Sequence variant analysis of RNA sequences in severe equine asthma

Laurence Tessier[1,2], Olivier Côté[1,3] and Dorothee Bienzle[1]

[1] Department of Pathobiology, University of Guelph, Guelph, ON, Canada
[2] BenchSci, Toronto, ON, Canada
[3] BioAssay Works, Ijamsville, MD, USA

ABSTRACT

Background: Severe equine asthma is a chronic inflammatory disease of the lung in horses similar to low-Th2 late-onset asthma in humans. This study aimed to determine the utility of RNA-Seq to call gene sequence variants, and to identify sequence variants of potential relevance to the pathogenesis of asthma.

Methods: RNA-Seq data were generated from endobronchial biopsies collected from six asthmatic and seven non-asthmatic horses before and after challenge (26 samples total). Sequences were aligned to the equine genome with Spliced Transcripts Alignment to Reference software. Read preparation for sequence variant calling was performed with Picard tools and Genome Analysis Toolkit (GATK). Sequence variants were called and filtered using GATK and Ensembl Variant Effect Predictor (VEP) tools, and two RNA-Seq predicted sequence variants were investigated with both PCR and Sanger sequencing. Supplementary analysis of novel sequence variant selection with VEP was based on a score of <0.01 predicted with Sorting Intolerant from Tolerant software, missense nature, location within the protein coding sequence and presence in all asthmatic individuals. For select variants, effect on protein function was assessed with Polymorphism Phenotyping 2 and screening for non-acceptable polymorphism 2 software. Sequences were aligned and 3D protein structures predicted with Geneious software. Difference in allele frequency between the groups was assessed using a Pearson's Chi-squared test with Yates' continuity correction, and difference in genotype frequency was calculated using the Fisher's exact test for count data.

Results: RNA-Seq variant calling and filtering correctly identified substitution variants in *PACRG* and *RTTN*. Sanger sequencing confirmed that the *PACRG* substitution was appropriately identified in all 26 samples while the *RTTN* substitution was identified correctly in 24 of 26 samples. These variants of uncertain significance had substitutions that were predicted to result in loss of function and to be non-neutral. Amino acid substitutions projected no change of hydrophobicity and isoelectric point in PACRG, and a change in both for RTTN. For *PACRG*, no difference in allele frequency between the two groups was detected but a higher proportion of asthmatic horses had the altered *RTTN* allele compared to non-asthmatic animals.

Discussion: RNA-Seq was sensitive and specific for calling gene sequence variants in this disease model. Even moderate coverage (<10–20 counts per million) yielded correct identification in 92% of samples, suggesting RNA-Seq may be suitable to detect sequence variants in low coverage samples. The impact of amino acid

alterations in PACRG and RTTN proteins, and possible association of the sequence variants with asthma, is of uncertain significance, but their role in ciliary function may be of future interest.

INTRODUCTION

Severe equine asthma (recurrent airway obstruction, heaves) is a chronic inflammatory lung disease caused by inhalation of environmental dust and microbial components (*Couëtil et al., 2016*). Exacerbation of the disease triggers excessive mucus production, cough, neutrophilic airway inflammation, bronchial hyperreactivity, and bronchospasm. Recurrent exacerbations induce smooth muscle hyperplasia and hypertrophy, fibrosis and eventual irreversible airway remodeling (*Pirie, 2014*; *Martinez & Vercelli, 2013*; *Vargas et al., 2016*; *Setlakwe et al., 2014*).

Asthma in humans is recognized to be a heterogeneous disease that is classified considering genetic, molecular and clinical features (*Wenzel, 2012*; *Skloot, 2016*). Severe equine asthma is most similar to human severe, late-onset asthma characterized by absence of Th2 cytokines, and presence of neutrophilic inflammation and bronchial neutrophil chemokines (*Wenzel, 2012*; *Ilmarinen, Tuomisto & Kankaanranta, 2015*). Severely asthmatic horses do not have a hypersensitivity response (*Pirie, 2014*) and efforts to associate equine asthma with a Th2 cytokine profile have yielded inconsistent or inconclusive results (*Lavoie et al., 2001*; *Joubert, Silversides & Lavoie, 2001*; *Klukowska-Rötzler et al., 2012a*; *Ainsworth et al., 2003*, *2007*, *2009*; *Debrue et al., 2005*; *Lavoie-Lamoureux et al., 2010*; *Padoan et al., 2013*). Mechanisms leading to the development of both severe equine asthma and late-onset low-Th2 severe asthma in humans remain largely undefined.

Interactions between genes and environmental factors have been recognized to contribute to development of equine asthma for many years (*Marti et al., 1991*). Genetic factors likely reside in multiple gene sequence variants, and may be influenced by age and sex (*Marti et al., 1991*; *Couëtil & Ward, 2003*; *Ramseyer et al., 2007*). Several susceptibility sequence variants, haplotypes and regions have been associated with human asthma (*Moffatt et al., 2007*; *Choudhry et al., 2008*; *Galanter et al., 2008*; *Tavendale et al., 2008*; *Madore et al., 2008*; *Hancock et al., 2009*; *Himes et al., 2009*; *Bisgaard et al., 2009*; *Wu et al., 2009*; *Li et al., 2010*; *Sleiman et al., 2010*; *Moffatt et al., 2010*; *Ferreira et al., 2011*; *Torgerson et al., 2011*) but no specific markers have been identified in the late-onset low-Th2 sub-phenotype (*Wenzel, 2012*; *Ilmarinen, Tuomisto & Kankaanranta, 2015*). Similarly, genetic markers of equine severe asthma were identified in certain families, but were not significantly associated across different families and genetic backgrounds (*Ramseyer et al., 2007*; *Swinburne et al., 2009*; *Klukowska-Rötzler et al., 2012b*; *Jost et al., 2007*; *Klukowska-Rötzler, Gerber & Leeb, 2012*).

RNA-Seq is a promising approach for calling sequence variants concurrent with analysis of gene and allele-specific expression, alternative splicing, and pathways (*Guo et al., 2015*). In this study we investigated whether single nucleotide variant (SNV) detected by RNA-Seq were also present in Sanger-sequenced amplicons. We hypothesized that RNA-Seq would identify gene sequence variants with high accuracy.

METHODS

Animals and procedures

Animal and sample procedures were previously described (*Tessier et al., 2017*). Briefly, six asthmatic and seven non-asthmatic horses without signs of asthma exacerbation belonging to the institutional research herd (mean ages of 15 and 12 years, respectively, $P = 0.352$, unpaired t-test) were placed indoor in a dust-free environment. After 24 h, asthmatic horses were exposed to dusty hay until exacerbation (range 1–3 days, average 2.2 days), while non-asthmatic horses were exposed for 3 days. Before and after the dusty hay asthmatic challenge, physical examination, pulmonary function test, and bronchoalveolar lavage were performed, and endoscopic bronchial biopsies were collected from lung lobes contralateral between first and second samples. Samples from an additional four asthmatic and seven non-asthmatic horses were used for PCR-amplification of specific sequence variant regions and Sanger sequencing. All procedures were approved by the Institutional Animal Care Committee of the University of Guelph (protocol R10-031) and conducted in compliance with Canadian Council on Animal Care guidelines.

RNA-Seq sample preparation and sequence alignment

RNA extraction, preparation, and sequencing procedures were as described previously (*Tessier et al., 2017*). In brief, total RNA was extracted from endobronchial biopsies (Qiagen, Toronto, ON, Canada) and tested for quality and concentration with the Bioanalyzer RNA Nanochip (Agilent, Mississauga, ON, Canada) and capillary electrophoresis. RNA-Seq library preparation (unstranded) and sequencing were performed using the Illumina TruSeq RNA sample preparation and appropriate sequencing protocols (Illumina, San Diego, CA, USA) at The Centre for Applied Genomics (Toronto, ON, Canada). Sequencing of 100-base paired-end reads was performed following the manufacturer's instructions on an Illumina HiSeq 2500 instrument.

FastQC software version 0.10.1 (bioinformatics.babraham.ac.uk/projects/fastqc/) was used to assess quality of raw reads, and alignment to the horse reference genome (*Wade et al., 2009*) (Ensembl v70) was performed with STAR version 2.4 (*Dobin et al., 2013*). Specifically, the STAR_pass2 alignment protocol was followed using the horse Ensembl version 70 GTF annotation file for first- and second-pass, and the junction SJ.tab file generated by STAR for the second-pass after non-canonical junctions were removed. Default settings were used except for: –runThreadN 8 –outFilterScoreMinOverLread 0.5 –outFilterMatchNminOverLread 0.5. Details and results for read alignment were previously described (*Tessier et al., 2017*).

Sequence variant calling and filtering

Read processing, sequence variant calling, and initial filtering were performed following the Genome Analysis ToolKit (GATK) best practice guide for variant calling on RNA-Seq, except for the Indel realignment step considering the pass-2 STAR alignment initially performed. Initial read processing was first performed with Picard tools version 1.114 (broadinstitute.github.io/picard/) to add read groups and mark duplicates. Split n' Trim as well as base recalibration were performed using the GATK software version 3.2.2 (*McKenna et al., 2010*) and the *-T SplitNCigarReads*, *-rf ReassignOneMappingQuality*, *-RMQF 255*, *-RMQT 60*, and *-U ALLOW_N_CIGAR_READS* options.

The GATK variant calling and filtering workflow yielded 2,823 and 1,788 sequence variants present in all horses of the asthmatic group pre- and post-challenge, respectively (Fig. S1). Sequence variants were subsequently called using the Haplotype Caller function in GATK with the same genome annotation file used in the read alignment phase and the following options: *-recoverDanglingHeads*, *-dontUseSoftClippedBases*, *-stand_call_conf 20.0*, and *-stand_emit_conf 20.0* options. Resultant sequence variants were processed with the variant filtration function of GATK software and the following options to establish a confidence threshold of reported variants: *-window 35*, *-cluster 3*, *-filterName FS*, *-filter "FS > 30.0"*, *-filterName QD*, and *-filter "QD < 2.0."* Sequence variants were analyzed individually in each of 26 samples (six asthmatics and seven non-asthmatics, before and after asthmatic challenge).

