Limited diversity associated with duplicated class II MHC-*DRB* genes in the red squirrel population in the United Kingdom compared with continental Europe

Keith T. Ballingall[1], Angeline McIntyre[1,2,*], Zhenzhen Lin[1,3], Naomi Timmerman[1,6], Erik Matthysen[6], Peter W.W. Lurz[3], Lynsey Melville[1], Amy Wallace[1], Anna L. Meredith[3], Claudia Romeo[4], Lucas A. Wauters[5], Anthony W. Sainsbury[2] and Colin J. McInnes[1]

[1] Moredun Research Institute, Midlothian, Scotland, UK,

[2] Institute of Zoology, Zoological Society of London, UK,

[3] The Royal (Dick) School of Veterinary Studies, The University of Edinburgh, UK,

[4] Department of Veterinary Sciences and Public Health, University of Milan, Italy,

[5]Department of Theoretical and Applied Sciences, University of Insubria, Varese, Italy,

[6] Evolutionary Ecology Group, University of Antwerp, Belgium.

Address correspondence to: Keith T. Ballingall, Moredun Research Institute, Pentlands Science Park, Bush Loan, Penicuik, Midlothian, EH26 OPZ, Scotland, UK.

E-mail keith.ballingall@moredun.ac.uk, Tel.:+44 (0) 131 445 5111, FAX: +44 (0) 131 445 6235

* Current address, Department of Ecosystem and Public Health, University of Calgary, Canada

Acknowledgements

The authors acknowledge all those who contributed genetic material to this study. KB and CM are supported by the Scottish Government Rural and Environment Science and Analytical Services (RESAS) Division.

Abstract

The red squirrel (*Sciurus vulgaris*) population in the United Kingdom has declined over the last century and is now on the UK endangered species list. This is the result of competition from the eastern grey squirrel (*S. carolinensis*) which was introduced in the 19th century. However, recent evidence suggests that the rate of population decline is enhanced by squirrelpox disease, caused by a viral infection carried asymptomatically by grey squirrels but to which red squirrels are highly susceptible. Population genetic diversity provides some resilience to rapidly evolving or exotic pathogens. There is currently no data on genetic diversity of extant UK squirrel populations with respect to genes involved in disease resistance. Diversity is highest at loci involved in the immune response including genes clustered within the major histocompatibility complex (MHC). Using the class II *DRB* locus as a marker for diversity across the MHC region we genotyped 110 red squirrels from locations in the UK and continental Europe. Twenty four *Scvu-DRB* alleles at two functional loci; *Scvu-DRB1* and *Scvu-DRB2*, were identified. High levels of diversity were identified at both loci in the continental populations. In contrast, no diversity was observed at the *Scvu-DRB2* locus in the mainland UK population while a high level of homozygosity was observed at the *Scvu-DRB1* locus. The red squirrel population in the UK appears to lack the extensive MHC diversity associated with continental populations, a feature which may have contributed to their rapid decline.

Keywords: Red squirrel, MHC *DRB*, Population, UK, diversity, Squirrelpox virus, disease

Introduction

The Eurasian red squirrel (*Sciurus vulgaris*) is currently on the endangered species list in the United Kingdom (UK) although not in the rest of its pan Eurasian range. Within the UK the majority of the population is restricted to Scotland with fragmented populations remaining in England and Wales, while the distribution of the eastern grey squirrel (*S. carolinensis*) has expanded to match that vacated by the red squirrels. As recently described in detail by Signorile et al (2016) the North American eastern grey squirrel was introduced and subsequently translocated across the UK and Ireland on at least 30 occasions from the 1870's until the 1920's (Middleton 1930; Shorten, 1954, Barratt et al. 1999). Grey squirrel numbers have increased ever since and have been estimated at around 2.5 million while red squirrel numbers have declined to approximately 120,000 (Harris et al. 1995). In continental Europe the grey squirrel has also been introduced to Northern Italy on at least three occasions between 1948 and the 1990s, followed by numerous translocations and undocumented releases (Martinoli et al. 2010; Bertolino et al. 2008, 2014). However, no evidence of the SQPV has been reported which may partially explain the slower rate of decline in Northern Italian red squirrels compared with those in the UK. The principal factors that underlie the rapid decline of the red squirrel and replacement by grey squirrels in the UK include competition from the grey squirrel (Gurnell et al. 2004; Kenward and Holm 1993; Tompkins et al. 2002; Wauters and Gurnell 1999) and disease caused by infection with the squirrelpox virus (SQPV) (Thomas et al. 2003; La-Rose et al. 2010). SQPV, a member of the *Poxviridae* family (Thomas et al. 2003; McInnes et al. 2006; Darby et al. 2014) is thought to be transmitted by asymptomatic grey squirrels (Sainsbury et al. 2000; Tompkins et al. 2002) to highly susceptible red squirrels. It has been estimated that on average 61% of grey squirrels in the UK are seropositive for SQPV (McInnes et al. 2006), although this fluctuates between 100% and 0% depending on the density of squirrels supported by different types of woodland.

Infection of red squirrels with SQPV generally results in death within 2-3 weeks of infection which is likely to be a result of starvation and dehydration due to the inability to forage for food and water and the combined effect of secondary, mainly bacterial, adventitious infections. In areas where red and grey squirrels coexist the decline of red squirrels is up to twenty five times faster if the grey squirrels are carrying SQPV than if they are free from the virus (Rushton et al. 2006). As a consequence, the red squirrel is unlikely to survive in the UK unless populations are maintained in favourable conifer habitats that reduce competition and immigration by grey squirrels (Gurnell et al. 2002).

