SPONTANEOUSLY OCCURRING CLONAL HEMATOPOIESIS IN THE CANINE

By

Kimberley Sebastian

A THESIS

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Comparative Medicine and Integrative Biology – Master of Science

2022

ABSTRACT

SPONTANEOUSLY OCCURRING CLONAL HEMATOPOIESIS IN THE CANINE

By

Kimberley Sebastian

Clonal hematopoiesis of indeterminate potential (CHIP) is a clinical entity of aging humans that is characterized by cancer-associated mutations in white blood cells, without evidence of overt neoplasia. CHIP has been associated with an increased risk of hematologic cancers, cardiovascular disease, and all-cause mortality. We hypothesized that somatic mutations in specific genes associated with human CHIP would be detectable in the blood of aged dogs not known to have hematologic disorders. DNA from paired germline and whole blood samples from 93 geriatric canine patients affected by solid cancer were subjected to targeted next generation sequencing. Impact of the variants was predicted using Polymorphism Phenotyping version 2 software (PolyPhen-2, Harvard). Clinical and demographic data were extracted from medical records. Somatic variants were detected in peripheral blood of four (4.3 %) female dogs aged 12-15 years. Affected genes were *ASXL1, KIT, SF3B1, TET2, RUNX1*, and *PPM1D*. The variant in *PPM1D* was a nonsense mutation, while the other five variants were single nucleotide non-synonymous variants in protein coding regions of the genes. The single nucleotide variants in *KIT* and *SF3B1* were predicted to be benign, while the variants in *ASXL1, TET2,* and *RUNX1* were predicted to be damaging. A mutant *RUNX1* cell line was designed and constructed with CRISPR technology, and the mutant cells had an increased growth rate compared to controls. These results support the presence of variants in CHIP-associated genes in geriatric canids similar to those observed in people, and the dog represents the first species in which the genetic lesion of CHIP has been documented. Further investigations are needed to confirm the association of this genetic lesion with clinical outcomes.

ACKNOWLEDGMENTS

I would like to express my sincere gratitude to my advisor, Dr. Bonnie Harrington, for the continuous support, motivation, enthusiasm and guidance throughout this thesis. I could not have imagined a better mentor.

In addition to my advisor, I would like to thank the rest of my thesis committee: Prof. Katheryn Meek, Dr. Cynthia Lucidi, and Dr. Leanne Magestro for their encouragement and insightful comments.

I thank my fellow lab mates, Cory Howard and Satyendra Singh, for their patience and their endless willingness to help.

Finally, I would like to thank my family, especially Christopher, for always believing in my dreams.

TABLE OF CONTENTS

LIST OF TABLES v

LIST OF FIGURES vi

Chapter 1: Describing CHIP variants in the dog 1
Introduction 1
Methods 7
Case Selection 7
DNA Extraction and Sequencing 7
Visual Inspection of Variants and Impact Assessment 9
Results 11
Discussion 16

Chapter 2: Generation of *RUNX1* point mutation 21
Introduction 21
Methods 23
sgRNA and primer design 23
Cloning of sgRNA into the PX458 VQR-SpCas9 plasmid 24
Transformation of competent cells 25
Transfection of 17-71 dog cell line 26
Single cell sorting 27
Genotyping of clones 28
Cell growth rate 29
Results 30
Discussion 32

APPENDICES 34
Appendix A. Signalments and germline tissue for study dogs 35
Appendix B. Available CBC data for study dogs 41
Appendix C. Coverage of target genes 46

REFERENCES 52

LIST OF TABLES

Table 1. Twelve canine genes selected for sequencing and investigating for the presence of CHIP. 10

Table 2. Clinical histories of canine CHIP carriers and summary of the mutations identified by targeted next generation sequencing. 13

Table 3. (Appendix) Signalments and germline tissue for study dogs....35

Table 4. (Appendix) Available CBC data for study dogs. 41

Table 5. (Appendix) Coverage of target genes. 46

LIST OF FIGURES

Figure 1. Model of CHIP progression to hematopoietic malignancy..... 2

Figure 2. Schematic *Canis familiaris* genes with predicted deleterious CHIP-associated mutations. 14

Figure 3. Kaplan-Meier survival curves for dogs with and without CHIP-associated variants.....15

Figure 4. Plasmid PX458 VQR-SpCas9. 25

Figure 5. FACS scatter plots of control 17-71 cells and CRISPR 17-71 cells. 28

Figure 6. Sanger sequencing confirmed the V105E-RUNX-1 17-71 cell line. 30

Figure 7. Bar chart showing the mutant cell line had a higher cell count after 72 hours of growth. 31

Chapter 1: Describing CHIP variants in the dog

Introduction

Clonal hematopoiesis (CH) refers to the overrepresentation of blood cells originating from a single clone, the frequency of which increases with age. CH is not a phenomenon inextricably linked to hematopoietic stem and precursor cells (HCPCs), as age-related somatic mutation accumulation also occurs in the skin, lung, and esophagus. Multiple studies have indicated environmental exposure during aging promotes mutant clones.[1,2] The prototypical and first known example of nonneoplastic clonal hematopoiesis was published in 1996 when Busque, et al discovered nonrandom patterns of X-chromosome inactivation (XCI) in geriatric yet healthy women. [3] In 2012, the same group identified mutations in the cancer driver gene *TET2,* thus explaining the skewed hematopoietic activity.[4]

C*lonal hematopoiesis of indeterminate potential* (CHIP) is a recently discovered condition of humans characterized by the expansion of a subpopulation of hematopoietic cells in the blood arising from a single hematopoietic stem cell bearing a cancer-associated gene mutation.[5-8] The current working definition for CHIP extends to include all patients whose white blood cells possess a mutation in a cancer driver gene at a variant allele fraction (VAF) >2 % and whom are not known to have any other hematologic disorder.[9] The frequency of this phenomenon increases with age, and its impact on health is multifold. First, it is a pre-malignant state, with a low, but significant risk of progressing to certain blood cancers.[7,8] Second, it predisposes individuals to non-neoplastic conditions, such as cardiovascular events, by enhancing macrophage activation and inflammation (**Figure 1**).[10] For these reasons there is a 30-40 % increased risk of all-cause mortality in CHIP patients.[5-8]

Three large scale epidemiologic studies of CHIP have been conducted to date, all of which were published in 2014. These studies collectively involved over 30,000 human patients, and all of them used whole exome sequencing of peripheral blood to identify mutation-driven

CH.[6-8] Interestingly, these studies resulted in the identification of a restricted set of genes. In particular, the most frequently mutated genes were the epigenetic modifiers *DNMT3A* (encoding DNA (cytosine-5-)-methyltransferase 3α), *TET2* (encoding tet methylcytosine dioxygenase 2), and *ASXL1* (encoding additional sex combs-like transcriptional regulator 1). Mutations in splicing factors, such as *SF3B1* (encoding splicing factor 3d subunit 1), were also frequent. Of importance, somatic clones increased in frequency with age; by the age of 70 or older, 10 to 20 % of persons harbored mutant clones, in contrast to <1 % of persons younger than age 40.[6-8]

The commonly seen mutations in CHIP are also recurrent drivers of myelodysplastic syndromes (MDS), myeloproliferative neoplasms (MPN), acute myeloid leukemias (AML), and lymphomas.[11,12] Thus, individuals with CHIP have the "first hit" needed for malignant transformation. Indeed, subjects with clonal hematopoiesis had ~10-fold increase of developing a hematologic malignancy which appeared to be associated with larger mutant clones.[6,7]

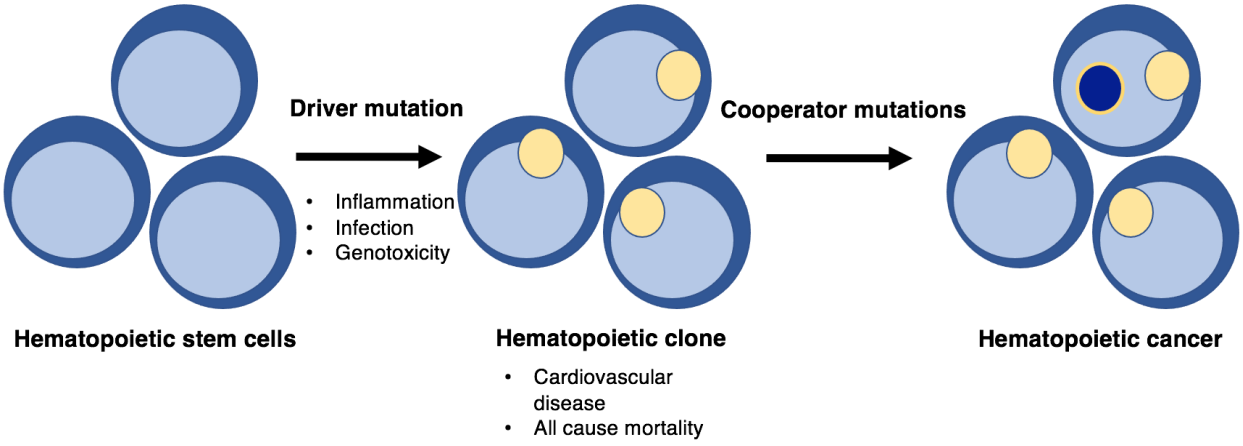

Figure 1. Model of CHIP progression to hematopoietic malignancy.
Mutations in certain genes confer a competitive advantage leading to clonal expansion (yellow circle). If further cooperator mutations (blue circle with yellow rim) occur (e.g., RUNX1, FLT3), hematopoietic cancer results.

Several studies have reported a 30-40 % increase in all-cause mortality, associated with CHIP.[5-7] The risk was related to cardiovascular mortality, and further analysis confirmed the future risk of ischemic stroke and coronary heart disease was more than doubled in persons

with CHIP.[10] The causal association of CHIP with cardiovascular disease has been confirmed using mouse models with loss of *Tet2* in bone marrow stem cells. Hyperlipidemic mice with *Tet2* loss had an increased size of atherosclerotic plaques that could not be explained by blood cell parameters such as low-density lipoprotein (LDL) and cholesterol, alone. The mutant bone marrow-derived macrophages up-regulated many proinflammatory molecules such as the NLRP3 inflammasome-mediated interleukin-β1.[10,13] Few studies have also demonstrated equivocal links with type 2 diabetes[6] and chronic obstructive pulmonary disease,[5] though these conditions may represent reverse causation; it is anticipated that further studies in these areas may reveal more consolidated associations. There is also growing evidence that CHIP may underlie, at least in some part, the phenomenon referred to as "inflammaging," a term used to describe chronic, age-related, low-grade, sterile inflammation in response to stimuli, such as non-self-pathogens, endogenous cell debris, and gut microbiota (quasi-self). This is thought to be the result of activation of the nuclear factor-κβ/inflammasome pathway.[14-17]

Studies have shown that CHIP is predominantly sporadic, and genetic factors are not thought to contribute significantly.[18,19] The aging hematopoietic microenvironment has been implicated with regards to influencing clonality of stem cells,[20] as has a pro-inflammatory milieu.[21,22] Genotoxic exposure has also been implicated in the progression of CHIP, where clones appear to survive under selective pressure post chemotherapy and are a risk factor for therapy-associated AML and MDS.[23] In addition, a subsequent smaller epidemiologic study has shown that humans with a history of solid cancers have an increased frequency of CHIP-associated mutations.[24] The authors of this manuscript described two major hypotheses for the increased risk of CHIP. They hypothesized the increased risk of CHIP was either 1) secondary to DNA-damaging therapies used to treat the primary cancer, or 2) resultant from the same DNA-damaging event that led to the solid tumor.

In spite of these associations, knowledge of factors (other than aging) that predispose to CHIP and its progression is largely incomplete. There is also an unmet need to develop

practical assays for the early detection of CHIP and explore methods of medical intervention that may abrogate progression to more serious disease. To implement these investigations, an appropriate animal model that closely replicates the physiologic consequences of CHIP, including its progression to blood cancer, is essential.

Animal models provide invaluable information in the pursuit of knowledge of many disease states. Rodents, in particular, are heavily utilized due to their small size, short life span, and relatively low cost compared to other animal models. In addition, there are many genetically engineered strains of mice available to study specific diseases and their genome is relatively easy to edit. Mice with loss-of-function mutations in *Tet2* and *Dnmt3a* have been generated by several groups. As discussed earlier, atherosclerotic mouse models undergoing autologous bone marrow transplantation from mice with a loss-of-function mutation in *Tet2* show acceleration of atherosclerosis via activation of the Nlrp3-mediated inflammasome, further supporting the link between cardiovascular disease and CHIP.[10] Mouse models with loss-of-function mutations in *Tet2* or *Dnmt3a* have bone marrow precursor myeloid cells that have a survival advantage and enhanced self-renewal properties.[25-27] In addition, those with *Dnmt3a* loss-of-function mutations have a propensity for both lymphoid and myeloid malignancies.[28,29] Recently, spontaneous clonal hematopoiesis was identified in the bone marrow of elderly mice, albeit at much lower frequencies (~5 fold less) than humans probably due to their much shorter life span and, as such, will likely not provide a model for spontaneous CHIP.[30] Non-human primates (NHPs), most commonly the rhesus macaque, are the animal model genetically most similar to humans; however, they are less frequently used due to their high cost, size, and unpredictable behavior in captivity. As far as the author is aware, there have been no studies reporting spontaneous CHIP in NHPs; however, the idea is conceivable. Macaques with loss-of-function mutations in *TET2* have been generated using CRISPR/Cas9 lentiviral-induced mutations.[31] Long-term studies documenting this group's findings are not yet published.

As there is an inability to model useful spontaneous clonal hematopoiesis in mice and thus far, NPHs, companion dogs (*Canis familiaris*) offer a unique translational model. Unlike laboratory animals, companion dogs share an environment and lifestyle similar to humans. They are exposed to the same pollutants, and similar pathogens and infections which may play important parts in age-related disease. There are over 450 reported diseases in domestic dogs, and 360 of those have an analogous human counterpart.[32,33] As a species, the dog has wide variation among breeds with regards to size, morphology, behavior and lifespan. There are ranges between breeds of more than two-fold in longevity, and lifespan is also strongly related to breed.[34-36] Due to dog breeding practices towards a specific and recognized breed, the genetic variation between dog breeds is high, but within breeds is restrictive.[37] This level of inbreeding makes genetic mapping of complex traits easier and may provide insights with regards to aging and associated diseases in humans.

Spontaneously occurring cancers in canines are similar to those in humans. An additional advantage of the canine model is the high frequency with which this species is affected by blood cancers. Hematopoietic neoplasms in dogs have morphologic similarities to humans and are classified according to human World Health Organization (WHO) criteria.[38,39] Dogs also share similar molecular features of MDS and AML with studies showing recurrent somatic mutations in *JAK2*, *NRAS*, *KRAS*, *C-KIT*, *DNMT3L* and *FLT3* in dogs.[40-42] In addition, dogs have an intact immune system and natural co-evolution of tumor and microenvironment. The domestic dog offers an excellent translational model for human cancers, whilst supporting discovery and development of novel therapeutics for both human and veterinary cancer patients.[43]

Despite recent advances in canine-specific genomics, transcriptomic, and proteomic investigations, much is still unknown regarding the structural and functional relationships between dog and human oncogenes and tumor suppressor genes.[44-46] The recognition of the companion dog in comparative oncology is rapidly increasing with an increased growth of the

scientific literature and various collaborative efforts to enhance the clinical translation by supporting access and enrollment of pets with cancer into clinical trials. An excellent recent review of current comparative oncology clinical trials was recently published by LeBlanc and Mazcko and highlights the importance of novel model development.[43]

Thus far, CHIP has not been reported in the dog, which led to the major goal of this study: to demonstrate that somatic mutations in specific genes associated with human CHIP would be detectable in the blood of aged dogs not known to have hematologic disorders. An additional goal was to design and generate a dog cell line with a described CHIP mutation for use in future functional studies.