PCR

Primers for amplification of sequence variant regions from bronchial DNA were parkin co-regulated (*PACRG*) forward (5′-CTC TGA ACC TCC GAA ACC GAC-3′) and reverse (5′-CTC CTG GGA TAA CTC ACC ATT C-3′), and rotatin (*RTTN*) forward (5′-TCC TGA GTT GTA TCA AGA AGT G-3′) and reverse (5′-CCA GCC TGC AAT TCC TTT CT-3′). A Taq polymerase PCR kit (Invitrogen, Mississauga, ON, Canada) was used for PCR amplifications. Each reaction was performed in a 25 μL final volume, including five μL of 10× PCR buffer, 0.2 mM dNTPs, two mM MgSO$_4$, 0.3 μM of each primer, two U of Platinum Taq, and five μL (100 ng) of template DNA. PCR conditions for amplification were 3 min at 94 °C followed by 35 cycles of 94 °C for 45 s, 60 °C or 58 °C for 30 s for *PACRG* and *RTTN*, respectively, and 72 °C for 90 s, followed by final elongation for 10 min at 72 °C. A total of 20 μL of each PCR product was separated by electrophoresis in a 1% agarose gel stained with SYBR Safe (Invitrogen, Mississauga, ON, Canada). Amplicons of appropriate size were cut out and DNA extracted and purified (QIAquick; Qiagen, Toronto, ON, Canada). Extracted and purified PCR products were Sanger sequenced (Laboratory Services Division, Guelph, ON, Canada).

RESULTS

Sequence variant calling and filtering

The mean of the total number of RNA-Seq reads for all samples was 36252701.08, and the mean of uniquely mapped number of reads was 33127466.35. The number of
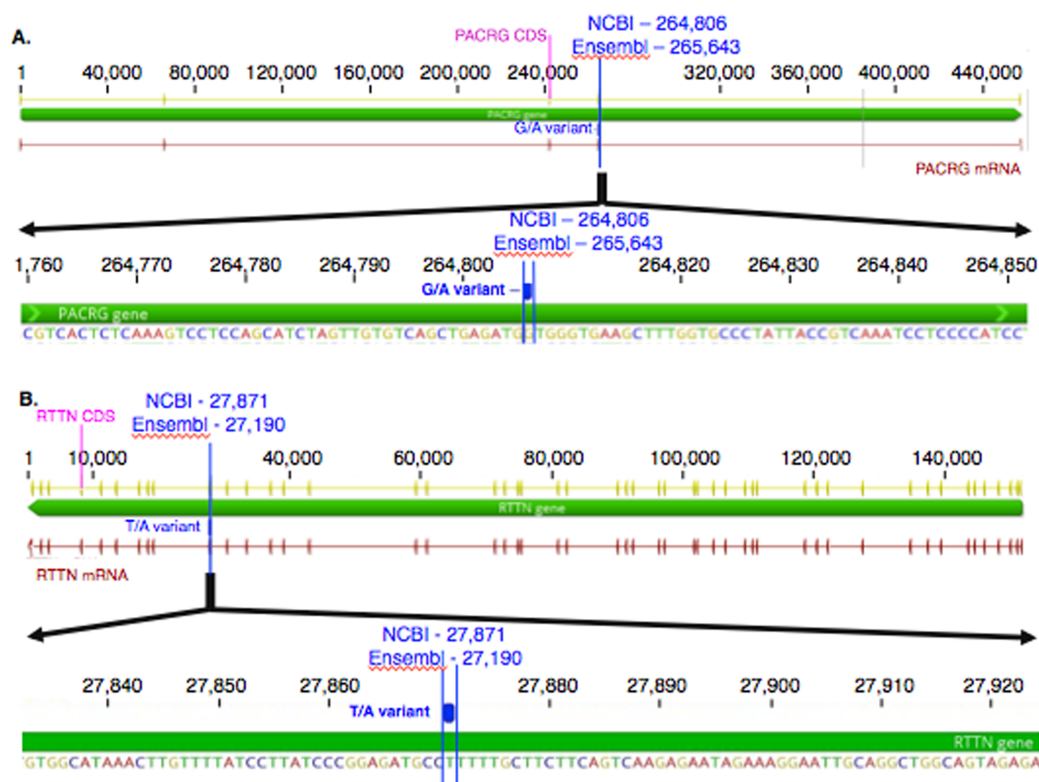
Figure 1 Substitution variants in *PACRG* (A) and *RTTN* (B) genes. Diagrams show position of gene (thick green line), mRNA (red line) and coding region (yellow line). The whole gene is outlined at the top and details of the nucleotide sequence around the substitution variant are shown at the bottom. Location of the substitution variant is indicated in blue for NCBI and Ensembl databases.

Full-size 🖼 DOI: 10.7717/peerj.5759/fig-1

individual total reads, total mapped reads, uniquely mapped reads and multiple mapped reads is detailed elsewhere (39). The GATK workflow resulted in 2,823 and 1,788 sequence variants present specifically in the asthmatic group pre- and post-challenge, respectively (Fig. S1). Of these, 10 were missense substitution variants, coded for proteins and had Sorting Intolerant from Tolerant (SIFT) scores <0.01. Substitution variants in *PACRG* (Fig. 1A) and *RTTN* (Fig. 1B) were detected at higher proportion in asthmatic compared to non-asthmatic horses. A missense G/A substitution was detected in the coding sequence of *PACRG* at position 265,643 (Ensembl sequence ENSECAG00000014308)/264,806 (NCBI accession number 100050378) (Fig. 1A). A missense T/A substitution was detected in the coding sequence of *RTTN* at position 27,190 (Ensembl sequence ENSECAG00000009711)/27,871 (NCBI accession number 100052029) (Fig. 1B).

Amino acid sequence alignment

In PACRG, the G/A substitution resulted in replacement of valine (V) for methionine (M) at position 182 (Fig. 2A). PACRG sequence alignment of wild type (WT) and altered proteins predicted changes from beta-strand to alpha-helix structure in the altered protein a few amino acids distant from the site of substitution (182) at positions 187 and 188
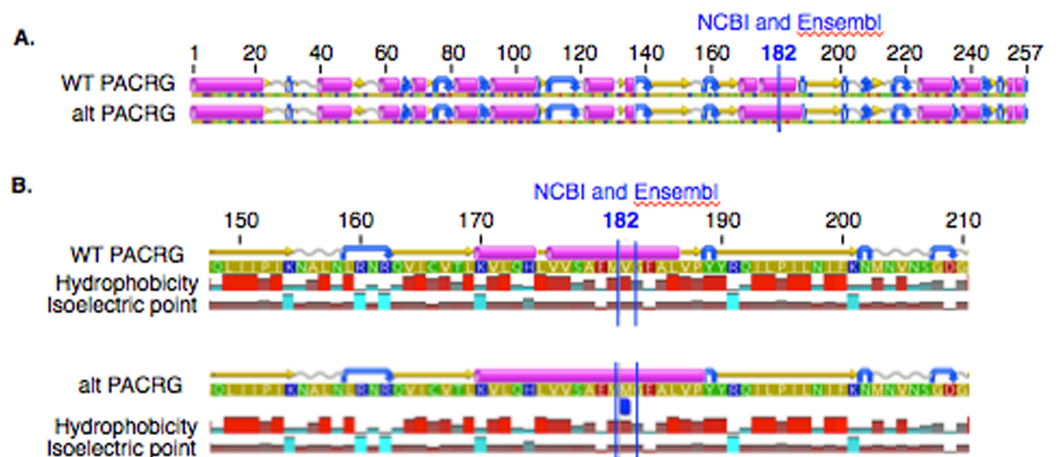
Figure 2 Alignment of wild type (WT) and mutant (mut) PACRG proteins with associated predicted hydrophobicity and isoelectric point. Low (A) and high magnification (B) of predicted protein changes. Replacement of methionine for valine at position 182 changes a beta strand to an alpha helix at position 187 and 188. Alpha helices (pink), coils (gray line), turns (blue arrows), and beta strands (yellow arrows).
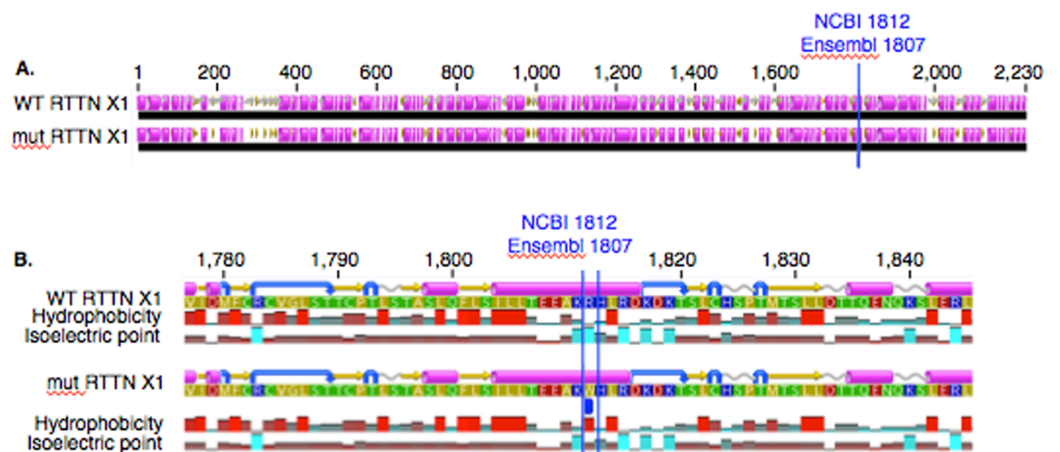
Full-size ☑ DOI: 10.7717/peerj.5759/fig-2

Figure 3 Alignment of WT and mutant RTTN protein with predicted hydrophobicity and isoelectric point. Low (A) and high magnification (B) of predicted protein changes. Replacement of arginine with tryptophan at position 1,812 changes alpha helix to beta strand at position 1,816. Alpha helices (pink), coils (gray line), turns (blue arrows), and beta strands (yellow arrows).

