Article type : General Article

Exploring the genetics underpinning dynamic laryngeal collapse associated with poll flexion in Norwegian-Swedish Coldblooded Trotter racehorses

B. D. Velie^{*†1,2}, P. M. Smith^{†3}, C. T. Fjordbakk³, M. Solé², K. Jäderkvist Fegraeus², M. K. Rosengren², K. H. Røed⁴, C. F. Ihler³, G. Lindgren^{a2,5} and E. Strand^{a3}

¹University of Sydney, School of Life and Environmental Sciences, Faculty of Science, Sydney, Australia; ²Swedish University of Agricultural Sciences, Department of Animal Breeding and Genetics, Uppsala, Sweden; ³Department of Companion Animal Clinical Sciences, Norwegian University of Life Sciences, Oslo, Norway; ⁴Department of Basic Sciences and Aquatic Medicine, Norwegian University of Life Sciences, Oslo, Norway and ⁵Livestock Genetics, Department of Biosystems, KU Leuven, Leuven, Belgium.

^aShared senior authorship

[†]Shared first authorship

*Corresponding author email: brandon.velie@sydney.edu.au

Keywords: horse; genomic; harness racing; performance; upper respiratory tract disorder

Running head: Genetics of dynamic laryngeal collapse

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1111/evj.13171

This article is protected by copyright. All rights reserved.

Summary

Background: Dynamic laryngeal collapse associated with poll flexion (DLC) is the most common disorder of the upper respiratory tract in the Norwegian-Swedish Coldblooded Trotter (NSCT). The disorder, which has also been diagnosed in other breeds of trotters and gaited horses, appears to be related to anatomic phenotypes and only occurs during poll flexion when the horse is exercised “on the bit”.

Objectives: Identify genomic regions associated with DLC in the NSCT by combining a rigorous phenotyping protocol with genomic data from a high density equine genotyping array.

Study design: Prospective case/control study.

Methods: High-speed treadmill endoscopy was used to phenotype horses (n = 61) for DLC, distinguishing between cases and controls. Genome-wide association analysis of DLC status was then performed using a principal components approach with haplotype analyses subsequently performed for regions containing single nucleotide polymorphisms (SNPs) above the suggestive genome-wide significance threshold ($P < 1.0 \times 10^{-5}$).

Results: One region containing 10 SNPs (*Equus caballus* chromosome [ECA] 7: 89601935-94647192) was above the suggestive genome-wide significance threshold. Two inferred haplotypes in this region demonstrated significant differences ($p < 0.001$) between cases and controls, with the most frequent haplotype resulting in a significantly increased risk of DLC.

Main limitations: Small sample size as a result of stringent phenotyping protocols.

Conclusions: The current study highlights a candidate genomic region on ECA7 as potentially important with regard to the manifestation of DLC. Further exploration of this region and the genes included within it will bring veterinarians and researchers closer to fully understanding the biological mechanisms underlying DLC in horses.

Introduction

Dynamic laryngeal collapse associated with poll flexion (DLC) is a recently described disorder of the upper respiratory tract (URT) diagnosed in Norwegian-Swedish Coldblooded trotter (NSCT) racehorses, Standardbreds, and other gaited breeds of performance horses that exercise with a high head/neck carriage [1-4]. The condition can cause severe clinical obstruction of the URT tract and affected horses have been known to collapse during racing due to airflow obstruction [1,5]. While similar to other dynamic disorders in that it cannot be appreciated during resting endoscopic evaluation, uniquely poll flexion contributes to the disorder in susceptible horses [1,5]. As a result, dynamic endoscopy with poll flexion has been the gold standard for diagnosis in harness racehorses [2,5-6]. Thus, diagnosis often requires significant time and resources (e.g. personnel, equipment) [5,7].

As DLC appears to be associated with a combination of anatomical phenotypes, it may be possible to identify genomic regions associated with this disorder, thus facilitating genetic testing for DLC in the future. An increased knowledge of the genes involved in the manifestation of DLC in NSCTs would not only be expected to broaden understanding of the biology underlying this disorder, but it may also assist in simplifying diagnosis of the disorder. Therefore, the aim of the current study was to identify candidate genomic regions associated with DLC in a pilot sample of NSCTs using the 670K Affymetrix Axiom high-density equine genotyping array.

Materials and Methods

Participant recruitment

Sixty-one horses recruited for the study were presented to the Equine Teaching Hospital at the Norwegian School of Veterinary Science in Oslo, Norway between 1999 and 2018 for one or more of the following reasons: perceived poor racing performance, fitness

testing, abnormal respiratory noise during exercise, or physical collapse during racing. Owner consent for inclusion in this study was obtained after completion of the phenotyping protocol for each horse.

