Published in final edited form as:

Immunogenetics. 2016 March ; 68(3): 205–217. doi:10.1007/s00251-015-0894-6.

Evaluation of a DLA-79 allele associated with multiple immune-mediated diseases in dogs

Steven G. Friedenberg,

Department of Clinical Sciences, College of Veterinary Medicine, North Carolina State University, Raleigh, NC 27607. Comparative Medicine Institute, North Carolina State University, Raleigh, NC 27607

Greg Buhrman,

Department of Molecular and Structural Biochemistry, North Carolina State University, Raleigh, NC 27695

Lhoucine Chdid,

Department of Clinical Sciences, College of Veterinary Medicine, North Carolina State University, Raleigh, NC 27607

Natasha J. Olby,

Department of Clinical Sciences, College of Veterinary Medicine, North Carolina State University, Raleigh, NC 27607. Comparative Medicine Institute, North Carolina State University, Raleigh, NC 27607

Thierry Olivry,

Department of Clinical Sciences, College of Veterinary Medicine, North Carolina State University, Raleigh, NC 27607. Comparative Medicine Institute, North Carolina State University, Raleigh, NC 27607

Julien Guillaumin,

Department of Veterinary Clinical Sciences, College of Veterinary Medicine, The Ohio State University, Columbus, OH 43210

Theresa O'Toole,

Department of Clinical Sciences, Cummings School of Veterinary Medicine, Tufts University, North Grafton, MA 01536

Robert Goggs,

Department of Clinical Sciences, College of Veterinary Medicine, Cornell University, Ithaca, NY 14853

Lorna J. Kennedy,

Address correspondence to: S. G. Friedenberg at steven_friedenberg@ncsu.edu, +1 919 515 3277 (telephone), +1 919 513 6452 (fax).

Compliance with Ethical Standards

Ethical approval: All applicable international, national, and/or institutional guidelines for the care and use of animals were followed. All procedures performed in studies involving animals were in accordance with the ethical standards of the institution or practice at which the studies were conducted.

Informed consent: Informed consent was obtained from the owners of all animal participants included in the study.

Conflict of Interest: The authors declare that they have no conflict of interest.

Centre for Integrated Genomic Medical Research, University of Manchester, Manchester, UK M13 9PT

Robert B. Rose, and

Department of Molecular and Structural Biochemistry, North Carolina State University, Raleigh, NC 27695. Comparative Medicine Institute, North Carolina State University, Raleigh, NC 27607

Kathryn M. Meurs

Department of Clinical Sciences, College of Veterinary Medicine, North Carolina State University, Raleigh, NC 27607. Comparative Medicine Institute, North Carolina State University, Raleigh, NC 27607

Abstract

Immune-mediated diseases are common and life-threatening disorders in dogs. Many canine immune-mediated diseases have strong breed predispositions and are believed to be inherited. However, the genetic mutations that cause these diseases are mostly unknown. As many immune-mediated diseases in humans share polymorphisms among a common set of genes, we conducted a candidate gene study of 15 of these genes across four immune-mediated diseases (immune-mediated hemolytic anemia, immune-mediated thrombocytopenia, immune-mediated polyarthritis (IMPA), and atopic dermatitis) in 195 affected and 206 unaffected dogs to assess whether causative or predictive polymorphisms might exist in similar genes in dogs.

We demonstrate a strong association (Fisher's exact $p = 0.0004$ for allelic association, $p = 0.0035$ for genotypic association) between two polymorphic positions (10 bp apart) in exon 2 of one allele in DLA-79, DLA-79*001:02, and multiple immune-mediated diseases. The frequency of this allele was significantly higher in dogs with immune-mediated disease than in control dogs (0.21 vs. 0.12), and ranged from 0.28 in dogs with IMPA to 0.15 in dogs with atopic dermatitis. This allele has two non-synonymous substitutions (compared to the reference allele, DLA-79*001:01) resulting in F33L and N37D amino acid changes. These mutations occur in the peptide binding pocket of the protein, and based upon our computational modeling studies are likely to affect critical interactions with the peptide N-terminus. Further studies are warranted to confirm these findings more broadly, and to determine the specific mechanism by which the identified variants alter canine immune system function.

Keywords

canine; immune-mediated disease; dog leukocyte antigen; major histocompatibility complex class Ib

Introduction

Immune-mediated diseases such as immune-mediated hemolytic anemia (IMHA), immune-mediated thrombocytopenia (ITP), and type I diabetes (T1D) are common and life-threatening conditions in dogs with short-term mortality rates as high as 50% (Carr et al. 2002; O'Marra et al. 2011). Most immune-mediated diseases result from an aberrant immune system response in which normal defense mechanisms attack the body's own cells,

tissues, and organs (Temajo and Howard 2014). Therapy for many of these diseases in dogs involves blanket suppression of the immune system, which can cause complications ranging from polyuria to severe secondary infections (Whitley and Day 2011). Other immune-mediated diseases require lifelong supplementation to replace hormones such as insulin, pancreatic enzymes, or mineralocorticoids, requiring strict compliance and regular monitoring that can be burdensome for owners and patients alike. For all of these reasons, finding ways to decrease the prevalence of immune-mediated disorders is an important challenge in canine medicine.

Many immune-mediated diseases in dogs have strong breed associations and are believed to be inherited (Famula et al. 2003; Chase et al. 2006; Wilbe et al. 2010a; Wilbe et al. 2010c; Salzmann et al. 2011; Wang et al. 2011; Pedersen et al. 2012a; Massey et al. 2014). One approach to decreasing the frequency of these diseases involves breeding programs that consider the role of predisposing genetic polymorphisms. In humans, variants in a common group of genes are believed to contribute to the development of many different immune-mediated diseases (Cho and Gregersen 2011; Voight and Cotsapas 2012; Cotsapas and Hafler 2013; Castiblanco et al. 2013; Parkes et al. 2013; Sorrentino 2013). Analysis of this network of genes has led to the theory that an accumulated burden of genetic variants causes an underlying predisposition to multiple immune-mediated diseases (Cotsapas et al. 2011; Gregersen et al. 2012; Cotsapas and Hafler 2013). Some of these genes, such as the Major Histocompatibility Complex (MHC) class II genes, PTPN22, and CTLA4, have been associated with specific canine immune-mediated diseases (Kennedy et al. 2006b; Kennedy et al. 2006a; Short et al. 2010; Greer et al. 2010; Barber et al. 2011; Massey et al. 2013b; Massey et al. 2013a; Catchpole et al. 2013; Short et al. 2013; Schrauwen et al. 2014), but the importance of broader gene networks across multiple immune-mediated diseases is not well understood in dogs.

In this study, we hypothesized that given the commonality of variants across immune-mediated diseases in humans, a candidate gene approach could allow us to identify polymorphisms that are common to multiple immune-mediated diseases in dogs. We identified 15 candidate genes that have been associated with multiple immune-mediated diseases in humans, and sequenced them initially in 16 dogs of different breeds with IMHA. Based upon our findings suggesting an association between IMHA and a non-synonymous variant in the canine MHC class Ib gene DLA-79, we evaluated this variant in a total of 195 dogs with various immune-mediated diseases and 206 dogs without immune-mediated disease in a total of 24 breeds. Our findings support a strong association of this non-synonymous variant in DLA-79 with multiple immune-mediated diseases in dogs.

Methods

Animal Selection and Phenotyping

We initially selected 16 dogs with IMHA for targeted resequencing of 15 candidate genes (see below), and validated our findings in 12 additional dogs with IMHA (Figure 1). We then selected an additional 173 dogs with various immune-mediated diseases for further genotyping (Figure 1). We included 195 dogs across 24 breeds with 4 immune-mediated diseases in the final analysis. These included: 94 dogs with IMHA, 41 dogs with immune-

mediated thrombocytopenia (ITP), and 16 dogs with immune-mediated polyarthrititis (IMPA) as examples of autoimmune diseases, and 44 dogs with atopic dermatitis (AD) as an example of an environmentally triggered immune-mediated disease. Six affected dogs were eliminated from the final analysis in order to balance within-breed cases and controls. In order to include samples from a particular dog breed in the final analysis, we required at least three within-breed cases and two breed-matched controls across all diseases.