In response to the threat posed by SQPV, a number of red squirrel strongholds have been established in the UK which combine measures to control exposure to the grey squirrels with habitat improvement. However, small isolated populations often suffer from reductions in genetic diversity due to inbreeding depression and the effect of genetic drift (Keller and Waller 2002; Charlesworth and Willis 2009). This reduces the ability of such populations to respond to rapidly evolving endemic and exotic pathogens compared with larger more genetically diverse populations (Frankham and Ralls 1998; Bernatchez and Landry 2003). Maintaining existing red squirrel diversity while developing strategies that allow diversity to increase within isolated populations will be important for the long term sustainability of the red squirrel strongholds. Historical evidence indicates that red squirrels may have experienced severe population declines and bottlenecks and there is a complete lack of knowledge on genetic diversity of extant UK populations especially with respect of genes involved in disease resistance. Previous analyses of genetic diversity in the red squirrel have targeted nuclear, neutral microsatellite and mitochondrial markers providing important information on the population structure (Barrett et al. 1999; Grill et al. 2009, Hale et al. 2001) but limited information on the role of diversity in the response to SQPV infection.

The highest levels of genetic diversity within mammalian populations are located within genes involved in the immune response including those clustered together within the major histocompatibility complex (MHC), (Horton et al. 2004; Robinson et al. 2013). As a consequence, MHC loci are frequently used as a source of genetic markers in studies of population diversity and population health (Sommer 2005; Osborne et al. 2015). The MHC is divided into three major clusters of closely linked genes, class I, II and III. MHC class I and II genes encode proteins responsible for the presentation of small fragments of pathogen proteins for recognition by antigen specific receptors on CD8 or $CD4^{+ve}$ T cells respectively (Bjorkman 1987; Germain and Margulies 1993). The specificity of the immune response is influenced by the range of pathogen peptides presented by MHC molecules. The majority of MHC diversity associated with the class II MHC loci is found in the second exon which determines part of the peptide binding groove. As a consequence, allelic diversity influences the range of peptides recognised by the immune system (Hughes and Yeager 1988; Hughes and Nei 1989) and many associations with susceptibility to autoimmune and infectious disease have been described (reviewed by Trowsdale 2011).

Earlier analyses of fragmented populations of European ground squirrel (*Spermophilus citellus,* Ricanova et al. 2011) and spotted suslik, *(Spermophilus suslicus,* Biedrzycka and Radwan 2008) described high

levels of allelic diversity at the class II MHC-*DRB* locus. Therefore, this study aims to characterise the *DRB* locus in red squirrels which will allow a comparison of diversity in fragmented UK red squirrel populations with populations from continental Europe.

Materials and methods

Red squirrel samples

Genomic DNA was prepared from 42 tissue samples obtained from red squirrels selected from archived material held at the Zoological Society of London (ZSL). These animals were found dead and submitted to the ZSL between 1996 and 2006 and represent three locations within mainland UK; Central Scotland, North West England, North East England and two island populations, the Isle of Wight and Jersey in the Channel Islands. Twelve road kill samples were obtained from the stronghold population on the Isle of Arran located of the West coast of Scotland, six samples from South West Scotland, six samples from North Central Scotland, thirteen from Northern Scotland and three from Northern Ireland. Eighteen samples of continental European red squirrels were obtained from Belgium and Northern Italy. The location and number of animals sampled at each location is detailed in Figure 1. For comparative purposes, DNA was also prepared from an eastern grey squirrel from the South West of Scotland.

Preparation of DNA

Genomic DNA was extracted from muscle or spleen samples using the DNeasy blood and tissue kit (Qiagen) following the manufacturer's instructions. The quantity and quality of DNA was estimated using a nanodrop spectrophotometer.

Preparation of RNA

Pseudogenes and gene fragments are common features of MHC regions in other mammalian species (Kumanovics et al. 2003). To provide evidence that the *Scvu-DRB* loci are functional, cDNA was prepared from mRNA isolated from the spleen of a red squirrel (sample 15, supplementary Table 1) following euthanasia of a suspected case of squirrelpox in South West Scotland. The spleen was removed, suspended in RNAlater™ and archived at -20^{o}C. Total RNA was prepared from 20 mg of spleen tissue using the Precellis Ribolyser Tissue RNA kit. Genomic DNA was also prepared from the same sample.

Targeting the red squirrel DRB loci

PCR primers Scvu351F and Scvu338R which amplify a 243 bp fragment of the second exon of the red squirrel *DRB* locus were designed using a *DRB* cDNA sequence from the tassel-eared squirrel (accession number M97616) as the template. Both primers are located within the second exon. The primer sequences are listed in

Table 1. Each PCR reaction was carried out in a final volume of 50 µl containing 200 nM of each primer, 1U *Taq* polymerase (Promega, Paisley, UK) and 50 ng of DNA template. Amplification reactions were performed under the following cycling conditions; 94ºC for 4 minutes followed by 30 cycles of 94ºC for 30 s, 60ºC for 30 s and72ºC for 30 sec. A final cycle of 72ºC for 5 min was added to complete the reaction.

Analysis of PCR products

The products of each PCR reaction were separated on a 1% agarose gel, stained with gel red and visualised under a UV transilluminator. PCR products were purified using the SV Gel and PCR Clean-Up System (Promega), quantified and sequenced in both directions using primers Scvu351F and Scvu338R. The forward and reverse sequences were aligned using the SeqManII™ program of the DNASTAR package and polymorphic positions identified. As the primers amplify the products of two polymorphic *DRB1* loci in order to define the allelic diversity at each locus the PCR fragments are cloned.

Cloning and sequence analysis

Scvu-DRB alleles were cloned into the pGEM-T-easy vector (Promega) and individual clones identified by colony PCR. Digestion of the colony PCR product with the restriction enzyme *Rsa* I followed by resolution of the fragments on an 8% polyacrylamide gel allowed the selection of clones with identical restriction patterns for sequencing. Depending on the complexity of the direct sequence analysis, up to 12 clones were sequenced in both directions. Sequencing or Taq induced errors were eliminated through comparison with the direct sequence of the PCR product. The majority of alleles including those that differ by single nucleotide substitutions were identified multiple times from different DNA samples and in some cases from cDNA as well as genomic DNA. Those alleles identified from single samples were cloned and sequenced independently from two different PCR reactions to eliminate possible artefacts associated with amplification and cloning.

Red squirrel Class II DRB nomenclature

We followed the accepted convention of MHC allelic nomenclature proposed by Klein et al. (1990) - which uses the first two letters of the genus and species (*Scvu*) followed by the locus (*Scvu-DRB1*) and then an allele designation (*Scvu-DRB1a, 1b, 1c,* based on the order of their identification). *DRB* alleles were assigned to either the *DRB1* or *DRB2* locus depending on sequence similarity and phylogenetic clustering. The allelic nomenclature shown in Table 2 is used throughout.

