HHS Public Access

Author manuscript

Anim Genet. Author manuscript; available in PMC 2018 June 01.

Published in final edited form as:

Anim Genet. 2017 June ; 48(3): 287–294. doi:10.1111/age.12531.

Deletion of 2.7 kb near *HOXD3* in an Arabian horse with occipitoatlantoaxial malformation

MH Bordbari[*], **MCT Penedo**[†], **M. Aleman**[‡], **SJ Valberg**[§], **J. Mickelson**[¶], and **CJ Finno**[*]

[*]Department of Population Health and Reproduction, School of Veterinary Medicine, University of California-Davis, Davis, CA 95616, USA

[†]Veterinary Genetics Laboratory, School of Veterinary Medicine, University of California-Davis, Davis, CA 95616, USA

[‡]Department of Veterinary Medicine and Epidemiology, School of Veterinary Medicine, University of California-Davis, Davis, CA 95616, USA

[§]Department of Large Animal Clinical Sciences, College of Veterinary Medicine, Michigan State University, East Lansing, MI 48824, USA

[¶]Department of Veterinary Population Medicine, College of Veterinary Medicine, University of Minnesota, St. Paul, MN 55108, USA

Summary

In the horse, the term occipitoatlantoaxial malformation (OAAM) is used to describe a developmental defect in which the first cervical vertebra (atlas) resembles the base of the skull (occiput) and the second cervical vertebra (axis) resembles the atlas. Affected individuals demonstrate an abnormal posture and varying degrees of ataxia. The homeobox (HOX) gene cluster is involved in the development of both the axial and appendicular skeleton. *Hoxd3*-null mice demonstrate a strikingly similar phenotype to Arabian foals with OAAM. Whole-genome sequencing was performed in an OAAM-affected horse (OAAM1) and seven unaffected Arabian horses. Visual inspection of the raw reads within the region of *HOXD3* identified a 2.7-kb deletion located 4.4 kb downstream of the end of *HOXD4* and 8.2 kb upstream of the start of *HOXD3*. A genotyping assay revealed that both parents of OAAM1 were heterozygous for the deletion. Additional genotyping identified two of 162 heterozygote Arabians, and the deletion was not present in 371 horses of other breeds. Comparative genomics studies have revealed that this region is highly conserved across species and that the entire genomic region between *Hoxd4* and *Hoxd3* is transcribed in mice. Two additional Arabian foals diagnosed with OAAM (OAAM 2 and 3) were genotyped and did not have the 2.7-kb deletion. Closer examination of the phenotype in these cases revealed notable variation. OAAM3 also had facial malformations and a patent ductus

Address for correspondence: C. J. Finno, University of California, Davis SVM, Room 4206 Vet Med 3A, One Shields Ave, Davis, CA 95616, USA. cjfinno@ucdavis.edu.

Conflicts of interest

The authors have no conflicts of interest to disclose.

Supporting information

Additional supporting information may be found online in the supporting information tab for this article:

arteriosus, and the actual malformation at the craniocervical junction differed. Genetic heterogeneity may exist across the *HOXD* locus in Arabian foals with OAAM.

Keywords

atlas; axis; cervical; equine; homeobox; whole-genome sequencing

Introduction

Craniocervical junction abnormalities are rare diseases encompassing a range of phenotypes that have been identified in humans (Menezes *et al.* 1980) and many domestic animals species (Leipold *et al.* 1972; Mayhew *et al.* 1978; Watson *et al.* 1985, 1988; Schmidt *et al.* 1993; Helie *et al.* 1999; Seva *et al.* 2008). In the horse, craniocervical junction abnormalities are termed occipitoatlantoaxial malformations (OAAM) and are characterized by an abnormal occiput, atlas (first cervical vertebra) and axis (second cervical vertebra). Equine OAAM has been classified into six groups (Chowdhary 2013). The first, occipitalization of the atlas with atlantalization of the axis, appears to be an inherited form of OAAM in Arabian horses (Mayhew *et al.* 1978). Subluxation of the atlantooccipital joint and fusion of the atlas and axis with lateral deviation of the atlantoaxial joint and 20-degree rotation of the atlas was reported in a one half-Arabian colt (Blikslager *et al.* 1991). Other groups include: congenital asymmetrical OAAM (Rosenstein *et al.* 2000), asymmetric atlanto-occipital fusion (Mayhew *et al.* 1978), duplication of axis or atlas (De Lahunta *et al.* 1989) and symmetrical OAAM in non-Arabian horses (Wilson *et al.* 1985; Bell *et al.* 2007).

In familial OAAM of Arabian horses, the atlas resembles the occiput and the axis resembles the atlas (Mayhew *et al.* 1978). The dens is often hypoplastic, with a partial or complete absence of articulation into the atlas. Clinical signs include abnormal head position, progressive ataxia, extended neck posture and visible irregularities along the cervical spine. In OAAM-affected Arabian horses, there is a characteristic ‘clicking’ sound, most likely due to the hypoplastic dens luxating during neck extension and subsequently properly repositioning during neck flexion. Equine OAAM affects the mobility of the neck, with loss of extension and flexibility. Affected foals may be found dead at birth or alive with tetraparesis, or progressive ataxia may develop due to vertebral canal stenosis (Mayhew *et al.* 1978).

