11.018 (2021).
20 Bertone, E. R., Snyder, L. A. & Moore, A. S. Environmental and lifestyle risk factors for oral squamous cell carcinoma in domestic cats. *J Vet Intern Med* **17**, 557-562 (2003).
21 Barrs, V. R. & Beatty, J. A. Feline alimentary lymphoma: 1. Classification, risk factors, clinical signs and non-invasive diagnostics. *J Feline Med Surg* **14**, 182-190, doi:10.1177/1098612X12439265 (2012).
22 Hayes, H. M., Jr., Milne, K. L. & Mandell, C. P. Epidemiological features of feline mammary carcinoma. *Vet Rec* **108**, 476-479, doi:10.1136/vr.108.22.476 (1981).
23 OMIA. Online Mendelian Inheritance in Animal. (2020).
24 Xu, C. *et al.* Human ADPKD primary cyst epithelial cells with a novel, single codon deletion in the PKD1 gene exhibit defective ciliary polycystin localization and loss of flow-induced Ca2+ signaling. *Am J Physiol Renal Physiol* **292**, F930-945, doi:10.1152/ajprenal.00285.2006 (2007).
25 Young, A. E., Biller, D. S., Herrgesell, E. J., Roberts, H. R. & Lyons, L. A. Feline polycystic kidney disease is linked to the PKD1 region. *Mamm Genome* **16**, 59-65, doi:10.1007/s00335-004-2412-2 (2005).
26 Rodney, A. R. *et al.* A domestic cat whole exome sequencing resource for trait discovery. *Sci Rep* **11**, 7159, doi:10.1038/s41598-021-86200-7 (2021).
27 Mauler, D. A. *et al.* Precision Medicine in Cats: Novel Niemann-Pick Type C1 Diagnosed by Whole-Genome Sequencing. *J Vet Intern Med* **31**, 539-544, doi:10.1111/jvim.14599 (2017).
28 Buckley, R. M. *et al.* A new domestic cat genome assembly based on long sequence reads empowers feline genomic medicine and identifies a novel gene for dwarfism. *PLoS Genet* **16**, e1008926, doi:10.1371/journal.pgen.1008926 (2020).
29 Vascellari, M., Baioni, E., Ru, G., Carminato, A. & Mutinelli, F. Animal tumour registry of two provinces in northern Italy: incidence of spontaneous tumours in dogs and cats. *BMC Vet Res* **5**, 39, doi:10.1186/1746-6148-5-39 (2009).
30 Simon, S. *Facts & Figures 2019: US Cancer Death Rate has Dropped 27% in 25 Years*, <https://www.cancer.org/latest-news/facts-and-figures-2019.html> (2019).
31 Nunley, J. *et al.* Primary pulmonary neoplasia in cats: assessment of computed tomography findings and survival. *J Small Anim Pract* **56**, 651-656, doi:10.1111/jsap.12401 (2015).
32 LeBlanc, A. K. & Mazcko, C. N. Improving human cancer therapy through the evaluation of pet dogs. *Nat Rev Cancer* **20**, 727-742, doi:10.1038/s41568-020-0297-3 (2020).

33 Gardner, H. L., Fenger, J. M. & London, C. A. Dogs as a Model for Cancer. *Annu Rev Anim Biosci* **4**, 199-222, doi:10.1146/annurev-animal-022114-110911 (2016).

34 N.Ettinger, S. Principles of treatment for feline lymphoma. *Clinical Techniques in Small Animal Practice* **18**, 98-102, doi:https://doi.org/10.1053/svms.2003.36623 (2003).

35 David M. Vail, D. H. T., Julias M. Liptak,. 33 - Hematopoietic Tumors. *Withrow and MacEwen's Small Animal Clinical Oncology* **Sixth Edition**, 688-772, doi:https://doi.org/10.1016/B978-0-323-59496-7.00033-5. (2019).

36 Gieger, T. Alimentary lymphoma in cats and dogs. *Vet Clin North Am Small Anim Pract* **41**, 419-432, doi:10.1016/j.cvsm.2011.02.001 (2011).

37 Louwerens, M., London, C. A., Pedersen, N. C. & Lyons, L. A. Feline lymphoma in the post-feline leukemia virus era. *J Vet Intern Med* **19**, 329-335, doi:10.1892/0891-6640(2005)19[329:flitpl]2.0.co;2 (2005).

38 Shelton, G. H. *et al.* Feline immunodeficiency virus and feline leukemia virus infections and their relationships to lymphoid malignancies in cats: a retrospective study (1968-1988). *J Acquir Immune Defic Syndr (1988)* **3**, 623-630 (1990).

39 Lingard, A. E. *et al.* Low-grade alimentary lymphoma: clinicopathological findings and response to treatment in 17 cases. *J Feline Med Surg* **11**, 692-700, doi:10.1016/j.jfms.2009.05.021 (2009).

40 Tidd, K. S. *et al.* Outcomes in 40 cats with discrete intermediate- or large-cell gastrointestinal lymphoma masses treated with surgical mass resection (2005-2015). *Vet Surg* **48**, 1218-1228, doi:10.1111/vsu.13281 (2019).

41 Gabor, L. J., Malik, R. & Canfield, P. J. Clinical and anatomical features of lymphosarcoma in 118 cats. *Aust Vet J* **76**, 725-732, doi:10.1111/j.1751-0813.1998.tb12300.x (1998).

42 Kano, R., Sato, E., Okamura, T., Watanabe, S. & Hasegawa, A. Expression of Bcl-2 in feline lymphoma cell lines. *Vet Clin Pathol* **37**, 57-60, doi:10.1111/j.1939-165X.2008.00013.x (2008).

43 Hartmann, K. Clinical aspects of feline retroviruses: a review. *Viruses* **4**, 2684-2710, doi:10.3390/v4112684 (2012).

44 Henry, C. J. Cancer Managment in Small Animals (2010).

45 Besmer, P. *et al.* The Hardy-Zuckerman 2-FeSV, a new feline retrovirus with oncogene homology to Abelson-MuLV. *Nature* **303**, 825-828, doi:10.1038/303825a0 (1983).

46 Vascellari, M., Melchiotti, E., Bozza, M. A. & Mutinelli, F. Fibrosarcomas at presumed sites of injection in dogs: characteristics and comparison with non-vaccination site fibrosarcomas and feline post-vaccinal fibrosarcomas. *J Vet Med A Physiol Pathol Clin Med* **50**, 286-291, doi:10.1046/j.1439-0442.2003.00544.x (2003).

47 Romanelli, G., Marconato, L., Olivero, D., Massari, F. & Zini, E. Analysis of prognostic factors associated with injection-site sarcomas in cats: 57 cases (2001-2007). *J Am Vet Med Assoc* **232**, 1193-1199, doi:10.2460/javma.232.8.1193 (2008).

48 Porcellato, I. *et al.* Feline Injection-Site Sarcoma. *Vet Pathol* **54**, 204-211, doi:10.1177/0300985816677148 (2017).

49 Aberdein, D. *et al.* Comparison of the histology and immunohistochemistry of vaccination-site and non-vaccination-site sarcomas from cats in New Zealand. *N Z Vet J* **55**, 203-207, doi:10.1080/00480169.2007.36769 (2007).

