Multiple independent de novo mutations are associated with the development of schistosoma reflexum, a lethal syndrome in cattle

J.G.P. Jacinto, I.M. Häfliger, A. Letko, J. Weber, M. Freick, A. Gentile, C. Drögemüller, J.S. Agerholm

PII: S1090-0233(24)00008-X

DOI: https://doi.org/10.1016/j.tvjl.2024.106069

Reference: YTVJL106069

To appear in: *The Veterinary Journal*

Received date: 11 August 2023
Revised date: 22 January 2024
Accepted date: 23 January 2024

Please cite this article as: J.G.P. Jacinto, I.M. Häfliger, A. Letko, J. Weber, M. Freick, A. Gentile, C. Drögemüller and J.S. Agerholm, Multiple independent de novo mutations are associated with the development of schistosoma reflexum, a lethal syndrome in cattle, *The Veterinary Journal,* (2024) doi:https://doi.org/10.1016/j.tvjl.2024.106069

Original Article

Multiple independent de novo mutations are associated with the development of schistosoma reflexum, a lethal syndrome in cattle

J. G. P. Jacinto [a,b], I. M. Häfliger [b], A. Letko [b], J. Weber [c], M. Freick [d] [d], A. Gentile[a], C. Drögemüller [b,#,*], J. S. Agerholm [e,#]

[a] *Department of Veterinary Medical Sciences, University of Bologna, Via Tolara di Sopra 50, 40064 Ozzano dell'Emilia (Bologna), Italy*
[b] *Institute of Genetics, Vetsuisse Faculty, University of Bern, Bremgartenstrasse 109a, 3012 Bern, Switzerland*
[c] *Clinic for Ruminants, Department of Clinical Veterinary Medicine, Vetsuisse Faculty, University of Bern, Bremgartenstrasse 109a, 3012 Bern, Switzerland*
[d] *Faculty of Agriculture/Environment/Chemistry, HTW Dresden-University of Applied Sciences, 01326 Dresden, Germany*
[e] *Department of Veterinary Clinical Sciences, University of Copenhagen, Højbakkegaard Allé 5A, 2630, Taastrup, Denmark*

[#] Shared senior authorship

* Corresponding author. Tel.:+41 31 684 25 29
E-mail address: cord.droegemueller@unibe.ch (C. Drögemüller).

Abstract

Schistosoma reflexum (SR) is a lethal congenital syndrome characterized by U-shaped dorsal retroflexion of the spine and exposure of abdominal viscera. SR is usually associated with severe dystocia. The syndrome is thought to be inherited as a Mendelian trait. We collected a series of 23 SR-affected calves from four breeds (20 Holstein, one Red Danish, one Limousin, one Romagnola) and performed whole-genome sequencing (WGS). WGS was performed on 51 cattle, including 14 cases with parents (trio-based; Group 1) and nine single cases (solo-based; Group 2).

Sequencing-based genome-wide association studies with 20 Holstein cases and 154 controls showed no association (above Bonferroni threshold; P-value<$3x10^{-09}$). Assuming a monogenic recessive inheritance, no region of shared homozygosity was observed, suggesting heterogeneity. Alternatively, the presence of possible dominant acting de novo mutations

were assessed. In Group 1, heterozygous private variants, absent in both parents, were found in seven cases. These involved the *ACTL6A*, *FLNA*, *GLG1*, *IQSEC2*, *MAST3*, *MBTPS2*, and *MLLT1* genes. In addition, heterozygous private variants affecting the genes *DYNC1LI1*, *PPP2R2B*, *SCAF8*, *SUGP1*, and *UBP1* were identified in five cases from Group 2. The detected frameshift and missense variants are predicted to cause haploinsufficiency. Each of these 12 affected genes belong to the class of haploinsufficient loss-of-function genes or are involved in embryonic and pre-weaning lethality or are known to be associated with severe malformation syndromes in humans and/or mice. This study presents for the first time a detailed genomic evaluation of bovine SR, suggesting that independent de novo mutations may explain the sporadic occurrence of SR in cattle.

Keywords: Bovine; Dystocia; Haploinsufficiency; Reflexus; Schistosomus

Introduction

Schistosoma reflexum (SR) is a lethal syndrome characterized by congenital U-shaped dorsal retroflexion of the spine, exposure of abdominal viscera due to thoracoschisis and abdominoschisis, arthrogryposis and ankylosis of the hind limbs, and occasionally palatoschisis and other defects. SR has mainly been reported in cattle (Online Mendelian Inheritance in Animals (OMIA000890-9913) but reports of a similar syndrome have also been published in goats (OMIA000890-9925), donkeys (OMIA000890-9793), horses (OMIA000890-9796), sheep (OMIA000890-9940), cats (Mateo and Camón, 2008), dogs (Özsoy et al., 2009), and pigs (Martín-Alguacil and Avedillo, 2020). The cause of SR has not been investigated at the molecular level.

SR is a cause of foetal dystocia and appears to be more common in cattle than in other animal species (Roberts, 1986; Richter and Götze, 1993; Cavalieri and Farin, 1999). SR has

been reported to account for 1.3% of dystocia cases in cattle attended by veterinarians (Knight, 1996). Although SR was first reported in cattle more than 170 years ago (Fleming, 1850), little is known about the aetiology. A genetic cause with an autosomal recessive mode of inheritance has been hypothesized by pedigree analysis in Holstein cattle (Citek, 2012), but conclusive evidence for the inheritance of SR has not been provided and molecular investigation of the pathogenesis has not been performed.

