

Research paper

A review of canine B cell clonality assays and primer set optimization using large-scale repertoire data



Mei-Hua Hwang^a, Nikos Darzentas^{b,c}, Dorothee Bienzle^a, Peter F. Moore^d, Franco Guscelli^e, Jodi Morrison^a, Stefan M. Keller^{a,*}

^a Department of Pathobiology, Ontario Veterinary College, University of Guelph, Guelph, Ontario, Canada

^b Department of Internal Medicine II, University Hospital Schleswig-Holstein, Kiel, Germany

^c Central European Institute of Technology, Masaryk University, Brno, Czech Republic

^d Department of Pathology, Microbiology and Immunology, School of Veterinary Medicine, University of California, Davis, California, USA

^e Institute of Veterinary Pathology, Vetsuisse Faculty, University of Zurich, Zurich, Switzerland

ARTICLE INFO

Keywords:

Clonality testing
PARR
Lymphoma
Dog
B cell lymphoma
Sensitivity
Specificity

ABSTRACT

Several molecular clonality assays have been developed to assess canine B cell proliferations. These assays were based on different sequence data, utilized different assay designs and employed different testing strategies. This has resulted in a complex body of literature and complicates evidence-based selection of primer sets. In addition, further refinement of primer sets is difficult because it is unknown how well current primer sets cover the expressed sequence repertoire. The objectives of this study were 1) to provide an overview of published IGH clonality assays that highlights key differences in assay design and testing strategy and 2) to propose a novel method for optimizing primer sets that leverages large-scale sequencing data. A review of previously published assays highlighted confounding factors that hamper a direct comparison of performance metrics between studies. These findings illustrate the need for a multi-institutional effort to harmonize veterinary clonality testing. A novel *in silico* analysis of primer sequences using a large dataset of expressed sequences identified shortfalls of existing primer sets and was used to guide primer optimization. Three optimized primer sets were tested and yielded qualitative sensitivity values between 80–90%. The qualitative sensitivity ranged from 1% to over 50% and was dependent on the size of the neoplastic clone and the sample DNA used. These findings illustrate that inclusion of high-throughput sequencing data for primer design can be a useful tool to guide primer design and optimization. This strategy could be applied to other antigen receptor loci or species to further improve veterinary clonality assays.

1. Introduction

Lymphoma is the most common hematopoietic neoplasm in dogs, but its diagnosis can be challenging. Molecular clonality testing is an adjunct to diagnosing canine lymphoma when conventional methods such as cytology or histopathology are inconclusive (van Dongen et al., 2003; Lana et al., 2006; Keller et al., 2016). Clonality testing exploits the phenomenon that lymphocytes generate antigen receptor genes by rearrangement of gene segments and addition of random nucleotides resulting in genes with variable in nucleotide composition and length. Clonality testing for B cell proliferations amplifies the hypervariable immunoglobulin heavy chain (IGH) gene by PCR followed by size separation of amplicons by slab-gel or capillary electrophoresis (van Dongen et al., 2003). The size distribution of amplicons is reflective of

the genetic diversity of rearranged IGH genes among lymphocytes within a given sample and can help differentiate between reactive and neoplastic lymphoproliferations (van Dongen et al., 2003; Gentilini et al., 2009; Langerak et al., 2012; Keller et al., 2016).

Primer sets are a pivotal component of clonality assays since they facilitate amplification of antigen receptor genes to levels that are detectable by gel electrophoresis. To design primer sets, knowledge of antigen receptor gene sequences is required. If a primer set does not cover all rearranged genes, it is prone to false negative results. Identification of antigen receptor gene sequences is not straightforward, and researchers have relied on two complementary strategies. Before publication of the dog genome assembly in 2005, researchers identified antigen receptor genes through cloning and sequencing of rearranged antigen receptor genes, which was effort intensive and low throughput.

* Corresponding author at: Ontario Veterinary College, 50 Stone Road, Guelph, Ontario, N1G 2W1, Canada.

E-mail address: smkeller@uoguelph.ca (S.M. Keller).

After publication of the dog genome assembly, studies could mine the genome assembly for germline genes, *i.e.* non-rearranged antigen receptor sequences. In 2010, the first annotation of the canine IGH heavy chain locus was published, which identified additional, previously unknown genes (Bao et al., 2010). A more recent IGH locus annotation from 2017 further expanded the number of known gene segments (Martin et al., 2017). In addition to identifying germline genes based on the dog genome assembly, few datasets of expressed sequences obtained by high-throughput sequencing have been published in recent years (Steiniger et al., 2014; Hwang et al., 2018). Together, these studies have gradually expanded our knowledge of canine IGH gene sequences and created the opportunity for further refinement of primer sets.

An additional challenge in primer design is the fact that the total number of gene sequences is further augmented by diversification of germline sequences by somatic hypermutation (SHM). Somatic hypermutation is a physiological process that increases antibody affinity by mutating antigen receptor genes (Kim et al., 1981; Tonegawa, 1983; Volpe and Kepler, 2008). Mismatches caused by SHM have been suggested to impede primer binding and amplification and have been implicated in false negative results (Burnett et al., 2003; Tamura et al., 2006; Valli et al., 2006; Gentilini et al., 2009; Stadhouders et al., 2010; Waugh et al., 2016). Studies in humans and mice have shown that certain nucleotide motifs, known as ‘mutational hotspots’, are particularly prone to SHM, but this phenomenon has yet to be studied in dogs (Odegard and Schatz, 2006; Yaari et al., 2013).

At least 16 different canine IGH primer sets have been described over the course of almost two decades, resulting in a complex body of literature. While most studies provided some indicator of test performance, it is unclear to what extent these metrics can be used to compare assays since no standardized testing procedure exists. Consequently, evidence-based selection of primer sets is currently challenging. Furthermore, existing literature suggests that a subset of neoplastic clones are missed by existing primer sets. Further optimization of primer sets might improve assay performance but the reasons for false negative results are unknown and targeted primer revisions hence difficult.

The objectives of this study were 1) to provide an overview of published IGH clonality assays that highlights key differences in assay design and testing strategy and 2) to propose a novel method for assessing and optimizing primer sets that leverages large-scale sequencing data. We show that 1) published clonality assays vary in assay design and testing scheme, which confounds direct comparison of performance metrics and 2) *in silico* primer testing of primer sets using large scale sequencing data can provide valuable clues for primer optimization.

2. Methods

2.1. Review of the literature

The literature was searched for canine clonality assays targeting the IGH, focusing on the first descriptions of the assays as well as a recent combined assay derived from previously described primer sets (Burnett et al., 2003; Tamura et al., 2006; Valli et al., 2006; Gentilini et al., 2009; Waugh et al., 2016). All studies were assessed in regards to (1) the sequence data used for primer design, (2) the number of tubes per primer set, (3) the use of replicate samples, (4) the samples used for testing, (5) the electrophoresis method and (6) the definition of clonality.