Methods

Case Selection

Paired germline and whole blood samples from 93 canine patients were obtained from The Ohio State University Veterinary Biospecimen Repository (IACUC protocol # 2010A0015-R3) (**Appendix A**). To enhance our ability to detect CHIP, patients were selected based on geriatric age (i.e., greater than 8 years) and/or to have a previous history of solid cancer but no concurrent evidence of circulating hematopoietic neoplasia, as assessed by the clinicians at The Ohio State University with a physical exam and when indicated, a complete blood count (CBC). A CBC was available at the time of DNA extraction for 73 dogs (**Appendix B**). Follow up clinical information was available for 29 dogs and was analyzed for concurrent inflammatory diseases and cause of death.

DNA Extraction and Sequencing

DNA was extracted from fresh, frozen germline and whole blood samples using a DNeasy Blood & Tissue Kit (Qiagen, 69504, Frederick, Maryland) according to manufacturer instructions. DNA was quantified using a spectrophotometric method (Nanodrop 2000, Thermo Scientific, Waltham, Massachusetts). Targeted sequencing of twelve prechosen genes (**Table 1**), based on the most common mutations in human CHIP and canine leukemia, was performed using next generation sequencing with the Ion Torrent platform and reagents from Life Technologies (Carlsbad, California). The canine whole genome DNA sequence CanFam 3.1 was downloaded from useast.ensembl.org, and the protein coding regions of 12 candidate genes were extracted using the Ensembl canine genome browser annotation. Primers for candidate genes were custom designed with the Ion Ampliseq Designer (ampliseq.com) and validated *in silico* with the Ion Torrent White Glove team. The library was prepared with Ion AmpliSeq Library kit 2.0 (4475345) with a custom designed panel of AmpliSeq primers (panel

design IAD165539). The target coverage of 166 regions was 96 %, and targets were covered in 404 amplicons, in 2 pools and of a 61.19 kb panel size (**Appendix C**); regions that lacked coverage were not in known hotspots. To run multiple samples on the same chip, we used IonExpress (Thermo Scientific, Waltham, Massachusetts) barcode adapters (kit #4471250 and #4474009). DNA was amplified on a GeneAmp PCR system 9700 Dual 96-well thermal cycler from Applied Biosystems (Thermo Scientific, Waltham, Massachusetts), using the following protocol: Initial hold 99 °C for 2 minutes, followed by 21 cycles at 99 °C for 15 seconds, 60 °C for 4 minutes, and a final hold 10 °C infinity. The PCR product was purified with Agencourt AMPure XP kit (A63881 Beckman Coulter, Indianapolis, Indiana). The library was quantified using real time PCR with Ion Library TAQMAN Quantitation kit 44688022 on an Applied Biosystems ViiA7 Real Time PCR System (Thermo Scientific, Waltham, Massachusetts) instrument to allow for optimal final dilution of the library for template preparation on OneTouch OT2 instrument (Thermo Scientific, Waltham, Massachusetts) with Ion 540 Kit OT2 (A27753). The ion sphere particle enrichment and purification were performed on Ion OneTouch2 ES. Purified ion sphere particles were sequenced on Ion Torrent Personal Genome Machine using an S5 instrument with 540 Kit OT2 (A27753) and 540 Chip Kit (A27766). Data were collected and analyzed using Torrent Server (4462616) with Torrent 5.6.0. Final analysis of the sequence data was performed using a combination of software Torrent Variant Caller v.5.6.8-1 (ThermoFisher, Waltham, Massachusetts) and Integrated Genomics Viewer 5.01 (Broad Institute, Cambridge, Massachusetts). Because canine polymorphisms are poorly characterized, we used pairwise comparison between germline derived DNA and blood derived DNA for every animal tested in our study. Positive CHIP calls were made only for variants that were detected in blood derived DNA and were absent from, or present at a much higher frequency than, matching germline DNA. The Canfam 3.1 reference sequence was used for analysis. The entire length of sequences was reviewed manually using these programs to assess for deviation from reference sequence and to evaluate the quality of sequence and the depth of coverage.

Variants and impact were annotated using snpEff v4.3t based on Ensembl CanFam 3.1.86 annotation. Results from all samples were integrated using in-house shell and R scripts. Ion Torrent run summaries were evaluated for loading density and quality of sequence, and in all cases, the majority of the reads were usable. Low quality sequence, when present, was discarded prior to analysis. Greater than 90 % of reads were mapped to on target regions for all dogs, and the average depth of coverage ranged from approximately 1,000 to 21,000. All synonymous variants and variants in non-protein coding regions were discarded.

Visual Inspection of Variants and Impact Assessment

Remaining variants were visually inspected in the Integrated Genomics Viewer 5.01 (Broad Institute, Cambridge, Massachusetts). Variants were discarded if the variant allele frequency was <2 %, the coverage at the variant location was less than 10 reads, the QV (Phred) score was less than 25, the variant was an insertion or deletion of a specific nucleotide in a string of at least 6 bases of that specific nucleotide (homopolymer region), the variant occurred in all samples, there were many visible variants surrounding the variant in question, or the variant was within 20 base pairs of the end of a read. The impact of the variants was predicted using Prediction of Functional Effects of Human nsSNPs software (PolyPhen-2, Harvard, Cambridge, Massachusetts). Sequencing files are publicly available in the National Center for Biotechnology Information Sequence Read Archive under the BioProject ID number PRJNA789308.

Table 1. Twelve canine genes selected for sequencing and investigating for the presence of CHIP.
Check marks indicate reasons for gene selection and double check marks denote the most frequently mutated genes in human CHIP.

Candidate Gene	Common in Human CHIP	Documented in Canine Cancer
ASXL1	✓✓	
DNMT3A	✓✓	
IDH1	✓	
IDH2	✓	
JAK2	✓✓	✓
KIT	✓	✓
PPM1D	✓	✓
RUNX1	✓✓	
SRSF2	✓	
SF3B1	✓	
TET2	✓✓	✓
TP53	✓	✓

Results

Of the 93 dogs included, 43 were male castrated, 4 were intact male, 42 were spayed female, and 4 were intact female; they ranged in age from 5 to 16 years (median = 12 years). Twenty-seven breeds, comprising 33 mixed breed dogs and 60 pure breeds were affected by 22 types of solid cancers. Four out of 93 dogs (4.3 %) carried variants in CHIP-associated genes in the blood (**Table 2**). Two of the dogs carried two variants. All dogs were female and ranged in age from 12-15 years. Previous solid cancer diagnoses were oral squamous cell carcinoma, thyroid carcinoma, soft tissue sarcoma, and osteosarcoma. The latter three diagnoses were made on the same day as DNA collection. The dog diagnosed with oral squamous cell carcinoma had DNA collected 480 days post cancer diagnosis. This dog received cyclophosphamide chemotherapy and radiotherapy prior to collection of the sample in which CHIP variants were detected. One dog received carboplatin chemotherapy only, and this occurred after blood collection for sequencing.

The mutations identified were annotated as either moderate or high impact, and allele frequencies ranged from 8 % - 67 %. Five single non-synonymous nucleotide variants were identified, three of which were predicted to have deleterious effects. A truncating mutation affecting *PPM1D* was also identified. Schematical representation of genes with predicted deleterious effects is shown in **Figure 2**.

A single dog carried mutations in both *ASXL1* and *KIT*. The N949G substitution in *ASXL1* occurred in exon 12, the most frequently affected exon in human CHIP,[6-8] while the M234V substitution in *KIT* occurred in exon 4, a locus which is not commonly affected in cases of human or canine cancer.[6-8] Co-occurrence of variants in *TET2* and *SF3B1* was noted in another dog. In a pattern similar to the previous dog, the *TET2* S15N substitution with predicted deleterious effects occurred in exon 3, which is consistent with loci of human variants, while the *SF3B1* E402G substitution affected exon 10, which differs from cases of human CHIP where

exon 14 and 15 are affected.[6-8] The two remaining variants in *RUNX1* and *PPM1D* each occurred in one dog. The *RUNX1* variant was predicted to alter splicing and lead to a V105E substitution in the anti-parallel strands of the β-sheet in the *Runt* domain; this point mutation was in close proximity to published human mutations that result in the diminished heterodimerization within the β-subunit and, therefore, reduced function of the gene.[47] The *PPM1D* mutation resulted in a stop codon at residue K535* of exon 6, also consistent with human CHIP.[48] *PPM1D* variants are often acquired secondary to cancer therapy;[49] it is unknown if this dog had received prior chemotherapy at the time of DNA collection.

Progression to blood cancer was not observed in any of the dogs for which follow-up was available (29 dogs out of 93). The median survival for dogs carrying variants in CHIP-associated genes was 128 days compared to 270 days for dogs without variants (not significant, **Figure 3**).

Table 2. Clinical histories of canine CHIP carriers and summary of the mutations identified by targeted next generation sequencing.

All dogs were aged 12 years or older and had been previously diagnosed with a solid cancer. Two dogs received radiotherapy and chemotherapy or chemotherapy, alone. Six unique mutations were identified, and their functional impact was scored using PolyPhen II software (Harvard), where a score of 0 is suggestive that the mutation is tolerated and a score of 1.0 is suggestive the mutation is deleterious.

	Case 1		Case 2		Case 3	Case 4
Sex	FS		FS		FI	FI
Age	15		12		12	12
Breed	Mixed		Mixed		Saluki	Border collie
Other medical conditions	Pollakiuria, metastatic pulmonary SCC		None		Food atopy, microfilariasis, pyometra, mammary masses	Reactive histiocytosis, ocular lipid precipitates, epistaxis
Cancer diagnosis	Oral SCC		Thyroid carcinoma		STS	OSA
Cancer treatment	Chemotherapy, RT, surgical excision, lung lobectomy		None		Surgical excision	Chemotherapy, limb amputation
Outcome	Euthanized 483 days post cancer diagnosis due to post lung lobectomy complications		Euthanized 1 day post cancer diagnosis due to QoL concerns		Lost to follow up	Euthanized 128 days post cancer diagnosis due to QoL concerns
Gene mutated	ASXL1	KIT	SF3B1	TET2	RUNX1	PPM1D
Type of mutation	Missense	Missense	Missense	Missense	Splice region variant (suspect)	Stop codon
Exon	12	4	10	3	5	6
Nucleotide position	chr24:21801101:A:G	chr13:47148941:A:G	chr37:6937498:T:C	chr32:26104689:G:A	chr31:30254159:T:A	chr9:35857110:A:T
VAF (%)	16	12	62	67	34	8
Codon position	p.Asp949Gly	p.Met234Val	p.Glu402Gly	p.Ser15Asn	p.Val105Glu	p.Lys535*
PolyPhen II prediction	Deleterious (0.901)	Tolerated	Tolerated	Deleterious (1.000)	Deleterious (1.000)	NA

FS, female spayed; FI, female intact; SCC, squamous cell carcinoma; RT, radiotherapy; dx, diagnosis; QoL, quality of life; STS, soft tissue sarcoma; OSA, osteosarcoma; NA, not applicable as the PolyPhen II software only analyzes SNPs

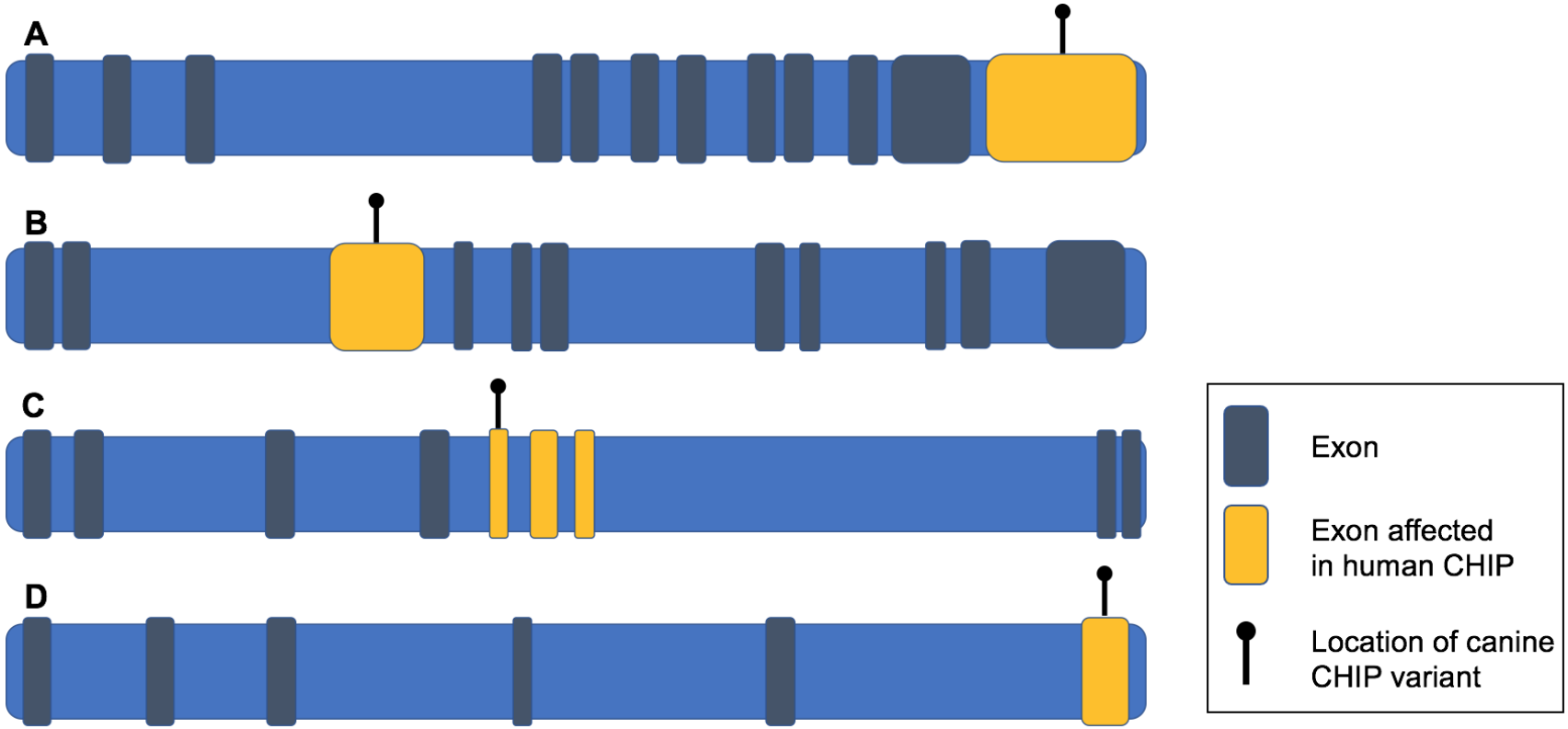

Figure 2. Schematic *Canis familiaris* genes with predicted deleterious CHIP-associated mutations.
The diagrams represent *ASXL1* **(A)**, *TET2* **(B)**, *RUNX1* **(C)**, and *PPM1D* **(D)** genes. Grey vertical bars represent exons. Yellow exons show those most commonly affected in human CHIP. Dot and line represent the variant found in this cohort of dog

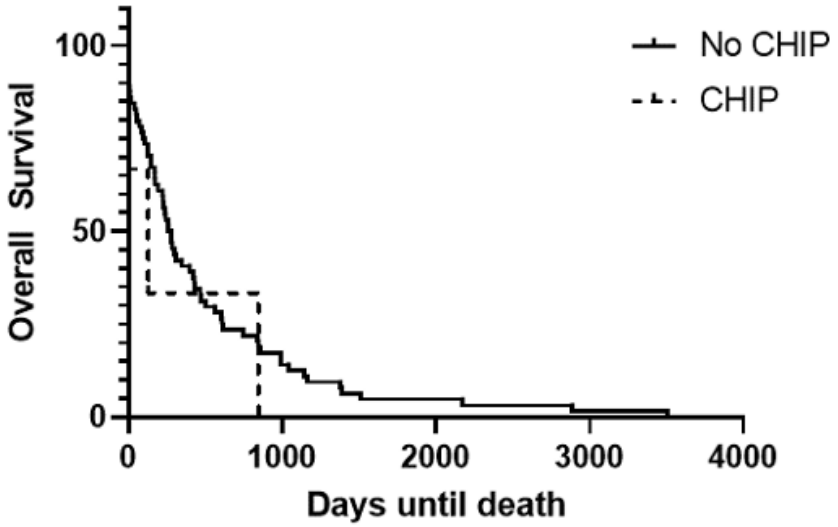

Figure 3. Kaplan-Meier survival curves for dogs with and without CHIP-associated variants.
The median survival for dogs with CHIP was 128 days, compared to 270 days without CHIP. (Three dogs with CHIP and 29 dogs without CHIP.) Using the Mantel-Cox test, there was no significant difference between survival of the two groups (p = 0.5177).

