RESEARCH ARTICLE

The genetic architecture of helminth-specific immune responses in a wild population of Soay sheep (*Ovis aries*)

Alexandra M. Sparks[1,2]*, **Kathryn Watt**[1], **Rona Sinclair**[1], **Jill G. Pilkington**[1], **Josephine M. Pemberton**[1], **Tom N. McNeilly**[3], **Daniel H. Nussey**[1], **Susan E. Johnston**[1]

1 Institutes of Evolutionary Biology and Immunology and Infection Research, School of Biological Sciences, University of Edinburgh, Edinburgh, United Kingdom, **2** Faculty of Biological Sciences, School of Biology, University of Leeds, Leeds, United Kingdom, **3** Moredun Research Institute, Pentlands Science Park, Bush Loan, Midlothian, United Kingdom

* A.M.Sparks@leeds.ac.uk

Abstract

Much of our knowledge of the drivers of immune variation, and how these responses vary over time, comes from humans, domesticated livestock or laboratory organisms. While the genetic basis of variation in immune responses have been investigated in these systems, there is a poor understanding of how genetic variation influences immunity in natural, untreated populations living in complex environments. Here, we examine the genetic architecture of variation in immune traits in the Soay sheep of St Kilda, an unmanaged population of sheep infected with strongyle gastrointestinal nematodes. We assayed IgA, IgE and IgG antibodies against the prevalent nematode *Teladorsagia circumcincta* in the blood plasma of > 3,000 sheep collected over 26 years. Antibody levels were significantly heritable (h^2 = 0.21 to 0.57) and highly stable over an individual's lifespan. IgA levels were strongly associated with a region on chromosome 24 explaining 21.1% and 24.5% of heritable variation in lambs and adults, respectively. This region was adjacent to two candidate loci, Class II Major Histocompatibility Complex Transactivator (*CIITA*) and C-Type Lectin Domain Containing 16A (*CLEC16A*). Lamb IgA levels were also associated with the immunoglobulin heavy constant loci (*IGH*) complex, and adult IgE levels and lamb IgA and IgG levels were associated with the major histocompatibility complex (MHC). This study provides evidence of high heritability of a complex immunological trait under natural conditions and provides the first evidence from a genome-wide study that large effect genes located outside the MHC region exist for immune traits in the wild.

Author summary

Understanding how immune responses vary in natural populations can give an insight into how infection affects the ability of hosts and parasites to survive and reproduce, and how this drives evolutionary and ecological dynamics. Yet, very little is known about how immune responses vary over an individual's lifetime and how genes contribute to this

variation under natural conditions. Our study investigates the genetic architecture of variation in three antibody types, IgA, IgE and IgG in a wild population of Soay sheep on the St Kilda archipelago in North-West Scotland. Using data collected over 26 years, we show that antibody levels have a heritable basis in lambs and adults and are stable over an individual's lifetime. We also identify several genomic regions with large effects on immune responses. Our study offers the first insights into the genetic control of immunity in a wild population, which is essential to understand how immune profiles vary in challenging natural conditions and how natural selection maintains genetic variation in complex immune traits.

Introduction

Individual differences in immune responses are widely observed in nature, and are likely to be a major factor underlying variation in disease resistance, health and evolutionary fitness in vertebrates [1]. Understanding the role of genes and the environment in generating this variation over the lifetime of an individual is therefore of considerable importance across disciplines including immunology, human and veterinary medicine, and evolutionary biology [2,3]. Yet, efforts to quantify the genetic basis of immunological variation have been largely limited to studies of humans, domesticated livestock and laboratory rodents [4–8]. Experimental studies provide a controlled environment in which to address mechanistic questions about immune function, but offer limited insight into immunological variation under more complex and challenging environmental conditions, as illustrated by dramatic differences in immune phenotypes between wild and laboratory rodents [9,10]. Whilst domesticated livestock and humans are subject to more complex environments, individuals generally receive regular treatment against prevalent infectious agents, which is likely to modify the natural immune response and the contributions of underlying genetic variation. Therefore, widening immunological studies to encompass more natural, untreated vertebrate systems will allow us to obtain an accurate understanding of the role of genetics in driving variation in immune phenotypes more broadly [1,2]. Here, we harness an extensive data set of immune phenotypes and genotypes from an unmanaged and untreated population of Soay sheep to determine the relative contribution of genetic, environmental and individual variation in shaping differences in immunity, and how this varies over the lifetime of individuals.

Studies in humans and livestock have shown that variation in immune traits is often heritable; that is, a significant proportion of phenotypic variance can be attributed to additive genetic effects [3,8,11–15]. Genome-wide association studies (GWAS) in these systems have identified a number of genes of relatively large effect contributing to heritable variation, most notably the major histocompatibility complex (MHC) and cytokine genes [4,5,16–19]. In wild populations, studies have investigated the heritability of immune traits, most often in birds [20–25], with candidate gene approaches further implicating MHC and cytokine regions in cases where significant associations are observed [25–30]. However, these studies often focus on broad, non-specific immune phenotypes such as the phytohaemagglutinin (PHA) response, haematocrit levels and/or parasite burden, rather than specific immune responses to ecologically-relevant parasites [31–33]. In addition, candidate gene studies focus on a small proportion of the genome and may fail to identify previously undiscovered coding or regulatory regions associated with immune trait variation [34,35]. The application of genome-wide association studies of immune variation allows us to identify genomic regions and their relative contribution to immune phenotypes without *a priori* selection of a candidate gene or gene set. This could

identify novel important genes and regulatory regions influencing parasite resistance, and test whether genes identified in human, livestock and laboratory rodent studies also shape immunological variation under natural, untreated conditions. Yet, to our knowledge, there are no genome-wide association studies of specific immune phenotypes in the wild.

In addition to genetic effects driving persistent among-individual differences in immune phenotypes, within-individual variation in immunity associated with recent exposure to parasites, nutritional state, and age are well documented in many systems [36,37]. Immunological studies in humans, livestock and laboratory rodents tend to focus on specific age groups, meaning that little is known about how temporally stable or "repeatable" immune traits are, and whether the genetic architecture underlying these traits remains consistent across age groups. To separate among- and within-individual contributions to phenotypic variation, longitudinal data across the entire lifetime is required. A growing number of longitudinal studies in humans have found that high immunological diversity is maintained by high inter-individual variation, but that immune profiles of individuals are stable across longitudinal sampling [4,38–40]. Studies in cattle found that cellular based immune traits are highly repeatable, while antibody-based traits were less so [15,41]. Therefore, the ability to dissect the relative contributions of among- and within-individual variation in immune phenotype using longitudinal data collected in natural systems is a crucial step towards our understanding of the evolutionary and ecological causes and consequences of observed variation in immunity in the wild.

Domestic sheep (*Ovis aries*) and their gastrointestinal strongyle nematodes represent a well-understood host-parasite system, due to their agricultural and economic importance, with much recent interest in determining the genes underlying host resistance to these parasites [42]. Of the strongyle parasites, *Teladorsagia circumcincta* is of major economic importance for domestic sheep in temperate regions [43] and has a simple direct life-cycle, with an infective L3 stage which develops to L4 stage within the gastric glands before emerging as sexually mature adult parasites which reside in the abomasum. Defence against *T. circumcincta* in lambs is associated with parasite-specific IgA antibody responses directed at worm growth and subsequent female fecundity [12,44], while in older animals a hypersensitive response, involving IgE antibodies, results in expulsion of incoming larvae from the mucosa [43]. Anti-*T. circumcincta* IgA levels are moderately heritable in lambs and adults [12,45,46]. Candidate gene and genome-wide studies have identified regions associated with faecal egg counts (FEC) or protective immunological traits related to gastrointestinal nematodes, with candidate gene studies primarily focussed on interferon gamma (*IFNγ*) and the MHC [42]. However, due to the focus on identifying individuals for selective breeding and the greater impact of parasite infections in lambs, most studies focus only on lambs, with only a few studies of adult ewes [6,46–48]. As a consequence, we know relatively little about age-dependent genetic effects; indeed, differences in resistance loci between lambs and adults suggest that the genetic control of these mechanisms may differ [6]. Furthermore, domestic sheep populations are almost always regularly treated with anthelmintic drugs, and resistance to all available major drug classes has been documented [49,50]. Understanding the genetic factors regulating host resistance in the absence of treatment (i.e. natural conditions) could facilitate artificial selection programmes that promote resistance and productivity without reliance on drugs [42,51,52].

The long-term individual-based study of the wild Soay sheep of St Kilda provides a powerful opportunity to understand the genetic architecture of immune traits at different ages under natural conditions. Soay sheep are infected with several gastrointestinal strongyle nematodes common to domestic sheep, predominantly *Teladorsagia circumcincta*, *Trichostrongylus axei* and *Trichostrongylus vitrinus* [53,54]. Strongyle nematode burden, in combination with harsh winter weather and low food availability, is a strong selective force on the sheep [53–56]. Parasite-specific antibody responses are moderately heritable [28,57] and parasite-specific IgA

levels and parasite-specific pan-isotype antibody levels have been shown to be negatively associated with strongyle faecal egg count [26,57]. Recent examination of anti-nematode antibody isotypes (namely IgA, IgE and IgG) over a 26-year period showed that levels of IgG are positively associated with adult survival, and negative associations are observed between antibodies and FEC for all isotypes in lambs, but only for IgG in adults [58–60].

Previous examination of the genetic architecture of immune traits in Soay sheep using QTL mapping and candidate gene approaches failed to identify loci associated with parasite egg counts and pan-isotype antibody levels [28,61]; however, a microsatellite polymorphism at the IFNγ locus in lambs had previously been associated with reduced faecal egg counts and increased parasite-specific IgA levels [26]. Today, the majority of study individuals have been genotyped on the Illumina 50K OvineSNP50 BeadChip, and genome-wide association studies have identified genomic regions associated with traits such as horn morphology, body size and recombination rate [62–64]. Here, we investigate the heritability and conduct genome-wide association studies of anti-*T. circumcincta* IgA, IgE and IgG levels from > 5,800 plasma samples collected from > 3,000 Soay sheep over a 26-year period. We show that antibody levels are heritable and temporally stable over an individual's lifetime, and that several genomic regions explain heritable variation in both lambs and adults.

Results

Phenotypic variation and heritability

Lambs in their first year of life had considerably lower antibody levels compared to older animals, so we investigated these two groups separately in our analyses (Figs 1 & S1). August *T. circumcincta*-specific antibody levels of IgA, IgG and IgE were weakly positively correlated with each other, with slightly stronger correlations in lambs (adjusted R^2 values from 0.078 to 0.175 in lambs, and from 0.005 to 0.012 in adults, $P < 0.001$; S2 Fig, S1 Table). Males had lower IgA levels as lambs and lower IgG levels as lambs and adults compared to females (Wald test $P < 0.001$, Figs 1 & S3, S2 Table). All three antibody isotypes were positively associated with age in days in lambs and age in years in adults, except for adult IgG levels, which were negatively associated with age (Wald test $P < 0.001$, Fig 1, S2 Table). Each antibody isotype was temporally stable or 'repeatable' over the lifespan of adults, as shown by high between-individual variation attributed to additive genetic effects and permanent environmental differences (proportion of phenotypic variance: IgA = 0.76, IgE = 0.72, IgG = 0.52; Fig 2, Table 1). This was further illustrated by a strong positive correlation between antibody measures taken in two consecutive years (IgA and IgE: slopes > 0.8, Adjusted R^2 > 0.66; IgG: slope = 0.52, Adjusted R^2 = 0.29; Fig 3, S4 Table).