2.2. *In silico* optimization of primer sets

The ability of five previously described primers sets to bind to a diverse repertoire of IGH sequences was gauged *in silico* by recording the mismatches between every sequence of a NGS dataset and the best-matching primer sequence using a Perl script (Burnett et al., 2003;

Tamura et al., 2006; Valli et al., 2006; Gentilini et al., 2009; Waugh et al., 2016). The NGS dataset has been described previously and consisted of 2,835,707 IGH reads with identifiable V and J genes that were derived from mRNA of spleen, lymph node and bone marrow of three healthy dogs (Hwang et al., 2018; GenBank Sequence Read Archive (SRA) accession number SUB4105450). IGH gene sequences were amplified using a 5'RACE approach with multiplexed reverse primers targeting the 5' end of all canine IGH constant regions. Sequencing was carried out on an Illumina MiSeq with PE600. The results of the *in silico* assessment were then used to guide the re-design of primer sets to better fit the expressed repertoire of rearranged canine IGH genes. For positions with a high number of mismatches between primer and dataset, primer sets were modified either by introduction of degenerate positions or addition of new primers. The IMGT gene nomenclature was used in this study.

2.3. *In vitro* testing of primer sets

Three optimized primer sets with different amplicon sizes were selected for *in vitro* assessment: optimized Gentilini 3FWR1, optimized Burnett IgH major/minor and Valli (Burnett et al., 2003; Valli et al., 2006; Gentilini et al., 2009). All primer sets were tested using DNA from formalin-fixed/paraffin-embedded (FFPE) tissues as well as DNA extracted from cytology slides. The FFPE samples consisted of 34 reactive lymph nodes and 30 B cell neoplasms without recognizable non-neoplastic B cells or plasma cells. The diagnosis was based on a combination of clinical history, histopathology and immunohistochemistry. The B cell lineage of all neoplastic cases was determined by reactivity for CD79 and/or MUM-1 and lack of reactivity for CD3 as assessed by immunohistochemistry. Only cytoplasmic reactivity of CD79 was interpreted as specific and positive. Of the neoplastic samples, 25 were lymph nodes with a diagnosis of diffuse large-B cell lymphoma, and five were cutaneous plasmacytomas. Reactive cases consisted of lymph nodes from dogs with osteosarcoma. Sample materials stemmed from the archives of the Animal Health Laboratory/Department of Pathobiology, Ontario Veterinary College, Canada or the Vetsuisse Faculty, University of Zürich, Switzerland. DNA extracted from unfixed, frozen mandibular lymph node tissue from a healthy animal was used as polyclonal control. The cytology samples consisted of 10 cytology slides from lymph node fine needle aspirates of dogs with B cell lymphoma, which were obtained from the diagnostic flow cytometry service at the Animal Health Laboratory, University of Guelph. A diagnosis of lymphoma had been made based on cytology and immunophenotyping was carried out for the purpose of further characterizing the neoplastic cells. The flow cytometry methods have been described previously (Reggeti and Bienzle, 2011).

DNA was extracted from either paraffin scrolls (FFPE samples) or cells that were scraped off slides (cytology samples) (Blood & Tissue kit and QIAamp DNA Micro Kit, Qiagen, CA). All primer sets were initially tested in a gradient PCR to optimise annealing temperature. The final cycling conditions consisted of an initial activation step of 95 °C for 15 min; 35 cycles of 95 °C for 30 s, 65 °C for 30 s, 72 °C for 30 s and a final extension time at 72 °C for 2 min. Each 25 µL reaction contained 0.38 µL HotStarTaq (Qiagen, CA), 5.0 µL of MgCl₂ (Qiagen, CA), 1.5 µL of 10x buffer (Qiagen, CA), 1.5 µL of 10 mM dNTPs (MilliporeSigma, MO), 0.5 µL each of 10 µM forward and reverse primers, 10.38 µL of water and 5 µL of 10 ng/µL DNA template.

Amplicons were visualized by capillary electrophoresis (Qiaxcel, Qiagen) using the High Resolution Cartridge, the run parameters OL500 with 30 s injection time, and the Qiaxcel ScreenGel software. The electrophoresis patterns were described as follows: ‘clonal’: one or two reproducible peaks of the expected size without concomitant polyclonal background; ‘clonal in polyclonal background’ one or two peaks within a polyclonal curve; ‘polyclonal’: a normally distributed curve. If a sample failed to amplify with any of the IGH primer sets, the DNA integrity was assessed using a multiplexed primer set that produced 4

products of differing size (Supplemental Table 1). If a sample failed to yield a band of the appropriate size in this amplification, the sample was excluded from further analysis.

Qualitative sensitivity was calculated as the fraction of samples with a clonal result out of all neoplastic test cases. Specificity was calculated two ways: 1) the fraction of samples with a polyclonal or ‘no amplification’ result out of all reactive test cases; and 2) the fraction of samples with a polyclonal result out of all reactive test cases. For the primer set with the best combination of sensitivity and specificity, the quantitative sensitivity was assessed using a serial dilution of clonal DNA in polyclonal DNA at a concentration of 10 ng/μL each. This experiment was performed three times using clonal DNA from three neoplastic samples (samples 21, 23, and 26) and polyclonal DNA from one reactive sample (sample 48). Proportions of clonal DNA tested were 100%, 50%, 25%, 10%, 5%, 1%, 0.1% and 0%. The quantitative sensitivity was defined as the lowest concentration of clonal DNA for which the height of the clonal peak was at least twice the height of the polyclonal background.

