Received: 20 June 2019 | Revised: 8 August 2019 | Accepted: 9 August 2019

DOI: 10.1111/jbg.12436

Check for updates

ORIGINAL ARTICLE

Journal of Animal Breeding and Genetics WILEY

Genome-wide association study for insect bite hypersensitivity susceptibility in horses revealed novel associated loci on chromosome 1

Merina Shrestha[1,2] | **Marina Solé**[1] | **Bart J. Ducro**[2] | **Marie Sundquist**[3] | **Ruth Thomas**[4] | **Anouk Schurink**[2] | **Susanne Eriksson**[1] | **Gabriella Lindgren**[1,5]

[1]Department of Animal Breeding and Genetics, Swedish University of Agricultural Sciences, Uppsala, Sweden

[2]Animal Breeding and Genomics, Wageningen University and Research, Wageningen, the Netherlands

[3]Östra Greda Research Group, Borgholm, Sweden

[4]The Exmoor Pony Society, Woodmans, Deveon, UK

[5]Department of Biosystems, KU Leuven, Leuven, Belgium

Correspondence
Gabriella Lindgren, Department of Animal Breeding and Genetics, Swedish University of Agricultural Sciences, P.O. Box 7023, 75007 Uppsala, Sweden.
Email: gabriella.lindgren@slu.se

Funding information
Research Executive Agency, Grant/Award Number: FP7/2007-2013

Abstract

Equine insect bite hypersensitivity (IBH) is a pruritic skin allergy caused primarily by biting midges, *Culicoides* spp. IBH susceptibility has polygenic inheritance and occurs at high frequencies in several horse breeds worldwide, causing increased costs and reduced welfare of affected horses. The aim of this study was to identify and validate single nucleotide polymorphisms (SNPs) associated with equine IBH susceptibility. After quality control, 33,523 SNPs were included in a Bayesian genome-wide association study on 177 affected and 178 unaffected Icelandic horses. We report associated regions in *E. caballus* (ECA) 1, 3, 15 and 18, overlapping with known IBH QTLs in horses, and novel regions containing several genes, together explaining 11.46% of the total genetic variance. For validation, three SNPs on ECA 1 and ECA X (explaining the largest percentage of genetic variance) within 1-mb genomic windows for IBH were genotyped in an independent population of 280 Exmoor ponies. The associated genomic region (152–153 mb) on ECA 1 was confirmed in Exmoor ponies and contains the *AQR* gene involved in splicing processes and a long non-coding RNA. This study confirms the polygenic nature of IBH susceptibility and suggests a role of transcriptional regulatory mechanisms (e.g., alternative splicing) for IBH predisposition in these horse breeds.

KEYWORDS

Exmoor pony, genetic risk factors, genome-wide association study, Icelandic horse, insect bite hypersensitivity

1 | INTRODUCTION

Equine insect bite hypersensitivity (IBH) is an allergic skin reaction occurring worldwide in various breeds (Wagner et al., 2006). IBH is seasonally recurrent as the prime cause is allergens in the saliva of biting midges of *Culicoides* spp. Some affected horses also react to other blood-feeding insects such as black flies (*Simulium* spp.) (Fadok & Greiner, 1990; Quinn, Baker, & Morrow, 1983). The prevalence varies under different conditions: previous studies found that less than 10% of Swedish-born Icelandic horses (Eriksson et al., 2008) and up to 50% of horses exported from Iceland to European countries (Bjornsdottir, Sigvaldadottir, Brostrom, Langvad, & Sigurdsson, 2006) were affected. Since *Culicoides* are absent in Iceland,

Merina Shrestha and Marina Solé contributed equally to this work.

Susanne Eriksson and Gabriella Lindgren contributed equally to this work

 © 2019 Blackwell Verlag GmbH

horses born on the island cannot develop a normal immune response to its antigen. Affected horses have been shown to react to allergens from native as well as exotic species of *Culicoides* (Anderson, Belton, & Kleider, 1993; Schaffartzik et al., 2012), but with stronger IgE response to the species the horses had previously been exposed to (van der Meide et al., 2012). Intense itching, lesions, skin thickening, crusting and scaling characterize the condition that creates extreme discomfort for the horses (Fadok & Greiner, 1990). Current preventive and therapeutic measures against IBH are costly and non-curative (Pilsworth & Knottenbelt, 2004; Schaffartzik et al., 2012). A more sustainable approach to reduce prevalence of IBH could be selection and breeding of horses less prone to develop the condition. Most heritability estimates for IBH susceptibility status on the observable scale have been low (ranging from 0.07 to 0.30) (Eriksson et al., 2008; Peeters, Janssens, Brebels, & Buys, 2015; Schurink, Ducro, Heuven, & van Arendonk, 2011; Schurink, van Grevenhof, Ducro, & van Arendonk, 2009), and reducing the prevalence through traditional breeding would thus require extensive recording over many generations. If genomic information could be utilized, this may help overcome such limitations. Genomic analyses of IBH susceptibility could also improve our understanding of biological mechanisms underlying allergic disorders.

A few candidate gene studies or genome-wide association (GWA) studies of IBH susceptibility in horses have previously been published. The equine leucocyte antigen (ELA) class II region (Andersson et al., 2012; Klumplerova et al., 2013), as well as some non-ELA genes regulating immunity and allergy (Klumplerova et al., 2013; Vychodilova et al., 2013), was suggested in candidate gene studies to influence IBH susceptibility. In a study on Icelandic horses and Exmoor ponies (Andersson et al., 2012), the same microsatellite marker in the ELA class II region was found to be associated with IBH susceptibility in both breeds, indicating that the genetic background of IBH susceptibility may be at least partly similar in these breeds. Other genomic regions were reported to be associated with IBH susceptibility in GWA studies (Schurink et al., 2012; Shrestha et al., 2015; Velie et al., 2016), but these were based on less numbers of horses than this study.

The aim of the present study was to find genomic regions associated with equine IBH susceptibility. To increase the statistical power of the GWA study, we combined data from previously published IBH studies on Icelandic horses in the Netherlands (Schurink et al., 2012) and Sweden (Shrestha et al., 2015). To validate non-randomly associated SNPs identified in this study, Exmoor ponies with known IBH status were genotyped for the markers explaining the largest percentage of genetic variance.