All antibody measures were heritable in lambs and adults, with IgA levels showing the highest heritabilities (h^2 = 0.39 & 0.57 for lambs and adults, respectively; Tables 1 & S3, Fig 2). Heritabilities in lambs and adults were 0.21 and 0.47 for IgE, and 0.29 and 0.23 for IgG, respectively (Tables 1 & S3, Fig 2). Permanent environment effects were also significant, explaining between 18.6 and 28.5% of the phenotypic variance in adults, resulting in repeatabilities of 0.76, 0.72 and 0.52 for IgA, IgE and IgG, respectively. There was significant variation in antibody levels among birth years in lamb IgA and IgG measures, although the effect was small ($\leq$ 7% of the phenotypic variance) and there was a weakly significant maternal effect explaining < 4% of variation in lamb IgA levels (Fig 2, Tables 1 & S3). In adults, capture year explained < 1.5% of the phenotypic variance in all antibody measures. The full results of the animal models are provided in S2 Table (fixed effect structures) and S3 Table (random effect structures). Genetic correlations among antibody isotypes in lambs were strongly positive ($r_A \geq 0.695$; Fig 4), but much weaker in adults ($r_A \leq 0.232$; Fig 4). Genetic correlations between
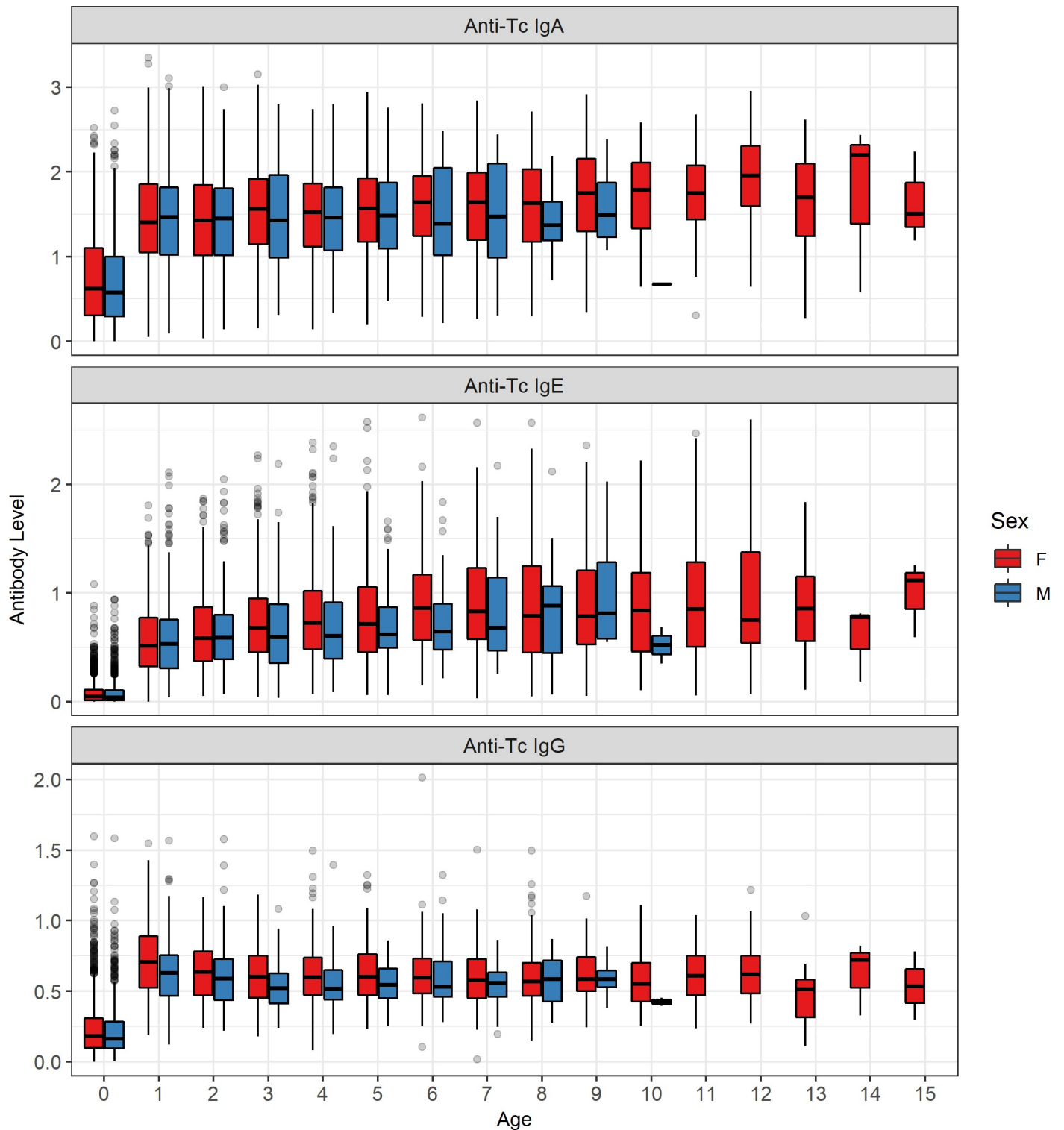
Fig 1. Distribution of anti-*Teladorsagia circumcincta* IgA, IgE and IgG levels with age and sex in Soay sheep.
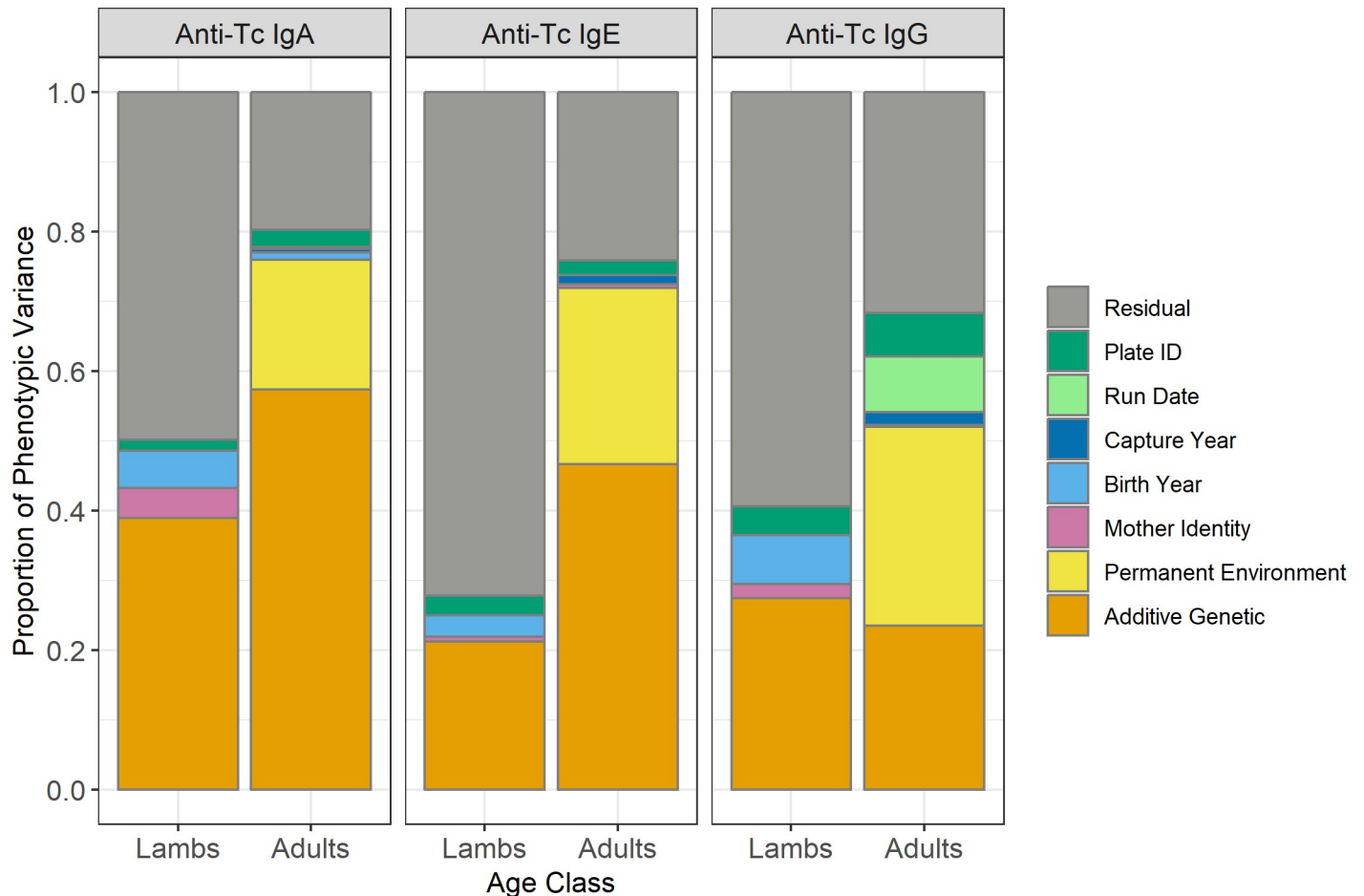
Fig 2. Proportion of phenotypic variance explained by random effects in animal models of anti-*T. circumcincta* IgA, IgE and IgG levels in lamb and adult Soay sheep. Data is provided in Table 1.

lambs and adults when considering the same antibody isotype were positive and ranged from 0.33–0.61 (Fig 4), whilst adult-adult and lamb-adult genetic correlations across different isotypes were generally weaker and often not significantly different from zero (Fig 4).

Genome-wide association studies

Genome-wide association studies (N_{SNPs} = 39,176) and regional imputation approaches identified several genomic regions associated with variation in anti-*T. circumcincta* IgA, IgE and IgG levels in lambs and adults (Figs 5 and S5, Tables 2, S5 & S6). SNP genotypes were fit as a two or three level factor (AA, AB or BB), rather than as an additive effect (0, 1 or 2). All test statistics were corrected using the genomic control parameter λ; this value was low for all 6 GWAS (λ < 1.074), indicating that population structure was adequately captured by fitting pedigree relatedness. Below, we discuss associations for each antibody separately, with summary information in Table 2. Full association results for genome-wide and imputed SNPs are provided in S5 and S6 Tables, respectively. Information on genes and orthologues within associated regions are provided in S7 Table and immune GO terms associated with these genes are provided in S8 Table. For the most highly associated SNPs (S6 Fig, Tables 2 & S9), we also examined sex by SNP interactions (S7 Fig & S10 Table). Whilst these interactions were

Table 1. Mean and variance estimates, and the proportion of variance explained for anti-*T. circumcincta* IgA, IgE and IgG levels measured in St. Kilda Soay sheep lambs and adults. Mean and V_{OBS} are the mean and variance of the raw data measures, N is the number of measures in N_{IDS} unique individuals. V_P is the phenotypic variance as a sum of all variance components as estimated by an animal model. The additive genetic effect (h^2) indicates the narrow sense heritability of the trait. Non-significant estimates are indicated in grey text. Full results of all variance components are provided in S3 Table. Figures in parentheses are standard errors. Effects are plotted in Fig 2.

Trait	Age	V_{OBS}	Mean	N	N_{IDS}	V_P	Proportion of V_P explained							
							Additive Genetic (h^2)	Permanent Environment	Birth Year	Capture Year	Mother Identity	Plate ID	Run Date	Residual
Anti-Tc IgA	Lambs	0.2529	0.741	2030	2030	0.2483	0.3890	NA	0.0529	NA	0.0436	0.0157	0.0000	0.4989
						(0.0099)	(0.0372)	NA	(0.0202)	NA	(0.0188)	(0.0098)	(0.0000)	(0.0366)
	Adults	0.3051	1.507	3793	1321	0.3048	0.5732	0.1863	0.0102	0.0052	0.0000	0.0242	0.0032	0.1977
						(0.0134)	(0.0363)	(0.0303)	(0.0068)	(0.0035)	(0.0000)	(0.0075)	(0.0068)	(0.0101)
Anti-Tc IgE	Lambs	0.0138	0.086	2035	2035	0.0135	0.2122	NA	0.0305	NA	0.0067	0.0288	0.0000	0.7219
						(0.0005)	(0.0334)	NA	(0.0153)	NA	(0.0174)	(0.0132)	(0.0000)	(0.0360)
	Adults	0.1835	0.733	3798	1321	0.1739	0.4662	0.2531	0.0000	0.0134	0.0047	0.0208	0.0000	0.2418
						(0.0071)	(0.0385)	(0.0368)	(0.0000)	(0.0057)	(0.0182)	(0.0055)	(0.0000)	(0.0117)
Anti-Tc IgG	Lambs	0.0364	0.236	2032	2032	0.0354	0.2739	NA	0.0703	NA	0.0203	0.0411	0.0000	0.5944
						(0.0015)	(0.0344)	NA	(0.0266)	NA	(0.0184)	(0.0161)	(0.0000)	(0.0381)
	Adults	0.0462	0.630	3776	1319	0.0468	0.2347	0.2854	0.0027	0.0180	0.0000	0.0618	0.0803	0.3172
						(0.0021)	(0.0330)	(0.0296)	(0.0048)	(0.0101)	(0.0000)	(0.0172)	(0.0300)	(0.0159)

https://doi.org/10.1371/journal.pgen.1008461.t001

significant for nearly all of these loci, the direction of these effects were similar unless otherwise stated below.

Anti-*T. circumcincta* IgA. There was a strong association between IgA levels and a region between 6.89 and 14.95 Mb on sheep chromosome 24, with the highest association observed at the SNP locus OAR24_12006191.1 in both lambs and adults (Wald test P = 1.01 x10^{-31} and 2.23x10^{-51} in lambs and adults, respectively; Figs 5, 6, S8 & S9; Tables 2 & S5). This SNP had an approximately additive effect on IgA levels in both lambs and adults (Table 2), with the region explaining 20.0% and 27.2% of the additive genetic variance in lambs and adults, respectively, equating to 7.8% and 15.3% of the phenotypic variance in lambs and adults, respectively. Associations at imputed SNPs in this region showed the strongest association at SNPs between 10.62Mb and 10.86Mb (maximum Wald test P = 4.08 x10^{-39} and 5.74x10^{-71} in lambs and adults, respectively; Fig 6, S6 Table), again with an additive effect on IgA levels (Table 2, S6 Fig). This region corresponded to a novel gene (ENSOARG00000007156) orthologous to the protein coding gene Sorting Nexin 29 (*SNX29*; Fig 6, S7 Table); GO terms indicated that this locus is associated with red blood cell phenotypes in humans and mice, including variation in haematocrit, erythrocyte cell number and circulating alkaline phosphate levels (International Mouse Phenotyping Consortium data; S8 Table). Whilst this gene has no clear role in driving IgA levels, the associated SNPs were downstream of two candidate genes (Fig 6; S7 & S8 Tables; distances of ~1.021Mb and ~709Kb, respectively): the Class II Major Histocompatibility Complex Transactivator (*CIITA*), which is described as a "master control factor" for gene expression at the major histocompatibility complex [65,66]; and C-Type Lectin Domain Containing 16A (*CLEC16A*), variants at which have been associated with common variable immunodeficiency disorder and IgA deficiency [67–69]. An unmapped SNP was significantly associated with IgA levels in both lambs and adults (Fig 5, chromosome '0'); this locus was originally mapped to the same chromosome 24 region in version 2.0 of the sheep genome.
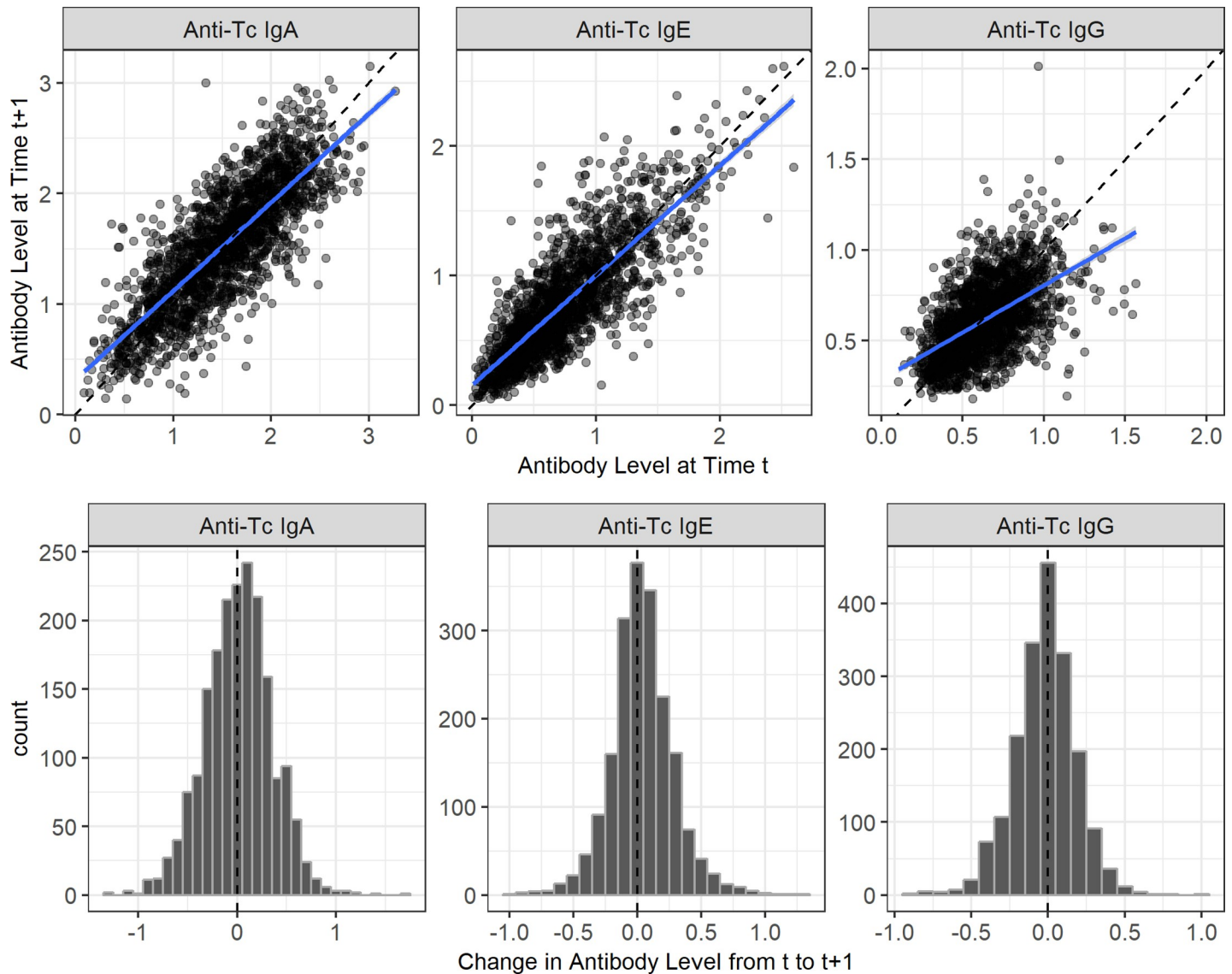
Fig 3. Temporal correlations in anti-*Teladorsagia circumcincta* IgA, IgE and IgG levels in adult Soay sheep. Scatterplots of all raw data in adults for which there are two antibody measures in two consecutive years with a dashed line indicating a perfect 1:1 relationship and the solid line indicating the regression slope. Histograms show the frequency of the change in antibody levels for adults in consecutive years with a dashed line indicating no change.

