Received: 20 December 2021 | Revised: 28 March 2022 | Accepted: 26 April 2022

DOI: 10.1111/vco.12825

ORIGINAL ARTICLE

Veterinary and Comparative Oncology WILEY

Using digital RNA counting to establish flow cytometry diagnostic criteria for subtypes of CD34+ canine acute leukaemia

R. Adam Harris | Emily D. Rout | Janna A. Yoshimoto | Paul R. Avery | Anne C. Avery

Department of Microbiology, Immunology and Pathology, College of Veterinary Medicine and Biomedical Sciences, Colorado State University, Fort Collins, Colorado, USA

Correspondence
R. Adam Harris, 200 West Lake Street, Colorado State University, Fort Collins, CO, 80523, USA.
Email: adam.dugger@colostate.edu

Funding information
Internal funding derived from diagnostic service was used for this project.

Abstract

Canine acute leukaemia is a heterogeneous neoplasm with multiple phenotypes. Criteria to subtype acute leukaemia by flow cytometry have not been validated. The goal of this study was to develop a panel of antibodies and objective antigen expression criteria for the assignment of lymphoid or myeloid lineage by flow cytometry. We isolated mRNA from the blood of 45 CD34+ acute leukaemia cases and measured expression of 43 genes that represent lymphoid and myeloid lineages using NanoString technology. We determined differentially expressed genes between major groups identified by unsupervised hierarchical clustering. We then evaluated the expression of antigens by flow cytometry to determine if cases could be assigned to a lineage. Two groups were identified by gene expression. Group 1/LYMPH overexpressed lymphoid-associated genes (ex. *DNTT*) and had a higher percentage of CD5 + CD3- cells by flow cytometry. Group 2/MYELO overexpressed myeloid-associated genes (ex. *ANPEP*/CD13) and had a higher percentage of class II major histocompatibility complex (MHCII)- CD14+ and/or CD18 + CD4- cells. We proposed that >12.5% CD5 + CD3- cells in the blood was indicative of lymphoid lineage, and > 3.0% CD14 + MHCII- cells or > 18% CD18 +- MHCII-CD4- cells was indicative of myeloid lineage. 15/15 cases that met the proposed criteria for acute lymphocytic leukaemia were in LYMPH group and 12/15 cases that met the proposed criteria for acute myeloid leukaemia were in MYELO group. The majority of CD34+ cases that did not meet either immunophenotyping lineage criterion (12/13) clustered within the LYMPH group. In conclusion, currently available antibodies can be useful for determining canine acute leukaemia subtypes.

KEYWORDS
acute leukaemia, dog, lymphoma, nanostring, oncology

1 | INTRODUCTION

Acute leukaemia (AL) is a group of heterogeneous, clinically aggressive haematopoietic neoplasms that arise from precursor lymphoid or myeloid cells that have not completed their program of differentiation. In human medicine, a combination of diagnostic tests, including cytology, histology, flow cytometry, and molecular genetics/mutational analysis are commonly used to characterize the neoplasm and assign lineage.

Acute leukaemias typically express antigens associated with immature haematopoietic cells, including the stem cell antigen CD34.[1,2] However, the expression of CD34 can be variable and lack of CD34 expression does not rule out AL as a diagnosis.[3,4] Acute myeloid leukaemias (AML) in humans are initially identified using immunophenotyping and are characterized by expression of CD14, CD11c, and/or myeloperoxidase expressed by neoplastic cells.[5] Acute lymphocytic leukaemias (ALL) are subdivided into B cell lineage (B-ALL) and T cell lineage (T-ALL) and typically express surface CD19 and/or cytoplasmic CD79a and CD22 or surface or cytoplasmic CD3, respectively.[5] CD5 expression in human T-ALL is variable.[6,7] Acute unclassifiable leukaemias (AUL) lack lineage specific antigens whereas mixed phenotype acute leukaemias (MPAL) express two or more lineage associated antigens.[5]

The diagnosis of AL in dogs begins with the identification of immature leukocytes in the blood and/or bone marrow. Some reports use >20% circulating blasts in the blood or > 30% in the marrow as diagnostic criteria for AL in the dog.[8,9] The expression of CD34, as assessed by flow cytometry, can further help distinguish immature or precursor cells from lymphoproliferative disorders of mature leukocytes and has been used independently of blast count as an indicator of AL.[10] Determination of lineage (myeloid versus lymphoid) requires additional tools. Adam et al used expression of cell surface CD14, intracellular MAC387, and myeloperoxidase to diagnose AML, and lack of expression of these together with positive intracellular expression of CD79a to indicate B-ALL and intracellular expression of CD3 and surface expression of CD5 to indicate T-ALL.[9] Similar criteria for identifying AMLs were employed by Stokol et al, who also demonstrated the use of alkaline phosphatase staining as an indicator of myeloid lineage.[11]

While the use of these antibodies to define lineage is consistent with practices in human hematopathology, species differences in antigen expression raise the possibility that different criteria for lineage assignment may be necessary in dogs. Furthermore, while CD79a expression has been used as an indicator of B cell lineage, it is expressed on precursor myeloid cells in people and mature myeloid cells in dogs and therefore does not reliably identify precursor B cells.[12–14] Other groups have relied on additional antigens such as CD21, and/or CD22 to identify B-cell lineage AL by flow cytometry.[8,11] In this study we attempt to address some gaps in our ability to determine the lineage of CD34+ canine AL.

In human medicine, gene expression profiling reliably classifies leukaemia subclasses and mRNA and DNA sequencing have revealed numerous genetic abnormalities that are critical for prognosis and prediction of treatment response.[15,16] Differentiation between tumour subclasses is imperative in human medicine because of disease-specific therapeutic options.[17] Similar clinical significances likely exist in veterinary medicine but have not been fully explored. The utility of microarray and digital RNA technology in a diagnostic setting is limited in veterinary medicine, partially attributed to cost and lack of supportive data. However, gene expression profiles of haematopoietic neoplasms can be used to design flow cytometry criteria to distinguish tumour subtypes. The goal of this study was to use gene expression to classify CD34+ AL as myeloid or lymphoid, and then correlate that classification with the immunophenotype observed by flow cytometry. We focused on CD34+ AL cases because, in the authors' experience, CD34+ AL cases are typically difficult to assign to a lineage, given their ambiguous protein expression and biological behaviour. This study allowed us to assign lineage to AL based on commercially available directly conjugated cell surface antibodies. Currently, outcomes in canine AL are poor and improving classification of canine AL may allow for more targeted therapy and improve clinical outcomes.

