

Does individual genetic diversity of red squirrels (*Sciurus vulgaris*) influence susceptibility to squirrelpox virus (SQPV) disease?

Melissa M. Marr,¹ Kate Ciborowski,² John Gurnell,³ Dada Gottelli,² Peter W. W. Lurz,⁴ Elspeth Milne,⁴ Katy Morgan,⁵ Anthony W. Sainsbury,² Jinliang Wang,² and William C. Jordan,²

¹Department of Earth Sciences, Natural History Museum London, Cromwell Road, London, SW7 5BD

²Institute of Zoology, Zoological Society of London, Regent's Park, London NW1 4RY

³School of Biological and Chemical Sciences, Queen Mary, University of London, Mile End Road, London, E1 4NS

⁴Royal (Dick) School of Veterinary Studies, The University of Edinburgh, Easter Bush, Midlothian, EH25 9RG

⁵Anthony Lab, University of New Orleans, 2000 Lakeshore Dr, New Orleans, LA 70148, United States

Introduction

British populations of the Eurasian red squirrel *Sciurus vulgaris* have experienced severe declines coupled with multiple local extinctions over the past 150 years (Lloyd 1983; Gurnell & Pepper 1993). While habitat loss and fragmentation have been suggested as contributory factors in historical declines (Shorten 1964), more recent declines are thought to be strongly influenced by interspecific competition with an invasive squirrel species, the North American grey squirrel *Sciurus carolinensis* (Gurnell & Pepper 1993; Gurnell et al. 2004). This species has replaced *S. vulgaris* over much of its former range and evidence strongly suggests that parasite-mediated competition through squirrelpox virus (SQPV) has played a major role in this process (Rushton et al. 2000; Sainsbury et al. 2000; Tompkins et al. 2003). While there is some uncertainty with respect to the origins of the disease, the facts that no clinical signs of SQPV were detected prior to the introduction of *S. carolinensis* and that serum of grey squirrels from Wisconsin, USA tested positive to antibodies of the virus suggests that SQPV represents an exotic infectious agent that was introduced with greys from North America (Sainsbury & Gurnell 1995; McInnes et al. 2006). Grey squirrels are largely unaffected by SQPV but their populations act as reservoirs of infection for the virus that spill over to infect red squirrels in which SQPV causes a lethal disease (Sainsbury et al. 2000; Tompkins et al. 2002; Thomas et al. 2003; Tompkins et al. 2003). In mainland Wales and northern England red squirrels are now largely restricted to small, fragmented populations while in Scotland they are found in larger, more contiguous woodland habitat, although Scottish populations are also in decline.

Previous genetic analyses of British red squirrel populations have uncovered patterns of low within-population diversity and high among population differentiation, indicative of severe historical bottlenecks and little contemporary gene flow (Barratt et al. 1999; Hale et al. 2004). Low levels of genetic diversity in wild populations have been associated with reductions in fitness at both the individual and population level while positive associations between higher levels of genetic diversity have been reported for several components of individual fitness such as body weight, survival, fecundity, behaviour, reproductive success and physiology (reviewed in Keller & Waller 2002; Reed & Frankham 2003). The observed loss of genetic diversity associated with small population size and low levels of immigration could reduce the ability of red squirrel populations to cope with a novel virus such as SQPV via the negative effects of genetic drift and reduced mating opportunities. Under a scenario where population size is small and immigration rates low, deleterious and recessive alleles will tend to accumulate as genome-wide homozygosity (including loci affecting fitness) increases (Frankham 1996; Keller & Waller 2002; Szulkin et al. 2010). This could potentially increase both the susceptibility of individual red squirrels to SQPV infection and the ability of populations to evolve adaptations in response to the virus.

Heterozygosity-Fitness Correlations (HFC's) aim to correlate levels of individual genetic diversity with key components of fitness. As pedigree information is unavailable for most natural populations, molecular estimates - typically using co-dominant, presumed selectively neutral markers such as microsatellites - provide a proxy measure of an individual's inbreeding coefficient (f , Grueber et al. 2008; Chapman et al. 2009; Szulkin et al. 2010). Where non-coding markers are used these associations are hypothesised to arise through two mechanisms which both assume that heterozygosity at neutral loci reflect heterozygosity at loci directly affecting fitness: i) 'local effects' - where the markers scored are in linkage disequilibrium (LD);

David 1998) with fitness influencing loci (David et al. 1995; Lynch & Walsh 1998; Hansson & Westerberg 2002; Hansson et al. 2004) and ii) general effects – where markers scored reflect genome-wide heterozygosity at fitness influencing loci with which they are in identity disequilibrium (ID; David 1998), effectively inbreeding (David et al. 1995; Lynch and Walsh 1998; Hansson & Westerberg 2002). The observation that HFC's in large, outbred populations are often weak and inconsistent led to the proposition that linkage disequilibrium was the main driver of such associations (Hansson et al. 2004; Acevedo-Whitehouse et al. 2006; Lieutenant-Gosselin & Bernatchez 2006; Malo & Coulson 2009). However, local effects can only strengthen HFC's, not generate them (Szulkin et al. 2010; Ruiz-López et al. 2012) and in contexts where population size is small, historical bottlenecks have occurred and matings among close relatives are common - such as has been reported for British red squirrels - HFC's due to genome-wide inbreeding may be easier to detect. Accordingly, this study aims to test the hypothesis that lower levels of genetic diversity, measured through microsatellite markers in red squirrels are associated with i) higher likelihood of acquiring SQPV disease and ii) lower body condition.