Homeobox genes direct the formation of many body structures during early embryonic development. In humans and mice, there are currently 38 identified homeobox genes assigned to clusters (A through D) across four different chromosomes. With targeted knockout mice, it has been determined that *HOXD3* encodes for the development of the occiput, axis and atlas and *HOXD4* for the development of cervical vertebrae 1–3 and components of the sternum. *Hoxd3*$^{-/-}$ mice display a similar phenotype to familial Arabian horse OAAM, with fusion of the atlas and occiput and anterior transformations of the atlas and axis (Condie & Capecchi 1993). Although these mice display slight variations in their skeletal phenotype, basioccipital-anterior arch fusion is always apparent along with a complete absence of the dens. Based on this mouse model and the previously described

autosomal recessive mode of inheritance of OAAM in Arabian horses (Mayhew *et al.* 1978), we hypothesized that a functional variant in *HOXD3* resulted in the development of OAAM in Arabian horses.

Materials and methods

Animals

All animal procedures were approved by the University of California–Davis and University of Minnesota Institutional Animal Care and Use Committees, and owners' consent was obtained for all horses. The index case (OAAM1) was a 1-week old Arabian filly, examined at the University of Minnesota, College of Veterinary Medicine for extension of the neck and ataxia. The filly demonstrated an abnormal extended neck posture when nursing and, following nursing, she would attempt to flex at the poll at which time a 'clicking' noise was audible. A symmetric grade 3/5 general proprioceptive ataxia of all four limbs was apparent on neurologic evaluation, according to a published grading scale (Lunn & Mayhew 1989). Gross asymmetry of the atlas was apparent (Fig. 1), and equine OAAM was diagnosed on cervical radiographs, documenting malformation of the occipital condyles and C1. Although the vertebral body of C2 appeared relatively unremarkable, the dens was hypoplastic. The owners elected to maintain the filly on pasture and did not use her for any performance-related or breeding activity. At 9 years of age, a repeat examination revealed abnormal head carriage and an unchanged symmetric grade 3/5 proprioceptive ataxia of the pelvic and thoracic limbs. When circled, the mare demonstrated difficulty in turning her neck. Both the sire and the dam of this mare were available for examination and demonstrated a normal head carriage and no neurologic abnormalities. Whole blood samples were collected from the trio for DNA preparation.

For next-generation sequencing, whole blood samples were collected from an additional seven healthy Arabian horses used for performance (five geldings and two mares; age range 9–29 years). To resolve the missing reference sequence in the EquCab2.0 assembly (http://www.ncbi.nlm.nih.gov/genome/145) extending from chr18:54 619 627– 54 620 959, whole blood was collected from three unaffected Quarter Horse mares at the UC Davis Center for Equine Health research facility for Sanger sequencing. Additionally, samples from two post-mortem confirmed additional OAAM-affected Arabian foals (OAAM2, a 20-day old colt, and OAAM3, a 1-month old colt) were available from one of the investigators (MA).

Genomic DNA preparation

DNA isolations were prepared from whole blood samples. DNA isolations were performed according to the protocol provided through the ArchivePure™ DNA Blood Kit (5′). The sample DNA concentrations were then determined using QIAxpert (QIAGEN) and diluted to 20 ng/μl in water for use in end-point PCR.

Sanger sequencing

Primers were designed using PRIMER3PLUS software (Untergasser *et al.* 2012) (Table S1) to amplify annotated equine exons 1 and 2 of *HOXD3*, along with the 100-bp flanking intronic sequence, in OAAM1 and its sire and dam. Additionally, primers were created to

amplify the orthologous region of the annotated non-coding exon 1 in human *HOXD3* (http://www.ncbi.nlm.nih.gov/genome/51?genome_assembly_id=273342). PCR reactions comprised 2 U of Hot-start TAQ and 2.0 μl of 10× Buffer (Applied Biosystems), 0.25 mM of dNTPs (Thermo Fisher), 0.5 μM of both forward and reverse primers (Invitrogen) and 20 ng of genomic DNA. Standard PCR conditions were performed as follows: 95 °C for 15 min, 35 cycles of 95 °C denaturation for 30 s, primer-specific annealing temperature for 1 min, 72 °C extension for 45 s and a final extension at 72 °C for 20 min. PCR products were purified using ExoSAP-IT® PCR Product Cleanup Kit (Affymetrix). Sanger sequencing was performed using Applied Biosystems 2500 automated sequencers. The resulting sequences were aligned to Equ-Cab2.0 (http://www.ncbi.nlm.nih.gov/genome/145) and analyzed with SEQUENCER® software (Gene Codes Corp.). As segregating variants following the proposed autosomal recessive mode of inheritance within this trio were not identified, DNA from OAAM1 underwent whole-genome sequencing.

Whole-genome next-generation sequencing

Using Illumina's TruSeq DNA PCR-free Library Preparation Kit and following the manufacturer's instructions, libraries were prepared with a median insert size of 355 bp from the OAAM-affected case and seven unaffected Arabian horses. The eight libraries were barcoded and pooled across eight lanes of a 100PE flow cell on an Illumina HiSeq2000, generating an average of 13× coverage per horse. Following quality trimming, reads were mapped to the EquCab2.0 reference genome using the BURROWS-WHEELER ALIGNER (BWA) version 0.7.5a (Li & Durbin 2009) using default settings. After sorting the mapped reads by the coordinates of the sequence, PCR duplicates were labeled with Picard tools (http://sourceforge.net/projects/picard/).

