# The genetics and expression of canine antigen receptors analysed with a novel RNA-seq strategy.

Jolyon N E Martin[1], Hannes Ponstingl[1], David Sargan[2] and Allan Bradley[1]
1 Wellcome Trust Sanger Institute, United Kingdom Univ. of Cambridge, United Kingdom
2 Univ. of Cambridge, United Kingdom

**INTRODUCTION** Dogs are an excellent model for human disease, e.g. the treatment of canine lymphoma is often predictive of the human response to that treatment. However, an incomplete picture of their antigen-receptor (AR) gene loci and a limited set of available experimental tools has restricted their use. This work advances the annotation of the canine AR loci, develops methods to interrogate their expression patterns, and looks into breed-specific features of the loci.
**METHODS** A bioinformatic approach alongside unbiased RNA-seq was used to complete the annotation of the canine AR genes, and these sequences were used to query 107 whole genome sequences from 19 breeds. A 5'RACE strategy was developed to selectively amplify ARs for next generation sequencing, and this was applied to RNA isolated from canine peripheral blood samples.
**RESULTS** >5,500 novel alleles were identified across the ~550 gene segments (of which 326 were newly annotated) of the AR loci, yielding insight into AR evolution as well as confirming the greater conservation between dog and human than mouse with either. The 5'RACE strategy was validated with a pilot study that cast light onto the baseline expression and mutation patterns of healthy dogs, finding they closely mirrored their human counterparts.
**CONCLUSION** This work brings our understanding of the genetics and expression of ARs in dogs to the same high standard as mice and men. The large number of genome sequences serves as a reference for future studies, and has allowed statistically powerful conclusions to be drawn on the pressures that have shaped these loci. Finally, it confirms the dog as a superior immunological model, particularly in the context of the high conservation seen in both sequence and expression features.