


RESEARCH ARTICLE

Open Access



# Exploring the genetics of trotting racing ability in horses using a unique Nordic horse model

Brandon D. Velie<sup>1,2\*</sup> , Mette Lillie<sup>3,4†</sup>, Kim Jäderkvist Fegraeus<sup>1</sup>, Maria K. Rosengren<sup>1</sup>, Marina Solé<sup>1</sup>, Maja Wiklund<sup>5</sup>, Carl-Fredrik Ihler<sup>6</sup>, Eric Strand<sup>6</sup> and Gabriella Lindgren<sup>1,7</sup>

## Abstract

**Background:** Horses have been strongly selected for speed, strength, and endurance-exercise traits since the onset of domestication. As a result, highly specialized horse breeds have developed with many modern horse breeds often representing closed populations with high phenotypic and genetic uniformity. However, a great deal of variation still exists between breeds, making the horse particularly well suited for genetic studies of athleticism. To identify genomic regions associated with athleticism as it pertains to trotting racing ability in the horse, the current study applies a pooled sequence analysis approach using a unique Nordic horse model.

**Results:** Pooled sequence data from three Nordic horse populations were used for  $F_{ST}$  analysis. After strict filtering,  $F_{ST}$  analysis yielded 580 differentiated regions for trotting racing ability. Candidate regions on equine chromosomes 7 and 11 contained the largest number of SNPs ( $n = 214$  and  $147$ , respectively). GO analyses identified multiple genes related to intelligence, energy metabolism, and skeletal development as potential candidate genes. However, only one candidate region for trotting racing ability overlapped a known racing ability QTL.

**Conclusions:** Not unexpected for genomic investigations of complex traits, the current study identified hundreds of candidate regions contributing to trotting racing ability in the horse. Likely resulting from the cumulative effects of many variants across the genome, racing ability continues to demonstrate its polygenic nature with candidate regions implicating genes influencing both musculature and neurological development.

**Keywords:** Athleticism, Conformation, Genomic, Performance, Racehorse

## Background

As genomics improves and enables the design of more targeted studies relating genotypes to phenotypes, the opportunity for non-model organisms continues to expand - facilitating greater opportunities to gain novel insight into the mechanisms regulating biological homeostasis and health [1, 2]. Genomic studies of natural model species, domestic species in particular, give a complimentary view of genotype-phenotype relationships compared with

the knowledge gained from the study of humans and experimental organisms [2]. Since the onset of domestication, horses have been strongly selected for, among other things, speed, strength, and endurance-exercise traits [3]. This diverse and, at times, divergent selection has ultimately led to the development of highly specialized horse breeds. Within the last 400 years, breed specialization has focused primarily on preserving and improving traits related to aesthetics and performance [3]. As a result, most horse breeds today are closed populations with high phenotypic and genetic uniformity within breed. However, a great deal of variation continues to exist among breeds [3]. This variation, combined with breed specialization, has made the horse particularly well suited for genetic studies of locomotion patterns and provides a unique opportunity for genetic studies of

\* Correspondence: [brandon.velie@sydney.edu.au](mailto:brandon.velie@sydney.edu.au)

<sup>†</sup>Brandon D. Velie and Mette Lillie contributed equally to this work.

<sup>1</sup>Department of Animal Breeding and Genetics, Swedish University of Agricultural Sciences, Uppsala, Sweden

<sup>2</sup>School of Life and Environmental Sciences, University of Sydney, Sydney, Australia

Full list of author information is available at the end of the article



athleticism [1–5]. Generally speaking, athleticism describes the physical qualities that are characteristic of athletes and typically refers to traits such as strength, fitness, and agility. Many modern day horse breeds exemplify some if not all of these traits, with shared selective pressures within breeds (e.g. health, fertility traits, conformation) and divergent selection between breeds (e.g. speed vs strength) yielding a wide range of athletic phenotypes [3, 6].

Here we apply a genomic approach to investigate athletic phenotypes associated with trotting racing ability (TRA) using a unique Nordic horse model consisting of the Norwegian-Swedish Coldblooded trotter (NSCT), the North Swedish Draught horse (NSD), and the Standardbred trotter (SB) (Fig. 1). Although both the NSCT and the NSD are horse breeds derived from the original North-Swedish horse, a small, heavy horse traditionally used in agriculture and forestry work, selection for traits beneficial to agricultural work only continues in the NSD [7–11]. Since the 1960s, the NSCT has been intensively selected for harness racing performance and is now considered a true racing breed [10–13]. As a result, a remarkable improvement in the racing performance of NSCTs has occurred during the last half-century. However, it is also well established that some degree of cross-breeding occurred between NSCT and SBs, a significantly faster breed of horse from a different gene pool, before obligatory paternity testing was introduced in Sweden in 1969 [10–13]. Consequently, the improvement in NSCT racing performance may be partially explained by a marked increase of favorable genetic variants originating from SBs. It is this specific attempt at gaining a competitive racing advantage that makes the NSCT ideal for genomic studies investigating TRA phenotypes.