The aim of this study was to briefly describe the phenotype of a series of SR-affected calves and to perform a genomic investigation in the cases and, where available, their parents, by whole-genome sequencing (WGS) to evaluate whether SR in cattle has a possible genetic cause.

Materials and methods

Ethics statement

This study did not require regulatory or institutional ethical approval, as it was not experimental. Cases were submitted by owners for research and bovine genetic disease surveillance purposes.

Animals and phenotypical investigation

SR-affected calves submitted to the Department of Veterinary Clinical Sciences, University of Copenhagen, Denmark, to the Department of Veterinary Medical Sciences, University of Bologna, Italy and to the Faculty of Agriculture/Environment/Chemistry, HTW Dresden-University of Applied Sciences, Germany from 2002 to 2022 were considered and tissues were stored at -20°C until analysis. The cases were divided into two groups for the purpose of DNA analysis: Group 1 consisted of cases where tissue was available from both

the SR affected calf and its parents, and Group 2 consisted of cases where only tissue from the affected calf itself was available.

Length of gestation of dams and birth weight of SR cases were recorded when possible and calves from twin pregnancies were excluded from the birth weight calculations. Necropsy as diagnostic method was performed when possible. Otherwise, cases were diagnosed on the basis of photographic documentation provided by the veterinarian attending the birth.

DNA extraction

Genomic DNA was extracted from SR cases (from ear cartilage or liver) and also from their parents when available (whole blood – EDTA anticoagulant – from dams and semen from sires) using the Maxwell RSC DNA System (Promega).

Whole-genome sequencing and variant calling

WGS data were generated for Group 1 and Group 2 using the Illumina NovaSeq6000 (Illumina Inc.). Trio-based analysis was performed in Group 1, while the solo-based approach was used in Group 2. Sequenced reads were aligned to the ARS-UCD1.2 reference genome (Rosen et al., 2020). The average coverage obtained for each genome is shown in Supplementary Table S1. Single-nucleotide variants (SNVs) and small indel variants were called. The applied software and steps to process fastq files into binary alignment map (BAM) and genomic variant call format (GVCF) files followed the 1000 Bull Genomes Project (run 7) processing guidelines (Hayes and Daetwyler, 2019), except for trimming, which was performed using fastp (Chen et al., 2018). The reason for using fastp for trimming was to support updates on the bioinformatics server. Fastp proved to be very efficient and much faster than trimmomatic, reducing data processing time.

Further processing of the genomic data was performed as previously reported (Jacinto et al., 2021). To find private variants, we compared the genotypes of SR case(s) with 896 bovine genomes of different breeds sequenced as part of the ongoing Swiss Comparative Bovine Resequencing project. All its data are available in the European Nucleotide Archive (project accession number PRJEB18113). Regarding the method of inheritance, three different scenarios were hypothesized: i) autosomal recessive mode of inheritance common to all SR cases of the same breed, ii) autosomal recessive mode of inheritance considering each case individually, or iii) dominant mode of inheritance acting de novo considering each SR case as an isolated event. Integrative Genomics Viewer (IGV) software version 2.0 (Robinson et al., 2017) was used for visual assessment of genomic regions containing potential candidate genes.

Sequencing-based genome-wide association study

A sequencing-based genome-wide association study (seqGWAS) was performed in order to test the hypothesis of a causal recessive variant common to all SR cases of the same breed. Biallelic variants were selected from the vcf file using plink command --biallelic-only as a common quality control step. This filter was used because multiallelic sites are problematic for subsequent analyses due to not being handled consistently by different bioinformatic tools. Furthermore, true multiallelic variants are observed very rarely and therefore, considered a sign of low-quality regions where artifacts and falsely called variants are likely to occur as documented in the Genome Analysis Toolkit guidelines (McKenna et al., 2010). For the control cohort, 154 phenotypically normal, not closely related Holstein cattle (identical by descent <0.5 as detected by PLINK v1.9) (Chang et al., 2015) were selected from the Swiss Comparative Bovine Resequencing project. In addition, a genome-

wide search for homozygous regions shared by the 20 cases was performed using the R package detectRUNS v.0.9.6 (Biscarini et al., 2019).

Evaluation of larger structural variants

To assess possible larger structural variants, including chromosomal, numerical, and structural abnormalities, the read depth along all chromosomes was calculated. A sliding window approach was used with two different window sizes (10 kb, 200 kb). The bedcov function of the Samtools program (Li, 2011) was used to generate the number of reads within each specified window. Coverage plots were generated using the Manhattan function of the qqman package in R (R Core Team).

Occurrence of variants in a global control cohort

The comprehensive variant catalogue from run 9 of the 1000 Bull Genomes Project was available to investigate the allelic distribution of variants within a global control cohort (Hayes and Daetwyler, 2019). The full dataset includes 5116 bovine genomes, including 576 from the Swiss Comparative Bovine Resequencing Project, from a wide variety of >130 breeds. Within the dataset, there were 1209 purebred Holstein, 74 Red Danish, 101 Limousin, and 25 Romagnola cattle allowing for the exclusion of variants common to these breeds.

In silico assessment of the molecular consequences

PredictSNP1 (Bendl et al., 2014), PolyPhen-2 (Adzhubei et al., 2010), and SIFT (Ng and Henikoff, 2003) were used to predict the biological consequences of the candidate variants. GnomAD was used to predict the probability of a gene being intolerant (pLI) (Karczewski et al., 2020).