2 | METHODS

2.1 | Case selection

Samples included EDTA peripheral blood samples from dogs (n = 45) diagnosed with AL by flow cytometry through the Colorado State University's (CSU) Clinical Hematopathology Laboratory between August 2015 and May 2017. Samples were not solicited for this study. These cases were not breed or age restricted. A diagnosis of acute leukaemia was made using a combination of laboratory abnormalities and protein expression assessed by flow cytometry. AL were considered CD34+ if >10% of total leukocytes expressed CD34 or there was >1000 CD34+ cells/μl. Some B cell lymphomas can express CD34 but these cases are readily identified by expression of class II MHC (MHCII) and the B cell antigen CD21, and were excluded from analysis.[18,19] Rarely plasma cell tumours can express CD34, but IRF4 (MUM1) was included in the gene expression panel to try to identify and exclude potential plasma cell cases.[20] AL cases were chosen to represent samples that were tentatively classified as lymphoid, myeloid, unclassified, or mixed phenotype based on subjective flow cytometric criteria. Blood samples from two dogs with no suspicion of haematopoietic neoplasia were used as controls.

2.2 | Immunophenotyping

Routine CBCs were performed and submitted concurrently with the patient's sample. CBCs were either performed in the submitting clinic with a bench top analyser, by a veterinary diagnostic reference laboratory, or performed at the clinical pathology laboratory at CSU Veterinary Diagnostic Laboratory. Samples (n = 14) submitted to CSU Clinical Pathology Laboratory were submitted for automated cell counts (Advia 120 Haematology Analyser, Siemans, Tarrytown, NY), blood smear evaluation, and a manual white blood cell differential count. For CSU Clinical Pathology Laboratory CBC cases, a leukocytosis was defined as > 15000 cells/μl, neutropenia as < 2600 neutrophils/μl, neutrophilia as >11000 neutrophils/μl. Additionally, patients were considered thrombocytopenic and anaemic if they had a platelet count less than 200000/μl without platelet clumps and a haematocrit less than 40%, respectively. For all other CBC data provided by the clinic, we interpreted those results using the reported analyzers' reference interval. A manual review of the blood film was not available in a subset (n = 14) of these cases.

Flow cytometry was performed as previously described and analysed using Kaluza software (Beckman Coulter, www.beckmancoulter.com).[21]

Samples were acquired over 21 months and the staining panel was changed during this time. The two panels are provided in Supplement Table 1. The panel does not detect intracellular antigens. Diagnosis of AL was based on CD34+ expression as noted above. All neoplasms expressed the pan-leukocyte marker, CD45. Cases were selected for further study based on subjective flow cytometric criteria with the goal of having a mixture of lineages (Figure 1). These included cases with a population of cells expressing CD5 but no other T cell antigens (surface CD3, CD4, or CD8) (n = 12), cases with a population of CD14+MHCII- cells (n = 19), cases where neoplastic cells expressed no lineage-specific antigens (n = 11), and cases with mixed phenotypes (CD5+ population and CD14+MHCII- population) (n = 3). These cases had variable expression of CD18 and none of these cases had surface expression of CD21 or surface CD3. The CD18 antibody used in this study is a cross-reactive anti-human CD18. This is used because of the wider variety of fluorochromes available. A prior study demonstrated that this antibody binds primarily to canine neutrophils and monocytes, and only at low levels to lymphocytes.[22]

2.3 | Gene expression profiling

Red blood cells were lysed with ACK lysis buffer using the same protocol as used for preparing cells for flow cytometry. After red cell lysis, leukocytes were pelleted and total RNA was isolated from the leukocyte pellet using the PureLink RNA Mini kit (ThermoFisher, Waltham, MA) as previously described.[23] RNA samples were analysed for concentration and integrity using a Nanodrop 2000c (ThermoFisher, Waltham, MA) and Agilent RNA ScreenTape assay (Agilent Technologies, Santa Clara, CA), respectively. Samples had concentrations ranging from 15–2177 ng/μl (median, 221 ng/μl; interquartile range [IQR], 99–370 ng/μl) and RNA integrity numbers (RINs) ranging from 3.1 to 8.4 (median: 6.0; IQR: 4.5–7.3) (Supplement Table 2). mRNA counts were obtained through the University of Arizona's Genetics Core using an nCounter Digital Analyser (Nanostring Technologies, Seattle, WA) with a custom-built panel of genes. We compiled a list of 43 genes to represent haematopoietic lineages, including genes commonly expressed by myeloid, T cell, and B cell lineages, by reviewing available literature for human and dog (Table 1 and Supplement Document 1). Additional genes were included that are expressed in haematopoietic precursors. Seven of the genes selected were previously shown to be differentially expressed between AML and ALL in humans.[15] Six housekeeping genes were originally included as internal normalization controls. Probes were designed by NanoString based on proprietary algorithms designed to provide the most efficient detection of mRNA using the known *Canine lupus familiaris* annotated gene sequences using the CanFam3.1 assembly (Supplement Table 3).