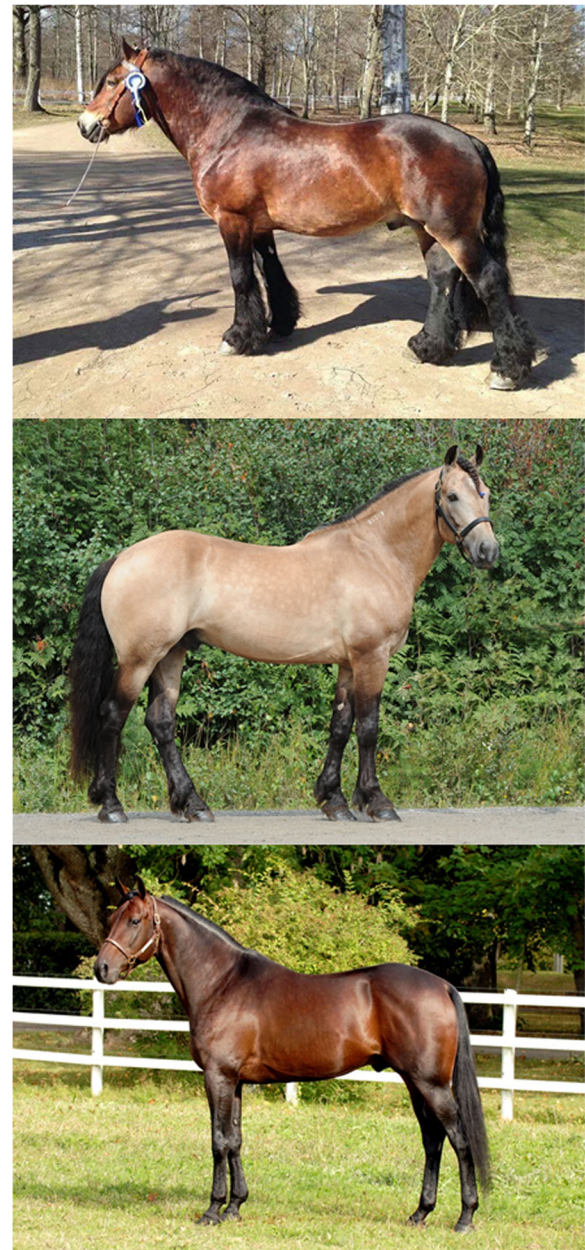
Despite a dispersed history of crossbreeding with SBs, the relationship between the NSCT and NSD remains closer than either of the breeds with the SB [10]. While both the NSCT and the SB are selected for racing performance, the Norwegian and Swedish breed organizations have remained highly committed to preserving the historical work-horse appearance of the NSCT breed [7, 8]. As a result, both NSCTs and NSDs can be classified as heavy horse breeds, with NSCTs sometimes referred to as “draft trotters” (Fig. 1) [14]. Any lingering genetic similarities between the NSCT and the SB are therefore highly likely to be associated with favorable traits for TRA. Our aim was to identify these similarities using pooled whole genome sequence data from a carefully selected sample of NSCTs, NSDs, and SBs.

## Results

A summary of the population statistics for the three populations is provided in Tables 1 and 2.

### $F_{ST}$ analyses

After clustering of windows located less than 0.1 Mb apart and filtering out regions containing less than 2 SNPs or



**Fig. 1** Breed classifications. Top panel = North Swedish Draught Horse; middle panel = Norwegian-Swedish Coldblooded trotter; bottom panel = Standardbred trotter. \*Images used in the figure are privately owned and thus were not taken from previously published sources requiring written permission for use

windows measuring only 1000 bp, 580 differentiated regions were retained (Fig. 2; Additional file 1). Candidate regions ranged in length from 1.5 kb to 773.5 kb (mean/median length: 80/56 kb; total cumulative length: 46.389 Mb).

Candidate regions on *Equus caballus* chromosome (ECA) 7 and 11 contained the largest number of SNPs at 214 (length: 151 kb) and 147 (length: 683.5 kb), respectively. Multiple regions on ECA 11, 13, and 15 were greater than

**Table 1** Pooled population samples and genome information

Breed	Number of individuals <sup>a</sup>	Coverage	Nucleotide diversity ( $\pi$ ) (%) <sup>b</sup>
Norwegian-Swedish Coldblooded trotter	18	43.98	0.171
North Swedish Draught	25	35.96	0.175
Standardbred trotter	22	56.26	0.159

<sup>a</sup>Male/Female ratio: Norwegian-Swedish Coldblooded trotter = 2/1; North Swedish Draught = 25/0; Standardbred trotter = 7/4

<sup>b</sup>Nucleotide diversity calculated over 5000 bp windows, averaged across the genome (%)

500 kb in length (Fig. 2; Additional file 1). From GO analysis, candidate regions identified contained 271 candidate genes associated with known molecular functions, 519 with known biological processes, and 124 with known pathways (Figs. 3 and 4; Additional files 2 and 3). One candidate region for TRA overlapped with a previously characterized QTL for racing ability [15]. This region was located on ECA1 and overlapped a QTL for racing speed (Fig. 2; Table 3).

## Discussion

For all intents and purposes, racehorses are professional athletes. Like professional human athletes, racehorses must not only endure the day to day physical demands of their sport, but they must also have a genetic capacity for athleticism relative to their sport in order to ultimately achieve success (i.e. win). However, unlike in humans, racehorses have been carefully selected and bred for centuries, resulting in alleles with subtle effects on athleticism being enriched over time. As a result, racehorses in particular provide a unique opportunity to identify genes and subsequently the molecular mechanisms underpinning athletic ability. Rarely found outside of the Nordic countries, the NSCT is perhaps one of the most unique types of racehorse in the modern era – originating not from historic racing breeds such as the Thoroughbred or Standardbred, but instead tracing its lineage back to the original North Swedish horse [10, 11]. Using whole-genome re-sequencing of pooled DNA from a carefully selected group of NSCTs, NSDs, and SBs, we capitalized on this unusual ancestry of the NSCT and identified 580 candidate regions for TRA in the horse.