50 Phelps, H. A., Kuntz, C. A., Milner, R. J., Powers, B. E. & Bacon, N. J. Radical excision with five-centimeter margins for treatment of feline injection-site sarcomas: 91 cases (1998-2002). *J Am Vet Med Assoc* **239**, 97-106, doi:10.2460/javma.239.1.97 (2011).
51 Egeblad, M. & Werb, Z. New functions for the matrix metalloproteinases in cancer progression. *Nat Rev Cancer* **2**, 161-174, doi:10.1038/nrc745 (2002).
52 Akkoc, A., Inan, S. & Sonmez, G. Matrix metalloproteinase (MMP-2 and MMP-9) and steroid receptor expressions in feline mammary tumors. *Biotech Histochem* **87**, 312-319, doi:10.3109/10520295.2011.652173 (2012).
53 Julius M. Liptak, N. I. C. 22 - Soft Tissue Sarcomas. *Withrow and MacEwen's Small Animal Clinical Oncology* **Sixth Edition**, 404-431, doi:https://doi.org/10.1016/B978-0-323-59496-7.00022-0. (2019).
54 Cronin, K. *et al.* Radiation therapy and surgery for fibrosarcoma in 33 cats. *Vet Radiol Ultrasound* **39**, 51-56, doi:10.1111/j.1740-8261.1998.tb00325.x (1998).
55 Bray, J. & Polton, G. Neoadjuvant and adjuvant chemotherapy combined with anatomical resection of feline injection-site sarcoma: results in 21 cats. *Vet Comp Oncol* **14**, 147-160, doi:10.1111/vco.12083 (2016).
56 Bloch, J., Rogers, K., Walker, M., Dawson, J. & Wilson-Robles, H. Treatment of feline injection-site sarcoma with surgery and iridium-192 brachytherapy: retrospective evaluation of 22 cats. *J Feline Med Surg* **22**, 313-321, doi:10.1177/1098612X19844345 (2020).
57 MacNeill, A. L. The potential of the combined use of targeted type I interferon pathway inhibitors and oncolytic viruses to treat sarcomas. *Vet Comp Oncol* **18**, 36-42, doi:10.1111/vco.12547 (2020).
58 Jourdier, T. M. *et al.* Local immunotherapy of spontaneous feline fibrosarcomas using recombinant poxviruses expressing interleukin 2 (IL2). *Gene Ther* **10**, 2126-2132, doi:10.1038/sj.gt.3302124 (2003).
59 Overley, B., Shofer, F. S., Goldschmidt, M. H., Sherer, D. & Sorenmo, K. U. Association between ovarihysterectomy and feline mammary carcinoma. *J Vet Intern Med* **19**, 560-563, doi:10.1892/0891-6640(2005)19[560:aboafm]2.0.co;2 (2005).
60 Karin U. Sorenmo, D. R. W., Valentina Zappulli,. 28 - Tumors of the Mammary Gland,. *Withrow and MacEwen's Small Animal Clinical Oncology* **Sixth Edition**, doi:https://doi.org/10.1016/B978-0-323-59496-7.00028-1. (2019).
61 Egenvall, A. *et al.* Mortality of life-insured Swedish cats during 1999-2006: age, breed, sex, and diagnosis. *J Vet Intern Med* **23**, 1175-1183, doi:10.1111/j.1939-1676.2009.0396.x (2009).
62 Misdorp, W., Romijn, A. & Hart, A. A. Feline mammary tumors: a case-control study of hormonal factors. *Anticancer Res* **11**, 1793-1797 (1991).
63 Gimenez, F., Hecht, S., Craig, L. E. & Legendre, A. M. Early detection, aggressive therapy: optimizing the management of feline mammary masses. *J Feline Med Surg* **12**, 214-224, doi:10.1016/j.jfms.2010.01.004 (2010).
64 Mills, S. W. *et al.* Prognostic value of histologic grading for feline mammary carcinoma: a retrospective survival analysis. *Vet Pathol* **52**, 238-249, doi:10.1177/0300985814543198 (2015).

65 Gemignani, F. *et al.* Association of surgical approach with complication rate, progression-free survival time, and disease-specific survival time in cats with mammary adenocarcinoma: 107 cases (1991-2014). *J Am Vet Med Assoc* **252**, 1393-1402, doi:10.2460/javma.252.11.1393 (2018).
66 Griffiths, C. L. & Olin, J. L. Triple negative breast cancer: a brief review of its characteristics and treatment options. *J Pharm Pract* **25**, 319-323, doi:10.1177/0897190012442062 (2012).
67 Marques, C., Correia, J. & Ferreira, F. HER2-positive feline mammary carcinoma. *Aging (Albany NY)* **8**, 1574-1575, doi:10.18632/aging.101015 (2016).
68 Caliari, D. *et al.* Triple-negative vimentin-positive heterogeneous feline mammary carcinomas as a potential comparative model for breast cancer. *BMC Vet Res* **10**, 185, doi:10.1186/s12917-014-0185-8 (2014).
69 De Maria, R. *et al.* Spontaneous feline mammary carcinoma is a model of HER2 overexpressing poor prognosis human breast cancer. *Cancer Res* **65**, 907-912 (2005).
70 Maniscalco, L. *et al.* Activation of AKT in feline mammary carcinoma: a new prognostic factor for feline mammary tumours. *Vet J* **191**, 65-71, doi:10.1016/j.tvjl.2010.12.016 (2012).
71 Mayr, B., Ortner, W., Reifinger, M. & Loupal, G. Loss of chromosome B2-material in three cases of feline mammary tumours. *Res Vet Sci* **59**, 61-63, doi:10.1016/0034-5288(95)90032-2 (1995).
72 Adega, F., Borges, A. & Chaves, R. Cat Mammary Tumors: Genetic Models for the Human Counterpart. *Vet Sci* **3**, doi:10.3390/vetsci3030017 (2016).
73 Nascimento, C. *et al.* Serum PD-1/PD-L1 Levels, Tumor Expression and PD-L1 Somatic Mutations in HER2-Positive and Triple Negative Normal-Like Feline Mammary Carcinoma Subtypes. *Cancers (Basel)* **12**, doi:10.3390/cancers12061386 (2020).
74 Nascimento, C., Gameiro, A., Ferreira, J., Correia, J. & Ferreira, F. Diagnostic Value of VEGF-A, VEGFR-1 and VEGFR-2 in Feline Mammary Carcinoma. *Cancers (Basel)* **13**, doi:10.3390/cancers13010117 (2021).
75 Belcher, R., Hayes, K., Fedewa, S. & Chen, A. Y. Current treatment of head and neck squamous cell cancer. *J Surg Oncol* **110**, 551-574, doi:10.1002/jso.23724 (2014).
76 Martin, C. K., Tannehill-Gregg, S. H., Wolfe, T. D. & Rosol, T. J. Bone-invasive oral squamous cell carcinoma in cats: pathology and expression of parathyroid hormone-related protein. *Vet Pathol* **48**, 302-312, doi:10.1177/0300985810384414 (2011).
77 Piegols, H. J., Takada, M., Parys, M., Dexheimer, T. & Yuzbasiyan-Gurkan, V. Investigation of novel chemotherapeutics for feline oral squamous cell carcinoma. *Oncotarget* **9**, 33098-33109, doi:10.18632/oncotarget.26006 (2018).
78 Soltero-Rivera, M. M., Krick, E. L., Reiter, A. M., Brown, D. C. & Lewis, J. R. Prevalence of regional and distant metastasis in cats with advanced oral squamous cell carcinoma: 49 cases (2005-2011). *J Feline Med Surg* **16**, 164-169, doi:10.1177/1098612X13502975 (2014).
79 Hayes, A. M., Adams, V. J., Scase, T. J. & Murphy, S. Survival of 54 cats with oral squamous cell carcinoma in United Kingdom general practice. *J Small Anim Pract* **48**, 394-399, doi:10.1111/j.1748-5827.2007.00393.x (2007).