Dogs with IMHA, ITP, and IMPA were selected from inpatient or outpatient populations at veterinary referral hospitals at North Carolina State University, The Ohio State University, Tufts University, and Cornell University; dogs with AD were recruited through referral veterinary dermatologists and breed clubs across the United States. Diagnoses of IMHA, ITP, and IMPA were confirmed by a diplomate of the American College of Veterinary Emergency & Critical Care or the American College of Veterinary Internal Medicine. Diagnosis of AD was confirmed by a diplomate of the American College of Veterinary Dermatology.

Inclusion criteria for dogs with IMHA were as follows: packed cell volume < 30%, spherocytosis or positive results on either a slide agglutination or direct Coombs test (minimum titer 1:64). Dogs with ITP were required to have a platelet count < $20 \times 10^9/L$, normal prothrombin time and partial thromboplastin time, and absence of significant active hemorrhage based upon physical examination or ultrasonographic cavitory fluid assessment. For dogs with IMPA inclusion criteria were: at least 2 joints containing a synovial fluid white blood cell count > $3 \times 10^9/L$ with at least 10% neutrophils and no observed intracellular or extracellular bacteria (as determined by a board-certified veterinary clinical pathologist), as well as radiographic evidence demonstrating non-erosive articular disease in the joints sampled. However, dogs with IMHA, ITP, or IMPA were excluded if they had any of these concurrent conditions: evidence of neoplasia on thoracic radiographs, evidence of neoplasia on abdominal ultrasound, presence of foreign metallic objects on abdominal radiographs, evidence of true infection with erythrocyte-borne or rickettsial disease based upon blood smear, blood culture, convalescent immunofluorescent assay, or polymerase chain reaction (PCR). Inclusion criteria for dogs with IMHA, ITP, and IMPA were similar to those described elsewhere (Clements et al. 2004; Weinkle et al. 2005; Huang et al. 2012). Dogs with AD were diagnosed as described previously (Salzmann et al. 2011). A summary of all 195 animals with immune-mediated disease included in the final analysis is provided in Table 1.

We also included 206 dogs across the same 24 breeds without immune-mediated disease as controls (Figure 1). Dogs without immune-mediated disease included 67 dogs with neurologic disease (primarily intervertebral disc disease, central nervous system neoplasms, or congenital malformations), 56 healthy dogs, 52 dogs with cardiac disease (primarily dilated cardiomyopathy, arrhythmogenic right ventricular cardiomyopathy, myxomatous mitral valve disease, or a congenital structural abnormality), 14 dogs with non-immune-mediated endocrine disorders, 12 dogs with corneal ulcers, and 5 dogs with gastrointestinal disease. These dogs are also described in Table 1.

Candidate Gene Selection

We reviewed meta-analyses of the genes involved in human immune-mediated disease and tabulated the frequency that each gene was associated with a disease (Cotsapas et al. 2011; Cho and Gregersen 2011; Visscher et al. 2012; Castiblanco et al. 2013). We also accessed the protein-coding genes from ImmunoBase version 1.7 (www.immunobase.org) and included these in our analysis (Parkes et al. 2013). After developing a preliminary list of candidate genes based upon the frequency of association with human autoimmune disease, we cross-referenced our findings with genes that have been implicated in canine immune-mediated diseases including IMHA, T1D, and systemic lupus erythematosus (Wilbe et al. 2009; Wilbe et al. 2010c; Pedersen et al. 2012a; Pedersen et al. 2012b; Massey et al. 2013a; Catchpole et al. 2013; Short et al. 2013). We ultimately selected 15 of these genes for targeted resequencing, which represented the maximum number of genes that could be covered within a 500 Kb capture region. The final list of 15 candidate genes selected is provided in Table 2.

DNA Isolation

Whole venous blood was obtained from each animal in ethylenediaminetetraacetic acid (EDTA). Genomic DNA samples were prepared by either phenol-chloroform or column-based extraction. For phenol-chloroform extraction, cells were osmotically lysed in 2× sucrose-Triton (pH 7.6), Tris-NH₄Cl buffer (pH 7.2) and nuclei were pelleted by centrifugation (2,000g for 20 min at 25 °C). Pellets were resuspended in saline-EDTA (pH 8.0) with 20% SDS and proteinase K, and incubated overnight at 56 °C. Samples were subjected to one phenol:chloroform:isoamyl alcohol (25:24:1, pH 8) and one chloroform extraction. Finally, the DNA was precipitated (95–100 % ethanol and 2 M NaCl), and resuspended in 100–200 µL of nuclease-free water. Column-based extraction was performed using a commercially available kit (QIAamp Midi Kit, Qiagen, Valencia, CA).

Next Generation Sequencing

Targeted resequencing was performed in 16 dogs with IMHA (TR genotyping cohort, Online Resource 1). Sequencing libraries were prepared from genomic DNA (approximately 3 µg per dog) using the SureSelectXT Custom MP4 Kit (Agilent Technologies, Santa Clara, CA). The kit contained a custom capture library of 11,117 RNA baits 121 bp in length and designed based on the CanFam3.1 reference sequence. The library was designed to capture both the exons and introns of each gene plus approximately 10% additional bases 5' and 3' to each gene, totaling 484,219 base pairs (bp). Baits were designed to give 2× coverage and exclude repeat-masked regions, resulting in coverage of 88.6% (368,234 bp). Library preparation was performed using the Accel-NCS 2S DNA Library Kit for Illumina (Swift Biosciences, Ann Arbor, MI), followed by probe hybridization and target enrichment per the SureSelectXT protocol. Shearing of genomic DNA was performed using a Covaris E220 ultrasonicator (Woburn, MA); library size and quality were analyzed using the LabChip GX microfluidics capillary electrophoresis system (PerkinElmer, Waltham, MA) according to manufacturer recommendations.

Paired-end sequencing resulting in 150 bp paired-end reads was conducted in a single flowcell on an Illumina MiSeq Desktop Sequencer (San Diego, CA), performed by the High

Throughput Sequencing Facility at the University of North Carolina at Chapel Hill. Sequence reads were trimmed using Trimmomatic 0.32 to a minimum phred-scaled base quality score of 30 at the start and end of each read with a minimum read length of 70 bp, and the resulting reads were aligned to the CanFam3.1 reference sequence using BWA 0.7.10 (Li and Durbin 2009; Bolger et al. 2014). Aligned reads were prepared for analysis using GATK 3.3-0 following the best practices for base quality score recalibration and indel realignment specified by the Broad Institute, Cambridge, MA. HaplotypeCaller was used to call SNP and indel variants (McKenna et al. 2010; DePristo et al. 2011; Van der Auwera et al. 2013). Reads were visualized using GenomeBrowse 2.0.6 software (Golden Helix, Bozeman, MT). On-target reads generated across all 16 samples represented a 675 Mb dataset. Average read depth across the targeted region for all samples ranged from 70× to 122×, and approximately 90% of the region was sequenced with at least 30× coverage.

Variant quality score recalibration was performed using GATK with canine SNPs from dbSNP 139 and SNPs on the Illumina CanineHD BeadChip used as training resources, with a tranche sensitivity cutoff of 99% for passing variants. Variants were removed from the passing callset that were also present in a subset of 25 control dogs with inherited or acquired cardiac or neurologic abnormalities (5 Great Danes, 3 Boxers, 3 Doberman Pinschers, 2 Beagles, 2 Schnauzers, 2 Saint Bernards, 2 Cavalier King Charles Spaniels, and one of each of the following: Irish Setter, Australian Shepherd, Mastiff, Coonhound, Labrador Retriever, Dalmatian) in order to eliminate common variants that were unlikely to be associated with immune-mediated disease. Whole genome sequencing or RNA-seq had been previously performed on these dogs as part of ongoing work in our laboratory. The remaining variants were analyzed using SnpEff 4.0 and Variant Effect Predictor (VEP) 75 to predict impact on protein structure and function (McLaren et al. 2010; Cingolani et al. 2012).

Variant Validation

Because our primary goal was to identify polymorphisms that contribute to disease across breeds, we selected the subset of variants present in at least 14 of 16 dogs with IMHA for further analysis. We also chose those variants that were likely to have an important structural or functional impact (based upon SnpEff and VEP predictions) and were present in at least a third of the dogs evaluated. We genotyped these variants in a separate cohort of 12 dogs with IMHA and 22 dogs without immune-mediated disease (genotyping cohort 1, Online Resource 1). Genotyping was performed using a custom Sequenom array and carried out by Neogene (Lincoln, NE) using standard protocols¹ with multiplexing at 24 SNPs per well. Genotype calls were made using Typer 4.0 software (Agena Bioscience, San Diego, CA).