Analysis of Scvu-DRB gene transcription

First strand cDNA was prepared using the ImProm-II RT system (Promega) in a 40 µl reaction using 200 ng of Total RNA. Using the full length *DRB* transcript from the tassel-eared squirrel (*Sciurus aberti*) as a template, primers Scvu363 and Scvu364 (listed in Table 1) were designed within exons 1 and exon 3 and tested for their capacity to amplify the *Scvu-DRB* transcripts. Reverse transcription-PCR was carried out in 50 µl reactions using each combination of forward and reverse primer, 3 µl of cDNA template and 200 nM of each primer in Go*Taq* polymerase master mix (Promega, Paisley, UK). Amplification reactions were performed under the following conditions; 94ºC for 4 minutes followed by 30 cycles of 94ºC for 1 min, 55ºC for 1min and 72ºC for 1 min. Fragments were visualised on 1% agarose gels and those of the expected size were gel purified and cloned into the pGEM-T-easy vector as detailed above.

Sequence analysis

Scvu-DRB gene sequences were assembled from each bi-directional sequence using the SeqManII program. All polymorphic sites were inspected manually. All sequences have been deposited in the European Nucleotide Archive and assigned accession numbers listed in Table 2. Multiple alignments of the nucleic acid and predicted amino acid sequences were produced using Clustal Omega available on the EMBL-EBI website (http://www.ebi.ac.uk/Tools/msa/clustalo/). Multiple alignments of the *Scvu-DRB* sequences generated here and other published sequences were used to estimate maximum likelihood trees using PhyML-aLRT (Version 2.4.5) (Anisimova and Gascuel 2006) launched from TOPALi v2.5 (Milne et al. 2008). Prior to phylogenetic tree estimation, the model selection feature in the TOPALi v2 package which produces improved estimates of likelihood values was used to select the nucleotide substitution model JC+G (Jukes and Cantor 1969). To test for positive selection we compared the average number of synonymous substitutions (dS) with the average number of non-synonymous substitutions (dN) for codons predicted to determine the antigen-binding sites (ABS), the remaining sites (non-ABS) and all sites. We used the modified Nei and Gojobori method with Jukes–Cantor correction as the substitution models. The codons predicted to determine amino acids associated with the APS were selected according to Reche and Reinherz (2003) and are shown in Figure 2. The average dN and dS and their variances estimated using 10000 bootstrap replicates were used to test the null hypothesis that H_0, dN=dS (test for neutrality) using a Z test. This analysis was carried out in MEGA version 6 (Tamura et al 2013). Rejection of the null hypothesis in favour of the alternative hypothesis where dN > dS where the probability values *P* are less than 0.05 is considered evidence for positive selection.

189

Results

Identification of two Scvu-DRB loci

A 243 bp fragment of the second exon of the *Scvu-DRB* locus was initially amplified from 10 red squirrel DNA samples from UK population 1 (Figure 1). Sequence analysis of the PCR fragments identified 29 identical polymorphic positions in each of these 10 animals. The presence of two distinct sequences which were identical in all ten animals was confirmed through analysis of individual clones obtained from four of these animals. The two sequences did not appear to segregate as expected for alleles at a single locus as no animal homozygous for either allele was identified. Therefore, rather than alleles at a single locus, we concluded that they are likely to represent two independent *DRB* loci, inherited together within a single haplotype. All 10 animals genotyped appeared homozygous for this one haplotype. The presence of two independent and polymorphic *DRB* loci was confirmed through identification of alleles at each locus in animals from populations 11 and 12 from Belgium and Italy respectively. The sequences identified from population 1 were used as reference sequences for each of these loci and termed *Scvu-DRB1a* and *Scvu-DRB2a* (Supplementary Figure 1).

Are both Scvu-DRB loci transcribed?

Using primers Scvu363 and Scvu364 located in exons 1 and 3, three correctly spliced transcripts representing two alleles at locus 1, (*Scvu-DRB1a* and *Scvu-DRB1b*) and a single allele at locus 2, (*Scvu-DRB2a*) were identified in sample 15 from population 2 (Supplementary Table 1)`, confirming that both loci are transcribed and therefore likely to be functional. No polymorphic sites were identified in the genomic DNA primer binding sites within exon 2 suggesting that the genotyping primers are likely to amplify the majority of *DRB* allelic diversity in red squirrels. The genotyping of a DNA sample from the same squirrel produced an identical result to the cDNA analysis.

Scvu-DRB sequence analysis

Sequence analysis of the PCR products from the remaining 90 samples identified a range of nucleotide substitutions not present in population 1. Where novel and multiple substitutions were identified, individual alleles were resolved through cloning. A total of 19 *Scvu-DRB1* alleles and 5 *Scvu-DRB2* alleles were identified. The alleles associated with each squirrel sample are shown in supplementary Table 1. The sequences have been assigned ENA database accession numbers LN832043 to LN832063 as shown in Table 2. The nucleotide sequences of the 24 *Scvu-DRB* variants are shown in supplementary Figure 1 while the predicted amino acid

sequences are shown in Figure 2. The *Scvu-DRB1* locus is the more polymorphic of the two with 19 of the 24 alleles. Twenty seven polymorphic nucleotide positions corresponding to 15 dN substitutions were identified within the second exon of the *Scvu-DRB1* locus compared with 16 polymorphic positions corresponding to eight dN substitutions within the second exon of the *Scvu-DRB2* locus. Allelic diversity at both *DRB1* and *DRB2* loci was generally associated with small numbers of nucleotide substitutions with many alleles differing at only one or two positions. Alleles *DRB1a* and *1b, DRB1m* and *1n* and *DRB2b* and *2c* differ at single dS substitutions. Alleles *DRB1e* and *1h* show the highest level of diversity with 90% identity while the most diverse *DRB2* alleles, *DRB2a* and *2e,* show 93% identity in pair-wise comparisons. Inter-locus diversity is greater with 85% identity between *DRB1a* and *DRB2a*.