80 Nasry, W. H. S. *et al.* CD147 and Cyclooxygenase Expression in Feline Oral Squamous Cell Carcinoma. *Vet Sci* **5**, doi:10.3390/vetsci5030072 (2018).
81 Sparger, E. E. *et al.* Investigation of immune cell markers in feline oral squamous cell carcinoma. *Vet Immunol Immunopathol* **202**, 52-62, doi:10.1016/j.vetimm.2018.06.011 (2018).
82 Chu, S. *et al.* A virome sequencing approach to feline oral squamous cell carcinoma to evaluate viral causative factors. *Vet Microbiol* **240**, 108491, doi:10.1016/j.vetmic.2019.108491 (2020).
83 Park, S. W. *et al.* The effect of nitric oxide on cyclooxygenase-2 (COX-2) overexpression in head and neck cancer cell lines. *Int J Cancer* **107**, 729-738, doi:10.1002/ijc.11498 (2003).
84 Bergkvist, G. T. *et al.* Expression of epidermal growth factor receptor (EGFR) and Ki67 in feline oral squamous cell carcinomas (FOSCC). *Vet Comp Oncol* **9**, 106-117, doi:10.1111/j.1476-5829.2010.00239.x (2011).
85 Sabattini, S. *et al.* Epidermal growth factor receptor expression is predictive of poor prognosis in feline cutaneous squamous cell carcinoma. *J Feline Med Surg* **12**, 760-768, doi:10.1016/j.jfms.2010.04.010 (2010).
86 Northrup, N. C. *et al.* Outcomes of cats with oral tumors treated with mandibulectomy: 42 cases. *J Am Anim Hosp Assoc* **42**, 350-360, doi:10.5326/0420350 (2006).
87 Boston, S. E. *et al.* Outcomes of eight cats with oral neoplasia treated with radical mandibulectomy. *Vet Surg* **49**, 222-232, doi:10.1111/vsu.13341 (2020).
88 Liptak, J. M. *et al.* Outcomes of cats treated with maxillectomy: 60 cases. A Veterinary Society of Surgical Oncology retrospective study. *Vet Comp Oncol* **19**, 641-650, doi:10.1111/vco.12634 (2021).
89 Bregazzi, V. S. *et al.* Response of feline oral squamous cell carcinoma to palliative radiation therapy. *Vet Radiol Ultrasound* **42**, 77-79, doi:10.1111/j.1740-8261.2001.tb00907.x (2001).
90 Marconato, L. *et al.* Toxicity and outcome in cats with oral squamous cell carcinoma after accelerated hypofractionated radiotherapy and concurrent systemic treatment. *Vet Comp Oncol* **18**, 362-369, doi:10.1111/vco.12557 (2020).
91 David M. Vail, D. H. T., Julias M. Liptak. 24 - Tumors of the Respiratory System,. *Withrow and MacEwen's Small Animal Clinical Oncology* **Sixth Edition**, Pages 492-523, doi:https://doi.org/10.1016/B978-0-323-59496-7.00024-4. (2019).
92 Trevail, T. Feline primary lung neoplasia challenge. *Vet Times* (2009).
93 Goldfinch, N. & Argyle, D. J. Feline lung-digit syndrome: unusual metastatic patterns of primary lung tumours in cats. *J Feline Med Surg* **14**, 202-208, doi:10.1177/1098612X12439267 (2012).
94 Hahn, K. A. & McEntee, M. F. Primary lung tumors in cats: 86 cases (1979-1994). *J Am Vet Med Assoc* **211**, 1257-1260 (1997).
95 Fowler, B. L. *et al.* Ecological level analysis of primary lung tumors in dogs and cats and environmental radon activity. *J Vet Intern Med* **34**, 2660-2670, doi:10.1111/jvim.15936 (2020).
96 Parikh, A. R. Lung Cancer Genomics. *Acta Med Acad* **48**, 78-83, doi:10.5644/ama2006-124.244 (2019).

97 Johnson, L. *et al.* Somatic activation of the K-ras oncogene causes early onset lung cancer in mice. *Nature* **410**, 1111-1116, doi:10.1038/35074129 (2001).
98 La Fleur, L. *et al.* Mutation patterns in a population-based non-small cell lung cancer cohort and prognostic impact of concomitant mutations in KRAS and TP53 or STK11. *Lung Cancer* **130**, 50-58, doi:10.1016/j.lungcan.2019.01.003 (2019).
99 de Sousa, V. M. L. & Carvalho, L. Heterogeneity in Lung Cancer. *Pathobiology* **85**, 96-107, doi:10.1159/000487440 (2018).
100 Maemondo, M. *et al.* Gefitinib or chemotherapy for non-small-cell lung cancer with mutated EGFR. *N Engl J Med* **362**, 2380-2388, doi:10.1056/NEJMoa0909530 (2010).
101 Hifumi, T., Miyoshi, N., Kawaguchi, H., Nomura, K. & Yasuda, N. Immunohistochemical detection of proteins associated with multidrug resistance to anti-cancer drugs in canine and feline primary pulmonary carcinoma. *J Vet Med Sci* **72**, 665-668, doi:10.1292/jvms.09-0519 (2010).
102 Bamshad, M. J. *et al.* Exome sequencing as a tool for Mendelian disease gene discovery. *Nat Rev Genet* **12**, 745-755, doi:10.1038/nrg3031 (2011).
103 Fresard, L. & Montgomery, S. B. Diagnosing rare diseases after the exome. *Cold Spring Harb Mol Case Stud* **4**, doi:10.1101/mcs.a003392 (2018).
104 Koboldt, D. C. *et al.* Exome-based mapping and variant prioritization for inherited Mendelian disorders. *Am J Hum Genet* **94**, 373-384, doi:10.1016/j.ajhg.2014.01.016 (2014).
105 An, Y. *et al.* De novo variants in the Helicase-C domain of CHD8 are associated with severe phenotypes including autism, language disability and overgrowth. *Hum Genet* **139**, 499-512, doi:10.1007/s00439-020-02115-9 (2020).
106 Anderson, D., Baynam, G., Blackwell, J. M. & Lassmann, T. Personalised analytics for rare disease diagnostics. *Nat Commun* **10**, 5274, doi:10.1038/s41467-019-13345-5 (2019).
107 Warr, A. *et al.* Exome Sequencing: Current and Future Perspectives. *G3 (Bethesda)* **5**, 1543-1550, doi:10.1534/g3.115.018564 (2015).
108 Fairfield, H. *et al.* Exome sequencing reveals pathogenic mutations in 91 strains of mice with Mendelian disorders. *Genome Res* **25**, 948-957, doi:10.1101/gr.186882.114 (2015).
109 Broeckx, B. J. *et al.* Improved canine exome designs, featuring ncRNAs and increased coverage of protein coding genes. *Sci Rep* **5**, 12810, doi:10.1038/srep12810 (2015).
110 Dillard, K. J. *et al.* A splice site variant in INPP5E causes diffuse cystic renal dysplasia and hepatic fibrosis in dogs. *PLoS One* **13**, e0204073, doi:10.1371/journal.pone.0204073 (2018).
111 Elvers, I. *et al.* Exome sequencing of lymphomas from three dog breeds reveals somatic mutation patterns reflecting genetic background. *Genome Res* **25**, 1634-1645, doi:10.1101/gr.194449.115 (2015).
112 Ramsey, S. A. *et al.* Cross-species analysis of the canine and human bladder cancer transcriptome and exome. *Genes Chromosomes Cancer* **56**, 328-343, doi:10.1002/gcc.22441 (2017).
113 Narfstrom, K., Holland Deckman, K. & Menotti-Raymond, M. The domestic cat as a large animal model for characterization of disease and therapeutic intervention in hereditary retinal blindness. *J Ophthalmol* **2011**, 906943, doi:10.1155/2011/906943 (2011).