Full-size ☑ DOI: 10.7717/peerj.5759/fig-3

(Fig. 2B). Hydrophobicity and isoelectric point were expected to remain similar despite the substitution (Geneious). In the RTTN sequence alignment, T/A substitution resulted in replacement of arginine (R) with tryptophan (W) at position 1,807 of the ENSECAT00000010304 protein isoform (Ensembl sequence, corresponding to position 1,812 of isoform X1 in NCBI [XP_001493238]) in NCBI sequence) (Fig. 3A). Sequence alignment of WT and altered proteins indicated a change from alpha helix to beta strand structure near the site of substitution (bp 1,807) at position 1,816 (Fig. 3B). In addition, increased hydrophobicity and decreased pI were projected at the site of substitution (1,807) in the altered compared to the WT protein.
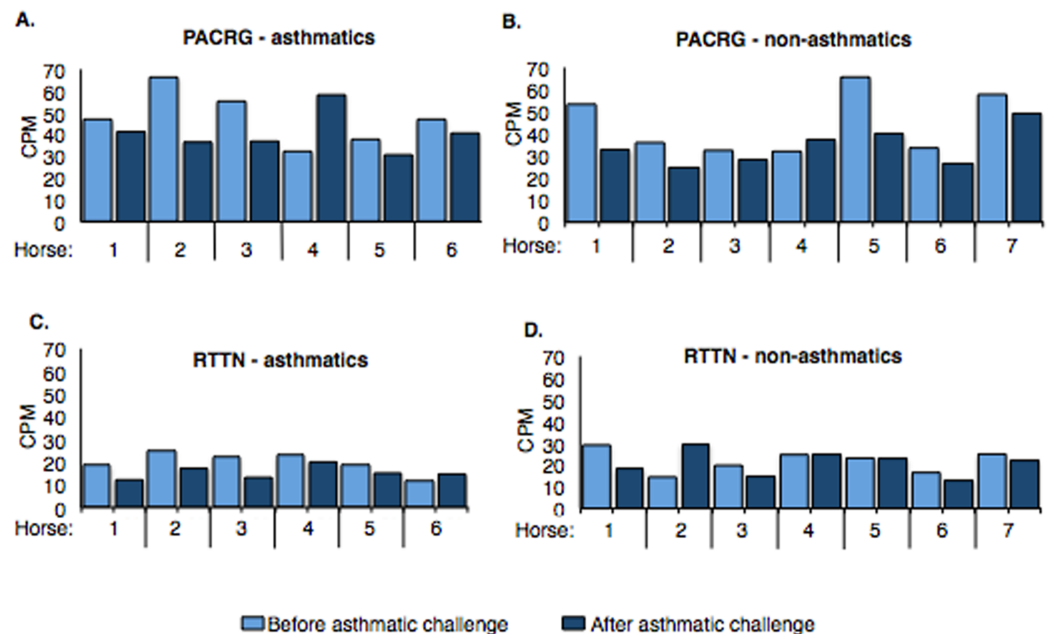
Figure 4 Expression of PACRG (A, B) and RTTN (C, D) in asthmatic and non-asthmatic horses in counts-per-million (CPM; *y*-axis) pre- and post-challenge. PACRG expression varied from ~30.7 to 66.3 CPM in asthmatic horses (A) and ~25.1 to 65.6 CPM on non-asthmatic horses (B), while RTTN expression varied from ~11.5 to 24.9 CPM and ~13.1 to 29.9 CPM in asthmatic (C) and non-asthmatic (D) horses, respectively. Full-size ◨ DOI: 10.7717/peerj.5759/fig-4

Expression of *PACRG* and *RTTN*

Counts per million (cpm) for *PACRG* ranged from ~30.7 to 66.3 (mean = 44.17) in asthmatic horses (Fig. 4A) and ~25.1 to 65.6 (mean = 39.31) in non-asthmatic horses (Fig. 4B), while expression of *RTTN* varied from ~11.5 to 24.9 (mean = 17.61) and ~13.1 to 29.9 (mean = 21.46) in asthmatic (Fig. 4C) and non-asthmatic (Fig. 4D) horses, respectively.

Confirmation of RNA-Seq with DNA Sanger sequencing

The *PACRG* substitution variants identified by RNA-Seq were confirmed on DNA with Sanger sequencing of PCR amplicons in 10 asthmatic (Fig. 5A) and 14 non-asthmatic (Fig. 5B) horses. In the asthmatic group, four horses were heterozygous [A/G] and two were homozygous with altered alleles [A/A] (Fig. 5A). In the non-asthmatic group, one horse was heterozygous with alleles [A/G], three horses were homozygous WT [G/G] and three horses were homozygous altered [A/A] (Fig. 5B). DNA was not available to confirm the genotype of horse 1. Hence, all genotypes derived from RNA-Seq were identical to Sanger sequencing results.

For *RTTN*, results of RNA-Seq and Sanger sequencing were very similar (Fig. 6). Four asthmatic horses were heterozygous [A/T] and two were homozygous altered [A/A] (Fig. 6A). In the non-asthmatic group, horse 5 was heterozygous [A/T] and four horses were identified as homozygous WT [T/T]. However, horse 1 was identified as homozygous before and heterozygous after the asthmatic challenge, while horse 4 was identified as heterozygous before and homozygous after the asthmatic challenge. Based on Sanger
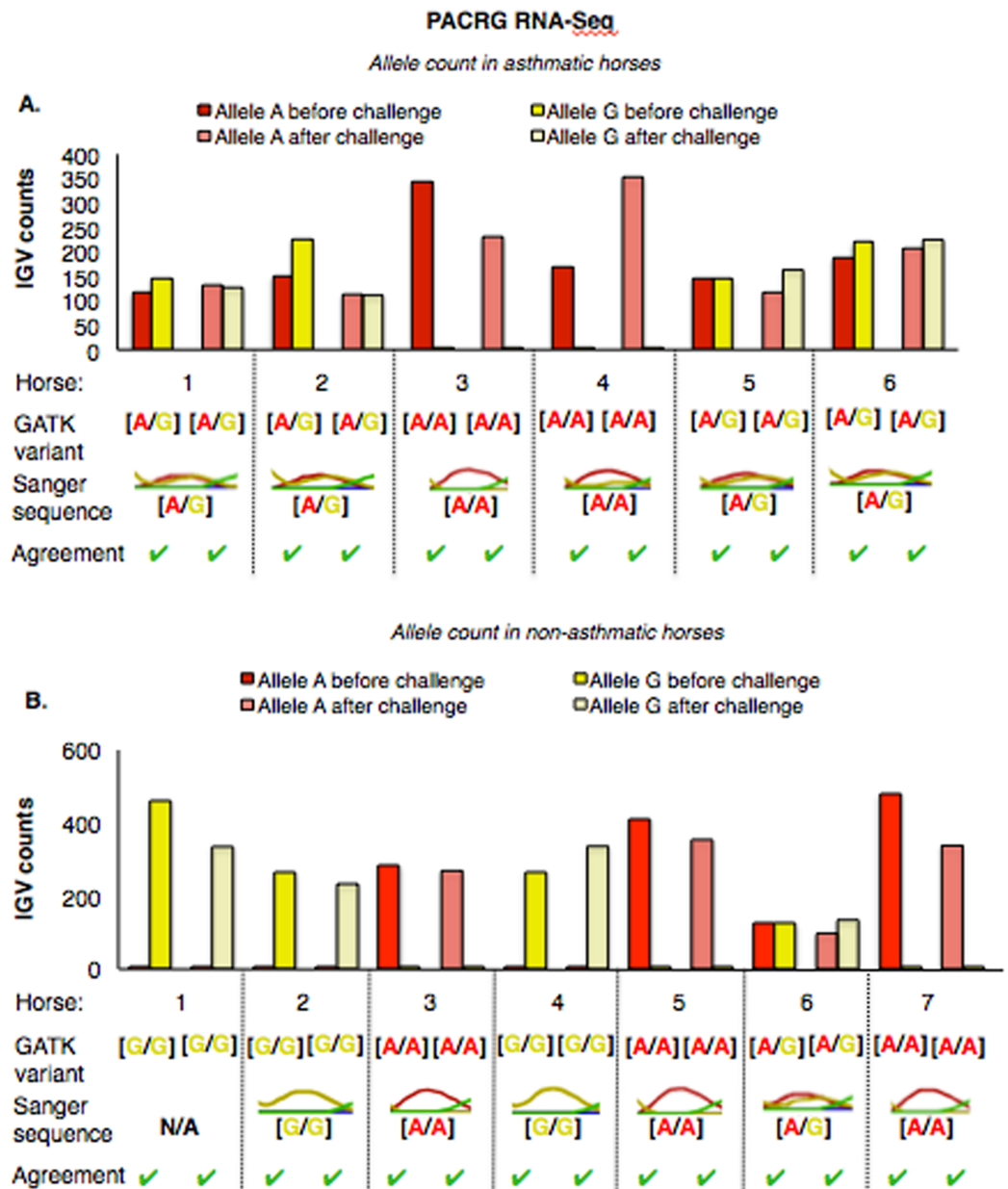
Figure 5 Comparison of GATK variant calls and Sanger sequencing results for *PACRG* in asthmatic (A) and non-asthmatic (B) horses. For both groups, the bar graph indicates the IGV count for each allele (A-red, G-yellow), horse and condition. Below the bar graph is the GATK variant call, the electropherogram of the Sanger sequence, and agreement. (A) Four asthmatic horses (1, 2, 5, 6) had heterozygous alleles [A/G] and two (3 and 4) were homozygous for the mutant allele [A/A]. (B) In non-asthmatic horses, one horse (6) had heterozygous alleles [A/G], three horses (1, 2, and 4) were homozygous for the wild-type allele [G/G] and three horses (3, 5, and 7) were homozygous for the mutant allele [A/A]. All genotypes were consistent across horses and methods. DNA was not available for non-asthmatic horse 1.

