Mamm Genome (2009) 20:504–515
DOI 10.1007/s00335-009-9214-5

A whole-genome scan for recurrent airway obstruction in Warmblood sport horses indicates two positional candidate regions

June E. Swinburne · Helen Bogle · Jolanta Klukowska-Rötzler · Michaela Drögemüller · Tosso Leeb · Elizabeth Temperton · Gaudenz Dolf · Vincent Gerber

Received: 9 April 2009 / Accepted: 17 August 2009 / Published online: 16 September 2009

Abstract Recurrent airway obstruction (RAO), or heaves, is a naturally occurring asthma-like disease that is related to sensitisation and exposure to mouldy hay and has a familial basis with a complex mode of inheritance. A genome-wide scanning approach using two half-sibling families was taken in order to locate the chromosome regions that contribute to the inherited component of this condition in these families. Initially, a panel of 250 microsatellite markers, which were chosen as a well-spaced, polymorphic selection covering the 31 equine autosomes, was used to genotype the two half-sibling families, which comprised in total 239 Warmblood horses. Subsequently, supplementary markers were added for a total of 315 genotyped markers. Each half-sibling family is focused around a severely RAO-affected stallion, and the phenotype of each individual was assessed for RAO and related signs, namely, breathing effort at rest, breathing effort at work, coughing, and nasal discharge, using an owner-based questionnaire. Analysis using a regression method for half-sibling family structures was performed using RAO and each of the composite clinical signs separately; two chromosome regions (on ECA13 and ECA15) showed a genome-wide significant association with RAO at $P < 0.05$. An additional 11 chromosome regions showed a more modest association. This is the first publication that describes the mapping of genetic loci involved in RAO. Several candidate genes are located in these regions, a number of which are interleukins. These are important signalling molecules that are intricately involved in the control of the immune response and are therefore good positional candidates.

Introduction

Recurrent airway obstruction (RAO), or heaves, is a naturally occurring asthma-like disease in horses that is related to sensitisation and exposure to mouldy hay. RAO has a familial basis and shows a complex mode of inheritance (Gerber et al. 2009; Marti et al. 1991; Ramseyer et al. 2007). Affected horses are typically middle-aged or older and show increased breathing effort, exercise intolerance, cholinergic bronchospasm, coughing, airway hyperreactivity, and neutrophil and mucus accumulation in the airways (Gerber et al. 2004; Robinson et al. 1995). Once a susceptible horse is sensitised it needs to be managed carefully in order to minimise exposure to allergens. In severe cases corticosteroids and/or bronchodilators are required to control clinical signs.

A genetic predisposition for RAO has been demonstrated with genetic epidemiological investigations in full- and half-sibling groups. The familial predisposition for

Electronic supplementary material The online version of this article (doi:10.1007/s00335-009-9214-5) contains supplementary material, which is available to authorized users.

J. E. Swinburne · H. Bogle · E. Temperton
Animal Health Trust, Newmarket, Suffolk CB8 7UU, UK

J. E. Swinburne (✉)
Centre for Preventive Medicine, Animal Health Trust, Lanwades Park, Kentford, Newmarket, Suffolk CB8 7UU, UK
e-mail: june.swinburne@aht.org.uk

J. Klukowska-Rötzler · V. Gerber
Equine Clinic, Department of Veterinary Clinical Studies, Vetsuisse-Faculty, University of Berne, Langgass-strasse 124, 3012 Berne, Switzerland

J. Klukowska-Rötzler · M. Drögemüller · T. Leeb · G. Dolf
Institute of Genetics, Vetsuisse-Faculty, University of Berne, Bremgartenstrasse 109A, 3001 Berne, Switzerland

equine chronic lower airway disease was first described by Schaeper (1939), and subsequently by Koch (1957) and Gerber (1989). Marti et al. (1991) then showed, in both Warmblood and Lipizzan horses, that the risk for developing moderate to severe equine chronic lower airway disease was significantly increased in offspring with one affected parent and even more with both parents affected. In Warmblood horses, the prevalence and clinical manifestation of moderate to severe equine chronic lower airway disease was found to be increased in offspring of two affected sires (referred to as sires 1 and 2) when compared to an unrelated control group and a group of maternal half-siblings (Ramseyer et al. 2007). The risk of offspring from sires 1 and 2 developing RAO (HOARSI grade 3 or 4) was increased 4.1- and 5.5-fold, respectively. That study developed the HOARSI (Horse Owner Assessed Respiratory Signs Index), a composite score based on owner-observed coughing frequency, nasal discharge, breathing effort, and performance. Multivariate regression analysis further showed that hay feeding and increasing age also increased the severity of the clinical signs. Jost et al. (2007) subsequently demonstrated association and linkage of HOARSI with microsatellite markers in the chromosome 13q13 region in the family of sire 1 but not in the family of sire 2. These results suggested a genetic background with locus heterogeneity for equine RAO. Furthermore, segregation analyses clearly showed the presence of a major gene playing a role in RAO (Gerber et al. 2009). In the family of sire 1 its mode of inheritance is autosomal recessive, whereas in the family of sire 2 it is autosomal dominant. Although the expression of RAO is influenced by age and exposure to hay, these findings suggest a strong, but complex genetic background for RAO.

There are many similarities between RAO and human asthma and RAO may prove to be a good animal model for asthma. Like RAO, asthma is a complex disease in which genetic and environmental factors interact to generate a continuous clinical phenotype which ranges from unaffected to severely affected. Many of the clinical symptoms and pathological features observed in RAO, as described above, are very reminiscent of asthma. Furthermore, the population structure of the horse breed under investigation here is better suited for the identification of genotypes associated with the RAO phenotype than a human population, as horse breeds can be considered isolated closed populations with limited heterogeneity. The availability of large families is another advantage of horses compared to humans.

The study of human asthma is hampered by its multigenic nature, genetic heterogeneity across populations, and variable disease expression. However, the development of high-resolution SNP chips, which can genotype up to 1,000,000 SNPs simultaneously, has facilitated association studies over the last few years and led to the identification of over 100 positional candidate genes for asthma (for review see Malerba and Pignatti 2005; Wills-Karp and Ewart 2004; Zhang et al. 2008). This plethora of candidate genes is no surprise considering the involvement of many different cell types and molecules in asthma pathogenesis, yet in only a modest number of cases has there been a replication of these associations (reviewed in Bossé and Hudson 2007). In comparison to man, domesticated animals are usually kept in closed populations derived from a small number of founders. As a consequence, the number of segregating haplotypes is likely to be much smaller and linkage disequilibrium much more extensive, which increases the chance of identifying genes involved in RAO. The use of related individuals should further increase the chances of detecting genetic variants against a homogeneous background.

