



Genome-wide association study of insect bite hypersensitivity in Dutch Shetland pony mares

A. Schurink*, B. J. Ducro*, J. W. M. Bastiaansen*, K. Frankena[†] and J. A. M. van Arendonk*

*Animal Breeding and Genomics Centre, Wageningen University, P.O. Box 338, 6700 AH, Wageningen, the Netherlands. [†]Quantitative Veterinary Epidemiology Group, Wageningen University, P.O. Box 338, 6700 AH, Wageningen, the Netherlands

Summary

Insect bite hypersensitivity (IBH) is the most common allergic disease present in horses worldwide. It has been shown that IBH is under genetic control, but the knowledge of associated genes is limited. We conducted a genome-wide association study to identify and quantify genomic regions contributing to IBH in the Dutch Shetland pony population. A total of 97 cases and 91 controls were selected and matched on withers height, coat colour and pedigree to minimise the population stratification. A blood sample was collected from participating Shetland pony mares, their IBH phenotype was scored and the owner filled in a questionnaire. A total of 40 021 single-nucleotide polymorphisms (SNPs) were fitted in a univariable logistic model fitting an additive effect. Analysis revealed no effects of population stratification. Significant associations with IBH were detected for 24 SNPs on 12 chromosomes [$-\log_{10}(P\text{-value}) > 2.5$]. Odds ratios of allele substitution effects of the unfavourable allele were between 1.94 and 5.95. The most significant SNP was found on chromosome 27, with an odds ratio of 2.31 and with an allele frequency of the unfavourable allele of 0.72 in cases and 0.53 in controls. Genome-wide association studies on additional horse populations are desired to validate the identified associations, to identify the genes involved in IBH and to develop genomic tools to decrease IBH prevalence.

Keywords genome-wide association study, insect bite hypersensitivity, matched case-control design

Introduction

Insect bite hypersensitivity (IBH) is the most common allergic disease in horses and found in many countries throughout the world (Anderson *et al.* 1988; Littlewood 1998; Steinman *et al.* 2003). Observed prevalence of IBH is, for instance, 8.1% in Swedish-born Icelandic horses (Eriksson *et al.* 2008), 8.8% in Shetland pony mares (Schurink *et al.* 2009) and 18.2% in Friesian broodmares (Schurink *et al.* 2011). Sensitive horses develop a severe itch after exposure to bites of certain midges; *Culicoides* spp. are reported as the most common cause of IBH (Riek 1954; Fadok & Greiner 1990). *Culicoides* spp. are most active at sunset (Van der Rijt *et al.* 2008) and usually disappear during winter in countries with a temperate

climate. The presence of *Culicoides* spp. is highly dependent on environmental conditions (Takken *et al.* 2008).

The welfare of affected horses is seriously reduced. Common clinical symptoms in affected horses are hair loss, bald spots, excoriation, crusting, scaling and thickening of the skin (Van den Boom *et al.* 2008). Severe cases might suffer from open wounds and secondary infections. Use of affected horses, like showing and riding, can be restricted owing to discomfort and disfigurement (Gortel 1998). Owners may encounter costs related to measures taken to prevent or limit the development of clinical symptoms, although fully effective measures are still unavailable (Van den Boom *et al.* 2008). The commercial value of affected horses is reduced (Broström *et al.* 1987).

Scoring clinical symptoms is still the gold-standard for detecting IBH (Baselgia *et al.* 2006). Analysing the severity of clinical symptoms hardly affects heritability estimates (Eriksson *et al.* 2008; Schurink *et al.* 2009) and therefore seems unnecessary, probably because only a small percentage of horses within a population are affected and preventive measures strongly influence the

Address for correspondence

A. Schurink, Animal Breeding and Genomics Centre, Wageningen University, P.O. Box 338, 6700 AH Wageningen, the Netherlands.
E-mail: Anouk3.Schurink@wur.nl

Accepted for publication 13 March 2012

severity of clinical symptoms. Specific *in vitro* and *in vivo* tests were developed to detect IBH (e.g. Baselgia *et al.* 2006; Wagner *et al.* 2008; Sloet van Oldruitenborgh-Oosterbaan *et al.* 2009), but test results often only support a clinical diagnosis. A case-control design, therefore, seems useful for treating the disorders like IBH that are difficult to phenotype. However, caution is needed when selecting cases and controls. Matching of cases with controls for certain features like ancestry is desired to avoid population stratification and thereby limit unwanted spurious associations in genome-wide association studies (e.g. Karlsson & Lindblad-Toh 2008; McCarthy *et al.* 2008).

A genetic component of IBH has been confirmed in various horse breeds (Unkel *et al.* 1987; Eriksson *et al.* 2008; Schurink *et al.* 2009, 2011). Insect bite hypersensitivity heritability estimates on the liability scale vary from 0.16 in Friesian horses (Schurink *et al.* 2011) to 0.33 in Swedish-born Icelandic horses (Eriksson *et al.* 2008). Knowledge of genes contributing to the genetic variation in IBH is limited. Recently, Andersson *et al.* (2012) conducted a candidate gene approach and identified an association between the equine leucocyte antigen (ELA) class II region and IBH in Icelandic horses and Exmoor ponies. The same allele (*COR112:274*) was associated with IBH in both breeds, and homozygosity across the entire ELA class II region increased the risk of developing IBH. Recently, the horse's (*Equus caballus*) genome was sequenced (Wade *et al.* 2009), which enabled the development of a genotyping chip with 50 000 single-nucleotide polymorphisms (SNPs) (Illumina Inc.), thereby presenting us with the opportunity to conduct a genome-wide association study to increase our knowledge about genes involved in IBH. Knowledge of genomic associations with IBH could increase the efficiency of breeding for decreased IBH prevalence through genomic selection using only significant markers or all markers. Moreover, an increased knowledge of genes involved in IBH will contribute to a better understanding of the underlying biology and thereby to improved prevention, diagnosis and therapy of IBH. The aims of our study were to identify and quantify genomic associations with IBH in the Dutch Shetland pony population, using a genome-wide association approach.