Diagnosis

After being appropriately trained to exercise on the treadmill, each horse was examined as follows: 1) The horse was equipped with standard racing tack including snaffle bit, bridle, harness, head check rein, and long reins. 2) The horse was then exercised/warmed-up for approximately 2500 metres at 6 m/s prior to insertion and securing of a videoendoscope at the level of the nasopharynx with optimal view of the larynx. 3) The horse was then exercised at 8-10 m/s on a 1.5° incline to ensure heart rate (HR) reached near maximum values and the horse was fatigued towards the end of a four to six minute period. During the first minute of this period, the horse was evaluated without tension on the reins (i.e. free head carriage), immediately followed by one minute of exercise with tension on the reins, while motivating the horse to move forward, thereby inducing poll flexion under exercise. The degree of poll flexion and tension on the bit were similar to what the horse would experience when routinely driven on a racetrack. One minute periods of free head carriage and poll flexion were then alternated in this manner until 2 to 3 one minute periods of poll flexion were evaluated endoscopically, with the number of periods evaluated depending on when the horse started to show signs of fatigue (Fig 1) [2,5,8]. Most trained NSCTs attained a heart rate over 200 bpm after the first minute of this protocol and peaked at between 220-240 bpm when they fatigued.

The continuous endoscopic videos were reviewed in slow motion to evaluate for URT abnormalities, and if present, match them to respective head positions at any given time point. The diagnosis of DLC was made on the basis of endoscopic findings at high-speed treadmill exercise as previously described [1,2,5,8]. Briefly, horses affected with DLC show signs of bilateral vocal fold collapse accompanied by some degree of arytenoid cartilage collapse (usually bilateral) at high-speed exercise during periods of poll flexion (Fig 1). The definitive diagnosis was made by two Diplomates of the European College of Veterinary Surgeons (ECVS), each with 10+ years of experience in the diagnosis of upper airway disorders in horses. Horses diagnosed with DLC were further graded as mildly, moderately, or markedly affected using a grading system described in a previous study [8]. For the purpose of this study, only moderately (n = 16) and markedly (n = 18) affected DLC horses were included as cases. Horses used as controls were free of all URT disorders, and had visually normal physiologic function of their larynx and caudal nasopharynx during all phases of the examination.

DNA isolation, genotyping, and quality control

For deoxyribonucleic acid (DNA) extraction from blood samples (collected from the horse via the left jugular vein by venipuncture), 350 μ L of blood was used and isolated on the Qiasymphony instrument using the Qiasymphony DSP DNA mini kit^a. Deoxyribonucleic acid obtained from hair roots (collected from the mane and/or tail) was extracted using a standard hair-preparation procedure. Briefly, 186 μ L Chelex 100 Resin^b and 14 μ L of proteinase K (20 mg/mL)^d were added to the sample. The mix was incubated at 56°C for 2 h and the proteinase K was inactivated for 10 min at 95°C.

Prior to quality control (QC) the data set consisted of 61 horses (27 controls, 34 cases) genotyped using the 670K+ Axiom Equine Genotyping Array. Iterative QC was then performed with the GenABEL package in R to remove population outliers, poorly genotyped data, and noisy data using the following thresholds: minor allele frequency (MAF) (<5%), missing genotypes per single nucleotide polymorphism (SNP) (>5%), missing SNPs per sample (>5%), and Hardy-Weinberg equilibrium (HWE) (first QC: $p < 1e^{-10}$; second QC applied to control group only: $FDR < 0.20$) [9].

Genome-wide association analysis

Genome-wide association (GWA) analysis was performed using the GenABEL package in R [9]. It was used to compute an autosomal genomic kinship matrix as well as perform standard K-means clustering. K-means clustering with $K = \{1, 2, \dots, 10\}$ were executed to determine the number of clusters (i.e. subpopulations). For each iteration, the sum of within-cluster sum of squares ($\sum WCSS$) was calculated and subsequently plotted vs. K. To define subpopulations, the number of clusters corresponding with the first inflection point ($K = 3$) was chosen [10]. The multidimensional scaling (MDS) plot yielded one outlier which was subsequently excluded. A visualisation of the final genomic-kinship matrix and subpopulations using MDS can be seen in Figure 2. To account for any population stratification, a GWA analysis of DLC status (controls vs. cases, 0/1) was performed using a principal components approach ('egscore' function) (PCA) with four components, determined by the screeplot function in R [9]. The first component explained 60.8% of the variation with the four components together explaining 86.7% of the variation observed. Although generally not believed to be a contributing factor for the disorder, sex of the horse was also included in the analysis as a fixed effect. Genome-wide significance for the analysis

was determined by Bonferroni correction ($P < 1.42 \times 10^{-7}$, corrected for total number of SNPs post QC) while a “suggestive” genome-wide significance threshold was set at 1.0×10^{-5} .