Based on the results of a segregation analysis in high-prevalence RAO families (Gerber et al. 2009) and on the locus heterogeneity suggested by the findings of Jost et al. (2007), we hypothesize that several major genes are responsible for RAO. In this study we use a genome-scan approach to identify the chromosome regions in which genetic variants are located that contribute to inherent RAO susceptibility. We used the same Warmblood families as were used in Jost et al. (2007) and Gerber et al. (2009) and a panel of microsatellite markers that spans the 31 horse autosomes with an average spacing of 10 Mb. This study expands greatly the study of Jost et al. (2007), which examined only four microsatellites in a 6-Mb region encompassing *IL4R* on ECA13, with a genome-wide panel that includes 22 markers spanning ECA13, one of which was included in Jost et al. (2007).

Materials and methods

Interview phenotype information and Horse Owner Assessed Respiratory Signs Index (HOARSI)

HOARSI 1-4 (healthy, mild, moderate, and severe clinical signs, respectively) has previously been described in detail (Ramseyer et al. 2007). Briefly, horse owners were contacted by phone and only horses with clinical signs that had persisted for at least 2 months were included in the study. All horses were 5 years or older with at least a 12-month history of hay-feeding. All horses included were born and lived in Switzerland and were registered in the Swiss Warmblood Studbook. A standardized questionnaire was used to gather information from the horse owners on the animals' history of chronic coughing [absent, occasional (intermittent occurrence of coughing with periods without cough of 1 week or more), regular (coughing consistently

at least every week but no more than once a day), frequent (coughing every day, several times)], increased breathing effort at rest (absent or present), increased breathing effort after exercise (absent or present), and nasal discharge (absent or present). The correlations between the component phenotypes were significant but weak to moderately strong (Table 1). This information was combined into a HOARSI 1-4. The classification refers to the period when the horses were exhibiting the most severe clinical signs. While HOARSI 1 comprises unaffected individuals (severity class 1), RAO in this study is represented in severity class 3, which comprises HOARSI 3 and 4 individuals (Ramseyer et al. 2007). Validation on 33 offspring of sire 1 and 36 offspring of sire 2 using comprehensive clinical examination showed that HOARSI 3 and 4 individuals in exacerbation are fully consistent with the RAO phenotype (Gerber and Laumen, unpublished results). The first questionnaire data were collected in 2005; rechecks in 2006 and in 2008 of HOARSI performed on over 200 questionnaires revealed a high consistency of categorisation with a rate of misclassification of less than 1% (Gerber et al. 2009; Ramseyer et al. 2007).

Sire families

The high-prevalence half-sibling sire families have also been described in detail elsewhere (Ramseyer et al. 2007). We selected two Warmblood sires that showed obvious clinical signs of respiratory distress (nostril flare, increased abdominal lift, or increased respiratory rate) and airway obstruction when stabled in stalls with straw bedding and fed hay and showed remission of these signs when stabled in a barn complex especially adapted to the requirements of RAO patients (bedding of dust-free shavings; haylage feeding). The stallions' RAO status was unambiguously documented by the stud veterinary service and confirmed by examination performed by the equine clinic of the University of Berne. The stallions, which were used mainly (>90%) in artificial insemination service, both stood at the same stud, but due to their popularity their semen was used throughout the Swiss Warmblood breeding population. The offspring of sire 1 included in this study were born from 118 different mares and included 103 paternal half-sibling offspring, 7 full-sibling pairs, 2 full-sibling triplets, and 7 maternal half-siblings. The offspring of sire 2 were born from 103 different mares and included 93 paternal half-sibling offspring, 4 full-sibling pairs, and 6 maternal half-siblings. Since both sires stood at the same stud and were used in the same breeding population, it was not surprising that there was considerable overlap between the dams-sires: more than 65% of the groups (2-15 offspring) with common grandsires on the dams' side were distributed across both families. There were no regional differences in proportion of affected progeny. Sire family 1 comprised 130 offspring of which 50 were considered unaffected and assigned to HOARSI 1, 28 were classified as HOARSI 2, 34 as HOARSI 3 and 18 as HOARSI 4. Sire family 2 comprised 107 offspring of which 34 were considered unaffected and assigned to HOARSI 1, 33 were classified as HOARSI 2, 27 as HOARSI 3, and 13 as HOARSI 4. Peripheral blood samples were taken from all individuals (sires and offspring) for DNA extraction.

Power of the study

The power of the study was calculated for each family separately (Table 2) using the power calculation facility in GRID QTL (Seaton et al. 2006). The heritability of RAO was previously calculated as extremely large ($h^2 = 1$) (Gerber et al. 2009) in both family 1 and family 2, where hay feeding (the major environmental risk factor) was an inclusion criterion for the study. Consequently, power calculations have been performed here with $h^2 = 1$, 0.9, and 0.8. The QTL heritability (q^2) was set at either 0.1 or 0.2, and the significance levels were set at 0.05 and 0.01. The method uses stochastic gene-drop to simulate identity-by-descent (IBD) sharing at a fully informative marker between all members of the pedigree and asymptotic theory

Table 1 Correlations between component phenotypes

	Coughing	Nasal discharge	Increased breathing effort at work	Increased breathing effort at rest
Coughing	1	0.52	0.19	0.15
		<0.0001	<0.0001	0.0020
Nasal discharge	0.52	1	0.23	0.23
	<0.0001		<0.0001	<0.0001
Increased breathing effort at work	0.19	0.23	1	0.20
	<0.0001	<0.0001		<0.0001
Increased breathing effort at rest	0.15	0.24	0.20	1
	0.0020	<0.0001	<0.0001	

First entry is the correlation coefficient (Spearman) and the second entry is the *P* value

Table 2 Power calculations for families 1 and 2

Total heritability of RAO (h^2)	QTL heritability (q^2)	Family 1			Family 2		
		Expected test statistic	Significance level	Power	Expected test statistic	Significance level α	Power
0.8	0.1	4.36	0.05	0.58	3.75	0.05	0.51
0.8	0.1	4.36	0.01	0.31	3.75	0.01	0.25
0.8	0.2	9.19	0.05	**0.89**	7.80	0.05	**0.83**
0.8	0.2	9.19	0.01	0.70	7.80	0.01	0.61
0.9	0.1	4.88	0.05	0.63	4.19	0.05	0.56
0.9	0.1	4.88	0.01	0.36	4.19	0.01	0.30
0.9	0.2	10.40	0.05	**0.92**	8.83	0.05	**0.88**
0.9	0.2	10.40	0.01	0.77	8.83	0.01	0.68
1	0.1	5.64	0.05	0.70	4.85	0.05	0.63
1	0.1	5.64	0.01	0.43	4.85	0.01	0.36
1	0.2	12.13	0.05	**0.95**	10.32	0.05	**0.92**
1	0.2	12.13	0.01	**0.84**	10.32	0.05	0.77

Power has been calculated in families 1 and 2 separately using various settings for total heritability, QTL heritability, and significance level. Power calculations over 80% are highlighted in bold

(Williams and Blangero 1999) to calculate the noncentrality parameter (NCP) conditional on the IBD sharing pattern. Power is calculated from the mean NCP across simulated IBD patterns.