https://doi.org/10.1371/journal.pgen.1008461.g003

Lamb IgA levels showed a further association at two more regions. A single imputed SNP on chromosome 20 had an approximately additive effect on IgA levels and was located ~ 157kb directly downstream of an orthologue of the MHC II locus *HLA-DQA1* (oar3_ OAR20_25196550, Wald test P = 1.96E-06; S6 & S10 Figs; Tables 2 & S6). A second SNP at the distal end of chromosome 18 had an approximately additive effect on IgA levels (s03219.1, Wald test P = 4.34 x10^{-07}; Figs 5, S6 & S11, Tables 2 & S5), and was located ~311kb to 454kb downstream of four novel genes (ENSOARG00000008862, ENSOARG00000008994, ENSOARG00000009143, ENSOARG00000009269) orthologous to various forms of immuno-globulin heavy constant alpha, epsilon, gamma and delta loci in humans (*IGHA*, *IGHE*, *IGHG* and *IGHD*, respectively; S7 & S8 Tables). These loci code for constituent proteins of immuno-globulins and have GO terms associated with variation in IgA, IgE and IgG levels in mice (S8
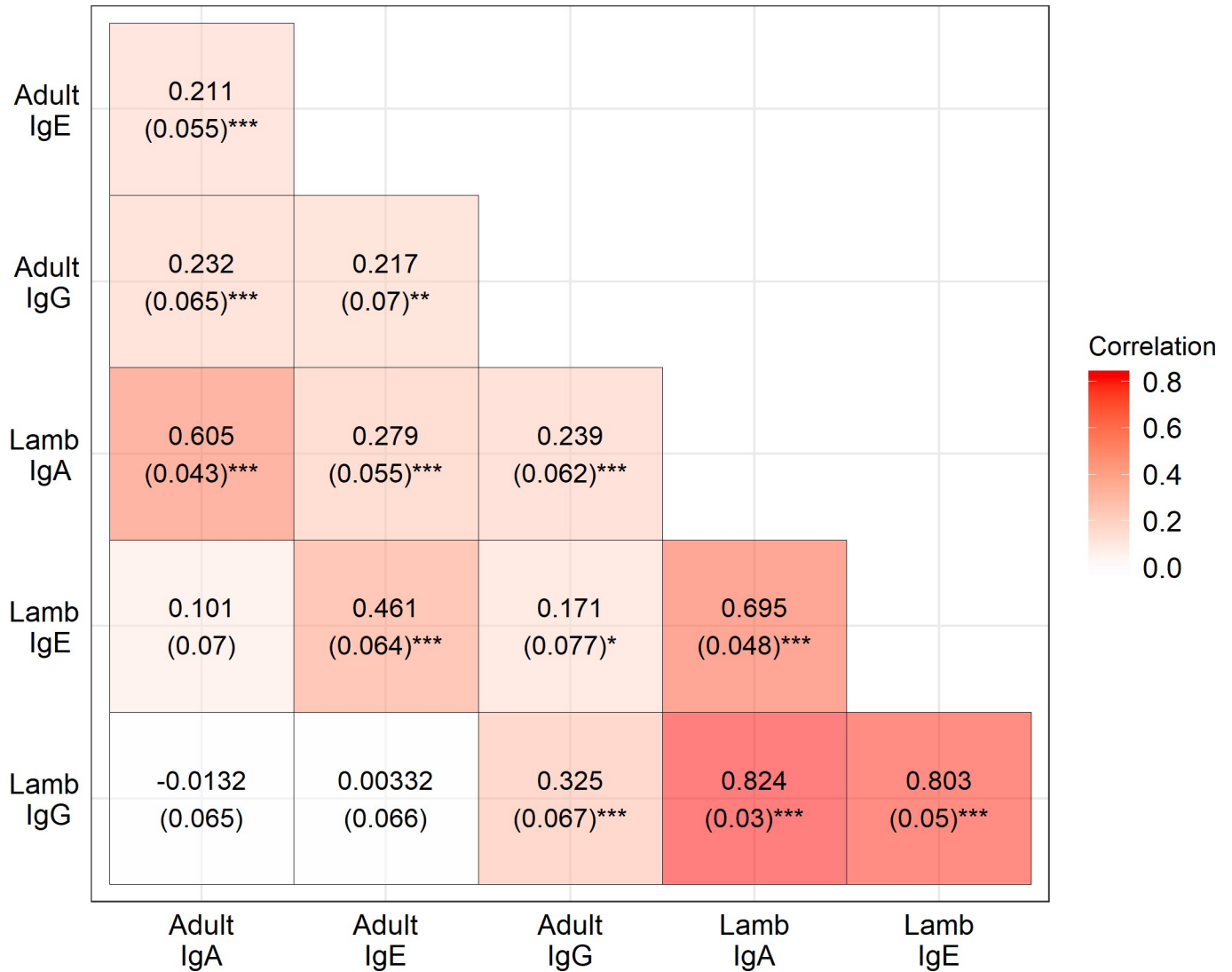
Fig 4. Genetic correlations between anti-*T. circumcincta* IgA, IgE and IgG levels in lamb and adult Soay sheep based on genomic relatedness. Numbers in parentheses are the standard error of the correlation estimate. Significantly different from zero: * P < 0.05, ** P < 0.01, *** P < 0.001.

Table). When the threshold was relaxed to include loci with < 0.95 imputation success, the most highly associated SNPs occurred at the immunoglobulin heavy chain complex, closest to *IGHA* and *IGHE* (highest associated SNP = oar3_OAR18_68320039, Wald test P = 1.82 x10^{-10}, imputation success = 0.754; S6 Table, S11 Fig).

Anti-*T. circumcincta* IgE. Lamb IgE levels were associated with a gene-poor region of chromosome 10, with the highest association observed at the imputed locus oar3_OAR10_10333145 (Figs 5 & S12, Table 2). The only protein-coding gene in this region, olfactomedin 4 (*OLFM4*), is associated with down-regulation of immune responses against bacterial infections in mice [70]. However, given the low minor allele frequency of this locus (S6 Fig), a lack of other associations at adjacent loci (S12 Fig, S5 & S6 Tables) and no contribution of the region to additive genetic variance (Table 2), we cannot rule out that association seen here is spurious and due to the sampling of rare alleles in individuals with extreme trait values.
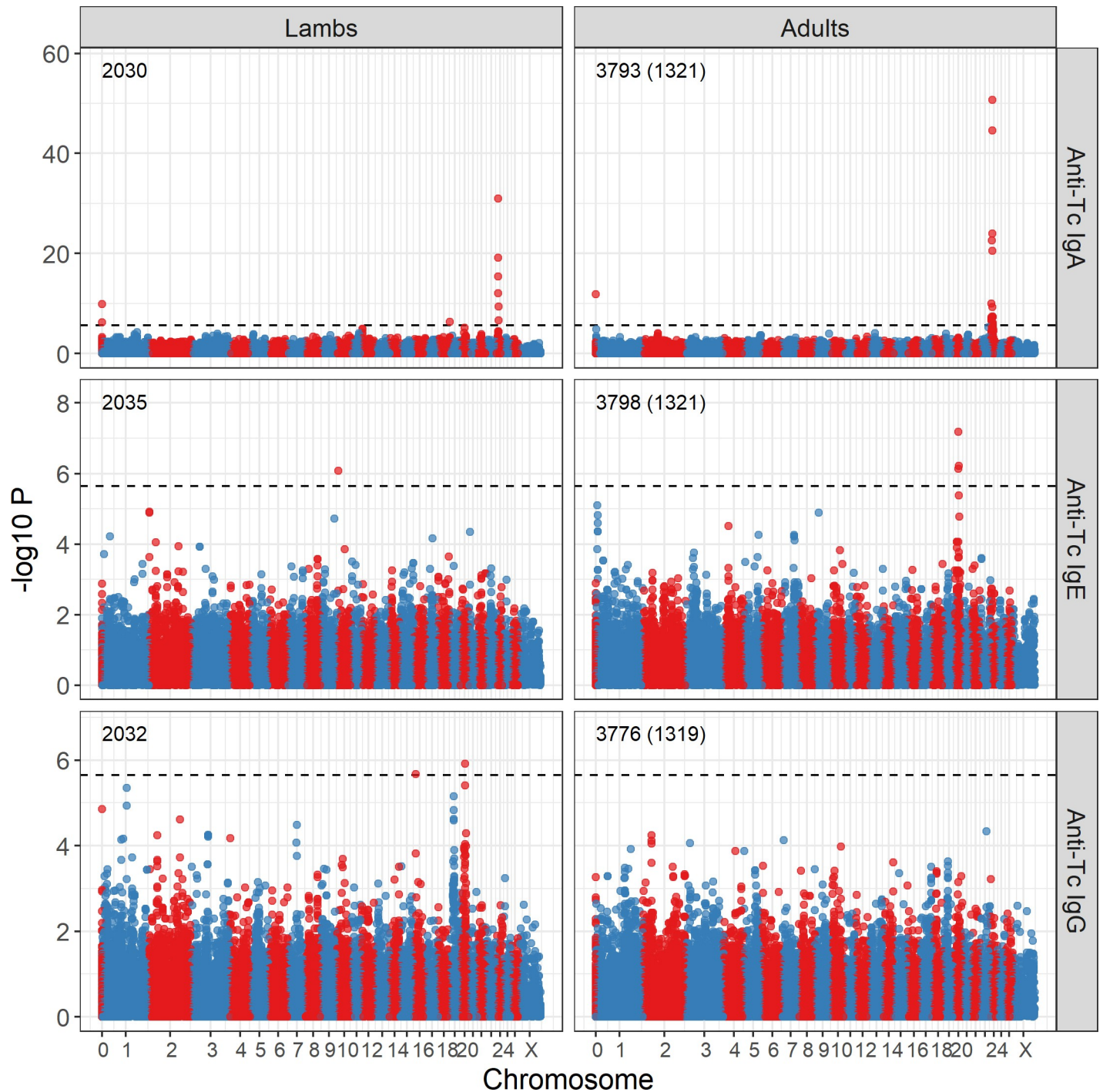
Fig 5. Genome-wide association of anti-*Teladorsagia circumcincta* IgA, IgE and IgG levels in lamb and adult Soay sheep with SNPs on the Ovine SNP50 BeadChip. Numbers indicate the number of measures and the number of unique individuals in parentheses. The dotted line indicates the genome-wide significance threshold equivalent to an experiment-wide threshold of P = 0.05. Points are colour-coded by chromosome. Positions are given relative to the sheep genome assembly Oar_v3.1. Underlying data, sample sizes and effect sizes are provided in S5 Table. P-values were corrected with genomic control λ, and comparisons with those expected under a null distribution (i.e. P-P plots) are provided in S5 Fig.

https://doi.org/10.1371/journal.pgen.1008461.g005

Adult IgE levels were associated with a region from 25.8Mb to 27.5Mb on chromosome 20, with the highest association seen at the locus OAR20_27259292.1 for both the SNP50 and

Table 2. SNPs showing the strongest association with anti-*T. circumcincta* IgA, IgE and IgG levels in lambs and adults. The P-values provided in this table have not been corrected using genomic control to allow comparisons between directly genotyped and imputed SNPs. Asterisks next to the SNP name indicate that the most highly associated SNP was imputed from the high-density SNP chip. N 50K and N HD indicate how many SNPs were significantly associated with the trait in the same region for the 50K and HD SNP chips, respectively. A and B indicate the reference and alternate alleles at each SNP. MAF indicates the minor allele frequency (allele B); for imputed SNPs, this was calculated using the HD chip data only and not imputed genotypes. Genotype effects AA, AB and BB are the effect sizes as calculated from the associated animal model. Full results including corrected P values are provided in Fig 5 and S5 and S6 Tables; gene and GO information is provided in S7 & S8 Tables. Lamb IgE associations are given for the $\log_{10}$ of the antibody measures (see Methods).