Candidate gene and candidate variant definitions

The term “candidate gene” was used to describe genes based on function and/or associated phenotype in mammalian species. The term candidate variant was used to describe variants considering the affected gene function and/or associated phenotype in mammalian species, rarity and the predicted effect of the variant on the encoded protein. In addition, the VarElect software (LifeMap Sciences Inc) was used for variant selection for phenotype-dependent gene variant prioritisation according to the phenotype query "skeletal OR dysplasia" to identify possible candidate dominant variants when more than 360 private (present in a SR case and absent in the controls) protein-changing variants were identified. All sequence accessions used for the candidate genes are listed in Supplementary Table S2.

Results

Phenotypical features

A total of 23 SR-affected calves were sampled. Twenty cases were submitted to the Department of Veterinary Clinical Sciences, University of Copenhagen, Denmark, two to the Department of Veterinary Medical Sciences, University of Bologna, Italy, and one to the Faculty of Agriculture/Environment/Chemistry, HTW Dresden-University of Applied Sciences, Germany. Fourteen calves were assigned to Group 1 and nine cases to Group 2.

Twenty cases were Holstein, while the Red Danish, Limousin, and Romagnola breeds were represented by one case each. Nine calves were female and 14 were male. Holstein SR cases were familiarly unrelated for at least two generations. Twenty-one cases were submitted by veterinarians assisting in a dystocia case, while two cases were submitted by abattoir officials recognizing an SR foetus at post-mortem inspection.

Necropsy was performed in 12 SR cases, while 11 cases were diagnosed on the basis of photographic documentation provided by the veterinarian attending the birth.

Gestation length at parturition was available for 16 SR-affected Holstein calves. The mean gestation length was 282 days (S.D. ±5.5), which is within the reference range for Holsteins in Denmark (Agerholm et al., 2023). The mean body weight of the necropsied Holsteins (n=8) was 28.3 kg (S.D. ±6.6) compared to the reported average birth weight of 42 kg for Holstein calves in Denmark (Agerholm et al., 2023). One case of SR from a twin pregnancy was excluded from the calculations as twin calves weigh less than single calves.

The SR cases had a common morphology with retroflexion of the spine at the level of the thoracolumbar junction and with involvement of segments of both the thoracic and lumbar spine as the main lesion.

In the cases that underwent necropsy, the lumbar spine was also twisted, so that the hind quarter of the calf was rotated relative to the anterior part of the calf, and one case had a torticollis. The thoracoabdominal schisis of the ventral midline had developed as a result of the retroflexion of the spine, resulting in a wide ventral opening through which the thoracic and abdominal organs herniated and were exposed externally (Fig. 1). The thoracic organs were surrounded by the diaphragm and parietal pleura, which together created a narrow cavity that prevented normal development of the lungs and heart. The lungs were hypoplastic and nonaerated, while the heart showed a variety of malformations such as interventricular septum defect, dilation of the truncus pulmonalis, myocardial hypertrophy, and reduced ventricular volume. The liver was abnormally shaped, indurated and sometimes interspersed with thin-walled cysts with a clear serous fluid attached to the serosa. The abdominal organs were

otherwise unremarkable, except for occasional occurrence of cysts in the renal cortex, anal or rectal atresia, and abdominal cryptorchidism. Lesions in the limbs were mostly limited to arthrogryposis of the fetlock joints but in one case there was extensive arthrogryposis of the right hind limb. Palatoschisis, asymmetry of the skull, and flattening of the calvarium associated with an Arnold-Chiari type 1-like malformation (Fig. 2) were observed in three necropsied cases (3/12 cases).

Whole genome sequencing

WGS data were generated for a total of 51 animals. Trio-based analysis was performed for 42 animals of the 14 trios from Group 1, while the solo-based approach was used for the nine individual cases from Group 2.

Exclusion of recessive inheritance

Assuming a simple recessive mode of inheritance, the WGS data were filtered for homozygous coding variants privately present in the 20 SR-affected Holstein calves, but without identifying any single-nucleotide or small indel variants common to all cases. Filtering for homozygous variants present only in the genomes of the individual cases using a global cohort of 5347 bovine control genomes revealed no evidence of rare pathogenic alleles. The same approach was used for the Red Danish, Limousin, and Romagnola SR cases, considering each case individually and similarly, no candidate recessive variants were identified.

For the 20 Holstein SR cases, the GWAS results, using only 15,228,104 biallelic SNVs, revealed 70 statistically significant single markers (above the Bonferroni threshold, P-value<$3x10^{-09}$; Supplementary Fig. S1; Supplementary Table S3). All these loci were further

visualized in IGV, revealing that the identified loci were most likely artefacts that were also present in unrelated unaffected animals from the global control cohort. Furthermore, no region of shared homozygosity was observed. These results suggest that simple recessive inheritance is unlikely to explain the development of SR in cattle.

Evidence for candidate causal de novo mutations

WGS data were filtered for heterozygous variants present in each case individually and absent in controls, including the parents (when available), under the assumption that a dominant acting de novo mutation was the cause. Because of the lethal effect, we hypothesised that loss-of-function variants affecting the coding sequence of a gene were most likely to be responsible for SR.

The results of the different filtering steps for the trio-based SR cases in Group 1 are shown in Table 1 and for the solo-based cases in Group 2 in Table 2. It was possible to identify private heterozygous candidate variants affecting a possible candidate gene for eight cases in the trio-based group and for eight cases in the solo-based group (Supplementary Table S4).

The variants in these sixteen genes were further analysed for their predicted effect on the encoded protein. For the trio-based Group 1, seven independent variants affecting different genes were predicted to be deleterious. Five of these identified variants were frameshift variants in the genes *ACTL6A* (p.Met92fs), *FLNA* (p.Val60fs), *GLG1* (p.Glu511fs), *MAST3* (p.Pro1202fs), and *MBTPS2* (p.Gly273fs), which are predicted to result in a completely different amino acid sequence when expressed compared to the wild-type proteins (Table 3). In addition, two missense variants were identified in *IQSEC2* (p.Leu957Pro) and in

MLLT1 (p.Arg20Cys), affecting a highly conserved amino acid across species, and predicted to be deleterious by various bioinformatic tools (Table 3).