Haplotype analysis

Haplotype detection was performed where possible for the regions containing SNPs above the suggestive genome-wide significance threshold. Inferred haplotype block construction was completed using the program Haploview with default values (v4.2) [11]. A linkage disequilibrium (LD) plot was constructed to visualise the haplotype block used in further analysis. The significance and risk prediction for DLC (as measured by the odds ratio) of the derived haplotype block combinations was determined by the haplo.cc program from the Haplo.stats package (v1.7.9) in R [9,12-13]. Simulated P-values were based on 10,000 replicates. The most frequent haplotype was used as the reference baseline and only haplotypes with frequencies greater than 5% were included. Candidate gene screening was subsequently carried out for the region(s) containing haplotype blocks using the bioinformatics database Ensembl (<http://www.ensembl.org/>). Haplotype block regions were used to generate a list of annotated genes using the Ensembl Biomart function. The resulting candidate gene list was then piped into the PANTHER Classification system in order to obtain an overview of the biological processes affected by the candidate genes [14-15]

Results

During QC, 14.4% (96796/670806), 32.7% (219516/670806), and 0.46% (3067/670806) of SNPs were excluded based on SNP call rate, MAF, and HWE, respectively. As a result, only 351,427 autosomal SNPs (52.4% of all SNPs) and 59 horses (26 controls, 33 cases) were available for association analysis. Examination of the corrected QQ plot revealed moderate overcorrection (Fig 3; Supplementary Item 1). No SNPs were above the Bonferroni

significance threshold; however, one region on *Equus caballus* chromosome (ECA) 7 was above the suggestive genome-wide significance threshold (Table 1; Fig 3). Figure 4 shows the haplotype LD block inferred in the genomic region (EquCab2 - chr 7: 89601935–94647192), containing ten of the most significant SNPs. Haplotype association analysis results belonging to the inferred haplotype LD block on chromosome 7 are summarised in Table 2. Only two inferred haplotypes showed statistical differences between cases and controls, resulting in significant associations (unadjusted and simulated p-values <0.001) for the opposite forms (Table 2). A significant protective association for DLC was observed for one haplotype combination (with frequencies less than 5% in cases versus frequencies above 26% in controls) assuming an additive genetic trait (Odds Ratios= 0.04, 95% CI = 0.01 to 0.23, unadjusted p-value = 0.0004). The most frequent haplotype (72.2% in cases versus 38.9% in controls) resulted in a significantly increased risk of DLC assuming additive genetic trait (unadjusted p-value = 0.0000). Fifteen candidate genes were identified in the genomic region containing the ECA7 haplotype block (Table 3). A summary of the known biological processes, molecular functions, and pathways associated with these candidate genes is presented in Figure 5.

Discussion

Although generally considered not as statistically robust as large principal studies, pilot studies provide crucial insight into both the feasibility and potential outcomes of future studies. Clinical research on DLC has established manifestation of the disorder as a complex phenotype with no apparent age or sex predisposition. However, despite advances in our understanding of both the manifestation and diagnosis of DLC, little is known about the genetic aetiology of the disorder [1-8]. As such, the current pilot study was performed with the primary aim of identifying candidate genomic regions associated with manifestation of

DLC in the NSCT breed. While no SNPs ultimately demonstrated genome-wide significance (GWS) in the current study, associations between DLC and 12 SNPs were above the suggestive GWS threshold. Associations with borderline GWS have been shown to be successfully replicated more than 70% of the time with many associations achieving substantially lower p-values when additional data is obtained. Considering what appears to be a slight over-correction for genomic inflation in the current study, 12 SNPs above the suggestive GWS threshold is an encouraging result for a genome scan of less than 100 individuals and suggest that the disorder may be less complex than anticipated [16].