Trait	Age	Chr	Position	Highest Associated SNP	N 50K	N HD	P	A	B	MAF	Effect AA	Effect AB	Effect BB	Prop V_A Explained	Closest Gene	Candidate Genes in Region
Anti-Tc IgA	Lambs	18	68137231	s03219.1	1	33	$1.47e^{-07}$	A	G	0.328	0.000	0.091	0.202	0.101	*CDCA4*	*IGH* complex
		20	25196550	oar3_OAR20_25196550*	0	1	$1.96e^{-06}$	A	G	0.490	0.000	-0.083	-0.175	0.136	*ELOVL5*	MHC II locus
		24	10616039	oar3_OAR24_10616039*	6	118	$4.08e^{-39}$	A	G	0.484	0.000	-0.192	-0.424	0.200	*GSPT1*	*CIITA, CLEC16A*
	Adults	24	10858856	oar3_OAR24_10858856*	25	383	$5.74e^{-71}$	A	G	0.472	0.000	-0.383	-0.718	0.272	*SNX29*	*CIITA, CLEC16A*
Anti-Tc IgE	Lambs	10	10333145	oar3_OAR10_10333145*	1	2	$2.91e^{-07}$	G	A	0.023	2.929	2.819	0.000	0.000	*OLFM4*	*OLFM4*
	Adults	20	25781566	OAR20_27259292.1*	3	25	$5.09e^{-09}$	A	G	0.386	0.000	-0.061	-0.220	0.080	*HLA-DRA*	MHC II locus
Anti-Tc IgG	Lambs	16	12632988	oar3_OAR16_12632988*	1	31	$5.20e^{-07}$	A	G	0.036	0.000	0.026	0.529	0.020	*MAST4*	*CD180*
		20	30876754	oar3_OAR20_30876754*	1	6	$2.44e^{-07}$	G	A	0.211	-0.104	-0.072	0.000	0.077	*TRIM38*	MHC I/II

https://doi.org/10.1371/journal.pgen.1008461.t002

imputed SNP loci (Figs 5 & S13, Tables 2, S5 and S6). This SNP is directly upstream of the major histocompatibility complex (MHC) class II locus *HLA-DRA*, as well as the MHC class II loci *DQA* and *HLA-DQB1*; the wider region contains ~46 annotated genes with GO terms associated with immune function (S13 Fig, S7 & S8 Tables).
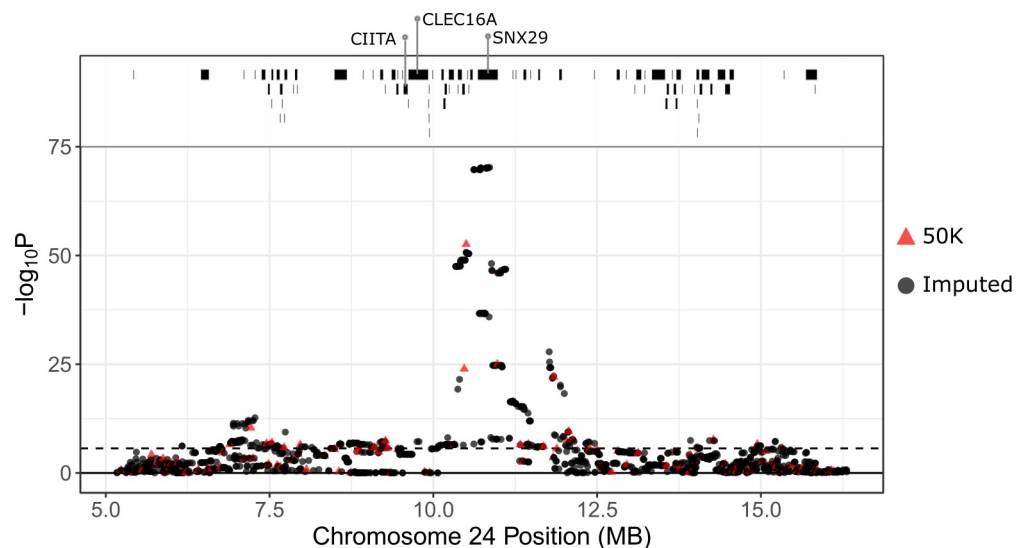
Fig 6. Local association of anti-*Teladorsagia circumcincta* IgA levels in adult Soay sheep with SNP50 and imputed SNP loci at the most highly associated region on chromosome 24. The dotted line indicates the genome-wide significance threshold equivalent to an experiment-wide threshold of P = 0.05. Points are colour-coded by their imputation status. Positions are given relative to the sheep genome assembly Oar_v3.1. Underlying data, sample sizes and effect sizes are provided in S6 Table. Gene positions were obtained from Ensembl (gene build ID Oar_v3.1.94) and are provided in S7 Table.

https://doi.org/10.1371/journal.pgen.1008461.g006

Anti-*T. circumcincta* IgG. Lamb IgG levels were significantly associated with a region from 29.6–30.9Mb on chromosome 20, with the highest association observed at the imputed locus oar3_OAR20_30876754 (Figs 5 & S14, Table 2). This region was ~4Mb from the region associated with IgE levels in adults and was close to protein coding regions orthologous to MHC Class I genes (S14 Fig, S7 & S8 Tables). This locus showed a sex by genotype interaction, with the G allele dominant to A for lower IgG levels in males, compared to an additive effect in females; AA males showed higher IgG levels than in females (S7 Fig, S10 Table). A further association was observed on chromosome 16, corresponding to a region containing *CD180*, a gene associated with variation in IgG2b levels in mice [71] (S8 Table, S15 Fig), although the minor allele frequency of the associated SNP is low (MAF = 0.036) and the association may again be partly driven by sampling effects (Tables 2 & S9 and S6 Fig). There was no association between adult IgG levels and the SNPs genotyped in this study (Fig 5).

Discussion

This study is the first to examine the genetic architecture of immune traits using a genome-wide association approach in a wild population. We have shown that anti-*Teladorsagia circumcincta* IgA, IgE and IgG levels in Soay sheep are highly repeatable within individuals' lifetimes and show substantial heritable variation that is underpinned by several genomic regions containing immune-associated genes. This suggests that antibody phenotypes have the potential to respond rapidly to selection, but also demonstrates that individual sheep develop distinct, temporally stable antibody phenotypes despite marked annual variation in exposure to nematode parasites, food availability and climate conditions [53,72,73]. Below, we discuss the genetic architecture of these traits in more detail and how our findings inform the broader field of understanding the evolution and adaptive potential of immune traits in both domestic and natural populations.

Temporal stability of antibody levels

We observed a large increase in anti-*Teladorsagia circumcincta* antibody levels between lambs (aged 4 months) and adults (aged >16 months). This was consistent with previous observations in this system and is probably due to the development of anti-helminth immunity with exposure over early life [26]. In adults, antibody levels were stable within individuals, as indicated by high repeatabilities and strong temporal correlations of antibody measures between years (Fig 3, S4 Table). This low intra-individual variation is notable given the temporally and spatially variable environment that individuals experience on St Kilda. The relatively small amount of variation explained by cohort, maternal and annual effects found here suggests that temporal variation in exposure to parasites, condition or early life effects had relatively little influence on antibody levels. It is also notable that repeatabilities for each antibody isotype were high despite different isotypes being only weakly correlated with one another, suggesting complex individualised immune phenotypes which are consistent over lifetimes. Our findings are consistent with the consensus emerging from human studies, which have also determined that variation in immune parameters is driven by high inter-individual and low intra-individual variation, indicative of stable immunological profiles of individuals [4,38–40]. Whilst most intra-individual variation in this study was attributed to additive genetic effects, the permanent environment effects were substantial, accounting for 19%, 25% and 29% of the phenotypic variance in IgA, IgE and IgG, respectively. At present, the factors contributing to this variation remain unknown, but may be driven by consistent spatial differences in exposure or individual disease history, or due to complex interactions between nutritional state, exposure to other parasites and life history during early life.

Heritable variation in antibody levels

Anti-*T. circumcincta* IgA, IgE and IgG levels were highly heritable in Soay sheep, ranging from 0.21 to 0.39 in lambs and from 0.23 to 0.57 in adults. These estimates are comparable to previous work estimating the pedigree heritability of an anti-*T. circumcincta* pan-isotype antibody measure (likely to be mainly comprised of IgG) in Soay sheep lambs (h^2 = 0.30) and adults (h^2 = 0.13–0.39) [28,57]. In domestic sheep, similar heritability estimates have been obtained for anti-*T. circumcincta* IgE in Texel lambs (h^2 = 0.39 and 0.50 against the third and fourth stage larvae, respectively [11]) and anti-*T. circumcincta* IgA in Scottish Blackface lambs (h^2 = 0.56 against fourth stage larvae [12]). The observation that immune traits in Soay sheep and domestic breeds appear to have substantial heritable variation is interesting from an evolutionary perspective, as selection for reduced parasite load is likely to be strong, which in turn is predicted to reduce underlying genetic variation and hence the heritability of quantitative traits [74]. In domestic sheep, anti-helminthic treatments may have relaxed the selection pressure on immune traits. Alternatively, the observed high heritabilities in both domestic and wild sheep may be in accordance with theory predicting that stabilising selection, rather than directional selection, is likely to be acting on immune traits [75], which in turn may lead to the maintenance of genetic variation at the underlying trait loci. In the Soay sheep, we have shown with the same dataset that there is little evidence for stabilising selection, with directional selection present for IgG in adults but not for other isotypes or age groups [60]. It is notable that adult IgG, as well as being under the strongest directional selection, also has the lowest heritability compared to other isotypes and age groups (Fig 2, Table 1). This is consistent with the prediction that directional selection should erode heritable variation, whilst the high observed heritabilities in general are consistent with observations of weak or variable selection on these antibody measures [60]. Nevertheless, a full understanding of the mechanisms maintaining this genetic variation will require examination of association between genotypes at significant loci with individual fitness, i.e. survival and reproductive success.

Genetic correlations among ages and antibody isotypes

Our study provides, to our knowledge, the first estimates of genetic correlations among immune phenotypes from any wild population. Our results illustrate the complexity of the genetic architecture underpinning immunological variation: even in a set of immune measures all associated with the same immune challenge (i.e. parasitic helminth antigens) and produced by the same arm of the immune system, genetic correlations among measures can still be weak. Our findings also highlight the importance of age-dependence in the genetic control of immune phenotypes, mirroring our recent finding that selection on these three isotypes is highly age-dependent in this study system [60]. Whilst the inter-isotype genetic correlations were highly positive in lambs ($r_A \geq 0.695$), they were much smaller in adults ($r_A \leq 0.232$). This is consistent with previously documented patterns at the phenotypic level [60]. Lambs measured in their first year have relatively naïve immune systems and have only very recently become exposed to parasitic helminths, whilst on St Kilda levels of exposure to infective larvae vary considerably over time and space [53]. The breakdown in adults of the tight genetic correlations among antibody isotypes observed in lambs and the general pattern of low correlations across age groups (Fig 4) may well reflect complex interactions between genotypes and the unique pattern of exposure and resource availability experienced by different individuals as they develop and mature. The notable exception here, the relatively high among-age genetic correlation observed for IgA antibodies ($r = 0.61$) presumably reflects the influence at both ages of the region identified in our GWAS on chromosome 24. From an evolutionary perspective, these results demonstrate that natural selection on different antibody isotypes against

parasitic helminths is relatively weakly constrained by genetic correlations in adults at least. The genetic architecture underpinning antibody levels also does not appear to be identical in lambs and adults, and future studies in wild and domestic animals, as well as humans, should note the possibility that immunological variation may be under the control of different genes and regions depending on age or life stage.

Genetic variants associated with antibody levels

In this study, we found associations between antibody levels and known immune loci, most notably at the MHC (lamb IgA and IgG, adult IgE) and IGH complex (lamb IgA). However, the strongest association observed in this study was between lamb and adult IgA levels and a region on chromosome 24 corresponding to the gene *SNX29*. This gene has no previous association with immune trait variation (see Results) but occurs downstream of two candidate genes. The first, *CIITA*, is a master regulator of MHC class II gene expression; overexpression of *CIITA* in rats can induce transcription of MHC Class II genes in nearly all cell types [76] and *CIITA* knockout mice show impaired MHC Class II expression [77]. Mutations in *CIITA* in humans are associated with bare lymphocyte syndrome type II, a severe primary immunodeficiency caused by the absence of MHC class II gene expression [78]. In addition, a human GWAS study showed an association with variants at *CIITA* and levels of activated T cells (i.e., HLA DR+ T lymphocytes) and is in linkage disequilibrium with disease variants associated with ulcerative colitis [4]. The second candidate, *CLEC16A*, is almost directly adjacent to *CIITA* and has been associated with IgA deficiency and common variable immunodeficiency disorder characterised by inadequate levels of multiple antibody isotypes [67–69] and *CLEC16A* knockdown mice have a reduced number of B cells and increased IgM levels compared with controls [68]. In addition to antibody variation, polymorphisms at *CLEC16A* and *CIITA* have been associated with numerous autoimmune diseases, including multiple sclerosis [79,80], diabetes [81], Crohn's disease [82], adrenal insufficiency [83] and arthritis [84], and *CLEC16A* has been functionally linked to auto-inflammation and autoimmunity in mice [85]. Previous work on Soay sheep found that auto-reactive antibody levels are variable and moderately heritable in this system and, despite their clear links with autoimmune diseases in humans, high levels were predictive of increased chances of over-winter survival [58,86]. However, we found only weak correlations between anti-*T. circumcincta* antibody levels and these auto-reactive antibodies [58], suggesting they may independently reflect very different components of the humoral immune system. The question of how natural selection in the wild acts on immune measures, and on the genes found in humans to be associated with autoimmune pathology, remains a fascinating and open topic. Further investigation of the function of these genes in natural systems, such as the Soay sheep, may help address this.

Despite *CIITA* and *CLEC16A* being strong candidate genes for IgA expression *a priori*, they lie ~1 & 0.7 Mb upstream from the GWAS peak, respectively (Fig 6). We cannot rule out that variants in protein-coding regions at *SNX29* and adjacent loci may drive IgA expression. However, a more plausible hypothesis is that the associated region contains cis-regulatory elements affecting the expression of *CIITA* and/or *CLEC16A*. Direct evidence of the precise cis-regulatory regions driving gene expression is scarce, but there is increasing evidence that genes can have multiple cis-regulatory regions driving expression [87], and that cis-regulatory regions can occur at distances of >1Mb from their target genes (see [88] and references therein).

The non-MHC variants identified in this study have not previously been associated with anti-*T. circumcincta* IgA, IgE or IgG levels in other sheep breeds investigated to date. A genome-wide association study in Scottish Blackface lambs failed to identify any SNPs associated with *T. circumcincta* IgA [45], while a study in Spanish Churra ewes found one genome-

wide significant SNP on chromosome 12 [6]. A quantitative trait locus (QTL) mapping study in Romney lambs found total IgE and anti-*Trichostrongylus colubriformis* IgG levels were each associated with a region on chromosome 23 [89]. Together with our results, it appears that QTL for parasite-specific antibody traits have not been consistently observed between sheep breeds. This may be due to different loci associated with immune responses at different ages, differences in host-parasite exposure, inherent differences between breeds driven by different selective breeding histories, and/or genetic drift [6,28,90]. Alternatively, there may be differences in the power to detect trait loci due to differences in patterns of linkage disequilibrium, effect sizes, sample sizes and/or analytical approaches between the studies. The loci identified in the current study may also be due to a genotype-by-environment effect that may only be manifested under natural conditions or could have been introduced with a historical admixture event with the Dunface breed [91]. Investigation of candidate causal mutations in the current study will shed light on the mechanisms driving antibody levels within Soay sheep, as well as their ubiquity and origin across different sheep breeds.