2 | MATERIALS AND METHODS

2.1 | Horses and phenotype assessment

The present GWA study comprised data on 355 Icelandic horses from Sweden and the Netherlands from previous studies within country (Schurink et al., 2012; Shrestha et al., 2015). Previously published IBH data on 280 Exmoor ponies in the UK (Velie et al., 2016) were used for validation of the markers explaining the largest percentage of genetic variance. Data on IBH symptoms were in all three countries gathered using similar questionnaires. In Sweden, the questionnaires were sent to Icelandic horse owners in 2005, and veterinarians later collected blood samples and confirmed the IBH symptoms (Eriksson et al., 2008; Shrestha et al., 2015). In the Netherlands, the questionnaire was filled out in 2010 by a researcher interviewing the owners, while a veterinarian examined the horse and collected blood samples (Schurink et al., 2012). In both countries, the Icelandic horses classified as unaffected had been exposed to midges in an area in which other horses had developed symptoms, showed no signs of IBH and were at least 4 years old to provide sufficient time for IBH to develop. The inspection of IBH symptoms was made after the horses had been exposed to *Culicoides* spp. during the warm season in the country in which they were sampled. Further details on data collection and phenotype assessment are described by Shrestha et al. (2015) and Schurink et al. (2012). In the UK, an owner questionnaire was used to gather information about IBH status, and pedigrees were provided by the Exmoor Pony Society. Samples from 336 Exmoor ponies collected between 2008 and 2011 were available and after removal of horses with unclear IBH status or missing pedigree, 280 were selected for genotyping as described in detail by Velie et al. (2016).

In total, 177 affected and 178 unaffected Icelandic horses were included in the study, and among the affected horses, there were slightly more females (52%) compared to males (48%). The Icelandic horses were 4–35 years at the time of data collection, with a mean age of 12.8 years ($SD = 4.8$) in affected horses and of 12.4 years ($SD = 5.2$) in unaffected horses. The vast majority (92%) of the Icelandic horses were born in Sweden or the Netherlands, the rest were imported from neighbouring countries (<3%) or from Iceland (<6%) (Table S1). Affected and unaffected horses were matched based on pedigree information from the international WorldFengur database (Lorange, 2011) to avoid the formation of subgroups that could lead to false-positive associations. The studied Icelandic horses were offspring of in total 136 sires and 333 dams, and the data contained 76.9% paternal half-sibs, 8.1% maternal half-sibs and 2.8% full sibs. Out of the 280 Exmoor ponies born and sampled in the UK, 39% were affected with IBH and 61% were unaffected. The age

information was incomplete for the Exmoor ponies. Almost two-thirds (65%) of the ponies were female, 34% were male and 1% had no information on gender. The distribution of affected (45%) and unaffected (55%) males was even, while there was fewer affected (37%) than unaffected (63%) females in the Exmoor pony data.

2.2 | SNP array genotyping and quality control

The 209 Icelandic horses sampled in Sweden were genotyped on the Illumina EquineSNP50 Genotyping BeadChip containing 54,602 single nucleotide polymorphisms (SNPs) (Shrestha et al., 2015). The 146 Icelandic horses sampled in the Netherlands were genotyped on the Illumina EquineSNP70 Genotyping BeadChip containing 65,157 SNPs (Schurink et al., 2012). A custom Perl script was used to find 45,986 SNPs in common between the EquineSNP50 and EquineSNP70 arrays. Quality control of the SNP data was performed using the GenABEL package (version 1.7-6) of R statistical environment (http://www.R-project.org (Aulchenko, Ripke, Isaacs, & van Duijn, 2007)). All the 355 Icelandic horses passed the quality control of ≥95% genotyping call rate. SNPs with a call rate of <95% (2.9%) and a minor allele frequency (MAF) of <5% (24.9%) were removed from the subsequent analysis. The remaining 33,523 SNPs were included in the GWA case–control study.

To investigate potential stratification due to country of origin between the Icelandic horses in Sweden and in the Netherlands, MAF of the post-quality-controlled SNPs was plotted. The correlation between the MAF for the two groups of horses was calculated and tested for significance using R statistical software (version 3.0.1) (R core Team, 2013). A multidimensional scaling (MDS) plot was also constructed based on a genomic relationship matrix with the combined data on Swedish and Dutch Icelandic horses. The matrix (355 by 355) containing pairwise identity by state values was constructed based on 31,967 autosomal SNPs using the GenABEL package and *ibs*( ) function. The genomic inflation factor (λ) was obtained by comparing *p*-values from a preliminary single SNP mixed-model approach with the expected distribution of *p*-values.

2.3 | Genome-wide association study

The Bayesian variable selection method Bayes C (Habier, Fernando, Kizilkaya, & Garrick, 2011; Kizilkaya, Fernando, & Garrick, 2010) using the GenSel v4R software (https://github.com/austin-putz/GenSel (Fernando & Garrick, 2009)) was used to perform the GWA study of IBH susceptibility as a binary trait (affected/unaffected) on the 355 Icelandic horses. We fitted all SNPs simultaneously as random effects in the model as described in a previous study (Shrestha et al., 2015). No covariates were used in the model because factors, such as sex and age, previously reported to be important (Eriksson et al., 2008), were not found to be significant in these data. This was likely due to the efforts to include a balanced number of unaffected and affected horses for each sex and the exclusion of young unaffected horses. The prior for marker variance was derived from the additive genetic variance (0.369) based on the estimated heritability (h^2) of 0.27 on the underlying liability scale for IBH susceptibility (Eriksson et al., 2008). Two hundred thousand Markov chain Monte Carlo cycles were performed. On each cycle, 0.1% of all SNPs were assumed to have an effect in the model. The first 50% of the cycles were discarded as a burn-in period, and estimates from every 100th cycle were saved.

The horse genome build (EquCab 2.0) was divided into 2,371 non-overlapping 1-mb windows. Windows explaining more than the expected percentage (100%/ 2,371 = 0.042%) of genetic variance were considered to be non-randomly associated with IBH susceptibility. In the GenSel software, the samples were directly obtained from the posterior distribution of parameters. It was therefore not necessary to obtain the significance level for the variance explained by the window (Wolc et al., 2012). Within windows, the SNP with the highest model frequency was taken as the candidate SNP associated with IBH susceptibility. The model frequency for a SNP refers to the proportion of post-burn-in cycles that include the SNP in the model. For SNPs associated with IBH susceptibility, the risk allele frequencies in unaffected and affected horses were investigated. This was done both in the combined and separate data sets from Sweden and the Netherlands. To test the sensitivity of the method, we also changed the value of additive genetic variance from 0.369 to 0.123 ($h^2 = 0.11$) and re-ran the analysis.

2.4 | Gene annotation and overlap with known IBH QTLs

The bioinformatics database Ensembl (http://www.ensembl.org/) was used to screen for candidate genes. Genomic coordinates of candidate SNPs associated with IBH 500 kb upstream and downstream were used as inputs to generate a list of annotated genes using the Ensembl Variant Effect Predictor online tool (McLaren et al., 2016). The identified genomic regions were also compared to previously reported IBH QTLs in the horse (downloaded from the horse QTL database; (Hu, Park, & Reecy, 2016)) using bed file comparisons in BEDTools software (-merge Bed command) (Quinlan & Hall, 2010).