Microsatellite marker panel

A panel of 286 markers, which had been previously developed as a comprehensive set with which to perform a low-density scan of all 31 horse autosomes and the X chromosome, was used for this study (Supplementary Table 1 lists those markers from which usable data were obtained). The location of each marker was identified on the horse genome sequence (http://www.broad.mit.edu/node/318) by comparing the sequence of the unique region flanking the microsatellite with the second assembly of the horse genome sequence (http://www.ensembl.org/Equus_caballus/Info/StatsTable) using BLAT. For all markers a unique match to the genome sequence was observed. In a handful of cases there were minor discrepancies between the original linkage map (Swinburne et al. 2006) and the physical position of the marker on the genome sequence. In only five instances (TKY344, TKY491, TKY806, TKY785, and TKY315) did the relative positions differ to a significant degree (Supplementary Table 1). In one instance, for ECA25, the linkage map and the corresponding genome sequence were inverted relative to each other. For positioning the markers for the QTL analysis described here, the physical position, in Mb, of the markers on the genome sequence was used.

Genotyping using the microsatellite marker panel

Genomic DNA was obtained from peripheral blood using a High Pure PCR Template Preparation Kit (Roche, Basel, Switzerland). PCR reactions were organised as multiplexes with three markers in each reaction. Four PCR reactions, each utilising a different fluorescent dye, were pooled together post-PCR to form a panel of 12 markers for analysis. This multiplexing and pooling regime was adopted to reduce costs and increase efficiency. Fluorescent labeling of the PCR amplicon was achieved using the 3-primer methodology described by Schuelke (2000), which provides a cost-effective means of labeling PCR fragments fluorescently. All aliquoting, PCR set up, and pooling steps were performed using a Thermo Scientific Matrix PlateMate 2 × 2 automated pipetting workstation (Thermo Fisher Scientific, Waltham, MA).

PCRs were performed in 384-well PCR plates (Axygen Scientific, Union City, CA) using 6-µl reaction volumes. All PCRs comprised 20 ng genomic DNA, 0.75 unit Ampli*Taq* Gold (Applied Biosystems, Foster City, CA), 1 × GeneAmp PCR buffer II (Applied Biosystems), 1.5 mM $MgCl_2$, and 200 µM each dNTP. Then 2.5 pmol of reverse, 1 pmol of tailed-forward, and 5 pmol of the labeled universal primer (either 6-FAM, VIC, NED, or PET) were added to the reaction. An MJ Tetrad PCR cycler (Bio-Rad Laboratories, Hercules, CA) and a PCR program of 94°C for 10 min, followed by 30 cycles of 94°C for 1 min, 55°C for 1 min, and 72°C for 1 min, followed by 8 cycles of 94°C for 1 min, 50°C for 1 min, and 72°C for 1 min, and then 72°C for 30 min was used. The four PCR

reactions belonging to each panel of 12 markers were pooled together and an aliquot of this used for analysis. Reactions were stored at –20°C prior to analysis on an ABI 3100 (Applied Biosystems) according to the manufacturer's instructions. Dye set G5 was used in conjunction with the LIZ500 size standard.

Genotyping data were collected and analysed with GeneMapper ver. 4.0 (Applied Biosystems). Alleles were assigned to bins and given a suitable integer name by the software. Data were exported into Excel and Mendelian inheritance confirmed for all genotypes. Any genotypes that did not display Mendelian inheritance were discarded.

Genotyping data cleaning

The error detection tool in MERLIN (Abecasis et al. 2002) was used to identify suspect genotypes by virtue of their implying double recombination. Forty individual genotypes were identified and removed in this way. In addition, the haplotyping tool in MERLIN was used to detect any markers where there was any gross disagreement with neighbouring markers implied by numerous double recombinants. Three markers were identified in this way and removed; it is likely that mislabeled data or reagents were to blame, or, alternatively, incorrect positioning of the marker on the chromosome.

QTL analysis

The genotyping data were analysed using QTL Express (http://qtl.cap.ed.ac.uk/) (Seaton et al. 2002) and GRID QTL (http://www.gridqtl.org.uk/) (Seaton et al. 2006), a program for the analysis of quantitative trait data from outbred populations, including collections of half-sibling families, that uses a regression approach. The software is accessed remotely using a Web-based user interface. Information content (IC) along each chromosome was also provided by this program; this information is given in Supplementary Table 1. A single QTL model was fitted at 1-Mb steps along the chromosome and initially 1000 permutations were run. Age was set as a covariate and sex as a fixed effect. HOARSI, "breathing at rest," "breathing at work," coughing, and nasal discharge were each examined as separate phenotypes. A QTL signal was suggested by peaks where the F statistic exceeded chromosome-wide significance at $P = 0.05$. For chromosomes where this occurred, the significance levels were recalculated by running 5000 permutations using the grid-based system offered by GRID QTL. All analyses were performed with each family separately. F statistics were also compared with genome-wide significance levels at $P = 0.05$, calculated by running 1000 permutations.

Addition of extra markers in selected chromosome regions

Once the initial panel of markers was analysed, 65 additional microsatellite markers were chosen for seven selected chromosomes to supplement the data. These were regions where preliminary analysis had suggested significant QTL signals. Suitable published markers were identified from those assigned a unique match with the horse genome sequence. Markers that were polymorphic in one or both of the sires were then assembled into groups of three and informative families were genotyped. These data were added to the initial panel of markers and reanalysed using GRID QTL.

Linkage analysis

As a complementary examination of the data, linkage analysis was performed for the phenotype HOARSI using the computer program package FASTLINK (Cottingham et al. 1993). Two-point linkage analysis was carried out using the option LODSCORE. For the disease locus, allele frequencies and penetrances for four age classes (1-9, 10-13, 14-19, and 20-25 years) were set according to the results of the two separate segregation analyses performed in the two sire half-sibling families (Gerber et al. 2009). In each half-sibling family, allele frequencies of the markers were estimated based on the alleles transmitted by the dams.