3. Results

3.1. Review of the literature

Five previously described clonality assays targeting the canine IGH locus were reviewed (Table 1, Fig. 1). The sequence data used for assay design ranged from few cDNA clones (Burnett) over 52 and 61 V genes deduced from the dog genome assembly (CanFam2.0) (Tamura, Gentilini) to one study that considered the first canine IGH locus description by Bao et al. (Waugh). Tamura and Valli combined all primers into one tube whereas Burnett split primers into two separate tubes (Fig. 1). Gentilini designed two new variable gene primers 5' of existing primers and combined these with a new and a previously described J primer each (Fig. 1). Waugh mixed previously described primers in various combinations distributed over 6 or 8 tubes (Fig. 1). For assay testing, studies used DNA extracted from either fresh samples (Burnett, Tamura, Gentilini) or formalin-fixed and paraffin-embedded (FFPE) samples (Valli) or both (Waugh). Diagnoses were either obtained by cytology, histology or both. One study classified lymphomas according to WHO criteria (Valli), all other studies did not further classify lymphoid neoplasms. The immunophenotype was determined for all samples in the studies by Valli and Tamura, the other studies determined the immunophenotype for a fraction of cases. The latter studies tested B and T cell assays concomitantly and reported assay sensitivity for B and T cell assays combined. Electrophoresis methods included agarose and PAGE slab gels (Burnett, Valli, Tamura) and capillary electrophoresis instruments (Gentilini, Waugh). The definition of clonality mainly differed by

Table 1

Review of previously described B cell clonality assays including aspects, which may impact the functional parameters of the assay. An “x” indicates that the design limitation was present in the study according to published literature.

| Experimental Design | Burnett et al. (2003) | Valli et al. (2006) | Tamura et al. (2006) | Gentilini et al. (2009) | Waugh et al. (2016) |
|--|-----------------------|--------------------------|------------------------------|----------------------------|----------------------|
| Sequence data for primer design | 5 cDNA clones | n.a. | GA ₁ , 6 human Js | GA ₁ , human Vs | GA ₂ |
| Number of tubes | 2 | 1 | 1 | 1 | 8/10 |
| Replicates mentioned | N | Y | Y | Y | N |
| Separation method | Slab gel (PAGE) | Slab gel (agarose, PAGE) | Slab gel (agarose) | ABI Prism 310 | ABI 3130 |
| Included neoplastic samples with unknown immunophenotype | Y | N | N | Y | Y |
| Number of tested samples | 77 + 24 | 42 + 1 | 5 | 96 | 185 |
| Qualitative sensitivity | 91% | 78.6% | 80% | 97.9% | 86.5% ³ |
| Specificity | n/a | n/a | n/a | n/a | 98.7% |
| Sensitivity/specificity pooled across multiple primer sets. ³ | 1B, 1T | n/a | n/a | 4B, 1T | 6B, 4 T ₄ |

¹CanFam2.0, AF354264, NW_139865.

²CanFam3.1, AF354264, NW_003726071.

³Two additional primer sets for both B and T cell loci were described (totalling 8 B cell and 6 T cell primer sets), and if also used, increased sensitivity to 87.6%.

⁴Whole-assay sensitivity was determined by the number of samples for which a clonal result was obtained with at least one of multiple B or T cell primer sets. The specificity reported by Waugh et al. (2016) was similarly determined by the absence of clonal results using either B cell or T cell primer sets in cases that were not diagnosed as lymphoid malignancies.

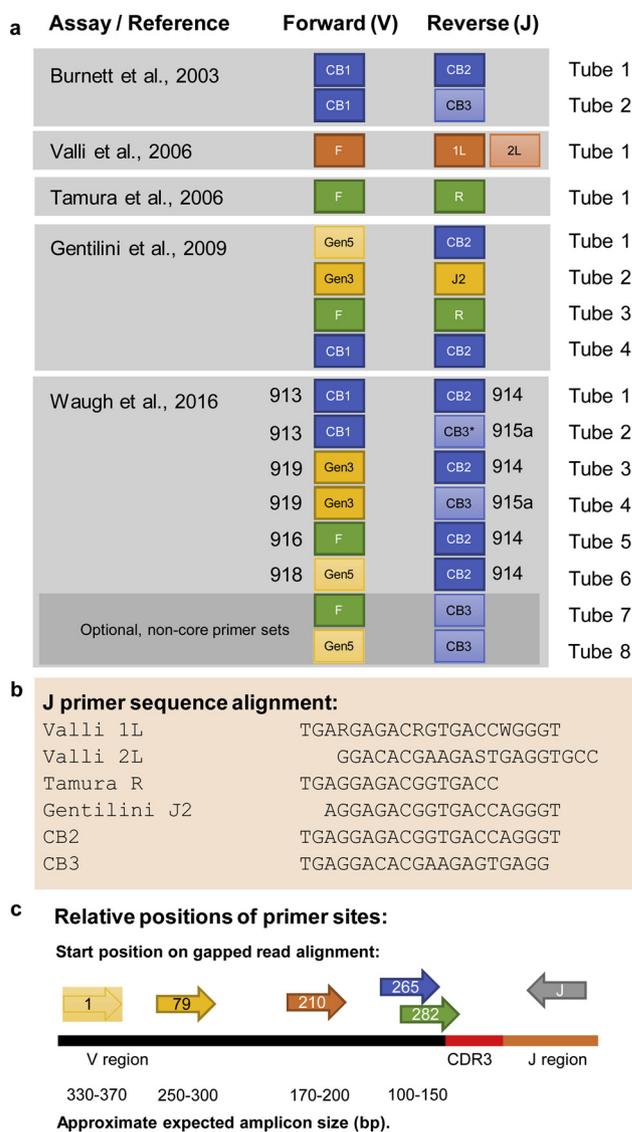


Fig. 1. Primer set analysis of 5 previously described clonality assays. (a) Schematic of forward and reverse primer combinations and their distribution over different tubes; (b) alignment of J reverse primer sequences; (c) position and expected amplicon size (bp) of the 5 forward primers with common J primers.