Material and Methods

Sample collection and microsatellite typing

Individuals were sourced from an existing database of wild British red squirrel carcasses from across northern England and southern Scotland collected and subjected to post-mortem examination from 1993 onwards. As inbreeding theory predicts that inbred individuals will show higher levels of genome-wide homozygosity relative to outbred individuals (Chapman et al. 2009), we would expect to see a difference in marker heterozygosity between diseased and healthy animals if inbreeding is associated with cases of the disease. This study therefore adopted a comparative case/control approach. Individuals were separated into two groups i) a case group consisting of 98 individuals which showed pathological lesions characteristic of SQPV and which tested positive for the disease via electron-microscopy and ii) a control group consisting of 101 individuals which had died of probable acute trauma (road traffic accidents; RTA's) and which showed no skin lesions or other signs of disease. The control group was selected to contain similar proportions of individuals for each age group, gender, season of death and region collected. During post-mortem examination body weight (BW) was measured to the nearest gram and shin length (SL) measured to the nearest mm. These data were used to construct a proxy measure of condition, Condition Factor (CF), calculated as:

$$CF = \left(\frac{BW}{SL^3} \right)$$

DNA was extracted from individual squirrels and each was genotyped at a panel of 9 previously published microsatellite loci (Hale et al. 2001; Supplementary Material a).

Determination of Population Structure

Individuals in each group may not be from a homogenous population, and any hidden genetic structure may affect the comparison between groups. In order to estimate how many populations (K) were in our sample and to assign each individual to one or more of these populations, a Bayesian clustering approach implemented in STRUCTURE v2.3.4 (Pritchard et al. 2000; Falush et al. 2003, 2007) was used to analyse the genotype data. To obtain the most likely number of clusters (K), an ad-hoc statistic, deltaK (ΔK ; Evanno et al. 2005), was employed in STRUCTURE Harvester (Earl & von Holdt 2012). Four separate STRUCTURE analyses were carried out, all with a burn-in length of 10^5 and a run length of 10^6 after burn-in. In each analysis, 10 independent simulations per K value were conducted with K varying from 1 to 8. The first and second analyses assumed the admixture model, with and without population included as a prior - where in this case population has been pre-defined geographically by a GIS land-cover linkage approach (see Supplementary Material b). The third and fourth analyses assumed the no admixture model, again with and without population included as a prior. In addition, and as a complementary

analysis to the STRUCTURE analysis, a pairwise matrix of genetic distances was calculated by Genepop v4.2 (Raymond & Rousset 1995) using the fixation index ' F_{ST} ' (Weir and Cockerham 1984) and the eight geographically defined populations as the grouping criterion.

Analysis of Data Controlled for Population Structure

Individuals were grouped into populations defined by the STRUCTURE analysis and tests for departures from Hardy-Weinberg equilibrium (HWE) were carried out in Genepop v4.2 (Raymond & Rousset 1995) with demorization number, number of batches and number of iterations of batches set to 1000. Three measures of inbreeding were calculated using the Rhh package (Alho et al. 2010) in Rv.3.0.2. Multilocus heterozygosity (MLH) was calculated for each individual as the proportion of typed loci at which an individual was heterozygous, Internal Relatedness (IR) was calculated as in Amos *et al.* 2001 and heterozygosity weighted by loci (HL) was calculated as in Aparicio *et al.* 2007. These measures were all strongly correlated in the direction expected (see Supplementary Material c) and showed highly congruent patterns of results therefore only the results for MLH are reported here.

To test for single locus effects an Excel-macro, GEPHAST (Amos & Acevedo-Whitehouse 2009), was used to calculate correlations between heterozygosity at individual loci and the presence of SQPV disease. This method identifies genotypes, quantifies which genotypes have unusual phenotypic values, classifies unique genotypes as either heterozygous or homozygous, assesses significance by randomization and performs a chi-square test to output a significance value. Data was partitioned into populations inferred by the STRUCTURE analysis and each population analysed separately in GEPHAST under the binary chi-squared option. Within population data was examined for the assumptions of parametric tests. For data that met these assumptions parametric tests were carried out and for those that did not, non-parametric tests were used. A Mann-Whitney U test or an Independent Samples T-test was used to test for significant differences in mean MLH between males and females and between case and control groups, an Independent Samples T-Test was used to test for significant differences in CF between males and females and between case and control groups and a simple correlation (Kendall's Tau-b) was used to test for significant relationships between CF and MLH in case and control groups.

Results

Under the first three STRUCTURE analyses the most likely value of K using the Evanno et al. 2005 ΔK method was $K=2$ regardless of whether population information was included as a prior in the model or not. In the fourth scenario, which assumed No Admixture but included population as a prior, the best-fit value was $K=3$. Taking into account the geographical distribution of these populations and the ecology of *S. vulgaris* it was decided to adopt the $K=3$ scenario as the most realistic representation of British red squirrel populations generated here (Figure 1). Partitioning our data into the populations inferred by the STRUCTURE analysis resulted in three large and well defined groups which, within our study area, roughly corresponds to a southern Scotland, a north/north-eastern England group and a west/south-western England group. This result seems to contradict previous studies which have reported strong population structure in this species (Barratt et al. 1999; Hale et al. 2004). However, in contrast to the STRUCTURE results a pairwise fixation index analysis (F_{ST} ;table 1) for all eight geographically defined populations detected very high levels of genetic subdivision between many population pairs with 16 pairs showing F_{ST} values > 0.20 , suggesting that there may be hidden internal sub-structure in each of the 3 populations inferred by STRUCTURE.