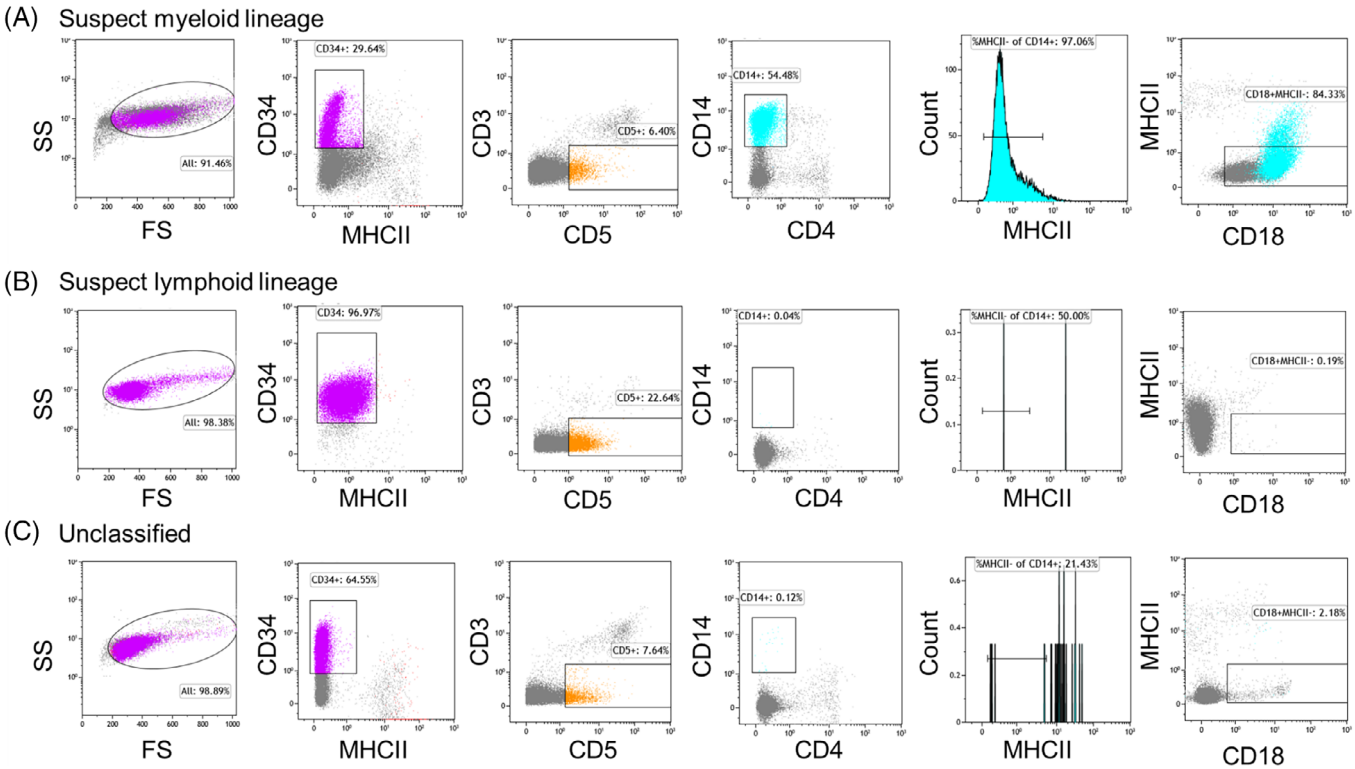

FIGURE 1 Flow cytometric plots for acute leukaemia subtypes: These are the most common subtypes of acute leukaemia seen by the Clinical Hematopathology Laboratory. (A) Suspected myeloid lineage neoplasms often have either CD14+CD4-cells that lack MHCII expression and/or a CD18+ population that is also MHCII- and lacks CD4 expression (indicating they are not mature neutrophils). (B) Suspected lymphoid lineages express CD5 and typically lack surface CD3 expression. (C) Lineage unclassified acute leukaemias do not express lineage associated antigens, including CD5, and lack CD14+MHCII- cells. CD18 expression is often variable in these cases. FS, linear forward light scatter. SS, log side light scatter

TABLE 1 Gene list used for investigating AL subtypes

Gene symbol[a]	Alternative symbol	Gene name[a]	Probe category[b]	Reference (s)[c]
CD19	CD19	CD19 molecule	B cell	1
CD22	CD22	CD22 molecule	B cell	1,2
CD79A	CD79a	CD79a molecule	B cell	1,2,3
CR2	CD21	Complement C3d receptor 2	B cell	1,3,4
IGHG	IgG	Immunoglobulin heavy constant gamma	B cell	5
IGHM	IgM	Immunoglobulin heavy constant Mu	B cell	2,6
MS4A1	CD20	Membrane Spanning 4-Domains A1	B cell	2,6
PAX5	Pax5	Paired Box 5	B cell	5
CD2	CD2	CD2 molecule	T cell	1
CD3E	CD3e	CD3e molecule	T cell	1,3,7
CD5	CD5	CD5 molecule	T cell	1,8
CD7	CD7	CD7 molecule	T cell	1
LYL1	LYL1	LYL1, basic helix–loop–helix family member	T cell	9
TRBC	TCR beta	T cell receptor beta locus	T cell	5
TRGC2	TCR gamma	T cell receptor gamma locus	T cell	10
TRGC3	TCR gamma	T cell receptor gamma locus	T cell	10
TRGC8	TCR gamma	T cell receptor gamma locus	T cell	10
CFD	Adipsin	Complement factor D	Myeloid	11
LGALS3	Lectin	Galaectin 3	Myeloid	11
MCL1	MCL1	BCL2 family apoptosis regulator	Myeloid	11
SQSTM1	P62	Sequestosome 1	Myeloid	11
ZYX	Zyxin	Zyxin	Myeloid	11
ANPEP	CD13	Alanyl aminopeptidase, membrane	Myeloid	6
CD14	CD14	CD14 molecule	Myeloid	2,3,4
CD33	CD33	CD33 molecule	Myeloid	1,2
ITGAM	CD11b	Integrin subunit alpha M	Myeloid	6,12
ITGAX	CD11c	Integrin subunit alpha X	Myeloid	2,11,12
LYZ	Lysozyme	Lysozyme	Myeloid	11
MPO	MPO	Myeloperoxidase	Myeloid	1
NPM1	NPM	Nucleophosmin	Myeloid	20
CD34	CD34	CD34 Molecule	Other	2,3,9
MPL	CD110	MPL proto-oncogene, thrombopoietin receptor	Other	13
IL3RA	CD123	Interleukin 3 receptor subunit alpha	Other	13
ITGB2	CD18	Integrin subunit beta 2	Other	14
PROM1	CD133	Prominin 1	Other	15
CIITA	MHC2TA/CIITA	Class II, major histocompatibility complex, transactivator	Other	1
DLA-DQA1	MHC class II/DLA-DQA1, dog ortholog of HLA-DQA2	Major histocompatibility complex, Class II, DQ Alpha 2	Other	2,5
DLA-DRA	MHC class II/DLA-DRA, dog ortholog of HLA-DRA	Major histocompatibility complex, Class II, DR Alpha	Other	2,5
DNTT	Tdt	DNA nucleotidylexotransferase	Other	2,6
GATA3	GATA3	GATA binding protein 3	Other	16
HLA-DRB1	MHC class II/HLA-DRB1	Major histocompatibility complex, Class II, DR Beta 1	Other	2,5
IRF4	MUM1	Interferon regulatory factor 4	Other	17
KIT	CD117	KIT proto-oncogene receptor tyrosine kinase	Other	1,2

(Continues)

TABLE 1 (Continued)

Gene symbol[a]	Alternative symbol	Gene name[a]	Probe category[b]	Reference (s)[c]
EEF1G	EEF1G	Eukaryotic translation elongation factor 1 gamma	Housekeeping	18
GUSB	GUSB	Glucuronidase beta	Housekeeping	18
HPRT1	HPRT1	Hypoxanthine phosphoribosyltransferase 1	Housekeeping	18
POLR2A	POLR2A	RNA polymerase II subunit A	Housekeeping	18
RPL19	RPL19	Ribosomal protein L19	Housekeeping	19
SDHA	SDHA	Succinate dehydrogenase complex flavoprotein subunit A	Housekeeping	18

[a]Gene symbols and gene names are provided for the 49 genes used in this study to determine AL lineages.
[b]The probe category and the associated hematopoeitic lineage were labelled accordingly as B cell, T cell, Myeloid, Other, and the housekeeping genes.
[c]The references provided were used to assign the probe categories.

2.4 | Statistical analysis

Results were analysed using nSolver 4.0 software (NanoString Technologies, Seattle, WA). No quality control flags were identified by the nSolver software for any of the probes in our panel. Gene counts were normalized using the four housekeeping genes that had the lowest variance (*SDHA*, *RPL19*, *POLR2A*, *HPRT1*). Gene counts were transformed into log2 for further analysis. Genes were first clustered using Pearson's correlation. To prevent bias in the expression profiles, multi-loci genes were removed if they clustered together (see below). Unsupervised hierarchical clustering of the normalized transformed log2 gene counts separated AL cases into two groups (Supplement Figure 2). Differential gene expression was determined between the two AL groups, excluding controls, using a cutoff of false discovery rate (FDR) < 0.05. Due to the fewer number of genes left post differential gene expression analysis, Spearman's correlation was then used to recluster AL samples, excluding controls, to generate a new heatmap (Figure 2). nCounter normalized gene counts were visually inspected between the controls and AL cases and a student's *t* test was performed between selected genes. For each of these groups defined by gene expression, the flow cytometry immunophenotyping data were summarized and descriptive statistics were generated in Prism (Graphpad, La Jolla, California) to determine new diagnostic criteria. nCounter normalized gene counts and median fluorescence intensity (MFI) of expression of the corresponding cell surface protein by flow cytometry for five different antigens were chosen for a comparative analysis using Spearman's correlation. Spearman's correlation was chosen because the MFI of all AL samples had a non-parametric distribution. Statistically significant (p value < .05) correlation coefficients were determined.