Strict phenotyping criteria were applied in the current study and while uncorrected QQ plots revealed some inflation due to population stratification, haplotype analyses of the regions encompassing the suggestive GWS SNPs indicate that further exploration into the genetic aetiology of DLC is likely warranted. Although failing to reach GWS, 11 of the top 12 SNPs were located on ECA7 with one haplotype displaying what appears to be a protective association for DLC. Previous studies exploring the genetic aetiology of other URT disorders in horses have identified candidate regions on ECA1, ECA3, ECA15, ECA18, ECA21, and ECA23; however, ECA7 has never previously been associated with a URT disorder in horses [17-19]. While the association of DLC with regions on ECA7 may be a side effect of the limited sample size in the current study, it may also reflect the complexity of URT disorders. Historically, studies exploring the genetic aetiology of URT disorders in horses tend to focus on recurrent laryngeal neuropathy (RLN) due to the relatively high frequency of the disorder across breeds and the unfavourable affect the disorder has on performance [17-19]. As a result, the vast majority of information available on the genetic aetiology of URT disorders in horses skews towards RLN. Genomic regions previously associated with RLN are likely to only overlap with regions associated with DLC if manifestation of the two disorders is similar. As the two disorders are actually quite different

in their underlying aetiology, the lack of overlap between previous studies of URT disorders and the current study is not entirely unexpected [1-8,17-20]. RLN is a distal axonopathy of the recurrent laryngeal nerve in large horse breeds, while DLC is seemingly a combination of several anatomic variants, such as a rostrally positioned larynx which leads to external compression of the larynx during poll flexion [7,17-19]. A recent study has even indicated that a number of different “dynamic” URT disorders are breed related, suggesting an anatomic or functional reason for the DLC airway collapse in response to the greatly increased airflows and pressure gradients generated during strenuous exercise [2]. Furthermore, it is important to note that previous studies of URT disorders in horses included substantially larger sample sizes ($n > 400$) than the current study, thus allowing them to perform more robust analyses and produce more concrete results [17-19].

Nevertheless, the results of the current study, although limited in strength and scope, provide useful information regarding the genetic aetiology of DLC. Despite what appears to be a propensity for DLC among gaited breeds, the results of the current study fail to indicate any obvious links between the “gait-keeper” gene (i.e. *doublesex and mab-3 related transcription factor 3*) or any other gene on ECA23 and manifestation of DLC in NSCTs [21-25]. It is possible that the high frequency of DLC observed in trotters and gaited breeds is more a product of the type of work typically performed by these breeds rather than a genetic predisposition. However, this cannot be said with any certainty without further study. Conversely, the genomic region (EquCab2: 87,060,344-94,647,192 bp) located on ECA7 highlighted by the genome-wide association and haplotype analyses contains multiple candidate genes for manifestation of DLC in the horse [26]. While additional studies containing larger sample sizes and more robust analyses are undoubtedly required to fully explore the significance of this region and the genes contained within it, two genes located within the haplotype region and another located just downstream to the region stand out as

clear candidate genes for future studies. Neural EGFL like 1 (*NELLI*), located slightly downstream of the haplotype block, has been associated with respiratory failure and skeletal abnormalities in mice, thus making it a plausible candidate gene for further exploration in the horse [26-28]. On the other hand, Anoctamin 3 (*ANO3*) and Anoctamin 5 (*ANO5*) are two candidate genes located within the haplotype block identified in the current study making them outstanding candidates for future studies of DLC. In humans, *ANO3* has been linked to craniocervical dystonia, a disorder affecting muscles of the upper limbs, neck, and larynx [29]. Furthermore, *ANO5* has been associated with disorders involving both bone and muscle weakness in humans [30-31]. As a result, both *ANO3* and *ANO5* appear to warrant further exploration in the horse – particularly in regard to genetic studies of DLC and other URT disorders.

Consequently, although the results of this study will require validation in a larger data set, the current study highlights a candidate genomic region as potentially important in regards to the manifestation of DLC. The supplementary knowledge gained from this pilot study together with other genomic studies of URT disorders will conceivably bring veterinarians and researchers closer to fully understanding the biological mechanisms underlying URT disorders in horses.

Authors' declaration of interests

No competing interests have been declared.

Ethical animal research

All experimental procedures and sample collection methods were approved by the Ethics Committee for Animal Experiments in Uppsala, Sweden [Number: C 121/14] and/or the Norwegian School of Veterinary Medicine, NMBU.

Owner informed consent

Informed consent from owners was obtained for all horses included in the study.

Data accessibility statement

The data that support the findings of this study are available from Swedish Trotting Association and the Norwegian Trotting Association. Restrictions apply to the availability of these data, which were used under license for this study. Data are available from the authors with the permission of Swedish Trotting Association (Stockholm, Sweden) and the Norwegian Trotting Association (Oslo, Norway).

Sources of funding

This work was supported by the Swedish-Norwegian Foundation for Equine Research, H-15-47-075, <http://hastforskning.se>.

Acknowledgements

We want to thank the owners of the horses included in the study for their and their horses' contributions.