Full-size 🖼 DOI: 10.7717/peerj.5759/fig-5

sequencing the genotype of horse 4 was homozygous. DNA was not available to confirm the genotype of horse 1. In all asthmatic horses, genotypes were consistent between pre- and post-challenge and sequencing methods.
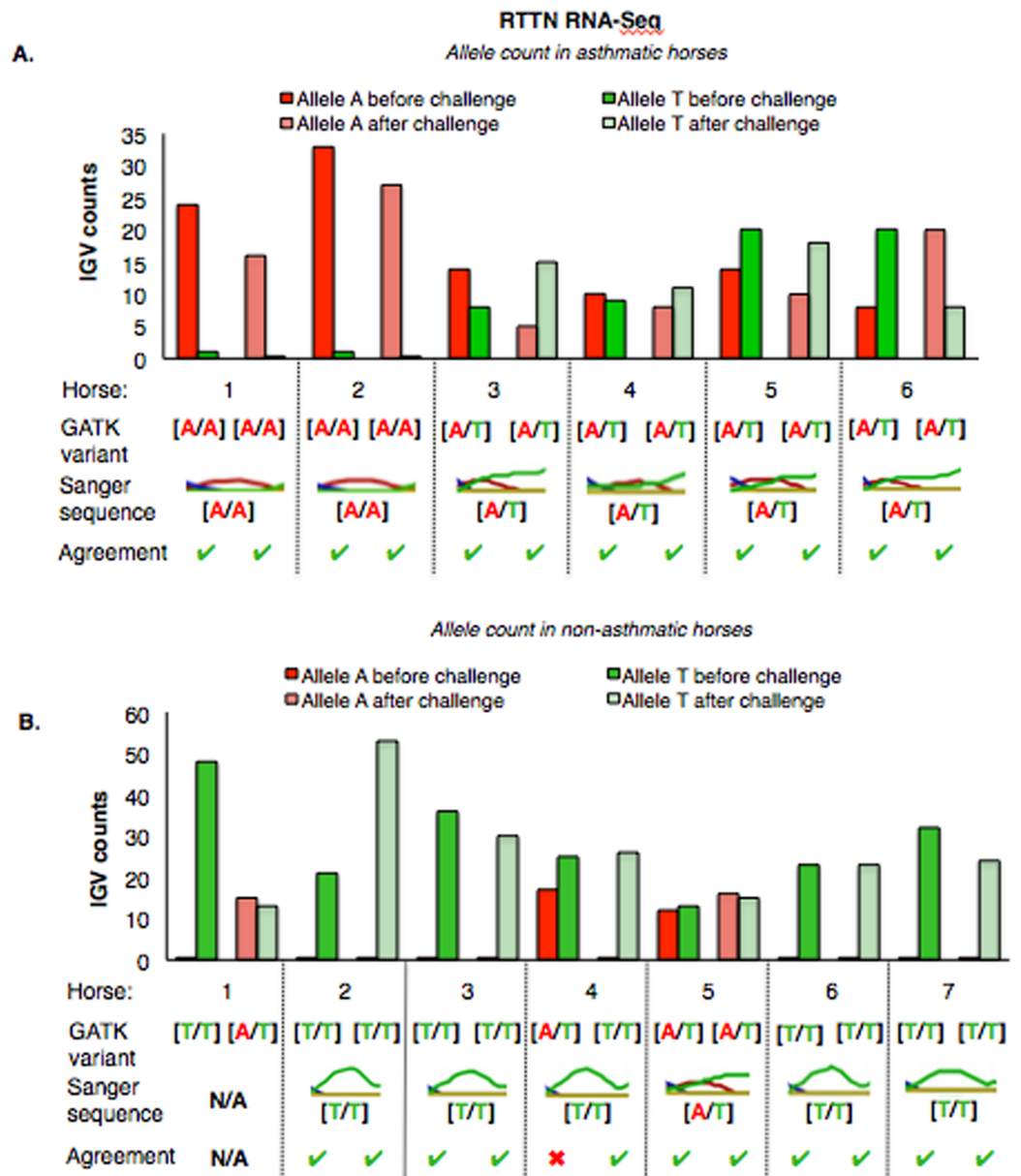
Figure 6 Comparison of GATK variant calls and Sanger sequencing results for *RTTN* in asthmatic (A) and non-asthmatic (B) horses. Details as in Fig. 5. (A) Four asthmatic horses (3–6) had heterozygous alleles [A/T] and two (1 and 2) were homozygous for the mutant allele [A/A]. Genotypes were consistent across horses and methods. (B) In non-asthmatic horses, one (5) had heterozygous [A/T] alleles, four horses had homozygous wild type [T/T] alleles, and two horses (1 and 4) were inconsistently identified as homozygous wild type and heterozygous in different samples. Sanger sequencing confirmed the genotype of horse 4 as heterozygous. DNA was not available for non-asthmatic horse 1.　　　　　　　　Full-size 🖼 DOI: 10.7717/peerj.5759/fig-6

Sequence alignment

Sanger DNA sequences of *PACRG* from 10 asthmatic horses and 14 non-asthmatic horses (including those that were analyzed by RNA-Seq) were aligned. Among asthmatic horses, eight had the heterozygous [A/G], two had the homozygous altered [A/A] and none had the homozygous WT [G/G] genotype. Among non-asthmatic horses, six had
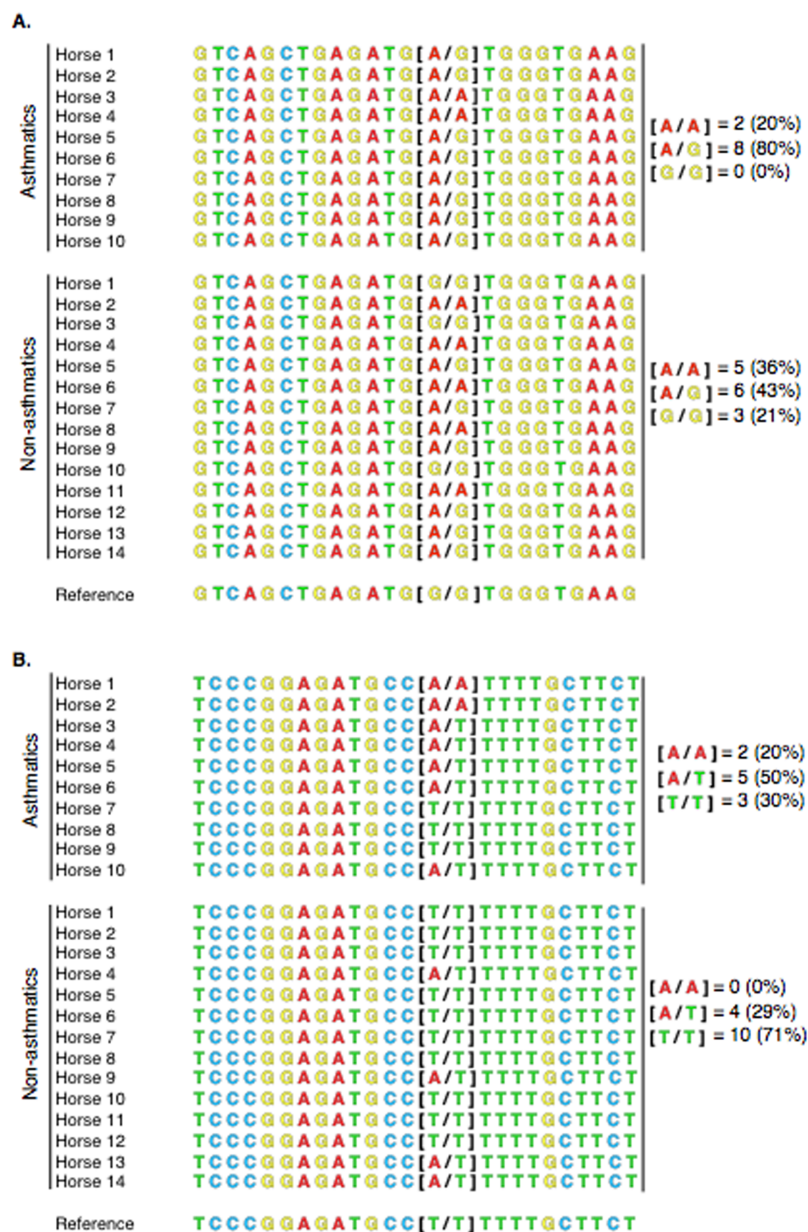
Figure 7 Alignment of PACRG (A) and RTTN (B) Sanger sequences for 10 asthmatic and 14 non-asthmatic horses with the reference genome. (A) For PACRG, eight asthmatic horses (80%) were heterozygous [A/G], 2 (20%) were homozygous mutant [A/A] and none was homozygous wild-type [G/G]. Six non-asthmatic horses (43%) were heterozygous [A/G], 5 (36%) were homozygous mutant [A/A] and 3 (21%) were homozygous wild type [G/G]. (B) For RTTN, in the asthmatic group, there were 5 (50%) heterozygous [A/T], 2 (20%) homozygous mutant [A/A], and 3 (30%) homozygous wild type [T/T] genotypes. In the non-asthmatic group, 4 (29%) were heterozygous [A/G], 10 (71%) were homozygous wild type [T/T], and none had the homozygous mutant [A/A] genotype.

Full-size ⊡ DOI: 10.7717/peerj.5759/fig-7

the heterozygous [A/G], five had the homozygous altered [A/A] and three had the homozygous WT [G/G] genotype (Fig. 7A). *RTTN* alignment in asthmatic horses yielded five heterozygous [A/T], two homozygous altered [A/A], and three homozygous WT [T/T]

genotypes. In non-asthmatic horses, four had the heterozygous [A/T], 10 had the homozygous WT [T/T], and no horse had the homozygous altered [A/A] genotype (Fig. 7B).

DISCUSSION

The goal of this study was to assess the reliability of an adapted RNA-Seq sequence variant calling workflow compared to Sanger sequencing. Sequence variant calling using RNA-Seq reads is recent practice, and reliability of results is a function of sequencing platform, depth, quality, precision of read mapping, and appropriate sequence variant calling and filtering methods. The reliability of identifying gene sequence variants using RNA-Seq has been considered uncertain. In some reports RNA-Seq was considered useful for identifying gene variants (*Quinn et al., 2013*; *Piskol, Ramaswami & Li, 2013*) while in other reports differences between RNA and DNA sequences were relatively frequent (*Sheng et al., 2016*; *Guo et al., 2017*; *Li et al., 2011*).

In this study we applied a modification of GATK best practices for sequence variant calling with RNA-Seq, and verified the results with Sanger sequencing. In 24 of 26 samples substitution variants in *PACRG* and *RTTN* were identified by both methods, while two horses' genotypes were discordant by RNA-Seq with inconsistent genotypes before and after challenge. Sanger sequencing confirmed one of the discordant genotypes, while the other could not be further assessed.

Two candidate substitution variants in the *PACRG* and *RTTN* coding sequence were identified after stringent filtering. Presence of the substitution variants was confirmed with PCR and Sanger sequencing in 24 samples. Correlation between RNA-Seq and Sanger sequencing showed that for *PACRG* both alleles of the gene were properly identified in all horses and conditions by the modified GATK workflow. For *RTTN*, two of the samples were misidentified by the workflow with alleles inconsistently identified before and after challenge. Lower mean expression suggesting lower sequencing coverage for *RTTN* might have affected the likelihood of inaccurate sequence variant calling. Nonetheless, the vast majority of alleles were identified appropriately, suggesting that the workflow is suitable for sequence variant calling in RNA-Seq at gene coverage in the 10–20 cpm range.

For supplementary analysis, SIFT was initially applied, followed by Polymorphism Phenotyping (PolyPhen) 2 and screening for non-acceptable polymorphism (SNAP) 2, to predict sequence variant effects on protein function for both substitution variants. SIFT uses phylogenetic data (*Ng & Henikoff, 2001*, *2002*, *2003*, *2006*; *Kumar, Henikoff & Ng, 2009*), while PolyPhen2 uses structural information and multiple alignments (*Adzhubei, Jordan & Sunyaev, 2013*) to predict whether or not a sequence variant may cause loss of function. The two methods often yield similar results, but limited specificity suggests that results should be interpreted with caution (*Flanagan, Patch & Ellard, 2010*). SNAP2, on the other hand, uses evolutionary, structural, solvent-access, and annotation information, as well as data from available homologs to predict whether a sequence variant is likely to have an effect on protein function (*Hecht, Bromberg & Rost, 2013*; *Bromberg & Rost, 2007*; *Hecht, Bromberg & Rost, 2015*; *McLaren et al., 2010*).

While these three approaches can yield different results (*Dong et al., 2015*; *Thusberg & Vihinen, 2009*), inferences regarding PACRG and RTTN amino acid substitutions were consistent. However, ultimately conclusions regarding the effect of substitution variants require stringent protein functional analysis, and results from this study should be considered preliminary.