Substantial allelic diversity within and between *DRB1* and *DRB2* loci is associated with positions predicted to directly interact with peptides bound within the peptide binding domain (Figure 2). Sixteen of the 18 amino acid positions estimated by Reche and Reinherz (2003) to directly interact with peptides bound within the class II MHC peptide binding domain are shown to be variable or adjacent to a variable amino acid in red squirrels (Figure 2). As positive selection is thought to drive and maintain diversity at MHC loci we tested the hypothesis that dN > dS at codons predicted to determine the antigen-binding sites (ABS), the remaining sites (non-ABS) and all sites. This hypothesis was rejected in the analysis of all sites (dN-dS = 0.96, p = 0.17) and the non ABS (dN-dS = -0.73, p = 1.0) and only at ABS sites was the hypothesis supported (dN-dS = 2.663, p=0.004).

Phylogenetic analysis

The relationship between *Scvu-DRB1* and *B2* sequences was further explored by phylogenetic analysis using the nucleic acid alignment shown in supplementary Figure 1. The tree topology (Figure 3) generally supports the two locus hypothesis as the two major clusters are formed by the *DRB1* and *DRB2* allelic lineages, the only exception being *Scvu-DRB1l* which clusters independently of the other *DRB1* alleles despite sharing many of the nucleotide and amino acid motifs characteristic of the *DRB1* locus. This may be due to a recombination event between *DRB1* and *DRB2* loci. The *S. aberti (Scab-DRB)* and the *S. carolinensis (Scca-DRB)* sequences all cluster with the *Scvu-DRB1* loci.

The distribution of Scvu-DRB1 and *Scvu-DRB2 allelic diversity in UK and continental European red squirrels*

247 The distribution and frequency of the 19 *Scvu-DRB1* alleles and 5 *Scvu-DRB2* alleles in UK and continental
248 European red squirrels is shown in Figure 4 and Table 3 respectively. Twelve *Scvu-DRB1* and 4 *Scvu-DRB2*
249 alleles were identified in the 18 animals from continental populations 11 and 12, while only 6 *Scvu-DRB1* and a
250 single *Scvu-DRB2* allele were found in 55 samples obtained from six UK mainland populations. Both *DRB* loci
251 were homozygous in 78% of animals from the mainland UK compared with 16% of the continental red
252 squirrels.

253 The highest level of allelic diversity with 12 *DRB* alleles associated with nine distinct haplotypes was
254 identified in the population from northern Italy, while the population with least diversity was population 1 from
255 central Scotland with only a single haplotype. These data indicate that the extensive allelic and haplotype
256 diversity associated with continental European red squirrels is not present in UK populations analysed.

257 With the exception of population 10 from the Isle of Arran, the *Scvu-DRB1a*/*Scvu-DRB2a* haplotype
258 dominates the UK population. This haplotype was not identified in the continental populations or in the small
259 number of samples from the Channel Islands. Given the proximity of the Channel Islands to the French coast, it
260 is not surprising that they share alleles with continental populations. However, population 10 shares allelic
261 diversity with samples from Belgium rather than with other UK populations. This suggests that this population
262 may have a more recent continental European origin.

263

Discussion

In response to selection by rapidly evolving pathogens, genes associated with protective immunity are often highly diverse (Barreiro and Quintana-Murci 2010). Such diversity increases the probability of population survival in the face of novel infections whereas populations with limited diversity are less secure. A major source of immunological diversity is within the MHC where substantial allelic diversity is thought to be maintained by a form of balancing selection (heterozygous advantage and/or frequency dependent selection) arising from the requirement to respond to rapidly evolving or novel pathogens (Hughes and Yeager 1998; Meyer and Thomson 2001). High levels of allelic diversity at MHC loci are often associated with large populations with high levels of genetic exchange whereas low levels are often associated with smaller, more isolated populations (reviewed in Sommer et al. 2005; Radwan et al. 2010).

Comparison of class II MHC DRB diversity in UK and continental squirrel populations

While comparing diversity at the class II MHC *Scvu-DRB* locus in UK red squirrels with populations from continental Europe, we identified a duplication of the *Scvu-DRB* locus, described *Scvu-DRB1* and *Scvu-DRB2* transcripts and sequenced families of alleles at each locus. We provide evidence of positive selection at sites associated with the binding of peptide antigens in agreement with orthologous loci in other species (Babik et al. 2005, Cizkova et al. 2011). Limited *Scvu-DRB1*, *Scvu-DRB2* allelic and haplotype diversity was identified in geographically distinct populations of red squirrel in the UK. A single *DRB* haplotype (*DRB1a/DRB2a*) appears to dominate the UK population with levels of homozygosity ranging between 68% and 100% depending on the population analysed. In contrast, substantial allelic diversity was identified in samples from continental Europe where Belgian and Italian populations provided 12 *Scvu-DRB1* and 4 *Scvu-DRB2* alleles from 18 animals compared with only 6 *Scvu-DRB1* and a single *Scvu-DRB2* allele in 55 samples from 6 populations from the UK mainland. While it is likely that some alleles present at lower frequencies will not have been recorded in both continental European and UK squirrels, it is clear that the extensive MHC diversity in continental European squirrels is not present in UK populations.

Origin of the Scvu-DRB1a/Scvu-DRB2a haplotype

The origin of the *Scvu-DRB1a/Scvu-DRB2a* haplotype which dominates the UK red squirrel populations is unclear. This haplotype may be a remnant from the original population that colonised the British Isles following the end of the last ice age between 7 and 10 thousand years ago when the UK remained connected with Western

Europe. The failure to identify this haplotype in the continental European or Channel Island populations supports this observation; however our analysis is limited to 18 animals from Italy and Belgium and is clearly not representative of the continental population as a whole. There is evidence that the original red squirrel population that colonised the British Isles was almost driven to extinction in the 18th century (summarised in Barratt et al. 1999). The lack of MHC diversity supports this extreme population bottleneck in which all but the most frequent alleles were lost due to inbreeding and drift. Historical records, confirmed by recent genetic analysis, indicate that animals from other parts of the UK and from Western Europe were re-introduced to restore lost or depleted UK populations including some from Scandinavia, re-introduced to secure populations in southern Scotland (Hale et al. 2004). The *Scvu-DRB1a/Scvu-DRB2a* haplotype may have originated with animals from Scandinavia which subsequently expanded throughout the UK. By extending future analyses to include samples from Scandinavia and other areas of Western Europe, the origin of the *Scvu-DRB1a/Scvu-DRB2a* haplotype may become clearer.