The identification of several large effect loci is in contrast with GWAS studies on body size and fitness-related traits in wild populations which have found few, if any, associations of SNPs with quantitative traits [29,63,92–96]. This is because wild studies are subject to limitations related to sample size, environmental heterogeneity and marker density, which may fail to identify trait loci, over-estimate effect sizes and/or generate spurious associations (e.g. as stated above for observed associations at rare variants for lamb IgE and IgG on chromosomes 16 and 10, respectively) [97]. We believe our overall findings are robust for the following reasons. This study has one of the highest sample sizes of any GWAS conducted in a wild system, with ~2,000 measures in lambs and ~3,800 measures in ~1300 unique adults, and sampling studies in this population suggest that causal variants contributing to heritable variation are adequately tagged by the Ovine SNP50 BeadChip [63,98]. The extent of LD between genotyped SNP loci allowed successful imputation of high-density SNP loci in almost all significant regions of the genome, providing sufficient power to fine-map loci of large effect on immune phenotypes [63]. We acknowledge that reduced LD in some regions (such as on chromosome 18 for lamb IgA) may mean that some regions of the genome are less able to tag heritable variation, potentially leading to reduced power to detect some trait loci. In addition, the Ovine SNP50 BeadChip has a low SNP density around the *DQA* and *DQB* loci in the MHC class II region, reducing power to detect associations (S10 & S13 Figs). Nevertheless: the quality of imputation was high within this region; other work has shown that there is no significant difference in patterns of LD and recombination rate compared to other locations within the genome [64]; and traits were successfully mapped to the MHC region within the current study.

Conclusion

This study provides evidence of a number of major effect loci and high additive genetic variation underlying complex immune traits in a wild population of Soay sheep, and provides a foundation for determining why genetic variation persists in immune traits by investigating associations with identified trait loci with individual fitness and genomic signatures of selection. The high heritability and repeatability of immune measures, as well as low correlations between them, suggests that strong targets for selection exist; a full understanding would require multivariate analysis with individual reproductive success and survival to understand the constraints on immune phenotype evolution. Previous studies of immunity in the wild often focussed on specific immune regions (e.g. the MHC) and candidate genes encoding proteins of known immune function. Our study reveals the importance of using a genome-wide

association, rather than candidate gene approach, for a clearer understanding of the genetic control of immune phenotypes. Overall, our study provides a rare example of multiple regions of large effect driving variation in immune phenotypes in the wild and presents strong evidence that immune profiles are temporally stable over an individual's lifetime.

Methods

Study population

The Soay sheep is a primitive breed of domestic sheep that was isolated on the island of Soay in the remote St Kilda archipelago several millennia ago, and has been living in unmanaged conditions since then [99]. In 1932, >100 Soay sheep were moved to the larger island of Hirta after the evacuation of all human residents. The population now fluctuates between 600 to 2,200 individuals. Approximately a third of the Hirta population lives in the Village Bay area, and these individuals have been the subject of a long-term study since 1985 [99]. In April, around 95% of all lambs born in the study area are caught each year and individually tagged. Each August, as many sheep as possible from the study population are re-captured using temporary traps [99]. At capture, whole blood samples are collected into heparin tubes, centrifuged at 3000 r.p.m. for 10 minutes, and plasma removed and stored at -20°C.

Ethics statement

All animal work has been carried out according to UK Home Office procedures and is licensed under the UK Animals (Scientific Procedures) Act of 1986 (license no. PPL60/4211).

Quantifying antibody levels

This study quantified antibody levels in animals that were caught and blood sampled in August between 1990 and 2015, comprising 6543 samples from 3190 individuals. Five samples from late-born lambs caught in August within 50 days of birth were excluded from the dataset, due to the potential presence of maternal antibodies and differences in development stage to other lambs. Levels of the antibodies IgA, IgG and IgE against antigens of the third larval stage of *Teladorsagia circumcincta* were measured using direct (IgA, IgG) and indirect (IgE) ELISAs. We used *T. circumcincta* L3 somatic antigen, provided by the Moredun Research Institute, as the capture antigen for all three assays diluted to 2μg/ml in 0.06M carbonate buffer at pH 9.6. 50μl of the diluted capture antigen was added to each well of a Nunc-immuno 96-microwell plate, which was covered and incubated at 4°C overnight. After washing the wells three times in Tris-buffered saline-Tween (TBST) using a plate washer, 50μl of the Soay sheep plasma sample diluted to 1:50 for IgA and IgE, and 1:12800 for IgG was added to each well. The plates were then covered and incubated at 37°C for 1 hour. Plates were then washed five times with TBST and 50μl per well of rabbit polyclonal anti-sheep IgA detection antibody conjugated to horseradish peroxidase (HRP) (AbD Serotec AHP949P) diluted 1:16000 was added to the anti-*T. circumcincta* IgA assay and 50μl per well of rabbit polyclonal anti-sheep IgG detection antibody conjugated to HRP (AbD Serotec 5184–2104) diluted 1:16000 was added to the anti-*T. circumcincta* IgG assay. For the anti-*T. circumcincta* IgE assay, 50μl per well of anti-sheep IgE (mouse monoclonal IgG1, clone 2F1, provided by the Moredun Research Institute) diluted 1:100 was added, followed by 1-hour incubation at 37°C, five washes with TBST and then 50μl per well of goat polyclonal anti-mouse IgG1-HRP detection antibody (AbD Serotec STAR132P) was added diluted to 1:8000 in TBST. All plates were then incubated at 37°C for 1 hour. Plates were then washed five times with TBST and 100μl of SureBlue TMB 1-Component microwell peroxidase substrate (KPL) was added per well and left to incubate for 5 minutes in

the dark at 37˚C. Reactions were stopped by adding 100μl per well of 1M hydrochloric acid and optical densities (OD) were read immediately at 450nm using a Thermo Scientific GO Spectrophotometer.

All results were measured as OD values due to the lack of standard solutions. To minimise confounding of capture year and age effects with plate to plate variation, each plate included samples from two years paired at random with different age groups on each plate. All plates were run in duplicate and duplicate sample ODs were removed if the coefficient of variation was > 0.2 or the difference between ODs was greater than 0.2. We also checked the correlation of ODs across duplicate plates and re-ran both plates if r < 0.8. We included two sample free wells (50μl TBST) as blanks and two wells of positive controls on each plate. The positive control for the IgE assay was pooled serum from ewes trickle-infected with *T. circumcincta* and for the IgA and IgG assays was pooled plasma from normal healthy non-immunised domestic sheep. For subsequent analyses, the mean optical density ratio of each sample was taken according to this formula:

$$OD = \frac{(\text{sample OD} - \text{blank OD})}{(\text{positive control OD} - \text{blank OD})}$$

where the numerator was set to zero if the blank OD was greater than the sample OD in order to avoid negative values. Distributions of antibodies are shown in S1 Fig. The number of samples that failed quality control per assay was 13 for IgA (7 lambs and 6 adults), 8 for IgE (6 lambs and 2 adults) and 27 for IgG (5 lambs and 22 adults). Correlations between antibody measures were modelled using linear regressions in R v3.4.3 (S2 Fig).

SNP data set

DNA was extracted from ear tissue or buffy coats using the Qiagen DNeasy blood and tissue kit according to the manufacturer's protocol, except that a single final elution with 50ul AE buffer was used to give DNA at a concentration ≥ 50ng/ul. A total of 7,386 Soay sheep have been genotyped at 51,135 SNPs on the Illumina Ovine SNP50 BeadChip. Quality control was carried out using the check.marker function in GenABEL version 1.8–0 [100] using the following thresholds: SNP minor allele frequency (MAF) > 0.01, SNP locus genotyping success > 0.95, individual sheep genotyping success > 0.95, identity by state with another individual ≥ 0.95. Following quality control, 39,176 SNPs from 7,268 sheep remained. A further 189 sheep have been genotyped at 606,066 SNP loci on the Ovine Infinium HD SNP BeadChip; these sheep were selected to maximise the genetic variation represented in the population and were subject to the same quality control thresholds as above (see [64] for individual selection criteria). All SNP locations were taken from their estimated positions on the sheep genome assembly Oar_v3.1 (GenBank assembly ID GCA_000298735.1 [101]). Pedigree relationships between individuals were inferred using data from 438 SNP loci in the R package Sequoia v1.02 [102] and from field observations between mothers and their offspring born within the study area (see [63] for SNP selection criteria).

Animal models

We modelled IgA, IgE and IgG levels in lambs and adults using a restricted maximum likelihood (REML) animal model approach [103] to determine the heritability of antibody levels in ASReml-R 3.0 [104] in R v3.4.3. We analysed lambs and adults separately due to a large difference observed in antibody levels (S2 Fig) and due to the expected immaturity of the immune response in 4-month-old lambs [60]. Using the above SNP dataset, a genomic relatedness matrix (GRM) at all autosomal markers was constructed for all genotyped individuals using

GCTA 1.90.2 beta0 [105] to determine the variance attributed to additive genetic effects (i.e. the narrow-sense heritability, h^2). Pedigree and GRM relatedness have been shown to be highly correlated in this system [98]. The GRM was adjusted using the argument --*grm-adj 0*, which assumes that allele frequencies of causal and genotyped loci are similar.

The fixed effect structure for the lamb models included sex and age in days as a linear covariate, while the random effects included the additive genetic component, maternal identity, birth year, ELISA plate number and ELISA run date. The fixed effect structure for the adult models included sex and age in years as a linear covariate, while the random effects included permanent environment (i.e. repeated measures within an individual) and capture year effects in addition to the random effects included in the lamb model. The proportion of the phenotypic variance explained by each random effect was estimated as the ratio of the relevant variance component to the sum of all variance components (i.e. the total phenotypic variance) as estimated by the animal model. The heritability of each measure was determined as the ratio of the additive genetic variance to the total phenotypic variance. The repeatability (i.e. the between-individual variation) of each measure in the adult and all age models was determined as the ratio of the sum of the additive genetic and permanent environment variance to the total phenotypic variance. Bivariate models were also run to determine the genetic correlation between all lamb and adult traits. For correlations within lambs or adults, the fixed effect structure of the model was as above; for correlations between lambs and adults, only sex was fit as a fixed factor. The additive genetic correlation, r_A, was calculated without constraint using the CORGH function (i.e. correlation with heterogenous variances) in ASReml-R 3.0. For adult models, a permanent environment was also fit as an additional random effect. The significance of the correlation estimate was determined using the reported Z-ratio with 1 degree of freedom.

Genome-wide association studies

Genome-wide association (GWA) was used to identify associations between individual single nucleotide polymorphisms (SNPs) and IgA, IgE and IgG levels in lambs and adults. This included SNPs on the X chromosome (N = 824) and those of unknown position (N = 313). For each trait and each class, a total of 39,176 individual animal models were run to determine the association with each SNP locus. Each model used the same fixed effect structures as above, with SNP genotype fitted as a two or three-level factor. To speed up computational time, the GRM was replaced with a relatedness matrix based on the pedigree (which is highly correlated with the GRM in this population [98]), and ELISA plate ID and run date were removed as random effects as they explained a very small proportion of the phenotypic variance (Fig 2). Models were run in ASReml-R 3.0 [104] in R v3.4.3. P-values were corrected for any additional unaccounted-for population structure by dividing them by the genomic control parameter λ [106] in cases where λ > 1, to reduce the incidence of false positives. λ was calculated as the median Wald test χ^2_2 divided by the median χ^2_2 expected from a null distribution. The significance threshold after multiple testing was determined using a linkage disequilibrium-based approach with a sliding window of 50 SNPs (outlined in [107]); for a false discovery rate of α = 0.05, the threshold P-value was set at 2.245×10^{-6} [64]. At the most highly associated SNPs, we repeated the animal models above including an interaction term between sex and SNP genotype to investigate if gene effects were sex-specific. The significance of this term was tested using a Wald test. Lamb IgE levels show strong right skew in their distribution (S1 Fig), which can increase spurious associations at rare alleles present in individuals with large trait values. To mitigate against this, all zero trait values were removed (N = 394), and the response variable $\log_{10}$ transformed; this correction had a negligible effect on the variance component estimates.

Variance explained by significantly associated regions

In regions of the genome where a SNP locus was significantly associated with an antibody measure, the proportion of phenotypic variation explained was modelled using a regional heritability approach [108]. Briefly, a second GRM was constructed as above using 50K SNP data from the most highly associated SNP in that region and the 9 SNP loci flanking that SNP on either side (i.e. 19 SNPs in total). This GRM was fitted as an additional random effect in the animal models and used to quantify the variance explained by variants within the associated region (see [64] for further details on the use of this method in Soay sheep).

Imputation of SNP genotypes in associated regions

Further investigation of significant associations from GWAS was carried out using an imputation approach using data from individuals typed on the Ovine Infinium HD SNP BeadChip. SNP genotypes were extracted from the HD chip ±2Mb on either side of all significantly associated regions. These data were then used to impute autosomal SNP genotypes in individuals typed on the 50K SNP chip alone, using a heuristic method in AlphaImpute 1.98 [109]. Briefly, this method uses pedigree information to phase the genotypes of the high-density individuals, which is followed by a haplotype library construction. The haplotype library is then used to iteratively impute missing genotypes around known genotypes in the 50K genotyped individuals. An example parameter file is provided in the data repository described below. Parameter files for each region are included in the analysis code repository (see below). SNPs with an imputation success >95% were retained and associations between antibody levels and genotypes at each imputed SNP was calculated using the same animal model structures as outlined for the GWAS above.

Gene and gene ontology annotation in associated regions

Gene annotations in significant regions were obtained from Ensembl (gene build ID Oar_v3.1.94). Gene ontology (GO) annotations for genes occurring within 1Mb of a significantly associated SNP were obtained from humans, mice, cattle and sheep gene builds using the function *getBM* in the R package biomaRt v2.34.2 [110]. For genes where the gene name was not known, orthologous genes were identified using the biomaRt function *getLDS*. For all the genes and orthologues identified within these regions, the gene names, phenotype descriptions and GO terms were queried for all terms associated with immune function and antibodies (using the strings immun* and antibod*).