The substitutions identified changed V182M (valine to methionine) and R1807W (arginine to tryptophan) in PACRG and RTTN, respectively. For PACRG, the V->M substitution minimally affected hydrophobicity and pI, while the R->W substitution in RTTN increased hydrophobicity and decreased pI. The substitution variants were considered to potentially cause loss of function and to have non-neutral effects (Tables S1 and S2). *PACRG* is a gene conserved across species (*Thumberger et al., 2012*) that shares a bi-directional promoter with parkin (*PARK2*) (*West et al., 2003*). PACRG is affiliated with axonemal doublet microtubules, and contributes to the signaling pathway that controls dynein-driven microtubule sliding (*Thumberger et al., 2012*; *Mizuno, Dymek & Smith, 2016*; *Wilson et al., 2010*; *Ikeda, 2008*). A SNV in *PACRG* was strongly associated with an increased risk of developing childhood asthma following early-life tobacco smoke exposure (*Scholtens et al., 2014*).

For the RTTN substitution variant, tryptophan is an aromatic, non-polar, and hydrophobic amino acid often buried in hydrophobic cores, while arginine is a polar and positively charged amino acid often found on outside chains (*Betts & Russell, 2003*). *RTTN* is a cilium-associated protein (*Faisst et al., 2002*) essential for assembly of centrosomes in non-motile and motile cilia (*Chen et al., 2017*). Absence of RTTN, or presence of gene sequence variants that disrupt the interaction of RTTN with SCL/TAL1 interrupting locus (*STIL*), abrogate proper ciliary development and function (*Chen et al., 2017*), and recessive mutations in *RTTN* are linked to abnormal primary ciliary development in humans (*Kheradmand Kia et al., 2012*). A change in the structural stability or binding affinity of the entire protein or the affected residue could impact ciliary structure and function. The R1807W substitution variant in the carboxy-terminal region is not immediately proximal to the suggested centrosome-targeting and STIL-binding site (*Chen et al., 2017*) but could nevertheless result in defective centrioles and hence affect cilium structure and function.

Substitution variant sequence determination in 13 RNA-Seq and 11 additional samples showed that 80% of asthmatic animals were heterozygous and 20% were homozygous altered for PACRG, and that no individual had the homozygous WT genotype [G/G]. Conversely, among non-asthmatic animals more than half were homozygous, whether WT or altered (5 [A/A] and 3 [G/G]). For *RTTN*, 20% of asthmatic horses were homozygous altered [A/A], 30% were homozygous WT [T/T] and 50% were heterozygous. Among non-asthmatics, none was homozygous altered [A/A], while 71% of horses were homozygous WT [T/T] and 29% were heterozygous. Therefore, the substitution was present in 70% of asthmatic horses and in only 30% of non-asthmatic horses (heterozygous or homozygous altered). Albeit, the variants have been identified in only a small sample of asthmatic and non-asthmatic animals, and have to be considered as variants of unknown significance (VUS). A comprehensive genome-wide association study would be

necessary to determine association between these VUS and asthma, and statistical analysis of potential associations would need to be performed prior to filtering of variants.

Pearson's Chi-squared test with Yates' continuity correction applied detected no difference in allele frequency for *PACRG*, or in genotype frequency for *RTTN* and *PACRG*, between asthmatic and non-asthmatic horses. A significantly higher frequency of the altered allele (A) in asthmatic compared to non-asthmatic horses was identified. For *PACRG*, although not significant, the *P*-value obtained after testing for differences in genotype frequency ($P = 0.213$) was lower than when testing for allele frequency ($P = 1$). This finding may be attributed to the higher proportion of asthmatic horses with a heterogeneous genotype (WT/alt for eight of 10 horses) compared to non-asthmatics (WT/alt for six of 14 horses). However, changes in allele frequency and potential roles in the pathogenesis of asthma are of unknown significance due to the small sample size in this study. Notwithstanding, a significant difference in the frequency of the *PACRG* heterozygous genotype has been reported in pulmonary tuberculosis in humans (*Udina et al., 2007*). A genome-wide interaction study also identified a *PACRG* SNP to be linked to an increased risk of developing childhood-onset asthma following early-life exposure to tobacco smoke (*Scholtens et al., 2014*). SNPs in *PACRG* also contributed to susceptibility to tuberculosis (*Bragina et al., 2016*).

For *RTTN*, the difference in allele frequency was encouraging and further analysis with a larger number of samples to assess association with asthma may be warranted. RTTN is a centrosome-associated protein first discovered for its role in axial rotation and left–right specification in the mouse embryo (*Faisst et al., 2002*). The R->W substitution altered the hydrophobicity and isoelectric point at position 1,807, and R <=> W substitutions were predicted to be most problematic in the human genome (*Majewski & Ott, 2003*). In addition, R->W substitution is generally disfavored in all protein types (*Betts & Russell, 2003*). Overrepresentation of mutated arginine was a prominent feature among disease-causing mutations in a range of conditions (*Khan & Vihinen, 2007*).

As landmarks in epithelial-environmental interaction, cilia are highly specialized cellular projections. Most vertebrate cells have a single non-motile ("primary") cilium that transduces signals from the environment or other cells, while motile cilia occur in multiples on specialized cells of the respiratory tract, oviduct and ventricles of the brain (*Reiter & Leroux, 2017*; *Szymanska & Johnson, 2012*). Motile cilia directionally propel cells or extracellular fluid through "metachronal wave" beating movements (*Eshel & Priel, 1987*; *Gheber & Priel, 1990*). The ability of motile cilia to beat in a synchronized manner requires specialized proteins that are absent in non-motile primary cilia, but otherwise both types of cilia have similar internal architecture. The main part of the cilium is the axoneme, which is comprised of nine outer microtubule doublets, one central microtubule pair (in motile multiple cilia) and a multitude of affiliated proteins. Prominent among these are tektins that stabilize microtubules and regulate axoneme length (*Steffen & Linck, 1988*), and protofilament ribbon proteins that are essential for sliding of adjacent microtubule doublets to generate ciliary movement (*Linck & Norrander, 2003*). Abnormalities in cilia are now appreciated as cause for the development of respiratory diseases, often through gene sequence variants associated with a loss of function

affecting unique ciliary proteins (*Reiter & Leroux, 2017*). Ultrastructural changes were previously reported in the ciliated epithelium of horses with severe asthma (formerly called chronic obstructive pulmonary disease), and included loss of ciliated cells (*Kaup, Drommer & Deegen, 1990*). Factors that affect beating, synchronization or orientation of motile cilia result in accumulation of mucus in airways (*Reiter & Leroux, 2017*), which is a prominent feature of equine asthma. Furthermore, hedgehog (HH) signaling is strongly linked to ciliary function, and many components of the HH signaling pathway localize to cilia (*Huangfu et al., 2003*; *Goetz & Anderson, 2010*). However, considering the relatively small number of individuals tested, allele frequencies identified in this manuscript, and their potential impact on ciliary function, remain to be confirmed on a larger scale.

PACRG may be linked to HH signaling in mice where patched1 (*PTCH1*) and *PACRG-PARK2* loci are thought to interact and regulate ciliary function in ependymal cells (*Gavino & Richard, 2011*). Interestingly, *PTCH1* is differentially expressed in asthmatic compared to non-asthmatic horses following challenge (*Tessier et al., 2017*), linking *PACRG* and an asthmatic response to environmental agents with the HH pathway. The PACRG protein associates with protofilaments (*Ikeda et al., 2007*) of the ciliary axoneme (*Thumberger et al., 2012*; *Lechtreck et al., 2009*; *Dawe et al., 2005*), has a role in ciliary morphogenesis and function (*Wilson et al., 2010*) and is directly involved in ciliary motility through control of dynein-driven microtubule sliding (*Mizuno, Dymek & Smith, 2016*). PACRG also has a variety of interacting partners such as microtubules, α- and β-tubulin and meiosis/spermiogenesis associated 1 (MEIG1) protein, heat shock protein (HSP) 70 and HSP 90 (*Ikeda, 2008*; *Li et al., 2015*; *Imai et al., 2003*). Impaired function or interaction of PACRG with its partners could weaken or impair ciliary stability and motility. The exact nature and function of methionine in protein structure remains incompletely understood, and substitutions involving methionine has been associated with several diseases (*Valley et al., 2012*). Both valine and methionine are hydrophobic residues grouped among the least polar amino acids (*Wolfenden, 2007*). Methionine is a sulfur-containing amino acid that is among the most hydrophobic residues and also easily oxidized if exposed (*Brosnan & Brosnan, 2006*). Although V->M substitutions are generally neutral, methionine's sulfur connected to a methyl group would make it less likely to interact with other proteins (*Betts & Russell, 2003*). Methionine was overrepresented as a mutant residue in several mutations associated with decrease or loss of function (*Khan & Vihinen, 2007*), including the human androgen receptor (*Kazemi-Esfarjani et al., 1993*). Although the effect of a V->M substitution is unknown, any change in PACRG structure or binding affinity could impact ciliary function, and may be of great interest in the context of severe asthma.

CONCLUSIONS

Sequence variants can be confidently called with RNA-Seq, although the required minimal coverage remains to be clearly defined and may be variable. Single point substitution variants in *PACRG* and *RTTN* were detected in all asthmatic horses, and although there was no significant difference in allele and genotype proportions between the two groups, the altered allele in the RTTN gene was more prevalent in asthmatic compared to

non-asthmatic horses. Functional cilia are crucial for lung health, and sequence variants resulting in impaired protein function are likely to have a negative impact. The significance of the substitutions in *PACRG* and *RTTN* remains to be determined but they are of potential interest for future investigations.

Competing Interests

Olivier Côté is employed by BioAssay Works LLC, Ijamsville, MD, USA. Laurence Tessier became an employee of BenchSci, Toronto, ON after completion of this study.
Dorothee Bienzle is an Academic Editor for PeerJ.

Author Contributions

- Laurence Tessier conceived and designed the experiments, performed the experiments, analyzed the data, contributed reagents/materials/analysis tools, prepared figures and/or tables, authored or reviewed drafts of the paper, approved the final draft.
- Olivier Côté conceived and designed the experiments, performed the experiments, contributed reagents/materials/analysis tools, authored or reviewed drafts of the paper, approved the final draft.
- Dorothee Bienzle conceived and designed the experiments, performed the experiments, analyzed the data, contributed reagents/materials/analysis tools, authored or reviewed drafts of the paper, approved the final draft.

Animal Ethics

The following information was supplied relating to ethical approvals (i.e., approving body and any reference numbers):

All procedures were approved by the Institutional Animal Care Committee of the University of Guelph (protocol R10-031) and conducted in compliance with the Canadian Council on Animal Care guidelines.