Authorship

B. Velie, K. Jäderkvist Fegraeus, C. Ihler, G. Lindgren and E. Strand designed research. P. Smith, C. Fjordbakk, K. Jäderkvist Fegraeus, M. Rosengren, K. Røed, C. Ihler and E. Strand collected the data and prepared the samples for analysis. B. Velie and M. Solé analysed the data. All authors prepared the manuscript and gave their final approval.

Manufacturers' addresses

^aQiagen, Hilden, Germany.

^bBio-Rad Laboratories, Hercules, California, USA.

^cMerck KGaA, Darmstadt, Germany.

Tables

Table 1: Top SNPs resulting from genome-wide association analysis for dynamic laryngeal collapse in Norwegian-Swedish Coldblooded trotter racehorses.

SNP	Location (ECA:bp) [†]	MAF	Raw p-value
AX-103214483	7:94,548,582	0.169	2.14E-06
AX-103205109	7:94,559,492	0.169	2.14E-06
AX-103401449	7:94,647,192	0.169	2.14E-06
AX-103708022	7:89,601,935	0.407	2.50E-06
AX-103116592	7:92,271,358	0.203	3.73E-06
AX-104105361	7:87,060,344	0.314	6.11E-06
AX-103891852	7:94,456,760	0.175	6.61E-06
AX-103809706	7:94,289,271	0.161	6.62E-06
AX-104940513	7:94,369,612	0.161	6.62E-06
AX-104106925	7:94,404,338	0.161	6.62E-06
AX-103914175	7:94,307,173	0.280	7.73E-06
AX-104012933	27:32,797,033	0.161	7.99E-06
AX-104909886	7:86,888,130	0.305	1.05E-05
AX-103187042	7:92,325,626	0.119	2.03E-05
AX-104481144	9:57,243,076	0.127	2.44E-05
AX-103954526	7:89,252,328	0.112	3.17E-05
AX-103940964	7:87,355,384	0.254	3.23E-05
AX-103469574	7:87,111,433	0.263	3.54E-05
AX-103964459	9:52,532,081	0.203	3.91E-05
AX-103351171	7:91,581,325	0.167	4.05E-05
AX-104614049	7:88,973,990	0.353	4.06E-05
AX-104709167	7:87,353,824	0.198	4.63E-05
AX-104900016	7:93,619,157	0.119	4.66E-05
AX-104708054	22:12,260,923	0.271	4.91E-05
AX-104913666	25:9,478,305	0.112	4.94E-05

Red line = Suggestive genome-wide significance threshold ($P < 1.0 \times 10^{-5}$)

[†]EquCab2

Table 2: Estimated haplotype scores, frequencies, odds ratios and association with dynamic laryngeal collapse of the inferred haplotype block on chromosome 7.

Haplotypes (SNPs numbers [†])										Score	Freq. (controls)	Freq. (cases)	Odds Ratio (95% CI)	p-value	Sim. p-value
2	3	4	5	6	7	8	9	1 0	1 1						
G	T	G	C	G	C	T	A	T	G	-3.53	26.63	4.55	0.04 (0.01-0.23)	<0.001	<0.001
G	C	T	C	T	T	C	G	C	A	-1.44	16.07	7.58	0.12 (0.02-0.60)	0.1	0.1
G	C	T	T	T	T	C	G	C	A	0.23	9.11	11.16	0.45 (0.08-2.49)	0.8	0.9
A	C	T	T	T	T	C	G	C	A	4.06	38.95	72.17	1.00	<0.001	<0.001

Freq. = frequencies in %

Significant results are bolded

Sim. P-Value = P-Value adjusted by using 100000 permutations

[†]SNP numbers correspond with Figure 4

Table 3: List of candidate genes from PANTHER analysis of dynamic laryngeal collapse in Norwegian-Swedish Coldblooded trotter racehorses.