Materials and methods

Study design

Shetland pony mares were selected according to a matched case-control design to minimise the effects of population stratification (the presence of multiple subgroups within the population that differ in both allele frequency and disease prevalence; for example, McCarthy *et al.* 2008). Cases were defined as mares showing IBH clinical symptoms (mild to severe), whereas controls were free of symptoms. Cases and controls were matched on

various factors (Table 1) to minimise potential population stratification. Mares were grouped by the Dutch Shetland Pony Studbook into four categories according to their withers height: mini (86 cm), small (87–92 cm), middle (93–98 cm) and tall (99–107 cm).

The number of available cases and controls determines the power to detect a specific locus; the power also depends on several (unknown) factors like inheritance, allele frequency and relative risk. We aimed at 150 cases and 150 controls to achieve reasonable power to detect genomic regions highly associated with IBH, but managed to collect data on only 97 cases and 91 controls.

Selection of cases and controls

In 2009, IBH phenotypes of Shetland pony mares were recorded by studbook inspectors during foal inspections. Owners of mares that scored positive (i.e. cases) were contacted later that year and asked to participate in our study.

Table 1 Distribution [% (number)] of characteristics of Dutch Shetland pony mares over cases and controls investigated in 2009 and 2010.

Trait	Case	Control	Total
Number of mares	97	91	188
Year of scoring			
2009	79.4 (77)	79.1 (72)	79.3 (149)
2010	20.6 (20)	20.9 (19)	20.7 (39)
Month of scoring			
September	53.6 (52)	57.1 (52)	55.3 (104)
October	43.3 (42)	39.6 (36)	41.5 (78)
November	3.1 (3)	3.3 (3)	3.2 (6)
Veterinarian			
1	94.8 (92)	95.6 (87)	95.2 (179)
2	5.2 (5)	4.4 (4)	4.8 (9)
Pedigree			
Number of sires	80	81	125
Number of dams	94	88	176
Age, years			
Mean (SD)	7.0 (4.3)	8.5 (4.5)	7.7 (4.4)
Range	0–21	4–22	0–22
Withers height category			
Mini	26.8 (26)	25.3 (23)	26.0 (49)
Small	15.5 (15)	16.5 (15)	16.0 (30)
Middle	30.9 (30)	29.7 (27)	30.3 (57)
Tall	26.8 (26)	28.5 (26)	27.7 (52)
Coat colour			
Bay	3.1 (3)	5.5 (5)	4.3 (8)
Bay paint	1.0 (1)	2.2 (2)	1.6 (3)
Black	50.5 (49)	52.7 (48)	51.6 (97)
Black paint	6.2 (6)	4.4 (4)	5.3 (10)
Black roan	3.1 (3)	1.1 (1)	2.1 (4)
Blue dun	–	1.1 (1)	0.5 (1)
Chestnut	24.7 (24)	24.2 (22)	24.5 (46)
Chestnut paint	8.3 (8)	6.6 (6)	7.5 (14)
Chestnut roan	2.1 (2)	1.1 (1)	1.6 (3)
Golden dun	1.0 (1)	–	0.5 (1)
Palomino	–	1.1 (1)	0.5 (1)

Participation consisted of filling in a questionnaire and allowing a veterinarian to collect blood from their affected mare(s). Enrolment of cases continued in 2010 when participants were recruited by the Dutch Shetland Pony Studbook through a publication in their magazine and on their website. This change of approach was needed because IBH phenotypes of Shetland pony mares were no longer recorded by studbook inspectors in 2010. The data collected from 2009 to 2010 contained 97 cases (Table 1).

Mares that had been exposed to *Culicoides* spp. but did not develop IBH clinical symptoms were used as controls. Controls were selected based on the following matching criteria: withers height, coat colour, proximity to a case and pedigree. Withers height clearly affects IBH in Shetland pony mares (Schurink *et al.* 2009), although its reason remains unclear. Coat colours might differ in attractiveness to *Culicoides* spp., as Braverman *et al.* (1983) found that horses with dark coat colour were more often affected by IBH. Although age at onset of IBH varies greatly (Anderson *et al.* 1988; Eriksson *et al.* 2008; Van den Boom *et al.* 2008), average age at onset is considered to be between 2 and 4 years of age. Therefore, controls were required to be at least 4 years of age and at least 1–2 years at risk of developing IBH clinical symptoms. Proximity to a case was required, that is, on the same premises, to ensure exposure of controls to *Culicoides* spp., and thereby increases the reliability of the control phenotype, as sensitive individuals would have developed clinical symptoms in the presence of *Culicoides* spp. Controls were also genealogically matched by selecting paternal half-sibs of cases wherever possible to minimise potential population stratification (e.g. Hirschhorn & Daly 2005; Karlsson & Lindblad-Toh 2008). We collected blood from 91 controls (Table 1).

Data came from 99 mares (52.7% of the data) descending from 89 sires with only case(s) or control(s) among their offspring, and 89 half-sib mares (47.3% of the data) descending from 36 sires with both case(s) and control(s) among their offspring (Table 2). Cases and controls descended from a similar number of ancestors (Table 1) (average inbreeding coefficient of cases was 0.047 and of controls was 0.046).

Cases and controls were preferably present for 2 or more years on their current premises to ensure constant management practices (i.e. environment) and exposure to *Culicoides* spp. for a reasonable time. Most Shetland pony mares (93.1%) were present for 2 or more years on the current premises, 5.3% of the mares were present <2 years, and information on length of presence was lacking for 1.6% of the mares.

Participation

In total, 115 (4.1%) out of 2809 mares scored IBH positive during foal inspections in 2009. Out of these 115

Table 2 Number of case and control offspring per sire.