Data Availability

The following information was supplied regarding data availability:

The raw sequences are publicly available under study PRJNA384774 (SRP106023).

Supplemental Information

Supplemental information for this article can be found online at http://dx.doi.org/10.7717/peerj.5759#supplemental-information.

REFERENCES

Adzhubei I, Jordan DM, Sunyaev SR. 2013. Predicting functional effect of human missense mutations using PolyPhen-2. *Current Protocols in Human Genetics* **76(1)**:7.20.1–7.20.41 DOI 10.1002/0471142905.hg0720s76.

Ainsworth DM, Grünig G, Matychak MB, Young J, Wagner B, Erb HN, Antczak DF. 2003. Recurrent airway obstruction (RAO) in horses is characterized by IFN-γ and IL-8 production in bronchoalveolar lavage cells. *Veterinary Immunology and Immunopathology* **96(1–2)**:83–91 DOI 10.1016/s0165-2427(03)00142-9.

Ainsworth DM, Matychak M, Reyner CL, Erb HN, Young JC. 2009. Effects of in vitro exposure to hay dust on the gene expression of chemokines and cell-surface receptors in primary bronchial epithelial cell cultures established from horses with chronic recurrent airway obstruction. *American Journal of Veterinary Research* **70(3)**:365–372 DOI 10.2460/ajvr.70.3.365.

Ainsworth DM, Wagner B, Erb HN, Young JC, Retallick DE. 2007. Effects of in vitro exposure to hay dust on expression of interleukin-17, -23, -8, and -1β and chemokine (C-X-C motif) ligand 2 by pulmonary mononuclear cells isolated from horses chronically affected with recurrent airway disease. *American Journal of Veterinary Research* **68(12)**:1361–1369 DOI 10.2460/ajvr.68.12.1361.

Betts MJ, Russell RB. 2003. Amino acid properties and consequences of substitutions. In: Barnes MR, Gray IC, eds. *Bioinformatics for Geneticists*. Hoboken: John Wiley & Sons, 289–316 DOI 10.1002/0470867302.ch14.

Bisgaard H, Bønnelykke K, Sleiman PM, Brasholt M, Chawes B, Kreiner-Møller E, Stage M, Kim C, Tavendale R, Baty F, Pipper CB, Palmer CNA, Hakonarsson H. 2009. Chromosome 17q21 gene variants are associated with asthma and exacerbations but not atopy in early childhood. *American Journal of Respiratory and Critical Care Medicine* **179(3)**:179–185 DOI 10.1164/rccm.200809-1436oc.

Bragina EY, Tiys ES, Rudko AA, Ivanisenko VA, Freidin MB. 2016. Novel tuberculosis susceptibility candidate genes revealed by the reconstruction and analysis of associative networks. *Infection, Genetics and Evolution* **46**:118–123 DOI 10.1016/j.meegid.2016.10.030.

Bromberg Y, Rost B. 2007. SNAP: predict effect of non-synonymous polymorphisms on function. *Nucleic Acids Research* **35(11)**:3823–3835 DOI 10.1093/nar/gkm238.

Brosnan JT, Brosnan ME. 2006. The sulfur-containing amino acids: an overview. *Journal of Nutrition* **136(6)**:1636S–1640S DOI 10.1093/jn/136.6.1636s.

Chen H-Y, Wu C-T, Tang C-JC, Lin Y-N, Wang W-J, Tang TK. 2017. Human microcephaly protein RTTN interacts with STIL and is required to build full-length centrioles. *Nature Communications* **8(1)**:247 DOI 10.1038/s41467-017-00305-0.

Choudhry S, Taub M, Mei R, Rodriguez-Santana J, Rodriguez-Cintron W, Shriver MD, Ziv E, Risch NJ, Burchard EG. 2008. Genome-wide screen for asthma in Puerto Ricans:

evidence for association with 5q23 region. *Human Genetics* **123(5)**:455–468
DOI 10.1007/s00439-008-0495-7.

Couëtil LL, Cardwell JM, Gerber V, Lavoie J-P, Léguillette R, Richard EA. 2016. Inflammatory
airway disease of horses-revised consensus statement. *Journal of Veterinary Internal Medicine*
30(2):503–515 DOI 10.1111/jvim.13824.

Couëtil LL, Ward MP. 2003. Analysis of risk factors for recurrent airway obstruction in North
American horses: 1,444 cases (1990–1999). *Journal of the American Veterinary Medical
Association* **223(11)**:1645–1650 DOI 10.2460/javma.2003.223.1645.

Dawe HR, Farr H, Portman N, Shaw MK, Gull K. 2005. The Parkin co-regulated gene
product, PACRG, is an evolutionarily conserved axonemal protein that functions in
outer-doublet microtubule morphogenesis. *Journal of Cell Science* **118(23)**:5421–5430
DOI 10.1242/jcs.02659.

Debrue M, Hamilton E, Joubert P, Lajoie-Kadoch S, Lavoie J-P. 2005. Chronic exacerbation of
equine heaves is associated with an increased expression of interleukin-17 mRNA in
bronchoalveolar lavage cells. *Veterinary Immunology and Immunopathology* **105(1–2)**:25–31
DOI 10.1016/j.vetimm.2004.12.013.

**Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M,
Gingeras TR. 2013.** STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29(1)**:15–21
DOI 10.1093/bioinformatics/bts635.

Dong C, Wei P, Jian X, Gibbs R, Boerwinkle E, Wang K, Liu X. 2015. Comparison and
integration of deleteriousness prediction methods for nonsynonymous SNVs in whole exome
sequencing studies. *Human Molecular Genetics* **24(8)**:2125–2137 DOI 10.1093/hmg/ddu733.

Eshel D, Priel Z. 1987. Characterization of metachronal wave of beating cilia on frog's palate
epithelium in tissue culture. *Journal of Physiology* **388(1)**:1–8
DOI 10.1113/jphysiol.1987.sp016597.

Faisst AM, Alvarez-Bolado G, Treichel D, Gruss P. 2002. Rotatin is a novel gene required for
axial rotation and left–right specification in mouse embryos. *Mechanisms of Development*
113(1):15–28 DOI 10.1016/s0925-4773(02)00003-5.

**Ferreira MA, Matheson MC, Duffy DL, Marks GB, Hui J, Le Souëf P, Danoy P, Baltic S,
Nyholt DR, Jenkins M. 2011.** Identification of IL6R and chromosome 11q13.5 as risk loci
for asthma. *Lancet* **378**:1006–1014.

Flanagan SE, Patch A-M, Ellard S. 2010. Using SIFT and PolyPhen to predict loss-of-function
and gain-of-function mutations. *Genetic Testing and Molecular Biomarkers* **14(4)**:533–537
DOI 10.1089/gtmb.2010.0036.

**Galanter J, Choudhry S, Eng C, Nazario S, Rodríguez-Santana JR, Casal J, Torres-Palacios A,
Salas J, Chapela R, Watson HG, Meade K, LeNoir M, Rodríguez-Cintrón W, Avila PC,
Burchard EG. 2008.** ORMDL3 gene is associated with asthma in three ethnically diverse
populations. *American Journal of Respiratory and Critical Care Medicine* **177(11)**:1194–1200
DOI 10.1164/rccm.200711-1644oc.

Gavino C, Richard S. 2011. Patched1 haploinsufficiency impairs ependymal cilia function of the
quaking viable mice, leading to fatal hydrocephalus. *Molecular and Cellular Neuroscience*
47(2):100–107 DOI 10.1016/j.mcn.2011.03.004.

Gheber L, Priel Z. 1990. On metachronism in ciliary systems: a model describing the dependence
of the metachronal wave properties on the intrinsic ciliary parameters. *Cell Motility and the
Cytoskeleton* **16(3)**:167–181 DOI 10.1002/cm.970160304.

Goetz SC, Anderson KV. 2010. The primary cilium: a signalling centre during vertebrate
development. *Nature Reviews Genetics* **11(5)**:331–344 DOI 10.1038/nrg2774.

Guo Y, Ding X, Shen Y, Lyon GJ, Wang K. 2015. SeqMule: automated pipeline for analysis of human exome/genome sequencing data. *Scientific Reports* **5(1)**:srep14283 DOI 10.1038/srep14283.

Guo Y, Zhao S, Sheng Q, Samuels DC, Shyr Y. 2017. The discrepancy among single nucleotide variants detected by DNA and RNA high throughput sequencing data. *BMC Genomics* **18(S6)**:690 DOI 10.1186/s12864-017-4022-x.

Hancock DB, Romieu I, Shi M, Sienra-Monge J-J, Wu H, Chiu GY, Li H, Del Rio-Navarro BE, Willis-Owen SAG, Willis-Owens SAG, Weiss ST, Raby BA, Gao H, Eng C, Chapela R, Burchard EG, Tang H, Sullivan PF, London SJ. 2009. Genome-wide association study implicates chromosome 9q21.31 as a susceptibility locus for asthma in mexican children. *PLOS Genetics* **5(8)**:e1000623 DOI 10.1371/journal.pgen.1000623.

Hecht M, Bromberg Y, Rost B. 2013. News from the protein mutability landscape. *Journal of Molecular Biology* **425(21)**:3937–3948 DOI 10.1016/j.jmb.2013.07.028.

Hecht M, Bromberg Y, Rost B. 2015. Better prediction of functional effects for sequence variants. *BMC Genomics* **16(Suppl 8)**:S1 DOI 10.1186/1471-2164-16-s8-s1.

Himes BE, Hunninghake GM, Baurley JW, Rafaels NM, Sleiman P, Strachan DP, Wilk JB, Willis-Owen SAG, Klanderman B, Lasky-Su J, Lazarus R, Murphy AJ, Soto-Quiros ME, Avila L, Beaty T, Mathias RA, Ruczinski I, Barnes KC, Celedón JC, Cookson WOC, Gauderman WJ, Gilliland FD, Hakonarson H, Lange C, Moffatt MF, O'Connor GT, Raby BA, Silverman EK, Weiss ST. 2009. Genome-wide association analysis identifies PDE4D as an asthma-susceptibility gene. *American Journal of Human Genetics* **84(5)**:581–593 DOI 10.1016/j.ajhg.2009.04.006.

Huangfu D, Liu A, Rakeman AS, Murcia NS, Niswander L, Anderson KV. 2003. Hedgehog signalling in the mouse requires intraflagellar transport proteins. *Nature* **426(6962)**:83–87 DOI 10.1038/nature02061.

Ikeda T. 2008. Parkin-co-regulated gene (PACRG) product interacts with tubulin and microtubules. *FEBS Letters* **582(10)**:1413–1418 DOI 10.1016/j.febslet.2008.02.081.

