9-21-2021

Whole-genome sequencing to investigate a possible genetic basis of perosomus elumbis in a calf resulting from a consanguineous mating

Alexa M. Barber

Alyssa Helms

Riley Thompson

Brian K. Whitlock

David J. Steffen

See next page for additional authors

Authors

Alexa M. Barber, Alyssa Helms, Riley Thompson, Brian K. Whitlock, David J. Steffen, and Jessica Lynn Petersen

Whole-genome sequencing to investigate a possible genetic basis of perosomus elumbis in a calf resulting from a consanguineous mating

Alexa M. Barber,[†,1] Alyssa Helms,[‡] Riley Thompson,[ǁ] Brian K. Whitlock,[§] David J. Steffen,[¶] and Jessica L. Petersen[†,]

[†]Department of Animal Science, University of Nebraska-Lincoln, Lincoln, NE 68503, USA [‡]Department of Large Animal Clinical Sciences, Virginia Tech, Blacksburg, VA 24061, USA [ǁ]College of Veterinary Medicine and Biomedical Sciences, Colorado State University, Fort Collins, CO 80523, USA [§]Department of Large Animal Clinical Sciences, University of Tennessee, Knoxville, TN 37996, USA [¶]School of Veterinary Medicine and Biomedical Science, University of Nebraska – Lincoln, Lincoln, NE 68503, USA

Transl. Anim. Sci. 2021.5:S1–S5
https://doi.org/10.1093/tas/txab171

INTRODUCTION

Perosomus elumbis (PE) is a lethal, congenital defect marked by aplasia of the lumbar and sacral spine and spinal cord. Contracture of the hind limbs is also commonly observed in affected individuals. PE has been reported in many domestic species, with numerous case reports in Holstein cattle in the past two decades (Jones, 1999; Karakaya et al., 2013; Agerholm et al., 2014). The etiology of PE remains unknown. In one instance, a stillborn Holstein calf with PE was found to be infected with Bovine Viral Diarrhea Virus (BVDV) (Karakaya et al., 2013), and thus, it is possible that PE may be due to genetic and/or environmental factors. Recently, a stillborn Angus calf was diagnosed with PE following an accidental mother-son mating (Helms et al., 2020). BVDV was not detected in the affected Angus calf, dam, nor sire. Due to the relationship between the sire and dam, it was hypothesized that a novel, recessive genetic variant may be responsible for the development of PE in this Angus calf. The objective of this study was to use whole-genome sequencing to address this hypothesis and identify candidate variants for PE in this calf.

MATERIALS AND METHODS

IACUC Statement

All procedures and protocols were performed following the University of Nebraska – Lincoln's Institutional Animal Care and Use Committee guidelines.

Sample Collection and DNA Isolation

Case presentation and diagnosis are reported in Helms et al. (2020). Tissue samples were collected from the affected calf following necropsy at the University of Tennessee Veterinary Medical Center. Blood samples were also taken from the dam, sire, and ten paternal half-siblings; tissue and blood were sent to the University of Nebraska – Lincoln. DNA was isolated from tissue and blood utilizing Qiagen Gentra Puregene Kits (Gentra Systems, Minneapolis, MN). Paternity was verified for all calves using the commercially available SeekSire parentage assay at Neogen GeneSeek (Lincoln, NE).

[1]Corresponding author: alexa.barber@huskers.unl.edu
Received May 4, 2021.
Accepted September 21, 2021.

Whole Genome Sequencing and Variant Filtering

DNA collected from the affected calf, the dam, the sire, and three paternal half-siblings was sent to Admera Health (South Plainfield, NJ) for KAPA library prep and 150 bp paired-end sequencing on an Illumina NovaSeq to a targeted sequencing depth of 12X. After trimming adapter sequences and poor quality bases (TrimGalore; Wu et al., 2011), sequence reads from the calf, dam, sire, and half siblings along with 25 other Angus and Angus-cross animals were mapped to the UOA_Angus_1 reference genome with BWA-MEM (Li, 2013).