The GENOME ANALYSIS TOOL KIT (GATK version v.2.7.4) was used to perform local realignment (McKenna *et al.* 2010). Variant calls were made across all eight samples simultaneously using standard hard filtering parameters or variant quality score recalibration with HAPLOTYPE CALLER according to GATK Best Practices Recommendations (DePristo *et al.* 2011; Van Der Auwera *et al.* 2013). SNPEFF (Cingolani *et al.* 2012b) and SNPSIFT (Cingolani *et al.* 2012a) were used to predict the functional effects of detected variants across the genome and within the HOXD cluster region of interest and filtered by segregation using Fisher's Exact Test. A second algorithm, DELLY (Rausch *et al.* 2012), was performed on the OAAM-affected horse to attempt to identify larger duplications, deletions and insertions. Functional homozygous variants segregating in OAAM1 with 'moderate' or 'high' effects, as defined by SNPEFF (Cingolani *et al.* 2012b), were further evaluated using all publically available mapped whole-genome sequences in the NCBI Sequence Read Archive (https://www.ncbi.nlm.nih.gov/sra). Any variant that remained unique to OAAM1 was further evaluated using Sanger sequencing, as described above (Table S1: primer sequences). Visual inspection of the raw reads using the INTEGRATED GENOMICS VIEWER (Robinson *et al.* 2011) within the region of HOXD was additionally performed. The whole-genome sequence from OAAM1 was deposited in the NCBI SRA (SRP077507).

Deletion confirmation

To confirm the putative deletion identified on whole-genome sequencing of OAAM1 between *HOXD4* and *HOXD3*, the approximately 1.3-kb gap in EquCab2.0 was first determined. Primers were designed to walk across sequence flanking the deletion through Sanger sequencing using genomic DNA from three healthy Quarter Horses. Primers were designed using PRIMER3PLUS software (Untergasser *et al.* 2012) (Table 2) and DNA oligonucleotides synthesized by Invitrogen (Life Technologies). Amplification of products was performed using end-point PCR and visualized with the QIAxcel Advanced System (QIAGEN) and the QIAxcel DNA Screening Kit (QIAGEN). The 20-μl PCR reactions comprised 2 U of Hot-start TAQ and 2.0 μl of 10× buffer (Applied Biosystems), 0.25 mM of dNTPs (Thermo Fisher), 0.5 μM of both forward and reverse primers (Invitrogen Life Technologies) and 20 ng of genomic DNA. Standard PCR conditions were performed as follows: 95 °C for 10 min, 35 cycles of 95 °C denaturation for 30 s, 60 °C annealing for 1 min, 72 °C extension for 1 min and a final extension at 72 °C for 10 min. PCR products were purified using the ExoSAP-IT® PCR Product Cleanup Kit (Affymetrix). Sanger sequencing was performed using ABI 2500 automated sequencers (Applied Biosystems). Resulting sequences were aligned to EquCab2.0 (http://www.ncbi.nlm.nih.gov/genome/145) and analyzed with SEQUENCER® software (Gene Codes Corp.).

Once the entire contig was established in healthy horses, OAAM1 and its sire and dam were assayed using the same primer sets.

Genotyping assay

To investigate the incidence of the deletion in Arabians and other breeds, genotyping of additional horses from the UC Davis Veterinary Genetics Laboratory database was done for 162 Arabians, 125 Colombian Creole horses, 37 Connemara Ponies, 112 Quarter Horses and 97 Thoroughbreds. The OAAM status of the Arabians was unknown, and all other horses were presumed to be unaffected. DNA samples were prepared from four to five tail or mane hair roots from each horse using a Proteinase-K digestion procedure (Locke *et al.* 2002). PCR amplification was performed with primers OAAM-F3: GCAGGGATCCCAAGAATACCA, OAAM-F4: GCCTGGGGAGGAGAAGAGC and OAAM-R2b: FAM-CCCGATCTTCCACTAGACTG, designed to detect the presence (wild-type allele, 178-bp product) and absence (variant allele, 217-bp product) of the deleted sequence. PCR reactions in 17 μl of total volume contained 2 μl of DNA template (ca. 10 ng total), 0.5 μM each of forward primers (F3 and F4), 1 μM of the reverse primer (R2b), 1× PCR buffer (750 nM of Tris-HCl pH 8.8, 200 nM of $(NH_4)_2SO_4$, 0.1% Tween 20) (Denville Scientific), 2.5 mM of $MgCl_2$, 200 μM of dNTPs, 0.33% DMSO and 1 U Choice Taq (Denville Scientific). Fluorescence-labeled PCR products were resolved by capillary electrophoresis on an ABI3730 DNA Analyzer (Applied Biosystems) and genotyped with STRAND software (available at https://www.vgl.ucdavis.edu/informatics/strand.php). DNA from a known carrier was included in all PCR runs as a positive control for wild-type and variant alleles.

Results

Variant identification

Sanger sequencing of annotated equine exons 1 and 2 of *HOXD3*, along with 100-bp flanking intronic sequence, in OAAM1 and its sire and dam identified 20 variants, including six previously annotated in dbSNP and 14 novel variants (Table S2; novel variants uploaded to dbSNP). None of these variants segregated in an autosomal recessive manner with the OAAM phenotype in this trio.

Whole-genome sequencing was performed on OAAM1 at ~13× coverage. Using HAPLOTYPE CALLER according to GATK Best Practices Recommendations (DePristo *et al.* 2011; Van Der Auwera *et al.* 2013), 516 598 360 variants [single nucleotide variants (SNVs) and insertions/deletions) were identified in OAAM1, of which 2587 were homozygous in OAAM1 as compared to the unaffected Arabian cohort (Fisher's Exact Test, $P < 0.009$). None of the 2587 variants were within the annotated HOXD region (chr18:54 000 000–55 000 000). Of these variants, one was classified as 'high effect' and five as 'moderate effect' by SNPEFF (Cingolani *et al.* 2012b) (Table 1).