For the solo-based Group 2, five private heterozygous variants affecting different genes were predicted to be deleterious. Two of these variants were frameshift insertions affecting *PPP2R2B* (p.Val245fs) and *SCAF8* (p.Val378fs; Table 4). In addition, three missense variants in *DYNC1LI1* (p.Arg505Trp), *SUGP1* (p.Arg326Cys) and *UBP1* (p.Arg388Gly), affecting a highly conserved amino acid across species, were identified and predicted to be deleterious by different bioinformatic tools (Table 4).

Exclusion of possible larger structural variants

No evidence of chromosomal abnormalities was found by analysing the obtained read depth or coverage along all chromosomes for all cases.

Discussion

The morphology of SR in cattle is dominated by the marked retroflexion of the thoracolumbar spine as observed in the 23 SR cases in this study. Considering spinal retroflexion as the primary lesion in SR may explain the thoracoabdominal schisis as a spinal retroflexion hypothetically disrupts the development of the body walls, since vertebrae, ribs and abdominal wall musculature originate from the same embryonic structure (somites). Exposure of the thoracic and abdominal organs then occurs spontaneously due to the lack of ventral closure of the thorax and abdomen. The syndrome develops during early foetal development and has been observed in mid-term foetuses (Windsor, 2019). SR should therefore be considered as a skeletal malformation rather than a primary defect of the body wall.

A Mendelian inheritance has been previously hypothesized for SR in cattle. Citek (2012) proposed a simple recessive mode of inheritance of SR in Holstein cattle based on the observation of affected offspring from phenotypically normal but related parents. Our genomic analysis of 20 cases of SR in the Holstein breed revealed the possibility that different single specific mutations can cause slightly different phenotypes. Similar results were obtained for the Red Danish, Limousin, and Romagnola breeds. Previously reported cases of SR in breeds other than Holstein have mostly been isolated, making it difficult to hypothesise the most likely mode of inheritance (Supplementary Table S5). The discovery that SR is most likely caused by mutations in different genes opens up new perspectives in the understanding of the syndrome and the function of the affected genes. Nevertheless, a more detailed study of SR cases combined with functional experiments will be still required to better understand the association between specific mutations and this syndromic pathology. Very recently, Park et al. (2023) reported a single Holstein foetus with SR. Interestingly, the foetus tested homozygous for the *APAF1* nonsense variant, previously identified as the causal variant for the deficient homozygous haplotype region 1 (HH1), which is known for its negative associations with female fertility traits, primarily due to embryonic lethality (Adams et al., 2016). However, this variant was not observed in the SR cases presented in our study. Therefore, it does not seem plausible as the cause of all SR Holstein cases.

The present study describes, for the first time, WGS-based molecular genetic findings for SR in cattle and provides evidence for an unexpected heterogeneous cause of SR by spontaneous mutations affecting different genes important for proper development, located on different chromosomes. Such de novo mutations may have been present in a single parental gamete or may have occurred during early embryonic development of the SR cases, and therefore need to be considered individually. In this study, 12 protein-changing heterozygous

variants located in genes involved in embryonic and pre-weaning lethality and/or in severe malformation syndromes (Table 5) are proposed as potential causes of SR in cattle. In addition, recent large-scale data from human genome sequencing studies presented in the Genome Aggregation Database (gnomAD; Karczewski et al., 2020) showed that the probability of a loss-of-function intolerance score for the 12 candidate genes ranged from 0.98 to 1, meaning that these genes belong to the class of haploinsufficient loss-of-function genes.

A congenital syndrome similar to SR has not been reported in humans. Variants of the 12 genes we report in SR-affected cattle are associated with congenital malformation syndromes in humans and mice, but none resemble SR, although some include body wall defects and scoliosis (Table 5). Taken together, the rarity of the identified candidate variants, their predicted deleterious effect on the respective genes that have known function in developmental malformations and embryonic death in mice or human, the 12 highlighted heterozygous variants could be the cause of the individual cases of SR. In view of the almost uniform morphology of the SR cases under investigation, an unexpectedly wide spectrum of genetic heterogeneity was observed. Similar genetic heterogeneity explaining a phenotype has previously been reported in several human disorders, such as Usher syndrome, which has been associated with identified mutations in 10 different genes (Fuster-García et al., 2021).

Our WGS approach was able to provide a molecular genetic diagnosis in approximately 50% of the SR cases. The efficiency of WGS for genetic diagnosis in cattle has not been investigated so far, but the results obtained are quite positive when compared with the efficiency of WGS-based genetic diagnosis in humans where a genetic diagnosis has been reported for 25% of the probands (Smedley et al., 2021). Furthermore, unresolved genetic diagnosis, which accounted for 11 SR cases in the present study, could be explained by 1)

limitations of the cattle genome annotation, 2) limitations of the short-read WGS approach, and/or 3) errors in read alignment and variant calling (Caspar et al., 2018), or 4) possible epi- or non-genetic factors.

Conclusions

The rather uniform pathology might suggest that SR is due to one or a few deleterious genes. However, this is not the case as WGS analysis of 14 trios and an additional nine solo cases revealed deleterious de novo variants in 12 of the SR-affected calves. Interestingly, all affected genes harbouring candidate causal variants belong to the class of haploinsufficient loss-of-function genes requiring two intact copies and are involved in embryonic and pre-weaning lethality or associated with severe malformation syndromes in humans and mice. Here we report for the first time that SR in cattle is most likely caused by independent heterozygous deleterious, probably pathogenic, variants in several genes important for proper foetal development. Therefore, independent de novo mutations present in the developing embryo may explain the sporadic occurrence of SR in cattle.