2.5 | Validation of most strongly associated SNPs in Exmoor ponies

The candidate SNPs within the windows that explained the largest percentage of the genetic variance for IBH

susceptibility in Icelandic horses were selected for validation in the 280 Exmoor ponies. We therefore genotyped three SNPs in the Exmoor ponies, using ABI custom-designed TaqMan SNP Genotyping Assay kits (Applied Biosystems). To design the assays, SNP information along with the sequences flanking the SNP of interest was imported from the horse genome build (EquCab 2.0). The sequences, designed primers and probes for the SNPs of interest are described in Table S2. Real-time polymerase chain reaction was performed using StepOnePlus Real-Time PCR System (Life Technologies). PCRs consisted of a pre-PCR holding stage at 60°C for 30 s, holding stage at 95°C for 10 min followed by 45 cycles of denaturation at 95°C for 15 s and annealing/extension at 60°C for 1 min. The final holding stage was at 60°C for 30 s. The association with IBH susceptibility in Exmoor ponies was investigated for each SNP using a chi-square test with 1 degree of freedom (*df*) in an online statistical software (www.vassarstats.net). No covariates were included in the test as the effect of sex was not found to be significant, and the age information was incomplete.

3 | RESULTS

3.1 | Population stratification analysis

Similar minor allele frequencies for the 33,523 SNPs were found in the Swedish and Dutch Icelandic horses (Figure 1), with a correlation of 0.91 (p-value = 2.26e-16). The multidimensional scaling plot identified two groups of Icelandic horses (Figure 2): a small group (PC1 > 0.10) of 38 horses with an even representation of affected and unaffected horses from Sweden related to a particular sire and a large group (PC1 < 0.10) of 317 horses (159 unaffected and 158 affected) from both countries. The first principal component, PC1 captured 3.8% and the second (PC2) captured 2.6% of the genetic variance in the combined data set. The genomic control inflation factor (λ) was 0.92 for the 355 Icelandic horses, supporting that the data were not stratified with regard to IBH susceptibility.

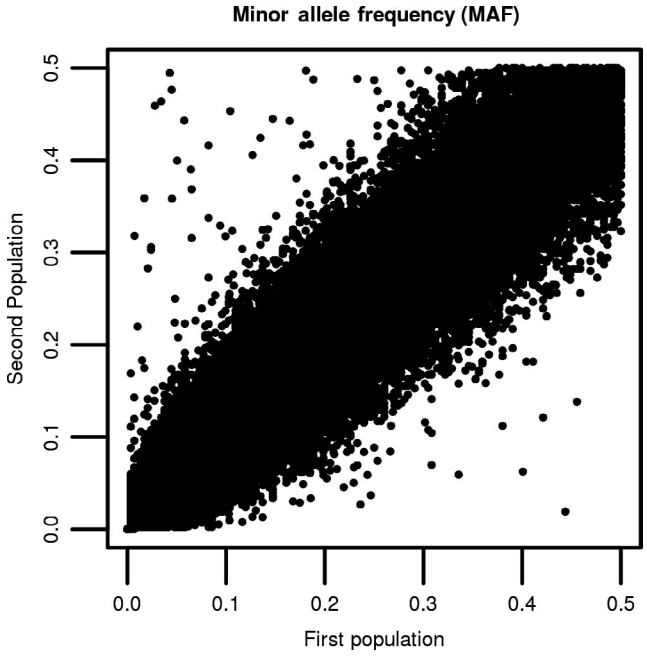

FIGURE 1 Frequency of minor alleles for 33,523 SNPs common between the Icelandic horse populations from the Netherlands and Sweden

3.2 | Genome-wide association study

In the GWA study including 355 Icelandic horses and 33,523 SNPs, on average 32.8 SNPs (SD = 5.8) were included in the model across the post-burn-in cycles. The GWA analysis of the combined data showed non-random associations of 1-mb windows with IBH susceptibility, and the 30 windows that explained the highest percentage of genetic variance are presented in Table 1. Together, these explained 11.46% of the genetic variance. The largest percentage of genetic variance (1.71%) was explained by a window located on the ECA X (82 mb, Figure 3). Two windows located on ECA 1 (152 and 153 mb, Figure 3) explained the second largest (1.44%) percentage of genetic variance. Changing the prior for additive genetic variance from 0.369 to 0.123 had generally very small effects on the results. The main difference was that the largest genetic variance explained by the window on ECA X (82 mb) decreased from 1.71% to 0.56%.

The risk allele frequency for the candidate SNP BIEC2_1138724 in the window that explained the highest percentage of genetic variation was 0.42 in the unaffected group compared to 0.58 in the affected group in the combined data (Table 2). Risk allele frequencies were on a similar level for Icelandic horses from Sweden and the Netherlands for candidate SNPs within all 30 windows.

3.3 | Validation of associated SNPs in Exmoor ponies

We tested three markers explaining the largest percentage of genetic variance in the GWA study in Icelandic horses in two windows on ECA 1 and one window X for association with IBH susceptibility in the Exmoor ponies. Among the tested genetic variants, we found significant associations with IBH susceptibility in the ponies for two SNPs located at ECA 1 (Table 3). The SNP BIEC2_65455 on ECA 1 (153 mb) was successfully genotyped in 277 ponies and associated with IBH (p-value = .019) with an odds ratio of 0.65. The risk allele frequency for this SNP was 0.50 in the unaffected group and 0.60 in the affected group of ponies. The other associated SNP at ECA 1 (p-value = .021) was BIEC2_65299 (152 mb), which was successfully genotyped in 278 ponies. The odds ratio for this SNP was 1.52. The risk allele frequency was 0.49 in the unaffected group and 0.60 in the affected group.

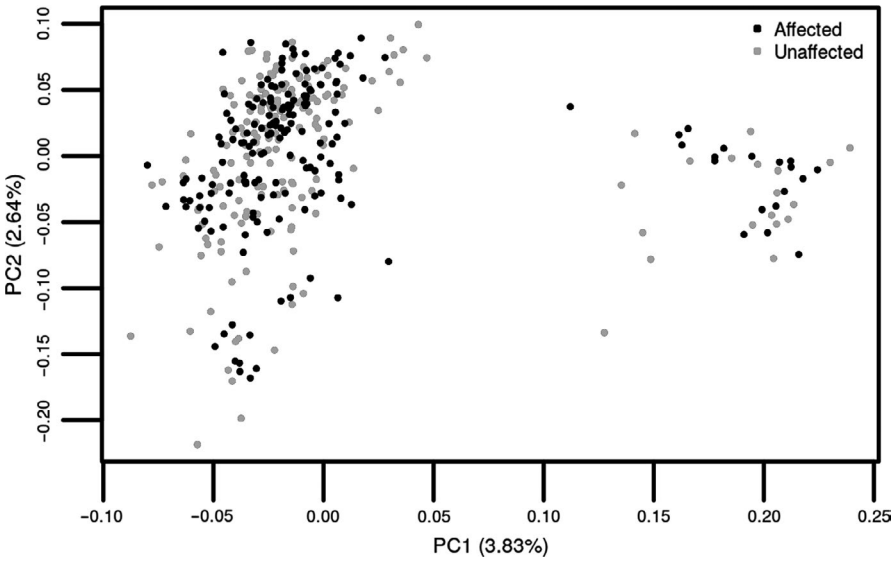

FIGURE 2 Multidimensional scaling plot for 355 Icelandic horses. The *x*-axis represents principal component (PC) 1 that explained 3.83% of the variation among the horses, and the *y*-axis represents PC2 that explained 2.64% of the variation

The SNP BIEC2_1138724 on ECA X (82 mb) was not associated with IBH susceptibility in Exmoors 1 (p-value = .887).