Visual inspection of the raw reads using the INTEGRATED GENOMICS VIEWER (Robinson *et al.* 2011) within the region of HOXD identified a 2.7-kb deletion 5′ to a gap in the reference sequence (Fig. 2). Although multiple gaps, representing regions where the assembly (EquCab2.0) is most likely incorrect, are noted in all horses, there was one gap spanning 2.7 kb that was present only in the OAAM-affected Arabian. This deletion was located 4.4 kb downstream of the end of *HOXD4* and 8.2 kb upstream of the start of *HOXD3*. Both variant-calling algorithms—HAPLOTYPE CALLER (SNV and small indel detection) and DELLY (duplications, deletions, insertions)—failed to detect this deletion.

The six unique SNVs (Table 1) and 2.7-kb deletion were further evaluated in all publically available mapped whole-genome sequences ($n = 121$) from the Sequence Read Archive (SRA; https://www.ncbi.nlm.nih.gov/sra) (Table 1). Heterozygote genotypes were observed for the five moderate SNVs across various breeds. The alternate allele for the SNV at chr9:82 287 859 and the 2.7-kb deletion were not observed in any other dataset. The SNV at chr9:82 287 859 is within an annotated gene (ENSECAT00000022023) encoding for a zinc finger (ZNF707). SNPEFF defined this SNV as resulting in a nonsense mutation. This chr9 SNV and the 2.7-kb deletion in the HOXD region were prioritized for further evaluation.

Genotyping of chr9:82 287 859

Primer pairs were designed to amplify and subsequently genotype the SNV at chr9:82 287 859 in OAAM1, OAAM2, OAAM3 and the sire and dam of OAAM1. Using Sanger sequencing, OAAM1 and its dam were determined to be heterozygous for the SNV, and all other horses were homozygous for the major allele. The SNV at chr9:82 287 859 was therefore excluded as a putative functional variant for OAAM in OAAM1. The 2.7-kb deletion remained prioritized for further assessment.

Defining the gap sequence of the 2.7-kb deletion

In unaffected horses, sequencing of amplicons generated a contiguous sequence in healthy unaffected horses with all base pairs defined along chr18:54 616 116–54 621 918. Although the IGR 9 primer set (Table 2) spanned the gap sequence, the fragment was too large (1658 bp) for standard Sanger sequencing. Therefore, primer sets IGR 10 and IGR 11 (Table 2) were designed based on the contiguous sequence developed using the sequenced reads from IGR 9 amplicons. This allowed for complete sequencing across the gap in EquCab2.0.

Deletion confirmation

After all the primer sets were amplified using end-point PCR, each set was assayed in OAAM1, along with its dam and sire, as presumed obligate carriers. Primer sets IGR 1 and IGR 9–13 amplified in OAAM1, whereas IGR 2–8 did not (Table 2). All primer sets amplified in the parents of OAAM1. Forward and reverse primers at the boundaries of the deletion (IGR1_F and IGR8_R) were then used for additional genotyping. The resulting sequence aligned to two different regions of the consensus sequence, defining the beginning and end of the deletion. The deletion in OAAM1 was localized to chr18:54 616 748–54 619 464.

Transcripts within deleted region

With the recent creation of an open-source, refined equine transcriptome (https://github.com/drtamermansour/horse_trans), the deleted region on OAAM1 was investigated. A custom transcript was identified (TCONS_00073653; chr18:54 619 060–54 632 596), with the first exon of this transcript located within the OAAM1 deletion. TCONS_00073653 was annotated as a variant of *HOXD3*, with exons 3 and 4 of TCONS_00073653 aligning with exons 1 and 2 of the Ensembl annotated *HOXD3* (ENSECAG00000022129) (Fig. 2). When this region was investigated in depth using the publically available reads from this open-source transcriptome, only one to two reads were evident over this putative exon in brainstem, cerebellum and spinal cord, suggestive of low expression levels in the adult horse.

Conservation across species

To investigate the potential functional significance of this deletion, conservation was evaluated across 10 species using the open-access software EVOLUTIONARY CONSERVED REGIONS (ECR; https://ecrbrowser.dcode.org/). The orthologous region spanning the HOXD cluster (*HOXD4*, *HOXD3* and *HOXD1*; chr18:54 605 908–54 650 486) was identified in human (chr2:177 011 292–177 058 499; 90.1% of bases and 100% of span) using the UCSC genome browser. This region was evaluated in ECR and was highly conserved across seven mammalian species (Fig. 3). The region of the deletion (human orthologous region chr2:177 022 138– 177 024 877; 95.2% of bases, 100% of span) included at least one highly conserved evolutionary peak across all 10 species.

Equine OAAM genotyping assays

The resulting PCR products were either a 756-bp band, representing the allele with the deletion, or a 3511-bp band, representing the wild-type allele (Fig. S1). The dam and sire of

OAAM1 genotyped for both the variant and wild-type bands. The two additional OAAM-affected Arabian foals (OAAM2 and 3) genotyped as homozygous wild type. A second genotyping assay was subsequently designed to screen a large number of horses for the identified deletion. Genotyping with the genotyping assay developed at the Veterinary Genetics Laboratory included 162 Arabians, 125 Colombian horses, 37 Connemara Ponies, 112 Quarter Horses and 97 Thoroughbreds. Two Arabians were identified as heterozygous for the deletion, and all other horses were homozygous wild type.