2.5 | Eliminating co-regulated genes

A subset of genes chosen in the panel included genes that are co-regulated. For example, the initial gene set included multiple class II MHC, immunoglobulin, and T cell receptor genes. Class II MHC genes are under the control of the CIITA protein and their expression is controlled as a group.[24] Similarly, IgG and IgM are under common transcriptional control.[25] To prevent overrepresenting these groups of genes that are co-regulated, we restricted our analysis to prevent overrepresenting each of these families of genes. To do this, we assessed how the genes we chose clustered together to identify potential bias. We first used Pearson's correlation to cluster the genes. Multi-locus probes for genes encoding MHCII (*DLA-DQA1*, *DLA-DRA*, and *HLA-DRB1*) and immunoglobulin heavy chains (*IGHG* and *IGHM*) clustered with their respective groups and were considered highly correlated (Pearson's r = 0.55–0.94). The genes that had the highest coefficients of variations (CV) (*DLA-DRA* CV = 240%; *IGHM* CV = 157%) were left to represent those families of genes for further analysis. Interestingly, the TRGC genes (*TRG2*, *TRG3*, and *TRG8*) did not cluster together, which could potentially be explained by the difference in expression of TRG-related genes seen between human ALL and AML.[26] Therefore, all TCR loci (*TRG2*, *TRG3*, and *TRG8*) were included for downstream analysis.

3 | RESULTS

3.1 | Patient population and clinical presentation

Digital RNA quantification was performed on peripheral blood from 45 cases and two healthy controls. Patient information, white blood cell concentrations, and flow cytometry data are available in Supplement Table 4. Ages of the 45 newly diagnosed AL patients ranged from 1.1 to 15.85 (median 8.7) years. There were 22 spayed females, 19 castrated males, and two intact males. Cases included mixed breed dogs (n = 10), Labrador Retrievers (n = 10), Golden Retrievers (n = 6), German Shepherds (n = 3), Australian Shepherds (n = 2), Labradoodles (n = 2), Pit bull terriers (n = 2), and one each of Belgian Shepherd, Bullmastiff, Portuguese Water dog, Keeshond, Boxer, Irish Setter, Greyhound, Maltese, and Beagle. Control samples were from one intact male English Bulldog and one castrated male Golden Retriever. Date of birth and sex status were not available for two samples and breed information was not available for one of the samples. 42/45 (93%) AL cases presented with a leukocytosis (median: 77300 leukocytes/μl; IQR: 35540–171650 cells/μl, range: 11300–417220

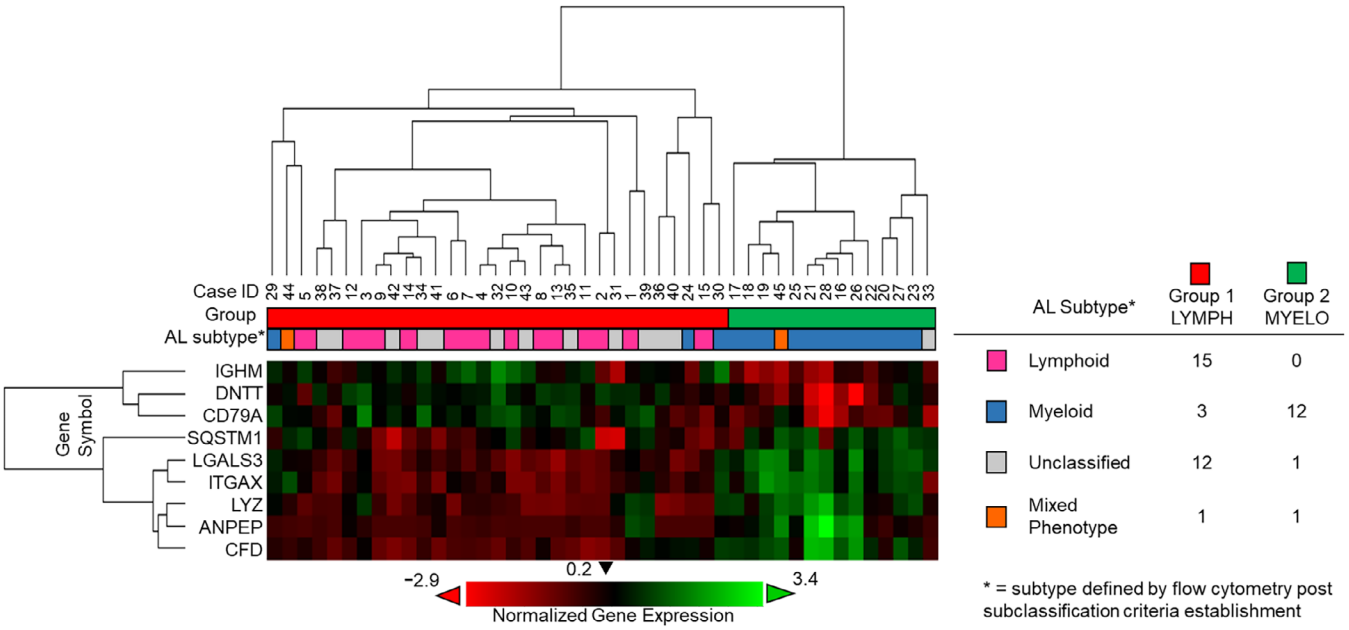

FIGURE 2 Hierarchical clustering of genes and cases: Samples were reclustered using the nine differentially expressed genes (FDR < 0.05) identified by unsupervised hierarchical clustering. By gene expression, cases were clustered into two groups: Group 1/LYMPH (red bar) overexpressed lymphoid genes and Group 2/MYELO (green bar) overexpressed myeloid genes. By flow cytometry, each case was retroactively classified as lymphoid, myeloid, unclassified or mixed phenotype using the developed immunophenotyping guidelines. The majority of CD5+ cases (lymphoid, pink) and CD14+/CD18+ MHCII- cases (myeloid, blue) separated into gene expression groups 1/LYMPH and 2/MYELO, respectively. Additionally, the majority of cases that did not express lineage specific antigens (unclassified, grey) also clustered with group 1/LYMPH. (Flow cytometry classified groups: lymphoid = pink, myeloid = blue, unclassified = grey, mixed phenotype = orange)

cells/uL), 29/45 (64%) were neutropenic, and 3/45 (6.6%) had a neutrophilia. AL cases had a median platelet count of 60000/μl (range: 10000–216000/μl) and 40/45 (88.8%) were thrombocytopenic. Additionally, they had a median haematocrit of 28% (range, 12%–46%) and 41/45 (91%) were considered anaemic.