Gene ID	Gene Name	Gene Symbol	PANTHER Family/Subfamily
ENSECAT00000016674	Protein lin-7 homolog	LIN7C	protein lin-7 homolog c (PTHR14063:SF5)
ENSECAT00000029072	Uncharacterised protein	unassigned	fanconi anaemia group f protein (PTHR14449:SF2)
ENSECAT00000013541	Coiled-coil domain containing 34	CCDC34	coiled-coil domain-containing protein 34 (PTHR23247:SF2)
ENSECAT00000013188	Gamma-butyrobetaine hydroxylase 1	BBOX1	gamma-butyrobetaine dioxygenase (PTHR10696:SF33)
ENSECAT00000008854	Uncharacterised protein	unassigned	leucine zipper protein 2 (PTHR22414:SF0)
ENSECAT00000010906	Mucin 15, cell surface associated	MUC15	mucin-15 (PTHR45427:SF1)
ENSECAT00000026937	Anoctamin	ANO5	anoctamin-5 (PTHR12308:SF23)
ENSECAT00000009410	Anoctamin	ANO3	anoctamin-3 (PTHR12308:SF16)
ENSECAT00000016720	Brain-derived neurotrophic factor	BDNF	brain-derived neurotrophic factor (PTHR11589:SF3)
ENSECAT00000017282	Kinesin-like protein	KIF18A	kinesin-like protein kif18a (PTHR24115:SF399)
ENSECAT00000027022	Solute carrier family 17 member 6	SLC17A6	vesicular glutamate transporter 2 (PTHR11662:SF201)
ENSECAT00000005073	Fin bud initiation factor homolog	FIBIN	fin bud initiation factor homolog (PTHR31185:SF0)
ENSECAT00000007548	Growth arrest specific 2	GAS2	growth arrest-specific protein 2 (PTHR46756:SF9)
ENSECAT00000014300	Leucine rich repeat containing G protein-coupled receptor 4	LGR4	leucine-rich repeat-containing g-protein coupled receptor 4 (PTHR24372:SF67)
ENSECAT00000011794	Solute carrier family 5 member 12	SLC5A12	sodium-coupled monocarboxylate transporter 2 (PTHR42985:SF15)

Figure Legends

Fig 1: Head positioning and level of flexion during diagnosis of dynamic laryngeal collapse in Norwegian-Swedish Coldblooded Trotter (NSCT) racehorses

- a) NSCT exercised without tension on the bit and reins (i.e. free head carriage)
- b) NSCT exercised with tension on the bit and reins (i.e. poll flexion)
- c) “Normal” videoendoscopic image of NSCT being exercised without tension on the bit and reins
- d) “Marked DLC” videoendoscopic image of NSCT being exercised with tension on the bit and reins

Fig 2: Visualisation of multidimensional scaling, stratified by dynamic laryngeal collapse status (case/control), performed on the genomic-kinship matrix

Fig 3: Corrected QQ plot and Manhattan plot of the genome-wide association analysis for dynamic laryngeal collapse in Norwegian-Swedish Coldblooded Trotter racehorses

Uncorrected lambda is presented in the QQ plot. The red horizontal line indicates the genome-wide significance level (1.42×10^{-7}) and the black horizontal line indicates the suggestive genome-wide significance level (1.0×10^{-5})

Fig 4: Linkage disequilibrium plot identified from the 11 suggestive SNPs on Chromosome 7. In the LD plot, those SNPs bolded are the SNPs that form the haplotype. Red diamonds represent D' values equal to 1 as a percentage, lower values of D' are given within the boxes in shades of pink to white

Fig 5: Summary information from the functional classification analysis of candidate genes in PANTHER.

a) PANTHER biological process classification: the function of the protein in the context of a larger network of proteins that interact to accomplish a process at the level of the cell or organism.

b) PANTHER molecular function classification: the function of the protein by itself or with directly interacting proteins at a biochemical level.

c) PANTHER pathway classification: similar to biological process, but a pathway also explicitly specifies the relationships between the interacting molecules.

Percentage of gene hits against total number of genes is boxed.

Percentage of gene hits against total number of process hits is circled.

References

1. Strand, E., Hanche-Olsen, S., Gronvold, A.M.R. and Mellum, C.N. (2004) Dynamic bilateral arytenoid and vocal fold collapse associated with head flexion in 5 Norwegian Coldblooded Trotter racehorses. *Equine Vet. Educ.* **16**, 242-250.
2. Strand, E., Fjordbakk, C.T., Sundberg, K., Spangen, L., Lunde, H. and Hanche-Olsen, S. (2012) Relative prevalence of upper respiratory tract obstructive disorders in two breeds of harness racehorses (185 cases: 1998-2006). *Equine Vet. J.* **44**, 518-523.
3. Hanche-Olsen, S., Rannem, L. and Strand, E. (2010) Bilateral dynamic laryngeal collapse associated with collection in “high poll flexion” in a gaited Icelandic horse. *Pferdeheilkunde* **26**, 810-813.
4. McCarrel, T.M. and Woodie, J.B. (2015) Update on laryngeal disorders and treatment. *Vet. Clin. North Am. Equine Pract.* **31**, 13-26.