Discussion

Variants in CHIP-associated genes were present in this cohort of dogs with a history of solid cancers but no known evidence of hematologic disorders. Of the six variants identified, two (33%) occurred in the second and third most commonly mutated genes in human CHIP (*TET2* and *ASXL1*), and three (50%) occurred in the top 10 most frequently mutated genes in human CHIP.[6-8] The genomic loci of these variants were also similar to humans; the three single nucleotide variants predicted to have deleterious effects and the truncating mutation in *PPM1D* all occurred in exons and domains that are frequently affected in cases of human CHIP.[6-8] In contrast to publications involving human CHIP, none of the dogs harbored mutations in *DNMT3A,* the most commonly mutated gene in human CHIP.[5-8] But this is not surprising as mutations in *DNMT3A* have not been detected in canine AML. Taken together, these data suggest some similarities between canine and human CHIP, though the frequencies with which certain genes are mutated may vary between the two species. It is also likely that variants exist in additional genes not sequenced in the current study. A larger sample size and more complete sequencing are needed to determine the true frequency with which certain genes are mutated, as well as identify the most commonly mutated genes in the dog.

In humans, mutations in *DNMT3A* and *TET2* are thought to convey a competitive survival advantage and enhanced self-renewal.[7,8] Whereas, mutations in *PPM1D* allow entry into the cell cycle despite the presence of DNA damage, conferring the stem cell advantage in the setting of cytotoxic drugs.[9] The effects of other driver mutations, such as *SF3B1* and *ASXL1* are currently unknown. Mutations in *KIT* and *RUNX1*, thought to be cooperating mutations in CHIP, are commonly associated with MDS and AML.[10,11]

Mechanisms of clonal expansion in CHIP are largely unknown; cells with mutations in *PPM1D* (encoding Mg^{2+}/Mn^{2+}-dependent protein phosphatase 1D), a regulator of p53, can enter the cell cycle despite the presence of DNA damage, conferring the stem cell advantage in the setting of

cytotoxic drugs.[48] A study has recently shown that treatment with a *PPM1D* inhibitor reverses a chemotherapy-resistance phenotype and selectively kills *PPM1D* mutant cells.[49] Truncating mutations typically occur in exon 6 in humans, and the affected dog in this study also had a truncating mutation in exon 6.

DNMT3A and *TET2* have opposite biochemical functions where the *DNMT3A* enzyme is responsible for de novo methylation of the fifth position in cytosine bases of DNA,[50] and *TET2* is one of three enzymes responsible for demethylation via oxidation of 5-methylcytosine to 5-hydroxymethylcytosine.[51,52] Loss-of-function mutations in either of these genes in mouse models lead to a competitive survival advantage with enhanced self-renewal properties, and also propensity towards leukemia, with cooperating mutations.[25-28] Of note, *TET2* has also been identified as a mediator of transcriptional regulation for inflammatory cytokines, such as interleukin-6 (IL-6). Normally, *TET2* recruits histone deactylase 2 (HDAC2) to deacetylate IL-6, repressing its transcription, and thus IL-6 levels.[53] This is an important step in the termination of an inflammatory response in macrophages and dendritic cells. A recent study showed that upregulated expression of IL-6 from hematopoietic stem and progenitor cells (HSPCs) in *Tet2*-knockout mice led to apoptotic resistance, suggesting that *Tet2*-mutated clones may propagate in an inflammatory bone marrow microenvironment.[54,55] In humans, most commonly nonsense or frameshift mutations occur before, occasionally within, the catalytic domain, and missense mutations or in-frame deletions occur in the catalytic domain. All of these lead to inactivation of *TET2.* [56,57] In our study, a missense mutation occurs in the non-catalytic domain.[6]

ASXL1 is an epigenetic modifier, like *DNMT3A* and *TET2*, and is involved in multiple histone modifications which suggests its function is of a scaffolding protein. Most mutations that have been detected in CHIP are frameshift or nonsense mutations of the exon 12.[6-8] Deletion of *ASXL1* or mutant *ASXL1* in mice alters histone modifications and ultimately results in myeloid transformation.[58] The mice also show decreased functioning of HSPCs which is unexpected to result in a clonal advantage.[59] Somatic mutations in *ASXL1* are frequently detected in various

types of myeloid malignancies such as MDS,[60,61] chronic myelomonocytic leukemia (CMML),[62] MPN,[63] and AML.[61,64,65] The variant detected in this study occurs in the *Asx* homology domain which studies in *Drosophila* show is indispensable for histone modifications by Trithorax group (TrxG) and Polycomb group (PcG) proteins.[66]

RUNX1 is a sequence-specific DNA binding protein that requires its non-DNA binding partner, core binding factor beta (CBFβ). A translocation of *RUNX1*, t(8;21)(q22;q22) results in acute myeloid leukemia-1. Other translocations, important in AML, involve those of the genes that encode CBFβ (inv(16)(p13;q22) and t(16;16)(p13;q22)) and are referred to as CBF-AML.[67] Point mutations in *RUNX1* have been found in *de novo* and therapy-related AML, MDS, CMML and acute lymphocytic leukemia (ALL).[68,69] Germline mutations are associated with the autosomal dominant “pre-leukemia” syndrome familial platelet disorder with predisposition to AML (FDP/AML). Loss of *RUNX1* in mouse HPSCs confers increased self-renewal activity.[70] Many missense mutations have been identified in the *Runt* domain (exons 3-5) that result in these neoplasms, many of which involves residues at the DNA binding interface.[71] The variant in one of the study dogs also occurred in the *Runt* domain in one of the ten anti-parallel strands of the β-sheet, in particular, β_D.

SF3B1, a component of the spliceosome that is involved in 3’-splice site recognition during pre-messenger RNA processing, is the most commonly mutated gene found in MDS, especially in patients with refractory anemia with ringed sideroblasts (RARS).[60,72,73] In a recent study, *SF3B1* knockdown in human myeloid cell lines resulted in inhibition of cell growth, cell cycle arrest, and impairment of erythroid differentiation.[74] Another study, in knockdown mice showed decreased numbers of HPSC.[75,76] Mutations in *SF3B1* must confer a survival advantage, but how this occurs is still unknown. In human CHIP, point missense mutations most commonly occur in exons 14 and 15.[6-8]

KIT is occasionally mutated in human CHIP[6-8] and often in core binding factor AML (CBF-AML).[77,78] It is a receptor tyrosine kinase (RTK) that has an important role in

hematopoiesis and over expression is common in myeloid leukemias. Mutations in *C-KIT* using human cells in vitro has been shown to increase cellular proliferation.[79] *C-KIT* mutations have also been identified in some dogs with acute myeloid leukemia (affecting exon 12 at codons 815 and 817).[42] Mutations in *C-KIT* are well documented in canine mast cell tumors[80,81] and gastrointestinal stromal tumors.[82] Mutations in human CHIP have been reported in exon 13[6] and the most commonly mutated sites in human AML affect exon 17 at codons 816 and 822,[78,83] and exon 7 at codons 715 and 815.[83,84]

The association of CHIP with inflammatory diseases such as atherosclerosis was not seen in this pilot study for which follow up data was available. Atherosclerosis is an uncommon spontaneous disease in species other than humans and non-human primates.[85] However, other chronic inflammatory diseases in dogs are similar to humans, such as some endocrinopathies, glomerulonephropathies, and osteoarthritis. Findings of inflammatory diseases in 3 of the 4 dogs with CHIP-associated variants were pollakiuria, food atopy, pyometra, and reactive histiocytosis. These conditions may conceivably have been associated with chronic stimulation of the inflammatory response.

All dogs were selected for advanced age and a history of solid cancer which was strongly selective for CHIP. This study was purposefully designed to identify CHIP in the canine population, therefore individuals considered as high-risk for CHIP were chosen to increase the probability that CHIP would be detected in dogs. Future directions are to establish a cohort of healthy and non-healthy dogs to document CHIP-associated mutations over time, as well as further establish concurrent clinical diagnoses associated with inflammatory disease and development or progression of CHIP. Many of our cases had incomplete clinicopathologic data or medical records due to sampling from a referral population where primary care veterinarians assume animal care once tertiary referral clinicians discharge the patients. Ideally, following patients from a primary care setting would be beneficial in recording and management of complete and thorough medical data.

The pathogenicity of the variants found was assessed using a software developed by Harvard University, PolyPhen-II. This software predicts deleterious or tolerated mutations in genes with complex algorithms to compare wild-type and mutant sequences using the sequence, phylogenetic, and structural information available. To prove the predicted deleterious variants are pathogenic, future plans are to assess pathogenicity with *in vitro* techniques.

In conclusion, our results support the presence of variants in CHIP-associated genes in geriatric canids. A major deficiency of this study is the inability to associate CHIP carrier status with clinical outcome, which was largely due to small sample size and the pre-existing diagnosis of solid cancer in all dogs. To justify use of the dog as a model of human CHIP, it will be necessary to document the association of this genetic lesion with clinical disease, such as blood cancer and inflammatory conditions, and future studies will accomplish this. Regardless, this represents the first domestic species in which the genetic lesion of CHIP has been documented, and results of this study may have significant relevance to multiple fields of medicine where this precursor aberration leads to pathogenic diseases. With further clinical investigations, it is possible that the companion dog may become an additional model for CHIP.

Chapter 2: Generation of *RUNX1* point mutation

Introduction

As previously discussed, the variant in *RUNX1* in this cohort was a non-synonymous suspected splice region variant and predicted to have deleterious effects. Errors during the splicing process can lead to improper removal of introns and cause alterations of the open reading frame. As such, this variant was chosen to be replicated in a dog cell line for future mechanistic studies. It was hypothesized that the mutant cell line would have an increased growth rate compared to wild-type controls. Cloning attempts were made to generate CRISPR cell lines with the *TET2* variant mutation observed in the pilot study; however, they were unsuccessful.

RUNX1 is an important transcription factor and is required for precise hematopoiesis; the high incidence of *RUNX1* mutations in multiple types of hematologic malignancies, previously discussed, provides strong support for its role in orderly hematopoiesis. *RUNX1* mutations alone are not associated with CHIP, but are often identified with driver mutations.[6-8] In addition, *RUNX1* mutations are observed in ~8 % of individuals with clonal cytopenia of undetermined significance (CCUS) and are associated with increased evolution to myeloid malignancy.[86-88] Importantly, malignancies with *RUNX1* mutations are more likely to have decreased survival.[69]

A recent publication specifies curation rules for *RUNX1* variant annotation which include DNA binding, heterodimerization with CBFβ, transactivation, and cellular localization.[89] Missense mutation in the runt homology domain is associated with decreased or complete loss of DNA-binding ability; however, it maintains the ability to bind to CBFβ. It is suspected that *RUNX1* mutated clones have a competitive advantage through two major functions. The first is intracellular mutant clone induction of inflammatory signaling pathways and hypoxia-inducible factor-1A (HIF1A) signaling pathways, and the suppression of TP53-related cellular responses

and ribosome biogenesis. The second is the function of mature mutant clones that produce inflammatory cytokines and, therefore, contribute to the inflammatory microenvironment.[90-94]

Exactly how the inflammatory microenvironment promotes hematopoietic malignancies is largely unknown. In addition, the specific inflammatory pathways and where exactly in the multistep process of hematopoiesis the mutation is gained remain to be elucidated. The exact temporal relationship of *RUNX1* mutations with other driver mutations is also incompletely understood and may or may not be related to this inflammatory microenvironment. Looking at the specific inflammatory pathways *RUNX1* mutant cells use to gain a selection advantage and thrive in the microenvironment could lead to the development of therapeutics to abrogate this inflammatory response before full blown malignancy is achieved. Identifying additional factors that predispose to development of hematopoietic malignancy in *RUNX1* mutant cells could potentially identify those patients at a higher risk for developing cancer, and therapeutics could be used proactively, rather than reactively.

RNA-guided Cas9 nucleases from the microbial clustered regularly interspaced short palindromic repeats (CRISPR) adaptive immune system are used as an efficient tool for genomic engineering in eukaryotic cells by specifying a 20-nucleotide targeting sequence within its guide RNA. Here, we describe a system for introducing a point mutation in a canine primitive B-cell lymphoma cell line (17-71) and hypothesize that the mutant cell line will have a growth advantage compared to the non-mutated cell line.

Methods

sgRNA and primer design

The exon 5 genome of the dog (CanFam 3.1) *RUNX1* was downloaded in FASTA format from Ensembl (ensembl.org) and the region surrounding the base pair mutation was selected for design of a single guide RNA (sgRNA). The construct was designed for the PX458 VQR-SpCas9 plasmid (**Figure 4**) as published by the Zhang lab[95] as follows, where N is a nucleotide:

5'- CACCGNNNNNNNNNNNNNNNNNNN -3'

3 - CNNNNNNNNNNNNNNNNNNNCAAA - 5'

The sgRNA was designed using the CRISPOR design tool (crispor.tefor.net) and a suitable sgRNA was chosen. The tool gives a score; the higher the score, the more faithful the sgRNA will be. The score incorporates off-target loci and therefore, a sgRNA can be chosen that does not have any, or minimal, off-target sites. For homologous direct repair (HDR), which is favored for point mutations, the sgRNA is designed to target the Cas9 to cut as close to the point at which the single base change is to be introduced; here the protospacer adjacent motif (PAM) site was designed to cut 3 bp proximal to the cut site, which for the SpCas9 is 5'-NGG. The designed sgRNA were as follows:

5'- **CACCG**TGTCCCTTGGTTGGCTTAGG -3'

3 - **C**ACAGGGAACCAACCGAATCC**CAAA** - 5'

To encourage HDR a single-stranded oligo DNA nucleotide (ssODN) was designed. A ssODN is a repair template that contains the desired base change flanked by 60 nucleotide homology arms, and a silent PAM mutation. The silent PAM mutation is shown here in bold orange (AGG to AAG) which prevents further Cas9 cleavage after the mutant is generated and avoids introduction of an unwanted missense mutation. The ssODN design was as follows:

5’-

CAATGGCATAAACGTTTACAGCATTTCTGATGTCTGCATC**TGTCCCTTGGTTGGCTTAGG**A

AGTGGCTCTGGGGGATGTCCCCGACGGCACTCTGGTCACTGTAATGGCGGGCAATGATG

AA - 3’

Primers were designed flanking the sgRNA genomic target site using the Primer-BLAST tool by the NIH U.S. National Library of Medicine as follows: The sequence of the forward primer was 5’- AAGCAATTAATACACCAATGGCA -3’ and reverse primer 5’- CTGCGGTGGATTTCTAA -3’.