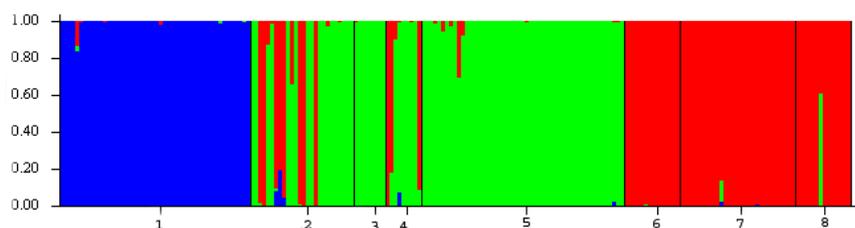


Figure 1. Inferred population structure under the no admixture model with geographically defined populations included as a prior. Most likely value of $K=3$.

Table 1. F_{ST} values between all pairs of geographically defined populations with all values > 0.20 highlighted.

	1	2	3	4	5	6	7	8
1	-							
2	0.16	-						
3	0.29	0.06	-					
4	0.21	0.01	0.01	-				
5	0.25	0.05	0.01	0.03	-			
6	0.17	0.07	0.23	0.09	0.2	-		
7	0.23	0.14	0.27	0.17	0.26	0.08	-	
8	0.36	0.23	0.33	0.2	0.34	0.2	0.18	-

No single locus effects were detected in any population bar population 1 in which Scv1, Scv32 and Scv19 all had p values < 0.05 (Table 2). In population 1 all microsatellite marker loci were in HWE bar Scv32, Scv31 and Scv16. In populations 2 and 3 highly significant deviations from HWE were observed at nearly all loci suggestive of severe heterozygote deficiency (Tables 3-8). There were no significant differences in mean MLH or CF between males and females in any population therefore sexes were pooled for all further analyses. In population 1 mean MLH differed significantly between case and control samples with the case group having lower mean MLH ($p = 0.002$; Case, 0.25; Control, 0.39) but this result was not observed in population 2 ($p = 0.837$) or 3 ($t = 0.533$, $p = 0.596$). Condition factor (CF) differed significantly between case and control groups in population 3 only ($t = -4.59$, $p = 0.000$) and a significant, negative correlation between CF and MLH was observed in the case group for this population ($r = -0.381$, $p = 0.016$) but not in the control group ($r = 0.153$, $p = 0.276$). However, no significant correlation between CF and MLH was observed in any other population (Tables 3-8, figures 2-7).

Table 2. Significance values for single locus effects by population generated by the GEPHAST analysis (Amos et al. 2009)

	Scv1	Scv9	Scv32	Scv3	Scv6	Scv31	Scv13	Scv16	Scv19
Population 1	0.043*	0.154	0.259	0.001**	0.467	0.307	0.346	0.12	0.005**
Population 2	0.601	0.204	0.214	0.888	0.972	0.064	0.585	0.147	0.779
Population 3	0.21	0.24	0.64	0.35	0.13	0.26	0.67	0.65	0.82

Tables 3 and 4 Multiplex results, tests from departure from HWE and statistical tests for population 1 ($n = 48$; case $n = 24$, control $n = 24$)

Multiplex	Locus	Repeat sequence	N alleles	N individuals typed	Freq obs heterozygotes	Freq exp heterozygotes	HWE
A	Scv1	[CA] ₂₈	5	47	0.52	0.79	0.18
A	Scv9	[ACAG] ₂ [AC] ₁₄	3	48	0.21	0.22	0.47
A	Scv32	[TC] ₅ ...[GA] ₂₉	6	45	0.38	0.49	0.02**
B	Scv3	[GA] ₂₈	5	48	0.48	0.54	0.18
B	Scv6	[TG] ₂₂	3	48	0.21	0.26	0.09
B	Scv31	[AG] ₂₉	4	43	0.17	0.38	0***
C	Scv13	[GA] ₁₅	2	48	0.25	0.28	0.37
C	Scv16	[ATAC] ₂ [AC] ₁₂	2	48	0.23	0.27	0***
C	Scv19	[TG] ₁₁	2	48	0.38	0.43	0.27
Statistical Test		Null hypothesis			Test Statistic	p-value	
Mann-Whitney U test		No significant differences in mean MLH between case and control samples			-	0.002**	
Mann-Whitney U test		No significant differences in mean MLH between male and females			-	0.43	
Independent Samples T-test		No significant differences in mean CF between case and control samples			$t = -0.934$	0.355	
Independent Samples T-test		No significant differences in mean CF between males and females			$t = -1.246$	0.22	
Simple correlation (Kendalls tau-b)		No significant correlation between CF and MLH in case group			$r = 0.098$	0.526	
Simple correlation (Kendalls tau-b)		No significant correlation between CF and MLH in control group			$r = -0.082$	0.595	

Tables 5 and 6. Multiplex results, tests from departure from HWE and statistical tests for population 2 ($n = 94$; case $n = 47$, control $n = 47$)