Conflict of interest statement

None of the authors has any financial or personal relationships that could inappropriately influence or bias this study.

Acknowledgements

The authors thank the Interfaculty Bioinformatics Unit of the University of Bern for providing the computational infrastructure. This study was financially supported by the Arbeitsgemeinschaft Schweizerischer Rinderzüchter (ASR), Zollikofen, Switzerland, the Federal Office for Agriculture (FOAG), Bern, Switzerland, and the Danish Bovine Genetic

Disease Program. Breeders and veterinarians in Denmark, Italy, and Germany are thanked for submission of cases.

Supplementary material

Supplementary data associated with this article can be found, in the online version, at doi: …

References

Adzhubei, I.A., Schmidt, S., Peshkin, L., Ramensky, V.E., Gerasimova, A., Bork, P., Kondrashov, A.S., Sunyaev, S.R., 2010. A method and server for predicting damaging missense mutations. Nat Methods. https://doi.org/10.1038/nmeth0410-248

Agerholm, J.S., Dahl, M., Herskin, M., Nielsen, S.S., 2023. Forensic age assessment of late-term bovine fetuses. Acta Vet Scand 65, 27. https://doi.org/10.1186/s13028-023-00691-0

Bendl, J., Stourac, J., Salanda, O., Pavelka, A., Wieben, E.D., Zendulka, J., Brezovsky, J., Damborsky, J., 2014. PredictSNP: Robust and Accurate Consensus Classifier for Prediction of Disease-Related Mutations. PLoS Comput Biol 10, 1–11. https://doi.org/10.1371/journal.pcbi.1003440

Biscarini, F.;, Cozzi, P.;, Gaspa, G.;, Marras, G., 2019. DetectRUNS: An R Package to Detect Runs of Homozygosity and Heterozygosity in Diploid Genomes. CRAN (The Comprehensive R Archive Network). [WWW Document]. URL https://CRAN.R-project.org/package=detectRUNS (accessed 7.4.23).

Caspar, S.M., Dubacher, N., Kopps, A.M., Meienberg, J., Henggeler, C., Matyas, G., 2018. Clinical sequencing: From raw data to diagnosis with lifetime value. Clin Genet 93, 508–519. https://doi.org/10.1111/cge.13190

Cavalieri, J., Farin, P., 1999. Birth of a holstein freemartin calf co-twinned to a schistosomus reflexus fetus. Theriogenology 52, 815–826. https://doi.org/10.1016/S0093-691X(99)00174-0

Chang, C.C., Chow, C.C., Tellier, L.C., Vattikuti, S., Purcell, S.M., Lee, J.J., 2015. Second-generation PLINK: rising to the challenge of larger and richer datasets. Gigascience 4, 7. https://doi.org/10.1186/s13742-015-0047-8

Chen, S., Zhou, Y., Chen, Y., Gu, J., 2018. Fastp: An ultra-fast all-in-one FASTQ preprocessor. Bioinformatics 34, i884–i890. https://doi.org/10.1093/bioinformatics/bty560

Cheng, C., Mu, J., Farkas, I., Huang, D., Goebl, M.G., Roach, P.J., 1995. Requirement of the Self-Glucosylating Initiator Proteins Glg1p and Glg2p for Glycogen Accumulation in *Saccharomyces cerevisiae*. Mol Cell Biol 15, 6632–6640. https://doi.org/10.1128/MCB.15.12.6632

Citek, J., 2012. Pedigree analysis of Czech Holstein calves with schistosoma reflexum. Acta Vet Scand 54, 22. https://doi.org/10.1186/1751-0147-54-22

Doty, R.T., Vanasse, G.J., Disteche, C.M., Willerford, D.M., 2002. The leukemia-associated gene Mllt1/ENL: characterization of a murine homolog and demonstration of an essential