Cloning of sgRNA into the PX458 VQR-SpCas9 plasmid

Each oligo was resuspended to a concentration of 100 μm with DNA-free water. 1 μg of the plasmid (PX458-VQR) (Addgene, Watertown, Massachusetts) was digested with 1 μL BbsI (New England Biolabs, Ipswich, Massachusetts), 2 μL 10X NEbuffer (New England Biolabs, Ipswich, Massachusetts), and 13 μL of distilled water (ddH_2O). The digested plasmid was purified with a QIAquick DNA purifying kit (Qiagen, Hilden, Germany), as per the manufacturer’s instructions. The oligos were phosphorylated as follows: 1 μL forward oligo, 1 μL reverse oligo, 2μL 10X T4 ligation buffer (New England Biolabs, Ipswich, Massachusetts), 2μL ATP (New England Biolabs, Ipswich, Massachusetts), 1 μL ligase (PNK) (New England Biolabs, Ipswich, Massachusetts), 13 μL ddH_2O, and annealed at 37 °C for 30 minutes, followed by 95 °C for 5 minutes, ramped down to 25 °C at 5 °C/minute. The plasmid and phosphorylated and annealed oligo duplex were ligated at room temperature for 10 minutes as follows: 0.5 ng digested plasmid, 37.5 ng oligo duplex (1:200 dilution), T4 DNA ligase reactive buffer (10X) (New England Biolabs, Ipswich, Massachusetts), and ddH_2O to a total volume of 20 μL.

Figure 4. Plasmid PX458 VQR-SpCas9.
The plasmid was used to introduce a DNA double strand break at a PAM site close to the single nucleotide mutation. The PAM site for this plasmid is 5'-NGG. The plasmid was a gift from Feng Zhang (Addgene plasmid # 48138; http://n2t.net/addgene:48138; RRID:Addgene_48138)

Transformation of competent cells

Competent cells (*Escherichia coli* DH5α) were thawed on ice. Twenty μL of the ligation reaction were added to 50 μL of competent cells, mixed by flicking, and incubated on ice for 30 minutes. The cells were heat shocked for 45 seconds in a water bath at 42 °C and the tube was immediately returned to ice for 2 minutes. 1000 μL of LB was added and incubated at 37 °C for one hour with shaking. 90 μL and 10 μL were plated on LB agar-ampicillin plates and incubated

overnight at 37 °C. The following day, 2 or 3 colonies were picked and inoculated into separate 5 mL cultures with LB-ampicillin liquid medium. The cultures were incubated at 37 °C overnight with shaking. The plasmid DNA was isolated from the cultures using QIAprep spin minkit (Qiagen, Hilden, Germany), according to the manufacturer's instructions. The sequence of each plasmid was verified by Sanger sequencing using the U6 sequencing primer (5'-ATACGATACAAGCTGTTAGAGA -3').

Transfection of 17-71 dog cell line

17-71 cells were grown in complete medium (10 % FBS, 1 % of 10,000 IU penicillin-10, 000 μg streptomycin- 29.2 mg/mL glutamine (Thermo Fischer Scientific, Waltham, Massachusetts), RPMI media (Thermo Fischer Scientific, Waltham, Massachusetts)) and kept in an incubator at 37 °C. The day before transfection, approximately 1 x 10^5 cells/well were plated in a 6 well plate with 2 mL/well of complete medium to achieve a desired confluence of 50-80 %.

For electroporation, the cells were washed twice with Opti-MEM medium (Thermo Fischer Scientific, Waltham, Massachusetts), and resuspended at approximately 1 x 10^8 cells/μL. 20 μg of the vector plasmid and 1 μL *of the ssODN (10* μM) and 2 x 10^7 cells were added to a cuvette designed for use with the Gemini BTX electroporator (BTX Harvard Apparatus, Holliston, Massachusetts) and the preset human protocol for T-cells was used (360 V for 1 ms and 1 pulse). The cells were incubated for 48 hours and prepared for flow cytometry cell sorting. Control cells were treated the same way, minus the vector.

Single cell sorting

For single cell sorting of the GFP-positive cells, 96-well plates were prepared with 150 μL/well of conditioned medium (50 % fresh medium and 50 % medium from a 75 cm^2 flask with cells at approximately 70 % confluence, followed by filtration with a 0.2 μm filter). The cells were pelleted at 1000 rpm for 5 minutes. The supernatant was discarded, the cells were washed with flow media buffer (1mM EDTA, 25 mM HEPES pH 7.0, 5 % heat-inactivated FBS and phosphate buffered saline were filtered through a *0.2* μm filter, sterilized, and stored at 4 °C), and centrifuged at 1000 rpm for 5 minutes. The supernatant was removed, the pellet was resuspended in 4 mL of flow media buffer, and the cells were passed through a 70 μm strainer. The tubes were kept on ice and proceeded to fluorescence-activated cell sorting (FACS) (**Figure 5**). The GFP-positive live cells were immediately sorted into single cells and plated into a well in the 96-well plate. The plate was spun at 1000 rpm for 2 minutes and returned to the incubator for 2 weeks at 37 °C. The colonies were checked every 48 hours for 2 weeks to identify cell colonies. Once confluent, the cell colony was transferred to a well in a 12-well plate with 1 mL of complete media. Once the 12-well plate was confluent, the cell colonies were transferred to a 25 cm^2 flask. The cells were passaged into two 25 cm^2 flasks and once the flasks reached approximately 70 % confluence, one flask of cells was frozen, and one flask was used to genotype the clones.

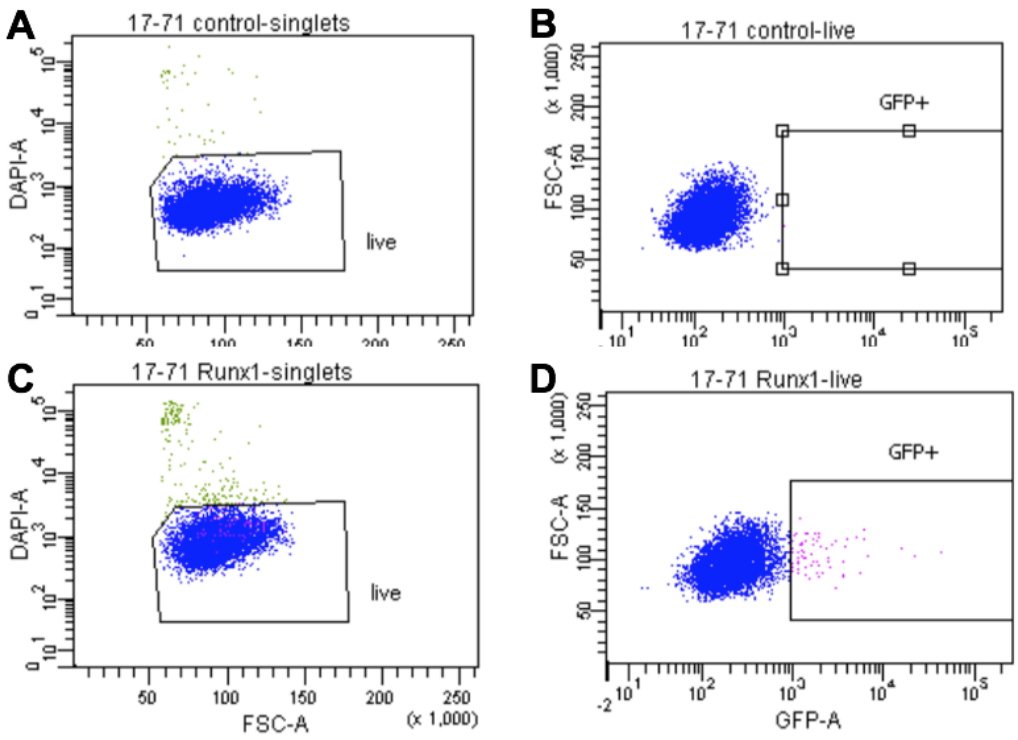

Figure 5. FACS scatter plots of control 17-71 cells and CRISPR 17-71 cells.
The gate is set for the live DAPI-labeled single cells (A and C). Live single cells expressing GFP (i.e., the inserted plasmid) are shown in B and D.

Genotyping of clones

1 μL of each cell colony was added to 10 μL of 10X Master PCR mix (Thermo Fischer Scientific, Waltham, Massachusetts), 4.5 μL of the previously designed forward primer, and 4.5 μL of the reverse primer. DNA was amplified on a GeneAmp PCR system 9700 Dual 96-well thermal cycler from Applied Biosystems (Thermo Fischer Scientific, Waltham, Massachusetts), using the following protocol: Initial hold 94 °C for 2min, followed by 40 cycles at 94 °C for 15sec, 60 °C for 15sec, 68 °C for 20sec, and a final hold 4 °C infinity. The PCR product was purified with GeneJET protocol (Thermo Fischer Scientific, Waltham, Massachusetts). 1 μL of the PCR product and buffer were Sanger sequenced. The mutations were visually inspected in 4peaks (nucleobytes.com).

Cell growth rate

Three separate frozen aliquots of V105E-RUNX-1 17-71 cells and three frozen aliquots of 17-71 cells that underwent the sample procedures outlined above but did not have a mutation, as identified by genotyping, were grown from the same concentration in 25 cm^2 flasks for three days and cells were counted from each flask using an automated cell counter (Countess 3, Thermo Fischer Scientific, Waltham, Massachusetts). Briefly, the cells were mixed to ensure homogeneity, 10 μL of each was added to 10 μL of 0.4 % trypan blue, 10 μL of this mixture was added to the sample well of the Countess cartridge, and the cartridge was loaded into the machine. The average number of cells was recorded over the three flasks for each cell line. A one-tailed t-test with the null hypothesis of the mutant cells will have no difference in growth rate compared to the control and an alternative hypothesis of the mutant cells will have a faster growth rate than control was performed in Word Excel with degrees of freedom = 4, and $\alpha = 0.1$.

Results

Sanger sequencing was repeated twice and confirmed the valine (GTG) to glutamine (GAA) codon mutation in the dog 17-71 cell line (**Figure 6**).


```
# Length: 50
# Identity:      47/50 (94.0%)
# Similarity:    47/50 (94.0%)
# Gaps:           1/50 ( 2.0%)
# Score: 217.0
#
#
#=======================================

EMBOSS_001         1 atgtctgcatctgtcccttggttggcttagGTGGTGGCTCTGGGGGATGT     50
                     ||||||||||||||| |||||||||||||||.|||||||||||||||||
EMBOSS_001        18 ATGTCTGCATCTGT-CCTTGGTTGGCTTAGGAAGTGGCTCTGGGGGATGT     66
```

Figure 6. Sanger sequencing confirmed the V105E-RUNX-1 17-71 cell line.
The codon was replaced from the wide type GTG, encoding valine, to GAA, encoding glutamine.

V105E-RUNX-1 17-71 cells had a statistically significant increased growth rate over 72 hours compared to control 17-71 cells (**Figure 7**). The calculated t value was 4.116 and the critical t value was 2.131. As the calculated t value was higher, we accepted the alternative hypothesis that the mutant cells have a faster growth rate than controls. The calculated p value was 0.007.

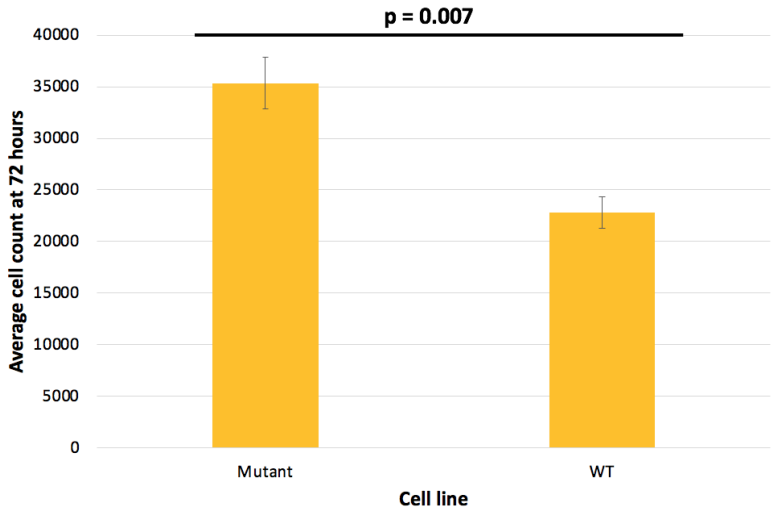

Figure 7. Bar chart showing the mutant cell line had a higher cell count after 72 hours of growth.
Three aliquots from each cell line were grown separately via routine cell culture and automated cell counts were obtained after 72 hours of uninterrupted growth (one tailed t-test, p = 0.007).

Discussion

The mutant V105E-*RUNX1* 17-71 cell line was successfully created, and the mutant cells had a faster growth rate than controls. This potentially indicates that this point mutation confers a competitive survival advantage and is similar to point mutations in human AML and MDS.

RUNX1 point mutations in humans are found in MDS and AML with minimal differentiation (M0), AML-MDS, therapy-associated AML, and ALL.[69,71] The most recent WHO classification of tumors also added a provisional category for AML associated with *RUNX1* mutations highlighting its importance in myeloid malignancies.[96] Point mutations have been reported throughout the whole Runt domain and also the C-terminal moiety where the transactivation domain Is located. Most of the cases are heterozygous, and as such, haploinsufficiency appears to be the basis for the pathogenesis. Interestingly, the regions most affected by mutations correspond to the loop-containing regions responsible for DNA binding; however, the regions are still capable of CBFβ, which may explain the dominant negative effects of these mutants.[97,98] Assessing DNA and CBFβ binding in this cell line would further elucidate its functional effects.

The point mutation was not achieved in a human cell line; ideally, this would be performed in human pluripotent stem cells to better replicate the precursor cells affected by CHIP. In addition, further functional studies as discussed by the ClinGen Myeloid Malignancy Variant Curation Expert Panel[89] would be required to prove definitively that the mutation is deleterious.

The molecular characterization of leukemia in dogs has been limited. Nevertheless, mutational analysis of dogs with acute leukemias has confirmed similar mutations in genes involved in human AML.[40-42] To the best of our knowledge, the prevalence of *RUNX1* mutations has not been assessed in canine leukemia or lymphoma. This study shows a point mutation in a

region similar to those described in human AML and as such, additional studies could potentially focus on identifying similar mutations in this human hot spot region via PCR in canine blood or tissue samples with hematologic malignancies. If the dog proves to frequently have this mutation in certain blood cancers, it could be a useful model to study potential therapeutic targets, such as CBFβ inhibitors,[99] or be used to develop additional small molecule inhibitors of *RUNX1* or CBF.

In conclusion, we document for the first time a likely functional pathologic point mutation variant in *RUNX1* in a dog cell line. If the variant is functional in a human cell line, the dog could prove relevant as an additional animal model for studies of CHIP and hematopoietic malignancies.

APPENDICES

Appendix A. Signalments and germline tissue for study dogs

Table 3. (Appendix) Signalments and germline tissue for study dogs.