Ikeda K, Ikeda T, Morikawa K, Kamiya R. 2007. Axonemal localization of Chlamydomonas PACRG, a homologue of the human Parkin-coregulated gene product. *Cell Motility and the Cytoskeleton* **64(11)**:814–821 DOI 10.1002/cm.20225.

Ilmarinen P, Tuomisto LE, Kankaanranta H. 2015. Phenotypes, risk factors, and mechanisms of adult-onset asthma. *Mediators of Inflammation* **2015**:1–19 DOI 10.1155/2015/514868.

Imai Y, Soda M, Murakami T, Shoji M, Abe K, Takahashi R. 2003. A product of the human gene adjacent to parkin is a component of Lewy bodies and suppresses Pael receptor-induced cell death. *Journal of Biological Chemistry* **278(51)**:51901–51910 DOI 10.1074/jbc.m309655200.

Jost U, Klukowska-Rötzler J, Dolf G, Swinburne JE, Ramseyer A, Bugno M, Burger D, Blott S, Gerber V. 2007. A region on equine chromosome 13 is linked to recurrent airway obstruction in horses. *Equine Veterinary Journal* **39(3)**:236–241 DOI 10.2746/042516407x171110.

Joubert P, Silversides DW, Lavoie JP. 2001. Equine neutrophils express mRNA for tumour necrosis factor-α, interleukin (IL)-1β, IL-6, IL-8, macrophage-inflammatory-protein-2 but not for IL-4, IL-5 and interferon-γ. *Equine Veterinary Journal* **33(7)**:730–733 DOI 10.2746/042516401776249246.

Kaup FJ, Drommer W, Deegen E. 1990. Ultrastructural findings in horses with chronic obstructive pulmonary disease (COPD) I: alterations of the larger conducting airways. *Equine Veterinary Journal* **22(5)**:343–348 DOI 10.1111/j.2042-3306.1990.tb04287.x.

Kazemi-Esfarjani P, Beitel LK, Trifiro M, Kaufman M, Rennie P, Sheppard P, Matusik R, Pinsky L. 1993. Substitution of valine-865 by methionine or leucine in the human androgen receptor causes complete or partial androgen insensitivity, respectively with distinct

androgen receptor phenotypes. *Molecular Endocrinology* **7(1)**:37–46
DOI 10.1210/mend.7.1.8446106.

Khan S, Vihinen M. 2007. Spectrum of disease-causing mutations in protein secondary structures.
BMC Structural Biology **7(1)**:56 DOI 10.1186/1472-6807-7-56.

Kheradmand Kia S, Verbeek E, Engelen E, Schot R, Poot RA, De Coo IFM, Lequin MH,
Poulton CJ, Pourfarzad F, Grosveld FG, Brehm A, De Wit MCY, Oegema R, Dobyns WB,
Verheijen FW, Mancini GMS. 2012. RTTN mutations link primary cilia function to
organization of the human cerebral cortex. *American Journal of Human Genetics* **91(3)**:533–540
DOI 10.1016/j.ajhg.2012.07.008.

Klukowska-Rötzler J, Gerber V, Leeb T. 2012. Association analysis of SNPs in the IL21R gene
with recurrent airway obstruction (RAO) in Swiss Warmblood horses. *Animal Genetics*
43(4):475–476 DOI 10.1111/j.1365-2052.2011.02289.x.

Klukowska-Rötzler J, Marti E, Lavoie J-P, Ainsworth DM, Gerber V, Zurbriggen A, Janda J.
2012a. Expression of thymic stromal lymphopoietin in equine recurrent airway obstruction.
Veterinary Immunology and Immunopathology **146(1)**:46–52
DOI 10.1016/j.vetimm.2012.01.019.

Klukowska-Rötzler J, Swinburne JE, Drögemüller C, Dolf G, Janda J, Leeb T, Gerber V. 2012b.
The interleukin 4 receptor gene and its role in recurrent airway obstruction in Swiss Warmblood
horses. *Animal Genetics* **43(4)**:450–453 DOI 10.1111/j.1365-2052.2011.02277.x.

Kumar P, Henikoff S, Ng PC. 2009. Predicting the effects of coding non-synonymous variants on
protein function using the SIFT algorithm. *Nature Protocols* **4(7)**:1073–1081
DOI 10.1038/nprot.2009.86.

Lavoie JP, Maghni K, Desnoyers M, Taha R, Martin JG, Hamid QA. 2001. Neutrophilic airway
inflammation in horses with heaves is characterized by a Th2-type cytokine profile.
American Journal of Respiratory and Critical Care Medicine **164(8)**:1410–1413
DOI 10.1164/ajrccm.164.8.2012091.

Lavoie-Lamoureux A, Moran K, Beauchamp G, Mauel S, Steinbach F, Lefebvre-Lavoie J,
Martin JG, Lavoie J-P. 2010. IL-4 activates equine neutrophils and induces a mixed
inflammatory cytokine expression profile with enhanced neutrophil chemotactic mediator
release ex vivo. *American Journal of Physiology-Lung Cellular and Molecular Physiology*
299(4):L472–L482 DOI 10.1152/ajplung.00135.2009.

Lechtreck K-F, Luro S, Awata J, Witman GB. 2009. HA-tagging of putative flagellar proteins
in *Chlamydomonas reinhardtii* identifies a novel protein of intraflagellar transport complex B.
Cell Motility and the Cytoskeleton **66(8)**:469–482 DOI 10.1002/cm.20369.

Li X, Howard TD, Moore WC, Ampleford EJ, Li H, Busse WW, Calhoun WJ, Castro M,
Chung KF, Erzurum SC, Fitzpatrick AM, Gaston B, Israel E, Jarjour NN, Teague WG,
Wenzel SE, Peters SP, Hawkins GA, Bleecker ER, Meyers DA. 2011. Importance of hedgehog
interacting protein and other lung function genes in asthma. *Journal of Allergy and Clinical
Immunology* **127(6)**:1457–1465 DOI 10.1016/j.jaci.2011.01.056.

Li X, Howard TD, Zheng SL, Haselkorn T, Peters SP, Meyers DA, Bleecker ER. 2010.
Genome-wide association study of asthma identifies RAD50-IL13 and HLA-DR/DQ regions.
Journal of Allergy and Clinical Immunology **125(2)**:328–335.e11 DOI 10.1016/j.jaci.2009.11.018.

Li Y, Xiao X, Ji X, Liu B, Amos CI. 2015. RNA-seq analysis of lung adenocarcinomas reveals
different gene expression profiles between smoking and nonsmoking patients. *Tumor Biology*
36(11):8993–9003 DOI 10.1007/s13277-015-3576-y.

Linck RW, Norrander JM. 2003. Protofilament ribbon compartments of ciliary and flagellar
microtubules. *Protist* **154(3–4)**:299–311 DOI 10.1078/143446103322454086.

Madore A-M, Tremblay K, Hudson TJ, Laprise C. 2008. Replication of an association between 17q21 SNPs and asthma in a French-Canadian familial collection. *Human Genetics* **123(1)**:93–95 DOI 10.1007/s00439-007-0444-x.

Majewski J, Ott J. 2003. Amino acid substitutions in the human genome: evolutionary implications of single nucleotide polymorphisms. *Gene* **305(2)**:167–173 DOI 10.1016/s0378-1119(03)00379-2.

Marti E, Gerber H, Essich G, Oulehla J, Lazary S. 1991. The genetic basis of equine allergic diseases 1. Chronic hypersensitivity bronchitis. *Equine Veterinary Journal* **23(6)**:457–460 DOI 10.1111/j.2042-3306.1991.tb03761.x.

Martinez FD, Vercelli D. 2013. Asthma. *Lancet* **382(9901)**:1360–1372 DOI 10.1016/s0140-6736(13)61536-6.

McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, Garimella K, Altshuler D, Gabriel S, Daly M, DePristo MA. 2010. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Research* **20(9)**:1297–1303 DOI 10.1101/gr.107524.110.

McLaren W, Pritchard B, Rios D, Chen Y, Flicek P, Cunningham F. 2010. Deriving the consequences of genomic variants with the Ensembl API and SNP Effect Predictor. *Bioinformatics* **26(16)**:2069–2070 DOI 10.1093/bioinformatics/btq330.

Mizuno K, Dymek EE, Smith EF. 2016. Microtubule binding protein PACRG plays a role in regulating specific ciliary dyneins during microtubule sliding. *Cytoskeleton* **73(12)**:703–711 DOI 10.1002/cm.21340.

Moffatt MF, Gut IG, Demenais F, Strachan DP, Bouzigon E, Heath S, Von Mutius E, Farrall M, Lathrop M, Cookson WOCM. 2010. A large-scale, consortium-based genomewide association study of asthma. *New England Journal of Medicine* **363(13)**:1211–1221 DOI 10.1056/nejmoa0906312.

Moffatt MF, Kabesch M, Liang L, Dixon AL, Strachan D, Heath S, Depner M, Von Berg A, Bufe A, Rietschel E, Heinzmann A, Simma B, Frischer T, Willis-Owen SAG, Wong KCC, Illig T, Vogelberg C, Weiland SK, Von Mutius E, Abecasis GR, Farrall M, Gut IG, Lathrop GM, Cookson WOC. 2007. Genetic variants regulating ORMDL3 expression contribute to the risk of childhood asthma. *Nature* **448(7152)**:470–473 DOI 10.1038/nature06014.

Ng PC, Henikoff S. 2001. Predicting deleterious amino acid substitutions. *Genome Research* **11(5)**:863–874 DOI 10.1101/gr.176601.

Ng PC, Henikoff S. 2002. Accounting for human polymorphisms predicted to affect protein function. *Genome Research* **12(3)**:436–446 DOI 10.1101/gr.212802.

Ng PC, Henikoff S. 2003. SIFT: predicting amino acid changes that affect protein function. *Nucleic Acids Research* **31(13)**:3812–3814 DOI 10.1093/nar/gkg509.

Ng PC, Henikoff S. 2006. Predicting the effects of amino acid substitutions on protein function. *Annual Review of Genomics and Human Genetics* **7(1)**:61–80 DOI 10.1146/annurev.genom.7.080505.115630.

Padoan E, Ferraresso S, Pegolo S, Castagnaro M, Barnini C, Bargelloni L. 2013. Real time RT-PCR analysis of inflammatory mediator expression in recurrent airway obstruction-affected horses. *Veterinary Immunology and Immunopathology* **156(3–4)**:190–199 DOI 10.1016/j.vetimm.2013.09.020.

Pirie RS. 2014. Recurrent airway obstruction: a review. *Equine Veterinary Journal* **46(3)**:276–288 DOI 10.1111/evj.12204.

Piskol R, Ramaswami G, Li JB. 2013. Reliable identification of genomic variants from RNA-seq data. *American Journal of Human Genetics* **93(4)**:641–651 DOI 10.1016/j.ajhg.2013.08.008.