Supporting information

S1 Fig. Histograms of anti-*Teladorsagia circumcincta* IgA, IgE and IgG levels in lamb (left column) and adult (right column) Soay sheep.
(TIF)

S2 Fig. Correlations between anti-*T. circumcincta* IgG, IgA, and IgE levels in lamb (A-C) and adult (D-F) Soay sheep. Model results are provided in S1 Table.
(TIF)

S3 Fig. Boxplots comparing anti-*T. circumcincta* IgG, IgA, and IgE levels between the sexes in lamb and adult Soay sheep.
(TIF)

S4 Fig. Anti-*T. circumcincta* IgG, IgA, and IgE levels in lambs with age in days (left) and in adults with age in years (right). Animal model results are provided in S2 Table.
(TIF)

S5 Fig. Distribution of observed vs expected P-values under a null χ^2 with 2 degrees of freedom for the GWAS of anti-*Teladorsagia circumcincta* IgA, IgE and IgG levels in lambs and adults. The dotted line indicates the genome-wide significance threshold, and the solid line indicates a 1:1 correspondence between the observed and expected values.
(TIF)

S6 Fig. Estimates of genotype effects in animal models for the most highly associated SNPs in Table 2. The model intercept is for genotype A/A. The full model results are provided in S9 Table. Sample sizes are provided above each set of points.
(TIF)

S7 Fig. Estimates of genotype effects in females (circles) and males (triangles) in animal models containing a SNP genotype by sex interaction for the most highly associated SNPs in Table 2. The model intercept is A/A females, except for locus oar3_OAR10_10333145, which is A/G females. The model results are provided in S10 Table. Sample sizes are provided above each set of points.
(TIF)

S8 Fig. Local association of anti-*Teladorsagia circumcincta* IgA levels in lambs with SNP50 and imputed SNP loci at the most highly associated region on chromosome 24. The dotted line indicates the genome-wide significance threshold equivalent to an experiment-wide threshold of P = 0.05. Points are colour-coded by their imputation status i.e. from the SNP50 chip (red triangles) or imputed from the Ovine HD chip (black points). Underlying data, sample sizes and effect sizes are provided in S6 Table. Gene positions are shown in the grey panel at the top of each plot and were obtained from Ensembl (gene build ID Oar_v3.1.94) and are provided in S7 Table. Genes coloured red have GO terms associated with immune traits (S8 Table).
(TIF)

S9 Fig. Local association of anti-*Teladorsagia circumcincta* IgA levels in adults with SNP50 and imputed SNP loci at the most highly associated region on chromosome 24. The dotted line indicates the genome-wide significance threshold equivalent to an experiment-wide threshold of P = 0.05. Points are colour-coded by their imputation status i.e. from the SNP50 chip (red triangles) or imputed from the Ovine HD chip (black points). Underlying data, sample sizes and effect sizes are provided in S6 Table. Gene positions are shown in the grey panel at the top of each plot and were obtained from Ensembl (gene build ID Oar_v3.1.94) and are provided in S7 Table. Genes coloured red have GO terms associated with immune traits (S8 Table).
(TIF)

S10 Fig. Local association of anti-*Teladorsagia circumcincta* IgA levels in lambs with SNP50 and imputed SNP loci at the most highly associated region on chromosome 20. The dotted line indicates the genome-wide significance threshold equivalent to an experiment-wide threshold of P = 0.05. Points are colour-coded by their imputation status i.e. from the SNP50 chip (red triangles) or imputed from the Ovine HD chip (black points). Underlying data, sample sizes and effect sizes are provided in S6 Table. Gene positions are shown in the grey panel at the top of each plot and were obtained from Ensembl (gene build ID Oar_v3.1.94) and are provided in S7 Table. Genes coloured red have GO terms associated with

immune traits (S8 Table).
(TIF)

S11 Fig. Local association of anti-*Teladorsagia circumcincta* IgA levels in lambs with SNP50 and imputed SNP loci at the most highly associated region on chromosome 18. The dotted line indicates the genome-wide significance threshold equivalent to an experiment-wide threshold of P = 0.05. Points are colour-coded by their imputation status i.e. from the SNP50 chip (red triangles) or imputed from the Ovine HD chip (black points). Underlying data, sample sizes and effect sizes are provided in S6 Table. Gene positions are shown in the grey panel at the top of each plot and were obtained from Ensembl (gene build ID Oar_v3.1.94) and are provided in S7 Table. Genes coloured red have GO terms associated with immune traits (S8 Table).
(TIF)

S12 Fig. Local association of anti-*Teladorsagia circumcincta* IgE levels in lambs with SNP50 and imputed SNP loci at the most highly associated region on chromosome 10. The dotted line indicates the genome-wide significance threshold equivalent to an experiment-wide threshold of P = 0.05. Points are colour-coded by their imputation status i.e. from the SNP50 chip (red triangles) or imputed from the Ovine HD chip (black points). Underlying data, sample sizes and effect sizes are provided in S6 Table. Gene positions are shown in the grey panel at the top of each plot and were obtained from Ensembl (gene build ID Oar_v3.1.94) and are provided in S7 Table. Genes coloured red have GO terms associated with immune traits (S8 Table).
(TIF)

S13 Fig. Local association of anti-*Teladorsagia circumcincta* IgE levels in adults with SNP50 and imputed SNP loci at the most highly associated region on chromosome 20. The dotted line indicates the genome-wide significance threshold equivalent to an experiment-wide threshold of P = 0.05. Points are colour-coded by their imputation status i.e. from the SNP50 chip (red triangles) or imputed from the Ovine HD chip (black points). Underlying data, sample sizes and effect sizes are provided in S6 Table. Gene positions are shown in the grey panel at the top of each plot and were obtained from Ensembl (gene build ID Oar_v3.1.94) and are provided in S7 Table. Genes coloured red have GO terms associated with immune traits (S8 Table).
(TIF)

S14 Fig. Local association of anti-*Teladorsagia circumcincta* IgG levels in lambs with SNP50 and imputed SNP loci at the most highly associated region on chromosome 20. The dotted line indicates the genome-wide significance threshold equivalent to an experiment-wide threshold of P = 0.05. Points are colour-coded by their imputation status i.e. from the SNP50 chip (red triangles) or imputed from the Ovine HD chip (black points). Underlying data, sample sizes and effect sizes are provided in S6 Table. Gene positions are shown in the grey panel at the top of each plot and were obtained from Ensembl (gene build ID Oar_v3.1.94) and are provided in S7 Table. Genes coloured red have GO terms associated with immune traits (S8 Table).
(TIF)

S15 Fig. Local association of anti-*Teladorsagia circumcincta* IgG levels in lambs with SNP50 and imputed SNP loci at the most highly associated region on chromosome 16. The dotted line indicates the genome-wide significance threshold equivalent to an experiment-wide threshold of P = 0.05. Points are colour-coded by their imputation status i.e. from the

SNP50 chip (red triangles) or imputed from the Ovine HD chip (black points). Underlying data, sample sizes and effect sizes are provided in S6 Table. Gene positions are shown in the grey panel at the top of each plot and were obtained from Ensembl (gene build ID Oar_v3.1.94) and are provided in S7 Table. Genes coloured red have GO terms associated with immune traits (S8 Table).
(TIF)

S1 Table. Correlations between anti-*Teladorsagia circumcincta* antibody levels in lambs and adults. Slope, intercept, adjusted R^2 and P-values are given for linear regressions.
(DOCX)

S2 Table. Fixed effects results from animal models of anti-*Teladorsagia circumcincta* IgA, IgE and IgG for lambs, and adults. Age is the age in days during the August catch for lambs, and age in years for adults. Wald statistics are given for the significance of each effect as included in the model.
(DOCX)

S3 Table. Random effects results from animal models of anti-*Teladorsagia circumcincta* IgA, IgE and IgG for lambs and adults. Wald statistics are given for the significance of each effect as included in the model. Sample sizes are provided in Table 1. Fixed effect structures and results are provided in S2 Table.
(DOCX)

S4 Table. Temporal correlations in anti-*Teladorsagia circumcincta* IgA, IgE and IgG levels at time t and t+1 (in years) as shown in Fig 3. Results are from a linear regression with t+1 levels as the response variable.
(DOCX)

S5 Table. Full GWAS results for animal models of anti-*Teladorsagia circumcincta* IgA, IgE and IgG in lambs and adults, fitting SNP genotype as a factor. A and B indicate the reference and alternate allele at each SNP. CallRate is the genotyping success of the locus on the SNP50 BeadChip. MAF is the frequency of allele B (minor allele frequency). Wald P and Wald P Corrected are the association P-values before and after correction with genomic control λ, respectively. Significant indicates if the SNP was significantly associated with trait variation after correcting for multiple testing. Effect AA, AB and BB are the effect sizes from the animal model for each genotype relative to the model intercept.
(CSV)

S6 Table. Full association results for animal models of anti-*Teladorsagia circumcincta* IgA, IgE and IgG in lambs and adults, fitting imputed SNP genotypes as a factor. SNP.Type indicates whether the SNP was imputed from the HD chip or from the SNP50 BeadChip (unknown genotypes are also imputed for the SNP50 BeadChip in this analysis meaning that results will not exactly match those of S5 Table). A and B indicate the reference and alternate allele at each SNP. ImputeSuccess is the imputation success reported from the AlphaImpute analysis. MAF is the frequency of allele B (minor allele frequency). Wald P are the association P-values that have not been corrected for genomic control (see main text). Effect AA, AB and BB are the effect sizes from the animal model for each genotype relative to the model intercept.
(CSV)

S7 Table. Gene information in regions significantly associated with anti-*Teladorsagia circumcincta* IgA, IgE and IgG in lambs and adults, obtained from the Ensembl Gene build Oar_v3.1.94. Start and stop indicate the gene start and stop positions. Strand indicates

whether transcription occurs in the forward or reverse strand. Gene_id is the Ensembl identifier for the gene. Gene_name is the gene name associated with the gene_id. Gene_biotype indicates the type of gene (i.e. protein coding, RNA etc). Orthologue is the gene name of orthologues associated with the gene ID, with orthologue count giving the number of unique orthologues. Consensus locus is the gene name or likely gene name based on orthology. (CSV)

S8 Table. Gene Ontology information for loci (including orthologues) in S7 Table that are associated with immune and antibody phenotypes in humans (hsapiens), mice (mmusculus), cattle (btaurus) and sheep (oaries) obtained using biomaRt. Column names are as for S7 Table, including the following: gene_id is the sheep gene ID; Species = species as previous; ensembl_gene_id is the gene ID within that Species; external_gene_name is the gene name for that species; description is the full gene name; phenotype_description is a description of phenotypes associated with the gene; go_id is the GO term identifier; name_1006 is the GO term name; definition_1006 is the GO term definition. (CSV)

S9 Table. Genotypic effects at the most significant GWAS loci (Table 2). These results are visualized in S6 Fig. Associated Wald statistics and P values are provided in Tables 2, S5 and S6. (DOCX)

S10 Table. Sex-specific SNP effects at the most significant GWAS loci (Table 2). The Wald statistic is given for the sex by genotype interaction term. These results are visualized in S7 Fig. (DOCX)

Acknowledgments

Author Contributions

Conceptualization: Alexandra M. Sparks, Tom N. McNeilly, Daniel H. Nussey, Susan E. Johnston.

Formal analysis: Alexandra M. Sparks, Susan E. Johnston.

Funding acquisition: Josephine M. Pemberton, Tom N. McNeilly, Daniel H. Nussey.

Investigation: Alexandra M. Sparks, Kathryn Watt, Rona Sinclair, Jill G. Pilkington.

Methodology: Alexandra M. Sparks, Kathryn Watt, Tom N. McNeilly, Daniel H. Nussey, Susan E. Johnston.

Project administration: Jill G. Pilkington, Josephine M. Pemberton, Daniel H. Nussey.

Resources: Josephine M. Pemberton, Tom N. McNeilly, Daniel H. Nussey, Susan E. Johnston.

Supervision: Tom N. McNeilly, Daniel H. Nussey, Susan E. Johnston.

Visualization: Alexandra M. Sparks, Susan E. Johnston.

Writing – original draft: Alexandra M. Sparks, Daniel H. Nussey, Susan E. Johnston.

Writing – review & editing: Alexandra M. Sparks, Josephine M. Pemberton, Tom N. McNeilly, Daniel H. Nussey, Susan E. Johnston.

References

1. Pedersen AB, Babayan SA. Wild immunology. Mol Ecol. 2011; 20: 872–880. https://doi.org/10.1111/j.1365-294X.2010.04938.x PMID: 21324009

2. Maizels RM, Nussey DH. Into the wild: digging at immunology's evolutionary roots. Nat Immunol. 2013; 14: 879–883. https://doi.org/10.1038/ni.2643 PMID: 23959175

3. Liston A, Carr EJ, Linterman MA. Shaping variation in the human immune system. Trends Immunol. 2016; 37: 637–646. https://doi.org/10.1016/j.it.2016.08.002 PMID: 27692231

4. Orrù V, Steri M, Sole G, Sidore C, Virdis F, Dei M, et al. Genetic variants regulating immune cell levels in health and disease. Cell. 2013; 155: 242–256. https://doi.org/10.1016/j.cell.2013.08.041 PMID: 24074872

5. Roederer M, Quaye L, Mangino M, Beddall MH, Mahnke Y, Chattopadhyay P, et al. The genetic architecture of the human immune system: a bioresource for autoimmunity and disease pathogenesis. Cell. 2015; 161: 387–403. https://doi.org/10.1016/j.cell.2015.02.046 PMID: 25772697

6. Atlija M, Arranz J-J, Martinez-Valladares M, Gutiérrez-Gil B. Detection and replication of QTL underlying resistance to gastrointestinal nematodes in adult sheep using the ovine 50K SNP array. Genet Sel Evol. 2016; 48: 4. https://doi.org/10.1186/s12711-016-0182-4 PMID: 26791855

7. Yang I V, Wade CM, Kang HM, Alper S, Rutledge H, Lackford B, et al. Identification of novel genes that mediate innate immunity using inbred mice. Genetics. 2009; 183: 1535–1544. https://doi.org/10.1534/genetics.109.107540 PMID: 19805818

8. Thompson-Crispi KA, Sewalem A, Miglior F, Mallard BA. Genetic parameters of adaptive immune response traits in Canadian Holsteins. J Dairy Sci. 2012; 95: 401–409. https://doi.org/10.3168/jds.2011-4452 PMID: 22192219

9. Abolins S, King EC, Lazarou L, Weldon L, Hughes L, Drescher P, et al. The comparative immunology of wild and laboratory mice, *Mus musculus domesticus.* Nat Commun. 2017; 8: 14811. https://doi.org/10.1038/ncomms14811 PMID: 28466840

10. Beura LK, Hamilton SE, Bi K, Schenkel JM, Odumade OA, Casey KA, et al. Normalizing the environment recapitulates adult human immune traits in laboratory mice. Nature. 2016; 532: 512–516. https://doi.org/10.1038/nature17655 PMID: 27096360