Number of case offspring per sire	Number of control offspring per sire	Number of sires
1	0	38
2	0	4
3	0	2
0	1	43
0	2	2
1	1	25
1	2	3
1	3	1
2	1	3
2	2	2
3	1	1
3	2	1

affected mares, 77 (67.0%) were included as cases in our study. In 2010, data on 20 additional cases were collected. The 97 cases and 91 controls belonged to 69 owners. Almost three-quarters (71.0%) of the owners had both case(s) and control(s), 23.2% of owners had only case(s), and 5.8% of owners had only control(s). Exposure to *Culicoides* spp. in these latter controls was plausible because of current or previous presence of cases that were unsuitable for our study, for instance, because they were males, which are more often kept indoors and therefore potentially less exposed to *Culicoides* spp.

Blood sample collection and phenotypes

Participating owners were visited by a veterinarian and researcher to take blood samples, to score phenotypes of mares and to conduct an IBH-related questionnaire. Blood collection from Dutch Shetland pony mares was approved by the Board on Animal Ethics and Experiments from Wageningen University (experiment 2009055 and 2010109). Most blood samples were collected in 2009 during September and October (Table 1). Blood was collected by jugular venipuncture in two 10-ml EDTA(K₂) Terumo® VenoSafe blood collection tubes per mare. Owners answered questions about the history and management of IBH to increase the information about and reliability of case phenotypes. Questions included the age at onset of IBH, the month during which symptoms began and disappeared, and which preventive and therapeutic measures were taken. During the visit, an experienced veterinarian scored the mare's IBH phenotype using a score sheet for hair loss, hyperkeratosis, scaling, crusting, excoriation, open wounds and secondary infections on the crest, base of the tail, hindquarters, abdomen, shoulder, neck and head. Severity of symptoms was graded in four categories from mild (category 1) to severe (category 4). However, severity of symptoms was not used during our analysis. To ensure uniform classification, most cases and controls (95.2%) were scored by the same veterinarian

(Table 1). The remaining five cases and four controls were scored by a second veterinarian.

Genotypes and quality control

DNA was extracted from leucocytes in the blood samples. Lack of degradation of DNA was confirmed by electrophoresis, carried out using a 1- μ l sample on 1.5% multi-purpose agarose gels in Tris–borate–EDTA buffer at 120 volts.

Genotypes from all Shetland pony mares were obtained using the Illumina® EquineSNP50 Genotyping BeadChip containing 54 602 SNPs with an average SNP spacing of 43.2 kb.

One homozygote genotype was coded as 0, the heterozygote genotype as 1 and the other homozygote genotype as 2. An additive effect of SNPs was assumed in the regression analysis. Consequently, missing genotypes were centred, that is, they were set to the average value of the genotypes obtained for a single SNP [e.g. in the cases of 50 genotypes 0 (AA), 80 genotypes 1 (AB) and 30 genotypes 2 (BB) and for 28 missing genotypes, the missing genotypes were replaced with the average, which was calculated by multiplying genotype frequencies with genotype values ($0.875 = \frac{50}{160} \times 0 + \frac{80}{160} \times 1 + \frac{30}{160} \times 2$)]. The GenABEL package in R (Aulchenko *et al.* 2007) was used to perform the quality control applying the check.marker function. The first step of the iterative procedure of the check.marker function excluded SNPs based on per SNP statistics (call rate ≤ 0.80 , minor allele frequency $\leq 2\%$, no exclusion based on *P*-value Hardy–Weinberg equilibrium) using the summary.snp.data function. The second step excluded SNPs based on per individual and between individual statistics (call rate ≤ 0.90 , no exclusion based on identity-by-state matrix) using the perid.summary function. The procedure was applied recursively until no further SNPs were eliminated. Genotyping failed for 364 SNPs. Monomorphic SNPs ($n = 7957$), SNPs with a call rate ≤ 0.80 ($n = 172$) and SNPs with minor allele frequency $\leq 2\%$ ($n = 6088$) were removed. Of 54 602 SNPs, 40 021 (73.3%) were used in the analyses. Call rate per sample varied from 0.930 to 0.999: 173 Shetland pony mares (92.0%) had a call rate above 0.99, 13 Shetland pony mares (6.9%) had a call rate between 0.95 and 0.99, and only two Shetland pony mares (1.1%) had a call rate between 0.90 and 0.95.

After the genome-wide association analysis, quality of SNP clusters and clustering of all significant SNPs was checked using cluster plots, Illumina® GenTrain and GenCall scores. When quality of the SNP cluster and clustering was poor (i.e. visual observation of cluster plots, Illumina® GenTrain score < 0.55 , Illumina® GenCall score < 0.7), SNPs were removed ($n = 2$ out of 26). Based on *P*-values Hardy–Weinberg equilibrium in controls (all > 0.15), no SNPs were removed.

Statistical analysis

To check whether the matching of cases with controls had been successful, the relation between IBH (case or control, binary trait) and matching factors was assessed by means of logistic regression. Fixed effects of withers height category, coat colour, month and year of scoring, region and veterinarian as well as the covariates of the inbreeding coefficient and age of the mare were tested for significance using the LOGISTIC procedure (SAS Institute, Inc.). Factors were tested in both univariable and multivariable models.

For the genome-wide association analysis, a univariable logistic model was used that fitted SNP genotypes as a main effect. The linear predictor (η_i) of the single-SNP model was:

$$\eta_i = \mu + b_1 \times \text{SNP}_i$$

where μ = the population mean and b_1 = the allele substitution effect of SNP_i , where $i = 0, 1$ or 2 (for 40 021 SNPs). Estimated allele substitution effects (b_1) were transformed to odds ratios ($\text{OR} = e^{b_1}$) for interpretation. An additional analysis was conducted on part of the genome with an extended model, which included a random polygenic effect. The covariance matrix of the random effect was based either on pedigree information or on relationship matrix (=genomic kinship matrix $\times 2$) calculated from SNP genotypes. The additional analysis was conducted to test whether matching of cases with controls on pedigree removed the effect of polygenic variance on SNP effects. For all analyses, we used the !CYCLE option in the ASREML program (Gilmour *et al.* 2009).