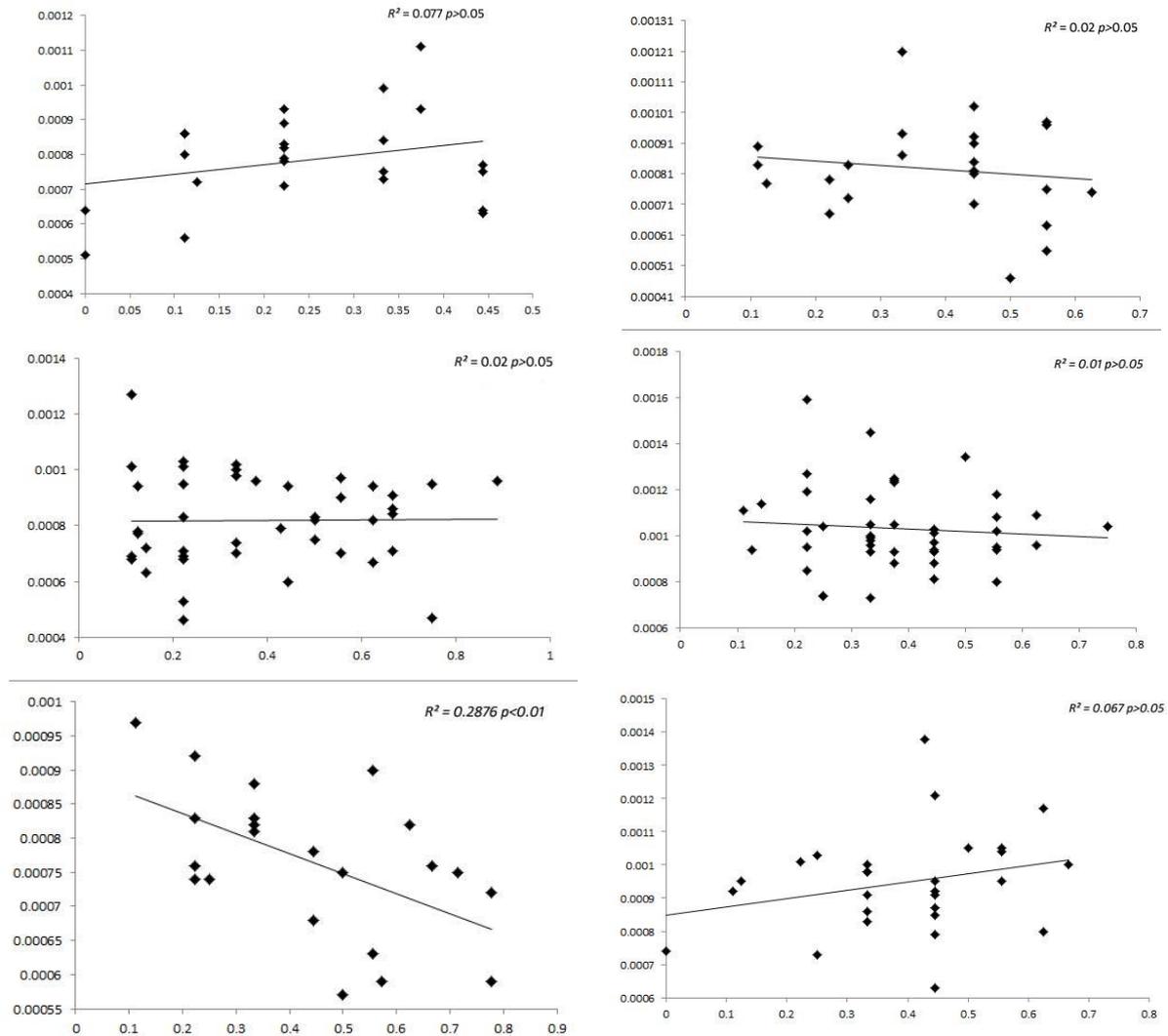
Multiplex	Locus	Repeat sequence	N alleles	N individuals typed	Freq obs heterozygotes	Freq exp heterozygotes	HWE
A	Scv1	[CA] ₂₈	6	93	0.30	0.38	0.003**
A	Scv9	[ACAG]2[AC]14	4	92	0.38	0.56	0***
A	Scv32	[TC]5...[GA]29	10	89	0.65	0.74	0.014**
B	Scv3	[GA]28	8	89	0.33	0.51	0***
B	Scv6	[TG]22	5	90	0.44	0.62	0.001***
B	Scv31	[AG]29	7	77	0.21	0.58	0***
C	Scv13	[GA]15	2	92	0.11	0.22	0***
C	Scv16	[ATAC]2[AC]12	4	93	0.39	0.48	0.012**
C	Scv19	[TG]11	3	93	0.48	0.51	0.27

Statistical Test	Null hypothesis	Test Statistic	p-value
Mann-Whitney U test	No significant differences in mean MLH between case and control samples	-	0.837
Mann-Whitney U test	No significant differences in mean MLH between male and females	-	0.154
Independent Samples T-test	No significant differences in mean CF between case and control samples	$t = -1.583$	0.121
Independent Samples T-test	No significant differences in mean CF between case and control samples	$t = 0.960$	0.342
Simple correlation (Kendalls tau-b)	No significant differences in mean CF between case and control samples	$r = 0.032$	0.768
Simple correlation (Kendalls tau-b)	No significant differences in mean CF between case and control samples	$r = -0.038$	0.728

Tables 7 and 8. Multiplex results, tests from departure from HWE and statistical tests for population 3 ($n = 57$; case $n = 27$, control $n = 30$)

Multiplex	Locus	Repeat sequence	N alleles	N individuals typed	Freq obs heterozygotes	Freq exp heterozygotes	HWE
A	Scv1	[CA] ₂₈	7	55	0.47	0.65	0.001**
A	Scv9	[ACAG]2[AC]14	6	56	0.39	0.68	0***
A	Scv32	[TC]5...[GA]29	8	55	0.53	0.72	0***
B	Scv3	[GA]28	8	55	0.49	0.65	0***
B	Scv6	[TG]22	4	54	0.33	0.55	0***
B	Scv31	[AG]29	11	52	0.39	0.75	0***
C	Scv13	[GA]15	4	55	0.28	0.47	0***
C	Scv16	[ATAC]2[AC]12	6	57	0.44	0.60	0.006**
C	Scv19	[TG]11	3	56	0.30	0.59	0***

Statistical Test	Null hypothesis	Test Statistic	p-value
Independent Samples T-test	No significant differences in mean MLH between case and control samples	$t = 0.595$	0.554
Independent Samples T-test	No significant differences in mean MLH between male and females	$t = -0.898$	0.373
Independent Samples T-test	No significant differences in mean CF between case and control samples	$t = -4.69$	0***
Independent Samples T-test	No significant differences in mean CF between males and females	$t = 0.013$	0.99
Simple correlation (Kendalls tau-b)	No significant correlation between CF and MLH in case group	$r = -0.381$	0.016**
Simple correlation (Kendalls tau-b)	No significant correlation between CF and MLH in control group	$r = 0.153$	0.276



Figures 2-7. Relationship between Multilocus Heterozygosity (MLH) and Condition Factor (CF) for populations 1 (top), 2 (middle) and 3 (bottom) for case (left) and control (right) groups. The only significant result returned was a negative correlation in the case group of population 3.