114 Gurda, B. L., Bradbury, A. M. & Vite, C. H. Canine and Feline Models of Human Genetic Diseases and Their Contributions to Advancing Clinical Therapies. *Yale J Biol Med* **90**, 417-431 (2017).
115 Lyons, L. A. *et al.* Whole genome sequencing in cats, identifies new models for blindness in AIPL1 and somite segmentation in HES7. *BMC Genomics* **17**, 265, doi:10.1186/s12864-016-2595-4 (2016).
116 Ofri, R. *et al.* Characterization of an Early-Onset, Autosomal Recessive, Progressive Retinal Degeneration in Bengal Cats. *Invest Ophthalmol Vis Sci* **56**, 5299-5308, doi:10.1167/iovs.15-16585 (2015).
117 Brashear, W. A., Raudsepp, T. & Murphy, W. J. Evolutionary conservation of Y Chromosome ampliconic gene families despite extensive structural variation. *Genome Res* **28**, 1841-1851, doi:10.1101/gr.237586.118 (2018).
118 Li, G. *et al.* Comparative analysis of mammalian Y chromosomes illuminates ancestral structure and lineage-specific evolution. *Genome Res* **23**, 1486-1495, doi:10.1101/gr.154286.112 (2013).
119 Ning, Z., Cox, A. J. & Mullikin, J. C. SSAHA: a fast search method for large DNA databases. *Genome Res* **11**, 1725-1729, doi:10.1101/gr.194201 (2001).
120 Michael R. Green, J. S. Molecular Cloning: A Laboratory Manual **Fourth**.
121 Reuben M. Buckley, B. W. D., Wesley A. Brashear, Fabiana H. G. Farias, Kei Kuroki, Tina Graves, LaDeana W. Hillier, Milinn Kremitzki, Gang Li, Rondo Middleton, Patrick Minx, Chad Tomlinson, Leslie A. Lyons, William J. Murphy, Wesley C. Warren. in *bioRxiv* (2020).
122 Li, H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. doi:arXiv preprint arXiv:1303.3997 (2013).
123 Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078-2079, doi:10.1093/bioinformatics/btp352 (2009).
124 DePristo, M. A. *et al.* A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat Genet* **43**, 491-498, doi:10.1038/ng.806 (2011).
125 Jun, G., Wing, M. K., Abecasis, G. R. & Kang, H. M. An efficient and scalable analysis framework for variant extraction and refinement from population-scale DNA sequence data. *Genome Res* **25**, 918-925, doi:10.1101/gr.176552.114 (2015).
126 Van der Auwera, G. A. *et al.* From FastQ data to high confidence variant calls: the Genome Analysis Toolkit best practices pipeline. *Curr Protoc Bioinformatics* **43**, 11 10 11-11 10 33, doi:10.1002/0471250953.bi1110s43 (2013).
127 Cunningham, F. *et al.* Ensembl 2019. *Nucleic Acids Res* **47**, D745-D751, doi:10.1093/nar/gky1113 (2019).
128 McLaren, W. *et al.* The Ensembl Variant Effect Predictor. *Genome Biol* **17**, 122, doi:10.1186/s13059-016-0974-4 (2016).
129 Knaus, B. J. & Grunwald, N. J. vcfr: a package to manipulate and visualize variant call format data in R. *Mol Ecol Resour* **17**, 44-53, doi:10.1111/1755-0998.12549 (2017).
130 Li, M. M. *et al.* Standards and Guidelines for the Interpretation and Reporting of Sequence Variants in Cancer: A Joint Consensus Recommendation of the Association for Molecular Pathology, American Society of Clinical Oncology, and College of American Pathologists. *J Mol Diagn* **19**, 4-23, doi:10.1016/j.jmoldx.2016.10.002 (2017).

131 Grahn, R. A. *et al.* Genetic testing for feline polycystic kidney disease. *Anim Genet* **35**, 503-504, doi:10.1111/j.1365-2052.2004.01213.x (2004).
132 Lyons, L. A. *et al.* Feline polycystic kidney disease mutation identified in PKD1. *J Am Soc Nephrol* **15**, 2548-2555, doi:10.1097/01.ASN.0000141776.38527.BB (2004).
133 Gershony, L. C. *et al.* Who's behind that mask and cape? The Asian leopard cat's Agouti (ASIP) allele likely affects coat colour phenotype in the Bengal cat breed. *Anim Genet* **45**, 893-897, doi:10.1111/age.12206 (2014).
134 Lyons, L. A., Foe, I. T., Rah, H. C. & Grahn, R. A. Chocolate coated cats: TYRP1 mutations for brown color in domestic cats. *Mamm Genome* **16**, 356-366, doi:10.1007/s00335-004-2455-4 (2005).
135 Imes, D. L., Geary, L. A., Grahn, R. A. & Lyons, L. A. Albinism in the domestic cat (Felis catus) is associated with a tyrosinase (TYR) mutation. *Anim Genet* **37**, 175-178, doi:10.1111/j.1365-2052.2005.01409.x (2006).
136 Ishida, Y. *et al.* A homozygous single-base deletion in MLPH causes the dilute coat color phenotype in the domestic cat. *Genomics* **88**, 698-705, doi:10.1016/j.ygeno.2006.06.006 (2006).
137 Drögemüller, C., Rüfenacht, S., Wichert, B. & Leeb, T. Mutations within the FGF5 gene are associated with hair length in cats. *Animal genetics* **38**, 218-221, doi:10.1111/j.1365-2052.2007.01590.x (2007).
138 Buckley, R. M. *et al.* Werewolf, There Wolf: Variants in Hairless Associated with Hypotrichia and Roaning in the Lykoi Cat Breed. *Genes (Basel)* **11**, doi:10.3390/genes11060682 (2020).
139 Cogne, B. *et al.* Mutations in the Kinesin-2 Motor KIF3B Cause an Autosomal-Dominant Ciliopathy. *Am J Hum Genet*, doi:10.1016/j.ajhg.2020.04.005 (2020).
140 Yu, Y., Creighton, E. K., Buckley, R. M., Lyons, L. A. & Lives, C. A Deletion in GDF7 is Associated with a Heritable Forebrain Commissural Malformation Concurrent with Ventriculomegaly and Interhemispheric Cysts in Cats. *Genes (Basel)* **11**, doi:10.3390/genes11060672 (2020).
141 Kehl, A. *et al.* Molecular characterization of blood type A, B, and C (AB) in domestic cats and a CMAH genotyping scheme. *PLoS One* **13**, e0204287, doi:10.1371/journal.pone.0204287 (2018).
142 Gandolfi, B. *et al.* A Novel Variant in CMAH Is Associated with Blood Type AB in Ragdoll Cats. *PLoS One* **11**, e0154973, doi:10.1371/journal.pone.0154973 (2016).
143 Holl, H. *et al.* A Frameshift Mutation in KIT is Associated with White Spotting in the Arabian Camel. *Genes (Basel)* **8**, doi:10.3390/genes8030102 (2017).
144 Kim, D. Y. & Park, J. H. Genetic Mechanisms of ADPKD. *Adv Exp Med Biol* **933**, 13-22, doi:10.1007/978-981-10-2041-4_2 (2016).
145 Gandolfi, B. *et al.* The naked truth: Sphynx and Devon Rex cat breed mutations in KRT71. *Mamm Genome* **21**, 509-515, doi:10.1007/s00335-010-9290-6 (2010).
146 Mizukami, K., Raj, K. & Giger, U. Feline cystinuria caused by a missense mutation in the SLC3A1 gene. *J Vet Intern Med* **29**, 120-125, doi:10.1111/jvim.12501 (2015).
147 Yoshikawa, R. *et al.* A Naturally Occurring Domestic Cat APOBEC3 Variant Confers Resistance to Feline Immunodeficiency Virus Infection. *J Virol* **90**, 474-485, doi:10.1128/JVI.02612-15 (2016).