Variants were called using Freebayes (Garrison and Marth, 2012) and annotated using SnpEff (Cingolani et al., 2012). SnpSift was also used to filter variants in which the affected calf was homozygous and both the dam and sire were heterozygous. With the assumption that PE is rare in Angus cattle, variants were further pruned using VCFtools (Danecek et al., 2011) to select only variants in which the alternative allele count was between four and seven to account for a homozygous calf, two heterozygous parents, and allow for the half-siblings to be heterozygous. Variants were further reduced to include only those predicted to have a moderate to high impact. Variants fitting the aforementioned criteria were further investigated. Variants were remapped to the ARS-UCD1.2 reference genome using NCBI's Remap tool to determine if the variants had been previously reported.

PCR and Sanger Sequencing

Primers for regions of interest were developed using sequence from the UOA_Angus_1 reference genome. Oligonucleotides were designed using IDT's PrimerQuest Tool. PCR products were amplified using an annealing temperature ranging between 54 and 58 °C and visualized on 1.2% agarose gels. PCR products were sent to ACGT Inc. (Wheeling, IL) for Sanger sequencing. Sequence results were visualized using Gene Codes Corporation's Sequencher.

Sequence Read Archive Search

A search of NCBI's Sequence Read Archive (SRA) was conducted using a variant Search pipeline (https://github.com/SichongP/SRA_variant_search); NCBI's Remap function was used to identify coordinates across genome assemblies. The frameshift variant was not assessed in the SRA search due to difficulty interpreting indels using this method.

RESULTS

Candidate Variant Filtering

Variant calling across the 31 Angus and Angus cross individuals, including the affected calf, the dam, the sire, and three half siblings, identified 21,223,927 variants across the genome. Using SnpSift to filter for variants in which the calf was homozygous for the variant and the dam and sire were heterozygous yielded 506,813 variants. Removing variants at high frequency in the data set reduced candidate variants to 14,011.

Filtering by predicted impact as annotated in SnpEff resulted in 77 variants with a predicted moderate impact and 5 with predicted high impact. Predicted high-impact variants were excluded from further analysis if they were previously annotated and a carrier was found in the original 31 screened animals or if the variant was found in the homozygous state in any individual(s) other than the affected calf. After removing variants fitting those criteria, the final candidate variant list consisted of 18 missense variants and one frameshift resulting from a one base pair deletion. Three of the 19 candidate variants were not previously annotated on Ensembl (Table 1). The frameshift variant was located in exon 4 of protein tyrosine kinase 7 (*PTK7*) and is predicted to result in a premature stop codon prior to the end exon. Due to its putative deleterious impact on gene function, this variant was further studied as a candidate causal variant.

Sanger Sequencing Verification of Frameshift Mutation in PTK7 and SRA Results

Sanger sequencing confirmed the presence of a homozygous, one base pair deletion in the affected calf. Additionally, six of 10 half-siblings were heterozygous for the deletion (Figure 1).

The search of the Sequence Read Archive (SRA) resulted in genotypes of 883 additional cattle including 96 Angus and Angus cross. Through this analysis, individuals homozygous for variants were identified at 15 of the 18 missense loci; the indel in *PTK7* was not able to be queried.

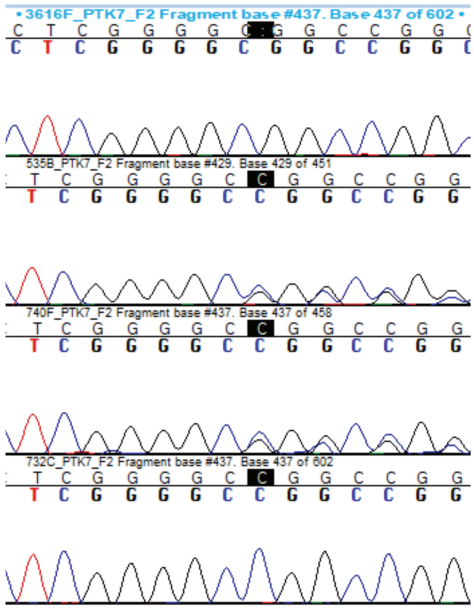
The three remaining missense variants in *KDM1A, C2H2orf66,* and *ZSCAN26,* and one frameshift variant in *PTK7* remained as candidate causal variants (Table 1). From the SRA data, 1 Holstein was heterozygous for the *KDM1A*