Consequence of limited DRB diversity in UK red squirrels

Consistent with functional class II MHC-*DRB* orthologues in other vertebrates, much of the allelic diversity is associated with non-synonymous substitutions at locations predicted to interact with peptides held within the peptide binding groove (Hughes and Nei 1989). Such diversity influences the range of peptides presented to CD4+ve T cells, one of the key regulatory cell types controlling both antibody and cellular responses to viral infection. Any reduction in the range of pathogen antigens available for recognition by the immune system may influence subsequent responses to infection at individual and population levels. However, the diversity between *DRB* loci suggests that each may present a distinct range of peptides for recognition by the immune system (Brown et al. 1993). Haplotypes with two diverse *DRB* loci will allow a wider array of peptides to be presented to T cells compared with haplotypes with only a single functional *DRB* locus. While this study has focused on the *Scvu-DRB* loci as a marker for MHC diversity, additional class II and class I loci will be included in future analyses, allowing a more complete picture of MHC haplotype diversity in squirrel populations from the UK and continental Europe.

Levels of MHC diversity in continental European red squirrels are consistent with a robust population associated with frequent genetic exchange between populations. This is in contrast to the limited diversity in the UK squirrels which is consistent with a strong founder effect which has led to low levels of diversity in the remaining small isolated populations in the UK. Inbred wildlife populations are often susceptible to

environmental change including the introduction of new pathogens and SQPV appears to be responsible for much of the decline of the UK red squirrel population (Rushton et al. 2006). Wildlife populations within a stable environment are generally resilient to the endemic pathogens; a range of which (adenovirus; Sainsbury et al. 2001), (hepatozoon species; Simpson et al. 2006) (mycobacteria; Meredith et al. 2014) have been described in red squirrels in the UK. However, the impact of these infections appears limited compared with the exotic SQPV, although they might have a stronger impact on captive collections (Everest et al. 2014; Shuttleworth et al. 2014).

Providing evidence for a direct link between MHC diversity and squirrelpox disease susceptibility remains challenging as samples from healthy animals with evidence of SQPV exposure for comparison with samples from animals known to have been killed by the virus are required. While limited MHC diversity may contribute directly to the decline of the UK red squirrel population through a failure to present protective antigens for recognition by the immune system, it may also reflect a general decrease in diversity across the genome (reviewed by Sommer et al. 2005). Previous analysis of UK red squirrel population diversity using neutral markers such as the mitochondrial d-loop (Barratt et al. 1999) and a range of microsatellites (Hale et al. 2004; Grill et al. 2009) also identified limited diversity compared with continental populations.

Limited MHC diversity has been described in other species and populations which have gone through population bottlenecks. These include the cheetah, where limited diversity at the MHC has been linked to susceptibility to viral infection (O'Brien et al. 1985) and in the Tasmanian devil, where it has been linked with susceptibility to a transmissible tumour (Siddle et al. 2007). Limited MHC diversity has also been recorded in expanding populations following a population bottleneck, including the European Beaver (Ellergren et al. 1993) the European and North American Moose (Miko and Anderson 1995) and the Mountain Goat (Mainguy et al. 2007). These populations are however predicted to remain susceptible to novel pathogen infections. The red squirrel population of the UK may provide a warning to such populations as it may be the first recorded example of a wildlife population with limited genetic diversity that expanded following a population bottleneck in the 18th century as a result of reforestation efforts (Shorten 1954) only to be decimated by an exotic viral infection in the 20th century.

It may be fortuitous, but no evidence of SQPV has been reported in continental European red squirrels despite the introduction on at least three occasions of eastern grey squirrels to Northern Italy between 1948 and the 1990s, followed by numerous translocations and undocumented releases (Martinoli et al. 2010; Bertolino et al.

2008, 2014). The Italian red squirrel population is the most genetically diverse population analysed in this study with a large number of diverse MHC haplotypes associated with high levels of heterozygosity. The absence of SQPV along with a genetically diverse red squirrel population and low levels of diversity in Italian grey squirrels (Signorile et al. 2014) may have contributed to the relatively slow spread of grey squirrels in Northern Italy compared with those in the UK (Bertolino et al. 2014).

MHC diversity and red squirrel conservation

Surprisingly, the distribution of alleles in the red squirrel population on the Isle of Arran, located off the West coast of Scotland, suggests that they are more closely related to those from continental Europe than to other squirrels from the UK. The *Scvu-DRB1a/Scvu-DRB2a* haplotype dominant in mainland UK populations was not recorded and existing records indicate that red squirrels were introduced to the island between the 1930s and 1950s. This supports an earlier study which identified two mitochondrial haplotypes in the Arran population, one of which was also found in Belgium populations (Barratt et al. 1999). As the Arran population appears unique in the UK we suggest that animals from this population could be used to expand levels of diversity and contribute to long term population health in other Scottish red squirrel strongholds (and potentially other areas of mainland UK) with established red squirrel populations with limited MHC diversity. This approach may be preferable to introductions from continental Europe with the risk of introducing exotic pathogens. Currently red squirrel reintroduction strategies in the UK are focused on controlling the grey squirrel population and habitat restoration which favours red squirrels with little regards to population genetic diversity. We suggest that by incorporating a simple measure of MHC diversity in the reintroduction strategy overall population health would be improved in the longer term.

Figure Legends

Fig. 1

Map of Western Europe showing the population number and number of red squirrels sampled from each location in parenthesis.