role in embryonic development. Blood Cells Mol Dis 28, 407–17. https://doi.org/10.1006/bcmd.2002.0525
Fleming, 1850. Schistosomus scoliosus beim Kalb. Magazin f. d. ges. Tierheilk. 479.
Fuster-García, C., García-Bohórquez, B., Rodríguez-Muñoz, A., Aller, E., Jaijo, T., Millán, J.M., García-García, G., 2021. Usher Syndrome: Genetics of a Human Ciliopathy. Int J Mol Sci 22. https://doi.org/10.3390/ijms22136723
Gregersen, L.H., Mitter, R., Ugalde, A.P., Nojima, T., Proudfoot, N.J., Agami, R., Stewart, A., Svejstrup, J.Q., 2019. SCAF4 and SCAF8, mRNA Anti-Terminator Proteins. Cell 177, 1797-1813.e18. https://doi.org/10.1016/j.cell.2019.04.038
Hayes, B.J., Daetwyler, H.D., 2019. 1000 Bull Genomes Project to Map Simple and Complex Genetic Traits in Cattle: Applications and Outcomes. Annu Rev Anim Biosci 7, 89–102. https://doi.org/10.1146/annurev-animal-020518-115024
He, N., Liu, M., Hsu, J., Xue, Y., Chou, S., Burlingame, A., Krogan, N.J., Alber, T., Zhou, Q., 2010. HIV-1 Tat and host AFF4 recruit two transcription elongation factors into a bifunctional complex for coordinated activation of HIV-1 transcription. Mol Cell 38, 428–38. https://doi.org/10.1016/j.molcel.2010.04.013
Jacinto, J.G.P., Häfliger, I.M., Akyürek, E.E., Sacchetto, R., Benazzi, C., Gentile, A., Drögemüller, C., 2021. KCNG1-related syndromic form of congenital neuromuscular channelopathy in a crossbred calf. Genes (Basel). https://doi.org/10.3390/genes12111792
Karczewski, K.J., Francioli, L.C., Tiao, G., Cummings, B.B., Alföldi, J., Wang, Q., Collins, R.L., et. al., 2020. The mutational constraint spectrum quantified from variation in 141,456 humans. Nature 581, 434–443. https://doi.org/10.1038/s41586-020-2308-7
Knight, R.P., 1996. The occurrence of schistosomus reflexus in bovine dystocia. Aust Vet J 73, 105–107. https://doi.org/10.1111/j.1751-0813.1996.tb09988.x
Li, H., 2011. A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. Bioinformatics 27, 2987–2993. https://doi.org/10.1093/bioinformatics/btr509
Lin, C., Smith, E.R., Takahashi, H., Lai, K.C., Martin-Brown, S., Florens, L., Washburn, M.P., Conaway, J.W., Conaway, R.C., Shilatifard, A., 2010. AFF4, a component of the ELL/P-TEFb elongation complex and a shared subunit of MLL chimeras, can link transcription elongation to leukemia. Mol Cell 37, 429–37. https://doi.org/10.1016/j.molcel.2010.01.026
Marom, R., Jain, M., Burrage, L.C., Song, I.-W., Graham, B.H., Brown, C.W., Stevens, S.J.C., et al., 2017. Heterozygous variants in *ACTL6A* , encoding a component of the BAF complex, are associated with intellectual disability. Hum Mutat 38, 1365–1371. https://doi.org/10.1002/humu.23282
Martín-Alguacil, N., Avedillo, L., 2020. Body stalk anomalies in pig—Definition and classification. Mol Genet Genomic Med 8. https://doi.org/10.1002/mgg3.1227
Mateo, I., Camón, J., 2008. Schistosoma reflexum in a cat: Insights into aetiopathogenesis. J Feline Med Surg 10, 376–379. https://doi.org/10.1016/j.jfms.2007.12.010
McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M., DePristo, M.A., 2010. The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res 20, 1297–1303. https://doi.org/10.1101/gr.107524.110
Mu, J., Cheng, C., Roach, P.J., 1996. Initiation of Glycogen Synthesis in Yeast. Journal of Biological Chemistry 271, 26554–26560. https://doi.org/10.1074/jbc.271.43.26554
Ng, P.C., Henikoff, S., 2003. SIFT: Predicting amino acid changes that affect protein function. Nucleic Acids Res 31, 3812–3814. https://doi.org/10.1093/nar/gkg509
Nijman, S.M.B., Huang, T.T., Dirac, A.M.G., Brummelkamp, T.R., Kerkhoven, R.M., D'Andrea, A.D., Bernards, R., 2005. The Deubiquitinating Enzyme USP1 Regulates the

Fanconi Anemia Pathway. Mol Cell 17, 331–339. https://doi.org/10.1016/j.molcel.2005.01.008

Özsoy, Ş.Y., Oto, Ç., Haziroglu, R., 2009. Schistosoma reflexum in a dog. Ankara Üniversitesi Veteriner Fakültesi Dergisi 56, 225–226.

Park, W., Chai, H.-H., Lim, D., Dang, C., Lee, J., Kim, J., Jeong, H., Lee, T., Lee, K.-C., Lee, K., 2023. Case report: Investigation of genetic mutations in a case of schistosomus reflexus in a Holstein dairy cattle fetus in Korea. Front Vet Sci 10. https://doi.org/10.3389/fvets.2023.1238544

R Core Team, n.d. R: A language and environment for statistical computing.

Rappaport, N., Fishilevich, S., Nudel, R., Twik, M., Belinky, F., Plaschkes, I., Stein, T.I., Cohen, D., Oz-Levi, D., Safran, M., Lancet, D., 2017. Rational confederation of genes and diseases: NGS interpretation via GeneCards, MalaCards and VarElect. Biomed Eng Online 16, 72. https://doi.org/10.1186/s12938-017-0359-2

Richter, J., Götze, R., 1993. Tiergeburtshilfe. Parey, Berlin .

Roberts, S.J., 1986. Veterinary obstetrics and genital diseases (Theriogenology). Theriogenology 26, 551–552. https://doi.org/https://doi.org/10.1016/0093-691X(86)90160-3

Robinson, J.T., Thorvaldsdóttir, H., Wenger, A.M., Zehir, A., Mesirov, J.P., 2017. Variant review with the integrative genomics viewer. Cancer Res 77, e31–e34. https://doi.org/10.1158/0008-5472.CAN-17-0337

Rosen, B.D., Bickhart, D.M., Schnabel, R.D., Koren, S., Elsik, C.G., Tseng, E., Rowan, T.N., et al., 2020. De novo assembly of the cattle reference genome with single-molecule sequencing. Gigascience 9, 1–9. https://doi.org/10.1093/gigascience/giaa021

Sampson, N.D., Hewitt, J.E., 2003. SF4 and SFRS14, two related putative splicing factors on human chromosome 19p13.11. Gene 305, 91–100. https://doi.org/10.1016/S0378-1119(02)01230-1

Sivaram, M.V.S., Wadzinski, T.L., Redick, S.D., Manna, T., Doxsey, S.J., 2009. Dynein light intermediate chain 1 is required for progress through the spindle assembly checkpoint. EMBO J 28, 902–914. https://doi.org/10.1038/emboj.2009.38