Table 1. Candidate variants for perosomus elumbis

Chr	Position	Gene	Reference	Variant	Type	Variant ID
2	*UOA: 6567555*	*KDM1A*	*C*	*T*	*Missense*	*Novel*
	ARS: 129835952				*Missense*	*Novel*
2	*UOA: 50768895*	*C2H2orf66*	*T*	*C*	*Missense*	–
	ARS: 85400473				*Intergenic*	*rs719944515*
4	UOA: 6313462	ASIC3	C	T	Missense	–
	ARS: 113625394				Missense	rs466455595
15	UOA: 63709344	QSER1	A	G	Missense	–
	ARS: 63626227				Intronic	rs380723979
15	UOA: 63709425	QSER1	CA	GG	Missense	–
	ARS: 63626308				Intronic	rs799405617
17	UOA: 51162578	NCOR2	C	T	Missense	–
	ARS: 51449160				Missense	rs472931263
17	UOA: 51555593	DNAH10	T	C	Missense	–
	ARS: 51850428				Missense	rs136088999
22	UOA: 59081586	EFCC1	G	T	Missense	Novel
	ARS: 58980413				Missense	Novel
23	UOA: 22015758	ZNF165	G	A	Missense	–
	ARS: 30587728				Missense	rs526649482
23	UOA: 22140420	ZSCAN9	C	T	Missense	–
	ARS: 30463098				Downstream	rs463835998
23	*UOA: 22181166*	*ZSCAN26*	*C*	*A*	*Missense*	–
	ARS: 30422277				*Missense*	rs521986257
23	UOA: 22192722	PGBD1	G	A	Missense	–
	ARS: 30410722				Missense	rs432139616
23	UOA: 22192789	PGBD1	C	T	Missense	–
	ARS: 30410655				Missense	rs449832006
23	UOA: 23246829	OR1O9	C	T	Missense	–
	ARS: 29295898				CNV	nsv835503
23	UOA: 23452894	OR2H1D	A	C	Missense	–
	ARS: 29099147				Downstream/CNV	rs800181923
23	**UOA: 35361113**	**PTK7**	**CG**	**G**	**Frameshift**	**Novel**
	ARS: 16744942				**Frameshift**	**Novel**
28	UOA: 25626059	TSPAN15	A	G	Missense	–
	ARS: 25846885				Missense	rs469369204
28	UOA: 26696544	ADAMTS14	C	T	Missense	–
	ARS: 26918561				Missense	rs135381293
28	UOA: 30658181	DUSP13	C	T	Missense	–
	ARS: 30876012				Missense	rs379594626

Italicized rows indicate that no individuals were homozygous for the variant in the Sequence Reads Archive (SRA) search. Bolded rows indicate variants with a predicted high impact on gene function from SnpEff (Cingolani et al., 2012). Positions labeled UOA correspond to the UOA_Angus_1 reference genome, and positions labeled ARS correspond to the ARS_UCD1.2 reference genome. Previously annotated variants are noted under Variant ID. Type represents the predicted position/outcome observed in the UOA_Angus_1 reference genome (top) and the ARS_UCD1.2 reference genome (bottom).

A)



B)

Individual	Genotype
Affected Calf	-/-
Dam	C/-
Sire	C/-
Half Sibling 1	C/-
Half Sibling 2	C/C
Half Sibling 3	C/-
Half Sibling 4	C/-
Half Sibling 5	C/-
Half Sibling 6	C/-
Half Sibling 7	C/C
Half Sibling 8	C/C
Half Sibling 9	C/C
Half Sibling 10	C/-

Figure 1. Sanger sequencing confirms the presence of a one base pair deletion in *PTK7*. (A) Sequence data from exon 4 of *PTK7*, as viewed in Sequencher (Gene Codes Corporation), depict the presence of a deletion for which the affected calf was homozygous, the dam and half-sibling heterozygous, and second half-sibling wild-type. (B) Genotypes of the affected calf, the dam, the sire, and ten half-siblings at the candidate locus in *PTK7*. A dash (-) indicates the 1bp deletion.

variant; 2 Tyrolean Grey cattle, 1 Chianina, and 1 Romagnola were heterozygous for the *C2H2orf66* variant; and 2 Angus, 1 Chi-Angus cross, and 1 Holstein were heterozygous for the *ZSCAN26* variant.