Discussion

In this study, a 2.7-kb deletion of highly conserved sequence within the *HOXD3/4* region was identified in an Arabian mare clinically affected with OAAM (OAAM1). Both parents were heterozygous for the deletion, as were two additional Arabian horses on a larger screening panel. The deletion has not been identified in any other breeds to date. The deletion appeared to be private to the OAAM1 family, with OAAM2 and OAAM3 identified as homozygous wild-type for this mutation. Although functional significance for this deletion was not attainable, evidence of a similar phenotype in *Hoxd3*-null mice supports a potential association in this one horse. A whole-genome approach to variant detection identified six other potential variants (Table 1), one of which appeared unique to OAAM1 but was in fact heterozygous at that particular locus. Whole-genome sequencing over the chr9 SNV was only at 4× coverage, and this SNV was most likely incorrectly genotyped using whole-genome sequencing. Therefore, the 2.7-kb deletion within the *HOXD3/4* region remains the strongest candidate for a putative functional variant in OAAM1.

The intergenic region between *HOXD4* and *HOXD3* containing the identified deletion is highly conserved across species. Using expression-tiling arrays in mouse embryos at 12.5 days, it was demonstrated that transcription is active throughout this intergenic region in the brachial, thoracic and lumbosacral spinal cord (Tschopp *et al.* 2012). Additionally, targeted deletions that included this region in E12.5 mice led to the ectopic anterior activation of posterior Hoxd genes (*Hoxd10*, *Hoxd12* and *Hoxd13*), with downregulation of *Hoxd3*. Although these studies were not able to evaluate the resulting phenotype, we hypothesize that a similar craniocervical malformation that is observed in $Hoxd3^{-/-}$ mice would be expected. Therefore, there is strong evidence that the intergenic region deleted in OAAM1 is required for normal transcription of the HOXD locus, providing support that this deletion resulted in the phenotype observed in this horse.

The 2.7-kb deletion was not present in two additional postmortem-confirmed OAAM-affected Arabians (OAAM2 and OAAM3). Based on comparative information in humans, genetic heterogeneity may exist for this condition within the HOXD region. Craniocervical junction abnormalities in humans are typically classified into five categories: (i) fusion of the atlas and occipital bone, (ii) basilar invagination of the occipital condyles, (iii) atlantoaxial subluxation of dislocation, (iv) Klippel-Feil malformation (i.e. fusion of cervical vertebrae) and (v) platybasia or flattening of the skull base (Saltzman *et al.* 1991). Within one family, nine of 12 family members were reported to be affected with a craniocervical junction abnormality, with all members having abnormalities of C1 and some with involvement of C2 or C3 (Saltzman *et al.* 1991). When reevaluating the postmortem reports

of OAAM2 and OAAM3, it was observed that OAAM2 also had mandibular malformation and a patent ductus arteriosis. This phenotypic variability may be an indicator that genetic heterogeneity exists for these craniocervical junction abnormalities in the Arabian horse. Careful phenotyping of these malformations is required, even in the Arabian breed where OAAM has been presumed to be a monogenic Mendelian disease (Mayhew *et al.* 1978). The potential for genetic heterogeneity is strengthened by the reports of OAAM across horse breeds, including half-Arabians (Wilson *et al.* 1985; Bell *et al.* 2007).

This study highlighted the limitations of detecting large genetic variants using standard 100-bp paired-end whole-genome sequencing and available algorithms. Mate-pair libraries, in conjunction with paired-end libraries, are required for the DELLY program to detect larger structural variants (Rausch *et al.* 2012). Mate-pair libraries, which could have resulted in the ability to detect the deletion with this algorithm, were not used in this study. This finding highlights the importance of direct visualization of aligned sequences (i.e. bam files) using a viewer such as the INTEGRATED GENOMICS VIEWER (Robinson *et al.* 2011) to identify structural variants that were not called with the available variant detection algorithms.

In conclusion, with the insight provided by the *Hoxd3*$^{-/-}$ mouse model and the discovered deletion in this Arabian horse (OAAM1), *HOXD3* and *HOXD4* would be strong candidates for further investigation in human craniocervical junction abnormalities and future OAAM cases across equine breeds.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Research was supported by Morris Animal Foundation (D12EQ-401), NCATS (K01OD015134-01A1) and LRP (L40 TR001136). We would like to acknowledge the owner of OAAM1 for allowing sampling of this affected mare, sire and dam.

References

Bell S, Detweiler D, Benak J, Pusterla N. What is your diagnosis? Occipitoatlantoaxial malformation. Journal of the American Veterinary Medical Association. 2007; 231:1033–4. [PubMed: 17916027]

Blikslager AT, Wilson DA, Constantinescu GM, Miller MA, Corwin LA Jr. Atlantoaxial malformation in a half-Arabian colt. The Cornell Veterinarian. 1991; 81:67–75. [PubMed: 1993394]

Chowdhary, BP., editor. Equine Genomics. John Wiley & Sons; Ames, IA: 2013. Genetics of equine neurologic disease; p. 229-31.