3.2 | Validating and reviewing the nCounter probe gene set

We first determined that the samples were consistent with immature haematopoietic neoplasms. The normalized mRNA counts for genes expressed in precursor cells (*CD34*, *KIT*, and *DNTT*[TdT]) were compared between all the leukaemic samples and the normal controls. Genes specific for early haematopoietic precursors were significantly increased compared to controls: *CD34* was upregulated 73-fold (p value < .01), *KIT* was upregulated 187-fold (p value < .01), and *DNTT* was upregulated 168-fold (p value = .02). We evaluated the correlation between gene expression and flow cytometry by comparing the gene count and antigen expression MFI for CD34, CD5, CD18, CD14 and MHCII (Supplement Figure 1). Gene counts and protein expression of CD34, CD5, CD18 and MHC class II correlated (r = 0.38–0.68, p value < .05), but CD14 gene counts and protein were not correlated (r = 0.04, p value = .75). A possible explanation for this is that the NanoString CD14 probe was designed against the CanFam3.1 assembly, but more recent canine genome builds have a discordant CD14 sequence. Therefore, it's possible this probe did not bind well to CD14 transcripts.

3.3 | Differential gene expression between AL groups

The average normalized gene counts for the control samples and each of the two groups are provided in Supplement Table 5. Nine of the 43 genes were significantly differentially expressed (FDR < 0.05) between the two groups. Six of these differentially expressed genes were myeloid-associated (*SQSTM1*, *LGALS3*, *ITAGX*, *LYZ*, *ANPEP*, *CFD)* and three were lymphoid-associated (*IGHM*, *DNTT*, *CD79A)*. Cases were clustered using these nine genes and a heat map was generated (Figure 2). Cases overexpressing the lymphoid-associated genes were assigned to a LYMPH group (n = 31 cases) and cases overexpressing myeloid-associated genes were assigned to a MYELO group (n = 14 cases).

Two genes were included in the expression panel that are considered defining for the T cell and B cell lineage: *CD3* and *PAX5*. *CD3* was considered highly expressed in the LYMPH group compared to the MYELO group (Supplement Table 5), being 3.4 log2 fold change times higher in the LYMPH group. It was also 1.7 log2 fold change higher than in the control peripheral blood. The comparison in *CD3* expression between the two AL groups was significant (p = .005) but because of the variability in expression, the FDR was greater than 0.05, and this gene was not included in the lineage-defining gene list. *PAX5* is the defining B cell lineage gene and was

expressed at higher levels in the two control peripheral blood samples than either the LYMPH or MYELO group.

3.4 | Flow cytometric features of each group

Flow cytometric data from each of the groups were compared using a Mann–Whitney U test for each of the following: %CD34+, %CD5+, %CD14+MHCII-, and %CD18+MHCII-CD4- (Figure 3A). Expression of all surface proteins, as determined by MFI, were statistically different between the groups. Descriptive statistics for each group were used to establish flow cytometric diagnostic guidelines for myeloid, lymphoid, unclassified, and mixed phenotype AL (Table 2). The LYMPH group had higher percentages of CD34+ cells ($p < .01$) and CD5+ (surface CD3-/CD4-/CD8-) cells ($p < .03$). The MYELO group had higher percentages of CD14+MHCII- cells ($p < .01$). The percentage of CD18+MHCII-CD4- cells was also statistically higher (p value < .01) in the MYELO group. Even though the percentage of CD34+ cells was significantly different between the two groups, CD34 expression was not used in establishing the guidelines as the 95% confidence interval (CI) of the mean overlapped between the two groups.

3.5 | Establishing immunophenotypic criteria for lineage assignment

We established flow cytometry criteria to differentiate lymphoid from myeloid leukaemias based on the total percentage of CD5+ cells, CD14+MHCII- cells, and CD18+MHCII-CD4- cells in the two groups defined by gene expression (Figure 3B). We calculated the upper and lower limits of the 95% CI of the mean for each group and defined the diagnostic cutoff as the halfway point between the upper 95% CI of one group and the lower 95% of the other group (rounding to the nearest 0.5% for a practical diagnostic cutoff percentage). For example, the lower 95% CI of %CD5+ cells for the LYMPH group was 13.6% and the upper 95% CI for %CD5+ cells in the MYELO group was 11.0% (Figure 3B & Table 2). The point halfway between them was chosen as the diagnostic cutoff: 12.5% (rounded to the nearest 0.5). The %CD5+ and %CD14+MHCII- were considered primary diagnostic criteria. If cases did not meet the primary criteria, then the total %CD18+MHCII-CD4- cells was considered a secondary criterion. For instance, if a sample had <3.0% CD14+MHCII- cells and < 12.5 CD5+ cells, we would then consider CD18 expression as the final diagnostic criteria cutoff. In these cases, samples with >18% CD18+MHCII-CD4-