148 Clavero, S., Bishop, D. F., Giger, U., Haskins, M. E. & Desnick, R. J. Feline congenital erythropoietic porphyria: two homozygous UROS missense mutations cause the enzyme deficiency and porphyrin accumulation. *Mol Med* **16**, 381-388, doi:10.2119/molmed.2010.00038 (2010).
149 Clavero, S. *et al.* Diagnosis of feline acute intermittent porphyria presenting with erythrodontia requires molecular analyses. *Vet J* **198**, 720-722, doi:10.1016/j.tvjl.2013.10.008 (2013).
150 Asada, H. *et al.* Hepatic copper accumulation in a young cat with familial variations in the ATP7B gene. *J Vet Intern Med* **33**, 874-878, doi:10.1111/jvim.15399 (2019).
151 Yang, L., Chen, J. & Shen, B. Newborn Screening in the Era of Precision Medicine. *Adv Exp Med Biol* **1005**, 47-61, doi:10.1007/978-981-10-5717-5_3 (2017).
152 Robert, C. *et al.* Design and development of exome capture sequencing for the domestic pig (Sus scrofa). *BMC Genomics* **15**, 550, doi:10.1186/1471-2164-15-550 (2014).
153 Foley, J. F. *et al.* Whole exome sequencing in the rat. *BMC Genomics* **19**, 487, doi:10.1186/s12864-018-4858-8 (2018).
154 Yoshihara, M. *et al.* Design and application of a target capture sequencing of exons and conserved non-coding sequences for the rat. *BMC Genomics* **17**, 593, doi:10.1186/s12864-016-2975-9 (2016).
155 Buckley, R. M. *et al.* A new domestic cat genome assembly based on long sequence reads empowers feline genomic medicine and identifies a novel gene for dwarfism. *bioRxiv*, 2020.2001.2006.896258, doi:10.1101/2020.01.06.896258 (2020).
156 Broeckx, B. J. *et al.* Development and performance of a targeted whole exome sequencing enrichment kit for the dog (Canis Familiaris Build 3.1). *Sci Rep* **4**, 5597, doi:10.1038/srep05597 (2014).
157 Broeckx, B. J. G. *et al.* An exome sequencing based approach for genome-wide association studies in the dog. *Sci Rep* **7**, 15680, doi:10.1038/s41598-017-15947-9 (2017).
158 Du, C. *et al.* Explorations to improve the completeness of exome sequencing. *BMC Med Genomics* **9**, 56, doi:10.1186/s12920-016-0216-3 (2016).
159 Thomas, W. B. Hydrocephalus in dogs and cats. *Vet Clin North Am Small Anim Pract* **40**, 143-159, doi:10.1016/j.cvsm.2009.09.008 (2010).
160 Kittleson, M. D., Meurs, K. M. & Harris, S. P. The genetic basis of hypertrophic cardiomyopathy in cats and humans. *J Vet Cardiol* **17 Suppl 1**, S53-73, doi:10.1016/j.jvc.2015.03.001 (2015).
161 Sato, R. *et al.* Epidemiological evaluation of cats associated with feline polycystic kidney disease caused by the feline PKD1 genetic mutation in Japan. *J Vet Med Sci* **81**, 1006-1011, doi:10.1292/jvms.18-0309 (2019).
162 Patterson, D. F., Haskins, M. E. & Jezyk, P. F. Models of human genetic disease in domestic animals. *Adv Hum Genet* **12**, 263-339, doi:10.1007/978-1-4615-8315-8_4 (1982).
163 Zhigachev, A. I., Vladimirova, M. V. & Katser, I. [Phenotypic and genotypic characteristics of russian hairless cats]. *Genetika* **36**, 538-544 (2000).
164 A.I. Zhigachev, M. V. V., and I.J. Kaster. Phenotypic and Genotypic Characteristics of Russian Hairless Cats. *St. Petersburg State Academy of Veterinary Medicine* **4** (2000).

165 Gandolfi, B. *et al.* To the Root of the Curl: A Signature of a Recent Selective Sweep Identifies a Mutation That Defines the Cornish Rex Cat Breed. *PLoS One* **8**, e67105, doi:10.1371/journal.pone.0067105 (2013).
166 Maruyama, H., Hosoe, H., Nagamatsu, K., Kano, R. & Kamata, H. A novel missense mutation in the factor XII gene in a litter of cats with factor XII deficiency. *J Vet Med Sci* **79**, 822-826, doi:10.1292/jvms.16-0602 (2017).
167 Robinson, P. N. *et al.* Improved exome prioritization of disease genes through cross-species phenotype comparison. *Genome Res* **24**, 340-348, doi:10.1101/gr.160325.113 (2014).
168 Collins, F. S. & Varmus, H. A new initiative on precision medicine. *N Engl J Med* **372**, 793-795, doi:10.1056/NEJMp1500523 (2015).
169 Gahl, W. A., Boerkoel, C. F. & Boehm, M. The NIH Undiagnosed Diseases Program: bonding scientists and clinicians. *Dis Model Mech* **5**, 3-5, doi:10.1242/dmm.009258 (2012).
170 Kim, K.-K. *et al.* Whole-exome and whole-transcriptome sequencing of canine mammary gland tumors. *Scientific Data* **6**, 147, doi:10.1038/s41597-019-0149-8 (2019).
171 van Dijk, E. L., Auger, H., Jaszczyszyn, Y. & Thermes, C. Ten years of next-generation sequencing technology. *Trends Genet* **30**, 418-426, doi:10.1016/j.tig.2014.07.001 (2014).
172 Bilgic, O., Duda, L., Sanchez, M. D. & Lewis, J. R. Feline Oral Squamous Cell Carcinoma: Clinical Manifestations and Literature Review. *J Vet Dent* **32**, 30-40, doi:10.1177/089875641503200104 (2015).
173 Wiles, V. *et al.* Retrospective evaluation of toceranib phosphate (Palladia) in cats with oral squamous cell carcinoma. *J Feline Med Surg* **19**, 185-193, doi:10.1177/1098612X15622237 (2017).
174 Wypij, J. M. A naturally occurring feline model of head and neck squamous cell carcinoma. *Patholog Res Int* **2013**, 502197, doi:10.1155/2013/502197 (2013).
175 Veeramachaneni, R. *et al.* Analysis of head and neck carcinoma progression reveals novel and relevant stage-specific changes associated with immortalisation and malignancy. *Sci Rep* **9**, 11992, doi:10.1038/s41598-019-48229-7 (2019).
176 Ferlay, J. *et al.* Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. *Int J Cancer* **136**, E359-386, doi:10.1002/ijc.29210 (2015).
177 Ragin, C. C., Modugno, F. & Gollin, S. M. The epidemiology and risk factors of head and neck cancer: a focus on human papillomavirus. *J Dent Res* **86**, 104-114, doi:10.1177/154405910708600202 (2007).
178 Shaw, R. & Beasley, N. Aetiology and risk factors for head and neck cancer: United Kingdom National Multidisciplinary Guidelines. *J Laryngol Otol* **130**, S9-S12, doi:10.1017/S0022215116000360 (2016).
179 Kawakita, D. & Matsuo, K. Alcohol and head and neck cancer. *Cancer Metastasis Rev* **36**, 425-434, doi:10.1007/s10555-017-9690-0 (2017).
180 Zhou, G., Liu, Z. & Myers, J. N. TP53 Mutations in Head and Neck Squamous Cell Carcinoma and Their Impact on Disease Progression and Treatment Response. *J Cell Biochem* **117**, 2682-2692, doi:10.1002/jcb.25592 (2016).
181 Xu, M. J., Johnson, D. E. & Grandis, J. R. EGFR-targeted therapies in the post-genomic era. *Cancer Metastasis Rev* **36**, 463-473, doi:10.1007/s10555-017-9687-8 (2017).