Only one previously characterized QTL for racing ability was overlapped by the TRA candidate regions in the current study [15]. A SNP in the first intron of the sortilin related VPS10 domain containing receptor 3 (*SORCS3*) gene, previously associated with racing speed in endurance horses, was overlapped by a 107.5 kb candidate region on

ECA1 [15]. The gene encodes for a type-I transmembrane receptor protein that is a member of family of receptors with known pleiotropic functions [16]. Genetic variation in *SORCS3* has been associated with Alzheimer's disease in humans with more recent studies suggesting that additive epistatic effects of genetic variants within the gene may be important [17–19]. Furthermore, variation in *SORCS3* has been associated with attention deficit hyperactivity disorder and *SORCS3* knockout mice display defects in spatial learning and memory, as well as increased fear extinction [20, 21]. Although the function of this gene in horses remains unknown, as the transcript is generally expressed at high levels in the brain, it could perhaps alter an individual's perception of athletic competition (e.g. altered feedback loop in response to exercise, reduced fear extinction) [20].

Interestingly, while a mere 1.9% (46.39 Mb) of the genome was covered by TRA candidate regions, the regions on ECA11 not only accounted for 11.2% of all TRA candidate regions identified, but ECA11 candidate regions also contained some of the highest number of SNPs ( $n = 100+$ ) and the largest sweeps (> 50 kb in length). Furthermore, ECA11 had the highest concentration of candidate TRA regions when compared to the other equine chromosomes. However, despite the density of this TRA signal, no QTLs for racing ability have been mapped to ECA11 [22].

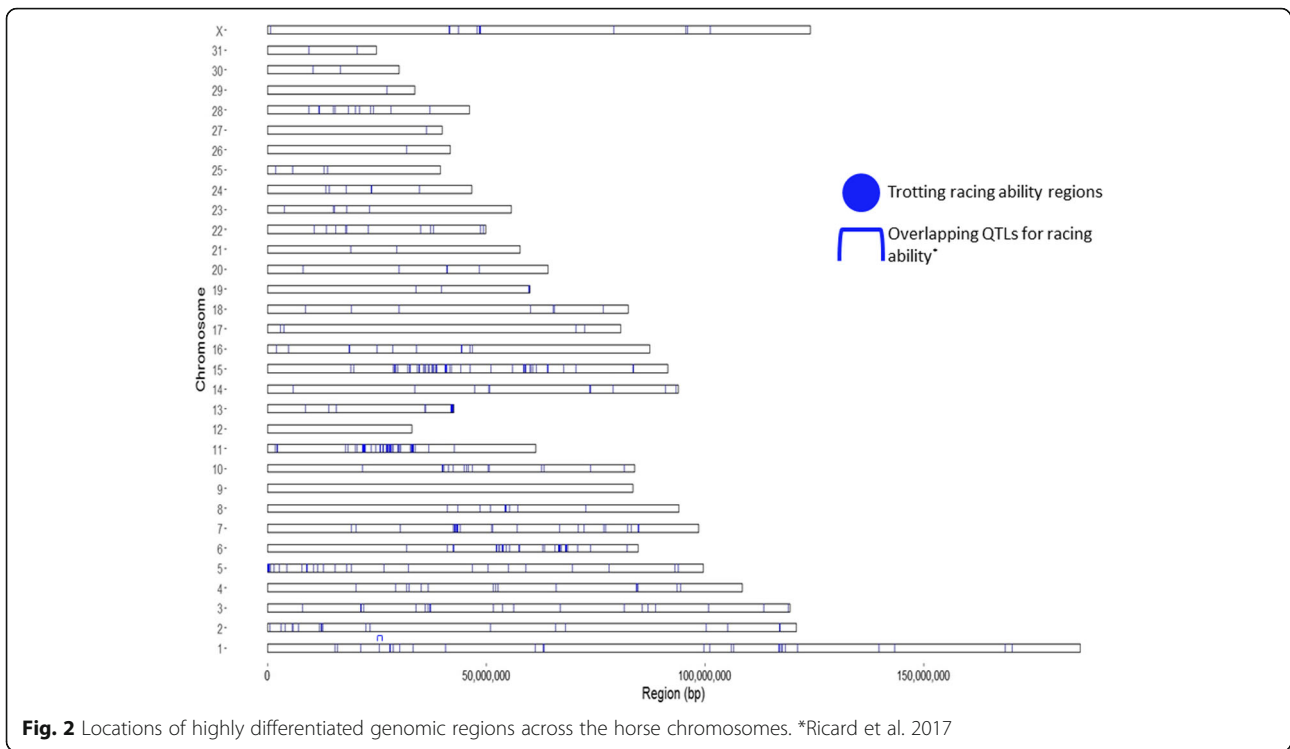
Consequently, regions on ECA11 have previously been associated with size (i.e. height and mass), which, given the prudent design of the current study, is particularly interesting [3, 23, 24]. Although similar in height, NSCTs, NSDs, and SBs differ in their physique. SBs tend to be leaner and more refined in their appearance compared with NSCTs, while NSCTs tend to be leaner and more refined than NSDs (Fig. 1). In order for a region to have been considered as a candidate region for TRA in the current study, the region had to be highly similar between NSCTs and SBs yet decidedly different from NSDs. It is possible that strict adherence to include only top performing NSCTs in the sequenced pool may have skewed the NSCT pool towards lighter framed horses; thereby demonstrating what conceivably is a competitive racing advantage for lighter horses. Perhaps even more interesting to note is that no previously reported QTLs for growth or conformation traits overlapped any of the candidate regions for TRA on ECA11; however, a recent study in American Quarter