11. Murphy L, Eckersall PD, Bishop SC, Pettit JJ, Huntley JF, Burchmore R, et al. Genetic variation among lambs in peripheral IgE activity against the larval stages of *Teladorsagia circumcincta*. Parasitology. 2010; 137: 1249–1260. https://doi.org/10.1017/S0031182010000028 PMID: 20233490

12. Strain SAJ, Bishop SC, Henderson NG, Kerr A, McKellar QA, Mitchell S, et al. The genetic control of IgA activity against *Teladorsagia circumcincta* and its association with parasite resistance in naturally infected sheep. Parasitology. 2002; 124: 545–552. https://doi.org/10.1017/s0031182002001531 PMID: 12049417

13. Clapperton M, Glass EJ, Bishop SC. Pig peripheral blood mononuclear leucocyte subsets are heritable and genetically correlated with performance. Animal. 2008; 2: 1575–1584. https://doi.org/10.1017/S1751731108002929 PMID: 22444008

14. Flori L, Gao Y, Oswald IP, Lefevre F, Bouffaud M, Mercat M-J, et al. Deciphering the genetic control of innate and adaptive immune responses in pig: a combined genetic and genomic study. BMC Proc. 2011; 5: S32. https://doi.org/10.1186/1753-6561-5-S4-S32 PMID: 21645313

15. Denholm SJ, McNeilly TN, Banos G, Coffey MP, Russell GC, Bagnall A, et al. Estimating genetic and phenotypic parameters of cellular immune-associated traits in dairy cows. J Dairy Sci. 2017; 100: 2850–2862. https://doi.org/10.3168/jds.2016-11679 PMID: 28131586

16. Davies G, Stear MJ, Benothman M, Abuagob O, Kerr A, Mitchell S, et al. Quantitative trait loci associated with parasitic infection in Scottish blackface sheep. Heredity. 2006; 96: 252–258. https://doi.org/10.1038/sj.hdy.6800788 PMID: 16391549

17. Edfors-Lilja I, Wattrang E, Marklund L, Moller M, Andersson-Eklund L, Andersson L, et al. Mapping quantitative trait loci for immune capacity in the pig. J Immunol. 1998; 161: 829–835. PMID: 9670960

18. Lu X, Liu JF, Fu WX, Zhou JP, Luo YR, Ding XD, et al. Genome-wide association study for cytokines and immunoglobulin G in swine. PLoS One. 2013; 8: 1–7. https://doi.org/10.1371/journal.pone.0074846 PMID: 24098351

19. Thompson-Crispi KA, Sargolzaei M, Ventura R, Abo-Ismail M, Miglior F, Schenkel F, et al. A genome-wide association study of immune response traits in Canadian Holstein cattle. BMC Genomics. 2014; 15: 559. https://doi.org/10.1186/1471-2164-15-559 PMID: 24996426

20. Pitala N, Gustafsson L, Sendecka J, Brommer JE. Nestling immune response to phytohaemagglutinin is not heritable in collared flycatchers. Biol Lett. 2007; 3: 418–421. https://doi.org/10.1098/rsbl.2007.0135 PMID: 17567550

21. Morrison ES, Ardia DR, Clotfelter ED. Cross-fostering reveals sources of variation in innate immunity and hematocrit in nestling tree swallows *Tachycineta bicolor*. J Avian Biol. 2009; 40: 573–578. https://doi.org/10.1111/j.1600-048X.2009.04910.x

22. Sakaluk SK, Wilson AJ, Bowers EK, Johnson LS, Masters BS, Johnson BG, et al. Genetic and environmental variation in condition, cutaneous immunity, and haematocrit in house wrens. BMC Evol Biol. 2014; 14: 242. https://doi.org/10.1186/s12862-014-0242-8 PMID: 25471117

23. Kim S-Y, Fargallo JA, Vergara P, Martínez-Padilla J. Multivariate heredity of melanin-based coloration, body mass and immunity. Heredity. 2013; 111: 139–146. https://doi.org/10.1038/hdy.2013.29 PMID: 23591519

24. Drobniak SM, Wiejaczka D, Arct A, Dubiec A, Gustafsson L, Cichoń M. Sex-specific heritability of cell-mediated immune response in the blue tit nestlings (*Cyanistes caeruleus*). J Evol Biol. 2010; 23: 1286–1292. https://doi.org/10.1111/j.1420-9101.2010.01993.x PMID: 20456564

25. Bonneaud C, Sinsheimer JS, Richard M, Chastel O, Sorci G. MHC polymorphisms fail to explain the heritability of phytohaemagglutinin-induced skin swelling in a wild passerine. Biol Lett. 2009; 5: 784–787. https://doi.org/10.1098/rsbl.2009.0435 PMID: 19671600

26. Coltman DW, Wilson K, Pilkington JG, Stear MJ, Pemberton JM. A microsatellite polymorphism in the gamma interferon gene is associated with resistance to gastrointestinal nematodes in a naturally-parasitized population of Soay sheep. Parasitology. 2001; 122: 571–582. https://doi.org/10.1017/s0031182001007570 PMID: 11393831

27. Turner AK, Begon M, Jackson JA, Bradley JE, Paterson S. Genetic diversity in cytokines associated with immune variation and resistance to multiple pathogens in a natural rodent population. PLoS Genet. 2011; 7: e1002343. https://doi.org/10.1371/journal.pgen.1002343 PMID: 22039363

28. Brown EA, Pilkington JG, Nussey DH, Watt KA, Hayward AD, Tucker R, et al. Detecting genes for variation in parasite burden and immunological traits in a wild population: testing the candidate gene approach. Mol Ecol. 2013; 22: 757–773. https://doi.org/10.1111/j.1365-294X.2012.05757.x PMID: 22998224

29. Wenzel MA, James MC, Douglas A, Piertney SB. Genome-wide association and genome partitioning reveal novel genomic regions underlying variation in gastrointestinal nematode burden in a wild bird. Mol Ecol. 2015; 24: 4175–4192. https://doi.org/10.1111/mec.13313 PMID: 26179597

30. Paterson S, Wilson K, Pemberton JM. Major histocompatibility complex variation associated with juvenile survival and parasite resistance in a large unmanaged ungulate population (Ovis aries L.). Proc Natl Acad Sci U S A. 1998; 95: 3714–3719. https://doi.org/10.1073/pnas.95.7.3714 PMID: 9520432

31. Kennedy MW, Nager RG. The perils and prospects of using phytohaemagglutinin in evolutionary ecology. Trends Ecol Evol. 2006; 21: 653–655. https://doi.org/10.1016/j.tree.2006.09.017 PMID: 17028055

32. Owen JP, Clayton DH. Where are the parasites in the PHA response? Trends Ecol Evol. 2007; 22: 228–229. https://doi.org/10.1016/j.tree.2007.02.003 PMID: 17296246

33. Schmid-Hempel P. Evolutionary Parasitology: The Integrated Study of Infections, Immunology, Ecology, and Genetics. New York: Oxford University Press; 2011.

34. Jepson A, Banya W, Sisay-Joof F, Hassan-King M, Nunes C, Bennett S, et al. Quantification of the relative contribution of major histocompatibility complex (MHC) and non-MHC genes to human immune responses to foreign antigens. Infect Immun. 1997; 65: 872–876. PMID: 9038290

35. Acevedo-Whitehouse K, Cunningham AA. Is MHC enough for understanding wildlife immunogenetics? Trends Ecol Evol. 2006; 21: 433–438. https://doi.org/10.1016/j.tree.2006.05.010 PMID: 16764966

36. Palacios MG, Winkler DW, Klasing KC, Hasselquist D, Vleck CM. Consequences of immune system aging in nature: a study of immunosenescence costs in free-living Tree Swallows. Ecology. 2011; 92: 952–966. https://doi.org/10.1890/10-0662.1 PMID: 21661557

37. Simon AK, Hollander GA, McMichael A. Evolution of the immune system in humans from infancy to old age. Proceedings Biol Sci. 2015; 282: 20143085. https://doi.org/10.1098/rspb.2014.3085 PMID: 26702035

38. Tsang JS, Schwartzberg PL, Kotliarov Y, Biancotto A, Xie Z, Germain RN, et al. Global analyses of human immune variation reveal baseline predictors of postvaccination responses. Cell. 2014; 157: 499–513. https://doi.org/10.1016/j.cell.2014.03.031 PMID: 24725414

39. Carr EJ, Dooley J, Garcia-Perez JE, Lagou V, Lee JC, Wouters C, et al. The cellular composition of the human immune system is shaped by age and cohabitation. Nat Immunol. 2016; 17: 461–468. https://doi.org/10.1038/ni.3371 PMID: 26878114

40. Brodin P, Jojic V, Gao T, Bhattacharya S, Angel CJL, Furman D, et al. Variation in the human immune system is largely driven by non-heritable influences. Cell. 2015; 160: 37–47. https://doi.org/10.1016/j.cell.2014.12.020 PMID: 25594173

41. Banos G, Wall E, Coffey MP, Bagnall A, Gillespie S, Russell GC, et al. Identification of immune traits correlated with dairy cow health, reproduction and productivity. PLoS One. 2013; 8: e65766. https://doi.org/10.1371/journal.pone.0065766 PMID: 23776543

42. Benavides MV, Sonstegard TS, Van Tassell C. Genomic regions associated with sheep resistance to gastrointestinal nematodes. Trends Parasitol. 2016; 32: 470–480. https://doi.org/10.1016/j.pt.2016.03.007 PMID: 27183838

43. McNeilly TN, Devaney E, Matthews JB. *Teladorsagia circumcincta* in the sheep abomasum: defining the role of dendritic cells in T cell regulation and protective immunity. Parasite Immunol. 2009; 31: 347–356. https://doi.org/10.1111/j.1365-3024.2009.01110.x PMID: 19527450

44. Stear MJ, Bishop SC, Doligalska M, Duncan JL, Holmes PH, Irvine J, et al. Regulation of egg production, worm burden, worm length and worm fecundity by host responses in sheep infected with *Ostertagia circumcincta*. Parasite Immunol. 1995; 17: 643–652. https://doi.org/10.1111/j.1365-3024.1995.tb01010.x PMID: 8834764

45. Riggio V, Matika O, Pong-Wong R, Stear MJ, Bishop SC. Genome-wide association and regional heritability mapping to identify loci underlying variation in nematode resistance and body weight in Scottish Blackface lambs. Heredity. 2013; 110: 420–429. https://doi.org/10.1038/hdy.2012.90 PMID: 23512009

46. Gutiérrez-Gil B, Pérez J, de la Fuente LF, Meana A, Martínez-Valladares M, San Primitivo F, et al. Genetic parameters for resistance to trichostrongylid infection in dairy sheep. Animal. 2010; 4: 505–512. https://doi.org/10.1017/S1751731109991431 PMID: 22444037

47. Bishop SC, Stear MJ. Inheritance of faecal egg counts during early lactation in Scottish Blackface ewes facing mixed, natural nematode infections. Anim Sci. 2001; 73: 389–395. https://doi.org/10.1017/S1357729800058355

48. Gutiérrez-Gil B, Pérez J, Alvarez L, Martínez-Valladares M, de la Fuente L-F, Bayón Y, et al. Quantitative trait loci for resistance to trichostrongylid infection in Spanish Churra sheep. Genet Sel Evol. 2009; 41: 46. https://doi.org/10.1186/1297-9686-41-46 PMID: 19863786

49. Nieuwhof GJ, Bishop SC. Costs of the major endemic diseases of sheep in Great Britain and the potential benefits of reduction in disease impact. Anim Sci. 2005; 81: 23–29. https://doi.org/10.1079/ASC41010023

50. Kaplan RM, Vidyashankar AN. An inconvenient truth: global worming and anthelmintic resistance. Vet Parasitol. 2012; 186: 70–78. https://doi.org/10.1016/j.vetpar.2011.11.048 PMID: 22154968

51. Jackson F, Miller J. Alternative approaches to control—quo vadit? Vet Parasitol. 2006; 139: 371–384. https://doi.org/10.1016/j.vetpar.2006.04.025 PMID: 16750600

52. De Cisneros JPJ, Stear MJ, Mair C, Singleton D, Stefan T, Stear A, et al. An explicit immunogenetic model of gastrointestinal nematode infection in sheep. J R Soc Interface. 2014; 11. https://doi.org/10.1098/rsif.2014.0416 PMID: 25121649

53. Wilson K, Grenfell BT, Pilkington JG, Boyd HEG, Gulland FM. Parasites and their impact. In: Clutton-Brock TH, Pemberton JM, editors. Soay Sheep: Dynamics and Selection in an Island Population. Cambridge, UK: Cambridge University Press; 2004. pp. 17–51.