4 | DISCUSSION

Knowledge is still lacking about genetic risk factors for IBH susceptibility, in spite of the large number of horses suffering from the condition, and the noticeable economic consequences for owners of affected horses. The multifactorial and polygenic nature of IBH susceptibility makes it challenging to study, as large data sets of high quality are needed. Still, the genetic background in horses remains interesting to study further. Increased knowledge about IBH susceptibility in horses may be of value for understanding allergic eczema in other species as well.

4.1 | Effects of combining data across countries

To increase data size and statistical power in the present study, we combined Icelandic horse data from the Netherlands and Sweden that were previously used in separate studies (Schurink et al., 2012; Shrestha et al., 2015). Icelandic horses are regularly exported from Iceland to both the Netherlands and Sweden, and the horses belong to the same international studbook (Lorange, 2011). Genetic similarities between horses in the two countries were clearly shown in the population stratification analysis. Similar risk allele frequencies in both populations for candidate SNPs in all 30 windows that explained the highest percentage of genetic variation imply that combining the two data sets increased the power in this study and assisted in finding the association of SNPs with IBH susceptibility. When using the combined data in the GWA study, a higher percentage of the genetic variation was explained (9.38%) by the 20 most associated windows than what was explained by the 20 most associated windows identified in the separate Dutch (4.59%) (Schurink et al., 2012) and Swedish data (4.78%) (Shrestha et al., 2015). A higher percentage of genetic variance was also explained in the combined data (0.54%) than in the study on Dutch data (0.39%) (Schurink et al., 2012) for a single window on ECA 3 (35 mb) that was among the four most associated windows

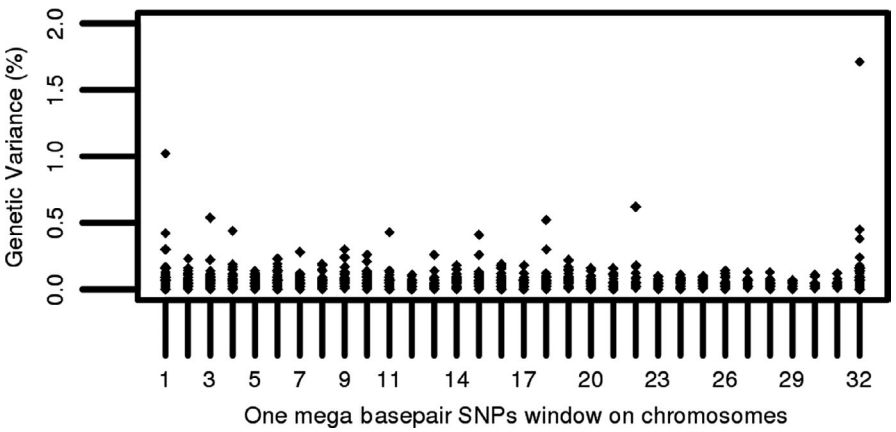

FIGURE 3 Manhattan plot of the percentage of additive genetic variance explained by each SNP by one mega-basepair (Mb) windows on the autosomal and X (32) chromosomes

TABLE 1 Genetic variance explained by 30 1-mb non-overlapping windows (that explained the highest percentage of genetic variation) along with SNPs associated with IBH susceptibility

1-mb windows explaining highest percentage of genetic variance					Associated SNP within 1-mb window	
ECA	mb	No. of SNP /window	Gen. var. (%)	Prop.> average	SNP name	Model freq.
1	7	15	0.30	0.028	BIEC2_2768	0.012
1	152	19	0.42	0.046	BIEC2_65299	0.013
1	153	14	1.02	0.066	BIEC2_65455	0.054
2	78	17	0.23	0.028	BIEC2_491393	0.005
3	35	17	0.54	0.068	BIEC2_776764	0.017
3	84	18	0.22	0.021	BIEC2_798106	0.016
4	80	15	0.19	0.026	BIEC2_871881	0.008
4	100	18	0.44	0.050	BIEC2_878810	0.031
6	24	19	0.19	0.036	BIEC2_945312	0.005
6	43	16	0.23	0.026	BIEC2_950930	0.011
7	55	16	0.28	0.043	BIEC2_1001712	0.007
8	39	17	0.19	0.028	BIEC2_1045967	0.004
					BIEC2_1045968	0.004
9	14	15	0.24	0.029	BIEC2_1075148	0.006
9	50	16	0.30	0.036	BIEC2_1093076	0.013
10	28	14	0.21	0.034	BIEC2_111858	0.005
10	29	8	0.26	0.029	BIEC2_111947	0.021
10	36	14	0.21	0.023	BIEC2_115985	0.009
11	0	13	0.43	0.046	BIEC2_133565	0.015
13	20	17	0.26	0.038	BIEC2_215349	0.006
15	19	17	0.41	0.044	BIEC2_293386	0.013
15	90	6	0.26	0.024	BIEC2_326263	0.023
16	51	22	0.19	0.032	BIEC2_349357	0.003
18	26	15	0.30	0.041	BIEC2_409581	0.008
18	32	21	0.52	0.046	BIEC2_410288	0.012
19	55	17	0.22	0.034	BIEC2_446571	0.017
22	25	12	0.62	0.052	BIEC2_589497	0.038
X	4	12	0.38	0.035	BIEC2_1108696	0.012
X	76	21	0.45	0.046	BIEC2_1133893	0.017
X	81	11	0.24	0.031	BIEC2_1138294	0.006
X	82	12	1.71	0.099	BIEC2_1138724	0.085

Abbreviation: Gen. var. (%), percentage of genetic variance explained by the window; mb, mega-basepair, 152 mb specifies a region from 152 to 153 (1-mb) on chromosome 1; Model freq., proportion of post-burn-in cycles where the SNP was included in the model and had an effect; Prop.> average, proportion of cycles where the window explained more than the expected genetic variance of 0.042%.