We tested the Sequenom assay genotypes for association with immune-mediated disease and selected two variants that defined one allele, in exon 2 of DLA-79, for genotyping in a larger group of dogs. Our genotyping assay for DLA-79*001:02 is based on one or two of the seven polymorphic positions in exon 2 and none of the five polymorphic positions in exon 3.

¹<http://agenabio.com/sites/default/files/iPLEX%20Application%20Note.pdf>

The assumption is that this assay will separate DLA-79*001:02 from the other 5 alleles that have been previously identified (Graumann et al. 1998; Venkataraman et al. 2013). Genotyping was performed by either Sanger sequencing or an endpoint genotyping assay in an additional 173 dogs with various immune-mediated diseases and 206 control dogs (validation cohort 2, Online Resource 1). In selecting dogs for this cohort, we attempted to balance within-breed cases and controls from among the samples available at our hospitals.

For Sanger sequencing, amplicons were generated by PCR using primers F-5'-gcactccctgaggttcttga and R-5'-gtgtgcaggctctggagttc with M13 forward and reverse sequences appended and were designed to generate a 209 bp fragment. PCR reaction conditions were as follows: 50 ng genomic DNA, 0.5 μ M each primer, 1 $\times$ AmpliTaq Gold 360 Master Mix (Life Technologies, Carlsbad, CA), 40 cycles, T_m 60 °C. PCR products were cleaned using ExoSAP-IT (Affymetrix, Santa Clara, CA) and sequenced by Sanger sequencing on an Applied Biosystems 3730xl DNA Analyzer (Carlsbad, CA) using standard reaction conditions. Analysis of the resulting sequence trace files was performed using Geneious 8.0.5 (Biomatters Limited, Auckland, NZ).

For endpoint genotyping, we used HEX- or FAM-labeled locked nucleic acid probes to genotype at the first polymorphic position in exon 2. Primer sequences were F-5'-gaggttcttgaacacggctcg and R-5'-ctcttcccagttactcaggcc, and probe sequences for the wild type and alternate allele were 5'-FAM-tTCGAaCGgtt and 5'-HEX-cgtTGGAaCgg, respectively. Both probe sequences had Iowa Black quenchers conjugated to the 3' end. PCR reaction conditions were as follows: 50 ng genomic DNA, 0.5 μ M each primer, 0.2 μ M each probe, 1 $\times$ LightCycler 480 Probes Master Mix (Roche Diagnostics, Indianapolis, IN), 40 cycles, T_m 60 °C. Fluorescence was detected using a Roche LightCycler 480 II real-time PCR system.

A summary of the dogs, genotyping methods, and flow of experiments used in this study is provided in Figure 1.

Statistical Analysis

In an effort to balance cases and controls by breed, we eliminated four dogs from the targeted resequencing cohort (all cases) and 24 dogs from validation cohort 1 (2 cases, 22 controls) from the final analysis as the breeds represented among these dogs did allow us to include at least 3 cases and 2 controls per breed. This left 401 dogs across 24 breeds (Table 1) for association testing in the final analysis. Phenotypes were analyzed in aggregate (all immune-mediated diseases vs. control) and by disease (IMHA, ITP, IMPA, AD, control). Allele and genotype frequencies were compared using $2 \times n$ or $3 \times n$ contingency tables and association with disease was tested using a χ^2 or Fisher's exact test with $p < 0.05$ considered significant. All statistical analyses were performed using JMP Pro 11 (SAS, Cary, NC).

Structural Modeling

Structural models of the reference (DLA-79*001:01) and DLA-79*001:02 proteins were developed in order to predict the impact of the genetic variants being analyzed. Homology models of both the reference and DLA-79*001:02-containing protein were created with the

SWISS-MODEL homology server using chain A, residues 1–276 of Protein Data Bank (PDB) 1Q94 as the template structure (Li and Bouvier 2004; Arnold et al. 2006; Kiefer et al. 2009; Guex et al. 2009; Biasini et al. 2014). 1Q94 was chosen as the template because it had the highest sequence identity (69%) of any deposited protein structure and because structure-based sequence alignment showed that the region of interest surrounding the mutation sites had the highest conservation compared to other available models. The template model is a protein structure of the human MHC class I allele HLA-A*1101. Ramachandran plots were generated by MolProbity as implemented in Phenix (Adams et al. 2010). Models were also validated with the SWISS-MODEL Qmean server, which predicts the local error (in Angstroms) of individual residues based on solvent accessibility, secondary structure, and model geometry (Benkert et al. 2011).

Computational Solvent Mapping

Computational solvent mapping for the reference and DLA-79:001:02 homology models was performed using a protein only model of each structure input to FTMap (ftmap.bu.edu) (Brenke et al. 2009; Kozakov et al. 2011). FTMap probes the protein surface with a defined set of small organic molecules; sites where significant clusters of organic solvents bind have been shown to correlate with ligand binding sites (Buhrman et al. 2011). Organic solvent clusters for each protein were superimposed and examined in Coot (Emsley et al. 2010).

Computational Peptide Docking

DLA-79-binding sequences were predicted using NetMHCpan 2.8 (Nielsen et al. 2007; Hoof et al. 2009) to derive peptide antigens from proteins implicated in the pathogenesis of canine IMHA or ITP (erythrocyte membrane protein band 3 [EMPB3], glycophorin A and C, and platelet integrins) as inputs (Lewis and Meyers 1996; Corato et al. 1997; Day 1999; Callan et al. 2013). NetMHCpan predicts peptide binding sequences and affinities to any MHC class I molecule using artificial neural networks, and is trained using inputs from human, mouse, pig, macaque, chimpanzee, gorilla, and cow MHC molecules and over 150,000 known peptide-protein interactions. Peptides that bind MHC class I molecules are typically 9–11 residues in length and are numbered from P1 at the N-terminus. Peptide sequences with a predicted protein binding affinity greater than 200 nM were selected for further evaluation (Online Resource 2)

Peptide structures were then developed in Coot by mutating the bound peptide from the 1Q94 template to match the NetMHCpan peptide sequences. Peptide-protein complexes for the wild type and mutant proteins were submitted to Rosetta FlexPepDock (flexpepdock.furmanlab.cs.huji.ac.il) (Raveh et al. 2010; London et al. 2011). Rosetta FlexPepDock uses Monte Carlo simulation with energy minimization to optimize peptide-protein interactions and allows for multiple protein side-chain rotamer conformations, a flexible peptide backbone, and multiple peptide side chain conformations. The 10 lowest energy complexes for each peptide were examined directly in Coot and PyMOL.² Visual inspection of these complexes showed that peptides with a hydrophobic amino acid at P1 adopted a wider range of conformations in the reference vs. the DLA-79:001:02 structure.

²The PyMOL Molecular Graphics System, Version 1.7.4 Schrödinger, LLC.

To quantitate conformational flexibility, we calculated the surface area occupied by all peptide conformations at the P1 position for both structures using AREAIMOL as implemented in the CCP4 suite (Lee and Richards 1971; Saff and Kuijlaars 1997).

Results

Genotyping

We found 963 variants passing our filter criteria through targeted resequencing of 16 dogs with IMHA. Of these variants, 875 were intragenic. The effect of these variants were classified by SnpEff as follows: intron 658, upstream 84, downstream 51, 3'-UTR 42, non-synonymous coding 22, synonymous coding 12, splice site region 4, splice site acceptor 1, 5'-UTR 1. STAT1 had by far the largest number of variants at 289, the vast majority of which were intronic (248). A summary of these 963 variants is provided in Online Resource 3.

We genotyped a subset of 22 of these variants (17 in at least 14 of 16 dogs and five in at least a third of dogs and predicted by SnpEff and VEP to have an important structural or functional impact) in a separate cohort of 34 dogs (validation cohort 1) and ruled out association with immune-mediated disease in all but two polymorphic SNPs in DLA-79. The DLA-79 polymorphisms are non-synonymous coding variants at chr18:41,142,888 and 41,142,898 and lie within exon 2 of the gene (Figure 2). The former polymorphism is a C>G transversion and the latter an A>G transition. Coding DNA coordinates for these variants are c.186C>G and c.196A>G, respectively; c.186C>G causes an F33L substitution and c.196A>G causes a N37D substitution.³ Together, these variants define the difference between the reference DLA-79 sequence and the DLA-79*001:02 allele that has been previously reported (Graumann et al. 1998; Venkataraman et al. 2013) and are in perfect linkage disequilibrium across all samples evaluated. We also noted an additional 5 polymorphisms (4 upstream of DLA-79, 1 in intron 1 of DLA-79) in perfect linkage disequilibrium with DLA-79*001:02, but these variants were not validated beyond the original 16 dogs (Figure 2).