Haplotype analysis

MERLIN was used to identify the most likely haplotypes in the two sires; the *–best* mode was used (Abecasis et al. 2002). For each offspring the inherited paternal haplotype was identified. In cases in which it was uncertain which allele had been inherited from the sire it was assumed that minimal recombination had occurred. The region of the chromosome that was genome-wide significant ($P < 0.05$) in QTL analysis was then selected for further scrutiny. Offspring with each haplotype in this selected region were counted for each HOARSI class. A Fisher's exact test was performed using an online calculator (http://www.graphpad.com/quickcalcs/contingency1.cfm); two-sided p values were calculated to assess significance.

Results

Genotyping

The preliminary panel of 286 markers was used to genotype the two sires and their half-sibling offspring. Of these

markers 36 were discarded, leaving a total of 250. The discarded markers included 10 X chromosome markers (since the inheritance of alleles from a sire by his offspring is under investigation here, these markers are of no interest); 17 markers that were either monomorphic in all samples ($n = 2$) or did not genotype well ($n = 15$); 2 markers that could not be positioned conclusively on the physical map; 4 markers segregating null alleles; and 3 markers that disagreed markedly with their neighbours, implying numerous double recombination events.

Two offspring were identified as being incorrectly assigned to the cited sire and their data were discarded, leaving a total of 239 genotyped individuals that included the two sires and 237 offspring (130 from family 1 and 107 from family 2). Of the 250 markers for which data were available, an average of 88% of the samples was scored. In addition, from the 239 DNA samples that were genotyped, an average of 87% of the markers was scored. The average spacing of the 250 utilised markers was 8.5 Mb, with a range of 0.03–42.97 Mb. Details of the markers used in this study can be found in Supplementary Table 1.

Next we selected a supplementary panel of 65 informative markers (indicated in Supplementary Table 1) for chromosomes for which putative associations with HOARSI were suggested (ECA3, 13, 15, 16, 21, 22, and 27). Fifty-seven of these markers had already been described in the literature and eight were developed *de novo*. All were genotyped and added to the data panel. Of the 315 markers ultimately genotyped, 206 (65.4%) were heterozygous and therefore informative in sire 1 and 212 (67.3%) in sire 2. Forty markers (12.7%) were homozygous in both sires.

QTL analysis

The panel of markers used in this study had an average information content across the genome of 0.69 [see Supplementary Table 1 for individual average IC values per chromosome which ranged from 0.45 (ECA25) to 0.9 (ECA13)]. The QTL analysis resulted in 13 signals significant at the chromosome-wide level at $P < 0.05$ (Table 3). The most significant signals for HOARSI were at 26 cM on ECA13 (in family 1) and at 46 cM on ECA15 (in family 2). Both of these were significant genome-wide at $P < 0.05$.

Figure 1 illustrates the significant signals obtained on ECA13 (for HOARSI) and ECA15 (for HOARSI and "breathing at work"). On ECA13 the signal for HOARSI reaches genome-wide significance ($P < 0.05$) in family 1 at 26 Mb; this signal exceeds chromosome-wide significance ($P < 0.05$) over 22 Mb (7–29 Mb). On ECA15 the signal for HOARSI reaches genome-wide significance ($P < 0.05$) in family 2 from 44 to 57 Mb, with the highest signal at 46 Mb; this signal exceeds chromosome-wide significance ($P < 0.05$) over 31 Mb (38–69 Mb). The signal for "breathing at work" exceeds chromosome-wide significance ($P < 0.05$) on ECA15 from 41 to 51 Mb.

There are less significant signals on other chromosomes; these are listed in Table 3. For example, on ECA21 there are signals that exceed chromosome-wide significance ($P < 0.05$) in both family 1 and family 2 (for "breathing at rest"). The signal in family 1 stretches over 36–43 Mb, and the signal in family 2 stretches over 38 Mb (6–44 Mb), with the highest signal at 16 Mb. For all three of the chromosomes described here, it is possible that there is more than one QTL on each chromosome.

Haplotype analysis

The results of haplotype analysis of ECA13 and ECA15 are given in Table 4. The region of the chromosome that was genome-wide significant ($P < 0.05$) in QTL analysis was selected for analysis (marked in italics in Table 4.a.i. and a.ii). The analysis of the data using a Fisher exact test (Table 4.b.i. and b.ii) indicates that for the ECA13 region, inheritance in family 1 is very significant when HOARSI 1 is compared with HOARSI 4. In addition, for the ECA15 region, inheritance in family 2 is extremely significant when HOARSI 1 is compared with HOARSI 4.

Linkage analysis of the phenotype HOARSI

The average LOD score over all markers was 0.09 for both sires. The highest LOD score for sire 1 was 1.82 on ECA13 and the highest LOD score for sire 2 was 2.47 on ECA15. Sire 1, but not sire 2, also showed a moderately increased LOD score of 1.75 on ECA21. All other increased LOD scores (on ECA6, 27, and 28) were less than 1.5. Complete FASTLINK results (theta and LOD scores) for all informative markers are on file (see Supplementary Table 2).

Discussion

The aim of this study was to locate chromosome regions associated with a history of clinical signs of RAO exhibited in two Warmblood horse families. The families consist of two severely RAO-affected stallions and their half-sibling offspring. The use of affected stallions causes a limitation for the study in that these sires will likely be homozygous for some of the regions controlling RAO; these regions will not be identified by this study. Each of the offspring was assessed for RAO symptoms and graded according to the HOARSI system (Ramseyer et al. 2007). QTLs, significant chromosome-wide ($P < 0.05$), were identified on 11 chromosomes (ECA6, 7, 12, 13, 14, 15, 16, 17, 21, 25, and

Table 3 Significant QTL analysis results from families 1 and 2 for the RAO composite phenotype HOARSI and related clinical signs of RAO, obtained using QTL Express