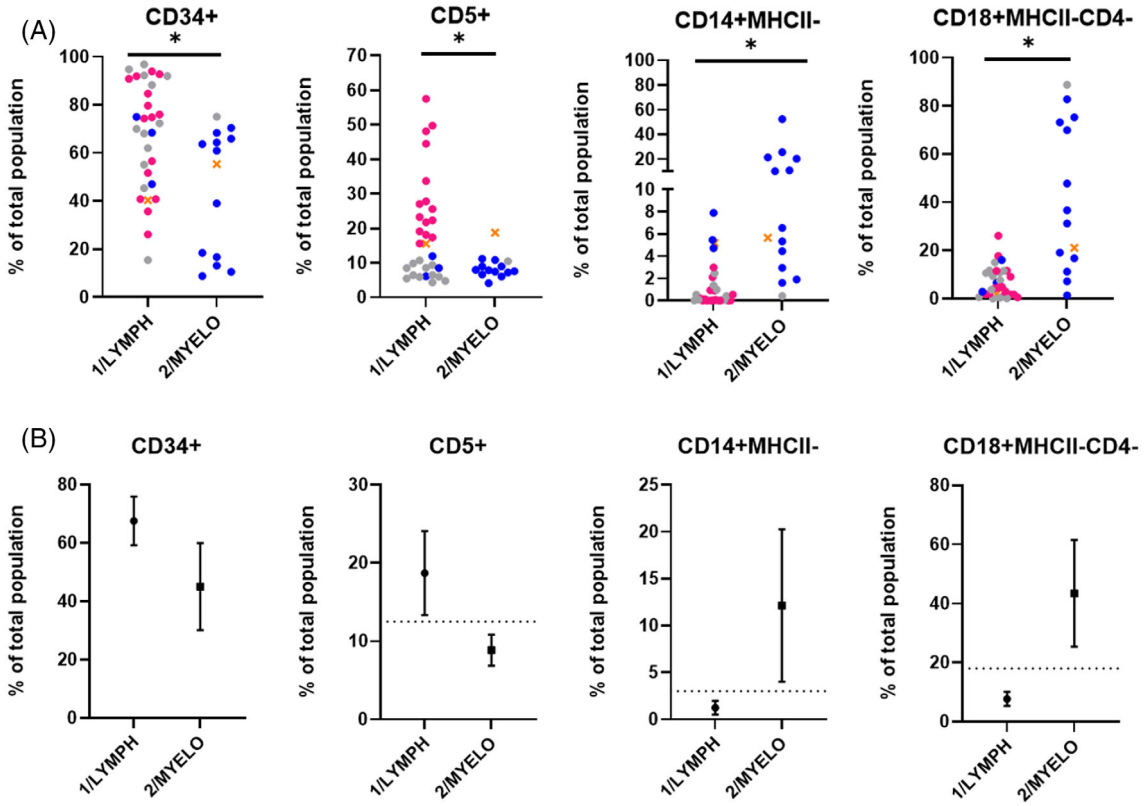

FIGURE 3 Comparison of flow cytometric data between groups. A) The percentage of positive cells for each surface marker assessed was statistically different (* = p value < .05) between gene expression groups 1/LYMPH and 2/MYELO. Each dot is an individual case, coloured by the flow cytometry subtype determined retroactively using the developed immunophenotyping classification scheme (myeloid = blue dots, lymphoid = pink dots, unclassified = grey dots, mixed phenotype = orange X). B) The 95% confidence interval (CI) of the mean for the percentages of CD5+ cells, CD14+MHCII- cells, and CD18+MHCII-CD4- cells are shown. The dashed line represents the proposed cutoff for subtyping acute leukaemias into lymphoid, myeloid, or unclassified. Note that even though the percentage of CD34+ cells was statistically different, the %CD34+ cells was highly variable, and the 95% CI overlapped substantially between the groups, so a cutoff was not established

TABLE 2 Descriptive statistics for each group

	%CD5+		%CD14+MHCII-		%CD18+MHCII-CD4-		%CD34+	
Gene expression group	**1/LYMPH**	**2/MYELO**	**1/LYMPH**	**2/MYELO**	**1/LYMPH**	**2/MYELO**	**1/LYMPH**	**2/MYELO**
n	31	14	31	14	31	14	31	14
Minimum	4.4	4.1	0.0	0.4	0.6	1.9	15.4	8.8
Maximum	57.5	18.8	7.9	52.4	27.1	90.8	96.9	75.1
Median	15.5	7.9	0.2	6.1	5.6	35.9	72.4	58.2
IQR of Median	6.5–25.6	7.1–10.5	0.0–1.4	2.7–20.6	2.4–11.9	16.7–77.5	47.1–90.9	15.8–66.6
Mean	19.0	9.0	1.2	12.1	7.7	43.4	67.6	45.1
Std. deviation of mean	14.5	3.5	2.0	14.1	6.4	31.2	22.7	25.8
Lower 95% CI of mean[a]	13.5	6.9	0.5	4.0	5.4	25.4	57.9	27.6
Upper 95% CI of mean[a]	24.3	10.8	2.0	20.3	10.1	61.5	77.2	62.6

[a]Criteria for subtyping acute leukaemia cases were generated by using the half-way point between each group's 95% confidence intervals of the mean for each category.

cells were identified as AML and those with <18% CD18+MHCII-CD4- cells were considered AUL. A proposed algorithm for subtyping canine AL into ALL, AML, AUL and MPAL is provided in Figure 4.

After developing the flow cytometry diagnostic criteria, the 45 AL cases were classified as ALL, AML, AUL or MPAL based on flow cytometry immunophenotyping (Figure 2 & Figure 3A). By flow cytometry, there were 15 ALL cases, 15 AML cases, 13 AUL cases and 2 MPAL cases. All cases segregated into the LYMPH group by gene expression and 12/15 AML cases segregated into the MYELO group by gene expression. Most of the cases classified as AUL (12/13) by flow cytometry segregated into the LYMPH gene expression group. In total, the LYMPH gene expression group contained all 15 ALL cases, 12 AUL cases, 3 AML, and 1 MPAL case (Figure 4C). The MYELO gene expression group contained 12 AML cases, 1 AUL cases and 1 MPAL case.

3.6 | Evaluating gene expression between flow cytometry-defined AL subtypes

Normalized mRNA counts are provided for the controls and each AL subtype as defined by flow cytometry (Supplement Table 6). In our probe list, 11 genes were differentially expressed (FDR < 0.05) between flow cytometry defined lymphoid and myeloid cases (Supplement Figure 3). Lymphoid associated genes were upregulated in the CD5+ cases compared to CD14+MHCII-/CD18+MHCII-CD4- cases. Myeloid genes, including *ITGAX* and *LGALS3*, were overexpressed in CD14+-MHCII-/CD18+MHCII-CD4- cases compared to both CD5+ and unclassified cases. CD5 was the only gene differentially expressed between CD5+ cases and unclassified cases.

4 | DISCUSSION

In this study, we used gene expression profiling to assign myeloid or lymphoid lineage to canine AL. We then determined the immunophenotypic features of these groups using commercially available, directly conjugated antibodies and developed an algorithm for lineage assignment of CD34+ AL by flow cytometry. Lymphoid lineage AL were identified by expression of CD5 and myeloid lineage AL were identified by expression of CD14+-MHCII- and CD18+MHCII-CD4- cells. The lineage of some cases could not be determined because they did not express any of these antigens or expressed both lymphoid and myeloid lineage antigens. Most unclassified cases clustered by gene expression with CD5+ AL cases and it is likely that some of these are of lymphoid origin.

After determining the differentially expressed genes between the two groups originally identified by unsupervised clustering, a heat map was generated to further classify the groups based on the differentially expressed genes of all the samples. The new groups identified (LYMPH vs. MYELO) were used to establish flow cytometric criteria. Nine genes defined the two clusters, the majority of which are associated with myeloid lineage (*SQSTM1*, *LGALS3*, *ITAGX*, *LYZ*, *ANPEP*, *CFD*). *DNTT*, *CD79A*, and *IGHM* are genes associated with lymphoid lineages. Once CD5+ and CD14+MHCII-/CD18+MHCII-CD4- AL were separated using the flow cytometric diagnostic criteria, the CD5+ neoplasms overexpressed lymphoid associated genes (*DNTT*, *CD2*, *CD5*, *CD79A*, and *IGHM*) compared to myeloid defined cases (Supplement Figure 3). The overexpression of B cell associated genes, including *CD79A* and *IGHM*, is surprising on initial examination.