182 Tannehill-Gregg, S. H., Levine, A. L. & Rosol, T. J. Feline head and neck squamous cell carcinoma: a natural model for the human disease and development of a mouse model. *Vet Comp Oncol* **4**, 84-97, doi:10.1111/j.1476-5810.2006.00096.x (2006).
183 in *Veterinary Oral and Maxillofacial Pathology* 3-12 (2019).
184 McLaren, W. *et al.* The Ensembl Variant Effect Predictor. *Genome Biology* **17**, 122, doi:10.1186/s13059-016-0974-4 (2016).
185 Ritterhouse, L. L. Tumor mutational burden. *Cancer Cytopathol* **127**, 735-736, doi:10.1002/cncy.22174 (2019).
186 Liu, S. H., Shen, P.C., Chen, C.Y., Hsu, A.N., Cho, Y.C., Lai, Y.L., Chen, F.H., Li, C.Y., Wang, S.C., Chen, M. et al. (2020) DriverDBv3: a multi-omics database for cancer driver gene research. Nucleic Acids Res, 48, D863-D870.
187 Chakravarty, D. *et al.* OncoKB: A Precision Oncology Knowledge Base. *JCO Precision Oncology*, 1-16, doi:10.1200/po.17.00011 (2017).
188 Guimaraes, D. P. & Hainaut, P. TP53: a key gene in human cancer. *Biochimie* **84**, 83-93, doi:10.1016/s0300-9084(01)01356-6 (2002).
189 Cancer Genome Atlas, N. Comprehensive genomic characterization of head and neck squamous cell carcinomas. *Nature* **517**, 576-582, doi:10.1038/nature14129 (2015).
190 Renzi, A. *et al.* Prevalence of p53 dysregulations in feline oral squamous cell carcinoma and non-neoplastic oral mucosa. *PLoS One* **14**, e0215621, doi:10.1371/journal.pone.0215621 (2019).
191 Supsavhad, W., Dirksen, W. P., Hildreth, B. E. & Rosol, T. J. p16, pRb, and p53 in Feline Oral Squamous Cell Carcinoma. *Vet Sci* **3**, doi:10.3390/vetsci3030018 (2016).
192 Snyder, L. A. *et al.* p53 expression and environmental tobacco smoke exposure in feline oral squamous cell carcinoma. *Vet Pathol* **41**, 209-214, doi:10.1354/vp.41-3-209 (2004).
193 Rizvi, N. A. *et al.* Cancer immunology. Mutational landscape determines sensitivity to PD-1 blockade in non-small cell lung cancer. *Science* **348**, 124-128, doi:10.1126/science.aaa1348 (2015).
194 Forschner, A. *et al.* Tumor mutation burden and circulating tumor DNA in combined CTLA-4 and PD-1 antibody therapy in metastatic melanoma – results of a prospective biomarker study. *Journal for ImmunoTherapy of Cancer* **7**, 180, doi:10.1186/s40425-019-0659-0 (2019).
195 Eder, T. *et al.* Interference of tumour mutational burden with outcome of patients with head and neck cancer treated with definitive chemoradiation: a multicentre retrospective study of the German Cancer Consortium Radiation Oncology Group. *Eur J Cancer* **116**, 67-76, doi:10.1016/j.ejca.2019.04.015 (2019).
196 Klinakis, A. & Rampias, T. TP53 mutational landscape of metastatic head and neck cancer reveals patterns of mutation selection. *EBioMedicine* **58**, 102905, doi:10.1016/j.ebiom.2020.102905 (2020).
197 Maleki Vareki, S. High and low mutational burden tumors versus immunologically hot and cold tumors and response to immune checkpoint inhibitors. *J Immunother Cancer* **6**, 157, doi:10.1186/s40425-018-0479-7 (2018).
198 Jiang, A. M. *et al.* Tumor Mutation Burden, Immune Cell Infiltration, and Construction of Immune-Related Genes Prognostic Model in Head and Neck Cancer. *Int J Med Sci* **18**, 226-238, doi:10.7150/ijms.51064 (2021).

199 Wapenaar, H. & Dekker, F. J. Histone acetyltransferases: challenges in targeting bi-substrate enzymes. *Clin Epigenetics* **8**, 59, doi:10.1186/s13148-016-0225-2 (2016).
200 Di Cerbo, V. & Schneider, R. Cancers with wrong HATs: the impact of acetylation. *Brief Funct Genomics* **12**, 231-243, doi:10.1093/bfgp/els065 (2013).
201 Sun, W. *et al.* Activation of the NOTCH pathway in head and neck cancer. *Cancer Res* **74**, 1091-1104, doi:10.1158/0008-5472.CAN-13-1259 (2014).
202 Froimchuk, E., Jang, Y. & Ge, K. Histone H3 lysine 4 methyltransferase KMT2D. *Gene* **627**, 337-342, doi:10.1016/j.gene.2017.06.056 (2017).
203 Scharenberg, J. G. *et al.* The inhibitory effect of deoxyadenosine and deoxyguanosine on in vitro lymphocyte function are expressed at different stages of lymphocyte activation. *Adv Exp Med Biol* **195 Pt A**, 541-546, doi:10.1007/978-1-4684-5104-7_91 (1986).
204 Guan, B., Wang, T. L. & Shih Ie, M. ARID1A, a factor that promotes formation of SWI/SNF-mediated chromatin remodeling, is a tumor suppressor in gynecologic cancers. *Cancer Res* **71**, 6718-6727, doi:10.1158/0008-5472.CAN-11-1562 (2011).
205 Lu, W. C. *et al.* miR-31 targets ARID1A and enhances the oncogenicity and stemness of head and neck squamous cell carcinoma. *Oncotarget* **7**, 57254-57267, doi:10.18632/oncotarget.11138 (2016).
206 Luo, Q. *et al.* TRIM32/USP11 Balances ARID1A Stability and the Oncogenic/Tumor-Suppressive Status of Squamous Cell Carcinoma. *Cell Rep* **30**, 98-111 e115, doi:10.1016/j.celrep.2019.12.017 (2020).
207 Zhang, S., O'Regan, R. & Xu, W. The emerging role of mediator complex subunit 12 in tumorigenesis and response to chemotherapeutics. *Cancer* **126**, 939-948, doi:10.1002/cncr.32672 (2020).
208 Syring, I. *et al.* Comprehensive analysis of the transcriptional profile of the Mediator complex across human cancer types. *Oncotarget* **7**, 23043-23055, doi:10.18632/oncotarget.8469 (2016).
209 Genome, M. C. MED12. *My Cancer Genome,* (2017).
210 Platais, C. *et al.* The role of HOX genes in head and neck squamous cell carcinoma. *J Oral Pathol Med* **45**, 239-247, doi:10.1111/jop.12388 (2016).
211 Lee, K. *et al.* HOXB5 acts as an oncogenic driver in head and neck squamous cell carcinoma via EGFR/Akt/Wnt/beta-catenin signaling axis. *Eur J Surg Oncol* **46**, 1066-1073, doi:10.1016/j.ejso.2019.12.009 (2020).
212 Brotto, D. B. *et al.* Contributions of HOX genes to cancer hallmarks: Enrichment pathway analysis and review. *Tumour Biol* **42**, 1010428320918050, doi:10.1177/1010428320918050 (2020).

Supplementary Table 3.1

Cat	TOTAL READS	Unique Reads	Duplicate Reads	% Unique Reads	Mapped Unique Reads	% Unique Reads Mapped	MEAN COVERAGE	MEDIAN COVERAGE
1	215022748	176614831	38407917	82.14	176570410	99.97	373	264
2	145242280	118605692	26636588	81.66	118567772	99.97	199	181
3	83764798	68089930	15674868	81.29	68065004	99.96	211	103
4	164681954	134926578	29755376	81.93	134884666	99.97	232	200
5	192670988	158132174	34538815	82.07	158090364	99.97	331	237
6	66980388	53730990	13249398	80.22	53700410	99.94	88	79
7	237374507	195097488	42277019	82.19	195050455	99.98	416	292
8	55650338	45654534	9995804	82.04	45633063	99.95	76	69
9	170319229	139649516	30669713	81.99	139610319	99.97	289	209
10	259726266	213580145	46146121	82.23	213530500	99.98	458	320
Mean	**159143350**	**130408188**	**28735162**	**81.94**	**130370296**	**99.97**	**267**	**195**
11	59689994	45040110	14649884	75.46	44947330	99.79	73.5	65
12	51213246	42192148	9021098	82.39	42135615	99.87	67.0	61
13	61610752	49451286	12159466	80.26	49384858	99.87	78.6	72
14	66562284	55076186	11486098	82.74	54994060	99.85	90.5	82
15	65035570	54927188	10108382	84.46	54853614	99.87	88.0	80
16	62943656	49742358	13201298	79.03	49665648	99.85	80.4	73
17	77857484	59696054	18161430	76.67	59618838	99.87	92.6	84
18	73192396	59937688	13254708	81.89	59848302	99.85	94.2	85
19	78645830	63544212	15101618	80.80	63467660	99.88	105.7	95
20	49270872	41010968	8259904	83.24	40948904	99.85	63.0	57