Since some genetic variants are common to many immune-mediated diseases and because of the potentially wide-ranging effects of DLA genes, we validated the association of DLA-79*001:02 with multiple immune-mediated diseases in aggregate. Independent validation in a separate cohort of 173 dogs with immune-mediated disease (IMHA, ITP, IMPA, or AD) and 206 control dogs (validation cohort 2) showed strong association using both allelic and genotypic tests (Fisher's exact $p = 0.001$ for allelic association, $p = 0.016$ for genotypic association). Combined results of all 401 dogs also showed strong association with immune-mediated disease (Fisher's exact $p = 0.0004$ for allelic association, $p = 0.0035$ for genotypic association, Figure 3). The DLA-79*001:02 allele frequency is 0.21 in patients with immune-mediated disease vs. 0.12 in patients without immune-mediated disease, and the relative risk of immune-mediated disease in patients with DLA-79*001:02 is 1.4 (95%

³Note that the numbering scheme of the cDNA starts at the initiation codon in the canine sequence, however the numbering scheme of the protein is adapted to align with human MHC class I molecules such that the first amino acid in exon 2 is counted as number one (Graumann et al. 1998).

confidence interval 1.2–1.6). Across the four diseases studied, the DLA-79*001:02 allele frequencies were 0.28 for IMPA, 0.24 for ITP, 0.21 for IMHA, and 0.15 for AD.

We further examined contingency tables using phenotype data separated by each immune-mediated disease rather than in aggregate (Table 3). The observed distribution of diseases in both allelic and genotypic categories was different from the expected distribution using the omnibus χ^2 test ($p = 0.0022$ for allelic association, $p = 0.04$ for genotypic association), consistent with our findings for immune-mediated diseases in aggregate.

Structural and Functional Modeling

Homology models of the DLA-79 reference and DLA-79*001:02 proteins were prepared as described above in order to predict the functional impact of DLA-79*001:02. Ramachandran plots for the template (1Q94), reference (DLA-79*001:01), and DLA-79*001:02 models are shown in Online Resource 4. These plots demonstrate that the peptide binding domain (residues 1–182) of both models are all within the allowable region; the only outliers are outside of the peptide binding domain. Global Qmean4 scores calculated by SWISS-MODEL were 0.591 for the wild type and 0.595 for the mutant model (Benkert et al. 2011). These scores are within two standard deviations of previously solved protein structures of similar size.

Structural alignment of the DLA-79 models with HLA-A*1101 shows a high degree of homology at important peptide-interacting residues (Online Resource 5). One of the four conserved class I tyrosines (Y59) is identical in DLA-79, while the other three tyrosine positions (Y7, Y159, Y171) have conserved aromatic substitutions (F7, F160, H172). (The residue numbering for Y159 and Y171 is offset by one because DLA-79 has an additional residue (S82) near the C-terminus of the long alpha helix 1.) F33, as well the surrounding residues W51, I52/L52, and W167/W168 (HLA-A*1101/DLA-79 numbering) are also conserved in HLA-A*1101.

We tested the assumption that the F33 change could affect peptide binding at the N-terminal peptide residue P1 using solvent mapping to evaluate differences in solvent binding between the wild type and mutant proteins. In both models, multiple solvent clusters were identified at known sites which bind peptide side chains and anchor the peptide to the MHC molecule (Burnett and Geraghty 1995). In the reference model, solvent mapping demonstrated clusters in the following binding sites: P1 (22 solvent molecules), P2 (15 solvent molecules), P3 (27 solvent molecules), and P9 (20 solvent molecules). In the DLA-79*001:02 model, solvent mapping demonstrated these same four clusters, but notably the change of F33 to L33 resulted in a solvent cluster that was no longer in hydrophobic contact with residue 33 (Figure 4).

We then evaluated whether the loss of this interaction could result in greater conformational flexibility for peptides at the P1 position by performing docking studies with two peptides identified by NetMHCpan that have hydrophobic residues at P1 (MMFASVLPALLL and VLMAGTFFF). The ten lowest energy complexes for each peptide were examined. The P1 position of these peptides adopts a much wider range of conformations in the mutant vs. the wild type protein, nearly doubling the amount of permissible surface area occupied by this

residue from 484 Å² to 774 Å² for the methionine in MMFASVLPALLL and from 404 Å² to 609 Å² for the valine in VLMAGTFFF (Figure 5). In all of the peptides, the C-terminal anchor remains intact.

Discussion

In this study, we demonstrate a strong association between the canine DLA-79*001:02 allele and multiple immune-mediated diseases in dogs. This allele results in two changes compared to the reference sequence, DLA-79*001:01. One of these changes (F33L) is located within the peptide binding pocket of the protein. Our computational modeling studies predict that this change may affect the interactions between DLA-79 and the N-terminus of peptides, creating increased flexibility at the P1 position. This flexibility may allow different peptides to bind the mutant vs. the wild type DLA-79 and/or change the conformation in which peptides are presented on the cell surface. Given the complex nature of immune-mediated disease, however, we do not suggest that DLA-79 by itself is the causal candidate for all of the diseases evaluated. Rather, our findings suggest that DLA-79 or its downstream pathways may play a role in the pathogenesis of autoimmune disease, and that this role merits further study.

DLA-79 is a canine class Ib MHC protein whose specific role in the immune system is unknown. The secondary structure of the protein has been characterized, yet differences in the amino acid sequence versus canine and orthologous class Ia proteins suggest that DLA-79 binds a unique set of peptides and may therefore have a distinct role in the immune response (Burnett and Geraghty 1995). The protein has also been shown to be widely expressed in skeletal muscle, which is a unique feature compared to other MHC class I proteins (Burnett and Geraghty 1995). Six alleles have been identified to date, however none have been associated with any specific disease (Burnett and Geraghty 1995; Venkataraman et al. 2013). Among a population of 407 healthy Beagles and a few mixed-breed dogs, 11.5% were heterozygous for DLA-79*001:02 and 0.3% of animals were homozygous for DLA-79*001:02 (Venkataraman et al. 2013). These percentages correspond well with the DLA-79*001:02 allele frequency we found in healthy dogs, but is much lower than the allele frequency we found in dogs with immune-mediated disease. However, it must be noted that the DLA-79*001:02 allele frequency described by Venkataraman et. al. may not be representative of the allele frequency in the larger canine population.

More broadly, the role of class Ib MHC proteins across all species is poorly understood compared to their classical class Ia counterparts. They are oligomorphic rather than polymorphic, can bind both lipids and peptides, are recognized by a wider range of immune receptors (e.g. T-cell receptors, lectin-type receptors), and are often upregulated in response to an inflammatory stimulus rather than being constitutively expressed (Rodgers and Cook 2005; Allen and Hogan). Because of this response to cellular stress, some have suggested a role for MHC class Ib molecules in immune-mediated or autoinflammatory conditions (Rodgers and Cook 2005; Allen and Hogan). Definitive evidence in support of this hypothesis is lacking, but some studies are available. For example, Jiang *et al.* have demonstrated that MHC class Ib molecules activate a key population of CD8⁺ T-regulatory cells (Jiang et al. 1995; Jiang et al. 2005). When lacking functional CD8⁺ T-regulatory cells,

mice develop a systemic lupus erythematosus-like syndrome, suggesting a role for these class Ib molecules in the suppression of autoimmunity (Kim et al. 2011; McPhee et al. 2011). While our study does not offer mechanistic insights, the demonstrated association between DLA-79 and immune-mediated disease supports further investigations regarding the role of class Ib molecules in these conditions. Because the dog is the only domestic species to have a single MHC class Ib paralog, this species may provide an ideal model for such research.

Additionally, it is noteworthy that the association between DLA-79*001:02 and immune-mediated disease is significant across all four disorders we evaluated despite some differences in their underlying pathophysiology. For example, IMHA and ITP are both type II autoimmune diseases in which an antibody response is generated against the cell membranes of erythrocytes or platelets, respectively (McCullough 2003; O'Marra et al. 2011). IMPA is a type III autoimmune disease in which antigen-antibody complexes accumulate in the joint space and lead to ongoing inflammation (Johnson and Mackin 2012). AD is an allergic syndrome caused by an aberrant immune response to an environmental exposure (Olivry et al. 2010; Nuttall et al. 2013). This suggests that DLA-79*001:02 may play a role in functions that are common across these diseases such as pathways which perpetuate an exuberant immune response. This role would be consistent with the previously described function of MHC class Ib genes in downregulating immune system effectors.