**Table 2** Differentiation between breeds

	NSD	NSCT	SB
NSD		0.07159	0.11077
NSCT	0.05755		0.10819
SB	0.08823	0.08555	

Mean  $F_{ST}$  above the diagonal, median  $F_{ST}$  below the diagonal

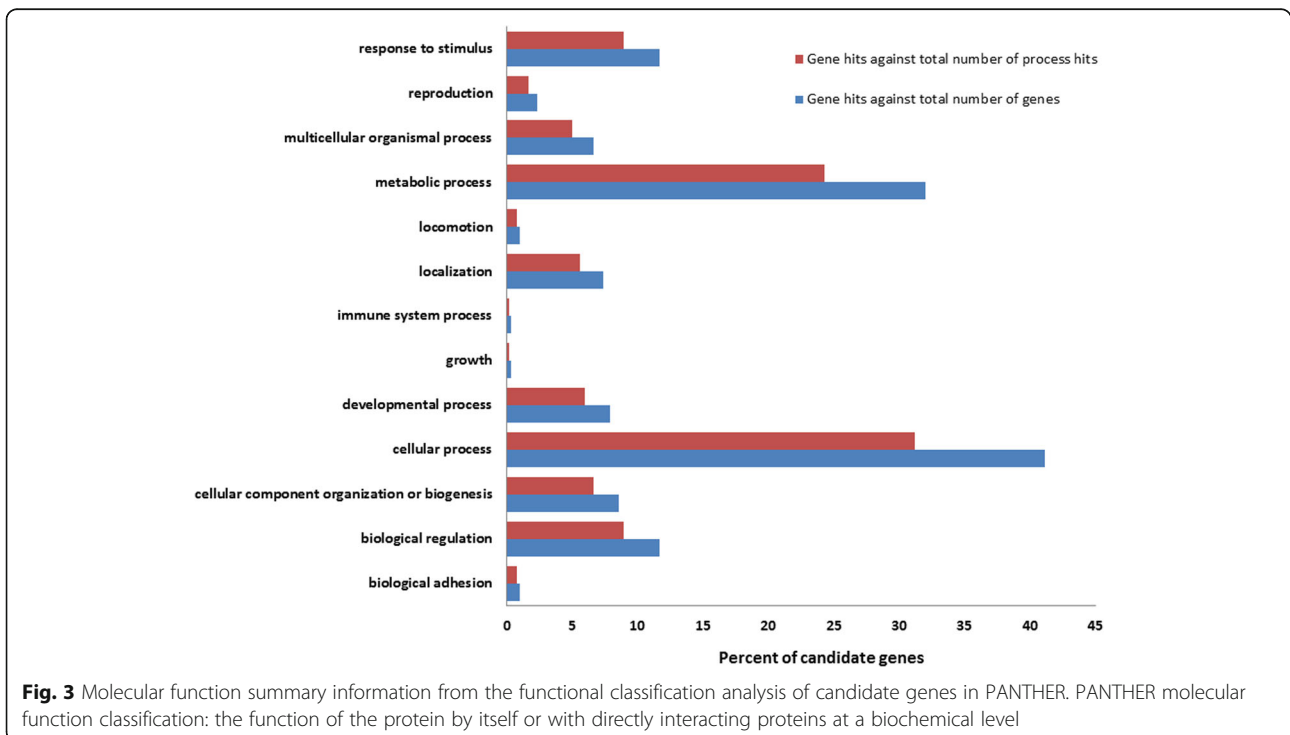
NSD North Swedish Draught Horse, NSCT Norwegian-Swedish Coldblooded Trotter, SB Standardbred trotter