21	57911238	48615406	9295832	83.95	48535472	99.84	78.8	71
22	82008038	64579700	17428338	78.75	64483066	99.85	108.2	97
23	92835910	72587612	20248298	78.19	72499058	99.88	103.5	94
24	63399300	52257320	11141980	82.43	52181036	99.85	85.1	77
25	53514212	44731352	8782860	83.59	44666286	99.85	75.1	68
26	46245536	38645300	7600236	83.57	38588802	99.85	60.0	54
27	66297558	55692274	10605284	84.00	55613569	99.86	88.7	80
28	63195078	52759900	10435178	83.49	52685856	99.86	83.6	76
29	54045256	44167266	9877990	81.72	44066794	99.77	70.1	64
30	49086732	40778946	8307786	83.08	40690449	99.78	66.5	60
31	46067970	37352402	8715568	81.08	37288379	99.83	60.1	54
32	61427656	47373818	14053838	77.12	47300394	99.85	77.2	70
33	72513806	56798782	15715024	78.33	56735443	99.89	90.7	83
34	46864618	39837652	7026966	85.01	39789028	99.88	63.7	57
35	57105380	45115204	11990176	79.00	45076149	99.91	62.2	57
36	49270872	41010968	8259904	83.24	40948904	99.85	63.0	57
37	57055464	47567866	9487598	83.37	47493680	99.84	75.9	69
38	72860408	60362784	12497624	82.85	60276637	99.86	97.2	88
39	65352134	52724620	12627514	80.68	52631834	99.82	83.8	76
40	61091136	49625330	11465806	81.23	49566582	99.88	77.7	69
41	53287642	41298322	11989320	77.50	41236732	99.85	65.9	57
Mean	**61853485**	**50145194**	**11708291**	**81.07**	**50071580**	**99.85**	**80**	**72**

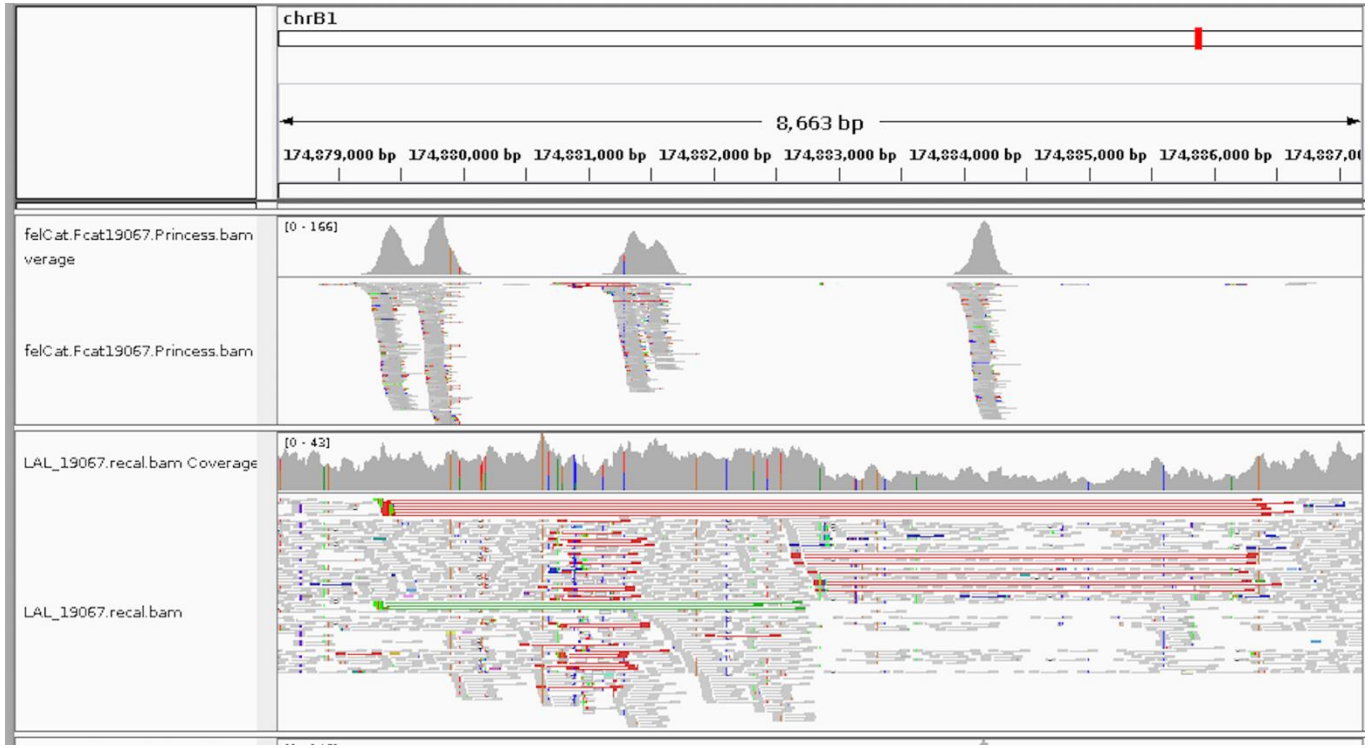
Supplementary Table 3.2

Cat	Median Coverage	% target bp covered >10x	% target bp covered >20x	% target bp covered >30x
1	264	99	100	103
2	181	99	99	98
3	103	99	98	97
4	200	99	99	99
5	237	99	100	101
6	79	99	97	94
7	292	100	101	104
8	69	99	97	92
9	209	99	99	100
10	320	100	101	105
Mean	**195**	**99**	**99**	**99**
31	54	98	95	86
26	54	98	95	87
41	57	98	93	83
34	57	98	96	88
35	57	98	96	89
20	57	98	96	89
36	57	98	96	89
30	60	98	96	90
12	61	98	96	90
29	64	98	96	91
11	65	99	96	90
25	68	98	97	92
40	69	99	96	91
37	69	98	97	93

32	70	99	97	93
21	71	96	91	85
13	72	99	97	94
16	73	99	97	94
28	76	99	97	94
39	76	99	97	94
24	77	99	97	94
15	80	99	97	94
27	80	99	97	94
14	82	99	98	95
33	83	99	98	96
17	84	99	98	95
18	85	97	94	91
38	88	99	98	96
23	94	99	98	97
19	95	99	98	96
22	97	99	98	96
Mean	**72.00**	**98.45**	**96.41**	**91.84**