The lineage (B vs. T cell) of the ALL was not determined in this study. While two of the three lymphoid-defining genes are *CD79A* and *IGHM*, which are expressed by mature B cells but not mature T cells, neither of these genes exhibits lineage fidelity in immature haematopoietic cells or AL. In particular, CD79a is considered non-specific for distinguishing B-ALL from T-ALL in humans.[6,13,27] CD79a is expressed in normal canine monocytes and T cells.[14] It is important to note that, at least in human haematopoietic progenitors, committed T cell lineage progenitors express antigens in the following order of development: CD7, CD2, CD5, and lastly surface CD3.[6] The authors hypothesize that the CD5+ ALL are a derivative of a T cell precursor neoplasm and along this biological spectrum. For instance, several T cell specific genes, including *CD2*, *CD3*, *CD7* and *GATA3* were upregulated in the LYMPH group of leukaemias compared to the MYELO group, reaching statistical

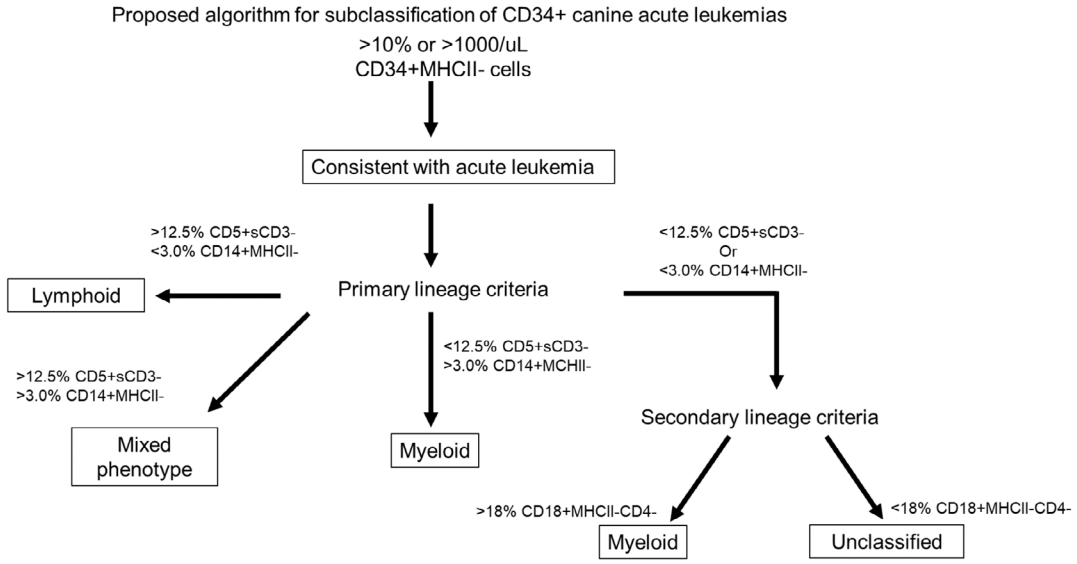

FIGURE 4 Algorithm for subclassification of CD34+ canine acute leukaemias by flow cytometry. A lineage classification scheme was developed using the cutoffs established in Figure 3B. Lineage is first determined based on the percentages of CD5 + CD3- cells and CD14 + MHCII- cells. Cases not defined by these primary lineage criteria are further evaluated for a CD18 + MHCII-CD4- population to differentiate myeloid and unclassified acute leukaemia cases

significance in individual comparisons, but not at the level of multiple comparisons (Supplement Table 5). By contrast, none of the B cell lineage-specific genes (*PAX5*, *CD19*, *CD20*, *CD21* or *CD22*) were significantly different between the groups. The conflicting gene expression data does not allow us to draw firm conclusions about the B- or T- cell origin of ALL in dogs. Although we favour the idea that most canine ALLs are T cell, this hypothesis will require further studies using additional samples, comparisons with normal canine B and T cells, as well as detection of proteins by ICC and IHC.

There are a few limitations to this study. One is that our study population only included CD34+ AL cases. There are a few reasons for this: (1) In the authors' experience CD34- precursor neoplasms are rare. There are cases that present like precursor neoplasms based on clinical presentation and aggressive clinical course that do not express CD34, but they express other lineage antigens and therefore lineage assessment is not challenging. The goal of this study was to focus on CD34+ ALs which are more ambiguous and challenging to assign lineage. (2) CD34 is highly suggestive for a precursor haematopoietic neoplasm (if CD45 is also expressed)[4] and the use of CD34 to define these cases provides an objective criterion for diagnosis. The criteria used to diagnosis CD34+ AL in this study (>1000 cells/μl or > 10% CD34+ cells) may have omitted AL cases with minor populations of CD34+ cells. Clinical presentation and CBC findings should be interpreted along with flow cytometry findings to diagnose AL. (3) CD34 expression is clinically relevant and known to be associated with shorter survival in dogs.[10] Therefore, we decided to validate our panel with CD34+ AL cases.

Another limitation is the limited probe gene list used in our study. Our probe list was based on a literature review of both veterinary and human medicine. We used probes that would identify a wide variety of haematopoietic lineages, primarily focusing on B, T, and myeloid lineages. We recognize the potential bias for hand-selecting probes. We attempted to remove any potential bias by identifying multi-locus genes that were expressed and clustered together. In this study set it was MHCII (*DLA-DQA1*, *DLA-DRA*, and *HLA-DRB1*) and immunoglobulin heavy chains (IgG and IgM). Only one of the genes from each of these gene families were chosen for the final clustering algorithm given that their transcription is similarly regulated. The selection of which TRGC loci was not unbiased either. Canines have eight TRGC loci that can be rearranged and incorporated into the T cell receptor.[28] We designed probes specifically for *TRGC2*, *TRGC3*, and *TRGC8* because previous groups have shown that these loci are preferentially rearranged.[29] It is conceivable that the other TRGC loci not included in our study could have been useful for differentiating these neoplasms.

Finally, we developed this classification scheme with antibodies directed at surface antigens that are used routinely for immunophenotyping in most laboratories.[30] It is likely that additional markers/combinations can elucidate AL subtypes similarly. There is some indirect evidence for this provided in our gene expression data. For instance, CD11c, CD11b, and Galectin 3 are overexpressed in AML compared to ALL and cross-reactive antibodies are commercially available. Interestingly, myeloperoxidase was not considered differentially expressed between subtypes (FDR = 0.42, p value = .03), even though other groups have shown it to be useful for distinguishing myeloid neoplasms.[31] Additional studies interrogating the specificity of these antigens for delineating AL subtypes are warranted.

Here, we show that AL subtypes (i.e., lymphoid vs. myeloid) can be differentiated using commercially available, directly conjugated antibodies. Additional work using a separate cohort is needed to validate our flow cytometry diagnostic criteria. We were able to show acceptable correlation between mRNA and expression of surface protein measured by flow cytometry using NanoString technology and we identified potential markers that can be attempted to further classify AL subtypes.

ACKNOWLEDGEMENTS

This study was made possible with support from the Clinical Hematopathology Laboratory at Colorado State University and by the veterinarians and clients that have provided valuable information and samples for these cases.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.