Genomic kinship

Genomic kinship among Shetland pony mares was computed to determine whether cases and controls differed in kinship (i.e. unsuccessful matching of pedigree). Differences in kinship between cases and controls might result from stratification and thereby lead to false-positive associations. Genomic kinship among mares was computed using the ibs function of the R package GenABEL (Aulchenko *et al.* 2007) as:

$$f_{ij} = \sum_k \frac{(x_{i,k} - p_k)(x_{j,k} - p_k)}{(p_k(1 - p_k))},$$

where f_{ij} = genomic kinship (identity-by-state) matrix between Shetland pony mare i and j , based on $k = 38\,226$ autosomal SNPs (X-chromosome excluded). $x_{i,k}$ or $x_{j,k}$ are the genotypes of the i th or j th Shetland pony mare for SNP k and p_k is the frequency of the allele (top strand).

The genomic kinship matrix was transformed to a distance matrix to perform classical multidimensional scaling (Gower 1966), which returned the first two principal components. These principal components were plotted to

visualise distances between the Shetland pony mares and more specifically between the group of cases and of controls.

Results

Phenotypes and questionnaire results

Blood samples from 97 cases and 91 controls were collected and included in our analysis. Age at sampling ranged from foal to 22 years (Table 1). Age at the onset of IBH varied from foal to 15 years, with the majority of the cases (59.8%) starting to show clinical symptoms at the age of 2, 3 or 4 years. For 16.5% of the cases, the year of data collection was the first year with clinical symptoms; for 44.3%, it was the second or third year, and 34.0% of the cases had clinical symptoms for 4 or more years. Most Shetland pony mares developed clinical symptoms during spring (71.1%) and some during summer (20.6%). Often, in 85.1%, symptoms disappeared during autumn. In some cases (7.4%), symptoms disappeared during summer. There was no clear pattern in severity of symptoms of affected mares in consecutive years. Many Shetland pony mares (67.0%) suffered from moderate to severe itch. Crest (99.0%) and base of the tail (82.5%) were mostly affected. In some cases, (11.3%) hindquarters were affected. The most frequently observed clinical symptoms were hair loss (97.9%) and thickening/ridged skin (86.6%), whereas crusting (22.7%), open wounds (8.2%) and scaling (7.2%) were observed less frequently. Observed clinical symptoms and most questionnaire results agreed with the typical course of IBH (e.g. Pilsworth & Knottenbelt 2004). Owners indicated that the clinical symptoms affected the use of their Shetland pony mares mainly in breeding and showing. Preventive and therapeutic measures (e.g. eczema blanket) were used in 78.4% of the IBH cases.

Success of matching to remove fixed and family effects

Mean age of Shetland pony mares differed significantly (P -value < 0.05) between cases and controls (Table 1). The difference in mean age was, however, a result of the experimental setup. For other risk factors, matching was applied. Statistical analysis revealed that these risk factors had no significant effect on IBH prevalence (data not shown). These results indicate that matching of cases with controls was successful in removing difference between cases and controls.

The multi-dimensional scaling plot (Fig. 1) visualised genetic distances between cases and controls based on the genomic kinship (identity-by-state) matrix. Within the population, subgroups based on withers height categories were identified, which is also reflected in Fig. 1. Shetland pony mares from mini and small withers height categories

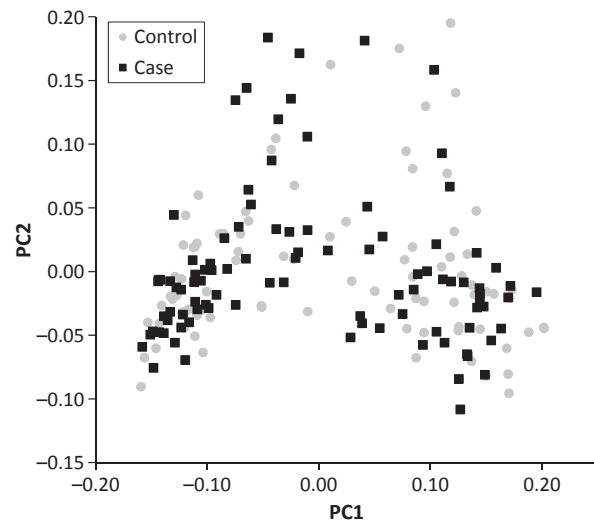


Figure 1 Multidimensional scaling plot of the distance between Shetland pony mares labelled as case (■) or control (●). Each point in the plot corresponds to one individual and indicates the distance between mares represented by the first two principal components (PC1 and PC2), which are based on the genomic kinship matrix.

grouped together (bottom left-hand side Fig. 1: $PC1 < 0$ and $PC2 < 0.05$) as did Shetland pony mares from middle and tall withers height categories (bottom right-hand side Fig. 1: $PC1 > 0$ and $PC2 < 0.05$). The distribution of mares over the two-dimensional space showed a high degree of overlap between cases and controls. The multi-dimensional scaling plot based on genomic kinship (Fig. 1) did not reveal population stratification between cases and controls.

Finally, the heritability for IBH in our sample of cases and controls was estimated with an animal model including either a pedigree or genomic-based covariance matrix. Results from both analyses (data not shown) indicated that genetic variance was absent ($h^2 < 0.01$), which reflects that matching of cases with controls on pedigree successfully removed the effect of polygenic variation on IBH.