ECA	Phenotype	Family	*F* statistic	Position of peak of signal (Mb)	Chromosome significance ($P = 0.05$)	All markers falling within chromosome significant ($P < 0.05$) QTL region	Genome significance ($P = 0.05$)
6	Breathing at work	2	6.88	66	5.71	COR070, TKY323	12.90
6	HOARSI	2	9.38	72	5.83	COR070, TKY323, TKY284	13.21
7	Breathing at rest	2	9.19	63	5.27	TKY005, TKY338, TKY282	10.95
12	HOARSI	1	5.78	11	4.87	UMNe191	12.85
13	HOARSI▼	1	12.90	26	6.40	COR069, VHL161, TKY031, TKY2450, UM030, TKY1219, VHL047, TKY581, TKY371, AHT095, TKY3325, TKY421	12.85
14	Breathing at rest	1	8.55	58	5.66	UM010	12.41
15	Breathing at work	2	7.57	42	6.55	UMNe219, UMNe156, TKY1724, TKY2036, TKY688	12.90
15	HOARSI▼	2	13.69	46	6.49	UMNe219, UMNe156, TKY1724, TKY2036, TKY688, UMNe608, TKY926, ASB002, TKY885	13.21
16	Breathing at rest	1	5.73	78	5.60	TKY341	12.41
16	Breathing at rest	2	7.58	42	6.04	TKY349, TKY1522, TKY2022	10.95
17	Nasal Discharge	1	8.06	70	6.32	LEX067, TKY379, AHT098	13.53
21	Breathing at rest	2	7.56	16	5.83	TKY806, COR073, TKY824, ECA21_001, ECA21_002, ECA21_003, ECA21_005, AHT 077, ECA21_007	10.95
21	Breathing at rest	1	6.05	38	6.03	ECA21_002	12.41
25	Nasal Discharge	1	7.59	2	4.24	TKY316	13.53
26	Coughing	2	10.02	14	5.13	UMNe066, LEX044	13.89

Significance levels chromosome- and genome-wide are shown at $P = 0.05$. The two genome-wide significant signals are indicated here with an inverted triangle and are illustrated in Fig. 1

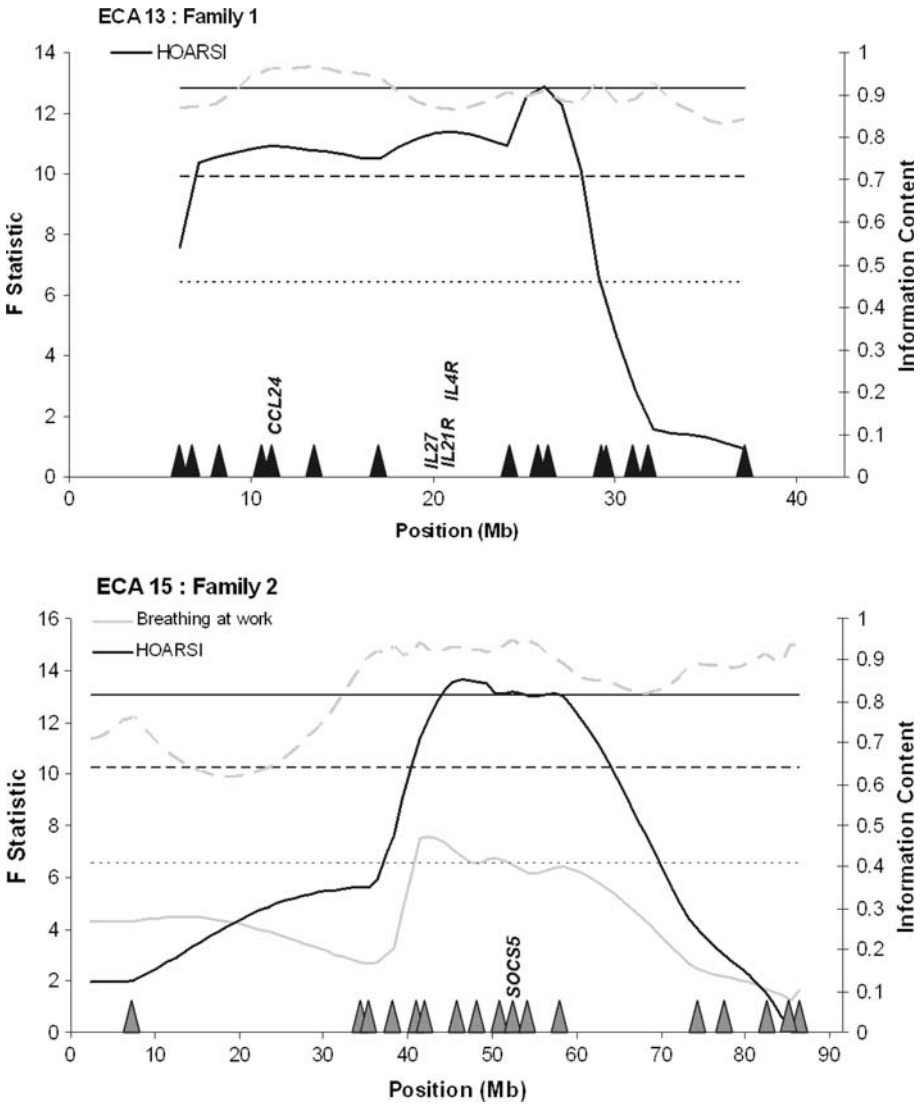

Fig. 1 *F* statistic plots for significant disease phenotypes on ECA13 and ECA15. The upper and lower broken horizontal lines indicate the chromosome-wide significance thresholds at $P = 0.01$ and $P = 0.05$, respectively. The *solid horizontal line* indicates the genome-wide significance threshold at $P = 0.05$ (the average of the individual phenotypes is shown for ECA15). Marker information content is shown on the right *Y*-axis and is indicated with a *dashed line*. The positions of informative markers in family 1 are shown with *black triangles* and for family 2 with *grey triangles*. The positions of genes referred to in the text are indicated

26) by regression analysis, and of these, the signals on ECA13 and 15 were significant at the genome-wide level ($P < 0.05$). The signals that are only chromosome-wide significant should be treated with caution as they may well be false positives. Linkage analysis of these two families using FASTLINK (Supplementary Table 2) also identified signals for HOARSI on ECA13 in family 1 and on ECA15 in family 2 and showed some evidence for a QTL on ECA21 in family 1.

Analysis was performed on each family separately because of the likelihood of observing genetic heterogeneity between the two families. In fact, this genetic heterogeneity had already been observed on ECA13 (Jost et al. 2007) where an association was seen in family 1 but not in family 2. Genetic heterogeneity exists when several genes are associated with the same disease; this can become evident when strictly defined phenotypic subgroups or individual families are examined separately (Rogaev et al. 1995) and the primary genetic factors causing disease in each family are identified independently. It is likely that with RAO there may be etiologically distinct subgroups that may not be distinguishable phenotypically; this is simplified in this study where the transmission of disease alleles from only one sire is analysed at a time. However, caution must be used in the extrapolation of these data to the general population; Warmblood horses are outbred and further study will be necessary to determine the extent of association seen at the population level between RAO and the regions indicated on ECA13 and ECA15.