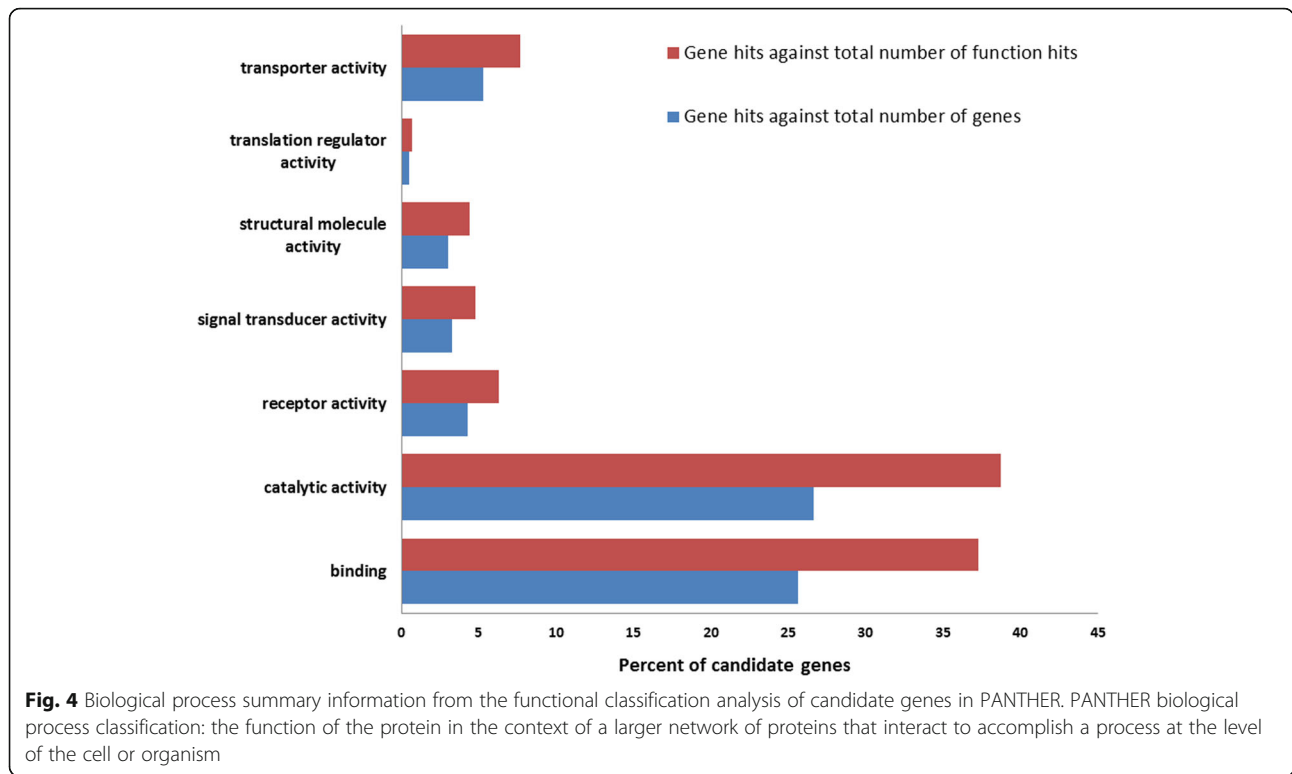


horses also suggested ECA11 as potentially important for racing ability [25].

ECA11 also contained the candidate region with the second largest number of SNPs ( $n = 147$ ) in the study. The region, located ECA11:32,874,000-33,557,500, encompasses 9

genes, (RAD51C, PPM1E, ENSECAG00000003590, ENSECAG00000015244, GDPD1, YPEL2, SKA2, PRR11) one of which is tripartite motif containing 37 (*TRIM37*). Mutations in *TRIM37* are associated with Mulibrey Nanism in humans, an extremely rare autosomal recessive disorder





characterized by profound growth delays and abnormalities of the muscles, liver, brain, and eyes [26, 27]. Instinctively this would further support the candidate region being associated with body size; however, even if this is the case, it does not necessarily mean the region is solely associated with body size and shape. It is highly plausible that haplotypes associated with body size differ by multiple substitutions with pleiotropic functional effects. Mutations that impact underlying mechanisms for muscle, ligament, and tendon development would certainly influence TRA – limiting racing ability in some instances, while enhancing racing ability in others [28–30]. Moreover, a large conserved haplotype containing tripartite motif containing 13 (*TRIM13*), a gene located on ECA17, has previously been suggested as having selective importance in the Thoroughbred [3]. *TRIM37*, while located on a different chromosome, is part of the same gene family as *TRIM13*.

**Conclusions**

This study identified hundreds of candidate genomic regions contributing to TRA in the horse, a result not unexpected for investigations into the genomics of such a

complex trait. The trait is undoubtedly polygenic, resulting from the cumulative effects of many variants across the genome. Candidates for TRA implicated both genes influencing musculature and conformation, as well as genes involved in neurological development, further suggesting that racing ability may not solely be a product of physical characteristics, but also mental characteristics. This study identified a strong racing ability signal on ECA11 that will be particularly interesting for follow-up.

**Methods**

**Animals**

Genomic DNA samples from 18 NSCTs, 25 NSDs, and 22 SBs were prepared from blood samples and pooled in equimolar ratios prior to library construction (Table 1). Each horse was selected based on strict breed specific criteria. All trotting horses racing in Norway and Sweden have breeding values estimated annually. For NSCTs, estimated breeding values (EBVs) are estimated using an animal model that includes the combined effect of country, sex, and birth year. EBVs are subsequently based on racing performance results (i.e. racing status

**Table 3** QTLs overlapped by trotting racing ability candidate regions

QTL ID	ECA	Start Position	End Position	Trait Name	Study
138,687	1	25,715,314	25,715,354	racing speed	Ricard et al. 2017

ECA *Equus caballus* chromosome

and earnings) occurring between 3 and 6 years of age [31]. For inclusion in the current study, NSCT horses were required to have an estimated breeding value of at least 115 and sire/progeny ratios were restricted to reflect the larger population as accurately as possible (Additional file 4) [7, 8].