Cingolani P, Patel VM, Coon M, Nguyen T, Land SJ, Ruden DM, Lu X. Using Drosophila melanogaster as a model for genotoxic chemical mutational studies with a new program, SNPSIFT. Frontiers in Genetics. 2012a; 3:35. [PubMed: 22435069]

Cingolani P, Platts A, Wang Le L, Coon M, Nguyen T, Wang L, Land SJ, Lu X, Ruden DM. A program for annotating and predicting the effects of single nucleotide polymorphisms, SNPEFF: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3. Fly (Austin). 2012b; 6:80–92. [PubMed: 22728672]

Condie BG, Capecchi MR. Mice homozygous for a targeted disruption of Hoxd-3 (Hox-4.1) exhibit anterior transformations of the first and second cervical vertebrae, the atlas and the axis. Development. 1993; 119:579–95. [PubMed: 7910549]

De Lahunta A, Hatfield C, Dietz A. Occipitoatlantoaxial malformation with duplication of the atlas and axis in a half Arabian foal. The Cornell Veterinarian. 1989; 79:185–93. [PubMed: 2924582]

DePristo MA, Banks E, Poplin R, et al. A framework for variation discovery and genotyping using next-generation DNA sequencing data. Nature Genetics. 2011; 43:491–8. [PubMed: 21478889]

Helie P, Odin M, Girard C. Atlanto-occipital malformation, osteoarthropathy, and myelopathy in a Duroc boar. Journal of Veterinary Diagnostic Investigation. 1999; 11:182–4. [PubMed: 10098694]

Leipold HW, Strafuss AC, Blauch B, Olson JR, Guffy M. Congenital defect of the atlantooccipital joint in a Holstein-Friesian calf. The Cornell Veterinarian. 1972; 62:646–53. [PubMed: 5077540]

Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics. 2009; 25:1754–60. [PubMed: 19451168]

Locke MM, Penedo MC, Bricker SJ, Millon LV, Murray JD. Linkage of the grey coat colour locus to microsatellites on horse chromosome 25. Animal Genetics. 2002; 33:329–37. [PubMed: 12354140]

Lunn DP, Mayhew IG. The neurologic evaluation of horses. Equine Veterinary Education. 1989; 1:94–101.

Mayhew IG, Watson AG, Heissan JA. Congenital occipitoatlantoaxial malformations in the horse. Equine Veterinary Journal. 1978; 10:103–13. [PubMed: 565704]

McKenna A, Hanna M, Banks E, et al. The GENOME ANALYSIS TOOLKIT: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Research. 2010; 20:1297–303. [PubMed: 20644199]

Menezes AH, VanGilder JC, Graf CJ, McDonnell DE. Craniocervical abnormalities: a comprehensive surgical approach. Journal of Neurosurgery. 1980; 53:444–55. [PubMed: 7420164]

Rausch T, Zichner T, Schlatt A, Stutz AM, Benes V, Korbel JO. DELLY: structural variant discovery by integrated paired-end and split-read analysis. Bioinformatics. 2012; 28:i333–9. [PubMed: 22962449]

Robinson JT, Thorvaldsdottir H, Winckler W, Guttman M, Lander ES, Getz G, Mesirov JP. Integrative genomics viewer. Nature Biotechnology. 2011; 29:24–6.

Rosenstein DS, Schott HC 2nd, Stickle RL. Imaging diagnosis—occipitoatlantoaxial malformation in a miniature horse foal. Veterinary Radiology & Ultrasound. 2000; 41:218–9. [PubMed: 10850870]

Saltzman CL, Hensinger RN, Blane CE, Phillips WA. Familial cervical dysplasia. Journal of Bone and Joint Surgery American. 1991; 73:163–71.

Schmidt SP, Forsythe WB, Cowgill HM, Myers RK. Acase of congenital occipitoatlantoaxial malformation(OAAM)in a lamb. Journal of Veterinary Diagnostic Investigation. 1993; 5:458–62. [PubMed: 8373866]

Seva JI, Gomez S, Pallares FJ, Sanchez P, Bernabe A. Occipitoatlantoaxial malformation in an adult goat. Journal of Veterinary Diagnostic Investigation. 2008; 20:654–6. [PubMed: 18776104]

Tschopp P, Christen AJ, Duboule D. Bimodal control of Hoxd gene transcription in the spinal cord defines two regulatory subclusters. Development. 2012; 139:929–39. [PubMed: 22278926]

Untergasser A, Cutcutache I, Koressaar T, Ye J, Faircloth BC, Remm M, Rozen SG. RIMER3—new capabilities and interfaces. Nucleic Acids Research. 2012; 40:e115. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/22730293. [PubMed: 22730293]

Van Der Auwera GA, Carneiro MO, Hartl C, et al. From FastQ data to high confidence variant calls: the genome analysis toolkit best practices pipeline. Current Protocols in Bioinformatics. 2013; 43:11.10.1–33. [PubMed: 25431634]

Watson AG, Wilson JH, Cooley AJ, Donovan GA, Spencer CP. Occipito-atlanto-axial malformation with atlantoaxial subluxation in an ataxic calf. Journal of the American Veterinary Medical Association. 1985; 187:740–2. [PubMed: 4055492]

Watson AG, De Lahunta A, Evans HE. Morphology and embryological interpretation of a congenital occipito-atlanto-axial malformation in a dog. Teratology. 1988; 38:451–9. [PubMed: 3238603]

Wilson WD, Hughes SJ, Ghoshal NG, McNeel SV. Occipitoatlantoaxial malformation in two non-Arabian horses. Journal of the American Veterinary Medical Association. 1985; 187:36–40. [PubMed: 4019299]