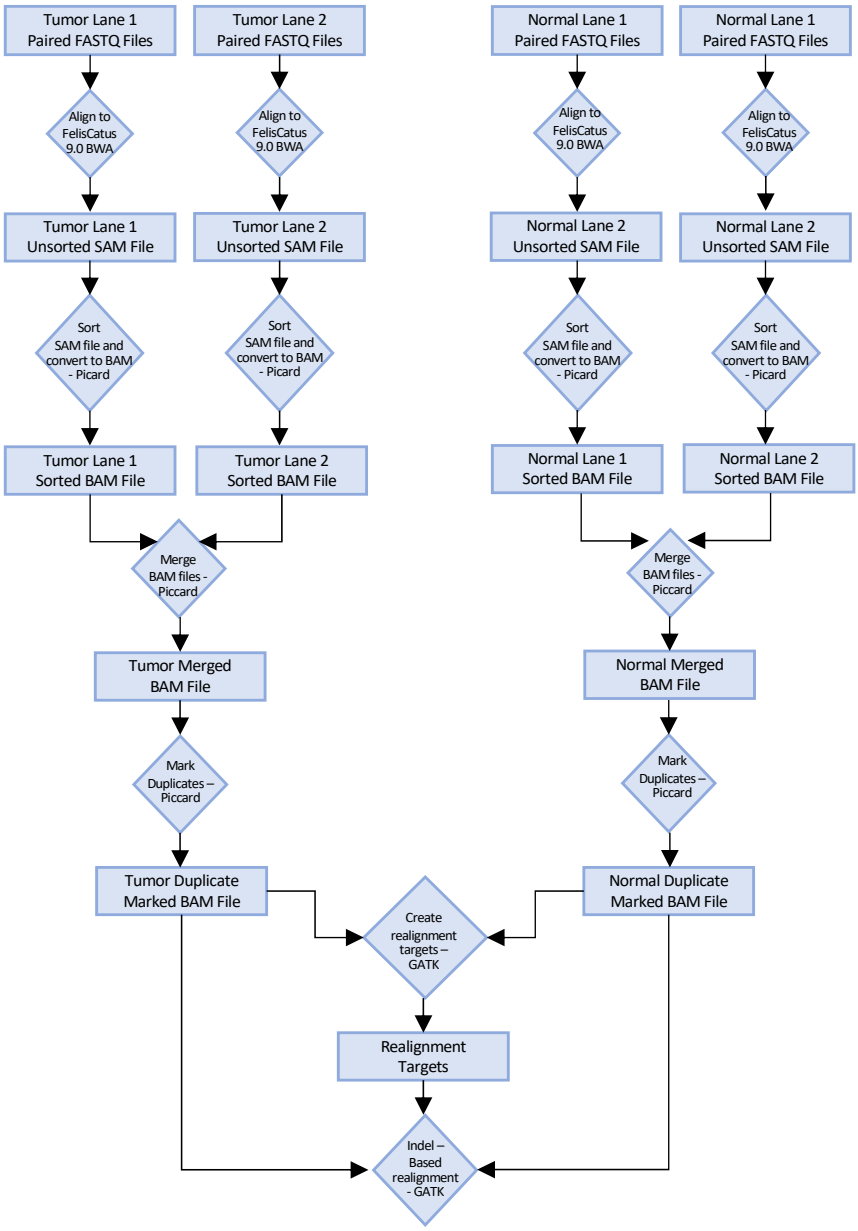
Supplementary Figure 3.1

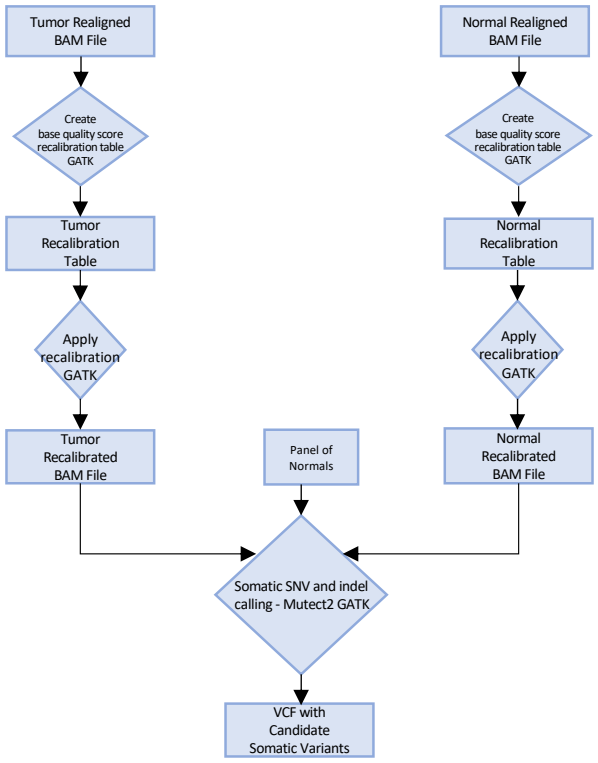
This is a visual of the reads that overlap the dwarfism structural variant in the dwarfism cat sample. The first row is the WES reads showing no evidence of a structural variant in the UDGH gene. Row 2 is the WGS reads showing a deletion and rearrangement (green and red). Thus, showing that WES does not adequately cover structural variants.



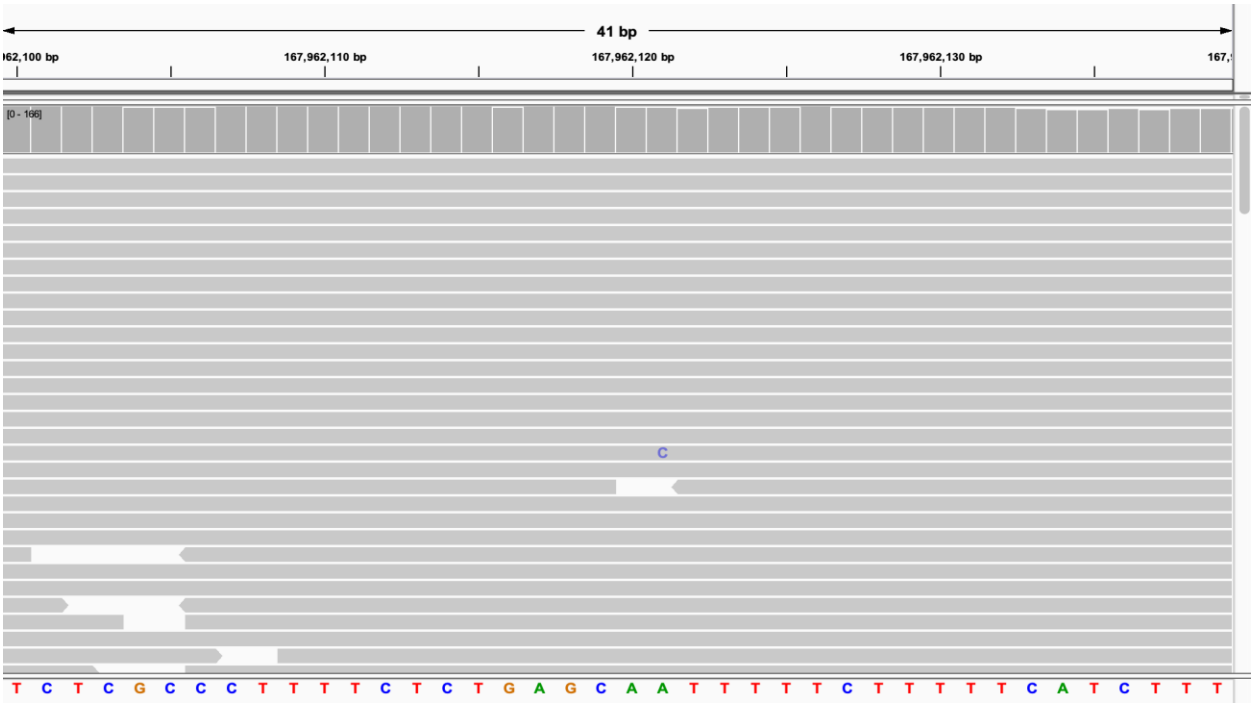
1698 **Supplemental Figure 4.1 Workflow of GATK-Mutect2 Pipeline**

1699




Tumor Realigned BAM File
Create base quality score recalibration table GATK
Tumor Recalibration Table
Apply recalibration GATK
Tumor Recalibrated BAM File
Normal Realigned BAM File
Create base quality score recalibration table GATK
Normal Recalibration Table
Apply recalibration GATK
Normal Recalibrated BAM File
Panel of Normals
Somatic SNV and indel calling - Mutect2 GATK
VCF with Candidate Somatic Variants

Supplementary Figure 4.2. GATK-Mutect2 called a SNV at position 167,962,120 but was determined to be a false call because there was only one variant called out of all the reads.



Supplementary Table 4.1. Tumor mutational burden calculated for all samples

Sample	TMB
7741	8.5
9895	2.6
26903	1.5
60551	1.4
23263	6
24147	2.5