For EBVs in SBs, the animal model includes genetic base group and a combination of sex and birth year with the evaluation based on racing performance results occurring between 2 and 5 years of age [31]. For inclusion in the current study, Standardbreds were also required to have estimated breeding values of at least 115 and both SBs and NSDs were not allowed to have a common ancestor within three generations (i.e. no shared sires, dams, grandsires, granddams) [7, 9]. An EBV requirement for NSDs was not possible as EBVs are not calculated for this breed.

#### Pool sequencing, genome alignments, variant calling, and population analyses

Genome sequencing library construction and sequencing was carried out by SciLifeLab (Uppsala, Sweden) using two lanes on the Illumina HiSeq2500 (150 bp paired-end). Sequencing libraries were prepared from 100 ng DNA using the TruSeq Nano DNA sample preparation kit targeting an insert size of 350 bp. Reads were aligned to the *Equus caballus* genome (EquCab2.70) using BWA (v0.7.15) [32]. Duplicates were marked with Picard (v1.118; <http://broadinstitute.github.io/picard/>) and GATK was used for realignment around indels [33]. Samtools (v1.8 [34, 35]) was used to generate the mpileup files needed for Popoolation (v1.2.2) and PoPoolation2 (v1.201) [36]. Nucleotide diversity ( $\pi$ ) was calculated across 5000 bp windows for each population pool using Popoolation [36]. PoPoolation2 was used to calculate  $F_{ST}$  over 1000 bp sliding windows with 50% overlap between the selected population samples using the Karlsson et al. method [36, 37]. Minimum count was set at 3, minimum coverage at 10, maximum coverage at 100, and minimum coverage fraction at 1.

#### Differentiated regions

Given the close relationship between NSCTs and NSDs, candidate regions for athletic traits were defined as genomic regions where  $F_{ST}$  values were relatively high between NSCTs and NSDs, but low between NSCTs and SBs. As such, stringent  $F_{ST}$  cutoffs ( $> 95\%$  percentile,  $F_{ST} = 0.179$  NSCT vs. NSD;  $< 5\%$  percentile,  $F_{ST} = 0.013$  NSCT vs. SB) were used when defining candidate regions. Windows with  $F_{ST}$  values that met these criteria were clustered into candidate sweep regions when they were less than 0.1 Mb from one another (custom R scripts) [38]. Clusters containing only a single 1000 bp window or less than 2 SNPs were excluded. Candidate gene screening was subsequently carried

out using the bioinformatics database Ensembl (<http://www.ensembl.org/>). Candidate regions from the  $F_{ST}$  analyses were used to generate a list of annotated genes using the Ensembl Biomart function. The resulting list of candidate genes was then piped into the PANTHER Classification system in order to obtain an overview of the molecular functions and biological processes affected by the candidate genes [39, 40]. Previously reported racing ability QTLs in the horse (downloaded from the horse QTL database; [22]) were also compared to differentiated regions to determine overlaps using bed file comparisons in BEDOPS [41].

#### Additional files

**Additional file 1:** Summary information for trotting racing ability candidate sweep regions. (XLSX 53 kb)

**Additional file 2:** Functional classification gene list from PANTHER analysis of trotting racing ability candidate sweep regions. (XLSX 42 kb)

**Additional file 3:** Pathway summary information from the functional classification analysis of trotting racing ability candidate sweep regions. (XLSX 12 kb)

**Additional file 4:** Pedigree breakdown for the 18 Norwegian-Swedish Coldblooded trotters (NSCT) included in the NSCT pool. (XLSX 11 kb)

#### Abbreviations

EBVs: Estimated breeding values; ECA: *Equus caballus* chromosome; GDPD1: Glycerophosphodiester phosphodiesterase domain containing 1; NSCT: Norwegian-Swedish Coldblooded trotter; NSD: North Swedish Draught horse; PPM1E: Protein phosphatase Mg<sup>2+</sup>/Mn<sup>2+</sup> dependent 1E; PRR11: Proline rich 11; RAD51C: RAD51 paralog C; SB: Standardbred trotter; SKA2: Spindle and kinetochore associated complex subunit 2; SORCS3: Sortilin related VPS10 domain containing receptor 3; TRA: Trotting racing ability; TRIM13: Tripartite motif containing 13; TRIM37: Tripartite motif containing 37; YPEL2: Yippee like 2

#### Acknowledgements

We want to thank Christina Olsson, Annica Edberg, Knut Larsen, Thorvaldur Árnason, the Swedish Trotting Association, and the Norwegian Trotting Association for providing their support and the data/samples for the study. Sequencing was performed by the SNP&SEQ Technology Platform in Uppsala. The facility is part of the National Genomics Infrastructure (NGI) Sweden and Science for Life Laboratory. The SNP&SEQ Platform is also supported by the Swedish Research Council and the Knut and Alice Wallenberg Foundation.

#### Funding

This work was supported by the Swedish Research Council for Environment, Agricultural Science and Spatial Planning (Formas), 2016–00947, (GL). [www.formas.se](http://www.formas.se). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

#### Availability of data and materials

The data that support the findings of this study are available from the Swedish Trotting Association (Stockholm, Sweden) and the Norwegian Trotting Association (Oslo, Norway), but restrictions apply to the availability of these data, which were used under license for the current study, and so are not publicly available. However, data are available from the authors upon reasonable request and with permission of the Swedish Trotting Association (Stockholm, Sweden) and the Norwegian Trotting Association (Oslo, Norway).