[a]Long non-coding RNA gene.

[b]Distance from the associated SNP to the nearest gene.

in both studies. Even after the increase, the percentage of variance explained by each top window remained low, however.

Several, but not all, genomic regions found to be associated with IBH susceptibility in either of the previous studies on IBH susceptibility in Dutch (Schurink et al., 2012) and Swedish (Shrestha et al., 2015) Icelandic horses were also found to be among the 30 windows that explained the highest percentage of genetic variation in the present study.

Position of SNP (bp) EquCab 2.0	Position of SNP (bp) EquCab 3.0	SNP effect	Nearest gene	Distance (bp)[b]	IBH QTL ID
7,759,159	7,879,776	−2.572e−03	ENSECAG00000029089[a]	37,880	29303
152,871,014	154,538,548	−2.941e−03	ENSECAG00000016399 (AQR)	0	
153,075,829	154,743,141	−1.597e−02	ENSECAG00000036544[a]	0	
78,264,466	78,433,704	7.269e−04	ENSECAG00000028993[a]	0	
35,408,432	36,131,080	−3.611e−03	ENSECAG00000003668 (CBFA2T3)	0	29304
84,757,007	86,881,691	−3.655e−03	ENSECAG00000022751 (GRXCR1)	137,392	
80,717,808	80,821,952	−1.475e−03	ENSECAG00000017360	1,709	
100,879,322	101,321,544	−7.747e−03	ENSECAG00000009606 (CUL1)	0	
24,315,560	24,250,966	8.633e−04	ENSECAG00000009886 (TRAF3IP1)	0	
43,005,654	44,139,170	−2.319e−03	ENSECAG00000015695 (PTPRO)	0	
55,887,882	57,939,282	1.242e−03	ENSECAG00000030449	248,362	
39,473,041	42,507,504	6.935e−04	ENSECAG00000039923[a]	0	
39,474,433	42,508,896	7.021e−04			
14,257,160	14,616,708	1.059e−03	ENSECAG00000039949[a]	0	
50,020,494	51,726,961	−2.734e−03	ENSECAG00000033571[a]	28,777	
28,870,560	29,398,152	8.861e−04	ENSECAG00000034972	9,504	
29,495,226	30,023,577	4.953e−03	ENSECAG00000038506[a]	333,482	
36,981,521	37,798,136	1.945e−03	ENSECAG00000004513 (DOP1A)	0	
300,347	302,364	3.480e−03	ENSECAG00000010105 (TBCD)	0	
20,145,342	21,226,046	9.623e−04	ENSECAG00000008601(XPO6)	0	
19,652,562	20,332,525	2.741e−03	ENSECAG00000011573 (DNAH6)/ ENSECAG00000010499 (TRABD2A)	0	29313
90,556,961	91,835,560	−5.881e−03	ENSECAG00000006257 (SNTG2)	0	
51,593,954	53,201,120	−4.337e−04	ENSECAG00000010584 (DYNC1LI1)	31,965	
26,766,516	26,831,155	−1.551e−03	ENSECAG00000012312 (ARHGAP15)/ ENSECAG00000014136 (GTDC1)	0	
32,615,336	32,647,983	2.581e−03	ENSECAG00000021310 (RND3)	122,809	29318
55,553,410	58,415,478	−3.936e−03	ENSECAG00000019508 (COL8A1)	92,610	
25,515,664	26,414,929	−1.037e−02	ENSECAG00000017510 (PIGU)	0	
4,923,258	5,810,591	−2.282e−03	ENSECAG00000035101 (ANOS1)	1,678	
76,174,026	80,565,502	3.133e−03	ENSECAG00000011844 (DIAPH2)	0	
81,974,630	86,288,723	−1.336e−03	ENSECAG00000036284	5,172	
82,842,940	87,097,446	−2.120e−02	ENSECAG00000039374[a]	100,125	

Overlapping associated genomic regions with the previous study on Dutch horses (Schurink et al., 2012) were found on ECA 1 (7 mb), ECA 3 (35 mb), ECA 15 (19 mb) and ECA 18 (32 mb). Overlapping or closely located (within 2 mb) regions with the study on Swedish horses (Shrestha et al., 2015) were found in the present study on ECA 7 (55 mb), ECA 9 (50 mb), ECA 10 (28, 29 and 36 mb) and ECA X (4, 76, 81 and 82 mb). Other regions among the associated windows

in the present study were not previously detected in either of the two studies on separate data sets, including the second (ECA 1, 153 mb) and third (ECA 22, 25 mb) most associated windows. Possible explanations for the limited overlap of associated regions found between this study and the previous studies on Icelandic horses (Schurink et al., 2012; Shrestha et al., 2015) include limited data size, use of different SNP arrays and exclusion of markers not present on both arrays used, and to some degree differences in environment and collection of phenotypes of horses in the Netherlands and Sweden.

4.2 | Comparison with other studies and validation in Exmoor ponies

Two SNPs found in a genomic region on ECA 1 (151–153 mb) associated with IBH susceptibility in the current GWA study on Icelandic horses were also found to be associated with IBH susceptibility when genotyped in the 280 Exmoor ponies. This region on ECA 1 was, however, not detected in a previous GWA study on IBH (Velie et al., 2016), including 268 of the Exmoor ponies in the present study. In that GWA study, no genome-wide significant associations were found, but a region on ECA 8 appeared to be of interest. The region on ECA 8 was not among the most associated regions in the current study on Icelandic horses, however. One region on ECA 4 (100 mb) in the present study was located nearby a region previously reported in a GWA study of IBH susceptibility in Shetland ponies (99 mb) (Andersson et al., 2012; Klumplerova et al., 2013; Schurink, Ducro, Bastiaansen, Frankena, & van Arendonk, 2013; Schurink et al., 2016). In contrast to several other studies (Andersson et al., 2012; Klumplerova et al., 2013; Schurink et al., 2016), we found no association with the ELA class II region. The coverage in the ELA class II region was low with only 26 SNPs (out of 34 SNPs) remaining after QC. This represents an overlapping number of SNPs between 50K (42 SNPs in ELA class II region) and 70K (65 SNPs in ELA class II region) SNP chip, in region ranging from ~32.5 Mb through ~33.7 Mb (Andersson et al., 2012). In addition to low number of SNPs in ELA class II region, there is an uneven distribution of these SNPs across the ELA region (Equcab 2.0) (figure 8; Shrestha, 2017). In our study, approximately five 1-Mb windows covered the ELA region and surroundings. One of the five 1-mb windows at a genomic region 33 mb explained 0.15% of the genetic variance compared to 0.05% for the same corresponding window in Shrestha et al. (2015). The difference could perhaps stem from the larger samples size in this study (355 horses) compared to 209 horses in the study by Shrestha et al. (2015).