- Accepted Article
5. Strand, E., Fjordbakk, C.T., Holcombe, S.J., Risberg, A. and Chalmers, H.J. (2009) Effect of poll flexion and dynamic laryngeal collapse on tracheal pressure in Norwegian Coldblooded Trotter racehorses. *Equine Vet. J.* **41**, 59-64.
 6. Strand, E. and Skjerve, E. (2012) Complex dynamic upper airway collapse: associations between abnormalities in 99 harness racehorses with one or more dynamic disorders. *Equine Vet. J.* **44**, 524-528.
 7. Fjordbakk, C.T., Chalmers, H.J., Holcombe, S.J. and Strand, E. (2013) Results of upper airway radiography and ultrasonography predict dynamic laryngeal collapse in affected horses. *Equine Vet. J.* **45**, 705-710.
 8. Fjordbakk, C.T., Strand, E. and Hanche-Olsen, S. (2008) Surgical and conservative management of bilateral dynamic laryngeal collapse associated with poll flexion in harness race horses. *Vet. Surg.* **37**, 501-507.
 9. R Development Core Team (2015) 'R: A Language and Environment for Statistical Computing'. R Foundation for Statistical Computing, Vienna Austria.
 10. Everitt, B.S. and Hothorn, T. (2010) A Handbook of Statistical Analyses Using R. Chapman & Hall/CRC, Boca Raton Florida.
 11. Barret, J.C., Fry, B., Maller, J. and Daly, M.J. (2005) Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* **21**, 263-265.
 12. Schaid, D.J., Rowland, C.M., Tines, D.E., Jacobson, R.M. and Poland, G.A. (2002) Score tests for association of traits with haplotypes when linkage phase is ambiguous. *Am. J. Hum. Genet.* **70**, 425-434.
 13. Lake, S.L., Lyon, H., Tantisira, K., Silverman, E.K., Weiss, S.T., Laird, N.M. and Schaid, D.J. (2003) Estimation and tests of haplotype-environment interaction when linkage phase is ambiguous. *Hum. Hered.* **55**, 56-65.

14. Thomas, P.D., Campbell, M.J., Kejariwal, A., Mi, H.Y., Karlak, B., Daverman, R., Diemer, K., Muruganujan, A. and Narechania, A. (2003) PANTHER: a library of protein families and subfamilies indexed by function. *Genome Res.* **13**, 2129-2141.
15. Mi, H., Poudel, S., Muruganujan, A., Casagrande, J.T. and Thomas, P.D. (2016) PANTHER version 10: expanded protein families and functions, and analysis tools. *Nucleic Acids Res.* **44**, D336-342.
16. Panagiotou, O.A. and Ioannidis, J.P.A. (2012) What should the genome-wide significance threshold be? Empirical replication of borderline genetic associations. *Int. J. Epidemiol.* **41**, 273-286.
17. Brooks, S.A., Stick, J., Braman, A., Palermo, K., Robinson, N.E., and Ainsworth, D.M. (2018) Identification of loci affecting sexually dimorphic patterns for height and recurrent laryngeal neuropathy risk in American Belgian Draft Horses. *Physiol. Genomics* **50**, 1051-1058.
18. Boyko, A.R., Brooks, S.A., Behan-Braman, A., Castelhana, M., Corey, E., Oliveira, K.C., Swinburne, J.E., Todhunter, R.J., Zhang, Z., Ainsworth, D.M. and Robinson N.E. (2014) Genomic analysis establishes correlation between growth and laryngeal neuropathy in Thoroughbreds. *BMC Genomics* **15**, 259.
19. Dupuis, M.C., Zhang, Z., Druet, T., Denoix, J.M., Charlier, C., Lekeux, P. and Georges, M. (2011) Results of a haplotype-based GWAS for recurrent laryngeal neuropathy in the horse. *Mamm. Genome* **22**, 613-620.
20. Fjordbakk, C.T., Revold, T., Goodwin, D. and Piercy, R.J. (2015) Histopathological assessment of intrinsic laryngeal musculature in horses with dynamic laryngeal collapse. *Equine Vet. J.* **47**, 603-608.
21. Andersson, L.S., Larhammar, M., Memic, F., Wootz, H., Schwochow, D., Rubin, C.J., Patra, K., Arnason, T., Wellbring, L., Hjälm, G., Imsland, F., Petersen, J.L.,