Genome-wide association study

In total, 24 SNPs on 12 chromosomes were significantly associated [$-\log_{10}(P\text{-value}) > 2.5$] with IBH (Table 3, Fig. 2). The most significant SNP was found on chromosome 27, with the frequency of the unfavourable allele at 0.72 for cases and 0.53 for controls. The odds ratio of allele substitution effect for this SNP was 2.31 (Table 3) meaning that the odds of IBH in mares with a heterozygous genotype is increased 2.31 times ($= e^{0.8351}$) compared to mares homozygous for the favourable allele. Further, the odds of IBH in mares homozygous for the unfavourable allele is increased 5.31 times ($= e^{2 \times 0.8351}$) compared to mares homozygous for the favourable allele. Odds ratios for the significant SNPs were between 1.94 and 5.95 (Table 3).

Table 3 Single-nucleotide polymorphisms (SNPs) showing most significant association with insect bite hypersensitivity in Dutch Shetland pony mares using a single-SNP model.

SNP name	ECA ¹	Position (bp)	P-value	OR (95% CI) ²	Allele frequency ³	
					Cases	Controls ⁴
<i>BIEC2-50714</i> ⁵	1	116 022 321	0.0030	2.31 (1.34–3.98)	0.27	0.14
<i>BIEC2-50725</i> ⁵	1	116 037 195	0.0029	2.53 (1.39–4.63)	0.22	0.09
<i>BIEC2-81773</i> ⁶	1	171 091 112	0.0019	5.95 (1.96–18.1)	0.98	0.89
<i>BIEC2-771602</i> ⁵	3	8 370 895	0.0030	1.96 (1.26–3.05)	0.59	0.43
<i>BIEC2-771605</i> ^{5,7}	3	8 392 689	0.0031	2.00 (1.27–3.15)	0.57	0.42
<i>BIEC2-786998</i> ⁷	3	64 632 334	0.0023	3.71 (1.62–8.52)	0.14	0.04
<i>BIEC2-877664</i>	4	99 194 126	0.0017	2.94 (1.52–5.69)	0.20	0.08
<i>BIEC2-953011</i>	6	47 258 572	0.0023	2.18 (1.33–3.57)	0.77	0.63
<i>BIEC2-980932</i>	7	13 582 164	0.0030	2.08 (1.29–3.35)	0.82	0.68
<i>BIEC2-1004427</i>	7	61 379 694	0.0026	2.31 (1.35–3.97)	0.29	0.16
<i>BIEC2-1027098</i> ⁷	8	8 138 630	0.0029	2.64 (1.41–4.95)	0.21	0.10
<i>BIEC2-1062146</i> ⁵	8	70 132 295	0.0031	2.05 (1.28–3.28)	0.80	0.66
<i>BIEC2-1062154</i> ⁵	8	70 132 806	0.0031	2.05 (1.28–3.28)	0.80	0.66
<i>BIEC2-1102005</i>	9	66 153 825	0.0021	2.04 (1.30–3.20)	0.63	0.47
<i>BIEC2-149137</i>	11	32 010 755	0.0012	2.48 (1.44–4.25)	0.28	0.13
<i>BIEC2-149356</i> ⁵	11	32 753 122	0.0010	2.93 (1.56–5.51)	0.22	0.09
<i>BIEC2-149592</i> ⁵	11	33 040 164	0.0014	2.90 (1.53–5.52)	0.21	0.09
<i>BIEC2-532511</i>	20	41 520 518	0.0023	2.01 (1.29–3.13)	0.44	0.28
<i>BIEC2-532524</i>	20	41 531 646	0.0009	2.17 (1.38–3.39)	0.45	0.27
<i>BIEC2-705454</i>	27	13 198 799	0.0004	2.31 (1.46–3.63)	0.72	0.53
<i>BIEC2-814955</i>	30	4 484 018	0.0010	2.21 (1.39–3.52)	0.75	0.58
<i>BIEC2-825651</i> ^{5,6}	30	20 973 099	0.0006	5.43 (2.10–14.0)	0.97	0.87
<i>BIEC2-825652</i> ^{5,6}	30	20 973 284	0.0006	5.43 (2.10–14.0)	0.97	0.87
<i>BIEC2-1114599</i>	X	23 168 684	0.0021	1.94 (1.28–2.94)	0.53	0.36

¹*Equus caballus* autosome.²Odds ratio of allele substitution effect of unfavourable allele calculated as e^{b_1} ; 95% confidence interval between parentheses.³Allele frequency of unfavourable allele.⁴No deviations from Hardy–Weinberg equilibrium (all P -values >0.15) were detected in controls.⁵Linkage disequilibrium (r^2) between SNP on same chromosome is >0.8 .⁶Only two out of three genotypes were present (i.e. homozygotes for one allele and heterozygotes).⁷Homozygote for one allele absent in controls.

Discussion

The aims of our study were to identify and quantify genomic associations with IBH in the Dutch Shetland pony population. For this purpose, we applied a genome-wide association study on a matched case–control sample of the Dutch Shetland pony population. To achieve reasonable power to detect genomic regions associated with IBH, we aimed at the collection of samples from 150 cases and 150 controls. Unfortunately, fewer cases than expected were available, which lowered the power to detect associations. We managed to collect blood samples from a large proportion of IBH cases in 2009, but IBH prevalence in 2009 (4.1%) was low compared to years before (8.6% in 2003, 9.5% in 2005 and 8.5% in 2006; Schurink *et al.* 2009). We collected 20 additional samples in 2010. As our sample size was limited, power to detect QTL with small effect was low and effects of significant QTL tend to be overestimated, which is known as the ‘Beavis effect’ (Beavis 1998; Xu 2003).

Genomic associations with IBH are potentially confounded by other effects relevant to IBH (e.g. region,

withers height category). However, these effects were nullified in our study by using a matched case–control design. Age of the mare was significantly associated with IBH but was not included in the analysis of genomic associations, as it was introduced by the study design and potentially could lower the power to detect associations between IBH and genotypes. Data comprised 89 paternal half-sibs (47.3% of the data). Previous genetic analysis of IBH in the Shetland pony population revealed significant genetic variance corresponding to an estimated heritability of 0.24 (Schurink *et al.* 2009). A similar genetic analysis of our case–control sample revealed no significant genetic variance. Also, the multi-dimensional scaling plot based on the genomic kinship (Fig. 1) showed a high degree of overlap between cases and controls over the two-dimensional space. These results indicate that cases and controls had been successfully matched to minimise effects of population stratification and unwanted spurious associations in the genome-wide association study.