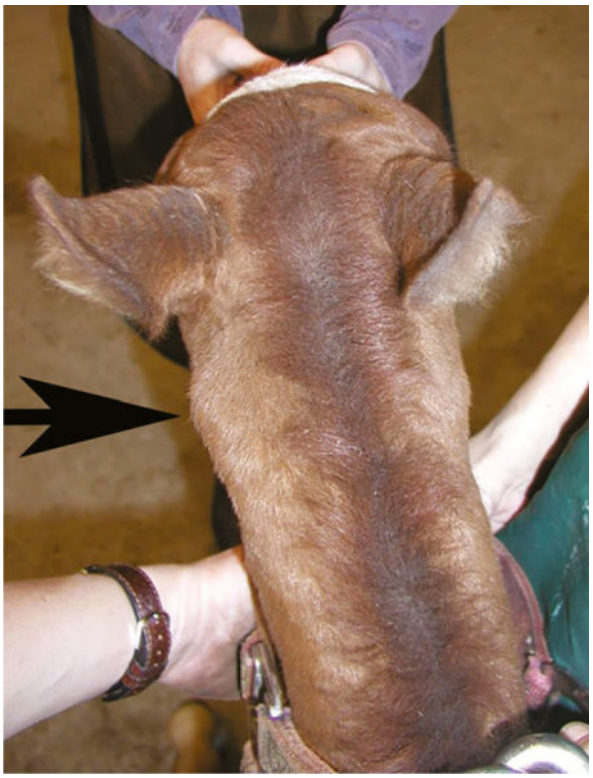

Figure 1.
One-week old Arabian filly with occipitoatlantoaxial malformation. The arrow indicates the asymmetric atlas.

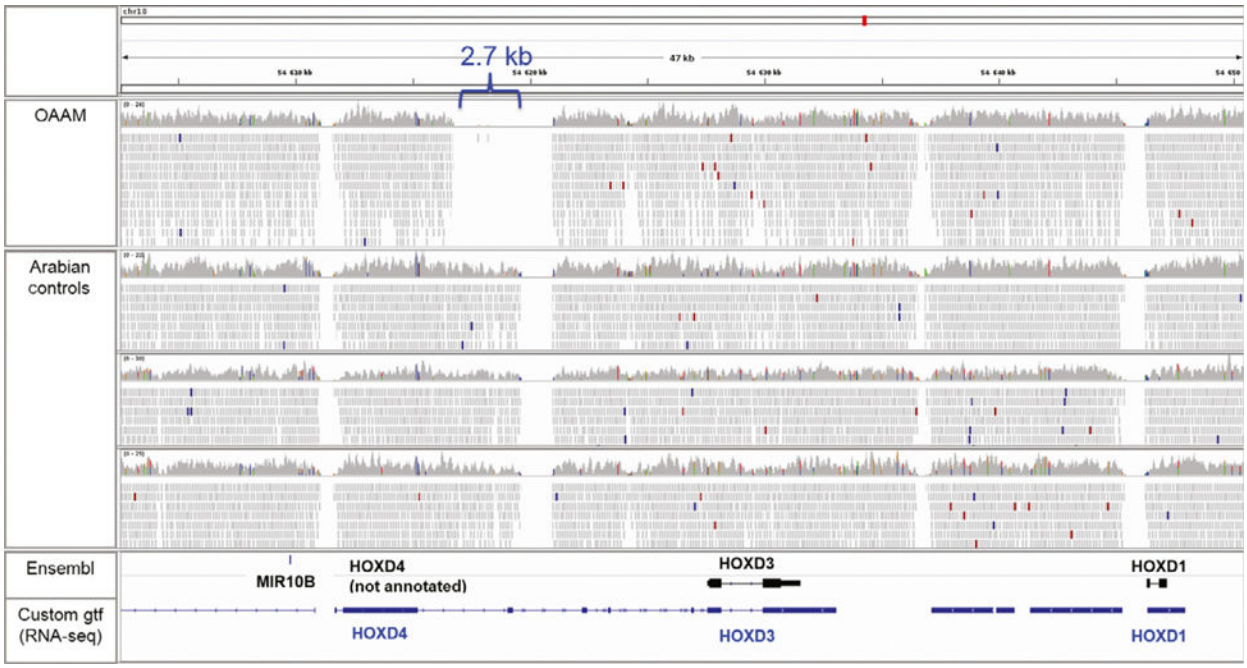

Figure 2.

Whole genome sequencing as visualized in the Integrated Genome Viewer (Robinson et al. 2011). Four horses are depicted (OAAM-affected Arabian and three unaffected Arabians) across the HOXD gene cluster on chr18. Annotations are provided in the bottom pane (Ensembl, with only HOXD3 and HOXD1 annotated in black and a custom annotation file using RNA-seq data (https://github.com/drtamermansour/horse_trans) with HOXD4, HOXD3 and HOXD1 annotated in blue). Although multiple gaps are noted in all horses, which represent regions where the assembly (EquCab2.0) is most likely incorrect, there is one gap spanning 2.7 kb that is present only in the OAAM-affected Arabian. MIR10B = microRNA

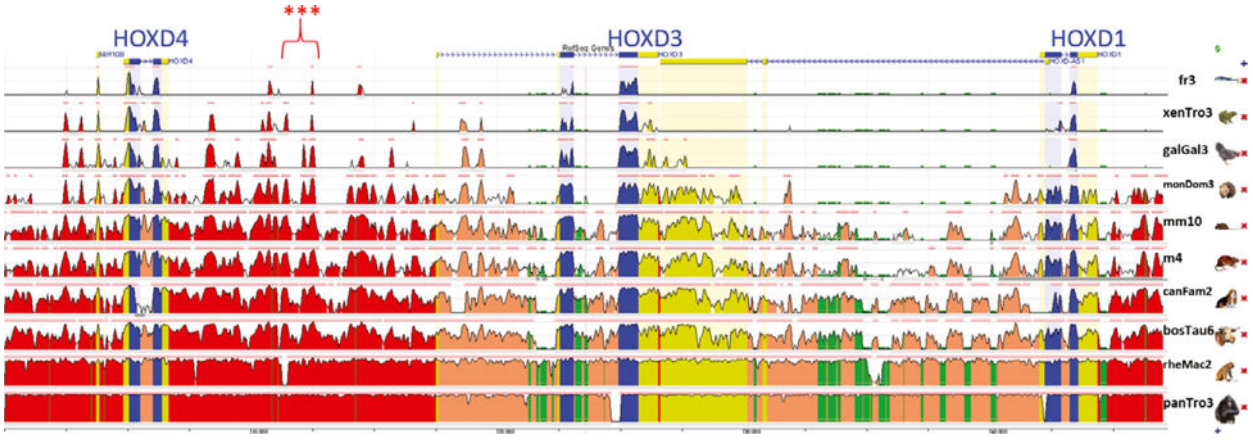

Figure 3.