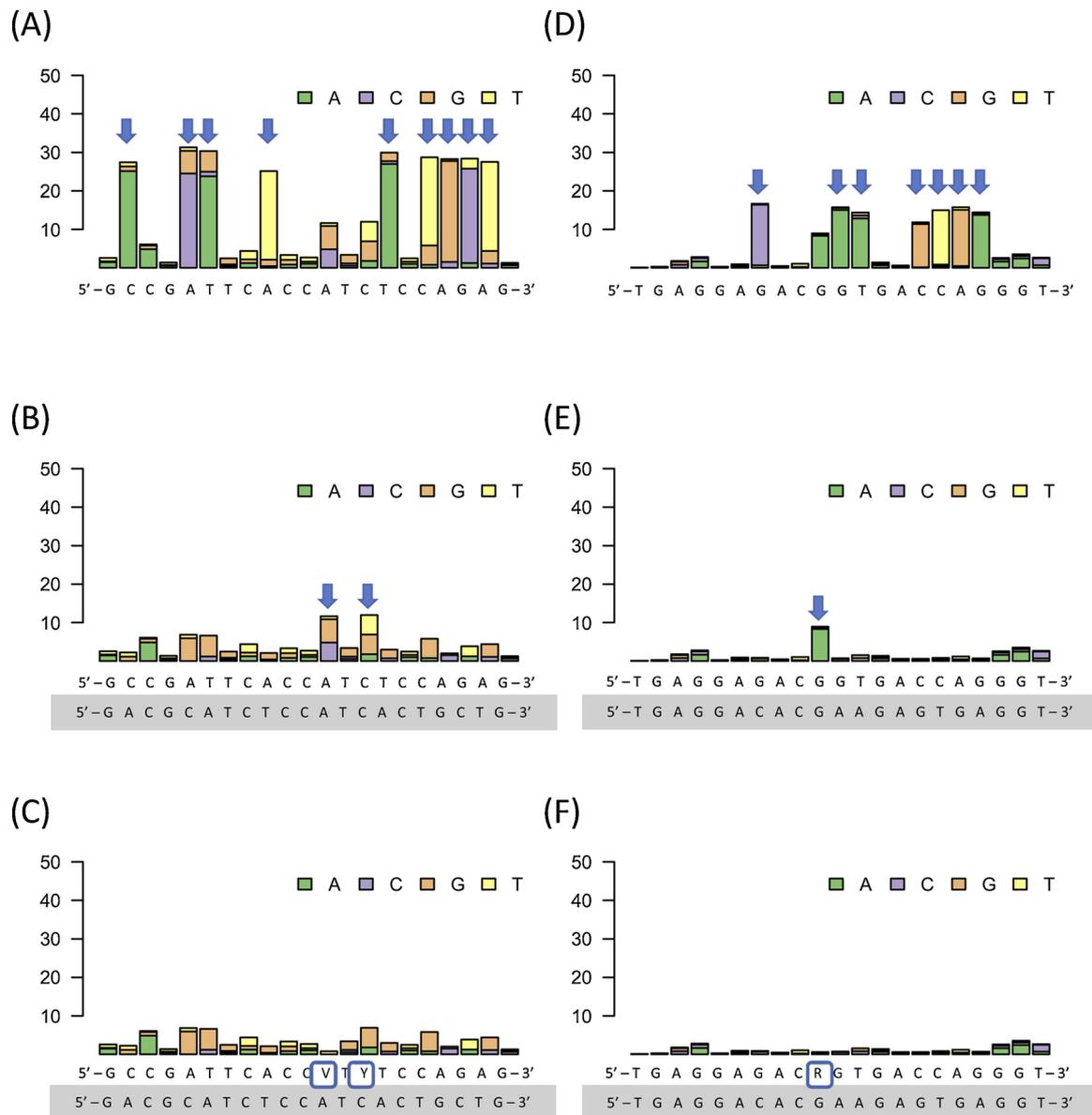


Fig. 2. *In silico* optimization of primers by reducing the mismatches between primer sequences and a NGS dataset of expressed IGH sequences of the Valli forward (a–c) and Valli reverse primers (d–f). Original primer sequences without degenerate bases (top); after addition of a primer for gene IGHV4-1 (b) or gene IGHJ6 (e) (middle); after incorporating of additional degeneracies (bottom; nucleotide changed in blue boxes). Arrow indicate sites with reduced mismatches with consecutive optimisation. X-axis: primer sequence 5'-3'. Y-axis: percent frequency of mismatches; A: adenine, C: cytosine, G: guanine, T: thymine; see Supplemental Fig. 1 for assessment of other primer binding sites (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article).

whether or not a numerical cut-off was given for the height of the neoplastic peak over the polyclonal background (which was not possible for studies using slab gels). Sensitivity values ranged from 86.7% to 98.7%; the quantitative sensitivity was assessed in one study only (Burnett).

3.2. *In silico* optimization of primer sets

The suitability of primer sets to recognize a diverse repertoire of IGH sequences was gauged using a set of expressed IGH gene sequences obtained by NGS (Hwang et al., 2018). Every sequence of the data set was aligned with the best-matching primer of a primer set and mismatches were recorded between the sequence and the primer for every position. Two main patterns were observed: First, certain positions exhibited high frequency (> 10%) of a single mismatch, and the magnitude of this mismatch was similar across affected positions (Fig. 2A and D). These mismatches arose because no primer set covered

gene IGHV4-1, the only gene in the V gene subgroup IGHV4. This effect was less marked for the Tamura primer set because the homology between V genes of different subgroups was greater at this position. Incorporating a primer covering IGHV4-1 reduced this type of mismatch drastically (Fig. 2B and E). Second, a low level of nucleotide mismatch (< 10%) was observed across all positions. These mismatches could be alleviated by introduction of degenerate bases into a primer (Fig. 2C and F, Supplemental Figs. 1 and 3). A primer for family IGHV1 was not included because it was not frequently used in the NGS dataset.

An alignment of reverse primers showed that primers from all assays were located within a 24 bp stretch of the J genes (Fig. 1B). Consequently, only the 'Ig 1 L-reverse' primer from Valli et al. (2006) was evaluated by *in silico* assessment since it spanned the most common start and end positions across J primers for all assays, and was reportedly compatible with multiplexing. Mismatches between the expressed sequences and the Ig 1 L-reverse primer-binding site arose primarily due to lack of coverage of the IGHJ6 gene (Fig. 2D). Adding a

Table 2
Original and modified V forward and J reverse primers for canine IGH.

| Primer Site | Assay / Reference | Genes Covered | Target Region | Sequences (5'-3') ¹ |
|---------------------------|-------------------------|------------------------------------|---------------|---|
| 5FWR1 | Gentilini et al. (2009) | IGHV3 IGHV3 IGHV4 | FR1 | GAGGTGCAGCTGGTGGAGTCT GAGGTGCAR <u>Y</u> TGGTGGAR <u>T</u> CT GA <u>A</u> CTCACACTGCAGGAGTCA |
| 3FWR1 ³ | Gentilini et al. (2009) | IGHV3 IGHV3 IGHV4 | FR1 | GCCTCTGGATTACCTTCAG GCCTCTGGATTACCTTCAR GTGTCCGGAGRCCTCCGTACCRV ² |
| Valli IGH 1U ³ | Valli et al. (2006) | IGHV3 IGHV3 IGHV4 | FR3 | GMCGVTTCCACATCTCCARRG GCCGATTCACCVTYTCCAGAG GACGCATCTCCATCACTGCTG |
| CB ³ | Burnett et al. (2003) | IGHV3 IGHV3 IGHV4 | FR3 | CAGCCTGAGAGCCGAGGACAC CAGCCTGAGAGYGA <u>V</u> GACAC CTCCATGACCA <u>C</u> CGAVGACAC |
| Tamura F | Tamura et al. (2006) | IGHV3/ IGHV4 IGHV3/ IGHV4 | FR3 | ACACGGCCVTGTATTACTGT ACACGGCYR <u>T</u> NATTA <u>T</u> Y <u>T</u> GT |
| Valli IGH 1L ³ | Valli et al. (2006) | IGHJ1-5 IGHJ6 IGHJ1-5 | J | TGARGAGACRGTGACCWGGGT TGAGGACACGAAGAGTGGGT TGAGGAGACRGTGACCAGGGT |

FR = framework region.