4.3 | A comment on window size

In the present GWA study on Icelandic horses, the same windows explained the highest percentage of genetic variance even when different prior values were used in the Bayesian analysis, which indicates a robustness of the results. It is possible that 1 mb is not the optimal window size, however. When merging the two neighbouring windows on ECA 1 (152 and 153 mb), the resulting 2-mb window explained more of the genetic variance (1.57%) than the combination of the two 1-mb windows separately. These two consecutive windows on the same chromosome may each be explaining a proportion of the same QTL. That the effect of a QTL could be spread across windows has been shown in a simulation study (Sun, Habier, Fernando, Garrick, & Dekkers, 2011), and they also observed more than one QTL within one window. One way to tackle this issue is to define the optimal window size based on LD patterns, but this is challenging due to the variation in LD between SNPs in the genome (Schurink et al., 2012). In future studies, a sliding window approach could be an option to tackle the problem of optimal window size.

4.4 | Associated genomic regions in this study

Two SNPs on ECA 1 (at 152 and 153 mb) were significantly associated with IBH susceptibility in both Icelandic horses and Exmoor ponies suggesting the existence of a common genetic variant of importance for IBH susceptibility across these breeds. The first SNP (at 152 mb) was located in the intronic region of a long non-coding RNA gene (lncRNA; ENSECAG00000036544) and the second SNP (at 153 mb) was located in the intronic region of the aquarius intron-binding spliceosomal factor gene (*AQR*). In human, lncRNAs are known to play a major role in gene regulation, including the regulation of alternative splicing and their dysfunction may have an impact on several pathologies and diseases (Fernandes, Acuña, Aoki, Floeter-Winter, & Muxel, 2019; Romero-Barrios, Legascue, Benhamed, Ariel, & Crespi, 2018). Alternative splicing is prevalent in the mammalian immune system to control lymphocyte homeostasis (Yabas, Elliott, & Hoyne, 2016). Moreover, lncRNAs regulate the inflammatory response following activation of innate immunity and have been related to autoimmune diseases such as psoriasis (Ahn et al., 2016; Heward & Lindsay, 2014). Interestingly, 25% of the annotated genes identified among the 30 1-mb non-overlapping windows most associated with IBH susceptibility correspond to lncRNAs. The SNP found in the genomic region on ECA X (82 mb) associated with IBH susceptibility in the current GWA study on Icelandic horses was not validated in Exmoor ponies. LD patterns can vary between horse breeds and genomic regions (McCue et al., 2012), and Exmoor ponies have been shown to display less genetic variation than the Icelandic horses (Petersen et al., 2013), which may explain the lack of association in this breed. However, it could also possibly be due to different environmental variation between the breeds studied. Overall, the current study highlights a candidate region on ECA 1 with a potential role on

TABLE 2 Risk allele frequency of SNPs associated with IBH susceptibility in Icelandic horses from Sweden, the Netherlands and for the combined two data sets within the 30 1-mb non-overlapping windows that explained the highest percentage of genetic variation

		Combined data		Swedish horses		Dutch horses	
SNP name	ECA	Unaffected	Affected	Unaffected	Affected	Unaffected	Affected
BIEC2_2768	1	0.45	0.55	0.48	0.52	0.40	0.60
BIEC2_65299	1	0.46	0.54	0.47	0.53	0.43	0.57
BIEC2_65455	1	0.46	0.54	0.47	0.53	0.44	0.56
BIEC2_491393	2	0.44	0.56	0.45	0.55	0.43	0.57
BIEC2_776764	3	0.44	0.56	0.45	0.55	0.41	0.43
BIEC2_798106	3	0.46	0.54	0.47	0.53	0.45	0.55
BIEC2_871881	4	0.46	0.54	0.45	0.55	0.48	0.52
BIEC2_878810	4	0.44	0.56	0.43	0.57	0.44	0.56
BIEC2−945312	6	0.44	0.56	0.44	0.56	0.43	0.57
BIEC2_950930	6	0.45	0.55	0.46	0.54	0.44	0.56
BIEC2_1001712	7	0.42	0.58	0.46	0.54	0.34	0.66
BIEC2_1045967	8	0.45	0.55	0.44	0.56	0.46	0.54
BIEC2_1045968	8	0.45	0.55	0.44	0.56	0.46	0.54
BIEC2-1075148	9	0.44	0.56	0.46	0.54	0.41	0.59
BIEC2_1093076	9	0.46	0.54	0.45	0.55	0.46	0.54
BIEC2-111858	10	0.42	0.58	0.43	0.57	0.41	0.59
BIEC2_111947	10	0.42	0.58	0.39	0.61	0.44	0.56
BIEC2_115985	10	0.41	0.59	0.34	0.66	0.46	0.54
BIEC2_133565	11	0.39	0.61	0.42	0.57	0.34	0.66
BIEC2_215349	13	0.43	0.57	0.45	0.55	0.41	0.59
BIEC2_293386	15	0.42	0.58	0.44	0.56	0.40	0.60
BIEC2_326263	15	0.46	0.54	0.47	0.53	0.45	0.55
BIEC2_349357	16	0.49	0.51	0.48	0.52	0.49	0.51
BIEC2_409581	18	0.46	0.54	0.43	0.57	0.49	0.51
BIEC2_410288	18	0.43	0.57	0.46	0.54	0.38	0.62
BIEC2-446571	19	0.46	0.54	0.47	0.53	0.46	0.54
BIEC2_589497	22	0.46	0.54	0.46	0.54	0.44	0.56
BIEC2_1108696	X	0.45	0.55	0.47	0.53	0.41	0.59
BIEC2_1133893	X	0.41	0.59	0.40	0.60	0.44	0.56
BIEC2_1138294	X	0.48	0.52	0.48	0.52	0.48	0.52
BIEC2_1138724	X	0.42	0.58	0.41	0.59	0.43	0.57

TABLE 3 Results for the SNPs validated in Exmoor ponies

			Risk allele frequency				95% Confidence interval	
ECA	SNP name	No. of ponies	Unaffected	Affected	*p*-value	Odds ratio	Lower	Upper
1	BIEC2_65455	277	0.50	0.60	.019	0.65	0.46	0.92
1	BIEC2_65299	278	0.49	0.60	.021	1.52	1.08	2.15
X	BIEC2_1138724	270	0.64	0.65	.887	0.95	0.66	1.37

transcriptional regulatory mechanisms to control genetic differences between IBH susceptibility in horses. Further exploration of this region and the genes included within it in other horse breeds and with a denser marker genetic map will bring a better understanding of the biological mechanisms underpinning IBH susceptibility in horses.