#### Authors' contributions

BDV, KJF, and GL conceived and designed the experiments; BDV, KJF, MKR, MW, and ES contributed to sampling. GL and ES contributed the reagents and KJF and MKR extracted the DNA; BDV and ML analyzed the data and drafted the manuscript; KJF, MKR, MW, ES, MS, CFI, ML, and GL discussed and contributed to data analysis; All authors read and approved the final manuscript.

### Ethics approval and consent to participate

All experimental procedures and sample collection methods were approved by the Ethics Committee for Animal Experiments in Uppsala, Sweden [Number: C 121/14]. Samples used in the study were already available at either the Animal Genetics Laboratory at SLU in Uppsala, Sweden or the Department of Basic Sciences and Aquatic Medicine at the Norwegian University of Life Sciences in Oslo, Norway as they previously had been used for parentage testing. Permission to use the samples was granted from the Swedish Trotting Association and the Norwegian Trotting Association (the owners of the samples per the rules/guidelines of the industry).

### Consent for publication

Not applicable.

### Competing interests

The authors have the following interest: GL is a co-inventor on a granted patent concerning commercial testing of the DMRT3 mutation: A method to predict the pattern of locomotion in horses. PCT EP 12747875.8. European patent registration date: 2011-05-05, US patent registration date: 2011-08-03. There are no further patents, products in development, or marketed products to declare.

### Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

### Author details

<sup>1</sup>Department of Animal Breeding and Genetics, Swedish University of Agricultural Sciences, Uppsala, Sweden. <sup>2</sup>School of Life and Environmental Sciences, University of Sydney, Sydney, Australia. <sup>3</sup>Department of Biological and Environmental Sciences, University of Gothenburg, Gothenburg, Sweden. <sup>4</sup>Department of Ecology and Genetics, Uppsala University, Uppsala, Sweden. <sup>5</sup>Department of Clinical Sciences, Swedish University of Agricultural Sciences, Uppsala, Sweden. <sup>6</sup>Department of Companion Animal Clinical Sciences, Norwegian University of Life Sciences, Oslo, Norway. <sup>7</sup>Livestock Genetics, Department of Biosystems, KU Leuven, Leuven, Belgium.