DISCUSSION

In this study, missense mutations in *KDM1A, C2H2orf66,* and *ZSCAN26*, as well as a frameshift mutation in *PTK7* could not be ruled out as causative of PE in this Angus calf. PE is a lethal congenital defect that results in aplasia of the lumbar spine and frequent contracture of the hind limbs. Although relatively rare in Angus cattle, numerous cases of PE have been reported in Holstein cattle. The cause of PE has yet to be determined with both environment and genetics suspected to play a role. In this case, the affected Angus calf was the result of a consanguineous mating suggesting that a recessive mutation may be the cause.

Of the four variants remaining after filtering out those that did not fit the hypothesized mode of inheritance, and those at high frequency in other cattle, the missense mutation in *KDM1A* and the frameshift mutation in *PTK7* are strong functional candidates due to their roles in early development. *KDM1A* is involved in epigenetic regulation of embryonic gene expression (Ancelin et al., 2016), whereas *PTK7* functions in the planar cell polarity (PCP) pathway that regulates cell movement and migration (Berger et al., 2017).

KDM1A is a histone 3 lysine 4 (H3K4) lysine demethylase that functions to remove enhancer marks from histones. These epigenetic marks influence early development in part by regulating the spaciotemporal activation of genes which orchestrates proper embryonic development (Ancelin et al., 2016). Dysregulation of *KDM1A* can result in developmental arrest and altered patterns of gene expression in the developing embryos (Ancelin et al., 2016).

PTK7, a member of the tyrosine kinase family, plays a role in the planar cell polarity (PCP) pathway. This pathway establishes polarity in cells and regulates cell movement and migration in embryonic development (Berger et al., 2017). This gene is of particular interest as it has been implicated in congenital scoliosis in zebrafish (Hayes et al., 2014) demonstrating a clear role in the development of the fetal spine. Additionally, another gene with a paralog in this pathway, *VANGL1*, has been implicated in an analogous human disorder called caudal regression syndrome (CRS) (Kibar et al., 2007; Porsch et al., 2016). Furthermore, *VANGL2*, which directly interacts with *PTK7* in the PCP pathway, has also been implicated in neural tube defects (Kibar et al, 2011). These studies demonstrate a clear role of *PTK7* and the PCP pathway in spinal development making a frameshift mutation in *PTK7* a strong functional candidate for PE in cattle.

Although *PTK7* provides a strong functional candidate for PE, this study is limited due to the availability of a single affected calf. This study should be supplemented with additional affected calves as cases are reported. Furthermore, as new sequence reads become available in the SRA database, additional animals can be screened for the associated variants found in this study. Due to the rarity of this condition, this study could be extended to consider affected calves from other breeds.

Implications

The accumulation of lethal recessive variation within breeds negatively impacts production and breed health. With the growing use of artificial insemination (AI), prolific carrier bulls can rapidly increase the allele frequency of recessive disorders within the breed. Through the use of whole-genome sequencing, disease-associated and disease-causing variation can be identified. Although a causative variant was not validated in this study, in the case that would occur, genetic testing could allow for informed matings to eliminate the production of affected individuals.