¹ Bolded sequences are the original primers described in the referenced literature; non-bolded sequences are modified primers with positions altered from the original primer sequence underlined. For the degree of mismatch at each position and the effect of modified primers, see Supplemental Fig. 1.

² The germline sequence for IGHV4-1 contains a 3bp insertion ('GGC') in positions 10–12 that are not found in other V genes at that site.

³ Modified primer sets were selected for *in vitro* assessment by PCR.

primer specific to IGHJ6 greatly reduced the frequency of mismatches (Fig. 2 E). Adding a degenerate base at position 10 accounted for the most common mismatches between the original reverse primer and the remaining J genes (Fig. 2F).

To assess the benefit of degeneracy in the original primer sets, these positions were reverted back to the most common nucleotide observed across all germline genes. In most cases, this did not have a substantial impact on the degree of primer mismatch (data not shown).

3.3. *In vitro* testing of primer sets

The optimized versions of the primer sets described by Burnett, Valli and Gentilini were tested on FFPE tissues from 30 neoplastic and 34 reactive samples and 10 neoplastic cytology samples (Table 2). From the FFPE cohort, three neoplastic samples and one reactive sample neither yielded amplification with any of the IGH primer sets nor with a primer set for assessment of DNA integrity. These samples were hence excluded from further analysis due to compromised DNA quality. Additionally, one neoplastic sample was excluded because of polyclonal background that had not been appreciated during the histological review.

Out of the 26 remaining neoplastic FFPE samples, the optimized Burnett, Gentilini and Valli primer sets yielded clonal results in 21 (80.8%), 11 (42.3%) and 22 (84.6%) samples, respectively (Fig. 3, Table 3, Supplemental tables 2–3, Supplemental Fig. 2). In instances where clonality was not detected, the electrophoresis result was 'no amplification'. No pseudoclonal results were observed. When stratified according to tumor type, the Burnett and Valli primer sets yielded clonal results in 2/3 plasmacytomas and the Gentilini primer sets yielded clonal results in 0/3 plasmacytomas. Out of 10 neoplastic cytology samples, clonal results were obtained in 8 (80.0%), 9 (90.0%) and 9 (90.0%) samples using the Burnett, Gentilini and Valli primer sets, respectively.

Out of 33 reactive samples, the original Burnett, Gentilini and Valli primer sets yielded polyclonal results in 30 (90.9%), 21 (63.6%) and 29 (87.9%) samples, respectively (Fig. 3, Table 3, Supplemental table 4,

Supplemental Fig. 2). As with neoplastic samples, the electrophoresis result was 'no amplification' in all cases in which clonality was not detected, and no pseudoclonal results were observed. Both reactive and neoplastic samples amplified with the Burnett primer sets yielded reproducible, faint bands at 96–100bp. Due to their consistent occurrence regardless of sample type, these bands were disregarded.

Out of all primer sets tested, the modified Valli primer set had the highest sensitivity and was hence selected for quantitative sensitivity testing. Clonal DNA from three neoplastic samples each (cases 21, 23, and 26) was diluted with polyclonal DNA from a reactive sample (case 48). The height of the clonal peak (electropherogram view) and the intensity of the band (gel view) decreased as the concentration of the neoplastic DNA template decreased (Fig. 4). The lowest concentration at which the clonal peak height was at least twice the polyclonal background varied from 10% to 50% clonal DNA, and was dependent on the location of the clonal peak relative to the crest of the polyclonal curve.

4. Discussion

The evolution of clonality testing for dogs has resulted in a multitude of B cell assays that were designed at different institutions using different sequence data, PCR strategies, electrophoresis methods and interpretational guidelines. As a result, the pertinent literature is disjointed and choosing 'the best' assay based on objective criteria is challenging. In addition, the results of previous studies suggest that a subset of neoplastic clones is not detected by current assays. However, the reasons for this phenomenon are unknown, which precludes further optimization of primer sets. The current study aimed at providing an overview of existing clonality assays and at exploring a new method to identify and mitigate shortfalls of current primer sets using large scale expressed repertoire data.

Since existing clonality assays have been developed over the course of almost two decades, one could speculate that the growing amount of sequence data had resulted in an improved primer coverage and hence better performance of assays over time. Based on nucleotide alignments of available primer sequences with the most current set of IMGT germline genes it appears that primers align well with the largest V gene subgroup IGHV3, that primers align moderately well with V gene subgroup IGHV1 genes depending on the gene and the primer set, and that primers align poorly with the single member subgroup gene IGHV4-1. These findings indicate that although assays differ in the location of primers within V genes, they do not differ fundamentally in regards the V gene subgroups they cover. Insufficient primer coverage of subgroup IGHV1 and IGHV4 genes might hence contribute to false negative results with existing canine IGH clonality assays. However, the magnitude of this effect is dependent on the rearrangement frequency of these genes *in vivo*. The fact that neither of the IGHV1 and IGHV4 subgroup genes was dominant in cDNA analyses by prior studies suggests that these genes represent a minor fraction of all rearranged genes. However, since previous studies only analysed a relatively low number of clones, these studies only provide a crude resolution of the rearrangement frequency of IGH genes.

Studies applied different strategies in regards to the distribution of primers across tubes. Primers were either combined in a single tube (Valli, Tamura) or distributed over multiple tubes (Burnett, Gentilini, Waugh). Distributing primers over multiple tubes reduces the likelihood of negative primer interactions or unspecific amplification, which is helpful when pooling large numbers of primers (VanDongen et al., 2003). On the other hand, amplifying a sample with primers that only recognize a subset of gene segments can disproportionately amplify insignificant clones and increase the likelihood of false positive results (Cushman-Vokoun et al., 2010). If the latter is of significance for assays currently used in veterinary medicine is unclear but warrants further investigation.