ID	Breed	Sex	Age (yr)	Cancer	Other disease	Germline tissue
1	Mixed Breed	FS	15	Squamous Cell Carcinoma	Pollakuria, metastatic pulmonary squamous cell carcinoma	Muscle
2	Mixed Breed	FS	12	Thyroid carcinoma	None	Skin
3	Saluki	F	12	Soft Tissue Sarcoma	Food atopy, microfilariasis, pyometra, mammary masses	Skin
4	Border Collie	F	12	Osteosarcoma	Reactive histiocytosis, ocular lipid precipitates, epistaxis	Intestine
5	Mixed Breed	FS	12	Mast Cell Tumor (grade II) STS (grade I)	Superficial pyoderma, flea bite hypersensitivity, seborrhea sicca, yeast otitis, spinal pain, chronic kidney disease	Skin
6	Mixed Breed	MC	13	Soft tissue sarcoma (Grade I)	Degenerative osteoarthritis, chronic kidney disease	Skin
7	Mixed Breed	MC	12	Soft tissue sarcoma (Hemangiopericytoma; grade I)	Osteoarthritis, periodontal disease, renal disease, nonerosive polyarthropathy, cranial cruciate ligament disease	Skin
8	Golden Retriever	FS	14	Hemangiosarcoma-splenic, liver and pulmonary nodules	Idiopathic epilepsy, hypothyroidism, pyoderma, otitis externa, hypertension	Muscle
9	Vizsla	F	12	Mast cell tumor, gastrointestinal stromal tumor, adrenal mass	Osteoarthritis	Skin
10	Rottweiler	MC	11	Mast Cell Tumor	Osteoarthritis, periodontal disease, chronic kidney disease, hypertension	Skin
11	Mixed Breed	FS	15	Melanoma	NA	Skin
12	Golden Retriever	MC	11	Soft Tissue Sarcoma	Pyoderma, otitis externa, pododermatitis	Skin
13	Mixed Breed	MC	11	Osteosarcoma	Asteroid hyalosis, periodontal disease, mitral endocardosis, osteoarthritis	Skin
14	Labrador Retriever	M	12	Hemangiosarcoma (primary bone)	Otitis externa, osteoarthritis	Skin
15	German Shepherd	MC	11	Apocrine gland anal sac adenocarcinoma	Bilateral hip dysplasia	Skin
16	St. Bernard	MC	7	Osteosarcoma	Osteoarthritis, chronic diarrhea, atopy	Muscle

Table 3. (cont'd)

17	Greyhound	MC	11	Osteosarcoma	Osteoarthritis	Muscle
18	Labrador Retriever	FS	12	Intestinal adenocarcinoma	Chronic nonregenerative anemia	Skin
19	Labrador Retriever	MC	12	Hemangiopericytoma	Osteoarthritis	Muscle
20	English Setter	FS	12	Soft Tissue Sarcoma	Chronic cough	Skin
21	Greyhound	FS	11	Osteosarcoma	Protein losing nephropathy	Skin
22	Labrador Retriever	FS	11	Squamous Cell Carcinoma	Otitis externa, seasonal atopy, osteoarthritis	Skin
23	Miniature Schnauzer	MC	12	Pulmonary carcinoma, papillary	Cystoliths	Skin
24	Mixed Breed	MC	12	Soft tissue sarcoma (Fibrosarcoma grade II)	NA	Skin
25	Lhasa Apso	FS	11	Carcinoma	Seasonal atopy	Skin
26	Golden Retriever	FS	13	Osteosarcoma	Osteoarthritis, hip dysplasia	Skin
27	Bouvier Des Flanders	MC	15	Osteosarcoma Squamous Cell Carcinoma (10/01/13) MCT (6/24/10) Lymphoma (Indolent) (6/8/12 - T cell)	Osteoarthritis, chronic cough, polyneuropathy, pyoderma, cataracts, hypothyroidism	Muscle
28	Labrador Retriever	MC	10	Oral Malignant Melanoma	Cranial cruciate ligament disease	Muscle
29	Beagle	FS	10	Soft tissue sarcoma (Peripheral Nerve Sheath Tumor; grade II/III)	Cranial cruciate ligament disease	Muscle
30	Mixed Breed	MC	12	Hemangiosarcoma	NA	Skin
31	Mixed Breed	FS	14	Thyroid carcinoma	Hypothyroidism, hypertension	Skin
32	Mixed Breed	MC	13	Hemangiosarcoma	Pyoderma, pulmonary hypertension	Muscle
33	Basset Hound	FS	11	Oral Malignant Melanoma	Compensated endocardiosis of mitral valve	Skin
34	Scottish Terrier	FS	12	Transitional Cell Carcinoma	Otitis externa	Skin

Table 3. (cont'd)

35	Greyhound	FS	12	Osteosarcoma	Mitral valve disease, osteoarthritis	Skin
36	Mixed Breed	MC	12	Salivary adenocarcinoma	Brainstem mass	Skin
37	Golden Retriever	MC	12	Soft tissue sarcoma (Peripheral Nerve Sheath Tumor)	T3-L3 spinal pain	Muscle
38	Greyhound	FS	12	Non-angiomatous, nonlymphomatous splenic sarcoma	Glomerulonephropathy, hypertension	Skin
39	Greyhound	MC	11	Osteosarcoma	Fracture of distal humerus	muscle
40	Golden Retriever	FS	12	Hemangiosarcoma	Leukopenia secondary to sepsis	Skin
41	Greyhound	FS	13	Osteosarcoma	Hypothyroidism, periodontal disease, osteoarthritis	Skin
42	Mixed Breed	FS	14	Soft tissue sarcoma (hemangiopericytoma), Osteosarcoma	NA	Skin
43	Mixed Breed	MC	14	Transitional Cell Carcinoma	Non regenerative anemia, keratoconjunctivitis sicca	Skin
44	Mixed Breed	MC	12	Osteosarcoma	Osteoarthritis, periodontal disease	Skin
45	Mixed Breed	MC	10	Carcinoma	Mitral valve degeneration, cystitis, idiopathic epilepsy	Muscle
46	Golden Retriever	MC	13	Multiple Soft Tissue Sarcomas, Metastatic Carcinoma	NA	Skin
47	Cocker Spaniel	MC	14	Chondrosarcoma	NA	Skin
48	Mixed Breed	FS	13	Soft tissue sarcoma (Hemangiopericytoma)	Osteoarthritis	Skin
49	Mixed Breed	MC	14	Prostatic carcinoma	Cranial cruciate ligament disease	Skin
50	Golden Retriever	MC	13	Osteosarcoma	Osteoarthritis	Urinary Bladder
51	Mastiff	MC	5	Osteosarcoma	NA	Skin
52	Whippet	MC	12	Thyroid carcinoma	Protein losing nephropathy	Skin
53	Shar-pei	FS	12	Mast Cell Tumor	Pyoderma	Skin

Table 3. (cont'd)

54	Border Collie	MC	12	Pulmonary Carcinoma	Cranial cruciate ligament disease	Skin
55	Cocker Spaniel	MC	12	Hemangiosarcoma	Anemia, thrombocytopenia, osteoarthritis	Skin
56	Labrador Retriever	FS	12	Osteosarcoma (Mandible)	Pyoderma	Skin
57	Mixed Breed	MC	12	Squamous Cell Carcinoma Soft tissue sarcoma - 2016 (Hemangiopericytoma)	Benign prostatic hyperplasia, periodontal disease, bilateral hip dysplasia	Skin
58	Rottweiler	FS	7	Melanoma	Cranial cruciate ligament disease	Gingiva
59	Irish Terrier	FS	12	Soft tissue sarcoma (Peripheral Nerve Sheath Tumor)- 8/12/11 Mast Cell Tumor -9/26/11	Seizures	Skin
60	German Shepherd	FS	8	Soft tissue sarcoma (Peripheral Nerve Sheath Tumor)	Exocrine pancreatic insufficiency	Muscle
61	Mixed Breed	FS	12	soft tissue sarcoma (Peripheral nerve sheath tumor) , Nonangiomatous, nonlymphomatous splenic sarcoma	NA	Skin
62	Labrador Retriever	MC	12	Mast Cell Tumor	Pyoderma, otitis externa	Muscle
63	Vizsla	FS	12	Histiocytic Sarcoma	NA	Skin
64	Labrador Retriever	FS	13	Histiocytic Sarcoma	NA	Skin
65	Mixed Breed	MC	14	Hemangiosarcoma	Non-regenerative anemia, thrombocytopenia	Skin
66	Miniature Schnauzer	MC	13	Soft Tissue Sarcoma, Hepatocellular carcinoma	Hypertension, gall bladder mucocele	Skin
67	Golden Retriever	MC	10	Maxillary Fibrosarcoma	Dermatitis	Skin
68	Mixed Breed	MC	13	perianal gland adenocarcinoma	Paradoxical vestibular disease	Skin
69	Beagle	MC	9	Carcinoma	NA	Oral mucosa
70	Mixed Breed	FS	12	Histiocytic Sarcoma	Osteoarthritis	Skin
71	Mixed Breed	FS	12	Transitional Cell Carcinoma	Renal disease	Skin

Table 3. (cont'd)

72	Labrador Retriever	FS	13	Soft Tissue Sarcoma	NA	Skin
73	Jack Russel Terrier	MC	12	Mast Cell Tumor	NA	Muscle
74	Mixed Breed	MC	5	Soft tissue sarcoma	Pyoderma, otitis externa, alopecia, pancreatitis	Skin
75	Mixed Breed	M	16	Maxillary fibrosarcoma	Periodontal disease	Skin
76	Mixed Breed	FS	12	Mast Cell Tumor	NA	Skin
77	Golden Retriever	FS	12	Squamous Cell Carcinoma	Peridontal disease, osteoarthritis, anemia, thrombocytopenia	Lip
78	Golden Retriever	MC	12	Mast Cell Tumor	NA	Skin
79	Labrador Retriever	FS	12	Soft Tissue Sarcoma	NA	Tongue
80	Mixed Breed	FS	12	Hepatocellular Carcinoma	NA	Skin
81	Boxer	M	12	Soft Tissue Sarcoma	NA	Skin
82	Greyhound	F	12	Osteosarcoma	NA	Liver
83	Mixed Breed	FS	11	Colon carcinoma, carcinoma of perineal skin	NA	Skin
84	Mixed Breed	FS	12	Mast Cell Tumor	NA	Muscle
85	Boston Terrier	MC	12	Soft tissue sarcoma (Hemangiopericytoma)	NA	Skin
86	Mixed Breed	M	14	Osteosarcoma	NA	Skin
87	ShihTzu	MC	12	Pulmonary carcinoma	Adrenal mass, periodontal disease, hyperadrenocorticism, polycystic kidneys	Muscle
88	ShihTzu	FS	8	Squamous Cell Carcinoma	NA	Muscle
89	Mixed Breed	FS	13	Insulinoma/mast cell tumor	Periodontal disease	Skin
90	Australian Shepherd	MC	13	Mast Cell Tumor	NA	Oral mucosa

Table 3. (cont'd)

91	Mixed Breed	MC	12	Hepatocellular Carcinoma	Protein losing nephropathy, mitral valve disease	Skin
92	Shetland Sheepdog	FS	13	Hemangiosarcoma	Osteoarthritis	Skin
93	Cocker Spaniel	MC	14	Chondrosarcoma	NA	Liver

F, female;
FS, female spayed;
M, male;
MC, male castrated;
NA, not available

Appendix B. Available CBC data for study dogs

Table 4. (Appendix) Available CBC data for study dogs.

ID	Hct %	Hgb g/dL	RBC **x** 10^6 /µL	MCV fL	MCHC g/dL	MCH pg	WBC **x** 10^3 /µL	Platelet count **x** 10^3 /µL	Microscopic findings
1	49%	16.9	7.23	68	34.5	23.4	10.6	541	None noted
3	55.9	19	8.62	64.8	34	22	8.14	208	None noted
4	51.3	17.1	7.56	67.9	33.3	22.6	9.41	179	None noted
5	39	13.4	5.4	72	34.8	NA	12.4	400	Reactive lymphocytes Slight anisocytosis and poikilocytosis
6	44	15.8	6.5	69	35.5	NA	11.9	488	Occasional reactive lymphocytes Slight anisocytosis Rare polychromasia, Slight poikilocytosis
7	39	13.8	6.1	64	35.3	NA	17	317	Occasional reactive lymphocytes Slight anisocytosis, polychromasia, and poikilocytosis
8	27.3	9.6	4.26	64.1	35.2	22.5	19.75	190	None reported
10	45	14.6	6.9	65	32.7	NA	8.1	431	Reactive lymphocytes Slight anisocytosis and poikilocytosis
11	40	14.2	6	66	35.9	NA	14.8	663	Occasional reactive lymphocytes Slight anisocytosis, polychromasia, and poikilocytosis Occasional target cells
12	41.8	14.4	6	69.7	34.4	24	12.2	252	None reported
13	39	13.5	6.1	64	35	NA	6.2	221	Occasional reactive lymphocytes Slight anisocytosis and poikilocytosis
14	42	13.9	5.9	71	33.2	NA	5.7	177	Reactive lymphocytes Slight anisocytosis Rare polychromasia Moderate poikilocytosis Rare acanthocytes
15	43.5	14.9	6.68	65	34.2	22.2	8.6	393	None reported
16	44.8	14.8	6.61	67.8	33	22.4	5.31	148	None reported

Table 4. (cont'd)

18	32	11.6	5	65	35.6	NA	6.1	104	Rare reactive lymphocytes Slight anisocytosis and poikilocytosis Rare target cells
19	39.9	11.9	6.24	64	29.7	19	2.37	277	3+ anisocytosis, 1+ hypochromasia
20	45	15.4	6.2	72	34.2	na	10	267	Slight anisocytosis and poikilocytosis
21	52.2	17.8	7.3	71.5	34.1	24.4	3.56	179	None reported
22	43	15.5	6.7	65	35.8	na	5.9	253	Slight anisocytosis and poikilocytosis
24	46	15.6	7.4	62	34.1	na	9.1	316	Slight anisocytosis and poikilocytosis
25	46.6	16.3	7.08	65.8	35	23	10.54	256	None noted
27	38	13	5.4	70	34.3	na	11.5	595	Slight anisocytosis and poikilocytosis
32	40	14.4	6.3	63	36.5	NA	15.3	122	Rare reactive lymphocytes Slight anisocytosis, poikilocytosis and polychromasia
34	45	16.2	6.9	66	36	NA	10.2	399	Reactive lymphocytes Slight anisocytosis and poikilocytosis Rare polychromasia
35	53.4	18.9	7.42	72	35.4	25.5	6.6	249	None reported
37	44	14.9	6	73	34.1	NA	7.7	219	Reactive lymphocytes Slight anisocytosis and poikilocytosis
38	43.4	15.7	6.26	69.3	36.2	25.07	12.79	317	None reported
39	55.5	18.4	7.43	74.6	33.1	24.72	10.78	301	None reported
40	35	11.7	5.11	59	33.4	22.9	5.4	197	None reported
41	56.4	19.6	7.35	76.8	34.8	26.72	6.13	236	None reported
42	47.2	15.5	6.71	70.3	32.8	23.1	6.1	469	Slight anisocytosis and poikilocytosis
43	36	12	5.1	70	33.8	NA	4.4	225	None reported
44	48	16.5	7.5	64	34.3	NA	11.3	396	Reactive lymphocytes Slight anisocytosis, poikilocytosis and polychromasia
45	33	11.5	4.85	68	34.8	23.7	16.23	660	Slight anisocytosis and poikilocytosis
46	31	11.2	4.5	68	36.3	NA	8.4	191	Slight anisocytosis and poikilocytosis
47	40	13.7	6.2	65	34.2	NA	9.3	529	Occasional reactive lymphocytes Slight poikilocytosis and occasional target cells

Table 4. (cont'd)

48	31	11	4.9	64	35.7	NA	7.6	365	Occasional reactive lymphocytes Rare to occasional acanthocytes, schistocytes, and keratocytes Moderate anisocytosis, poikilocytosis, and rare polychromasia
49	43	16.7	6.9	62	38.8	24.2	24.36	405	None reported
51	34	12	5.5	62	35	NA	27.1	300	Slight anisocytosis, polychromasia, and poikilocytosis
53	41.9	14.4	6.54	64.1	34.4	22	8.75	146	None noted
55	19.76	6.8	3.02	65	34.5	22.5	43.98	41	Spherocytes
56	50	17.1	7.6	65	34.6	NA	15.2	389	Occasional reactive lymphocytes Rare polychromasia
57	47	15.5	6.5	72	33	NA	9.2	351	Reactive lymphocytes Slight anisocytosis and poikilocytosis
58	38.8	13.4	5.83	66.6	34.6	23	7.51	437	Slight anisocytosis and poikilocytosis
59	37.3	12.6	5.14	72.6	33.8	24.5	16.46	485	Slight anisocytosis and poikilocytosis
60	42.5	14.8	6.58	64.6	34.8	22.5	7.3	171	Slight anisocytosis and poikilocytosis
61	34	11.9	5.6	61	34.8	NA	15.5	774	Reactive lymphocytes Slight anisocytosis, poikilocytosis Target cells
62	45	15.6	7.5	61	34.5	NA	7.3	239	Reactive lymphocytes Slight anisocytosis, poikilocytosis and polychromasia Occasional target cells
63	35	12.3	5	71	34.8	NA	10.3	178	Slight anisocytosis Rare poikilocytosis and target cells
64	40.7	14	6.18	65.9	34.4	22.7	6.24	278	Slight anisocytosis and poikilocytosis
65	29	10.3	4.6	63	36	NA	15.7	51.3	Occasional reactive lymphocytes Slight anisocytosis and poikilocytosis