ORCID

R. Adam Harris https://orcid.org/0000-0002-9471-886X
Emily D. Rout https://orcid.org/0000-0003-1435-8532

REFERENCES

1. Vernau W, Moore PF. An immunophenotypic study of canine leukemias and preliminary assessment of clonality by polymerase chain reaction. *Vet Immunol Immunopathol*. 1999;69(2–4):145-164.
2. Batinić D, Tindle R, Boban D, et al. Expression of haematopoietic progenitor cell-associated antigen BI-3C5/CD34 in leukemia. *Leuk Res*. 1989;13(1):83-85.
3. Avery AC. The genetic and molecular basis for canine models of human leukemia and lymphoma. *Front Oncol*. 2020;10:23.
4. Novacco M, Comazzi S, Marconato L, et al. Prognostic factors in canine acute leukaemias: a retrospective study. *Vet Comp Oncol*. 2016;14(4):409-416.
5. 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.
6. You MJ, Medeiros LJ, Hsi ED. T-lymphoblastic leukemia/lymphoma. *Am J Clin Pathol*. 2015;144(3):411-422.
7. Coustan-Smith E, Mullighan CG, Onciu M, et al. Early T-cell precursor leukaemia: a subtype of very high-risk acute lymphoblastic leukaemia. *Lancet Oncol*. 2009;10(2):147-156.
8. Bennett AL, Williams LE, Ferguson MW, et al. Canine acute leukaemia: 50 cases (1989-2014). *Vet Comp Oncol*. 2017;15(3):1101-1114.
9. Adam F, Villiers E, Watson S, Coyne K, Blackwood L. Clinical pathological and epidemiological assessment of morphologically and immunologically confirmed canine leukaemia. *Vet Comp Oncol*. 2009;7(3):181-195.
10. Williams MJ, Avery AC, Lana SE, Hillers KR, Bachand AM, Avery PR. Canine lymphoproliferative disease characterized by lymphocytosis: immunophenotypic markers of prognosis. *J Vet Intern Med*. 2008;22(3):596-601.
11. Stokol T, Schaefer DM, Shuman M, Belcher N, Dong L. Alkaline phosphatase is a useful cytochemical marker for the diagnosis of acute myelomonocytic and monocytic leukemia in the dog. *Vet Clin Pathol*. 2015;44(1):79-93.
12. Kozlov I, Beason K, Yu C, Hughson M. CD79a expression in acute myeloid leukemia t(8;21) and the importance of cytogenetics in the diagnosis of leukemias with immunophenotypic ambiguity. *Cancer Genet Cytogenet*. 2005;163(1):62-67.
13. Lai R, Juco J, Lee SF, Nahirniak S, Etches WS. Flow cytometric detection of CD79a expression in T-cell acute lymphoblastic leukemias. *Am J Clin Pathol*. 2000;113(6):823-830.
14. Pantelyushin S, Ranninger E, Bettschart-Wolfensberger R, Vom Berg J. OMIP-065: dog immunophenotyping and T-cell activity evaluation with a 14-color panel. *Cytometry A*. 2020;97(10):1024-1027.
15. Golub TR, Slonim DK, Tamayo P, et al. Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. *Science*. 1999;286(5439):531-537.
16. Bacher U, Schnittger S, Haferlach C, Haferlach T. Molecular diagnostics in acute leukemias. *Clin Chem Lab Med*. 2009;47(11):1333-1341.
17. Bullinger L, Valk PJ. Gene expression profiling in acute myeloid leukemia. *J Clin Oncol*. 2005;23(26):6296-6305.
18. Rao S, Lana S, Eickhoff J, et al. Class II major histocompatibility complex expression and cell size independently predict survival in canine B-cell lymphoma. *J Vet Intern Med*. 2011;25(5):1097-1105.
19. Gelain ME, Mazzilli M, Riondato F, Marconato L, Comazzi S. Aberrant phenotypes and quantitative antigen expression in different subtypes of canine lymphoma by flow cytometry. *Vet Immunol Immunopathol*. 2008;121(3–4):179-188.
20. Rout ED, Shank AM, Waite AH, Siegel A, Avery AC, Avery PR. Progression of cutaneous plasmacytoma to plasma cell leukemia in a dog. *Vet Clin Pathol*. 2017;46(1):77-84.
21. Seelig DM, Avery P, Webb T, et al. Canine T-zone lymphoma: unique immunophenotypic features, outcome, and population characteristics. *J Vet Intern Med*. 2014;28(3):878-886.
22. Chabanne L, Marchal T, Kaplanski C, et al. Screening of 78 monoclonal antibodies directed against human leukocyte antigens for cross-reactivity with surface markers on canine lymphocytes. *Tissue Antigens*. 1994;43(3):202-205.
23. Hughes KL, Labadie JD, Yoshimoto JA, Dossey JJ, Burnett RC, Avery AC. Increased frequency of CD45 negative T cells (T zone cells) in older Golden retriever dogs. *Vet Comp Oncol*. 2018;16(1):E109-E116.
24. Reith W, LeibundGut-Landmann S, Waldburger JM. Regulation of MHC class II gene expression by the class II transactivator. *Nat Rev Immunol*. 2005;5(10):793-806.
25. Staudt LM, Lenardo MJ. Immunoglobulin gene transcription. *Annu Rev Immunol*. 1991;9:373-398.
26. Chapiro E, Delabesse E, Asnafi V, et al. Expression of T-lineage-affiliated transcripts and TCR rearrangements in acute promyelocytic leukemia: implications for the cellular target of t(15;17). *Blood*. 2006;108(10):3484-3493.
27. Hashimoto M, Yamashita Y, Mori N. Immunohistochemical detection of CD79a expression in precursor T cell lymphoblastic lymphoma/leukaemias. *J Pathol*. 2002;197(3):341-347.
28. Massari S, Bellahcene F, Vaccarelli G, et al. The deduced structure of the T cell receptor gamma locus in Canis lupus familiaris. *Mol Immunol*. 2009;46(13):2728-2736.
29. Keller SM, Moore PF. Rearrangement patterns of the canine TCRγ locus in a distinct group of T cell lymphomas. *Vet Immunol Immunopathol*. 2012;145(1–2):350-361.
30. Meichner K, Stokol T, Tarigo J, et al. Multicenter flow cytometry proficiency testing of canine blood and lymph node samples. *Vet Clin Pathol*. 2020;49(2):249-257.
31. Villiers E, Baines S, Law AM, Mallows V. Identification of acute myeloid leukemia in dogs using flow cytometry with myeloperoxidase, MAC387, and a canine neutrophil-specific antibody. *Vet Clin Pathol*. 2006;35(1):55-71.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

How to cite this article: Harris RA, Rout ED, Yoshimoto JA, Avery PR, Avery AC. Using digital RNA counting to establish flow cytometry diagnostic criteria for subtypes of CD34+ canine acute leukaemia. *Vet Comp Oncol*. 2022;20(3):710-719. doi:10.1111/vco.12825