Discussion

This study attempted to investigate the relationship between individual genetic diversity measured by multi-locus heterozygosity (MLH), the presence of squirrelpox viral (SQPV) disease and a proxy measure of body condition (Condition Factor, CF) in British red squirrels sourced from populations in northern England and southern Scotland. Patterns of results were not congruent among populations. We found no support for the first hypothesis, that lower individual genetic diversity was associated with greater likelihood of developing SQPV-disease, in either population 2 or 3 but a significantly higher mean MLH was detected in the case group of population 1. A significantly higher mean CF in the case group was detected in population 3 which also showed a significant correlation between MLH and CF in the case, but not control, group. This correlation was in the opposite direction to that expected; *i.e* within SQPV-diseased squirrels CF was lower in those individuals with higher levels of MLH. It has been suggested that HFC's are stronger under conditions of increased stress such as when individuals suffer a debilitating disease (Keller & Waller 2002). As weight-loss is a symptom of progression of SQPV-disease in red squirrels (Tompkins et al. 2002; Thomas et al. 2003) it is tempting to interpret this result as an indication that more heterozygous

individuals survive for a longer duration after disease development, undergoing more prolonged periods of weight-loss than less heterozygous individuals as a result.

However, an overriding difficulty with interpretation of results obtained here is the likelihood that the STRUCTURE analysis did not detect the high levels of genetic population substructure known to exist within British red squirrel populations (Barratt et al. 1999; Hale et al. 2004) and which was also suggested by our GIS land-cover linkage analysis (see Supplementary Information b). The STRUCTURE analysis suggested three large, well defined populations but these are unlikely to be an accurate representation of the true extent of population sub-structure in this species, an assumption supported by the high values in the pairwise fixation index (F_{ST}) for many populations pairs and also the significant heterozygote deficits at the majority of marker loci in populations 2 & 3. As the fixation index, F_{ST} , relies only on the amount of genetic differentiation between populations (Weir and Cockerham 1984) and the STRUCTURE method incorporates both F_{ST} and marker information content (Pritchard et al. 2000) it is likely that either that the microsatellite loci used here did not contain sufficient information or that too few individuals were typed to reveal the complete population structure. As any hidden genetic structure would reduce the power of HFC analyses, our results of MLH on CF should be viewed as conservative.

Heterozygosity-fitness correlations (HFC's) in natural populations measured with neutral co-dominant markers are hypothesised to arise primarily through the 'general effects' hypothesis where marker loci are in Identity Disequilibrium (ID) with fitness influencing loci (Szulkin et al. 2010). The strength of the relationship generated under the general effects hypothesis depends entirely on the presence of a high variance in the inbreeding coefficient (f) within populations as these will show elevated levels of ID (Chapman et al. 2009). However, variance in f is influenced by the demographic history of a population with small population size, genetic bottlenecks and high levels of consanguineous mating all being major contributory factors in generating a higher variance in f (Chapman et al. 2009). Furthermore, levels of Linkage Disequilibrium (LD) may also be higher under scenarios where population size is small, where populations have expanded from recent bottlenecks and/or have been recently admixed - increasing the likelihood of detecting 'local effects' driven HFC's (Briscoe et al. 1994; Wall et al. 2002; Gaut and Long 2003; Chapman et al. 2009). Population structure and history is therefore a major factor influencing the strength and underlying mechanism of HFC's. British red squirrels have a complex demographic history. Previous studies have uncovered evidence of historical genetic bottlenecks and low levels of gene flow between isolated populations resulting in high among population differentiation and low within population diversity (Barratt et al. 1999; Hale et al. 2004). Habitat and, consequently, population size also varies considerably over the study area with populations in some areas of the north of England restricted to small, isolated patches and populations in the areas such as Kielder Forest and the south of Scotland inhabiting larger, more contiguous blocks. Additionally there have been several historical introductions from continental Europe and translocations from other areas of the UK into some areas (Harvie-Brown 1880-81; Lowe and Gardiner 1983). British red squirrel populations and sub-populations would therefore be expected to have unique demographic histories thus the existence, strength and drivers of HFC's may differ between them. This may account for the incongruence of patterns of results observed here such as the detection of a significant HFC only in population 3 and significant single locus effects only at three loci within population 1.

Heterozygosity-fitness correlations may well exist in British red squirrel populations but the ten microsatellite markers used in this study did not provide enough statistical power to detect hidden population substructure. The importance of considering population demography has been highlighted in recent studies (Chapman et al. 2009, Szulkin et al. 2010; Kardos et al. 2014; Miller et al. 2014). While many examples exist in the literature where data have been pooled from multiple populations and subpopulations (e.g Acevedo-Whitehouse et al. 2005; Gage et al. 2006; Ortego et al. 2007; Välimäki et al. 2007; see Chapman

et al. 2009 for a review) this approach can lead to either a dilution effect where strong relationships that may be present in small and/or inbred populations are obscured by weak or absent effects in large and/or outbred populations (Chapman et al. 2009). Alternatively, HFC's could be detected in pooled data which are not present in hidden subpopulations (Slate and Pemberton 2006). This suggests that the results obtained in this study may be spurious and highlights the importance of accurately delimiting population structure and giving due consideration to past demographic processes when testing for heterozygosity-fitness correlations in natural populations.