Table 4. (cont'd)

66	36	11.8	4.9	75	32.6	NA	10.8	734	Reactive lymphocytes Slight anisocytosis, polychromasia, poikilocytosis
67	46.2	16	7.15	64.6	34.6	22.4	10.15	398	Slight anisocytosis and poikilocytosis
68	31	11.1	4.7	70	34.1	NA	109.2	469	Slight anisocytosis and poikilocytosis
69	32.8	11.2	5.44	60.3	34.1	20.6	2.97	598	None reported
70	35	12.9	5.8	60	37.2	NA	17.2	565	Reactive lymphocytes Slight anisocytosis and occasional polychromasia
71	34.7	11.8	4.9	70.8	34	24.1	6.75	227	None reported
73	34	11.8	5	68	35	NA	8.7	856	Rare reactive lymphocytes Slight anisocytosis, poikilocytosis and polychromasia
74	54	18.8	7.3	74	34.9	NA	14.5	215	Rare reactive lymphocytes Slight anisocytosis and poikilocytosis and rare polychromasia
76	38	13.6	5.7	66	36.1	NA	21.4	566	Slight anisocytosis and poikilocytosis and target cells present
77	43	14.7	6.4	67	34	23.1	9.8	227	None reported
78	34	11.8	4.96	68.5	34.7	23.8	7.18	330	Slight anisocytosis and poikilocytosis
79	40	13.6	5.8	69	34.1	NA	11	329	Occasional reactive lymphocytes Slight anisocytosis and poikilocytosis Rare target cells
80	41	13.5	6.4	64	33.2	NA	7.3	392	Reactive lymphocytes Slight anisocytosis and poikilocytosis
81	43	14.4	5.8	74	33.6	NA	8.4	251	Slight anisocytosis and poikilocytosis
82	55	19.1	7.3	75	34.3	NA	10.9	103	Occasional reactive lymphocytes Slight anisocytosis and poikilocytosis
83	42	14	6.4	65	33.5	NA	8.5	325	Slight anisocytosis and poikilocytosis

Table 4. (cont'd)

84	50	15.9	6.3	80	31.9	NA	8.9	255	Occasional reactive lymphocytes Slight anisocytosis Rare polychromasia, Slight poikilocytosis
85	49	17.6	7.3	68	35.7	NA	12.4	581	Rare reactive lymphocytes Slight anisocytosis and poikilocytosis Rare polychromasia
86	42	13.9	6.7	63	33.4	NA	7	666	Occasional reactive lymphocytes Slight anisocytosis and poikilocytosis
88	42	14.4	6.2	68	34.4	NA	10.5	583	None reported
89	39	12.3	5.3	73	31.9	NA	7.9	213	Slight anisocytosis and poikilocytosis
91	39	12.8	5.4	73	32.5	NA	11.2	277	Occasional reactive lymphocytes Slight anisocytosis of RBC
92	22	6.7	3.2	68	30.7	NA	48.4	89	Slight anisocytosis, polychromasia, and poikilocytosis Mild cytoplasmic basophilia of leukocytes

Hct, hematocrit;
Hgb, hemoglobin;
RBC, red blood cell concentration;
MCV, mean corpuscular volume;
MCHC, mean corpuscular hemoglobin concentration;
MCH, mean corpuscular hemoglobin;
WBC, white blood cell concentration;
NA, not available

Appendix C. Coverage of target genes

Table 5. (Appendix) Coverage of target genes.

WG_IAD165539_region.20181130.results_coverage_summary

Request_ID	Type	Name	Chr	Chr_Start	Chr_End	#_Amplicons	Total_Bases	Covered_Bases	Missed_Bases	Overall_Coverage
WG_IAD165539_region	GENOME_REGION	ENSCAFE00000295916	chr1	93321921	93322186	1	266	153	113	0.575
WG_IAD165539_region	GENOME_REGION	ENSCAFE00000299563	chr1	93322493	93322572	1	80	80	0	1
WG_IAD165539_region	GENOME_REGION	ENSCAFE00000023004	chr1	93368076	93368350	3	275	275	0	1
WG_IAD165539_region	GENOME_REGION	ENSCAFE00000023005	chr1	93377034	93377157	1	124	124	0	1
WG_IAD165539_region	GENOME_REGION	ENSCAFE00000023006	chr1	93387550	93387667	1	118	118	0	1
WG_IAD165539_region	GENOME_REGION	ENSCAFE00000023007	chr1	93393198	93393343	2	146	146	0	1
WG_IAD165539_region	GENOME_REGION	ENSCAFE00000023009	chr1	93400304	93400625	3	322	322	0	1
WG_IAD165539_region	GENOME_REGION	ENSCAFE00000023012	chr1	93401365	93401484	2	120	120	0	1
WG_IAD165539_region	GENOME_REGION	ENSCAFE00000023014	chr1	93407369	93407526	2	158	158	0	1
WG_IAD165539_region	GENOME_REGION	ENSCAFE00000023016	chr1	93409415	93409526	1	112	112	0	1
WG_IAD165539_region	GENOME_REGION	ENSCAFE00000023017	chr1	93411592	93411778	2	187	187	0	1
WG_IAD165539_region	GENOME_REGION	ENSCAFE00000023021	chr1	93412653	93412780	1	128	115	13	0.898
WG_IAD165539_region	GENOME_REGION	ENSCAFE00000023025	chr1	93415218	93415352	2	135	135	0	1
WG_IAD165539_region	GENOME_REGION	ENSCAFE00000023029	chr1	93416434	93416521	1	88	88	0	1
WG_IAD165539_region	GENOME_REGION	ENSCAFE00000023033	chr1	93419678	93419805	2	128	128	0	1
WG_IAD165539_region	GENOME_REGION	ENSCAFE00000023036	chr1	93420576	93420714	2	139	139	0	1
WG_IAD165539_region	GENOME_REGION	ENSCAFE00000023039	chr1	93422517	93422668	1	152	152	0	1
WG_IAD165539_region	GENOME_REGION	ENSCAFE00000023040	chr1	93422808	93422958	2	151	151	0	1
WG_IAD165539_region	GENOME_REGION	ENSCAFE00000023042	chr1	93423797	93423933	1	137	137	0	1
WG_IAD165539_region	GENOME_REGION	ENSCAFE00000023045	chr1	93426620	93426809	1	190	190	0	1
WG_IAD165539_region	GENOME_REGION	ENSCAFE00000023049	chr1	93427407	93427531	1	125	125	0	1
WG_IAD165539_region	GENOME_REGION	ENSCAFE00000023050	chr1	93427704	93427876	2	173	173	0	1
WG_IAD165539_region	GENOME_REGION	ENSCAFE00000023052	chr1	93433066	93433183	2	118	118	0	1
WG_IAD165539_region	GENOME_REGION	ENSCAFE00000023061	chr1	93435111	93435224	1	114	114	0	1

Table 5. (cont'd)

WG_IAD16553 9_region	GENOME_ REGION	ENSCAFE00 000320624	chr1	93435667	93435774	1	108	108	0	1
WG_IAD16553 9_region	GENOME_ REGION	ENSCAFE00 000318704	chr13	47107802	47108604	8	803	803	0	1
WG_IAD16553 9_region	GENOME_ REGION	ENSCAFE00 000293717	chr13	47144084	47144120	1	37	37	0	1
WG_IAD16553 9_region	GENOME_ REGION	ENSCAFE00 000022598	chr13	47144447	47144716	2	270	270	0	1
WG_IAD16553 9_region	GENOME_ REGION	ENSCAFE00 000022599	chr13	47147501	47147782	3	282	282	0	1
WG_IAD16553 9_region	GENOME_ REGION	ENSCAFE00 000022600	chr13	47148867	47149003	1	137	137	0	1
WG_IAD16553 9_region	GENOME_ REGION	ENSCAFE00 000022601	chr13	47153583	47153754	2	172	172	0	1
WG_IAD16553 9_region	GENOME_ REGION	ENSCAFE00 000022602	chr13	47157660	47157849	1	190	190	0	1
WG_IAD16553 9_region	GENOME_ REGION	ENSCAFE00 000022603	chr13	47160927	47161042	1	116	116	0	1
WG_IAD16553 9_region	GENOME_ REGION	ENSCAFE00 000022604	chr13	47175058	47175172	1	115	115	0	1
WG_IAD16553 9_region	GENOME_ REGION	ENSCAFE00 000022605	chr13	47176922	47177103	1	182	182	0	1
WG_IAD16553 9_region	GENOME_ REGION	ENSCAFE00 000022606	chr13	47178312	47178418	1	107	107	0	1
WG_IAD16553 9_region	GENOME_ REGION	ENSCAFE00 000022607	chr13	47178512	47178638	2	127	127	0	1
WG_IAD16553 9_region	GENOME_ REGION	ENSCAFE00 000022608	chr13	47178924	47179028	1	105	105	0	1
WG_IAD16553 9_region	GENOME_ REGION	ENSCAFE00 000022609	chr13	47179115	47179225	1	111	111	0	1
WG_IAD16553 9_region	GENOME_ REGION	ENSCAFE00 000022611	chr13	47180416	47180566	1	151	151	0	1
WG_IAD16553 9_region	GENOME_ REGION	ENSCAFE00 000022612	chr13	47182449	47182540	1	92	92	0	1
WG_IAD16553 9_region	GENOME_ REGION	ENSCAFE00 000022613	chr13	47183018	47183145	1	128	128	0	1
WG_IAD16553 9_region	GENOME_ REGION	ENSCAFE00 000022614	chr13	47184250	47184372	1	123	123	0	1
WG_IAD16553 9_region	GENOME_ REGION	ENSCAFE00 000022615	chr13	47187863	47187974	1	112	112	0	1
WG_IAD16553 9_region	GENOME_ REGION	ENSCAFE00 000022616	chr13	47188088	47188187	1	100	100	0	1
WG_IAD16553 9_region	GENOME_ REGION	ENSCAFE00 000022617	chr13	47188541	47188646	1	106	106	0	1
WG_IAD16553 9_region	GENOME_ REGION	ENSCAFE00 000022618	chr13	47189703	47190029	4	327	327	0	1
WG_IAD16553 9_region	GENOME_ REGION	ENSCAFE00 000307341	chr13	47189703	47192057	19	2355	2355	0	1
WG_IAD16553 9_region	GENOME_ REGION	ENSCAFE00 000257596	chr17	19492193	19492334	1	142	142	0	1
WG_IAD16553 9_region	GENOME_ REGION	ENSCAFE00 000045109	chr17	19493109	19493227	1	119	119	0	1
WG_IAD16553 9_region	GENOME_ REGION	ENSCAFE00 000045104	chr17	19494050	19494119	1	70	70	0	1
WG_IAD16553 9_region	GENOME_ REGION	ENSCAFE00 000045096	chr17	19496046	19496131	2	86	86	0	1
WG_IAD16553 9_region	GENOME_ REGION	ENSCAFE00 000045089	chr17	19497240	19497388	1	149	149	0	1

Table 5. (cont'd)

WG_IAD165539_region	GENOME_REGION	ENSCAFE00000045083	chr17	19497578	19497668	1	91	91	0	1
WG_IAD165539_region	GENOME_REGION	ENSCAFE00000045077	chr17	19498504	19498655	1	152	152	0	1
WG_IAD165539_region	GENOME_REGION	ENSCAFE00000045071	chr17	19500469	19500553	1	85	85	0	1
WG_IAD165539_region	GENOME_REGION	ENSCAFE00000045065	chr17	19500795	19500978	3	184	184	0	1
WG_IAD165539_region	GENOME_REGION	ENSCAFE00000045060	chr17	19501152	19501264	2	113	113	0	1
WG_IAD165539_region	GENOME_REGION	ENSCAFE00000045058	chr17	19501873	19501952	1	80	80	0	1
WG_IAD165539_region	GENOME_REGION	ENSCAFE00000045056	chr17	19502705	19502749	1	45	45	0	1
WG_IAD165539_region	GENOME_REGION	ENSCAFE00000045055	chr17	19502857	19503006	1	150	150	0	1
WG_IAD165539_region	GENOME_REGION	ENSCAFE00000045053	chr17	19503342	19503498	1	157	157	0	1
WG_IAD165539_region	GENOME_REGION	ENSCAFE00000045051	chr17	19503793	19503900	1	108	108	0	1
WG_IAD165539_region	GENOME_REGION	ENSCAFE00000045047	chr17	19504176	19504334	1	159	159	0	1
WG_IAD165539_region	GENOME_REGION	ENSCAFE00000045043	chr17	19504599	19504814	2	216	216	0	1
WG_IAD165539_region	GENOME_REGION	ENSCAFE00000045041	chr17	19527910	19528056	1	147	147	0	1
WG_IAD165539_region	GENOME_REGION	ENSCAFE00000324742	chr17	19530681	19530727	1	47	47	0	1
WG_IAD165539_region	GENOME_REGION	ENSCAFE00000045035	chr17	19534652	19534922	3	271	271	0	1
WG_IAD165539_region	GENOME_REGION	ENSCAFE00000045031	chr17	19549441	19549545	1	105	105	0	1
WG_IAD165539_region	GENOME_REGION	ENSCAFE00000045029	chr17	19562840	19562902	1	63	63	0	1
WG_IAD165539_region	GENOME_REGION	ENSCAFE00000290364	chr24	21731567	21731649	1	83	83	0	1
WG_IAD165539_region	GENOME_REGION	ENSCAFE00000330261	chr24	21733974	21734085	1	112	112	0	1
WG_IAD165539_region	GENOME_REGION	ENSCAFE00000078429	chr24	21793571	21793691	1	121	121	0	1
WG_IAD165539_region	GENOME_REGION	ENSCAFE00000078432	chr24	21793768	21793865	1	98	98	0	1
WG_IAD165539_region	GENOME_REGION	ENSCAFE00000078437	chr24	21794434	21794527	1	94	94	0	1
WG_IAD165539_region	GENOME_REGION	ENSCAFE00000078443	chr24	21794986	21795138	1	153	153	0	1
WG_IAD165539_region	GENOME_REGION	ENSCAFE00000078451	chr24	21796392	21796555	1	164	164	0	1
WG_IAD165539_region	GENOME_REGION	ENSCAFE00000078457	chr24	21796651	21796747	1	97	97	0	1
WG_IAD165539_region	GENOME_REGION	ENSCAFE00000078465	chr24	21798334	21798439	1	106	106	0	1
WG_IAD165539_region	GENOME_REGION	ENSCAFE00000078476	chr24	21798722	21799355	5	634	634	0	1
WG_IAD165539_region	GENOME_REGION	ENSCAFE00000078488	chr24	21799891	21804772	39	4882	4882	0	1
WG_IAD165539_region	GENOME_REGION	ENSCAFE00000131755	chr3	53070718	53070802	1	85	85	0	1

Table 5. (cont'd)