Misclassification of case or control phenotypes – misclassification bias (e.g. McCarthy *et al.* 2008) – is expected to reduce allele frequency differences between cases and con-

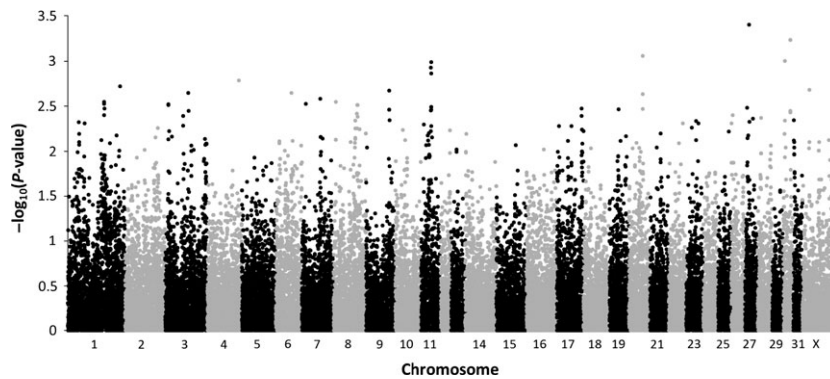


Figure 2 Genome-wide association plot of insect bite hypersensitivity in Dutch Shetland pony mares. The $-\log_{10}(P\text{-value})$ is plotted against the chromosomal location for each SNP tested across all chromosomes. Chromosome number is indicated on the x-axis, where the chromosomes are plotted in alternating black and grey for clarity.

trols and therefore reduces power of the association study (e.g. Edwards *et al.* 2005; Pearson & Manolio 2008). To diminish potential misclassification bias, Dutch Shetland pony mares were selected according to a strict protocol.

Candidate gene approach or genome-wide association studies on IBH in horses are limited. *Serine peptidase inhibitor, Kazal type 5 (SPINK5)* in humans has been related to disorders with equivalent clinical symptoms as IBH. Andersson *et al.* (2009), therefore, considered *SPINK5* on ECA14 to be a candidate gene, but found that it was not associated with IBH in Swedish-born Icelandic horses. Wagner *et al.* (2001) identified polymorphisms in the *immunoglobulin heavy constant epsilon (IGHC)* gene on ECA24 that encodes for the heavy chain constant region domains of equine IgE, which mediates IBH (Wagner *et al.* 2006). Wagner *et al.* (2001) suggested that polymorphisms in the *IGHC* gene might impose genetic variability to allergic diseases in horses, but they did not study associations. Hořín *et al.* (2004) carried out an association study and identified and tested polymorphic markers (both microsatellites and SNPs) in various immune response-related genes to identify associations with two important horse infections: *R. equi* and *L. intracellularis*. Alleles of, for instance, the microsatellite locus *HMS01* were significantly associated with *R. equi* infection. Immune response-related genes [e.g. *IL10* (*interleukin 10*)] were found within the associated regions, although fine-mapping is needed to assign candidate genes (Hořín *et al.* 2004). Locus *HMS01* on ECA15 previously has been identified to be associated with IBH (Marti *et al.* 2005 in Chowdhary & Raudsepp 2008). Above-mentioned candidate genes and regions associated with IBH, however, did not coincide with the genomic regions associated with IBH in our study.

The equine major histocompatibility complex (MHC), or ELA, recognises many foreign molecules and thereby evokes an immune response (Bailey *et al.* 2000). Equine ELA class I and II regions are positioned on ECA20 and were associated with allergen-specific immunoglobulin E (IgE) levels against two moulds (Eder *et al.* 2001) and IBH

(Halldórsdóttir *et al.* 1991; Marti *et al.* 1992; Andersson *et al.* 2012). Eder *et al.* (2001) investigated associations between ELA class I antigens and allergen-specific IgE levels against two moulds and found that horses carrying ELA A1 had a significantly higher IgE level against *Aspergillus f.* and that horses carrying ELA A14 had significantly lower IgE levels against both *Alternaria a.* and *Aspergillus f.* Moreover, serological research on ELA class II antigens identified a significant difference in distribution of the Be 8 antigen between IBH cases and controls (Halldórsdóttir *et al.* 1991). Also, the W23 antigen segregated with IBH predisposition in specific families (Marti *et al.* 1992). Recently, genomic research identified an association between the ELA class II region and IBH in two distinct horse breeds (Andersson *et al.* 2012). Although we found associations with IBH on ECA20 (Table 3), the identified genomic region was approximately 8 Mb away from ELA regions. Therefore, we cannot confirm nor exclude the impact of ELA on IBH. This requires a closer investigation of ELA regions.

From our study, using visual inspection of the Manhattan plot (Fig. 2) to identify interesting genomic regions, it appeared that 24 SNPs on 12 chromosomes were associated with IBH [$-\log_{10}(P\text{-value}) > 2.5$]. After calculating the false discovery rate using *q*-values (Storey & Tibshirani 2003), no SNP exceeded a false discovery rate of 0.5 (data not shown). Odds ratios of allele substitution effects of the unfavourable allele were between 1.94 and 5.95. Population attributable fractions (PAF), which indicate the reduction in IBH prevalence when eliminating the unfavourable allele, were calculated following Bertram *et al.* (2007). Based on the odds ratios and allele frequency of the unfavourable allele in controls, PAFs varied from 0.11 to 0.42. Hence, the IBH prevalence in the Dutch Shetland pony population, being 7.6% (Schurink *et al.* 2010), could be reduced to values between 4.4 and 6.8% (depending on PAF per SNP) when the unfavourable allele is eliminated. However, this reduction in prevalence is overestimated as the odds ratio overestimates relative risk

(Schmidt & Kohlman 2008). Second, the allele frequency of the unfavourable allele in the population estimated from our controls might not represent the real allele frequency in the Dutch Shetland pony population.