- McCue, M.E., Mickelson, J.R., Cothran, G., Ahituv, N., Roepstorff, L., Mikko, S., Vallstedt, A., Lindgren, G., Andersson, L. and Kullander, K. (2012) Mutations in DMRT3 affect locomotion in horses and spinal circuit function in mice. *Nature* **488**, 642-646.
22. Promerová, M., Andersson, L.S., Juras, R., Penedo, M.C., Reissmann, M., Tozaki, T., Bellone, R., Dunner, S., Horin, P., Imsland, F., Mikko, S., Modry, D., Roed, K.H., Schwochow, D., Vega-Pla, J.L., Mehrabani-Yeganeh, H., Yousefi-Mashouf, N., Cothran, G.E., Lindgren, G. and Andersson, L. (2014) Worldwide frequency distribution of the “Gait keeper” mutation in the DMRT3 gene. *Anim. Genet.* **45**, 274-282.
23. Jäderkvist Fegraeus, K., Hirschberg, I., Árnason, T., Andersson, L., Velie, B.D., Andersson, L.S. and Lindgren, G. (2017) To pace or not to pace: a pilot study of four- and five-gaited Icelandic horses homozygous for the DMRT3 ‘Gait Keeper’ mutation. *Anim. Genet.* **48**, 694-697.
24. Jäderkvist Fegraeus, K., Johansson, L., Mäenpää, M., Mykkänen, A., Andersson, L.S., Velie, B.D., Andersson, L., Árnason, T. and Lindgren, G. (2015) Different DMRT3 Genotypes Are Best Adapted for Harness Racing and Riding in Finnhorses. *J. Hered.* **106**, 734-740.
25. Jäderkvist Fegraeus, K., Lawrence, C., Petäjistö, K., Johansson, M.K., Wiklund, M., Olsson, M., Andersson, L., Andersson, L.S., Røed, K.H., Ihler, C.F., Strand, E., Lindgren, G. and Velie, B.D. (2017) Lack of significant associations with early career performance suggest no link between the DMRT3 “Gait Keeper” mutation and precocity in Coldblooded trotters. *PLoS One* **10**, e0177351.
26. Zerbino, D.R., Achuthan, P., Akanni, W., Amode, M.R., Barrell, D., Bhai, J., Billis, K., Cummins, C., Gall, A., Girón, C.G., Gil, L., Gordon, L., Haggerty, L., Haskell,

- E., Hourlier, T., Izuogu, O.G., Janacek, S.H., Juettemann, T., To, J.K., Laird, M.R., Lavidas, I., Liu, Z., Loveland, J.E., Maurel, T., McLaren, W., Moore, B., Mudge, J., Murphy, D.N., Newman, V., Nuhn, M., Ogeh, D., Ong, C.K., Parker, A., Patricio, M., Riat, H.S., Schullenburg, H., Sheppard, D., Sparrow, H., Taylor, K., Thormann, A., Vullo, A., Walts, B., Zadissa, A., Frankish, A., Hunt, S.E., Kostadima, M., Langridge, N., Martin, F.J., Muffato, M., Perry, E., Ruffler, M., Staines, D.M., Trevanion, S.J., Aken, B.L., Cunningham, F., Yates, A. and Flicek, P. (2018) Ensembl 2018. *Nucleic Acids Res.* **46**, 754-761.
27. Zhang, X., Zara, J., Siu, R.K., Ting, K. and Soo C. (2010) The role of NELL-1, a growth factor associated with craniosynostosis, in promoting bone regeneration. *J. Dent. Res.* **89**, 865-878.
28. Desai, J., Shannon, M.E., Johnson, M.D., Ruff, D.W., Hughes, L.A., Kerley, M.K., Carpenter, D.A., Johnson, D.K., Rinchik, E.M. and Culiati, C.T. (2006) Nell1-deficient mice have reduced expression of extracellular matrix proteins causing cranial and vertebral defects. *Hum. Mol. Genet.* **15**, 1329-1341.
29. Charlesworth, G., Plagnol, V., Holmström, K.M., Bras, J., Sheerin, U.M., Preza, E., Rubio-Agusti, I., Ryten, M., Schneider, S.A., Stamelou, M., Trabzuni, D., Abramov, A.Y., Bhatia, K.P. and Wood, N.W. (2012) Mutations in ANO3 cause dominant craniocervical dystonia: ion channel implicated in pathogenesis. *Am. J. Hum. Genet.* **91**, 1041-1050.
30. Jin, L., Liu, Y., Sun, F., Collins, M.T., Blackwell, K., Woo, A.S., Reichenberger, E.J. and Hu, Y. (2017) Three novel ANO5 missense mutations in Caucasian and Chinese families and sporadic cases with gnathodiaphyseal dysplasia. *Sci. Rep.* **7**, 40935.
31. Papadopoulos, C., Laforet, P., Nectoux, J., Stojkovic, T., Wahbi, K., Carlier, R.Y., Carlier, P.G., Leonard-Louis, S., Leturcq, F., Romero, N., Eymard, B. and Behin, A. (2017) Hyperckemia and myalgia are common presentations of anoctamin-5-related myopathy in French patients. *Muscle Nerve* **56**, 1096-1100.

Supporting Information

Supplementary Item 1: Uncorrected QQ plot

† Uncorrected lambda is presented in the QQ plot

Figure 2

Figure 3

Figure 5