Fig. 2

Multiple alignments of the predicted amino acid sequences derived from three red Squirrel *DRB1* and *DRB2* transcripts aligned with nineteen *DRB1* and *DRB2* allelic sequences derived from the genomic analysis of 100 red squirrels from the UK and continental Europe. Only unique allelic sequences are included. The full length *DRB* transcript derived from the tassel-eared squirrel (*Sciurus aberti, Scab-DRB)* is used as the reference sequence. Sequences are numbered from the first amino acid of the mature protein. The portion of the DRB-β1 domain encoded by the second exon is shaded and amino acid positions predicted by Reche and Reinherz (2003) to interact with peptides within the peptide binding domain are indicated with a *. Sequence identity is indicated by a . and missing sequence is indicated by a -.

Fig. 3

Maximum likelihood tree estimating the relationships between Squirrel *DRB* nucleotide sequences. The tree is generated using the HKY substitution model and rooted using the murine *DRB* orthologue, *H2-EB1*, (NM_010382). Only bootstrap values 60 or above are shown. Species designations are as follows; *Scab, Sciurus aberti* (tassel eared squirrel, M97616); *Scvu, Sciurus vulgaris* (Eurasian Red squirrel, LN832043 to LN832063), *Scca, Sciurus carolinensis* (eastern grey squirrel).

Fig. 4

Distribution of *Scvu-DRB1* and *DRB2* allelic diversity in each red squirrel population.

Tables

Table 1. PCR primers

Primer	Specificity	Template/Location	Sequence
Scvu351F	*DRB1* and *DRB2*	gDNA, exon 2	5'-AGTGCCATTTCTACAACGGGAC-3'
Scvu338R	*DRB1* and *DRB2*	gDNA, exon 2	5'-CTCTCCGCTCCACAGTGAAGC-3'
Scvu363F	*DRB1* and *DRB2*	cDNA, exon 1	5'-TCCTCTCCTGTTCTCCAGCAT-3'
Scvu364R	*DRB1* and *DRB2*	cDNA, exon 3	5'-CACAGTCACCTTCGGCTTAAC-3'

Table 2. *Scvu-DRB1/DRB2* allelic nomenclature and associated accession numbers

***Scvu-DRB* allele**	**Accession Number**	***Scvu-DRB* allele**	**Accession Number**
Scvu-DRB1a	LN832039	*Scvu-DRB1m*	LN832052
Scvu-DRB1b	LN832040	*Scvu-DRB1n*	LN832053
Scvu-DRB1c	LN832042	*Scvu-DRB1o*	LN832054
Scvu-DRB1d	LN832043	*Scvu-DRB1p*	LN832055
Scvu-DRB1e	LN832044	*Scvu-DRB1q*	LN832056
Scvu-DRB1f	LN832045	*Scvu-DRB1r*	LN832057
Scvu-DRB1g	LN832046	*Scvu-DRB1s*	LN832058
Scvu-DRB1h	LN832047	*Scvu-DRB2a*	LN832041
Scvu-DRB1i	LN832048	*Scvu-DRB2b*	LN832059
Scvu-DRB1j	LN832049	*Scvu-DRB2c*	LN832060
Scvu-DRB1k	LN832050	*Scvu-DRB2d*	LN832061
Scvu-DRB1l	LN832051	*Scvu-DRB2e*	LN832062

Table 3. *Scvu-DRB* allelic frequencies associated with individual populations

Population Allelic Frequencies												
Population	1	2	3	4	5	6	7	8	9	10	11*	12#
N	10	6	6	13	3	10	10	10	2	12	10	8
DRB1a	1.0	0.50	0.5	0.69	0.667	0.75	0.65	0.09	0.75	-	-	-
DRB1b		0.167	0.42	0.31	0.333	0.2	-	-	-	-	-	-
DRB1c	-	-	-	-	-	-	0.1	-	-	-	-	-
DRB1d	-	-	-	-	-	-	-	-	-	0.375	0.45	0.062
DRB1e	-	-	-	-	-	-	-	-	0.25	-	0.25	-
DRB1f	-	0.333	0.08	-	-	-	0.05	-	-	-	-	-
DRB1g	-	-	-	-	-	-	-	0.1	-	-	-	-
DRB1h	-	-	-	-	-	-	-	-	-	0.625	0.1	-
DRB1i	-	-	-	-	-	-	-	-	-	-	0.1	-
DRB1j	-	-	-	-	-	-	-	-	-	-	0.1	-
DRB1k	-	-	-	-	-	-	-	-	-	-	-	0.062
DRB1l	-	-	-	-	-	-	-	-	-	-	-	0.062
DRB1m	-	-	-	-	-	0.05	0.1	-	-	-	-	-
DRB1n	-	-	-	-	-	-	0.1	-	-	-	-	-
DRB1o	-	-	-	-	-	-	-	-	-	-	-	0.537
DRB1p	-	-	-	-	-	-	-	-	-	-	-	0.125
DRB1q	-	-	-	-	-	-	-	-	-	-	-	0.125
DRB1r	-	-	-	-	-	-	-	-	-	-	-	0.062
DRB1s	-	-	-	-	-	-	-	-	-	-	-	0.062
DRB2a	1.0	1.0	1.0	1.0	1.0	1.0	1.0	0.9	1.0	1.0	0.85	0.75
DRB2b	-	-	-	-	-	-	-	0.1	-	-	0.15	-
DRB2c	-	-	-	-	-	-	-	-	-	-	-	0.125
DRB2d	-	-	-	-	-	-	-	-	-	-	-	0.062
DRB2e	-	-	-	-	-	-	-	-	-	-	-	0.062

Legend Table 3, N = Number of individuals genotyped; *, Belgian population; #, Italian population

402 **References**

403 Anisimova M, Gascuel O (2006) Approximate likelihood ratio test for branches: A fast, accurate and powerful
404 alternative. Systematic Biology 55:539-552

405 Babik W, Durka W, Radwan J (2005) Sequence diversity of the MHC DRB gene in the Eurasian beaver (Castor
406 fiber). Mol Ecol 14: 4249–4257.