Smedley, D., Smith, K.R., Martin, A., Thomas, E.A., McDonagh, E.M., Cipriani, V., Ellingford, J.M., et al., 2021. 100,000 Genomes Pilot on Rare-Disease Diagnosis in Health Care — Preliminary Report. New England Journal of Medicine 385, 1868–1880. https://doi.org/10.1056/NEJMoa2035790

Tripathy, R., Leca, I., van Dijk, T., Weiss, J., van Bon, B.W., Sergaki, M.C., Gstrein, T., et al., 2018. Mutations in MAST1 Cause Mega-Corpus-Callosum Syndrome with Cerebellar Hypoplasia and Cortical Malformations. Neuron 100, 1354-1368.e5. https://doi.org/10.1016/j.neuron.2018.10.044

Windsor, P., 2019. Abnormalities of Development and Pregnancy, in: Veterinary Reproduction and Obstetrics. Elsevier, pp. 168–194. https://doi.org/10.1016/B978-0-7020-7233-8.00009-4

Table 1

Results of variant filtering of the trio-based Group 1, assuming a dominant acting de novo mutation as the cause.

Case ID	All variants	Private variants in the SR-calf using 896 cattle genome controls	Private protein changing variants in the SR-calf using 896 cattle genome controls	Remaining protein-changing private variants using a global control cohort of 4540 cattle genomes and subsequent IGV inspection	Candidate gene
Case 1	4,487,960	16,426	1205	1096*	*FLNA*
Case 2	4,797,543	293	3	2	*ACTL6A*
Case3	4,810,478	280	1	1	*MLLT1*
Case 4	4,677,082	1,046	6	5	*MAST3*
Case 5	4,545,328	21,267	366	332*	*MBTPS2*
Case 6	4,206,833	1,115	18	6	*GLG1*
Case 7	3,843,285	954	14	2	*IQSEC2*
Case 8	4,384,747	2,083	2	1	*ISL1*
Case 9	4,305,483	421	0	0	none
Case 10	4,784,539	431	3	2	none
Case 11	4,539,422	946	4	1	none
Case 12	4,874,189	1,082	4	0	none
Case 13	4,566,295	667	1	1	none
Case 14	4,314,663	279	0	0	none

ID, identification; SR, schistosoma reflexum; IGV, Integrative Genomics Viewer; *, Considering the high number of associated genes, VarElect software was used for variant selection for phenotype-dependent gene variant prioritization according to the phenotype query "skeletal OR dysplasia" where only one was identified as a possible candidate for SR

Table 2

Results of variant filtering of the solo-based Group 2, assuming a dominant acting de novo mutation as the cause.

Case ID	All variants	Private variants in the SR-calf using 896 cattle genome controls	Private protein changing variants in the SR-calf using 896 cattle genome controls	Remaining protein-changing private variants using a global control cohort of 4540 cattle genomes and subsequent IGV inspection	Candidate gene
Case 15	5,374,628	4,048	24	9	*ANO4*
Case 16	5,464,072	18,654	136	21	*MYH1*
Case 17	4,555,767	1,678	12	4	*DYNC1LI1*
Case 18	4,740,928	2,886	26	5	*PPP2R2B*
Case 19	4,614,321	1,701	5	1	none
Case 20	4,689,266	3,544	41	4	*UBP1*
Case 21	5,331,073	25,951	147	22	*SUGP1*
Case 22	4,787,151	7,979	39	12	*SCAF8*
Case 23	4,638,557	3,560	28	6	*DYRK1B*

ID, identification; SR, schistosoma reflexum; IGV, Integrative Genomics Viewer

Table 3

List of detected de novo mutations predicted to be deleterious in seven schistosoma reflexum cases of the trio-based Group 1.

Case ID	Gene	Variant type	Variant features (genomic, cDNA, protein)	pLI score[a]
Case 1	*FLNA*	frameshift insertion	Chr X: g.37419367C>CATGTGCCG c.177_178insCGGCACAT p.Val60fs	1
Case 2	*ACTL6A*	frameshift deletion	Chr 1: g.87555312TC>T c.273delG p.Met92fs	1
Case 3	*MLLT1*	missense	Chr 7 : g.18151483C>T c.58C>T p.Arg20Cys	0.99

Case 4	*MAST3*	frameshift deletion	Chr 7 : g.5076415TG>T c.3605delC p.Pro1202fs	1
Case 5	*MBTPS2*	frameshift insertion	Chr X: g.121442341C>CCG c.817_818insCG p.Gly273fs	1
Case 6	*GLG1*	frameshift insertion	Chr 18 : g.1958877T>TGAGTCAGTCGTCGC TAAACTTAAAGAAAACGGTATTCGTG CTATTTCTGACTTGAGAAATGAGAAAA TCGGCTTTAAGATTCGTGAACGCA c.1531_1532insTGCG TTCACGAATCTTAAAGCCG ATTTTCTCATTTCTCAAGTCA GAAATAGCACGAATACCGT TTTCTTTAAGTTTAGCGACGA CTGACTC p.Glu511fs	1
Case 7	*IQSEC2*	missense	Chr X: g. 90686127A>G c.2870T>C p.Leu957Pro	1

ID, identification; cDNA, complementary DNA; pLI, probability of being intolerant; Chr, chromosome; g., genomic position; c., cDNA position; p., protein position

[a] according to GnomAD, Karczewski et al., 2020

Table 4

List of possible de novo mutations predicted to be deleterious in five schistosoma reflexum cases of the solo-based Group 2.