Important performance metrics for biomedical laboratory tests are

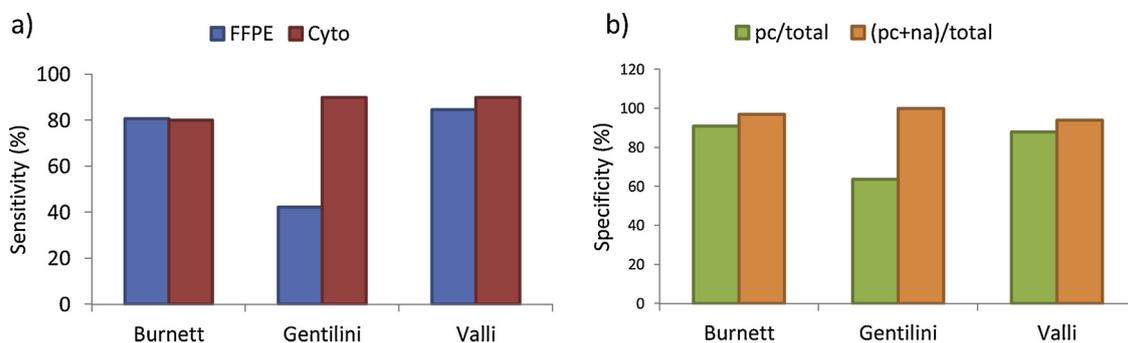


Fig. 3. Sensitivity (a) and specificity (b) for each tested primer set. Sensitivity was calculated for FFPE (n = 26) and cytology (n = 10) samples separately; Specificity was calculated either as fraction of polyclonal results (pc/total) or polyclonal plus ‘no amplification’ ((pc + na)/total) out of all reactive cases. All reactive cases (n = 33) were from FFPE samples. See supplemental data for individual sample results.

sensitivity and specificity. The reported sensitivity values for the reviewed assays ranged from 78.6% to 97.9%. However, several differences in testing strategies between studies were evident. First, these numbers represent additive values across all tubes of an assay, the number of which ranged from 1 tube (Valli, Tamura) to 4 tubes (Gentilini) and 6/8 tubes (Waugh). For multi-tube assays, sensitivity values are lower when considered on a per-tube basis. Second, studies differed in regards to whether or not replicate samples were used to identify pseudoclonality. Omitting replicate samples can result in overestimating assay sensitivity because pseudoclonal results might be interpreted as clonal (Böer et al., 2008). Third, samples differed in the extent to which immunophenotyping was done. This pertained to studies that assessed B cell and T cell assays simultaneously and where clonal test samples consisted of both B cell and T cell neoplasms without knowledge of the immunophenotype. Consequently, these studies reported the percentage of lymphomas that were clonal with the tested IGH primer set. This testing strategy allows establishing the percentage of clonal samples but cannot distinguish if a clonal IGH result was produced by a B cell lymphoma, a T cell lymphoma with a cross-lineage rearrangement or if it was a false positive result if the study did not use replicate samples to identify pseudoclonality. Fourth, studies differed in whether they distinguished between quantitative and qualitative sensitivity testing. In cases with significant polyclonal background, assay sensitivity might have been underestimated because a clonal amplification might have been obscured by polyclonal background. The magnitude and potential effect of non-neoplastic lymphocytes in neoplastic test samples was generally not discussed in any of the studies. Fifth, at least one study included ‘clinical’ case material for which a definite diagnosis had not been established. Lastly, studies varied in how clonality was defined. Some studies used a cut-off for the ratio of peak height to polyclonal background while others did not. To what degree these considerations confounded the reported sensitivity is unknown. However, the differences in assay testing suggest that direct comparison of performance metrics that were generated in different studies is problematic.

Table 3

PCR test results optimised primer set using neoplastic and reactive samples. Italicised numbers on the right of the columns are percentages (%) of total neoplastic or reactive cases.

| Primer set | Neoplastic FFPE samples (n = 26) | | | | | | Neoplastic cytology samples (n = 10) | | | | | | Reactive FFPE samples (n = 33) | | | | | | | |
|------------------------|----------------------------------|-------|----|-------|--------|------|--------------------------------------|-------|----|-------|--------|------|--------------------------------|------|----|-------|----|-------|--------|------|
| | Clonal | | NA | | Pseudo | | Clonal | | NA | | Pseudo | | Clonal | | PC | | NA | | Pseudo | |
| Gentilini ¹ | 11 | 42.3% | 15 | 57.7% | 0 | 0.0% | 9 | 90.0% | 1 | 10.0% | 0 | 0.0% | 0 | 0.0% | 21 | 63.6% | 12 | 36.4% | 0 | 0.0% |
| Burnett | 21 | 80.8% | 5 | 19.2% | 0 | 0.0% | 8 | 80.0% | 2 | 20.0% | 0 | 0.0% | 1 | 3.0% | 30 | 90.9% | 2 | 6.1% | 0 | 0.0% |
| Valli | 22 | 84.6% | 4 | 15.4% | 0 | 0.0% | 9 | 90.0% | 1 | 10.0% | 0 | 0.0% | 2 | 6.1% | 29 | 87.9% | 2 | 6.1% | 0 | 0.0% |

PC = Polyclonal result.

NA = No target amplification or amplification too weak for interpretation.

Pseudo = Pseudoclonal result.

¹Primer 3FWR1.