407 Barratt EM, Gurnell J, Malarky G, Deaville R, Bruford MW (1999) Genetic structure of fragmented populations
408 of red squirrel (*Sciurus vulgaris*) in the UK. Mol Ecol 8:55-63

409 Barreiro LB, Quintana-Murci L (2010) From evolutionary genetics to human immunology: how selection
410 shapes host defense genes. Nature Rev Genet 11:17-30

411 Bernatchez L, Landry C (2003) MHC studies in nonmodel vertebrates: what have we learned about natural
412 selection in 15 years? J Evol Biol 16:363-377

413 Bertolino S, Lurz PWW, Sanderson R, Rushton SP (2008) Predicting the spread of the American grey squirrel
414 (*Sciurus carolinensis*) in Europe: a call for a co-ordinated European approach. Biological Conservation
415 141:2564-2575

416 Bertolino S, Cordero di Montezemolo N, Wauters LA, Martinoli A (2014) A grey future for Europe: *Sciurus*
417 *carolinensis* is replacing red squirrels in Italy. Biological Invasions, 16:53-62

418 Biedrzycka A, Radwan J (2008) Population fragmentation and major histocompatibility complex variation in the
419 spotted suslik, *Spermophilus suslicus*. Mol Ecol 17:4801-4811

420 Brown JH, Jardetzky TS, Gorga JC, Stern LJ, Urban RG, Strominger JL, Wiley DC (1993) Three-dimensional
421 structure of the human class II histocompatibility antigen HLA-DR1. Nature 364:33-39

422 Bjorkman PJ, Saper MA, Samraoui B, Bennett WS, Strominger JL, Wiley DC (1987) Structure of human class I
423 histocompatibility antigen. Nature, 329:506-512

424 Charlesworth D, Willis JH (2009) The genetics of inbreeding depression. Nature Rev Genet 10**:**783-796

425 Cížková D, Gouy de Bellocq J, Baird SJ, Piálek J, Bryja J (2011) Genetic structure and contrasting selection
426 pattern at two major histocompatibility complex genes in wild house mouse populations. Heredity
427 106:727-740

428 Darby AC, McInnes CJ, Kjær KH Wood AR, Hughes M, Martensen PM, Radford AD, Hall N, Chantrey J
429 (2014) Novel Host-Related Virulence Factors Are Encoded by Squirrelpox Virus, the Main Causative
430 Agent of Epidemic Disease in Red Squirrels in the UK. PLoS ONE 9**:**e96439

431 Ellegren H, Hartman G, Johansson M, Andersson L (1993) Major histocompatibility complex monomorphism
432 and low-levels of DNA fingerprinting variability in a reintroduced and rapidly expanding population of
433 beavers. Proc Natl Acad Sci USA 90:8150-8153

434 Everest DJ, Shuttleworth CM, Stidworthy MF, Grierson SS, Duff JP, Kenward RE (2014) Adenovirus: an
435 emerging factor in red squirrel *Sciurus vulgaris* conservation. Mammal Rev 44**:**225-233

436 Frankham R, Ralls K (1998) Inbreeding leads to extinction. Nature 392:441-442

437 Germain RN, Margulies DH (1993) The biochemistry and cell biology of antigen processing and presentation.
438 Annu Rev Immunol 11:403-450

439 Grill A, Amori G, Aloise G, Lisi I, Tosi G, Wauters LA, Randi E (2009) Molecular phylogeography of
440 European *Sciurus vulgaris*: refuge within refugia? Mol Ecol 18:2687-2699

441 Gurnell J, Clark MJ, Lurz PWW, Shirley MDF, Rushton SP (2002) Conserving red squirrels (Sciurus vulgaris):
442 mapping and forecasting habitat suitability using a Geographic Information Systems Approach.
443 Biological Conservation 105:53-64

444 Gurnell J, Wauters LA, Lurz PW, Tosi G (2004) Alien species and interspecific competition: effects of
445 introduced eastern grey squirrels on red squirrel population dynamics J Anim Ecol 73:26-35

446 Hale ML, Lurz PWW, Shirley MDF, Rushton S, Fuller RM, Wolff K (2001) Impact of Landscape management
447 on the genetic structure of red squirrel populations. Science 293:2246-2248

448 Hale M.L, Lurz PWW, Wolff K (2004) Patterns of genetic diversity in the red squirrel: footprints of
449 biogeographic history and artificial introductions. *Conservation Genetics* 5**:**167-179

450 Harris S, Morris P, Wray S (1995) A Review of British Mammals: Population Estimates and Conservation
451 Status of British Mammals Other than Cetaceans. Report to the Joint Nature Conservation Committee,
452 Peterborough, UK

453 Horton R, Wilming L, Rand V Lovering RC, Bruford EA, Khodiyar VK, Lush MJ, Povey S, Talbot CC Jr,
454 Wright MW, Wain HM, Trowsdale J, Ziegler A, Beck S (2004) Gene map of the extended human
455 MHC. Nature Rev Genet 5:889-899

456 Hughes AL, Yeager M (1998) Natural selection at major histocompatibility complex loci of vertebrates. Annu
457 Rev Genet 32:415-434

458 Hughes AL, Nei M (1989) Nucleotide substitution at major histocompatibility complex class II loci: Evidence
459 for overdominant selection. Proc Natl Acad Sci USA 86:948-962

460 Jukes TH, Cantor CR (1969) Evolution of protein molecules In *Mammalian protein metabolism III* (HN Munro,
461 ed), Academic Press, New York: 21-132

462 Keller L, Waller D (2002) Inbreeding effects in wild populations. Trends in Ecology and Evolution 17:230-241

463 Kenward RE, Holm JL (1993) On the replacement of the red squirrel in Britain: a phytotoxic explanation. Proc
464 R Soc B- Biolog Sci 251:187-194

465 Klein J, Bontrop RE, Dawkins RL Erlich HA, Gyllensten UB, Heise ER, Jones PP, Parham P, Wakeland EK,
466 Watkins DI (1990) Nomenclature for the major histocompatibility complexes of different species: a
467 proposal. Immunogenetics, 31:217-219

468 Kumanovics A, Takada, T, Lindahl, KF (2003) Genomic organization of the mammalian MHC. Annu Rev
469 Immunol 21:629-657

470 La-Rose JP, Meredith AL, Everest DJ, Fiegna C, McInnes CJ, Shaw DJ, Milne EM (2010) Epidemiological and
471 postmortem findings in 262 red squirrels (*Sciurus vulgaris*) in Scotland, 2005 to 2009. Vet Rec
472 167:297-302