From the power calculations (Table 2) it can be seen that QTL signals obtained in these families with a genome-wide significance of $P < 0.05$ probably have a substantial QTL heritability of at least 0.2 because there is less than 80% power to detect QTL with a heritability of less than

Table 4 Haplotype analysis of regions where genome-wide significance was found: ECA13 in family 1 and ECA15 in family 2

a.i. Haplotype analysis of family 1 for ECA13

	Mb	Marker name	Related haplotype	Unrelated haplotype
Region significant genome-wide at $P < 0.05$ in QTL analysis	6.10	COR069	293	291
	6.90	VHL161	180	182
	8.16	TKY031[a]	A140,109	A109,104
	10.28	TKY2450	213	207
	10.71	UM030	150	157
	13.30	TKY1219	244	232
	16.89	VHL047	156	140
	24.05	TKY581	192	188
	25.65	*TKY371*	*175*	*163*
	26.50	*AHT095*	*254*	*238*
	28.84	TKY3325	180	182
	29.05	TKY421	279	283
	31.74	ASB001	170	176
	32.29	TKY594	133	139
	37.56	TKY693	223	233

a.ii. Frequency of related haplotype in phenotypic classes: ECA13 in family 1

HOARSI	Severity of RAO phenotype	Related haplotype	Unrelated haplotype	Recombinant haplotype within this region
4	Severe	14	4	0
3	Moderate	18	16	0
2	Mild	13	15	0
1	Unaffected	19	31	0

a.iii. Fisher's exact test: ECA13 in family 1

Comparison	*P* value
HOARSI 4 vs. 1	0.005
HOARSI 4 and 3 vs. 1	0.0289

b.i. Haplotype analysis of family 2 for ECA15

	Mb	Marker name	Related haplotype	Unrelated haplotype
Region significant genome-wide at $P < 0.05$ in QTL analysis	7.05	UMNe222	165	151
	34.23	TKY1091	154	156
	36.03	TKY2810	192	194
	38.28	UMNe219	175	176
	41.33	UMNe156	138	136
	41.53	TKY1724	155	153
	46.02	*TKY2036*	*196*	*194*
	48.06	*TKY688*	*180*	*182*
	50.58	*UMNe608*	*231*	*229*
	52.31	*TKY926*	*120*	*109*
	54.61	*ASB002*	*202*	*198*
	57.93	*TKY885*	*250*	*252*
	73.96	HTG006	103	113
	77.19	TKY2020	129	125
	82.84	TKY2167	172	170
	85.45	HMS001	190	192
	86.78	COR075	216	218

Table 4 continued

b.ii. Frequency of related haplotype in phenotypic classes: ECA15 in family 2

HOARSI	Severity of RAO phenotype	Related haplotype	Unrelated haplotype	Recombinant haplotype within this region
4	Severe	9	4	0
3	Moderate	14	13	0
2	Mild	13	18	2
1	unaffected	5	28	1

b.iii. Fisher's exact test: ECA15 in family 2

Comparison	*P* value
HOARSI 4 vs. 1	0.0008
HOARSI 4 and 3 vs. 1	0.0003

a. Haplotype analysis for ECA13 in family 1

b. Haplotype analysis for ECA15 in family 2

i. Haplotyes identified as related and unrelated to HOARSI are shown with markers listed in their order on the chromosome. Allele sizes are in base pairs and the genome-wide significant region ($P < 0.05$) in QTL analysis is indicated (highlighted in italics)

ii. Haplotype frequencies of related, unrelated, and recombinant haplotypes for each of the HOARSI classes. These refer only to the haplotype region which is genome-wide significant ($P < 0.05$) in QTL analysis, i.e., markers TKY371 and AHT095 in ECA13 and TKY2036, TKY688, UMNe608, TKY926, ASB002, and TKY885 in ECA15

iii. Results of Fisher's exact test comparing first HOARSI 4 against HOARSI 1 and second HOARSI 4 and 3 against HOARSI 1

[a] The haplotypes could not be identified unambiguously at marker TKY1031

0.2. This can be considered a large-effect QTL and is in agreement with previous study of these families in which a major gene is indicated for the RAO phenotype in both families 1 and 2 (Gerber et al. 2009).

The recent development of high-resolution mapping studies in human has led to the identification of chromosome regions and, in many instances, specific genes implicated in complex diseases of the immune system (asthma reviewed in Zhang et al. 2008). The obvious similarities between RAO and human atopic asthma led us to compare the results of this RAO mapping study with the selection of genes identified for human asthma. Certainly the clinical signs of these two diseases are very comparable in many ways; bronchospasm, hyperreactivity of the airways, increased mucus secretion, reduced gaseous exchange resulting in a chronic cough, and/or increased breathing effort are among the obvious similarities. The conditions are both recurrent yet reversible and respond to bronchodilators and corticosteroids. Both diseases are triggered and exacerbated by inhaled particulate allergens. RAO has a genetic basis and develops over a period of years; it offers a unique natural model for human asthma (Herszberg et al. 2006; Kurucz and Szelenyi 2006; Lavoie et al. 2001).

Investigations of cytokine profiles in RAO have yielded conflicting results and still do not associate RAO with one polarized immune response, i.e., Th1 vs. Th2. While some studies reported a Th1-type or mixed response (Ainsworth et al. 2003; Giguere et al. 2002), others have found a Th2-biased response (Horohov et al. 2005; Lavoie et al. 2001). It may be that RAO is a more complex disease and that genetic heterogeneity may lead to a Th1-type response in some individuals and a Th2-type response in others, as we have proposed based on our earlier results (Jost et al. 2007). Similarly, the underlying mechanisms of inflammation, including Th1 vs. Th2 type and innate immunity cytokine profiles, vary greatly in different types of asthma (Abdulamir et al. 2008).

Both equine RAO and human asthma are complex natural diseases resulting from an interaction of genetic background and environmental factors. In human asthma, candidate gene approaches and association analyses were performed in many studies. However, results have been contradictory, with associations noted in one population but not in another (Whittaker 2003).