One limitation of the work presented here is the lack of precise balance between cases and controls by breed within our study population. We limited our study to 24 dog breeds where we had both case and control samples in an effort to mitigate breed bias, however for some breeds the numbers of samples available were small. Further studies that include a larger number of within-breed samples may be helpful in determining whether certain breeds that develop immune-mediated disease have a higher frequency of the DLA-79*001:02 allele. Nevertheless, canine immune-mediated diseases tend to cluster strongly within breeds and because non-genetic effects (e.g., environmental exposures, viruses) are likely to play a key role in the onset of these diseases, completely controlling for breed may be disadvantageous as it limits insights into disease-causing mutations that may be fixed within breeds (Famula et al. 2003; Salzmann et al. 2011; Piek 2011; Bogdanos et al. 2012; Costenbader et al. 2012; Selmi et al. 2012; Catchpole et al. 2013). Additionally, because of these non-genetic influences on phenotype, truly describing an animal as “unaffected” is challenging as an animal that never develops clinical disease may in fact have a predisposing genetic variant but never have come in contact with a required environmental exposure.

Additionally, we cannot rule out that the association we have described between DLA-79*001:02 and autoimmune disease is in linkage disequilibrium with a more important causal variant. Studies that evaluate association between genetic polymorphisms and autoimmune disease on a genome-wide basis in dogs would be helpful in ruling out this possibility.

In summary, this study demonstrates an association between two tagging SNPs in DLA-79*001:02 and four common immune-mediated diseases in dogs. Based upon our computational modeling, the F33L mutation associated with this allele may alter key

peptide-protein interactions. Changes in these key interactions may alter the downstream immune response and predispose dogs to developing certain immune-mediated diseases, however further studies are required to definitively evaluate the impact of this polymorphism on protein function.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Funding: This study was funded in part by a National Institutes of Health T32 training award provided to Dr. Friedenberg (5T32OD011130-07). Some DNA samples and associated phenotypic data were provided by the Cornell Veterinary Biobank, a resource built with the support of NIH grant R24-GM082910 and the Cornell University College of Veterinary Medicine.

References

- Adams PD, Afonine PV, Bunkóczi G, et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr*. 2010; 66:213–221.10.1107/S0907444909052925 [PubMed: 20124702]
- Allen, RL.; Hogan, L. eLS. John Wiley & Sons Ltd; Chichester: Nov. 2013 Non-Classical MHC Class I Molecules (MHC-Ib). <http://www.els.net>
- Arnold K, Bordoli L, Kopp J, Schwede T. The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. *Bioinformatics*. 2006; 22:195–201.10.1093/bioinformatics/bti770 [PubMed: 16301204]
- Barber RM, Schatzberg SJ, Corneveaux JJ, et al. Identification of risk loci for necrotizing meningoencephalitis in Pug dogs. *J Hered*. 2011; 102(Suppl 1):S40–6.10.1093/jhered/esr048 [PubMed: 21846746]
- Bayer AL, Pugliese A, Malek TR. The IL-2/IL-2R system: from basic science to therapeutic applications to enhance immune regulation. *Immunol Res*. 2013; 57:197–209.10.1007/s12026-013-8452-5 [PubMed: 24214027]
- Benkert P, Biasini M, Schwede T. Toward the estimation of the absolute quality of individual protein structure models. *Bioinformatics*. 2011; 27:343–350.10.1093/bioinformatics/btq662 [PubMed: 21134891]
- Biasini M, Bienert S, Waterhouse A, et al. SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information. *Nucleic Acids Res*. 2014; 42:W252–W258.10.1093/nar/gku340 [PubMed: 24782522]
- Bogdanos DP, Smyk DS, Rigopoulou EI, et al. Twin studies in autoimmune disease: genetics, gender and environment. *Journal of Autoimmunity*. 2012; 38:J156–69.10.1016/j.jaut.2011.11.003 [PubMed: 22177232]
- Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*. 2014; 30:2114–2120.10.1093/bioinformatics/btu170 [PubMed: 24695404]
- Brenke R, Kozakov D, Chuang G-Y, et al. Fragment-based identification of druggable “hot spots” of proteins using Fourier domain correlation techniques. *Bioinformatics*. 2009; 25:621–627.10.1093/bioinformatics/btp036 [PubMed: 19176554]
- Buhrman G, O'Connor C, Zerbe B, et al. Analysis of binding site hot spots on the surface of Ras GTPase. *J Mol Biol*. 2011; 413:773–789.10.1016/j.jmb.2011.09.011 [PubMed: 21945529]
- Burnett RC, Geraghty DE. Structure and expression of a divergent canine class I gene. *J Immunol*. 1995; 155:4278–4285. [PubMed: 7594586]
- Callan MB, Werner P, Mason NJ, et al. Polymorphisms in canine platelet glycoproteins identify potential platelet antigens. *Comp Med*. 2013; 63:348–354. [PubMed: 24209971]

- Carr AP, Panciera DL, Kidd L. Prognostic factors for mortality and thromboembolism in canine immune-mediated hemolytic anemia: a retrospective study of 72 dogs. *J Vet Intern Med.* 2002; 16:504–509. [PubMed: 12322697]
- Castiblanco J, Arcos-Burgos M, Anaya J-M. What is next after the genes for autoimmunity? *BMC Med.* 2013; 11:197–206.10.1186/1741-7015-11-197 [PubMed: 24107170]
- Catchpole B, Adams JP, Holder AL, et al. Genetics of canine diabetes mellitus: are the diabetes susceptibility genes identified in humans involved in breed susceptibility to diabetes mellitus in dogs? *Vet J.* 2013; 195:139–147.10.1016/j.tvjl.2012.11.013 [PubMed: 23265864]
- Chase K, Sargan D, Miller K, et al. Understanding the genetics of autoimmune disease: two loci that regulate late onset Addison's disease in Portuguese Water Dogs. *Int J Immunogenet.* 2006; 33:179–184.10.1111/j.1744-313X.2006.00593.x [PubMed: 16712648]
- Cho JH, Gregersen PK. Genomics and the multifactorial nature of human autoimmune disease. *N Engl J Med.* 2011; 365:1612–1623.10.1056/NEJMra1100030 [PubMed: 22029983]
- Cingolani P, Platts A, Wang LL, et al. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w1118; iso-2; iso-3. *Fly (Austin).* 2012; 6:80–92.10.4161/fly.19695 [PubMed: 22728672]
- Clements DN, Gear RNA, Tattersall J, et al. Type I immune-mediated polyarthritis in dogs: 39 cases (1997–2002). *J Am Vet Med Assoc.* 2004; 224:1323–1327. [PubMed: 15112783]
- Corato A, Shen CR, Mazza G, et al. Proliferative responses of peripheral blood mononuclear cells from normal dogs and dogs with autoimmune haemolytic anaemia to red blood cell antigens. *Vet Immunol Immunopathol.* 1997; 59:191–204. [PubMed: 9477471]
- Costenbader KH, Gay S, Alarcón-Riquelme ME, et al. Autoimmunity Reviews. Autoimmunity Reviews. 2012; 11:604–609.10.1016/j.autrev.2011.10.022 [PubMed: 22041580]
- Cotsapas C, Hafler DA. Immune-mediated disease genetics: the shared basis of pathogenesis. *Trends Immunol.* 2013; 34:22–26.10.1016/j.it.2012.09.001 [PubMed: 23031829]
- Cotsapas C, Voight BF, Rossin E, et al. Pervasive sharing of genetic effects in autoimmune disease. *PLoS Genet.* 2011; 7:e1002254.10.1371/journal.pgen.1002254 [PubMed: 21852963]
- Day MJ. Antigen specificity in canine autoimmune haemolytic anaemia. *Vet Immunol Immunopathol.* 1999; 69:215–224. [PubMed: 10507306]
- de Almeida DE, Holoshitz J. MHC molecules in health and disease: At the cusp of a paradigm shift. *Self Nonself.* 2011; 2:43–48.10.4161/self.2.1.15757 [PubMed: 21776334]
- DePristo MA, Banks E, Poplin R, et al. A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat Genet.* 2011; 43:491–498.10.1038/ng.806 [PubMed: 21478889]
- Eames HL, Corbin AL, Udalova IA. Interferon regulatory factor 5 in human autoimmunity and murine models of autoimmune disease. *Transl Res.* 201510.1016/j.trsl.2015.06.018
- Emsley P, Lohkamp B, Scott WG, Cowtan K. Features and development of Coot. *Acta Crystallogr D Biol Crystallogr.* 2010; 66:486–501.10.1107/S0907444910007493 [PubMed: 20383002]
- Ettinger R, Kuchen S, Lipsky PE. Interleukin 21 as a target of intervention in autoimmune disease. *Ann Rheum Dis.* 2008; 67(Suppl 3):iii83–6.10.1136/ard.2008.098400 [PubMed: 19022821]
- Famula TR, Belanger JM, Oberbauer AM. Heritability and complex segregation analysis of hypoadrenocorticism in the standard poodle. *J Small Anim Pract.* 2003; 44:8–12. [PubMed: 12570345]
- Gee K, Guzzo C, Che Mat NF, et al. The IL-12 family of cytokines in infection, inflammation and autoimmune disorders. *Inflamm Allergy Drug Targets.* 2009; 8:40–52. [PubMed: 19275692]
- Ghoreschi K, Laurence A, O'Shea JJ. Janus kinases in immune cell signaling. *Immunol Rev.* 2009; 228:273–287.10.1111/j.1600-065X.2008.00754.x [PubMed: 19290934]
- Graumann MB, DeRose SA, Ostrander EA, Storb R. Polymorphism analysis of four canine MHC class I genes. *Tissue Antigens.* 1998; 51:374–381. [PubMed: 9583809]
- Greer KA, Wong AK, Liu H, et al. Necrotizing meningoencephalitis of Pug dogs associates with dog leukocyte antigen class II and resembles acute variant forms of multiple sclerosis. *Tissue Antigens.* 2010; 76:110–118.10.1111/j.1399-0039.2010.01484.x [PubMed: 20403140]