473 Mainguy J, Worley K, Cote SD (2007) Low MHC *DRB* class II diversity in the mountain goat: past bottlenecks
474 and possible role of pathogens and parasites. Conservation Genet 8:885–891

475 Martinoli A, Bertolino B, Preatoni DG Balduzzi A, Marsan A, Genovesi P, Tosi G, Wauters LA (2010)
476 Headcount 2010: the multiplication of the grey squirrel populations introduced in Italy. *Hystrix* Italian J
477 of Mammalogy 21:127-136

478 Meredith A, Del Pozo J, Smith S, Milne E, Stevenson K, McLuckie J (2014) Leprosy in red squirrels in
479 Scotland. Vet Rec 175:285-286

480 Meyer D, Thomson G (2001) How selection shapes variation on the human major histocompatibility complex: a
481 review. Annals of Human Genet 65:1-26

McInnes CJ, Wood AR, Thomas K, Sainsbury AW, Gurnell J, Dein FJ, Nettleton PF (2006) Genomic characterization of a novel poxvirus contributing to the decline of the red squirrel (*Sciurus vulgaris*) in the UK. J Gen Virol 87:2115-2125

Middleton A.D (1930) Ecology of the American gray squirrel in the British Isles. Proc Zool Soc Lond, 100:809–843

Mikko S, Andersson L (1995) Low major histocompatibility complex class-II diversity in European and North-American moose. Proc Natl Acad Sci USA 92:4259-4263

Milne I, Lindner D, Bayer M, Husmeier D, McGuire G, Marshall DF, Wright F (2008) TOPALi v2: a rich graphical interface for evolutionary analyses of multiple alignments on HPC clusters and multi-core desktops. Bioinformatics 25:126-127

O'Brien SJ, Roelke ME, Marker L, Newman A, Winkler CA, Meltzer D, Colly L, Evermann JF, Bush M, Wildt DE (1985) Genetic basis for species vulnerability in the cheetah. Science 227:1428-1434

Osborne AJ, Pearson J, Negro SS, Chilvers BL, Kennedy MA, Gemmell NJ (2015) Heterozygote advantage at MHC *DRB* may influence response to infectious disease epizootics. Mol Ecol 24:1419-1432

Radwan J, Biedrzyck A, Babik, W (2010) Does reduced MHC diversity decrease viability of vertebrate populations? Biological Conservation 143**:**537-544

Reche PA, Reinherz EL (2003) Sequence variability analysis of human class I and class II MHC molecules: functional and structural correlates of amino acid polymorphisms. J Mol Biol 331:623-641

Ricanova S, Bryja J, Cosson J-F Gedeon C, Choleva LS Ambros M, Sedlacek F (2011) Depleted genetic variation of the European ground squirrel in Central Europe in both microsatellites and the major histocompatibility complex gene: implications for conservation. Conservation Genet 12:1115-1129

Robinson J, Halliwell JA, McWilliam H, Lopez R, Marsh SG (2013) IPD - the Immuno Polymorphism Database. Nucleic Acids Res 41:D1234-1240

Rushton SP, Lurz PW, Gurnell J, Nettleton P, Bruemmer C, Shirley MD, Sainsbury AW (2006) Disease threats posed by alien species: the role of a poxvirus in the decline of the native red squirrel in Britain. Epidemiol Infect 134:521-533

Sainsbury AW, Adair B, Graham D (2001) Isolation of a novel adenovirus associated with splenitis, diarrhoea, and mortality in translocated red squirrels. *Sciurus vulgaris.* Verhandlungsberichte über Erkrankungen der Zootiere 40:265-270

Sainsbury AW, Nettleton P, Gilray J Gurnell J (2000) Grey squirrels have a high seroprevalence to a parapoxvirus associated with deaths in red squirrels. Animal Conservation 3:229-233.

Shorten M. (1954) Squirrels. Collins, London

Shuttleworth CM, Everest DJ, McInnes CJ, Greenwood A, Jackson NL, Rushton S, Kenward RE (2014) Inter-specific viral infections: Can the management of captive red squirrel collections help inform scientific research? Hystrix, the Italian J Mammalogy 25:18-24

Siddle HV, Kreiss A, Eldridge MD, Noonan E, Clarke CJ, Pyecroft S, Woods GM, Belov K (2007) Transmission of a fatal clonal tumor by biting occurs due to depleted MHC diversity in a threatened carnivorous marsupial. Proc Natl Acad Sci USA 104:6221-16226

Signorile AL, Wang J, Lurz PWW, Bertolino S, Carbone C, Reuman DC (2014) Do founder size, genetic diversity and structure influence rates of expansion of North American grey squirrels in Europe? Diversity Distrib 20:918-930

523 Simpson VR, Birtles RJ, Bown KJ, Panciera RJ, Butler H, Davison N (2006) *Hepatozoon* species infection in
524 wild red squirrels *(Sciurus vulgaris)* on the Isle of Wight. Vet Rec 159:202-205

525 Sommer S (2005) The importance of immune gene variability (MHC) in evolutionary ecology and conservation.
526 Frontiers in Zool 2:1-18

527 Tamura K, Stecher G, Peterson D, Filipski A, Kumar S (2013) MEGA6: Molecular Evolutionary Genetics
528 Analysis version 6.0. Mol Biol Evol 30:2725-2729

529 Thomas K, Tompkins D, Sainsbury A, Wood AR, Dalziel R, Nettleton PF, McInnes CJ (2003) A novel poxvirus
530 lethal to red squirrels (*Sciurus vulgaris*). J Gen Virol 84:3337-3341

531 Tompkins D, Sainsbury AW, Nettleton P, Buxton D, Gurnell J (2002) Parapoxvirus causes a deleterious disease
532 of red squirrels associated with UK population declines. Proc R Soc Lond 269:529-533

533 Trowsdale J (2011) The MHC, disease and selection. Immunol Letters 137:1-8

534 Wauters LA, Gurnell J (1999) The mechanism of replacement of red squirrels by grey squirrels: A test of the
535 interference competition hypothesis. Ethology 105:1053-1071