5 | CONCLUSIONS

The polygenic nature of the multifactorial condition equine IBH susceptibility was clearly supported in the present study. A large number of genomic regions, each explaining only a small fraction of the genetic variance, appear to be involved. Besides breed-specific genetic risk factors for equine IBH susceptibility, the results from this study also indicated common genetic variants on ECA 1 associated with IBH susceptibility in Icelandic horses and Exmoor ponies. The role of these loci for equine IBH susceptibility may be interesting to study further.

ACKNOWLEDGEMENTS

We thank all horse owners that answered questionnaires and the breeding organizations for their support. The study was financially supported by the European Union's Seventh Framework Program managed by the Research Executive Agency (REA; http://ec.europa.eu/research/rea [FP7/2007-2013]) under grant agreement no. 606142, and by the European Graduate School of Animal Breeding and Genetics (EGS-ABG).

CONFLICT OF INTERESTS

The authors declare no conflict of interest.

DATA AVAILABILITY STATEMENT

Relevant data for to the Swedish-born Icelandic horses are available at https://datadryad.org/resource/doi:10.5061/dryad.9r161/1. Genotypes for to the Dutch Icelandic horses are available from the authors upon reasonable request.

ORCID

Marina Solé https://orcid.org/0000-0001-7884-1051
Bart J. Ducro https://orcid.org/0000-0002-0866-0199
Susanne Eriksson https://orcid.org/0000-0003-3357-5065

REFERENCES

Ahn, R., Gupta, R., Lai, K., Chopra, N., Arron, S. T., & Liao, W. (2016). Network analysis of psoriasis reveals biological pathways and roles for coding and long non-coding RNAs. *BMC Genomics*, *17*, 841. https://doi.org/10.1186/s12864-016-3188-y

Anderson, G. S., Belton, P., & Kleider, N. (1993). Hypersensitivity of horses in British-Columbia to extracts of native and exotic species of Culicoides (Diptera, Ceratopogonidae). *Journal of Medical Entomology*, *30*(4), 657–663. https://doi.org/10.1093/jmedent/30.4.657

Andersson, L. S., Swinbune, J. E., Meadows, J. R. S., Broström, H., Eriksson, S., Fikse, W. F., … Lindgren, G. (2012). The same ELA class II risk factors confer equine insect bite hypersensitivity in two distinct populations. *Immunogenetics*, *64*(3), 201–208. https://doi.org/10.1007/s00251-011-0573-1

Aulchenko, Y. S., Ripke, S., Isaacs, A., & van Duijn, C. M. (2007). GenABEL: An R package for genome-wide association analysis. *Bioinformatics*, *23*(10), 1294–1296.

Bjornsdottir, S., Sigvaldadottir, J., Brostrom, H., Langvad, B., & Sigurdsson, A. (2006). Summer eczema in exported Icelandic horses: Influence of environmental and genetic factors. *Acta Veterinaria Scandinavica*, *48*, https://doi.org/10.1186/1751-0147-48-3

Eriksson, S., Grandinson, K., Fikse, W. F., Lindberg, L., Mikko, S., Broström, H., … Lindgren, G. (2008). Genetic analysis of insect bite hypersensitivity (summer eczema) in Icelandic horses. *Animal*, *2*(3), 360–365. https://doi.org/10.1017/S1751731107001413

Fadok, V. A., & Greiner, E. C. (1990). Equine insect hypersensitivity - Skin-test and biopsy results correlated with clinical-data. *Equine Veterinary Journal*, *22*(4), 236–240. https://doi.org/10.1111/j.2042-3306.1990.tb04259.x

Fernandes, J. C. R., Acuña, S. M., Aoki, J. I., Floeter-Winter, L. M., & Muxel, S. M. (2019). Long non-coding RNAs in the regulation of gene expression: Physiology and disease. *Non-Coding RNA*, *5*(1), 17. https://doi.org/10.3390/ncrna5010017

Fernando, R. L., & Garrick, D. J. (2009). *GenSel – User manual for a portfolio of genomic selection related analyses*. Retrieved from https://static-content.springer.com/esm/art%3A10.1186%2F1471-2105-12-186/MediaObjects/12859_2010_4655_MOESM1_ESM.PDF

GenABEL project developers (2013). *GenABEL: genome-wide SNP association analysis (Version R package version 1.8-0)*. Retrieved from http://CRAN.R-project.org/package=GenABEL

Habier, D., Fernando, R. L., Kizilkaya, K., & Garrick, D. J. (2011). Extension of the bayesian alphabet for genomic selection. *BMC Bioinformatics*, *186*, 12. https://doi.org/10.1186/1471-2105-12-186

Heward, J. A., & Lindsay, M. A. (2014). Long non-coding RNAs in the regulation of the immune response. *Trends in Immunology*, *35*(9), 408–419. https://doi.org/10.1016/j.it.2014.07.005

Hu, Z.-L., Park, C. A., & Reecy, J. M. (2016). Developmental progress and current status of the Animal QTLdb. *Nucleic Acids Research*, *44*(D1), D827–D833. https://doi.org/10.1093/nar/gkv1233

Kizilkaya, K., Fernando, R. L., & Garrick, D. J. (2010). Genomic prediction of simulated multibreed and purebreed performance using observed fifty thousand single nucleotide polymorphism genotypes. *Journal of Animal Science*, *88*(2), 544–551. https://doi.org/10.2527/jas.2009-2064

Klumplerova, M., Vychodilova, L., Bobrova, O., Cvanova, M., Futas, J., Janova, E., … Horin, P. (2013). Major histocompatibility complex and other allergy-related candidate genes associated with insect bite hypersensitivity in Icelandic horses. *Molecular Biology Reports*, *40*, 3333–3340. https://doi.org/10.1007/s11033-012-2408-z

Lorange, J. B. (2011). WorldFengur - the studbook of origin for the Icelandic horse. *Acta Veterinaria Scandinavica*, *53*(Suppl 1), S5. https://doi.org/10.1186/1751-0147-53-S1-S5

McCue, M. E., Bannasch, D. L., Petersen, J. L., Gurr, J., Bailey, E., Binns, M. M., … Mickelson, J. R. (2012). A high density SNP array for the domestic horse and extant Perissodactyla: utility for association mapping, genetic diversity, and phylogeny studies. *PLOS Genetics*, *8*(1), e1002451. https://doi.org/10.1371/journal.pgen.1002451

McLaren, W., Gil, L., Hunt, S. E., Riat, H. S., Ritchie, G. R. S., Thormann, A., … Cunningham, F. (2016). The Ensembl variant effect predictor. *Genome Biology*, *17*, https://doi.org/10.1186/s13059-016-0974-4

Peeters, L. M., Janssens, S., Brebels, M., & Buys, N. (2015). Genetic parameters and estimated breeding values of insect bite hypersensitivity in Belgian Warmblood horses. *Veterinary Journal*, *206*(3), 420–422. https://doi.org/10.1016/j.tvjl.2015.08.012