- Gregersen PK, Diamond B, Plenge RM. GWAS implicates a role for quantitative immune traits and threshold effects in risk for human autoimmune disorders. *Curr Opin Immunol.* 2012; 24:538–543.10.1016/j.coi.2012.09.003 [PubMed: 23026397]
- Guex N, Peitsch MC, Schwede T. Automated comparative protein structure modeling with SWISS-MODEL and Swiss-PdbViewer: a historical perspective. *Electrophoresis.* 2009; 30(Suppl 1):S162–S173.10.1002/elps.200900140 [PubMed: 19517507]
- Hoof I, Peters B, Sidney J, et al. NetMHCpan, a method for MHC class I binding prediction beyond humans. *Immunogenetics.* 2009; 61:1–13.10.1007/s00251-008-0341-z [PubMed: 19002680]
- Huang AA, Moore GE, Scott-Moncrieff JC. Idiopathic immune-mediated thrombocytopenia and recent vaccination in dogs. *J Vet Intern Med.* 2012; 26:142–148.10.1111/j.1939-1676.2011.00850.x [PubMed: 22151468]
- Iyer SS, Cheng G. Role of interleukin 10 transcriptional regulation in inflammation and autoimmune disease. *Crit Rev Immunol.* 2012; 32:23–63. [PubMed: 22428854]
- Jiang H, Ware R, Stall A, et al. Murine CD8+ T cells that specifically delete autologous CD4+ T cells expressing V beta 8 TCR: a role of the Qa-1 molecule. *Immunity.* 1995; 2:185–194. [PubMed: 7895175]
- Jiang H, Wu Y, Liang B, et al. An affinity/avidity model of peripheral T cell regulation. *J Clin Invest.* 2005; 115:302–312.10.1172/JCI23879 [PubMed: 15668735]
- Johnson KC, Mackin A. Canine immune-mediated polyarthritis: part 1: pathophysiology. *J Am Anim Hosp Assoc.* 2012; 48:12–17.10.5326/JAAHA-MS-5744 [PubMed: 22190602]
- Kennedy LJ, Barnes A, Ollier WER, Day MJ. Association of a common dog leucocyte antigen class II haplotype with canine primary immune-mediated haemolytic anaemia. *Tissue Antigens.* 2006a; 68:502–508.10.1111/j.1399-0039.2006.00715.x [PubMed: 17176441]
- Kennedy LJ, Quarmby S, Happ GM, et al. Association of canine hypothyroidism with a common major histocompatibility complex DLA class II allele. *Tissue Antigens.* 2006b; 68:82–86.10.1111/j.1399-0039.2006.00614.x [PubMed: 16774545]
- Kiefer F, Arnold K, Künzli M, et al. The SWISS-MODEL Repository and associated resources. *Nucleic Acids Res.* 2009; 37:D387–D392.10.1093/nar/gkn750 [PubMed: 18931379]
- Kim H-J, Wang X, Radfar S, et al. CD8+ T regulatory cells express the Ly49 Class I MHC receptor and are defective in autoimmune prone B6-Yaa mice. *Proceedings of the National Academy of Sciences.* 2011; 108:2010–2015.10.1073/pnas.1018974108
- Korman BD, Kastner DL, Gregersen PK, Remmers EF. STAT4: genetics, mechanisms, and implications for autoimmunity. *Curr Allergy Asthma Rep.* 2008; 8:398–403. [PubMed: 18682104]
- Kozakov D, Hall DR, Chuang G-Y, et al. Structural conservation of druggable hot spots in protein-protein interfaces. *Proceedings of the National Academy of Sciences.* 2011; 108:13528–13533.10.1073/pnas.1101835108
- Lee B, Richards FM. The interpretation of protein structures: estimation of static accessibility. *J Mol Biol.* 1971; 55:379–400. [PubMed: 5551392]
- Lewis DC, Meyers KM. Canine idiopathic thrombocytopenic purpura. *J Vet Intern Med.* 1996; 10:207–218. [PubMed: 8819045]
- Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics.* 2009; 25:1754–1760.10.1093/bioinformatics/btp324 [PubMed: 19451168]
- Li L, Bouvier M. Structures of HLA-A*1101 Complexed with Immunodominant Nonamer and Decamer HIV-1 Epitopes Clearly Reveal the Presence of a Middle, Secondary Anchor Residue. *J Immunol.* 2004; 172:6175–6184.10.4049/jimmunol.172.10.6175 [PubMed: 15128805]
- London N, London N, Raveh B, et al. Rosetta FlexPepDock web server--high resolution modeling of peptide-protein interactions. *Nucleic Acids Res.* 2011; 39:W249–W253.10.1093/nar/gkr431 [PubMed: 21622962]
- Massey J, Boag A, Short AD, et al. MHC class II association study in eight breeds of dog with hypoadrenocorticism. *Immunogenetics.* 2013a; 65:291–297.10.1007/s00251-013-0680-2 [PubMed: 23358933]
- Massey J, Rothwell S, Rusbridge C, et al. Association of an MHC class II haplotype with increased risk of polymyositis in Hungarian Vizsla dogs. *PLoS One.* 2013b; 8:e56490.10.1371/journal.pone.0056490 [PubMed: 23457575]