References

- Alho, J.S., Välimäki, K., Merilä, J.** (2010). R_h: an R extension for estimating multilocus heterozygosity and heterozygosity-heterozygosity correlations. *Molecular Ecology Resources* **10**: 720-722
- Amos, W., Worthington, W. J., Fullard, K., Burg, T. M., Croxall, J. P., Bloch, D., Coulson, T.** (2001). The influence of parental relatedness on reproductive success. *Proceedings of the Royal Society of London B* **268**: 2021-2027
- Amos, W., & Acevedo-Whitehouse, K.** (2009). A new test for genotype - fitness associations reveals a single microsatellite allele that strongly predicts the nature of tuberculosis infections in wild boar. *Molecular Ecology Resources* **9**: 1102-1111
- Aparicio, J. M., Ortego, J., Corden, P. J.** (2007). Can a simple algebraic analysis predict markers-genome heterozygosity correlations? *Journal of Heredity* **98**: 93-96
- Avecedo-Whitehouse, K., Vicente, J., Gortazar, C., Höfle, U., Fernández-de-Mera, I. G., Amos, W.** (2005). Genetic resistance to bovine tuberculosis in the Iberian wild boar. *Molecular Ecology* **14**: 3209-3217.
- Avecedo-Whitehouse, K., Spraker, T. R., Lyons, E., Melin, S. R., Gulland, F., DeLong, R. L., Amos, W.** (2006). Contrasting effects of heterozygosity on survival and hookworm resistance in California sea lion pups. *Molecular Ecology* **15**:1973-82
- Barratt, E. M., Gurnall, J. W., Malarky, G., Deaville, R., Bruford, M. W.** (1999). Genetic structure of fragmented populations of red squirrel (*Sciurus vulgaris*) in the UK. *Molecular Ecology* **8**: S55-S63
- Briscoe, D., Stephens, J. C., O'Brian, S. J.** (1994). Linkage disequilibrium in admixed populations: applications in gene mapping. *Heredity* **85**: 59-63
- Chapman, J. R., Nakagawa, S., Coltman, D. W., Slate, J., Sheldon, B. C.** (2009). A quantitative review of heterozygosity-fitness correlations in animal populations. *Molecular ecology* **18**: 2746-2765
- David, P. B., Delay, P., Jarne, P.** (1995). Alternative models for allozyme-associated heterosis in the marine bivalve *Spisula ovalis*. *Genetics* **139**: 1719-1726
- David, P.** (1998). Heterozygosity-fitness correlations: new perspectives on old problems. *Heredity* **80**: 531-537
- Earl, D. A., & von Holdt.** (2012). STRUCTURE HARVESTER: a website and program for visualising STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources* **4**: 359-361
- Evanno, G, Regnaut, S., Goudet, J.** (2005). Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology* **14**: 2644-2620

- Falush, D., Stephens, M., Pritchard, J. K.** (2003). Inference of population structure using multilocus genotype data; linked loci and correlated allele frequencies. *Genetics* **164**: 1567-1587
- Falush, D., Stephens, M., Pritchard, J. K.** (2007). Inference of population structure using multilocus genotype data; dominant markers and null alleles. *Molecular Ecology Notes* **7**: 574-578
- Frankham, R.** (1996). Relationship of genetic variation to population size in wildlife. *Conservation Biology* **10**:1500-1508
- Gage, M. J. G., Surridge, A. K., Tomkins, J. L., Green, E., Wiskin, L., Bell, D. J., Hewitt, G. M.** (2006). Reduced heterozygosity depresses sperm quality in wild rabbits, *Oryctolagus cuniculus*. *Current Biology* **16**: 612-617
- Gaut, B. S., & Long, A. D.** (2003). The lowdown on linkage disequilibrium. *Plant Cell* **15**: 612-617
- Gruerber, C. E., Walis, G. P., Jamieson, I. G.** (2008). Heterozygosity-fitness correlations and their relevance to studies on inbreeding depression in threatened species. *Molecular Ecology* **17**: 3978-3984
- Gurnell, J. & Pepper, H.** (1993). A critical look at conserving the British red squirrel, *Sciurus vulgaris*. *Mammal Review* **23**: 127-137
- Gurnell, J., Wauters, L. A., Lurz, P. W. W., Torsi, G.** (2004). Alien species and interspecific competition: effects of introduced eastern grey squirrels on red squirrel population dynamics. *Journal of Animal Ecology* **73**: 26-35
- Hale, M. L., Lurz, P. W.W., Wolff, K.** (2004). New polymorphic microsatellite markers for the red squirrel (*Sciurus vulgaris*) and their applicability to the grey squirrel (*S. carolinensis*). *Molecular Ecology Notes* **1**: 47-49
- Hale, M. L., Lurz, P. W.W., Wolff, K.** (2004). Patterns of genetic diversity in the red squirrel (*Sciurus vulgaris* L.): Footprints of biogeographic history and artificial introductions. *Conservation Genetics* **5**: 167-179
- Hansson, B., Westerdahl, D., Hasselquist, M., Akesson, M.** (2001). Microsatellite diversity predicts recruitment of sibling great reed warblers. *Proceedings of the Royal Society B* **268**: 1287-1291
- Hansson, B., & Westerberg, L.** (2002). On the correlation between heterozygosity and fitness in natural populations. *Molecular Ecology* **11**: 2467-2474
- Hansson, B., Westerdahl, H., Hasselquist, M., Akesson, M., Bensch, S.** (2004). Does linkage disequilibrium generate heterozygosity-fitness correlations in great reed warblers? *Evolution* **58**: 870-879
- Harvie-Brown, J. A. (1880-81).** The history of the squirrel in Great Britain. *Proceedings of the Royal Physical Society of Edinburgh* **5**: 343-348, **6**: 31-63, **6**: 155-183
- Kardos, M., Allendorf, F., Luikart, G.** (2014). Evaluating the role of inbreeding depression in heterozygosity-fitness correlations: how useful are tests for identity disequilibrium? *Molecular Ecology Resources* **14**: 519-530
- Keller, L. F., & Waller, D. M.** (2002). Inbreeding effects in wild populations. *Trends in Ecology and Evolution* **17**: 230-241
- Lieutenant-Gosselin, M., Bernatchez, L.** (2006). Local heterozygosity-fitness correlations with global positive effects on fitness in three-spine stickleback. *Evolution* **60**: 1658-1668