The most significant regions in which QTLs have been identified for RAO in this study are 6-28 Mb on ECA13 (corresponding to various rearranged segments of HSA7, 9, 12, 16, 19, and 22), 40-62 Mb on ECA15 (corresponding to HSA2, 36-61 Mb), and 6-44 Mb on ECA21 (corresponding to HSA5, 15-66 Mb). There are several possible candidate genes in these regions, notably *IL27*, *IL21R*, *IL4R*, *CCL24*, and *SOCS5*; their location is indicated in Fig. 1. *IL27* (interleukin 27), which has a role in the regulation of T-helper 1 cell differentiation, has been recognized in the downregulation of airway hyperreactivity and in lung inflammation during the development of allergic asthma. Chae et al. (2007) suggest that a polymorphism in *IL-27* is

associated with susceptibility to asthma. There is a significant association of an *IL21R* polymorphism with elevated IgE levels in females (Hecker et al. 2003). Elevated allergen-specific IgE levels are associated with RAO and influenced by genetic factors, but the role of IgE-mediated reactions in the pathogenesis of RAO is still unclear (Eder et al. 2000). Numerous studies have found associations between genetic variants in *IL4R* and risk of asthma in human populations (Cui et al. 2003; Howard et al. 2002; Hytönen et al. 2004; Mitsuyasu et al. 1998; Ober et al. 2000), although other studies have failed to find an association. A previous examination of SNPs in this region (ECA13, 15.05–20.87 Mb) found an association between one haplotype and RAO in family 1 but not in family 2 (Jost et al. 2007), indicating genetic heterogeneity in RAO. *CCL24* (eotaxin-2) recruits and activates cells carrying chemokine receptor 3, which plays a major role in asthma. However, since RAO, in contrast to some forms of human asthma, is not associated with significant eosinophilia, this gene may not be an obvious candidate for RAO. Recently it was shown that *SOCS5* acts in a classic negative feedback loop to inhibit IL-4 signaling (Kolesnik et al. 2008) and may play a role in allergic asthma (Inoue et al. 2007).

The QTL region on ECA21 contains three obvious candidate genes for RAO. The first is *PDE4D* (phosphodiesterase 4D), which is responsible for the degradation of cAMP, a key signal transduction molecule. Mice deficient in PDE4D are resistant to cholinergic airway contraction (Hansen et al. 2000). Phosphodiesterase inhibitors are under development as possible drug treatments for asthma (Kroegel and Foerster 2007). Two SNP variants in *IL7R* (interleukin 7 receptor) are associated with allergy caused by inhalation (Shamim et al. 2007). Finally, different forms of asthma in different human populations have been shown to be associated with a SNP in *PTGER4* (prostaglandin E receptor 4; Kim et al. 2007; Kurz et al. 2006).

The recent sequencing of the horse genome (http://www.nih.gov/news/pr/feb2007/nhgri-07.htm) has provided information on over one million SNPs. A selection has now been used to generate a SNP microarray of 54,000 SNPs which is suitable for genome-wide association studies (GWAS). Such a study will now be used to supplement the findings described here, using a population-wide sample collection, and will identify genetic regions associated with RAO in the general population.

It is very likely that a number of genetic loci contribute to the development of RAO. Single-gene tests will be limited in their diagnostic value; instead, a genetic profiling panel, which takes into account several genetic factors, will have greater value. Ideally, this will be combined with an assessment of environmental risk factors. Such assessment is already used in predicting the risk of human disease (Koppelman et al. 2008). In addition, the identification of genes that are responsible for RAO could lead to new therapeutic drugs specifically designed to target the implicated proteins. Such targeted pharmacological interventions are in development for asthma (Kroegel and Foerster 2007; Vendelin et al. 2005). In addition, and perhaps most importantly, the identification of genetic loci that are involved in the inheritance of RAO will contribute to the understanding of the pathological mechanisms underpinning this complicated condition.

Acknowledgments The Horse Trust, a UK horse charity committed to promoting education and welfare within the equine world, funded this project in the UK. In Switzerland, this study was supported by Vetsuisse and DKV grants, the Berne Equine Lung Research Group, and the Swiss National Science Foundation grant number 310000-116502. We thank Dr. Claire Wade of the Broad Institute for performing the BLAT search of known horse microsatellite sequences against the horse genome sequence, the information from which was used here to physically position the markers. Last but not least, we thank A. Ramseyer, U. Jost, N. Hasler, E. Laumen, P. Nussbaumer, and D. Burger and all horse owners for their help phenotyping the horses and collecting blood samples.

References

Abdulamir AS, Hafidh RR, Abubakar F, Abbas KA (2008) Changing survival, memory cell compartment, and T-helper balance of lymphocytes between severe and mild asthma. BMC Immunol 9:73

Abecasis GR, Cherny SS, Cookson WO, Cardon LR (2002) Merlin—rapid analysis of dense genetic maps using sparse gene flow trees. Nat Genet 30:97–101

Ainsworth DM, Grünig G, Matychak M, Young J, Wagner B et al (2003) Recurrent airway obstruction (RAO) in horses is characterized by IFN-γ and IL-8 production in bronchoalveolar lavage cells. Vet Immunol Immunopathol 96:83–91

Bossé Y, Hudson TJ (2007) Toward a comprehensive set of asthma susceptibility genes. Annu Rev Med 58:171–184

Chae SC, Li CS, Kim KM, Yang JY, Zhang Q et al (2007) Identification of polymorphisms in human interleukin-27 and their association with asthma in a Korean population. J Hum Genet 52:355–361

Cottingham RW Jr, Idury RM, Schäffer AA (1993) Faster sequential genetic linkage computations. Am J Hum Genet 53:252–263

Cui T, Wu J, Pan S, Xie J (2003) Polymorphisms in the IL-4 and IL-4R [α] genes and allergic asthma. Clin Chem Lab Med 41:888–892

Eder C, Crameri R, Mayer C, Eicher R, Straub R et al (2000) Allergen-specific IgE levels against crude mould and storage mite extracts and recombinant mould allergens in sera from horses affected with chronic bronchitis. Vet Immunol Immunopathol 73:241–253

Gerber H (1989) The genetic basis of some equine diseases (Sir Frederick Hobday Lecture). Equine Vet J 21:244–248

Gerber V, Straub R, Marti E, Hauptman J, Herholz C et al (2004) Endoscopic scoring of mucus quantity and quality: observer and horse variance and relationship to inflammation, mucus viscoelasticity and volume. Equine Vet J 36:576–582

Gerber V, Baleri D, Klukowska-Rötzler J, Swinburne JE, Dolf G (2009) Mixed inheritance of equine recurrent airway obstruction (RAO). J Vet Intern Med 23:626–630

Giguere S, Viel L, Lee E, MacKay RJ, Hernandez J et al (2002) Cytokine induction in pulmonary airways of horses with heaves and effect of therapy with inhaled fluticasone propionate. Vet Immunol Immunopathol 85:147–158

Hansen G, Jin S, Umetsu DT, Conti M (2000) Absence of muscarinic cholinergic airway responses in mice deficient in the cyclic nucleotide phosphodiesterase PDE4D. Proc Natl Acad Sci USA 97:6751–6756

Hecker M, Bohnert A, König IR, Bein G, Hackstein H (2003) Novel genetic variation of human interleukin-21 receptor is associated with elevated IgE levels in females. Genes Immun 4:228–233