Quinn EM, Cormican P, Kenny EM, Hill M, Anney R, Gill M, Corvin AP, Morris DW. 2013. Development of strategies for SNP detection in RNA-seq data: application to lymphoblastoid cell lines and evaluation using 1000 Genomes data. *PLOS ONE* **8(3)**:e58815 DOI 10.1371/journal.pone.0058815.

Ramseyer A, Gaillard C, Burger D, Straub R, Jost U, Boog C, Marti E, Gerber V. 2007. Effects of genetic and environmental factors on chronic lower airway disease in horses. *Journal of Veterinary Internal Medicine* **21(1)**:149–156 DOI 10.1111/j.1939-1676.2007.tb02941.x.

Reiter JF, Leroux MR. 2017. Genes and molecular pathways underpinning ciliopathies. *Nature Reviews Molecular Cell Biology* **18(9)**:533–547 DOI 10.1038/nrm.2017.60.

Scholtens S, Postma DS, Moffatt MF, Panasevich S, Granell R, Henderson AJ, Melén E, Nyberg F, Pershagen G, Jarvis D, Ramasamy A, Wjst M, Svanes C, Bouzigon E, Demenais F, Kauffmann F, Siroux V, Von Mutius E, Ege MJ, Braun-Fahrländer C, Genuneit J, Group GS, Brunekreef B, Smit HA, Wijga AH, Kerkhof M, Curjuric I, Imboden M, Thun GA, Probst-Hensch N, Freidin MB, Bragina EI, Deev IA, Puzyrev VP, Daley D, Park J, Becker A, Chan-Yeung M, Kozyrskyj AL, Pare P, Marenholz I, Lau S, Keil T, Lee Y-A, Kabesch M, Wijmenga C, Franke L, Nolte IM, Vonk J, Kumar A, Farrall M, Cookson WO, Strachan DP, Koppelman GH, Boezen HM. 2014. Novel childhood asthma genes interact with in utero and early-life tobacco smoke exposure. *Journal of Allergy and Clinical Immunology* **133**:885–888.

Setlakwe EL, Lemos KR, Lavoie-Lamoureux A, Duguay J-D, Lavoie J-P. 2014. Airway collagen and elastic fiber content correlates with lung function in equine heaves. *American Journal of Physiology-Lung Cellular and Molecular Physiology* **307(3)**:L252–L260 DOI 10.1152/ajplung.00019.2014.

Sheng Q, Zhao S, Li C-I, Shyr Y, Guo Y. 2016. Practicability of detecting somatic point mutation from RNA high throughput sequencing data. *Genomics* **107(5)**:163–169 DOI 10.1016/j.ygeno.2016.03.006.

Skloot GS. 2016. Asthma phenotypes and endotypes: a personalized approach to treatment. *Current Opinion in Pulmonary Medicine* **22(1)**:3–9 DOI 10.1097/mcp.0000000000000225.

Sleiman PMA, Flory J, Imielinski M, Bradfield JP, Annaiah K, Willis-Owen SAG, Wang K, Rafaels NM, Michel S, Bonnelykke K, Zhang H, Kim CE, Frackelton EC, Glessner JT, Hou C, Otieno FG, Santa E, Thomas K, Smith RM, Glaberson WR, Garris M, Chiavacci RM, Beaty TH, Ruczinski I, Orange JS, Orange JM, Allen J, Spergel JM, Grundmeier R, Mathias RA, Christie JD, von Mutius E, Cookson WOC, Kabesch M, Moffatt MF, Grunstein MM, Barnes KC, Devoto M, Magnusson M, Li H, Grant SFA, Bisgaard H, Hakonarson H. 2010. Variants of DENND1B associated with asthma in children. *New England Journal of Medicine* **362(1)**:36–44 DOI 10.1056/nejmoa0901867.

Steffen W, Linck RW. 1988. Evidence for tektins in centrioles and axonemal microtubules. *Proceedings of the National Academy of Sciences of the United States of America* **85(8)**:2643–2647 DOI 10.1073/pnas.85.8.2643.

Swinburne JE, Bogle H, Klukowska-Rötzler J, Drögemüller M, Leeb T, Temperton E, Dolf G, Gerber V. 2009. A whole-genome scan for recurrent airway obstruction in Warmblood sport horses indicates two positional candidate regions. *Mammalian Genome* **20(8)**:504–515 DOI 10.1007/s00335-009-9214-5.

Szymanska K, Johnson CA. 2012. The transition zone: an essential functional compartment of cilia. *Cilia* **1(1)**:10 DOI 10.1186/2046-2530-1-10.

Tavendale R, Macgregor DF, Mukhopadhyay S, Palmer CNA. 2008. A polymorphism controlling ORMDL3 expression is associated with asthma that is poorly controlled by

current medications. *Journal of Allergy and Clinical Immunology* **121**(4):860–863 DOI 10.1016/j.jaci.2008.01.015.

Tessier L, Côté O, Clark ME, Viel L, Diaz-Méndez A, Anders S, Bienzle D. 2017. Impaired response of the bronchial epithelium to inflammation characterizes severe equine asthma. *BMC Genomics* **18**(1):708 DOI 10.1186/s12864-017-4107-6.

Thumberger T, Hagenlocher C, Tisler M, Beyer T, Tietze N, Schweickert A, Feistel K, Blum M. 2012. Ciliary and non-ciliary expression and function of PACRG during vertebrate development. *Cilia* **1**(1):13 DOI 10.1186/2046-2530-1-13.

Thusberg J, Vihinen M. 2009. Pathogenic or not? And if so, then how? Studying the effects of missense mutations using bioinformatics methods. *Human Mutation* **30**(5):703–714 DOI 10.1002/humu.20938.

Torgerson DG, Ampleford EJ, Chiu GY, Gauderman WJ, Gignoux CR, Graves PE, Himes BE, Levin AM, Mathias RA, Hancock DB, Baurley JW, Eng C, Stern DA, Celedón JC, Rafaels N, Capurso D, Conti DV, Roth LA, Soto-Quiros M, Togias A, Li X, Myers RA, Romieu I, Van Den Berg DJ, Hu D, Hansel NN, Hernandez RD, Israel E, Salam MT, Galanter J, Avila PC, Avila L, Rodriquez-Santana JR, Chapela R, Rodriguez-Cintron W, Diette GB, Adkinson NF, Abel RA, Ross KD, Shi M, Faruque MU, Dunston GM, Watson HR, Mantese VJ, Ezurum SC, Liang L, Ruczinski I, Ford JG, Huntsman S, Chung KF, Vora H, Li X, Calhoun WJ, Castro M, Sienra-Monge JJ, Del Rio-Navarro B, Deichmann KA, Heinzmann A, Wenzel SE, Busse WW, Gern JE, Lemanske RF, Beaty TH, Bleecker ER, Raby BA, Meyers DA, London SJ, Gilliland FD, Burchard EG, Martinez FD, Weiss ST, Williams LK, Barnes KC, Ober C, Nicolae DL. 2011. Meta-analysis of genome-wide association studies of asthma in ethnically diverse North American populations. *Nature Genetics* **43**(9):887–892 DOI 10.1038/ng.888.

Udina IG, Kordicheva SI, Pospelov LE, Malenko AF, Gergert VI, Pospelov AL, Matrashkin AG, Kyzyl-Ool MM, Nachin AA, Zhivotovskiĭ LA. 2007. Study of the polymorphic markers—the PARK2 and PACRG genes due to the incidence of pulmonary tuberculosis in two districts of the Republic of Tyva. *Probl Tuberk Bolezn Legk* **7**:27–29.

Valley CC, Cembran A, Perlmutter JD, Lewis AK, Labello NP, Gao J, Sachs JN. 2012. The methionine-aromatic motif plays a unique role in stabilizing protein structure. *Journal of Biological Chemistry* **287**(42):34979–34991 DOI 10.1074/jbc.m112.374504.

Vargas A, Roux-Dalvai F, Droit A, Lavoie J-P. 2016. Neutrophil-derived exosomes: a new mechanism contributing to airway smooth muscle remodeling. *American Journal of Respiratory Cell and Molecular Biology* **55**(3):450–461 DOI 10.1165/rcmb.2016-0033oc.

Wade CM, Giulotto E, Sigurdsson S, Zoli M, Gnerre S, Imsland F, Lear TL, Adelson DL, Bailey E, Bellone RR, Blocker H, Distl O, Edgar RC, Garber M, Leeb T, Mauceli E, MacLeod JN, Penedo MCT, Raison JM, Sharpe T, Vogel J, Andersson L, Antczak DF, Biagi T, Binns MM, Chowdhary BP, Coleman SJ, Della Valle G, Fryc S, Guerin G, Hasegawa T, Hill EW, Jurka J, Kiialainen A, Lindgren G, Liu J, Magnani E, Mickelson JR, Murray J, Nergadze SG, Onofrio R, Pedroni S, Piras MF, Raudsepp T, Rocchi M, Roed KH, Ryder OA, Searle S, Skow L, Swinburne JE, Syvanen AC, Tozaki T, Valberg SJ, Vaudin M, White JR, Zody MC, Lander ES, Lindblad-Toh K, Broad Institute Genome Sequencing Platform, Broad Institute Whole Genome Assembly Team. 2009. Genome sequence, comparative analysis, and population genetics of the domestic horse. *Science* **326**(5954):865–867 DOI 10.1126/science.1178158.

Wenzel SE. 2012. Asthma phenotypes: the evolution from clinical to molecular approaches. *Nature Medicine* **18**(5):716–725 DOI 10.1038/nm.2678.

West AB, Lockhart PJ, O'Farell C, Farrer MJ. 2003. Identification of a novel gene linked to parkin via a bi-directional promoter. *Journal of Molecular Biology* **326(1)**:11–19 DOI 10.1016/s0022-2836(02)01376-1.

Wilson GR, Wang HX, Egan GF, Robinson PJ, Delatycki MB, O'Bryan MK, Lockhart PJ. 2010. Deletion of the Parkin co-regulated gene causes defects in ependymal ciliary motility and hydrocephalus in the quakingviable mutant mouse. *Human Molecular Genetics* **19(8)**:1593–1602 DOI 10.1093/hmg/ddq031.

Wolfenden R. 2007. Experimental measures of amino acid hydrophobicity and the topology of transmembrane and globular proteins. *Journal of General Physiology* **129(5)**:357–362 DOI 10.1085/jgp.200709743.

Wu H, Romieu I, Sienra-Monge J-J, Li H, Del Rio-Navarro BE, London SJ. 2009. Genetic variation in ORM1-like 3 (ORMDL3) and gasdermin-like (GSDML) and childhood asthma. *Allergy* **64(4)**:629–635 DOI 10.1111/j.1398-9995.2008.01912.x.