- Lloyd, H. G.** (1983). Past and present distribution of Red and Grey squirrels. *Mammal Review* **13**: 69-80
- Lowe, V. P. W., & Gardiner, A. S.** (1983). Is the British squirrel (*Sciurus leucourus* Kerr) British? *Mammal Review* **13**: 57-67
- Lynch, M., Walsh, B.** (1998). *Genetics and the analysis of quantitative traits*. Sinauer Associates Inc: Sunderland, MA
- Malo, A. F., & Coulson, T.** (2009). Heterozygosity-fitness correlations and associative overdominance: new detection method and proof of principle in the Iberian wild boar. *Molecular ecology* **18**: 2741-2742
- McInnes, C. J., Coulter, L., Dagleish, M.P., Fiegna, C., Gilray, J., Willoughby, K., Cole, M., Milne, E., Meredith, A., Everest, D. J., MacMaster, A-M.** (2006). First cases of squirrelpox in red squirrels (*Sciurus vulgaris*) in Scotland. *The Veterinary Record* **164**: 528-531
- Miller, J. M., Malenfant, R. M., David P., Davis, C. S., Poissant, J., Hogg, J. T., Festa-Bianchet, M., Coltman, D. W.** (2014). Estimating genome-wide heterozygosity: effects of demographic history and marker type. *Heredity* **112**: 240-247
- Ortego, J., Calabuig, G., Cordero, P. J., Aparicio, J. M.** (2007). Egg production and individual genetic diversity in lesser kestrels. *Molecular Ecology* **14**: 3383-3393
- Pritchard, J. K., Stephens, M., Donnelly, P.** (2000). Inference of population structure using multilocus genotype data. *Genetics* **155**: 945-959
- Raymond, M., & Rousset, F.** (1995). An exact test for population differentiation. *Evolution* **49**: 1280-1283
- Reed, D. H., & Frankham, R.** (2003). Correlation between fitness and genetic diversity. *Conservation Biology* **17**: 230-237
- Ruiz-Lopez, M. J., Roldan, E. R. S., Espeso, G., Gomendio, M.** (2009). Pedigrees and microsatellites among endangered ungulates: what do they tell us? *Molecular Ecology* **18**: 1352-1364
- Rushton, S. P, Lurz, P. W.W., Gurnell, J., Fuller, R.** (2000). Modelling the spatial dynamics of parapox disease in red and grey squirrels: a possible cause of the decline in the red squirrel in the UK? *Journal of Applied Ecology* **37**: 997-1012
- Sainsbury, A. W & Gurnell, J. W.** (1995). An investigation into the health and welfare of red squirrels, *Sciurus vulgaris*, involved in re-introduction studies. *Veterinary Record* **137**: 367-370
- Sainsbury, A. W, Nettleton, P., Gilray, J., Gurnell, J. W.** (2000). Grey squirrels have high seroprevalance to a parapox associated with deaths in red squirrels. *Animal Conservation* **3**: 229-233
- Shorten, M.** (1964). *Squirrels*. Collins: London, UK
- Slate, J., & Pemberton, J.** (2006). Does reduced heterozygosity depress sperm quality in wild rabbits (*Oryctolagus cuniculus*)? *Current Biology* **16**: R790-R791
- Szulkin, M., Bierne, N., David, P.** (2010). Heterozygosity-fitness correlations: a time for reappraisal. *Evolution* **64**:1201-1217
- Thomas, K., Tomkins, D. M., Sainsbury, A. W., Wood, A. R., Dalziel, R., Nettleton, P. F., McInnes, C.** (2003). A novel poxvirus lethal to red squirrels. *Journal of General Virology* **84**: 3337-3341
- Tompkins, D. M., Sainsbury, A. W., Nettleson, P., Buxton, D., Gurnell J. W.** (2002). Parapox virus causes a deleterious disease in red squirrels associated with UK declines. *Proceedings of the Royal Society of London B* **269**: 529-533

Tompkins, D. M., White, A. R., Boots, M. (2003). Ecological replacement of native red squirrels by invasive greys driven by disease. *Ecology Letters* **6**: 189-196

Välimäki, K., Hinten, G. N., Hanski, I. (2007). Inbreeding and competitive ability in the common shrew (*Sorex araneus*). *Behavioural Ecology and Sociobiology* **61**: 997-1005

Wall, J. D., Andolfatto, P., Przeworski, M. (2002). Testing models of selection and demography in *Drosophila simulans*. *Genetics* **162**: 203-216

Weir, B. S & Cockerham, C. C. (1984). Estimating *F*-statistics for the analysis of population structure. *Evolution* **38**:1358-1370

Supplementary Information

(a) DNA Extraction and Polymerase Chain Reaction

Red squirrel carcasses were thawed after freezing at -28°C and c. 1cm^3 of tissue removed from the quadriceps. Nuclear DNA was extracted using the Qiagen DNAeasy kit. Nine microsatellite loci were amplified from Hale *et al.* 2001 in three multiplex groups containing three loci each. Primers were diluted with nuclease-free water and mixed in multiplex groups to give final concentrations of 0.2mM. 0.8 μl primer mix, 4 μl multiplex mix and 0.2 μl nuclease-free water was added to 1 μl of each sample in a 6 μl reaction. PCR was carried out under the following conditions: 95°C for 12 mins, 20 cycles at 94°C for 30 s, 52°C for 90 s followed by 74°C for 12 mins and cooling to a final temperature of 4°C . PCR products were diluted 1:50 with nuclease-free water. A sample loading mix was made up of from 1000 μl formamide and 20 μl liz and 9 μl of this solution added to 1 μl of the diluted PCR products. This plate was heated for 3 mins at 94°C to denature the DNA and placed on ice to cool before being placed in an ABI 3100 sequencer.