54. Craig BH, Pilkington JG, Pemberton JM. Gastrointestinal nematode species burdens and host mortality in a feral sheep population. Parasitology. 2006; 133: 485–496. https://doi.org/10.1017/S0031182006000618 PMID: 16817995

55. Gulland FM. The role of nematode parasites in Soay sheep (*Ovis aries* L.) mortality during a population crash. Parasitology. 1992; 105: 493–503. https://doi.org/10.1017/s0031182000074679 PMID: 1461688

56. Hayward AD, Wilson AJ, Pilkington JG, Clutton-Brock TH, Pemberton JM, Kruuk LEB. Natural selection on a measure of parasite resistance varies across ages and environmental conditions in a wild

mammal. J Evol Biol. 2011; 24: 1664–1676. https://doi.org/10.1111/j.1420-9101.2011.02300.x PMID: 21658142

57. Hayward AD, Garnier R, Watt KA, Pilkington JG, Grenfell BT, Matthews JB, et al. Heritable, heterogeneous, and costly resistance of sheep against nematodes and potential feedbacks to epidemiological dynamics. Am Nat. 2014; 184: S58–76. https://doi.org/10.1086/676929 PMID: 25061678

58. Nussey DH, Watt KA, Clark A, Pilkington JG, Pemberton JM, Graham AL, et al. Multivariate immune defences and fitness in the wild: complex but ecologically important associations among plasma antibodies, health and survival. Proc R Soc London B Biol Sci. 2014; 281: 20132931. https://doi.org/10.1098/rspb.2013.2931 PMID: 24500168

59. Watson RL, McNeilly TN, Watt KA, Pemberton JM, Pilkington JG, Waterfall M, et al. Cellular and humoral immunity in a wild mammal: variation with age & sex and association with overwinter survival. Ecol Evol. 2016; 6: 8695–8705. https://doi.org/10.1002/ece3.2584 PMID: 28035261

60. Sparks AM, Watt K, Sinclair R, Pilkington JG, Pemberton JM, Johnston SE, et al. Natural Selection on Antihelminth Antibodies in a Wild Mammal Population. Am Nat. 2018; 192: 745–760. https://doi.org/10.1086/700115 PMID: 30444657

61. Beraldi D, McRae AF, Gratten J, Pilkington JG, Slate J, Visscher PM, et al. Quantitative trait loci (QTL) mapping of resistance to strongyles and coccidia in the free-living Soay sheep (Ovis aries). Int J Parasitol. 2007; 37: 121–129. https://doi.org/10.1016/j.ijpara.2006.09.007 PMID: 17067607

62. Johnston SE, Gratten J, Berenos C, Pilkington JG, Clutton-Brock TH, Pemberton JM, et al. Life history trade-offs at a single locus maintain sexually selected genetic variation. Nature. 2013; 502: 93–95. https://doi.org/10.1038/nature12489 PMID: 23965625

63. Bérénos C, Ellis PA, Pilkington JG, Lee SH, Gratten J, Pemberton JM. Heterogeneity of genetic architecture of body size traits in a free-living population. Mol Ecol. 2015; 24: 1810–1830. https://doi.org/10.1111/mec.13146 PMID: 25753777

64. Johnston SE, Bérénos C, Slate J, Pemberton JM. Conserved genetic architecture underlying individual recombination rate variation in a wild population of soay sheep (Ovis aries). Genetics. 2016; 203: 583–598. https://doi.org/10.1534/genetics.115.185553 PMID: 27029733

65. Reith W, LeibundGut-Landmann S, Waldburger J-M. Regulation of MHC class II gene expression by the class II transactivator. Nat Rev Immunol. 2005; 5: 793–806. https://doi.org/10.1038/nri1708 PMID: 16200082

66. Krawczyk M, Reith W. Regulation of MHC class II expression, a unique regulatory system identified by the study of a primary immunodeficiency disease. Tissue Antigens. 2006; 67: 183–197. https://doi.org/10.1111/j.1399-0039.2006.00557.x PMID: 16573555

67. Ferreira RC, Pan-Hammarström Q, Graham RR, Gateva V, Fontán G, Lee AT, et al. Association of IFIH1 and other autoimmunity risk alleles with selective IgA deficiency. Nat Genet. 2010; 42: 777–780. https://doi.org/10.1038/ng.644 PMID: 20694011

68. Li J, Jørgensen SF, Maggadottir SM, Bakay M, Warnatz K, Glessner J, et al. Association of CLEC16A with human common variable immunodeficiency disorder and role in murine B cells. Nat Commun. 2015; 6: 6804. https://doi.org/10.1038/ncomms7804 PMID: 25891430

69. Bronson PG, Chang D, Bhangale T, Seldin MF, Ortmann W, Ferreira RC, et al. Common variants at PVT1, ATG13-AMBRA1, AHI1 and CLEC16A are associated with selective IgA deficiency. Nat Genet. 2016; 48: 1425–1429. https://doi.org/10.1038/ng.3675 PMID: 27723758

70. Liu W, Yan M, Liu Y, Wang R, Li C, Deng C, et al. Olfactomedin 4 down-regulates innate immunity against *Helicobacter pylori* infection. Proc Natl Acad Sci U S A. 2010; 107: 11056–11061. https://doi.org/10.1073/pnas.1001269107 PMID: 20534456

71. Chaplin JW, Kasahara S, Clark EA, Ledbetter JA. Anti-CD180 (RP105) Activates B Cells To Rapidly Produce Polyclonal Ig via a T Cell and MyD88-Independent Pathway. J Immunol. 2011; 187: 4199–4209. https://doi.org/10.4049/jimmunol.1100198 PMID: 21918197

72. Coulson T, Catchpole EA, Albon SD, Morgan BJT, Pemberton JM, Clutton-Brock TH, et al. Age, sex, density, winter weather and population crashes in Soay sheep. Science. 2001; 292: 1528–1531. https://doi.org/10.1126/science.292.5521.1528 PMID: 11375487

73. Crawley MJ, Albon SD, Bazely DR, Milner JM, Pilkington JG, Tuke AL. Vegetation and sheep population dynamics. In: Clutton-Brock TH, Pemberton JM, editors. Soay Sheep: Dynamics and Selection in an Island Population. Cambridge: Cambridge University Press; 2004. pp. 89–112.

74. Falconer DS, Mackay TFC. Introduction to Quantitative Genetics. 4th ed. UK: Longman; 1996.

75. Seppälä O. Natural selection on quantitative immune defence traits: a comparison between theory and data. J Evol Biol. 2015; 28: 1–9. https://doi.org/10.1111/jeb.12528 PMID: 25400248

76. Mori-Aoki A, Pietrarelli M, Nakazato M, Caturegli P, Kohn LD, Suzuki K. Class II Transactivator Suppresses Transcription of Thyroid-Specific Genes. Biochem Biophys Res Commun. 2000; 278: 58–62. https://doi.org/10.1006/bbrc.2000.3769 PMID: 11071855

77. Chang C-H, Guerder S, Hong S-C, van Ewijk W, Flavell RA. Mice Lacking the MHC Class II Transactivator (CIITA) Show Tissue-Specific Impairment of MHC Class II Expression. Immunity. 1996; 4: 167–178. https://doi.org/10.1016/s1074-7613(00)80681-0 PMID: 8624807

78. DeSandro AM, Nagarajan UM, Boss JM. Associations and interactions between bare lymphocyte syndrome factors. Mol Cell Biol. 2000; 20: 6587–6599. https://doi.org/10.1128/mcb.20.17.6587-6599.2000 PMID: 10938133

79. Sawcer S, Hellenthal G, Pirinen M, Spencer CCA, Patsopoulos NA, Moutsianas L, et al. Genetic risk and a primary role for cell-mediated immune mechanisms in multiple sclerosis. Nature. 2011; 476: 214–219. https://doi.org/10.1038/nature10251 PMID: 21833088

80. Leikfoss IS, Keshari PK, Gustavsen MW, Bjolgerud A, Brorson IS, Celius EG, et al. Multiple Sclerosis Risk Allele in CLEC16A Acts as an Expression Quantitative Trait Locus for CLEC16A and SOCS1 in CD4+ T Cells. PLoS One. 2015; 10: e0132957. https://doi.org/10.1371/journal.pone.0132957 PMID: 26203907

81. Wang Y, Yuan W, Guo H, Jiang Y. High frequency of activated NKp46(+) natural killer cells in patients with new diagnosed of latent autoimmune diabetes in adults. Autoimmunity. 2015; 48: 267–273. https://doi.org/10.3109/08916934.2014.990629 PMID: 25495606

82. Márquez A, Varadé J, Robledo G, Martínez A, Mendoza JL, Taxonera C, et al. Specific association of a CLEC16A/KIAA0350 polymorphism with NOD2/CARD15- Crohn's disease patients. Eur J Hum Genet. 2009; 17: 1304–1308. https://doi.org/10.1038/ejhg.2009.50 PMID: 19337309

83. Skinningsrud B, Husebye ES, Pearce SH, McDonald DO, Brandal K, Wolff AB, et al. Polymorphisms in CLEC16A and CIITA at 16p13 Are Associated with Primary Adrenal Insufficiency. J Clin Endocrinol Metab. 2008; 93: 3310–3317. https://doi.org/10.1210/jc.2008-0821 PMID: 18593762

84. Skinningsrud B, Lie BA, Husebye ES, Kvien TK, Førre Ø, Flatø B, et al. A CLEC16A variant confers risk for juvenile idiopathic arthritis and anti-cyclic citrullinated peptide antibody negative rheumatoid arthritis. Ann Rheum Dis. 2010; 69: 1471–1474. https://doi.org/10.1136/ard.2009.114934 PMID: 19734133

85. Pandey R, Bakay M, Hain HS, Strenkowski B, Elsaqa BZB, Roizen JD, et al. CLEC16A regulates splenocyte and NK cell function in part through MEK signaling. PLoS One. 2018; 13: e0203952. Available: https://doi.org/10.1371/journal.pone.0203952 PMID: 30226884

86. Graham AL, Hayward AD, Watt KA, Pilkington JG, Pemberton JM, Nussey DH. Fitness correlates of heritable variation in antibody responsiveness in a wild mammal. Science. 2010; 330: 662–665. https://doi.org/10.1126/science.1194878 PMID: 21030656

87. Diao Y, Fang R, Li B, Meng Z, Yu J, Qiu Y, et al. A tiling-deletion-based genetic screen for cis-regulatory element identification in mammalian cells. Nat Methods. 2017; 14: 629–635. https://doi.org/10.1038/nmeth.4264 PMID: 28417999

88. Orsolya S, François S. From remote enhancers to gene regulation: charting the genome's regulatory landscapes. Philos Trans R Soc B Biol Sci. 2013; 368: 20120358. https://doi.org/10.1098/rstb.2012.0358 PMID: 23650632

89. Crawford AM, Paterson KA, Dodds KG, Diez Tascon C, Williamson PA, Roberts Thomson M, et al. Discovery of quantitative trait loci for resistance to parasitic nematode infection in sheep: I. Analysis of outcross pedigrees. BMC Genomics. 2006; 7: 178. https://doi.org/10.1186/1471-2164-7-178 PMID: 16846521

90. Kijas JW, Lenstra JA, Hayes B, Boitard S, Porto Neto LR, San Cristobal M, et al. Genome-wide analysis of the world's sheep breeds reveals high levels of historic mixture and strong recent selection. PLoS Biol. 2012; 10: e1001258. https://doi.org/10.1371/journal.pbio.1001258 PMID: 22346734

91. Feulner PGD, Gratten J, Kijas JW, Visscher PM, Pemberton JM, Slate J. Introgression and the fate of domesticated genes in a wild mammal population. Mol Ecol. 2013; 22: 4210–4221. https://doi.org/10.1111/mec.12378 PMID: 23786437

92. Husby A, Kawakami T, Ronnegard L, Smeds L, Ellegren H, Qvarnstrom A. Genome-wide association mapping in a wild avian population identifies a link between genetic and phenotypic variation in a life-history trait. Proc R Soc London B Biol Sci. 2015; 282: 20150156. https://doi.org/10.1098/rspb.2015.0156 PMID: 25833857

93. Johnston SE, Orell P, Pritchard VL, Kent MP, Lien S, Niemelä E, et al. Genome-wide SNP analysis reveals a genetic basis for sea-age variation in a wild population of Atlantic salmon (Salmo salar). Mol Ecol. 2014; 23: 3452–3468. https://doi.org/10.1111/mec.12832 PMID: 24931807

94. Kardos M, Husby A, McFarlane SE, Qvarnström A, Ellegren H. Whole-genome resequencing of extreme phenotypes in collared flycatchers highlights the difficulty of detecting quantitative trait loci in

natural populations. Mol Ecol Resour. 2016; 16: 727–741. https://doi.org/10.1111/1755-0998.12498 PMID: 26649993

95. Santure AW, Poissant J, De Cauwer I, Van Oers K, Robinson MR, Quinn JL, et al. Replicated analysis of the genetic architecture of quantitative traits in two wild great tit populations. Mol Ecol. 2015; 24: 6148–6162. https://doi.org/10.1111/mec.13452 PMID: 26661500

96. Silva CNS, McFarlane SE, Hagen IJ, Rönnegård L, Billing AM, Kvalnes T, et al. Insights into the genetic architecture of morphological traits in two passerine bird species. Heredity. 2017; 119: 197–205. https://doi.org/10.1038/hdy.2017.29 PMID: 28613280

97. Santure AW, Garant D. Wild GWAS—association mapping in natural populations. Mol Ecol Resour. 2018; 18: 729–738. https://doi.org/10.1111/1755-0998.12901 PMID: 29782705

98. Bérénos C, Ellis P, Pilkington JG, Pemberton JM. Estimating quantitative genetic parameters in wild populations: a comparison of pedigree and genomic approaches. Mol Ecol. 2014; 23: 3434–3451. https://doi.org/10.1111/mec.12827 PMID: 24917482

99. Clutton-Brock T, Pemberton J. Soay sheep: dynamics and selection in an island population. Cambridge, UK: Cambridge University Press; 2004.

100. Aulchenko YS, Ripke S, Isaacs A, van Duijn CM. GenABEL: an R library for genome-wide association analysis. Bioinformatics. 2007; 23: 1294–1296. https://doi.org/10.1093/bioinformatics/btm108 PMID: 17384015

101. Jiang Y, Xie M, Chen W, Talbot R, Maddox JF, Faraut T, et al. The sheep genome illuminates biology of the rumen and lipid metabolism. Science. 2014; 344: 1168–1173. https://doi.org/10.1126/science.1252806 PMID: 24904168

102. Huisman J. Pedigree reconstruction from SNP data: parentage assignment, sibship clustering and beyond. Mol Ecol Resour. 2017; 17: 1009–1024. https://doi.org/10.1111/1755-0998.12665 PMID: 28271620

103. Henderson CR. Best linear unbiased estimation and prediction under a selection model. Biometrics. 1975; 31: 423–447. https://doi.org/10.2307/2529430 PMID: 1174616

104. Butler DG, Cullis BR, Gilmour AR, Gogel BJ. Mixed Models for S language Environments: ASReml-R reference manual. 2009.

105. Yang J, Lee SH, Goddard ME, Visscher PM. GCTA: a tool for genome-wide complex trait analysis. Am J Hum Genet. 2011; 88: 76–82. https://doi.org/10.1016/j.ajhg.2010.11.011 PMID: 21167468

106. Devlin AB, Roeder K, Devlin B. Genomic Control for Association. Biometrics. 1999; 55: 997–1004. PMID: 11315092

107. Moskvina V, Schmidt KM. On multiple-testing correction in genome-wide association studies. Genet Epidemiol. 2008; 32: 567–573. https://doi.org/10.1002/gepi.20331 PMID: 18425821

108. Nagamine Y, Pong-Wong R, Navarro P, Vitart V, Hayward C, Rudan I, et al. Localising Loci underlying Complex Trait Variation Using Regional Genomic Relationship Mapping. PLoS One. 2012; 7: e46501. https://doi.org/10.1371/journal.pone.0046501 PMID: 23077511

109. Hickey JM, Kinghorn BP, Tier B, Werf JHJ Van Der, Cleveland MA. A phasing and imputation method for pedigreed populations that results in a single-stage genomic evaluation. Genet Sel Evol. 2012; 44: 9. https://doi.org/10.1186/1297-9686-44-9 PMID: 22462519

110. Durinck S, Spellman PT, Birney E, Huber W. Mapping identifiers for the integration of genomic datasets with the R/Bioconductor package biomaRt. Nat Protoc. 2009; 4: 1184–1191. https://doi.org/10.1038/nprot.2009.97 PMID: 19617889