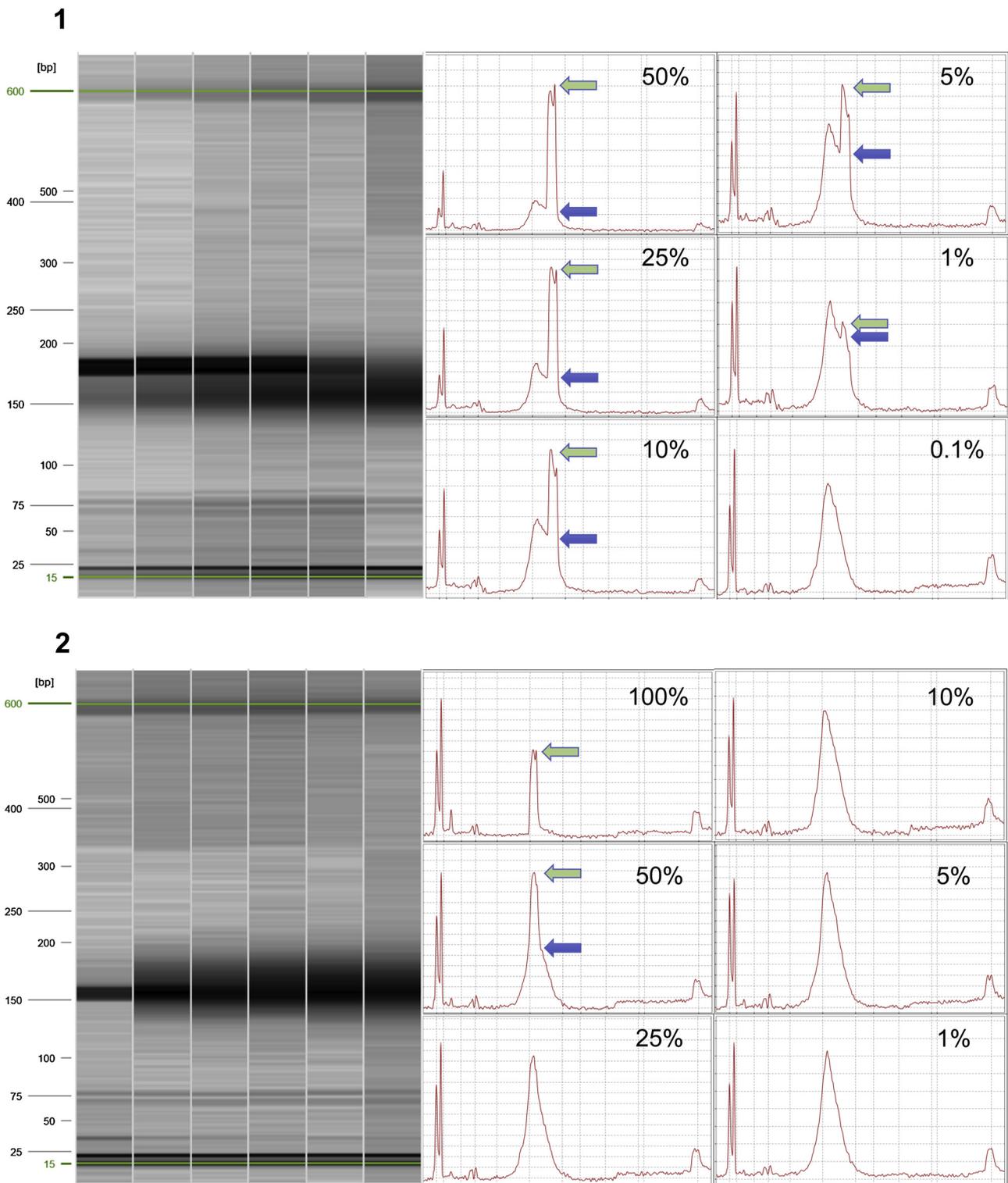


Fig. 4. Electrophoresis of PCR amplification of dilution series of DNA from clonal/neoplastic samples in DNA from a polyclonal/reactive sample using the modified Valli primer set. Gel view lanes on the left are arranged from highest concentration of clonal DNA to lowest, corresponding with electropherogram images on the right. 1) Case 21; the clonal peak is offset from the polyclonal smear giving a distinct clonal peak down to 10% neoplastic DNA. 2) Case 23; the clonal peak coincides with the crest of the polyclonal curve resulting in a distinct clonal peak down to 50% neoplastic DNA only. A clonal peak at least twice the size of the polyclonal background was considered ‘distinct’. For case 26, see Supplemental Fig. 3. blue arrow: height of the polyclonal background, green arrow: height of the neoplastic peak (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article).

primers were optimized based on a set of over 2 million unique rearranged IGH sequences from 3 different lymphoid organs from 3 different individuals. Given the lack of publicly available NGS IGH datasets for the dog, a limitation of this study is that it is unknown how

representative these sequence data are of the dog population in general. Further modification of primer sets might be indicated as more high-throughput sequencing data will become available over time.

Following primer optimization *in silico*, three optimized primer sets

were tested using a series of FFPE samples and a smaller cohort of neoplastic cytology samples. The sensitivity values obtained for the primer sets developed in the current study ranged from 80.0%–90.0%, with the exception of the optimized Gentilini primer set that yielded a sensitivity of 42.3% on FFPE samples. This is most likely attributable to the fact that the gene fragment amplified by the Gentilini primer set is larger than that of the Burnett and Valli primer sets, which makes this primer set less suitable for samples with compromised DNA integrity. This notion is supported by the fact that this primer set performed equally well on fresh samples as the other two primer sets.

The quantitative sensitivity was tested on one primer set and varied markedly depending on the location of the clonal peak relative to the crest of the polyclonal curve. Previously, only one study had assessed the quantitative sensitivity and the lowest detectable concentration of clonal DNA was 0.1% clonal DNA using slab gel electrophoresis (Burnett et al., 2003). In the current study, the lowest quantitative sensitivity was 10% using capillary electrophoresis and a clonal peak height of at least twice the height of the polyclonal curve. Of note, a peak height cut-off was used for the purpose of quantitative sensitivity testing but, in the opinion of the authors, it is not appropriate for use in clinical samples. With one sample, a clonal peak was quenched by polyclonal background at a proportion of greater than 50% clonal DNA. Neoplastic samples in which the clonal peak overlapped in amplicon size with the crest of the reactive polyclonal curve had higher detection limits than samples in which the clonal peak was closer to the edge of the polyclonal curve. The substantially lower detection threshold of the original Burnett assay could be due to the two-tube assay design. Since each tube only amplifies a subset of all rearrangements, polyclonal background is amplified less efficiently than in a tube that amplifies all possible rearrangements. Alternatively, DNA integrity differences between the clonal and polyclonal DNA could influence qualitative sensitivity. While slab gel electrophoresis yielded a lower detection threshold in this case, it should be stated that the length of the capillary when using capillary electrophoresis might affect the quantitative sensitivity. These findings illustrate how easily a neoplastic clone can be missed in samples with concomitant reactive lymphocytes. They also underscore the importance of integrating microscopy findings to adjust the level of confidence in a given electrophoresis result based on the degree of polyclonal background observed. Lastly, these findings illustrate that clonality testing is neither objective nor informative in all cases. Especially in cases with polyclonal background that result in peaks of unknown significance, the electrophoresis profile should be interpreted with caution. Next generation sequencing-based clonality testing might be helpful in resolving some of these complex cases in the future.

In summary, this study identified differences in assay design and testing strategy of published canine B cell clonality assays, which hamper a direct comparison of performance metrics. These findings illustrate the need for a multi-institutional and standardized effort to harmonize veterinary clonality testing and, in the authors' opinion, the confounding factors identified in this study should be considered in any harmonization efforts moving forward. Furthermore, the study identified potential shortfalls in primer coverage by existing assays using a novel *in silico* strategy that leverages high-throughput sequencing data. This strategy could be applied to guide primer optimization for other antigen receptor loci or species to further minimize false negative test results.

Declaration of conflicting interests

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

Acknowledgements

The authors acknowledge the Animal Health Laboratory, Guelph,

Canada for excellent support regarding histology and immunohistochemistry and the Bioinformatics Analysis Team (<http://bat.infspire.org/>) under the guidance of Nikos Darzentas, Brno, Czech Republic for their bioinformatics support. This study was funded by Pet Trust grant #051699. Computational resources were funded by the Ministry of Health of the Czech Republic grant nr. 16-34272A. Computational resources were provided by the CESNET LM2015042 and the CERIT Scientific Cloud LM2015085, provided under the program "Projects of Large Research, Development, and Innovations Infrastructures".