Received: 19 September 2018 Accepted: 28 January 2019

Published online: 04 February 2019

### References

- Meadows JRS, Lindblad-Toh K. Dissecting evolution and disease using comparative vertebrate genomics. *Nat Rev Genet.* 2017;18:624–36.
- Andersson L. Domestic animals as models for biomedical research. *Ups J Med Sci.* 2016;121:1–11.
- Petersen JL, Mickelson JR, Rendahl AK, Valberg SJ, Andersson LS, Axelsson J, et al. Genome-wide analysis reveals selection for important traits in domestic horse breeds. *PLoS Genet.* 2013;9:e1003211.
- Andersson L. How selective sweeps in domestic animals provide new insight into biological mechanisms. *J Intern Med.* 2012;271:1–14.
- Andersson LS, Larhammar M, Memic F, Wootz H, Schwochow D, Rubin CJ, et al. Mutations in *DMRT3* affect locomotion in horses and spinal circuit function in mice. *Nature.* 2012;30:642–6.
- Petersen JL, Mickelson JR, Cothran EG, Andersson LS, Axelsson J, Bailey E, et al. Genetic diversity in the modern horse illustrated from genome-wide SNP data. *PLoS One.* 2013;8:e54997.
- Svensk Travsport/Uppfödning. <https://www.travsport.se/artikel/uppfodning> (2018). Accessed 31 Aug 2018.
- Det Norske Travselvskap: Næring og Avl. <https://www.travsport.no> (2018). Accessed 31 Aug 2018.
- Föreningen Nordsvenska Hästen. <http://www.nordsvensken.org/> (2018). Accessed 31 Aug 2018.
- Jäderkvist Fegraeus K, Velie BD, Axelsson J, Ang R, Hamilton NA, Andersson L, et al. A potential regulatory region near the *EDN3* gene may control both harness racing performance and coat color variation in horses. *Physiol Rep.* 2018;6:e13700.
- Bohin O, Rönningen K. Inbreeding and relationship within the north-Swedish horse. *Acta Agric Scand.* 1975;25:121–5.
- Árnason T. Trends and asymptotic limits for racing speed in Standardbred trotters. *Livest Prod Sci.* 2001;72:135–45.
- Árnason T, Bendroth M, Philipsson J, Henriksson K, Darenius A. Genetic evaluations of Swedish trotters-state of breeding evaluation in trotters. Proceedings of the European Federation of Animal Science symposium of the commission on horse production. Wageningen, the Netherlands: Pudoc; 1989. p. 106–29.
- Pedigree Online All Breed Database. <https://www.allbreedpedigree.com/> (2018). Accessed 31 Aug 2018.
- Ricard A, Robert C, Blouin C, Baste F, Torquet G, Morgenthaler C, Riviére J, Mach N, Mata X, Schibler L, Barrey E. Endurance exercise ability in the horses: a trait with complex polygenic determinism. *Front Genet.* 2017;8:89.
- Hermey G. The Vps10p-domain receptor family. *Cell Mol Life Sci.* 2009;66:2677–89.
- Reitz C, Tosto G, Vardarajan B, Rogava E, Ghani M, Rogers RS, et al. Independent and epistatic effects of variants in *VPS10-d* receptors on Alzheimer disease risk and processing of the amyloid precursor protein (APP). *Transl Psychiatry.* 2013;3:e256.
- Reitz C. The role of the retromer complex in aging-related neurodegeneration: a molecular and genomic review. *Mol Gen Genomics.* 2015;290:413–27.
- Binzer S, Stenager E, Binzer M, Kyvik KO, Hillert J, Imrell K. Genetic analysis of the isolated Faroe Islands reveals *SORCS3* as a potential multiple sclerosis risk gene. *Mult Scler.* 2016;22:733–40.
- Breiderhoff T, Christiansen GB, Pallesen LT, Vaegter C, Nykjaer A, Holm MM, et al. Sortilin-related receptor *SORCS3* is a postsynaptic modulator of synaptic depression and fear extinction. *PLoS One.* 2013;8:e75006.
- Lionel AC, Crosbie J, Barbosa N, Goodale T, Thiruvahindrapuram B, Rickaby J, et al. Rare copy number variation discovery and cross-disorder comparisons identify risk genes for ADHD. *Sci Transl Med.* 2011;3:95ra75.
- Hu Z, Park CA, Reecy JM. Developmental progress and current status of the animal QTLdb. *Nucleic Acids Res.* 2016;44:D827–33.
- Frischknecht M, Jagannathan V, Plattet P, Neuditschko M, Signer-Hasler H, Bachmann I, et al. A non-synonymous *HMGA2* variant decreases height in Shetland ponies and other small horses. *PLoS One.* 2015;10:e0140749.
- Metzger J, Schrimpf R, Philipp U, Distl O. Expression levels of *LCORL* are associated with body size in horses. *PLoS One.* 2013;8:e56497.
- Avila F, Mickelson JR, Schaefer RJ, McCue ME. Genome-wide signatures of selection reveal genes associated with performance in American quarter horse subpopulations. *Front Genet.* 2018;9:249.
- Tang SL, Gao YL, Wen-Zhong H. Knockdown of *TRIM37* suppresses the proliferation, migration and invasion of glioma cells through the inactivation of *PI3K/Akt* signaling pathway. *Biomed Pharmacother.* 2018;99:59–64.
- Jobic F, Morin G, Vincent-Delorme C, Cadet E, Cabry R, Mathieu-Dramard M, et al. New intragenic rearrangements in non-Finnish mulibrey nanism. *Am J Med Genet A.* 2017;173:2782–8.
- Mosher DS, Quigon P, Bustamante CD, Sutter NB, Mellers CS, Parker HG, Ostrander EA. A mutation in the myostatin gene increases muscle mass and enhances racing performance in heterozygote dogs. *PLoS Genet.* 2007;3:e79.
- Hill EW, McGivney BA, Gu J, Whiston R, MacHugh DE. A genome-wide SNP-association study confirms a sequence variant (g.66493737C>T) in the equine myostatin (*MSTN*) gene as the most powerful predictor of optimum racing distance for thoroughbred racehorses. *BMC Genomics.* 2010;11:552.
- McGivney BA, Browne JA, Fonseca RG, Katz LM, MacHugh DE, Whiston R, Hill EW. *MSTN* genotypes in thoroughbred horses influence skeletal muscle gene expression and racetrack performance. *Anim Genet.* 2012;43:810–2.
- Árnason T. Genetic evaluation of Swedish Standardbred trotters for racing performance traits and racing status. *J Anim Breed Genet.* 1999;116:387–98.
- Li H, Durbin R. Fast and accurate short read alignment with burrows-wheeler transform. *Bioinformatics.* 2009;25:1754–60.
- McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernysky A, et al. The genome analysis toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* 2010;20:1297–303.
- Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. Genome project data P: the sequence alignment/map format and SAMtools. *Bioinformatics.* 2009;25:2078–9.
- Li H. A statistical framework for SNP calling, mutation discovery, association mapping and population genetic parameter estimation from sequencing data. *Bioinformatics.* 2011;27:2987–93.
- Kofler R, Orozco-terWengel P, De Maio N, Pandey RV, Nolte V, Futschik A, et al. PoPoolation: a toolbox for population genetic analysis of next generation sequencing data from pooled individuals. *PLoS One.* 2011;6:e15925.

37. Karlsson EK, Baranowska I, Wade CM, Salmon Hillbertz NHC, Zody MC, Anderson N, et al. Efficient mapping of mendelian traits in dogs through genome-wide association. *Nat Genet.* 2007;39:1321–8.
38. R Development Core Team: R-A Language and Environment for Statistical Computing. <https://www.r-project.org/>. (2018). Accessed 13 Jan 2018.
39. Thomas PD, Campbell MJ, Kejariwal A, Mi HY, Karlak B, Daverman R, et al. PANTHER: a library of protein families and subfamilies indexed by function. *Genome Res.* 2003;13:2129–41.
40. Mi H, Poudel S, Muruganujan A, Casagrande JT, Thomas PD. PANTHER version 10: expanded protein families and functions, and analysis tools. *Nucleic Acids Res.* 2016;44:D336–42.
41. Neph S, Kuehn MS, Reynolds AP, Haugen E, Thurman RE, Johnson AK, et al. BEDOPS: high-performance genomic feature operations. *Bioinformatics.* 2012;28:1919–20.

**Ready to submit your research? Choose BMC and benefit from:**

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

**At BMC, research is always in progress.**

Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)