- Massey J, Short AD, Catchpole B, et al. Genetics of canine anal furunculosis in the German shepherd dog. *Immunogenetics*. 2014; 66:311–324.10.1007/s00251-014-0766-5 [PubMed: 24626934]
- McCullough S. Immune-mediated hemolytic anemia: understanding the nemesis. *Vet Clin North Am Small Anim Pract*. 2003; 33:1295–1315.10.1016/S0195-5616(03)00123-2 [PubMed: 14664200]
- McKenna A, Hanna M, Banks E, et al. The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res*. 2010; 20:1297–1303.10.1101/gr.107524.110 [PubMed: 20644199]
- McLaren W, Pritchard B, Rios D, et al. Deriving the consequences of genomic variants with the Ensembl API and SNP Effect Predictor. *Bioinformatics*. 2010; 26:2069–2070.10.1093/bioinformatics/btq330 [PubMed: 20562413]
- McPhee CG, Sproule TJ, Shin D-M, et al. MHC class I family proteins retard systemic lupus erythematosus autoimmunity and B cell lymphomagenesis. *J Immunol*. 2011; 187:4695–4704.10.4049/jimmunol.1101776 [PubMed: 21964024]
- Nielsen M, Lundegaard C, Blicher T, et al. NetMHCpan, a method for quantitative predictions of peptide binding to any HLA-A and -B locus protein of known sequence. *PLoS One*. 2007; 2:e796.10.1371/journal.pone.0000796 [PubMed: 17726526]
- Nuttall T, Uri M, Halliwell R. Canine atopic dermatitis - what have we learned? *Vet Rec*. 2013; 172:201–207.10.1136/vr.f1134 [PubMed: 23436599]
- O'Marra SK, Delaforcade AM, Shaw SP. Treatment and predictors of outcome in dogs with immune-mediated thrombocytopenia. *J Am Vet Med Assoc*. 2011; 238:346–352.10.2460/javma.238.3.346 [PubMed: 21281218]
- Olivry T, DeBoer DJ, Favrot C, et al. Treatment of canine atopic dermatitis: 2010 clinical practice guidelines from the International Task Force on Canine Atopic Dermatitis. *Vet Dermatol*. 2010; 21:233–248.10.1111/j.1365-3164.2010.00889.x [PubMed: 20456716]
- Parkes M, Cortes A, van Heel DA, Brown MA. Genetic insights into common pathways and complex relationships among immune-mediated diseases. *Nat Rev Genet*. 2013; 14:661–673.10.1038/nrg3502 [PubMed: 23917628]
- Pedersen NC, Liu H, Greenfield DL, Echols LG. Multiple autoimmune diseases syndrome in Italian Greyhounds: preliminary studies of genome-wide diversity and possible associations within the dog leukocyte antigen (DLA) complex. *Vet Immunol Immunopathol*. 2012a; 145:264–276.10.1016/j.vetimm.2011.11.015 [PubMed: 22178273]
- Pedersen NC, Liu H, McLaughlin B, Sacks BN. Genetic characterization of healthy and sebaceous adenitis affected Standard Poodles from the United States and the United Kingdom. *Tissue Antigens*. 2012b; 80:46–57.10.1111/j.1399-0039.2012.01876.x [PubMed: 22512808]
- Piek CJ. Canine idiopathic immune-mediated haemolytic anaemia: a review with recommendations for future research. *Vet Q*. 2011; 31:129–141.10.1080/01652176.2011.604979 [PubMed: 22029883]
- Raveh B, London N, Schueler-Furman O. Sub-angstrom modeling of complexes between flexible peptides and globular proteins. *Proteins*. 2010; 78:2029–2040.10.1002/prot.22716 [PubMed: 20455260]
- Rawlings DJ, Dai X, Buckner JH. The role of PTPN22 risk variant in the development of autoimmunity: finding common ground between mouse and human. *J Immunol*. 2015; 194:2977–2984.10.4049/jimmunol.1403034 [PubMed: 25795788]
- Reich NC. STATs get their move on. *JAKSTAT*. 2013; 2:e27080–e27080.10.4161/jkst.27080 [PubMed: 24470978]
- Rodgers JR, Cook RG. MHC class Ib molecules bridge innate and acquired immunity. *Nat Rev Immunol*. 2005; 5:459–471.10.1038/nri1635 [PubMed: 15928678]
- Romo-Tena J, Gómez-Martín D, Alcocer-Varela J. CTLA-4 and autoimmunity: New insights into the dual regulator of tolerance. *Autoimmunity Reviews*. 2013; 12:1171–1176.10.1016/j.autrev.2013.07.002 [PubMed: 23851140]
- Saff EB, Kuijlaars ABJ. Distributing many points on a sphere. *The Mathematical Intelligencer*. 1997; 19:5–11.10.1007/BF03024331
- Salzmann CA, Olivry TJM, Nielsen DM, et al. Genome-wide linkage study of atopic dermatitis in West Highland White Terriers. *BMC Genet*. 2011; 12:37.10.1186/1471-2156-12-37 [PubMed: 21510878]

- Schrauwen I, Barber RM, Schatzberg SJ, et al. Identification of novel genetic risk loci in Maltese dogs with necrotizing meningoencephalitis and evidence of a shared genetic risk across toy dog breeds. *PLoS One*. 2014; 9:e112755–e112755.10.1371/journal.pone.0112755 [PubMed: 25393235]
- Selmi C, Lu Q, Humble MC. *Journal of Autoimmunity*. 2012; 39:249–252.10.1016/j.jaut.2012.07.011 [PubMed: 22980030]
- Sharma R, Fu SM, Ju S-T. IL-2: a two-faced master regulator of autoimmunity. *Journal of Autoimmunity*. 2011; 36:91–97.10.1016/j.jaut.2011.01.001 [PubMed: 21282039]
- Short AD, Boag A, Catchpole B, et al. A candidate gene analysis of canine hypoadrenocorticism in 3 dog breeds. *J Hered*. 2013; 104:807–820.10.1093/jhered/est051 [PubMed: 23997205]
- Short AD, Saleh NM, Catchpole B, et al. CTLA4 promoter polymorphisms are associated with canine diabetes mellitus. *Tissue Antigens*. 2010; 75:242–252.10.1111/j.1399-0039.2009.01434.x [PubMed: 20196832]
- Sorrentino R. Genetics of autoimmunity: An update. *Immunol Lett*. 2013; 158:116–119.10.1016/j.imlet.2013.12.005 [PubMed: 24370643]
- Stanford SM, Bottini N. PTPN22: the archetypal non-HLA autoimmunity gene. *Nat Rev Rheum*. 2014; 10:602–611.10.1038/nrrheum.2014.109
- Temajo NO, Howard N. The mosaic of environment involvement in autoimmunity: The abrogation of viral latency by stress, a non-infectious environmental agent, is an intrinsic prerequisite prelude before viruses can rank as infectious environmental agents that trigger autoimmune diseases. *Autoimmunity Reviews*. 2014; 13:635–640.10.1016/j.autrev.2013.12.003 [PubMed: 24418293]
- Van der Auwera GA, Carneiro MO, Hartl C, et al. From FastQ data to high confidence variant calls: the Genome Analysis Toolkit best practices pipeline. *Curr Protoc Bioinformatics*. 2013; 11:11.10.1–11.10.33.10.1002/0471250953.bi1110s43 [PubMed: 25431634]
- Venkataraman GM, Geraghty D, Fox J, et al. Canine DLA-79 gene: an improved typing method, identification of new alleles and its role in graft rejection and graft-versus-host disease. *Tissue Antigens*. 2013; 81:204–211.10.1111/tan.12094 [PubMed: 23510416]
- Visscher PM, Brown MA, McCarthy MI, Yang J. Five years of GWAS discovery. *Am J Hum Genet*. 2012; 90:7–24.10.1016/j.ajhg.2011.11.029 [PubMed: 22243964]
- Voight BF, Cotsapas C. Human genetics offers an emerging picture of common pathways and mechanisms in autoimmunity. *Curr Opin Immunol*. 2012; 24:552–557.10.1016/j.coi.2012.07.013 [PubMed: 23041452]
- Wang P, Zangerl B, Werner P, et al. Familial cutaneous lupus erythematosus (CLE) in the German shorthaired pointer maps to CFA18, a canine orthologue to human CLE. *Immunogenetics*. 2011; 63:197–207.10.1007/s00251-010-0499-z [PubMed: 21132284]
- Weinkle TK, Center SA, Randolph JF, et al. Evaluation of prognostic factors, survival rates, and treatment protocols for immune-mediated hemolytic anemia in dogs: 151 cases (1993–2002). *J Am Vet Med Assoc*. 2005; 226:1869–1880. [PubMed: 15934255]
- Whitley NT, Day MJ. Immunomodulatory drugs and their application to the management of canine immune-mediated disease. *J Small Anim Pract*. 2011; 52:70–85.10.1111/j.1748-5827.2011.01024.x [PubMed: 21265846]
- Wilbe M, Jokinen P, Hermanrud C, et al. MHC class II polymorphism is associated with a canine SLE-related disease complex. *Immunogenetics*. 2009; 61:557–564.10.1007/s00251-009-0387-6 [PubMed: 19636550]
- Wilbe M, Jokinen P, Truvé K, et al. Genome-wide association mapping identifies multiple loci for a canine SLE-related disease complex. *Nat Genet*. 2010a; 42:250–254.10.1038/ng.525 [PubMed: 20101241]
- Wilbe M, Sundberg K, Hansen IR, et al. Increased genetic risk or protection for canine autoimmune lymphocytic thyroiditis in Giant Schnauzers depends on DLA class II genotype. *Tissue Antigens*. 2010b; 75:712–719.10.1111/j.1399-0039.2010.01449.x [PubMed: 20210920]
- Wilbe M, Ziener ML, Aronsson A, et al. DLA Class II Alleles Are Associated with Risk for Canine Symmetrical Lupoid Onychodystrophy (SLO). *PLoS One*. 2010c; 5:e12332.10.1371/journal.pone.0012332.t004 [PubMed: 20808798]

Ymer SI, Huang D, Penna G, et al. Polymorphisms in the Il12b gene affect structure and expression of IL-12 in NOD and other autoimmune-prone mouse strains. *Genes Immun.* 2002; 3:151–157.10.1038/sj.gene.6363849 [PubMed: 12070779]

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Figure 1.