Genomic selection using only significant markers or all markers could increase the efficiency of breeding for decreased IBH prevalence. The increased efficiency results from increased accuracies of estimated breeding values and decreased generation interval (e.g. Meuwissen *et al.* 2001; Hayes & Goddard 2010) compared to traditional breeding using own performance or progeny information. Own performance for IBH depends highly on exposure to *Culicoides* spp. and can be recorded only from an age of 3 years. Progeny testing of sires will improve accuracy of breeding values but will increase generation interval, as sufficient progeny information will be available only late in life. Genomic selection can be applied at a young age because it requires collection of only marker genotypes but not of phenotypic information on the selection candidate or its progeny. Phenotypic information already collected in the Shetland pony population is exploited through its genomic association with markers.

Horses from various breeds in many countries worldwide are affected with IBH and prevalence varies between breeds, even within a similar environment. Breed differences in susceptibility to IBH are expected, although a common genetic background is anticipated. Genome-wide association analyses of IBH in additional samples from the Dutch Shetland pony population or in samples from other horse breeds are needed to validate the identified associations and to narrow down the genomic regions.

Acknowledgements

We are grateful to Bert Dibbits for extracting DNA and genotyping all samples, to veterinarians Drs. Remmen and Drs. Dalemans for scoring phenotypes and taking blood samples, and to Betteke van Eijk for fast data exchange during busy times. We thank the Dutch Shetland Pony Studbook (Zutphen) for providing part of the data. Many thanks to all Shetland mare owners who contributed to our research. This work was financially supported by technology foundation STW (Utrecht), the Dutch federation of horse breeding Vereniging Koepel Fokkerij (Ermelo) and the Dutch (bio)pharmaceutical company Artu Biologics part of ALK-Abelló (Almere).

References

- Anderson G.S., Belton P. & Kleider N. (1988) The hypersensitivity of horses to *Culicoides* bites in British Columbia. *Canadian Veterinary Journal* **29**, 718–23.
- Andersson L.S., Höglström C., Mikko S., Eriksson S., Grandinson K., Broström H., Frey R., Sundquist M. & Lindgren G. (2009) Polymorphisms in *SPINK5* do not associate with insect bite hypersensitivity in Icelandic horses born in Sweden. *Animal Genetics* **40**, 790–1.
- Andersson L.S., Swinbune J.E., Meadows J.R.S. *et al.* (2012) The same ELA class II risk factors confer equine insect bite hypersensitivity in two distinct populations. *Immunogenetics* **64**, 201–8.
- Aulchenko Y.S., Ripke S., Isaacs A. & van Duijn C.M. (2007) GenABEL: an R library for genome-wide association analysis. *Bioinformatics* **23**, 1294–6.
- Bailey E., Marti E., Fraser D.G., Antczak D.F. & Lazary S. (2000) Immunogenetics of the horse. In: *The Genetics of the Horse* (Ed. by A.T. Bowling & A. Ruvinsky), pp. 123–55. CABI Publishing, New York.
- Baselgia S., Doherr M.G., Mellor P., Torsteinsdóttir S., Jermann T., Zurbriggen A., Jungi T. & Marti E. (2006) Evaluation of an *in vitro* sulphidoleukotriene release test for diagnosis of insect bite hypersensitivity in horses. *Equine Veterinary Journal* **38**, 40–6.
- Beavis W.D. (1998) QTL analyses: power, precision, and accuracy. In: *Molecular Dissection of Complex Traits* (Ed. by A.H. Paterson), pp. 145–62. CRC Press, New York.
- Bertram L., McQueen M.B., Mullin K., Blacker D. & Tanzi R.E. (2007) Systematic meta-analyses of Alzheimer disease genetic association studies: the AlzGene database. *Nature Genetics* **39**, 17–23.
- Braverman Y., Ungar-Waron H., Frith K., Adler H., Danieli Y., Baker K.P. & Quinn P.J. (1983) Epidemiological and immunological studies of sweet itch in horses in Israel. *Veterinary Record* **112**, 521–4.
- Broström H., Larsson Å. & Troedsson M. (1987) Allergic dermatitis (sweet itch) of Icelandic horses in Sweden: an epidemiological study. *Equine Veterinary Journal* **19**, 229–36.
- Chowdhary B.P. & Raudsepp T. (2008) The horse genome derby: racing from map to whole genome sequence. *Chromosome Research* **16**, 109–27.
- Eder C., Curik I., Brem G., Crameri R., Bodo I., Habe F., Lazary S., Sölkner J. & Marti E. (2001) Influence of environmental and genetic factors on allergen-specific immunoglobulin-E levels in sera from Lipizzan horses. *Equine Veterinary Journal* **33**, 714–20.
- Edwards B.J., Haynes C., Levenstien M.A., Finch S.J. & Gordon D. (2005) Power and sample size calculations in the presence of phenotype errors for case/control genetic association studies. *BMC Genetics* **6**, 18.
- Eriksson S., Grandinson K., Fikse W.F., Lindberg L., Mikko S., Broström H., Frey R., Sundquist M. & Lindgren G. (2008) Genetic analysis of insect bite hypersensitivity (summer eczema) in Icelandic horses. *Animal* **2**, 360–5.
- Fadok V.A. & Greiner E.C. (1990) Equine insect hypersensitivity: skin test and biopsy results correlated with clinical data. *Equine Veterinary Journal* **22**, 236–40.
- Gilmour A.R., Cullis B.R. & Thompson R. (2009) *ASREML Update: What's new in Release 3.00*. VSN International Ltd., Hemel Hempstead, UK.
- Gortel K. (1998) Equine parasitic hypersensitivity. A review. *Equine Practice* **20**, 14–6.
- Gower J.C. (1966) Some distance properties of latent root and vector methods used in multivariate analysis. *Biometrika* **53**, 325–8.
- Halldórsdóttir S., Lazary S., Gunnarsson E. & Larsen H.J. (1991) Distribution of leucocyte antigens in Icelandic horses affected