LITERATURE CITED

Agerholm, J. S., W. Holm, M. Schmidt, P. Hyttel, M. Fredholm, and F. J. McEvoy. 2014. perosomus elumbis in Danish Holstein cattle. BMC Vet. Res. 10:227. doi:10.1186/s12917-014-0227-2

Ancelin, K., L. Syx, M. Borensztein, N. Ranisavljevic, I. Vassilev, L. Briseño-Roa, T. Liu, E. Metzger, N. Servant, E. Barillot, C. J. Chen, R. Schüle, and E. Heard. (2016). Maternal LSD1/KDM1A is an essential regulator of chromatin and transcription landscapes during zygotic genome activation. eLife, 5, e08851. https://doi.org/10.7554/eLife.08851

Berger, H., A. Wodarz, and A. Borchers. 2017. PTK7 faces the Wnt in development and disease. Front. Cell Dev. Biol. 5:31. doi:10.3389/fcell.2017.00031

Cingolani, P., A. Platts, l. e. L. Wang, M. Coon, T. Nguyen, L. Wang, S. J. Land, X. Lu, and D. M. Ruden. 2012. 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). 6:80–92. doi:10.4161/fly.19695

Danecek, P., A. Auton, G. Abecasis, C. A. Albers, E. Banks, M. A. DePristo, R. E. Handsaker, G. Lunter, G. T. Marth, S. T. Sherry, et al.; 1000 Genomes Project Analysis Group. 2011. The variant call format and VCFtools. Bioinformatics 27:2156–2158. doi:10.1093/bioinformatics/btr330.

Garrison E, G. Marth. Haplotype-based variant detection from short-read sequencing. 2012. arXiv preprint arXiv:1207.3907 [q-bio.GN].

Hayes, M., X. Gao, L. X. Yu, N. Paria, R. M. Henkelman, C. A. Wise, and B. Ciruna. 2014. ptk7 mutant zebrafish models of congenital and idiopathic scoliosis implicate dysregulated Wnt signalling in disease. Nat. Commun. 5:4777. doi:10.1038/ncomms5777.

Helms, A. B., R. E. Thompson, S. Lawton, J. L. Petersen, A. Watson, M. J. Sula, D. Steffen, and B. K. Whitlock. 2020. Uterine torsion dystocia complicated by perosomus elumbis in an Angus calf associated with a consanguineous mating. Case Rep. Vet. Med. 2020:6543037. doi:10.1155/2020/6543037.

Jones, C. J. 1999. Perosomus elumbis (vertebral agenesis and arthrogryposis) in a stillborn Holstein calf. Vet. Pathol. 36:64–70. doi:10.1354/vp.36-1-64.

Karakaya, E., G. Alpay, G. Yilmazbas-Mecitoglu, A. Alasonyalilar-Demirer, B. Akgül, S. Inan-Ozturkoglu, M. O. Ozyigit, D. Seyrek-Intas, K. Seyrek-Intas, K. Yesilbag, et al. 2013. Perosomus elumbis in a Holstein calf infected with bovine viral diarrhea virus. Tierarztl. Prax. Ausg. G. Grosstiere. Nutztiere. 41:387–391.

Kibar, Z., S. Salem, C.M. Bosoi, E. Pauwels, P. De Marco, E. Merello, A.G. Bassuk, V. Capra, and P. Gros. (2011). Contribution of VANGL2 mutations to isolated nerual tube defects. Clin. Genet. 80(1):76–82. doi: 10.1111/j.1399-0004.2010.01515.x

Kibar, Z., E. Torban, J.R. McDearmid, A. Reynolds, J. Berghout, M. Mathieu, I. Kirillova, P. De Marco, E. Merello, J.M. Hayes, J.B. Wallingford, and P. Drapeau. (2007). Mutations in *VANLG1* associated with neural-tube defects. N. Eng. J. Med. 356:1432–1437.

Li H. 2013. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. arXiv:1303.3997v2 [q-bio.GN].

Porsch, R. M., E. Merello, P. De Marco, G. Cheng, L. Rodriguez, M. So, P. C. Sham, P. K. Tam, V. Capra, S. S. Cherny, et al. 2016. Sacral agenesis: a pilot whole exome sequencing and copy number study. BMC Med. Genet. 17:98. doi:10.1186/s12881-016-0359-2.

Wu, Z., X. Wang, and X. Zhang. 2011. Using non-uniform read distribution models to improve isoform expression inference in RNA-Seq. Bioinformatics 27:502–508. doi:10.1093/bioinformatics/btq696.