Evolutionary conserved regions (https://ecrbrowser.dcode.org/) of the orthologous human region (chr2:177 011 292–177 058 499; 90.1% of bases and 100% of span) spanning the HOXD cluster in horse (HOXD4, HOXD3 and HOXD1; chr18:54 605 908–54 650 486). The region of the deletion (red stars) included at least one highly conserved evolutionary peak across all 10 species. Blue, exon; yellow, intron or untranslated region; red, intergenic; green, repeated sequence. Species (listed at right) include Fugu rubripes (pufferfish, genome build fr3), Xenopus (Silurana) tropicalis (western clawed frog, genome build xenTro3), Gallus gallus (chicken, genome build galGal3), Monodelphis domestica (opossum, genome build monDom3), Mus musculus (mouse, genome build mm10), Rattus norvegicus (rat, genome build m4), Canis lupus familiaris (dog, genome build canFam2), Bos Taurus (cow, genome build bosTau6), Rhesus macaque (rhesus macaque, genome build rheMac2) and Pan troglodytes (chimpanzee, genome build panTro3).

Table 1

Significantly segregating genome-wide variants identified in OAAM1 included 6 single nucleotide variants (SNVs) and one 2.7-kb deletion.

					Public database (.bam files)			
Variant	**Ref allele**	**Alt allele**	**OAAM1 genotype**	**SNPEFF**	**Homo Alt allele**	**Het Alt allele**	**Breeds**	**Ensembl transcript ID**
chr9:82 287 859	G	A	AA (WGS) AG (Sanger)	HIGH Nonsense	0	0	N/A	ENSECAT00000012645
chr20:27 996 691	T	C	CC	MODERATE Missense	0	5	Franches-Montagnes Standardbred Unspecified	ENSECAT00000022878
chr25:28 840 158	G	A	AA	MODERATE Missense	0	3	Icelandic Norwegian Fjord	ENSECAT00000016101
chr7:24 869 230	C	T	TT	MODERATE Missense	0	2	Haflinger Franches-Montagnes	ENSECAT00000005230
chr7:65 101 509	T	G	GG	MODERATE Missense	0	9	German Warmblood Haflinger Unspecified Franches-Montagnes Aedigenberger	ENSECAT00000007397
chr8:29 207 931	T	G	GG	MODERATE Missense	0	4	Franches-Montagnes Friesian dwarf Standardbred Norwegian Fjord	ENSECAT00000020160
chr18:54 616 748– 54 619 464		Del	Del/Del	N/A	0	0	N/A	Intergenic (*HOXD4/3*)

Of these, five were identified in the heterozygote state in other breeds as determined by publically accessible 121 mapped genomes of Equus caballus within the Sequence Read Archive database (SRA; https://www.ncbi.nlm.nih.gov/sra). This dataset included two Arabian horses. One SNV (chr9:82 287 859) was unique to OAAM1 and predicted to result in a nonsense mutation in ZNF707 (ENSECAT00000012645) but was found to be a heterozygous variant via Sanger sequencing.

Alt, alternate; Het, heterozygous; Homo, homozygous; Ref, reference; WGS, whole-genome sequencing.

Table 2

Forward and reverse primer sequences used for the confirmation of the deletion sequence.

Primer name	Sequence (5′–3′)	Amplification in OAAM1
IGR 1	F: TTGTGAAGTGGTGGGTGAGA R: TATTCTTGGGATCCCTGCAC	Yes
IGR 2	F: AGGGCCTTTCATAACCCAGA R: CCTCCTCACACACACCCTCT	No
IGR 3	F: AGTGGCCGAGTGTGTGTGT R: TTCATAACGGAGGAAACTTTGTC	No
IGR 4	F: CACCCCAGCAAGAAGTTTTC R: TCGCACCCTTTCAGGAATTA	No
IGR 5	F: AGAGAGGAACCCGCAAAAG R: AGTCTTGGACAGGCGAAGGT	No
IGR 6	F: CTGCCTTTGGTGGTTGTTTT R: ATCCCCAAAGTCTCGCTTTC	No
IGR 7	F: GCGGAGAAGCTTACCCTCTT R: CGGAGGTGTCTCTGAAGCA	No
IGR 8	F: GTACAGGTCACGGCAAACAG R: TGTTTTGCTTTCGGAATGTG	No
IGR 9	F: GAGGAGAAGAGCAGCGTGTC R: GACCCAGAAGCCCTAAACCT	Yes
IGR 10	F: TTCCGAAAGCAAAACAAACC R: GCCAACAGCAAGAGATAAAGG	Yes
IGR 11	F: GGGTCACGTGAACAAATATGC R: CATTTTTCAGCCCATTCTCA	Yes
IGR 12	F: TGCCAAAGTAAGATGGAGCA R: AATGCTGCTCGCAGTAATTG	Yes
IGR 13	F: ACCCCCAAACTACCCAAAAG R: AGTCCCCAGCGACTTCTCTT	Yes

All primer sets utilized an annealing temperature of 60 °C.