- with summer eczema compared to non-affected horses. *Equine Veterinary Journal* **23**, 300–2.
- Hayes B. & Goddard M. (2010) Genome-wide association and genomic selection in animal breeding. *Genome* **53**, 876–83.
- Hirschhorn J.N. & Daly M.J. (2005) Genome-wide association studies for common diseases and complex traits. *Nature Review Genetics* **6**, 95–108.
- Hořín P., Smola J., Matiašovic J. *et al.* (2004) Polymorphisms in equine immune response genes and their associations with infections. *Mammalian Genome* **15**, 843–50.
- Karlsson E.K. & Lindblad-Toh K. (2008) Leader of the pack: gene mapping in dogs and other model organisms. *Nature Review Genetics* **9**, 713–25.
- Littlewood J.D. (1998) Incidence of recurrent seasonal pruritus ('sweet itch') in British and German shire horses. *Veterinary Record* **142**, 66–7.
- Marti E., Gerber H. & Lazary S. (1992) On the genetic basis of equine allergic diseases: II. Insect bite dermal hypersensitivity. *Equine Veterinary Journal* **24**, 113–7.
- Marti E., Glowatzki-Mullis M.L., Curik I., Torsteinsdottir S. & Binns M.M. (2005) Investigating the genetic background for insect bite hypersensitivity in Icelandic horses. 6th International Equine Gene Mapping Workshop, Dublin, Ireland.
- McCarthy M.I., Abecasis G.R., Cardon L.R., Goldstein D.B., Little J., Ioannidis J.P.A. & Hirschhorn J.N. (2008) Genome-wide association studies for complex traits: consensus, uncertainty and challenges. *Nature Review Genetics* **9**, 356–69.
- Meuwissen T.H.E., Hayes B.J. & Goddard M.E. (2001) Prediction of total genetic value using genome-wide dense marker maps. *Genetics* **157**, 1819–29.
- Pearson T.A. & Manolio T.A. (2008) How to interpret a genome-wide association study. *Journal of the American Medical Association* **299**, 1335–44.
- Pilsworth R.C. & Knottenbelt D.C. (2004) Equine insect hypersensitivity. *Equine Veterinary Education* **16**, 324–5.
- Riek R.F. (1954) Studies on allergic dermatitis (Queensland Itch) of the horse: The aetiology of the disease. *Australian Journal of Agricultural Research* **5**, 109–29.
- Schmidt C.O. & Kohlman T. (2008) When to use the odds ratio or the relative risk? *International Journal of Public Health* **53**, 165–7.
- 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**, 484–90.
- Schurink A., Vogelzang R.H. & Ducro B.J. (2010) Relation between insect bite hypersensitivity and body condition score in Dutch Shetland mares. *Proceedings of the 9th World Congress on Genetics Applied to Livestock Production*, Leipzig, Germany, 1–6 August 2010.
- 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**, 1286–93.
- Sloet van Oldruitenborgh-Oosterbaan M.M., van Poppel M., de Raat I.J., van den Boom R. & Savelkoul H.F.J. (2009) Intradermal testing of horses with and without insect bite hypersensitivity in the Netherlands using an extract of native *Culicoides* species. *Veterinary Dermatology* **20**, 607–14.
- Steinman A., Peer G. & Klement E. (2003) Epidemiological study of *Culicoides* hypersensitivity in horses in Israel. *Veterinary Record* **152**, 748–51.
- Storey J.D. & Tibshirani R. (2003) Statistical significance for genome-wide experiments. *Proceedings of the National Academy of Sciences* **100**, 9440–5.
- Takken W., Verhulst N., Scholte E.-J., Jacobs F., Jongema Y. & van Lammeren R. (2008) The phenology and population dynamics of *Culicoides* spp. in different ecosystems in The Netherlands. *Preventive Veterinary Medicine* **87**, 41–54.
- Unkel M., Simon D., Mayer M. & Sommer H. (1987) Studies on the genetic basis of sweet itch in Island horses. *Zeitschrift für Tierzüchtung und Züchtungsbiologie* **104**, 217–30 (in German).
- Van den Boom R., Ducro B. & Sloet van Oldruitenborgh-Oosterbaan M.M. (2008) Identification of factors associated with the development of insect bite hypersensitivity in horses in the Netherlands. *Tijdschrift voor Diergeneeskunde* **133**, 554–9.
- Van der Rijt R., Van den Boom R., Jongema Y. & Sloet van Oldruitenborgh-Oosterbaan M.M. (2008) *Culicoides* species attracted to horses with and without insect hypersensitivity. *Veterinary Journal* **178**, 91–7.
- Wade C.M., Giulotto E., Sigurdsson S. *et al.* (2009) Genome sequence, comparative analysis, and population genetics of the domestic horse. *Science* **326**, 865–7.
- Wagner B., Siebenkotten G., Radbruch A. & Leibold W. (2001) Nucleotide sequence and restriction fragment length polymorphisms of the equine *Ce* gene. *Veterinary Immunology and Immunopathology* **82**, 193–202.
- 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**, 813–25.
- Wagner B., Childs B.A. & Erb H.N. (2008) A histamine release assay to identify sensitisation to *Culicoides* allergens in horses with skin hypersensitivity. *Veterinary Immunology and Immunopathology* **126**, 302–8.
- Xu S. (2003) Theoretical basis of the Beavis effect. *Genetics* **165**, 2259–68.