References

- Bao, Y., Guo, Y., Xiao, S., Zhao, Z., 2010. Molecular characterization of the VH repertoire in *Canis familiaris*. *Vet. Immunol. Immunopathol.* 137, 64–75.
- Böer, A., Tirumalae, R., Bresch, M., Falk, T.M., 2008. Pseudoclone in cutaneous pseudolymphomas: a pitfall in interpretation of rearrangement studies. *Br. J. Dermatol.* 159 (Aug. (2)), 394–402.
- Burnett, R.C., Vernau, W., Modiano, J.F., Olver, C.S., Moore, P.F., Avery, A.C., 2003. Diagnosis of canine lymphoid neoplasia using clonal rearrangements of antigen receptor genes. *Vet. Pathol.* 40, 32–41.
- Cushman-Vokoun, A.M., Connealy, S., Greiner, T.C., 2010. Assay design affects the interpretation of T-cell receptor gamma gene rearrangements: comparison of the performance of a one-tube assay with the BIOMED-2-based TCRG gene clonality assay. *J. Mol. Diagn.* 12, 787–796.
- Gentilini, F., Calzolari, C., Turba, M.E., Bettini, G., Famigli-Bergamini, P., 2009. GeneScanning analysis of Ig/TCR gene rearrangements to detect clonality in canine lymphomas. *Vet. Immunol. Immunopathol.* 127, 47–56.
- Hwang, M., Darzentas, N., Bienzle, D., Moore, P.F., Morrison, J., Keller, S.M., 2018. Characterization of the canine immunoglobulin heavy chain repertoire by next generation sequencing. *Vet. Immunol. Immunopathol.* 202, 189–190.
- Keller, S.M., Vernau, W., Moore, P.F., 2016. Clonality testing in veterinary medicine: a review with diagnostic guidelines. *Vet. Pathol.* 53, 711–725.
- Kim, S., Davis, M., Sinn, E., Patten, P., Hood, L., 1981. Antibody diversity: somatic hypermutation of rearranged VH genes. *Cell* 27, 573–581.
- Lana, S.E., Jackson, T.L., Burnett, R.C., Morley, P.S., Avery, A.C., 2006. Utility of polymerase chain reaction for analysis of antigen receptor rearrangement in staging and predicting prognosis in dogs with lymphoma. *J. Vet. Intern. Med.* 20, 329–334.
- Langerak, A.W., Groenen, P.J.T.A., Brüggemann, M., Beldjord, K., Bellan, C., Bonello, L., Boone, E., Carter, G.I., Catherwood, M., Davi, F., Delfau-Larue, M.-H., Diss, T., Evans, P.A.S., Gameiro, P., Garcia Sanz, R., Gonzalez, D., Grand, D., Håkansson, Å., Hummel, M., Liu, H., Lombardia, L., Macintyre, E.A., Milner, B.J., Montes-Moreno, S., Schuurings, E., Spaargaren, M., Hodges, E., van Dongen, J.J.M., 2012. EuroClonality/BIOMED-2 guidelines for interpretation and reporting of Ig/TCR clonality testing in suspected lymphoproliferations. *Leukemia* 26, 2159–2171.
- Martin, J., Pongstingl, H., Lefranc, M.P., Archer, J., Sargan, D., Bradley, A., 2017. Comprehensive annotation and evolutionary insights into the canine (*Canis lupus familiaris*) antigen receptor loci. *Immunogenetics* 1–14.
- Odegard, V.H., Schatz, D.G., 2006. Targeting of somatic hypermutation. *Nat. Rev. Immunol.* 6, 573–583.
- Reggati, F., Bienzle, D., 2011. Flow cytometry in veterinary oncology. *Vet. Pathol.* 48, 223–235.
- Stadhouders, R., Pas, S.D., Anber, J., Voermans, J., Mes, T.H.M., Schutten, M., 2010. The effect of primer-template mismatches on the detection and quantification of nucleic acids using the 5' nuclease assay. *J. Mol. Diagn.* 12, 109–117.
- Steinger, S.C.J., Dunkle, W.E., Bamert, G.F., Wilson, T.L., Krishnan, A., Dunham, S.A., Ippolito, G.C., Bainbridge, G., 2014. Fundamental characteristics of the expressed immunoglobulin VH and VL repertoire in different canine breeds in comparison with those of humans and mice. *Mol. Immunol.* 59, 71–78.
- Tamura, K., Yagihara, H., Isotani, M., Ono, K., Washizu, T., Bonkobara, M., 2006. Development of the polymerase chain reaction assay based on the canine genome database for detection of monoclonality in B cell lymphoma. *Vet. Immunol. Immunopathol.* 110, 163–167.
- Tonegawa, S., 1983. Somatic generation of antibody diversity. *Nature* 302, 575–581.
- Valli, V.E., Vernau, W., de Lorimier, L.-P., Graham, P.S., Moore, P.F., 2006. Canine indolent nodular lymphoma. *Vet. Pathol.* 43, 241–256.
- van Dongen, J.J.M., Langerak, A.W., Brüggemann, M., Evans, P.A.S., Hummel, M., Lavender, F.L., Delabesse, E., Davi, F., Schuurings, E., Garcia-Sanz, R., van Krieken, J.H.J.M., Droese, J., González, D., Bastard, C., White, H.E., Spaargaren, M., González, M., Parreira, A., Smith, J.L., Morgan, G.J., Kneba, M., Macintyre, E.A., 2003. Design and standardization of PCR primers and protocols for detection of clonal immunoglobulin and T-cell receptor gene recombinations in suspect lymphoproliferations: report of the BIOMED-2 Concerted Action BMH4-CT98-3936. *Leukemia* 17, 2257–2317.
- Volpe, J.M., Kepler, T.B., 2008. Large-scale analysis of human heavy chain V(D)J recombination patterns. *Immunome Res.* 4, 3.
- Waugh, E.M., Gallagher, A., Haining, H., Johnston, P.E.J., Marchesi, F., Jarrett, R.F., Morris, J.S., 2016. Optimisation and validation of a PCR for antigen receptor rearrangement (PARR) assay to detect clonality in canine lymphoid malignancies. *Vet. Immunol. Immunopathol.* 182, 115–124.
- Yaari, G., Vander Heiden, J., Uduman, M., Gadala-Maria, D., Gupta, N., Stern, J., O'Connor, K., Hafler, D., Laserson, U., Vigneault, F., Kleinstein, S., 2013. Models of somatic hypermutation targeting and substitution based on synonymous mutations from high-throughput immunoglobulin sequencing data. *Front. Immunol.* 4, 358.