WG_IAD16553 9_region	GENOME_ REGION	ENSCAFE00 000131732	chr3	53071120	53071212	1	93	93	0	1
WG_IAD16553 9_region	GENOME_ REGION	ENSCAFE00 000131720	chr3	53071320	53071417	1	98	98	0	1
WG_IAD16553 9_region	GENOME_ REGION	ENSCAFE00 000131717	chr3	53071592	53071704	1	113	113	0	1
WG_IAD16553 9_region	GENOME_ REGION	ENSCAFE00 000131715	chr3	53073528	53073679	1	152	152	0	1
WG_IAD16553 9_region	GENOME_ REGION	ENSCAFE00 000131714	chr3	53073874	53074010	2	137	137	0	1
WG_IAD16553 9_region	GENOME_ REGION	ENSCAFE00 000131712	chr3	53074826	53074969	3	144	144	0	1
WG_IAD16553 9_region	GENOME_ REGION	ENSCAFE00 000131706	chr3	53075052	53075212	3	161	161	0	1
WG_IAD16553 9_region	GENOME_ REGION	ENSCAFE00 000131702	chr3	53076447	53076612	1	166	166	0	1
WG_IAD16553 9_region	GENOME_ REGION	ENSCAFE00 000131696	chr3	53077475	53077566	2	92	92	0	1
WG_IAD16553 9_region	GENOME_ REGION	ENSCAFE00 000317459	chr3	53082217	53082680	4	464	464	0	1
WG_IAD16553 9_region	GENOME_ REGION	ENSCAFE00 000303975	chr31	30170967	30171442	4	476	476	0	1
WG_IAD16553 9_region	GENOME_ REGION	ENSCAFE00 000104076	chr31	30177933	30178094	1	162	162	0	1
WG_IAD16553 9_region	GENOME_ REGION	ENSCAFE00 000104110	chr31	30209968	30210159	3	192	192	0	1
WG_IAD16553 9_region	GENOME_ REGION	ENSCAFE00 000104062	chr31	30233276	30233380	2	105	105	0	1
WG_IAD16553 9_region	GENOME_ REGION	ENSCAFE00 000104049	chr31	30254040	30254196	1	157	157	0	1
WG_IAD16553 9_region	GENOME_ REGION	ENSCAFE00 000104034	chr31	30260450	30260703	3	254	254	0	1
WG_IAD16553 9_region	GENOME_ REGION	ENSCAFE00 000251675	chr31	30266612	30266705	1	94	94	0	1
WG_IAD16553 9_region	GENOME_ REGION	ENSCAFE00 000312331	chr31	30419856	30419972	1	117	117	0	1
WG_IAD16553 9_region	GENOME_ REGION	ENSCAFE00 000317947	chr31	30420093	30420332	2	240	240	0	1
WG_IAD16553 9_region	GENOME_ REGION	ENSCAFE00 000302105	chr32	26021552	26021994	3	443	427	16	0.964
WG_IAD16553 9_region	GENOME_ REGION	ENSCAFE00 000287772	chr32	26059693	26059837	2	145	145	0	1
WG_IAD16553 9_region	GENOME_ REGION	ENSCAFE00 000118877	chr32	26104600	26108084	28	3485	3485	0	1
WG_IAD16553 9_region	GENOME_ REGION	ENSCAFE00 000283371	chr32	26109778	26109868	1	91	91	0	1
WG_IAD16553 9_region	GENOME_ REGION	ENSCAFE00 000118909	chr32	26111293	26111386	1	94	94	0	1
WG_IAD16553 9_region	GENOME_ REGION	ENSCAFE00 000245329	chr32	26111985	26112193	2	209	209	0	1
WG_IAD16553 9_region	GENOME_ REGION	ENSCAFE00 000317146	chr32	26126985	26127135	1	151	151	0	1
WG_IAD16553 9_region	GENOME_ REGION	ENSCAFE00 000331029	chr32	26129145	26129234	1	90	90	0	1
WG_IAD16553 9_region	GENOME_ REGION	ENSCAFE00 000303624	chr32	26137124	26137261	1	138	138	0	1
WG_IAD16553 9_region	GENOME_ REGION	ENSCAFE00 000323679	chr32	26139602	26139956	3	355	355	0	1

Table 5. (cont'd)

WG_IAD165539_region	GENOME_REGION	ENSCAFE00000293142	chr32	26142047	26144313	18	2267	2267	0	1
WG_IAD165539_region	GENOME_REGION	ENSCAFE00000319785	chr37	6920834	6924501	29	3668	3668	0	1
WG_IAD165539_region	GENOME_REGION	ENSCAFE00000118469	chr37	6924808	6925024	2	217	217	0	1
WG_IAD165539_region	GENOME_REGION	ENSCAFE00000118467	chr37	6926751	6927023	3	273	273	0	1
WG_IAD165539_region	GENOME_REGION	ENSCAFE00000118465	chr37	6929028	6929159	2	132	132	0	1
WG_IAD165539_region	GENOME_REGION	ENSCAFE00000118463	chr37	6929503	6929623	1	121	121	0	1
WG_IAD165539_region	GENOME_REGION	ENSCAFE00000118460	chr37	6931099	6931210	3	112	112	0	1
WG_IAD165539_region	GENOME_REGION	ENSCAFE00000118458	chr37	6931304	6931486	2	183	183	0	1
WG_IAD165539_region	GENOME_REGION	ENSCAFE00000118455	chr37	6931763	6931984	2	222	222	0	1
WG_IAD165539_region	GENOME_REGION	ENSCAFE00000118451	chr37	6932914	6933039	1	126	126	0	1
WG_IAD165539_region	GENOME_REGION	ENSCAFE00000118447	chr37	6933175	6933321	1	147	147	0	1
WG_IAD165539_region	GENOME_REGION	ENSCAFE00000118444	chr37	6933428	6933573	1	146	146	0	1
WG_IAD165539_region	GENOME_REGION	ENSCAFE00000118443	chr37	6933849	6934119	3	271	271	0	1
WG_IAD165539_region	GENOME_REGION	ENSCAFE00000118441	chr37	6934247	6934333	2	87	87	0	1
WG_IAD165539_region	GENOME_REGION	ENSCAFE00000118438	chr37	6935050	6935229	2	180	180	0	1
WG_IAD165539_region	GENOME_REGION	ENSCAFE00000118426	chr37	6935781	6935882	1	102	102	0	1
WG_IAD165539_region	GENOME_REGION	ENSCAFE00000118416	chr37	6935979	6936176	2	198	198	0	1
WG_IAD165539_region	GENOME_REGION	ENSCAFE00000118412	chr37	6937464	6937585	1	122	122	0	1
WG_IAD165539_region	GENOME_REGION	ENSCAFE00000118407	chr37	6937829	6938041	1	213	195	18	0.915
WG_IAD165539_region	GENOME_REGION	ENSCAFE00000118400	chr37	6938941	6939178	3	238	238	0	1
WG_IAD165539_region	GENOME_REGION	ENSCAFE00000118392	chr37	6943306	6943476	2	171	171	0	1
WG_IAD165539_region	GENOME_REGION	ENSCAFE00000118385	chr37	6944427	6944506	1	80	80	0	1
WG_IAD165539_region	GENOME_REGION	ENSCAFE00000118375	chr37	6946341	6946455	1	115	115	0	1
WG_IAD165539_region	GENOME_REGION	ENSCAFE00000118361	chr37	6947063	6947167	1	105	105	0	1
WG_IAD165539_region	GENOME_REGION	ENSCAFE00000118349	chr37	6949248	6949414	2	167	167	0	1
WG_IAD165539_region	GENOME_REGION	ENSCAFE00000316981	chr37	6962328	6962848	4	521	521	0	1
WG_IAD165539_region	GENOME_REGION	ENSCAFE00000253596	chr37	16512895	16512985	1	91	91	0	1
WG_IAD165539_region	GENOME_REGION	ENSCAFE00000148526	chr37	16514285	16514447	1	163	163	0	1
WG_IAD165539_region	GENOME_REGION	ENSCAFE00000148514	chr37	16515382	16515522	1	141	141	0	1

Table 5. (cont'd)

WG_IAD165539_region	GENOME_REGION	ENSCAFE00000148505	chr37	16517599	16517750	1	152	152	0	1
WG_IAD165539_region	GENOME_REGION	ENSCAFE00000148492	chr37	16519669	16519846	2	178	178	0	1
WG_IAD165539_region	GENOME_REGION	ENSCAFE00000148485	chr37	16521138	16521243	2	106	106	0	1
WG_IAD165539_region	GENOME_REGION	ENSCAFE00000148469	chr37	16524275	16524566	2	292	292	0	1
WG_IAD165539_region	GENOME_REGION	ENSCAFE00000148446	chr37	16527154	16527291	1	138	138	0	1
WG_IAD165539_region	GENOME_REGION	ENSCAFE00000300858	chr37	16530017	16530083	1	67	67	0	1
WG_IAD165539_region	GENOME_REGION	ENSCAFE00000325685	chr37	16531098	16531352	3	255	255	0	1
WG_IAD165539_region	GENOME_REGION	ENSCAFE00000311903	chr5	32560598	32561487	7	890	890	0	1
WG_IAD165539_region	GENOME_REGION	ENSCAFE00000181406	chr5	32562109	32562215	1	107	107	0	1
WG_IAD165539_region	GENOME_REGION	ENSCAFE00000181403	chr5	32562594	32562667	1	74	74	0	1
WG_IAD165539_region	GENOME_REGION	ENSCAFE00000181399	chr5	32562954	32563090	1	137	137	0	1
WG_IAD165539_region	GENOME_REGION	ENSCAFE00000181397	chr5	32563352	32563461	2	110	110	0	1
WG_IAD165539_region	GENOME_REGION	ENSCAFE00000181396	chr5	32563684	32563796	2	113	113	0	1
WG_IAD165539_region	GENOME_REGION	ENSCAFE00000181392	chr5	32563878	32564064	2	187	187	0	1
WG_IAD165539_region	GENOME_REGION	ENSCAFE00000254320	chr5	32564567	32564806	2	240	240	0	1
WG_IAD165539_region	GENOME_REGION	ENSCAFE00000181390	chr5	32564897	32564918	1	22	22	0	1
WG_IAD165539_region	GENOME_REGION	ENSCAFE00000181389	chr5	32565071	32565172	2	102	102	0	1
WG_IAD165539_region	GENOME_REGION	ENSCAFE00000322991	chr5	32573982	32574109	2	128	128	0	1
WG_IAD165539_region	GENOME_REGION	ENSCAFE00000056063	chr9	4028695	4029056	3	362	362	0	1
WG_IAD165539_region	GENOME_REGION	ENSCAFE00000056064	chr9	4029266	4029569	2	304	304	0	1
WG_IAD165539_region	GENOME_REGION	ENSCAFE00000193584	chr9	35855967	35857452	11	1486	1480	6	0.996
WG_IAD165539_region	GENOME_REGION	ENSCAFE00000193581	chr9	35862571	35862813	3	243	243	0	1
WG_IAD165539_region	GENOME_REGION	ENSCAFE00000193575	chr9	35868946	35869136	1	191	191	0	1
WG_IAD165539_region	GENOME_REGION	ENSCAFE00000193567	chr9	35886212	35886336	1	125	125	0	1
WG_IAD165539_region	GENOME_REGION	ENSCAFE00000193565	chr9	35895600	35895828	3	229	229	0	1
WG_IAD165539_region	GENOME_REGION	ENSCAFE00000193562	chr9	35913413	35913884	3	472	472	0	1

REFERENCES

REFERENCES

1. Blokzijl F, De Ligt J, Jager M, et al. Tissue-specific mutation accumulation in human adult stem cells during life. *Nature*. 2016;538(7624):260-264.

2. Yizhak K, Aguet F, Kim J, et al. RNA sequence analysis reveals macroscopic somatic clonal expansion across normal tissues. *Science*. 2019;364(6444).

3. Fey MF, Liechti-Gallati S, Von Rohr A, et al. Clonality and X-inactivation patterns in hematopoietic cell populations detected by the highly informative M27 beta DNA probe. *Blood*. 1994;83(4):931-938.

4. Busque L, Patel JP, Figueroa ME, et al. Recurrent somatic TET2 mutations in normal elderly individuals with clonal hematopoiesis. *Nature genetics*. 2012;44(11):1179-1181.

5. Zink F, Stacey SN, Norddahl GL, et al. Clonal hematopoiesis, with and without candidate driver mutations, is common in the elderly. *Blood*. 2017;130(6):742-752.

6. Jaiswal S, Fontanillas P, Flannick J, et al. Age-related clonal hematopoiesis associated with adverse outcomes. *New England Journal of Medicine*. 2014;371(26):2488-2498.

7. Genovese G, Kähler AK, Handsaker RE, et al. Clonal hematopoiesis and blood-cancer risk inferred from blood DNA sequence. *New England Journal of Medicine*. 2014;371(26):2477-2487.

8. Xie M, Lu C, Wang J, et al. Age-related mutations associated with clonal hematopoietic expansion and malignancies. *Nature medicine*. 2014;20(12):1472-1478.

9. Steensma DP, Bejar R, Jaiswal S, et al. Clonal hematopoiesis of indeterminate potential and its distinction from myelodysplastic syndromes. *Blood*. 2015;126(1):9-16.

10. Jaiswal S, Natarajan P, Silver AJ, et al. Clonal hematopoiesis and risk of atherosclerotic cardiovascular disease. *New England Journal of Medicine*. 2017;377(2):111-121.

11. Couronné L, Bastard C, Bernard OA. TET2 and DNMT3A mutations in human T-cell lymphoma. *New England Journal of Medicine*. 2012;366(1):95-96.

12. Reddy A, Zhang J, Davis NS, et al. Genetic and functional drivers of diffuse large B cell lymphoma. *Cell*. 2017;171(2):481-494.

13. Fuster JJ, MacLauchlan S, Zuriaga MA, et al. Clonal hematopoiesis associated with TET2 deficiency accelerates atherosclerosis development in mice. *Science*. 2017;355(6327):842-847.

14. Franceschi C, Campisi J. Chronic inflammation (inflammaging) and its potential contribution to age-associated diseases. *Journals of Gerontology Series A: Biomedical Sciences and Medical Sciences*. 2014;69(Suppl_1):S4-S9.

15. Cook EK, Luo M, Rauh MJ. Clonal hematopoiesis and inflammation: Partners in leukemogenesis and comorbidity. *Experimental hematology*. 2020;83:85-94.

16. Franceschi C, Garagnani P, Parini P, Giuliani C, Santoro A. Inflammaging: a new immune–metabolic viewpoint for age-related diseases. *Nature Reviews Endocrinology*. 2018;14(10):576-590.

17. Ferrucci L, Fabbri E. Inflammageing: chronic inflammation in ageing, cardiovascular disease, and frailty. *Nature Reviews Cardiology*. 2018;15(9):505-522.

18. Fabre MA, McKerrell T, Zwiebel M, et al. Concordance for clonal hematopoiesis is limited in elderly twins. *Blood, The Journal of the American Society of Hematology*. 2020;135(4):269-273.

19. Hansen JW, Pedersen DA, Larsen LA, et al. Clonal hematopoiesis in elderly twins: concordance, discordance, and mortality. *Blood*. 2020;135(4):261-268.

20. Vas V, Senger K, Dörr K, Niebel A, Geiger H. Aging of the microenvironment influences clonality in hematopoiesis. *PloS one*. 2012;7(8):e42080.

21. Zhang B, Chu S, Agarwal P, et al. Inhibition of interleukin-1 signaling enhances elimination of tyrosine kinase inhibitor–treated CML stem cells. *Blood*. 2016;128(23):2671-2682.

22. Abegunde SO, Buckstein R, Wells RA, Rauh MJ. An inflammatory environment containing TNFα favors Tet2-mutant clonal hematopoiesis. *Experimental hematology*. 2018;59:60-65.

23. Lindsley RC, Saber W, Mar BG, et al. Prognostic mutations in myelodysplastic syndrome after stem-cell transplantation. *New England Journal of Medicine*. 2017;376(6):536-547.

24. Ptashkin RN, Mandelker DL, Coombs CC, et al. Prevalence of clonal hematopoiesis mutations in tumor-only clinical genomic profiling of solid tumors. *JAMA oncology*. 2018;4(11):1589-1593.