(b) GIS Populations

We used a Geographical Information Systems (GIS) approach based on landcover data (Büttner *et al.* 2004) at 50m resolution to define the potential extent of red squirrel populations and to give an a-priori estimate of the number of populations. We assumed that squirrels within a population may occupy adjacent, non-contiguous blocks of woodland or suburban (e.g. gardens and parks) habitat provided that the non-forest habitat in between can be crossed by individuals up to a certain distance, termed 'linking distance', as part of their daily movements (see also Bakker & Van Vuren 2004). Based on reported densities for red squirrels in the literature (Magris & Gurnell 2002; Lurz *et al.* 2005; Gurnell *et al.* 2009), we amalgamated deciduous and suburban habitat types and assumed minimum patch sizes of 5 (Verbeylen *et al.* 2003) and 10 ha (Rodriguez and Andrén 1999) for deciduous/urban and conifer habitat patches respectively. Data from the literature suggests that the perceptual range of North American squirrel species comparable to the size range of the European red squirrel is approximately 400 m (Mech & Zollner 2002). Translocation studies with the North American red squirrel (*Tamiasciurus hudsonicus*) and the grey squirrel (*S. carolinensis*) also suggested that a latency to move increased dramatically at distances > 400 m to forest patches (Goheen *et al.* 2003). Furthermore, a review of studies assessing the presence of red squirrels within a landscape matrix reflecting minimum habitat requirements found red squirrel presences up to a distance of < 600 m from a source woodland (Rodriguez and Andrén 1999). We therefore determined potential red squirrel populations on the basis of 2 linking distances of 400 and 600 m and assigned red squirrels to the nearest local population for the purpose of our analysis. At a 400m linking distance all habitat patches within the study area were contiguous. Results suggested eight potential geographic populations: 1. Dumfries &

Galloway, 2. Kielder Forest, 3. Northumberland, 4. Allendale/Weardale, 5. North-East, 6. Central Cumbria, 7. South Cumbria and 8. Lancaster.

(c) *Statistical Tests*

1. *Correlations between measures of inbreeding*

Correlations were carried out between the three measures of inbreeding, Multilocus Heterozygosity (MLH), Internal Relatedness (IR: Amos et al. 2001) and Heterozygosity Weighted by Locus (HL: Aparicio et al. 2006), using a two-tailed Pearson's Correlation (Table S1).

Table S1. Pearson's correlations between MLH, IR and HL

	MLH	IR	HL
MLH		-0.973	-0.986
Sig.		0	0
IR	-0.973		0.956
Sig.	0		0
HL	-0.986	0.956	
Sig.	0	0	

2. *Tests of normality*

Deviations from a normal distribution in MLH were assessed within populations using a one-sample Kolmogorov-Smirnov test (Table S2). MLH departed from normality in all populations bar population 3. Accordingly, non-parametric statistics were performed when using this variable in populations 1 and 2.

Table S2. Test for departures from a normal distribution in MLH within populations.

	Statistic	df	Sig.
Population 1	0.158	48	0.004**
Population 2	0.119	94	0.002**
Population 3	0.118	57	0.06

References

Bakker, V. J. & van Vuren, D. H. (2004). Gap-crossing decisions by the red squirrel, a forest dependant small mammal. *Conservation Biology* **18**: 689-697

Büttner G, Feranec J, Jaffrain G, Mari L, Maucha G, Soukup T. (2004) *The European CORINE land cover 2000 project*. CLC2000 data available at: <http://dataservice.eea.eu.int/>

- Goheen, J. R., Swihart, R. K., Gehring, T. M., Miller, M. S.** (2003). Forces structuring tree squirrel communities in landscapes fragmented by agriculture: species differences in perceptions of forest connectivity and carrying capacity. *OIKOS* **102**: 95-103
- Gurnell, J., Lurz, P.W.W., McDonald, R., Pepper, H.** (2009). *Practical techniques for surveying and monitoring squirrels*, Forestry Commission Technical Note, FCPN011
- Lurz, P.W.W., Gurnell, J. & Magris, L.** (2005). *Sciurus vulgaris*. *Mammalian Species* **769**, 1-10.
- Magris, L. & Gurnell, J.** (2002). Population ecology of the red squirrel (*Sciurus vulgaris*) in a fragmented woodland ecosystem on the island of Jersey, Channel Islands. *Journal of Zoology* **256**, 99-112.
- Mech, S. G. & Zollner, P. A.** (2002). Using body size to predict perceptual range. *OIKOS* **98**: 47-52
- Rodríguez, A. & Andrén, H.** (1999). A comparison of Eurasian red squirrel distribution in different fragmented forest landscapes. *Ecography* **24**: 588-600
- Verbeylen, G., Wauters, L. A., Bruyn, L. D., Matthysen, E.** (2009). Woodland fragmentation affects space use of Eurasian red squirrels. *Acta Oecologica* **35**: 94-103.