Herszberg B, Ramos-Barbón D, Tamaoka M, Martin JG, Lavoie JP (2006) Heaves, an asthma-like equine disease, involves airway smooth muscle remodelling. J Allergy Clin Immunol 118:382–388

Horohov DW, Beadle RE, Mouch S, Pourciau SS (2005) Temporal regulation of cytokine mRNA expression in equine recurrent airway obstruction. Vet Immunol Immunopathol 18:237–245

Howard TD, Koppelman GH, Xu J, Zheng SL, Postma DS et al (2002) Gene-gene interaction in asthma: IL4RA and IL13 in a Dutch population with asthma. Am J Hum Genet 70:230–236

Hytönen AM, Löwhagen O, Arvidsson M, Balder B, Björk AL et al (2004) Haplotypes of the interleukin-4 receptor alpha chain gene associate with susceptibility to and severity of atopic asthma. Clin Exp Allergy 34:1570–1575

Inoue H, Fukuyama S, Matsumoto K, Kubo M, Yoshimura A (2007) Role of endogenous inhibitors of cytokine signaling in allergic asthma. Curr Med Chem 14:181–189

Jost U, Klukowska-Rötzler J, Dolf G, Swinburne JE, Ramseyer A et al (2007) Association of microsatellite markers near interleukin-4 receptor α chain gene (IL4Rα) with chronic lower airway disease—locus heterogeneity in two high-prevalence horse families. Equine Vet J 39:236–241

Kim SH, Kim YK, Park HW, Jee YK, Kim SH et al (2007) Association between polymorphisms in prostanoid receptor genes and aspirin-intolerant asthma. Pharmacogenet Genomics 17:295–304

Koch P (1957) Heredity of chronic alveolar emphysema of the lungs in horses. Dtsch Tieraerztl Wochenschr 64:485–486

Kolesnik T, Columbus RE, Chakravorty A, Willson TA, Sprigg NS et al (2008) Suppressor of cytokine signalling (SOCS)-5 regulates IL-4 activity via interaction with the IL-4 receptor and inhibition of JAK1 kinase activity. Cytokine 43:327

Koppelman GH, te Meerman GJ, Postma DS (2008) Genetic testing for asthma. Eur Respir J 32:775–782

Kroegel C, Foerster M (2007) Phosphodiesterase-4 inhibitors as a novel approach for the treatment of respiratory disease: cilomilast. Expert Opin Investig Drugs 16:109–124

Kurucz I, Szelenyi I (2006) Current animal models of bronchial asthma. Curr Pharm Des 12:3175–3194

Kurz T, Hoffjan S, Hayes MG, Schneider D, Nicolae R et al (2006) Fine mapping and positional candidate studies on chromosome 5p13 identify multiple asthma susceptibility loci. J Allergy Clin Immunol 118:396–402

Lavoie JP, Maghni K, Desnoyers M, Taha R, Matin JG et al (2001) Neutrophilic airway inflammation in horses with heaves is characterized by a Th2-type cytokine profile. Am J Respir Crit Care Med 164:1410–1413

Malerba G, Pignatti PF (2005) A review of asthma genetics: gene expression studies and recent candidates. J Appl Genet 46:93–104

Marti E, Gerber H, Essich G, Oulehla J, Lazary S (1991) The genetic basis of equine allergic diseases. 1. Chronic hypersensitivity bronchitis. Equine Vet J 23:457–460

Mitsuyasu H, Izuhara K, Mao XQ, Gao PS, Arinobu Y et al (1998) Ile50Val variant of IL4R alpha upregulates IgE synthesis and associates with atopic asthma. Nat Genet 19:119–120

Ober C, Leavitt SA, Tsalenko A, Howard TD, Hoki DM et al (2000) Variation in the interleukin 4-receptor alpha gene confers susceptibility to asthma and atopy in ethnically diverse populations. Am J Hum Genet 66:517–526

Ramseyer A, Gaillard C, Burger D, Straub R, Jost U et al (2007) Effects of genetic and environmental factors on Horse Owner Assessed Respiratory Signs Index (HOARSI). J Vet Intern Med 21:149–156

Robinson NE, Derksen FJ, Olszewski MA, Buechner-Maxwell VA (1995) The pathogenesis of chronic obstructive pulmonary disease of horses. Br Vet J 152:283–286

Rogaev EI, Sherrington R, Rogaeva EA, Levesque G, Ikeda M et al (1995) Familial Alzheimer's disease in kindreds with missense mutations in a gene on chromosome 1 related to the Alzheimer's disease type 3 gene. Nature 376:775–778

Schaeper W (1939) Untersuchungen über die Erblichkeit und das Wesen des Lungendampfes beim Pferd [Investigation into the nature and heritability of heaves in the horse]. Tieraerztl Rundschau 31:595–599

Schuelke M (2000) An economic method for the fluorescent labeling of PCR fragments. Nat Biotechnol 18:233–234

Seaton G, Haley CS, Knott SA, Kearsey M, Visscher PM (2002) QTL Express: mapping quantitative trait loci in simple and complex pedigrees. Bioinformatics 18:339–340

Seaton G, Hernandez J, Grunchec JA, White I, Allen J et al (2006) GridQTL: A grid portal for QTL mapping of compute intensive datasets. In: Proceedings of the 8th world congress on genetics applied to livestock production, Belo Horizonte, Brazil, August 13–18, 2006

Shamim Z, Müller K, Svejgaard A, Poulsen LK, Bodtger U et al (2007) Association between genetic polymorphisms in the human interleukin-7 receptor α-chain and inhalation allergy. Int J Immunogenet 34:149–151

Swinburne JE, Boursnell M, Hill G, Pettitt L, Allen T et al (2006) Single linkage group per chromosome genetic linkage map for the horse, based on two three-generation, full-sibling, crossbred horse reference families. Genomics 87:1–29

Vendelin J, Pulkkinen V, Rehn M, Pirskanen A, Räisänen-Sokolowski A et al (2005) Characterization of GPRA, a novel G protein-coupled receptor related to asthma. Am J Respir Cell Mol Biol 33:262–270

Whittaker PA (2003) Genes for asthma: much ado about nothing? Curr Opin Pharmacol 3:212–219

Williams JT, Blangero J (1999) Power of variance component linkage analysis to detect quantitative trait loci. Ann Hum Genet 63:545–563

Wills-Karp M, Ewart SL (2004) Time to draw breath: asthma-susceptibility genes are identified. Nat Rev Genet 5:376–387

Zhang J, Paré PD, Sandford AJ (2008) Recent advances in asthma genetics. Respir Res 15:9–14