Petersen, J. L., Mickelson, J. R., Cothran, E. G., Andersson, L. S., Axelsson, J., Bailey, E., … McCue, M. E. (2013). Genetic diversity in the modern horse illustrated from genome-wide SNP data. *PLoS ONE*, *8*(1), e54997. https://doi.org/10.1371/journal.pone.0054997

Pilsworth, R. C., & Knottenbelt, D. C. (2004). Equine insect hypersensitivity. *Equine Veterinary Education*, *16*(6), 324–325. https://doi.org/10.1111/j.2042-3292.2004.tb00321.x

Quinlan, A. R., & Hall, I. M. (2010). BEDTools: A flexible suite of utilities for comparing genomic features. *Bioinformatics*, *26*(6), 841–842. https://doi.org/10.1093/bioinformatics/btq033

Quinn, P. J., Baker, K. P., & Morrow, A. N. (1983). Sweet itch: Responses of clinically normal and affected horses to intradermal challenge with extracts of biting insects. *Equine Veterinary Journal*, *15*(3), 266–272. https://doi.org/10.1111/j.2042-3306.1983.tb01788.x

R core Team (2013). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from http://www.R-project.org/

Romero-Barrios, N., Legascue, M. F., Benhamed, M., Ariel, F., & Crespi, M. (2018). Splicing regulation by long noncoding RNAs. *Nucleic Acids Research*, *46*(5), 2169–2184. https://doi.org/10.1093/nar/gky095

Schaffartzik, A., Hamza, E., Janda, J., Crameri, R., Marti, E., & Rhyner, C. (2012). Equine insect bite hypersensitivity: What do we know? *Veterinary Immunology and Immunopathology*, *147*(3–4), 113–126. https://doi.org/10.1016/j.vetimm.2012.03.017

Schurink, A., da Silva, V., Velie, B. D., Dibbits, B. W., Crooijmans, R. P. M. A., Francois, I., Ducro, B. J. (2016). *Copy number variations associated with insect bite hypersensitivity in Friesian horses*. Paper presented at the 67th Ann. Meet., Europ. Fed. Anim. Sci. (EAAP), Belfast, Northern Ireland.

Schurink, A., Ducro, B. J., Bastiaansen, J. W. M., Frankena, K., & van Arendonk, J. A. M. (2013). Genome-wide association study of insect bite hypersensitivity in Dutch Shetland pony mares. *Animal Genetics*, *44*(1), 44–52. https://doi.org/10.1111/j.1365-2052.2012.02368.x

Schurink, A., Ducro, B. J., Heuven, H. C. M., & van Arendonk, J. A. M. (2011). Genetic parameters of insect bite hypersensitivity in Dutch Friesian broodmares. *Journal of Animal Science*, *89*(6), 1952–1952. https://doi.org/10.2527/jas.2011-89-6-1952

Schurink, A., van Grevenhof, E. M., Ducro, B. J., & van Arendonk, J. A. M. (2009). Heritability and repeatability of insect bite hypersensitivity in Dutch Shetland breeding mares. *Journal of Animal Science*, *87*(2), 484–490. https://doi.org/10.2527/jas.2008-1129

Schurink, A., Wolc, A., Ducro, B. J., Frankena, K., Garrick, D. J., Dekkers, J. C. M., & van Arendonk, J. A. M. (2012). Genome-wide association study of insect bite hypersensitivity in two horse populations in the Netherlands. *Genetics Selection Evolution*, *44*(31), 1–12. https://doi.org/10.1186/1297-9686-44-31

Shrestha, M. (2017). *Genetics of equine insect bite hypersensitivity and genetic diversity in horses*. Retrieved from https://pub.epsilon.slu.se/13923/

Shrestha, M., Eriksson, S., Schurink, A., Andersson, L. S., Sundquist, M., Frey, R., … Lindgren, G. (2015). Genome-wide association study of insect bite hypersensitivity in Swedish-Born Icelandic Horses. *Journal of Heredity*, *106*(4), 366–374. https://doi.org/10.1093/jhered/esv033

Sun, X., Habier, D., Fernando, R. L., Garrick, D. J., & Dekkers, J. C. (2011). Genomic breeding value prediction and QTL mapping of QTLMAS2010 data using Bayesian Methods. *BMC Proceedings*, *5*(Suppl 3), S13. https://doi.org/10.1186/1753-6561-5-S3-S13

van der Meide, N. M. A., Meulenbroeks, C., van Altena, C., Schurink, A., Ducro, B. J., Wagner, B., … Tijhaar, E. (2012). Culicoides obsoletus extract relevant for diagnostics of insect bite hypersensitivity in horses. *Veterinary Immunology and Immunopathology*, *149*(3–4), 245–254. https://doi.org/10.1016/j.vetimm.2012.07.007

Velie, B. D., Shrestha, M., François, L., Schurink, A., Tesfayonas, Y. G., Stinckens, A., … Lindgren, G. (2016). Using an Inbred Horse Breed in a high density genome-wide scan for genetic risk factors of Insect Bite Hypersensitivity (IBH). *PLoS ONE*, *11*(4), e0152966. https://doi.org/10.1371/journal.pone.0152966

Vychodilova, L., Matiasovic, J., Bobrova, O., Futas, J., Klumplerova, M., Stejskalova, K., … Horin, P. (2013). Immunogenomic analysis of insect bite hypersensitivity in a model horse population. *Veterinary Immunology and Immunopathology*, *152*(3–4), 260–268. https://doi.org/10.1016/j.vetimm.2012.12.013

Wagner, B., Miller, W. H., Morgan, E. E., Hillegas, J. M., Erb, H. N., Leibold, W., & Antczak, D. F. (2006). IgE and IgG antibodies in skin allergy of the horse. *Veterinary Research*, *37*, 813–825. https://doi.org/10.1051/vetres:2006039

Wolc, A., Arango, J., Settar, P., Fulton, J. E., O'Sullivan, N. P., Preisinger, R., … Dekkers, J. C. M. (2012). Genome-wide association analysis and gentic architecture of egg weight and egg uniformity in layer chickens. *Animal Genetics*, *43*, 87–96. https://doi.org/10.1111/j.1365-2052.2012.02381.x

Yabas, M., Elliott, H., & Hoyne, G. F. (2016). The Role of alternative splicing in the control of immune homeostasis and cellular differentiation. *International Journal of Molecular Sciences*, *17*(1), https://doi.org/10.3390/ijms17010003

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

How to cite this article: Shrestha M, Solé M, Ducro BJ, et al. Genome-wide association study for insect bite hypersensitivity susceptibility in horses revealed novel associated loci on chromosome 1. *J Anim Breed Genet*. 2020;137:223–233. https://doi.org/10.1111/jbg.12436