25. Ko M, Bandukwala HS, An J, et al. Ten-Eleven-Translocation 2 (TET2) negatively regulates homeostasis and differentiation of hematopoietic stem cells in mice. *Proceedings of the National Academy of Sciences*. 2011;108(35):14566-14571.

26. Moran-Crusio K, Reavie L, Shih A, et al. Tet2 loss leads to increased hematopoietic stem cell self-renewal and myeloid transformation. *Cancer cell*. 2011;20(1):11-24.

27. Challen GA, Sun D, Jeong M, et al. Dnmt3a is essential for hematopoietic stem cell differentiation. *Nature genetics*. 2012;44(1):23-31.

28. Celik H, Mallaney C, Kothari A, et al. Enforced differentiation of Dnmt3a-null bone marrow leads to failure with c-Kit mutations driving leukemic transformation. *Blood*. 2015;125(4):619-628.

29. Mayle A, Yang L, Rodriguez B, et al. Dnmt3a loss predisposes murine hematopoietic stem cells to malignant transformation. *Blood*. 2015;125(4):629-638.

30. Chin DWL, Yoshizato T, Virding SC, et al. Aged healthy mice acquire clonal hematopoiesis mutations. *Blood*. 2022;139(4):629-634.

31. Shin T, Chen S, Cordes S, et al. Macaque CRISPR/Cas9 age-related clonal hematopoiesis model demonstrates expansion of TET2-mutated clones and applicability for testing mitigation approaches. *Blood*. 2020;136:27-28.

32. Parker HG, Ostrander EA. Canine genomics and genetics: running with the pack. *PLoS Genet*. 2005;1(5):e58.

33. Wayne RK, Ostrander EA. Lessons learned from the dog genome. *TRENDS in Genetics*. 2007;23(11):557-567.

34. Michell AR. Longevit of British breeds of dog and its relationships with-sex, size, cardiovascular variables and disease. *Veterinary Record*. 1999;145(22):625-629.

35. Parker HG, Kim LV, Sutter NB, et al. Genetic structure of the purebred domestic dog. *Science*. 2004;304(5674):1160-1164.

36. Egenvall A, Bonnett BN, Shoukri M, Olson P, Hedhammar Å, Dohoo I. Age pattern of mortality in eight breeds of insured dogs in Sweden. *Preventive veterinary medicine*. 2000;46(1):1-14.

37. Pollinger JP, Lohmueller KE, Han E, et al. Genome-wide SNP and haplotype analyses reveal a rich history underlying dog domestication. *Nature*. 2010;464(7290):898-902.

38. Valli VE, Myint MS, Barthel A, et al. Classification of canine malignant lymphomas according to the World Health Organization criteria. *Veterinary pathology*. 2011;48(1):198-211.

39. McManus PM. Classification of myeloid neoplasms: a comparative review. *Veterinary Clinical Pathology*. 2005;34(3):189-212.

40. Beurlet S, Krief P, Sansonetti A, et al. Identification of JAK2 mutations in canine primary polycythemia. *Experimental hematology*. 2011;39(5):542-545.

41. Bronzini I, Aresu L, Paganin M, et al. DNA methylation and targeted sequencing of methyltransferases family genes in canine acute myeloid leukaemia, modelling human myeloid leukaemia. *Veterinary and Comparative Oncology*. 2017;15(3):910-918.

42. Usher SG, Radford AD, Villiers EJ, Blackwood L. RAS, FLT3, and C-KIT mutations in immunophenotyped canine leukemias. *Experimental hematology*. 2009;37(1):65-77.

43. LeBlanc AK, Mazcko CN. Improving human cancer therapy through the evaluation of pet dogs. *Nature Reviews Cancer*. 2020:1-16.

44. Lindblad-Toh K, Wade CM, Mikkelsen TS, et al. Genome sequence, comparative analysis and haplotype structure of the domestic dog. *Nature*. 2005;438(7069):803-819.

45. Ostrander EA, Dreger DL, Evans JM. Canine cancer genomics: lessons for canine and human health. *Annual Review of Animal Biosciences*. 2019;7:449-472.

46. Ostrander EA, Wang G-D, Larson G, et al. Dog10K: an international sequencing effort to advance studies of canine domestication, phenotypes and health. *National science review*. 2019;6(4):810-824.

47. Nagata T, Gupta V, Sorce D, et al. Immunoglobulin motif DNA recognition and heterodimerization of the PEBP2/CBF Runt domain. *Nature structural biology*. 1999;6(7):615-619.

48. Hsu JI, Dayaram T, Tovy A, et al. PPM1D mutations drive clonal hematopoiesis in response to cytotoxic chemotherapy. *Cell Stem Cell*. 2018;23(5):700-713.

49. Kahn JD, Miller PG, Silver AJ, et al. PPM1D-truncating mutations confer resistance to chemotherapy and sensitivity to PPM1D inhibition in hematopoietic cells. *Blood*. 2018;132(11):1095-1105.

50. Schübeler D. Function and information content of DNA methylation. *Nature*. 2015;517(7534):321-326.

51. He Y-F, Li B-Z, Li Z, et al. Tet-mediated formation of 5-carboxylcytosine and its excision by TDG in mammalian DNA. *Science*. 2011;333(6047):1303-1307.

52. Ito S, Shen L, Dai Q, et al. Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine. *Science*. 2011;333(6047):1300-1303.

53. Zhang Q, Zhao K, Shen Q, et al. Tet2 is required to resolve inflammation by recruiting Hdac2 to specifically repress IL-6. *Nature*. 2015;525(7569):389-393.

54. Cai Z, Kotzin JJ, Ramdas B, et al. Inhibition of inflammatory signaling in Tet2 mutant preleukemic cells mitigates stress-induced abnormalities and clonal hematopoiesis. *Cell Stem Cell*. 2018;23(6):833-849.

55. Cull AH, Rauh MJ. Success in bone marrow failure? Novel therapeutic directions based on the immune environment of myelodysplastic syndromes. *Journal of Leukocyte Biology*. 2017;102(2):209-219.

56. Delhommeau F, Dupont S, Valle VD, et al. Mutation in TET2 in myeloid cancers. *New England Journal of Medicine*. 2009;360(22):2289-2301.

57. Langemeijer SMC, Kuiper RP, Berends M, et al. Acquired mutations in TET2 are common in myelodysplastic syndromes. *Nature genetics*. 2009;41(7):838-842.

58. Wang J, Li Z, He Y, et al. Loss of Asxl1 leads to myelodysplastic syndrome–like disease in mice. *Blood, The Journal of the American Society of Hematology*. 2014;123(4):541-553.

59. Abdel-Wahab O, Gao J, Adli M, et al. Deletion of Asxl1 results in myelodysplasia and severe developmental defects in vivoConditional deletion of Asxl1 results in MDS. *The Journal of experimental medicine*. 2013;210(12):2641-2659.

60. Haferlach T, Nagata Y, Grossmann V, et al. Landscape of genetic lesions in 944 patients with myelodysplastic syndromes. *Leukemia*. 2014;28(2):241-247.

61. Boultwood J, Perry J, Pellagatti A, et al. Frequent mutation of the polycomb-associated gene ASXL1 in the myelodysplastic syndromes and in acute myeloid leukemia. *Leukemia*. 2010;24(5):1062-1065.

62. Gelsi-Boyer V, Trouplin V, Roquain J, et al. ASXL1 mutation is associated with poor prognosis and acute transformation in chronic myelomonocytic leukaemia. *British journal of haematology*. 2010;151(4):365-375.

63. Lundberg P, Karow A, Nienhold R, et al. Clonal evolution and clinical correlates of somatic mutations in myeloproliferative neoplasms. *Blood*. 2014;123(14):2220-2228.

64. Döhner H, Weisdorf DJ, Bloomfield CD. Acute myeloid leukemia. *New England Journal of Medicine*. 2015;373(12):1136-1152.

65. Welch JS, Ley TJ, Link DC, et al. The origin and evolution of mutations in acute myeloid leukemia. *Cell*. 2012;150(2):264-278.

66. Milne TA, Sinclair DAR, Brock HW. The Additional sex combs gene of Drosophila is required for activation and repression of homeotic loci, and interacts specifically with Polycomb and super sex combs. *Molecular and General Genetics MGG*. 1999;261(4-5):753-761.

67. Liu P, Tarle SA, Hajra A, et al. Fusion between transcription factor CBF beta/PEBP2 beta and a myosin heavy chain in acute myeloid leukemia. *Science*. 1993;261(5124):1041-1044.

68. Grossmann V, Kern W, Harbich S, et al. Prognostic relevance of RUNX1 mutations in T-cell acute lymphoblastic leukemia. *Haematologica*. 2011;96(12):1874.

69. Gaidzik VI, Teleanu V, Papaemmanuil E, et al. RUNX1 mutations in acute myeloid leukemia are associated with distinct clinico-pathologic and genetic features. *Leukemia*. 2016;30(11):2160-2168.

70. Ichikawa M, Asai T, Saito T, et al. AML-1 is required for megakaryocytic maturation and lymphocytic differentiation, but not for maintenance of hematopoietic stem cells in adult hematopoiesis. *Nature medicine*. 2004;10(3):299-304.

71. Mangan JK, Speck NA. RUNX1 mutations in clonal myeloid disorders: from conventional cytogenetics to next generation sequencing, a story 40 years in the making. *Critical Reviews™ in Oncogenesis*. 2011;16(1-2).

72. Papaemmanuil E, Gerstung M, Malcovati L, et al. Clinical and biological implications of driver mutations in myelodysplastic syndromes. *Blood*. 2013;122(22):3616-3627.

73. Malcovati L, Papaemmanuil E, Bowen DT, et al. Clinical significance of SF3B1 mutations in myelodysplastic syndromes and myelodysplastic/myeloproliferative neoplasms. *Blood, The Journal of the American Society of Hematology*. 2011;118(24):6239-6246.

74. Dolatshad H, Pellagatti A, Fernandez-Mercado M, et al. Disruption of SF3B1 results in deregulated expression and splicing of key genes and pathways in myelodysplastic syndrome hematopoietic stem and progenitor cells. *Leukemia*. 2015;29(5):1092-1103.

75. Matsunawa M, Yamamoto R, Sanada M, et al. Haploinsufficiency of Sf3b1 leads to compromised stem cell function but not to myelodysplasia. *Leukemia*. 2014;28(9):1844-1850.

76. Wang C, Sashida G, Saraya A, et al. Depletion of Sf3b1 impairs proliferative capacity of hematopoietic stem cells but is not sufficient to induce myelodysplasia. *Blood*. 2014;123(21):3336-3343.

77. Goemans BF, Zwaan CM, Miller M, et al. Mutations in KIT and RAS are frequent events in pediatric core-binding factor acute myeloid leukemia. *Leukemia*. 2005;19(9):1536-1542.

78. Shimada A, Taki T, Tabuchi K, et al. KIT mutations, and not FLT3 internal tandem duplication, are strongly associated with a poor prognosis in pediatric acute myeloid leukemia with t (8; 21): a study of the Japanese Childhood AML Cooperative Study Group. *Blood*. 2006;107(5):1806-1809.

79. Kitayama H, Tsujimura T, Matsumura I, et al. Neoplastic transformation of normal hematopoietic cells by constitutively activating mutations of c-kit receptor tyrosine kinase. 1996.

80. London CA, Galli SJ, Yuuki T, Hu Z-Q, Helfand SC, Geissler EN. Spontaneous canine mast cell tumors express tandem duplications in the proto-oncogene c-kit. *Experimental hematology*. 1999;27(4):689-697.

81. Webster JD, Yuzbasiyan-Gurkan V, Miller RA, Kaneene JB, Kiupel M. Cellular proliferation in canine cutaneous mast cell tumors: associations with c-KIT and its role in prognostication. *Veterinary Pathology*. 2007;44(3):298-308.

82. Frost D, Lasota J, Miettinen M. Gastrointestinal stromal tumors and leiomyomas in the dog: a histopathologic, immunohistochemical, and molecular genetic study of 50 cases. *Veterinary Pathology*. 2003;40(1):42-54.

83. Beghini A, Ripamonti CB, Cairoli R, et al. KIT activating mutations: incidence in adult and pediatric acute myeloid leukemia, and identification of an internal tandem duplication. *Haematologica*. 2004;89(8):920-925.

84. Care RS, Valk PJM, Goodeve AC, et al. Incidence and prognosis of c-KIT and FLT3 mutations in core binding factor (CBF) acute myeloid leukaemias. *British journal of haematology*. 2003;121(5):775-777.

85. Hayashi Y, Harada Y, Huang G, Harada H. Myeloid neoplasms with germ line RUNX1 mutation. *International Journal of Hematology*. 2017;106(2):183-188.

86. Kwok B, Hall JM, Witte JS, et al. MDS-associated somatic mutations and clonal hematopoiesis are common in idiopathic cytopenias of undetermined significance. *Blood*. 2015;126(21):2355-2361.

87. Jajosky AN, Sadri N, Meyerson HJ, et al. Clonal cytopenia of undetermined significance (CCUS) with dysplasia is enriched for MDS-type molecular findings compared to CCUS without dysplasia. *European Journal of Haematology*. 2021;106(4):500-507.

88. Malcovati L, Gallì A, Travaglino E, et al. Clinical significance of somatic mutation in unexplained blood cytopenia. *Blood*. 2017;129(25):3371-3378.

89. Luo X, Feurstein S, Mohan S, et al. ClinGen myeloid malignancy variant curation expert panel recommendations for germline RUNX1 variants. *Blood advances*. 2019;3(20):2962-2979.

90. Bellissimo DC, Chen C-h, Zhu Q, et al. Runx1 negatively regulates inflammatory cytokine production by neutrophils in response to Toll-like receptor signaling. *Blood advances*. 2020;4(6):1145-1158.

91. Hayashi Y, Zhang Y, Yokota A, et al. Pathobiological pseudohypoxia as a putative mechanism underlying myelodysplastic syndromes. *Cancer discovery*. 2018;8(11):1438-1457.

92. Wu D, Ozaki T, Yoshihara Y, Kubo N, Nakagawara A. Runt-related transcription factor 1 (RUNX1) stimulates tumor suppressor p53 protein in response to DNA damage through complex formation and acetylation. *Journal of Biological Chemistry*. 2013;288(2):1353-1364.

93. Cai X, Gao L, Teng L, et al. Runx1 deficiency decreases ribosome biogenesis and confers stress resistance to hematopoietic stem and progenitor cells. *Cell stem cell*. 2015;17(2):165-177.

94. Muto T, Walker CS, Choi K, et al. Adaptive response to inflammation contributes to sustained myelopoiesis and confers a competitive advantage in myelodysplastic syndrome HSCs. *Nature immunology*. 2020;21(5):535-545.

95. Ran F, Hsu PD, Wright J, Agarwala V, Scott DA, Zhang F. Genome engineering using the CRISPR-Cas9 system. *Nature protocols*. 2013;8(11):2281-2308.

96. Arber DA, Orazi A, Hasserjian R, et al. The 2016 revision to the World Health Organization classification of myeloid neoplasms and acute leukemia. *Blood*. 2016;127(20):2391-2405.

97. Harada H, Harada Y, Tanaka H, Kimura A, Inaba T. Implications of somatic mutations in the AML1 gene in radiation-associated and therapy-related myelodysplastic syndrome/acute myeloid leukemia. *Blood*. 2003;101(2):673-680.

98. Vegesna V, Takeuchi S, Hofmann W-K, et al. C/EBP-β, C/EBP-δ, PU. 1, AML1 genes: mutational analysis in 381 samples of hematopoietic and solid malignancies. *Leukemia research*. 2002;26(5):451-457.

99. Illendula A, Gilmour J, Grembecka J, et al. Small molecule inhibitor of CBFβ-RUNX binding for RUNX transcription factor driven cancers. *EBioMedicine*. 2016;8:117-131.