Flow of experiments performed in this study, including the number of dogs genotyped in each cohort, the genotyping technology employed, and the number of variants screened in each population. Of the dogs genotyped, 12 of the 16 dogs with IMHA in the targeted resequencing cohort, 10 of the 12 dogs with IMHA in validation cohort 1, and zero of the 22 control dogs in validation cohort 1 were included in the final analysis. Animals were removed from the final analysis in order to balance within-breed cases and controls.

Figure 2.

Schematic representation of the region of interest in DLA-79. F33L and N37D indicate the amino acid substitutions resulting from the DLA-79*001:02 allele. The five labeled SNPs were noted to be in complete linkage disequilibrium with DLA-79*001:02 in the targeted resequencing analysis of 16 dogs, however they were not validated in further cohorts.

Figure 3.

Allelic and genotypic association of DLA-79*001:02 with immune-mediated disease in dogs. (A) Dogs with a G allele at the first base of DLA-79*001:02 are significantly more likely to have immune-mediated disease than dogs with a C allele (63% vs. 46%). Note that the first and second bases of DLA-79*001:02 are in complete linkage disequilibrium, therefore only results from the first base are shown. (B) Dogs homozygous (GG/GG) or heterozygous (GG/CA) for DLA-79*001:02 are significantly more likely to have immune-mediated disease than dogs with the reference genotype (CA/CA). For (A) and (B), asterisk represents significant difference vs. wild type. Fisher's exact $p = 0.0004$ for allelic association, $p = 0.0035$ for genotypic association. Abbreviation: immune-mediated disease (IMD).

Figure 4.

Solvent mapping of the peptide binding site of DLA-79 showing (A) the reference model and (B) the DLA-79*001:02 model. The protein backbone is shown in grey, the solvent cluster is shown in magenta, and the F33 or L33 residue is shown in cyan. In the wild type model, the nearest non-hydrogen atom in the solvent cluster is 3.6 Å from the C_Z atom of F33, and the solvent cluster is positioned for hydrophobic interactions with F33. In the mutant model, the nearest non-hydrogen atom in the solvent cluster is 5.4 Å from the C_{D2} atom of L33, making the solvent cluster inaccessible for hydrophobic interactions with L33. Solvent mapping was performed with FTMap.

Figure 5.

Peptide mapping of the peptide binding site of DLA-79 with (A) peptide MMFASVLPALL and (B) VLMAGTFFF. In both (A) and (B) the wild type model is shown on the left and the mutant model is shown on the right. In all figures, the surface depicting the conformational flexibility of the N-terminal residue and the top 10 predicted peptide poses are shown; F33 or L33 is shown in magenta. In (A) the surface area of the N-terminal methionine docked to the wild-type model is $484 \text{ \AA}^2$ compared to a surface area of $774 \text{ \AA}^2$ in the mutant model. In (B) the surface area of the N-terminal valine is $404 \text{ \AA}^2$ compared to a surface area of $609 \text{ \AA}^2$ in the mutant model. Surface construction was performed in PyMOL and surface area calculation was performed in AREAIMOL as implemented in the CCP4 Suite.

Table 1

Dogs, diseases, and phenotypes included in the final analysis.

Breed	Autoimmune Disease				Phenotype		
	AD	IMHA	ITP	IMPA	Case	Control	
Australian Shepherd	0	3	1	1	5	7	
Beagle	0	5	0	0	5	5	
Bichon Frise	0	2	2	0	4	2	
Border Collie	0	4	0	0	4	8	
Boxer	0	0	3	1	4	30	
Bulldog	0	1	1	0	2	8	
Chihuahua	0	3	0	0	3	5	
Cocker Spaniel	0	17	4	0	21	11	
Dachshund	0	1	3	0	4	8	
English Springer Spaniel	0	3	0	0	3	3	
German Shepherd	0	3	0	0	3	9	
Jack Russell Terrier	0	7	2	1	10	3	
Labrador Retriever	0	5	3	3	11	10	
Maltese	0	2	8	0	10	3	
Mastiff	0	1	3	0	4	2	
Miniature Poodle	0	1	1	0	2	7	
Miniature Schnauzer	0	5	1	1	7	3	
Mixed	0	13	4	4	21	9	
Pug	0	2	1	1	4	3	
Shetland Sheepdog	0	3	0	0	3	9	
Shih Tzu	0	7	0	2	9	6	
Standard Poodle	0	6	1	1	8	22	
Toy Poodle	0	0	2	1	3	3	
West Highland White Terrier	44	0	1	0	45	30	
Total	44	94	41	16	195	206	

Table 2

Candidate genes selected for targeted resequencing. Capture coordinates and size based upon CanFam 3.1 reference sequence. Total covered region is just over 484 kB. Probe coverage is measured in percent of bases. References are provided supporting the relevance of each gene to autoimmune disease in either humans or dogs.

Gene	Capture coordinates	Capture size (bp)	Probe coverage	References
CTLA4	chr37:12647015-12654496	7,481	100.00%	(Romo-Tena et al. 2013)
DLA-DQA1	chr12:2219986-2233024	13,038	81.70%	(Kennedy et al. 2006a; Wilbe et al. 2010b; de Almeida and Holoshitz 2011; Massey et al. 2013b; Massey et al. 2013a)
DLA-DQB1	chr12:2239851-2252008	12,157	68.65%	
DLA-DRB1	chr12:2150084-2165878	15,794	94.54%	
DLA-79	chr18:41142058-41146474	4,416	100.00%	(Rodgers and Cook 2005)
IL-10	chr7:5932263-5937569	5,306	100.00%	(Iyer and Cheng 2012)
IL-12B	chr4:51182567-51194586	12,019	99.71%	(Ymer et al. 2002; Gee et al. 2009)
IL-2	chr19:17750219-17756955	6,736	94.63%	(Sharma et al. 2011; Bayer et al. 2013)
IL-21	chr19:17640888-17652234	11,346	99.42%	(Eitinger et al. 2008)
IL-2RA	chr2:29888276-29954389	66,113	86.70%	(Bayer et al. 2013)
IRF5	chr14:7722018-7736335	14,317	94.48%	(Eames et al. 2015)
PTPN22	chr17:51616390-51689806	73,416	79.35%	(Stanford and Botini 2014; Rawlings et al. 2015)
STAT 1; STAT4	chr37:1492157-1702921	210,764	91.34%	(Korman et al. 2008; Reich 2013)
TYK2	chr20:50744507-50775809	31,302	84.80%	(Choreschi et al. 2009)

Table 3

Allele and genotype frequencies of DLA-79*001:02 for each of the four immune-mediated diseases evaluated in the study. Allele frequencies of two tagging SNPs for DLA-79*001:02 are shown on the left and genotype frequencies are shown on the right. G-G are the genotypes of the two tagging SNPs in DLA-79*001:02 and C-A are the genotypes of the reference allele. Both allele and genotype are significant predictors of disease ($\chi^2 p = 0.0022$ for allele frequencies, $p = 0.04$ for genotype frequencies). Abbreviations: immune-mediated hemolytic anemia (IMHA), immune-mediated thrombocytopenia (ITP), immune-mediated polyarthritis (IMPA), atopic dermatitis (AD).

allele frequencies for the two SNPs		two-SNP genotype frequencies				
Disease	G-G	C-A	GG/GG	GG/CA	CA/CA	Dogs
IMHA	0.31	0.22	0.33	0.28	0.21	94
ITP	0.15	0.09	0.14	0.17	0.09	41
IMPA	0.07	0.03	0.08	0.05	0.03	16
AD	0.10	0.11	0.06	0.15	0.11	44
Control	0.37	0.54	0.39	0.35	0.56	206
Total	132	670	36	60	305	401