Case ID	Gene	Variant type	Variant features (genomic, cDNA, protein)	pLI score [a]
Case 17	*DYNC1LI1*	missense	Chr 22 : g.6974203G>A c.1513C>T p.Arg505Trp	0.96

Case 18	*PPP2R2B*	frameshift insertion	Chr 7 : g.57855214C>CG c.732dupC p.Val245fs	0.98
Case 20	*UBP1*	missense	Chr 22 : g.7654852T>C c.1162A>G p.Arg388Gly	1
Case 21	*SUGP1*	missense	Chr 7 : g. 3984900C>T c.976C>T p.Arg326Cys	1
Case 22	*SCAF8*	frameshift insertion	Chr 9 : g. 91478233G>GCTTTAGATATAGAAAGAAGCTTTGACGAAAGTCTTTACGAAATGGCTAACGCTATAAACAAATAATTTTAGTTAATGACATAATAAATTTTTTAAATTTATTATTTATTATATATTTTAAAC c.1132_1133insCTTTAGATATAGAAAGAAGCTTTGACGAAAGTCTTTACGAAATGGCTAACGCTATAAACAAATAATTTTAGTTAATGACATAATAAATTTTTTAAATTTATTATTTATTATATATTTTAAAC p.Val378fs	1

ID, identification; cDNA, complementary DNA; pLI, probability of being intolerant; Chr, chromosome; g., genomic position; c., cDNA position; p., protein position

[a] according to GnomAD, Karczewski et al., 2020

Table 5

Reported associated human or mouse phenotypes or function for 12 bovine schistosoma reflexum candidate genes.

Gene	Associated phenotype in mice	MGI	Associated phenotype in humans or gene function	OMIM
ACTL6A	Premature death, decreased hematopoietic stem cell proliferation, and abnormal definitive hematopoiesis	1861453	Intellectual disability, developmental delay, dysmorphic features, limb and, urogenital and cardiac defects resembling	604958

DYNC1LI1	Preweaning lethality, abnormal heart, testis, skin, liver, lung and sternal morphology, and microphtalmia	2135610	brachymorphism-onychodysplasia-dysphalangism[a]; torticollis, umbilical and inguinal hernia[b] Non-catalytic accessory component of the cytoplasmic dynein 1 complex involved in linking dynein to cargos and to adapter proteins that regulate dynein function; motor for the intracellular retrograde motility of vesicles and organelles along microtubules; involved in the microtubule-dependent transport of pericentrin; required for progress through the spindle assembly checkpoint[c]	615890
FLNA	Prenatal lethality, thoracoabdominoschisis, cardiovascular, craniofacial and sternal abnormalities	95556	Cardiac valvular dysplasia; congenital short bowel syndrome; frontometaphyseal dysplasia 1; heterotopia, periventricular; intestinal pseudoobstruction, neuronal; Melnick-Needles syndrome; otopalatodigital syndrome, type I and II	300017
GLG1	Neonatal lethality, abnormal caudal vertebrae and craniofacial morphology, cleft palate, decreased body size, multiple skeletal malformations (e.g., of vertebra, rib, and intervertebral disk, and spinal curvature)	2444506	Self-glucosylating initiator of glycogen synthesis; catalyzes the formation of a short alpha (1,4)-glucosyl chain covalently attached via a glucose 1-O-tyrosyl linkage to internal tyrosine residues	600753
IQSEC2	Premature death, growth retardation, convulsive seizures, abnormal brain interneuron morphology, astrocytosis	3528396	Microcephaly, brachycephaly, strabismus, hypotonia, and intellectual disability	309530
MAST3	Decreased viability ratio, abnormal eyes and ears	2683541	Developmental and epileptic encephalopathy; mega-corpus-callosum syndrome with microcephaly, cerebellar hypoplasia and cortical malformation and short stature [d,e]	612258
MBTPS2	Ectodermal dysplasia, brain, skeletal, ear/eye, and renal	2444506	Olmsted syndrome; ichthyosis follicularis, atrichia, and photophobia	300294

	abnormalities, cleft palate, and cryptorchidism		with or without brain anomalies, retardation, ectodermal dysplasia, skeletal malformations, Hirschsprung disease, ear/eye anomalies, cleft palate/cryptorchidism, and kidney dysplasia/hypoplasia; osteogenesis imperfecta, type XIX	
MLLT1	Embryonic lethality	1927238	Chromatin reader component of the super elongation complex that is necessary to increase the catalytic rate of RNA polymerase II transcription by suppressing transient pausing by the polymerase at multiple sites along the DNA and plays an essential role in embryonic development [f]	159556
PPP2R2B	Spinocerebellar ataxia and abnormal cholesterol homeostasis	1920180	Spinocerebellar ataxia 12	604325
SCAF8	Anti-terminator protein required to prevent early mRNA termination during transcription	1925212	Anti-terminator protein required to prevent early mRNA termination during transcription; associates with the phosphorylated C-terminal heptapeptide repeat domain of the largest RNA polymerase II subunit (POLR2A), and binds nascent RNA upstream of early polyadenylation sites to prevent premature mRNA transcript cleavage and polyadenylation [g]	616024
SUGP1	Embryonic growth retardation, preweaning lethality, abnormal ovaries, uterus, kidneys, spleen and liver	1917866	Functions in pre-mRNA splicing mechanisms [h]	607992
UBP1	Embryonic growth retardation, preweaning lethality, abnormal angiogenesis and heart morphology, abnormal visceral yolk sac morphology and placental labyrinth vasculature morphology	104889	Negative regulator of DNA damage repair which specifically deubiquitinates monoubiquitinated FANCD2[i]	609784

MGI, Mouse Genome Informatics; OMIM, Online Mendelian Inheritance in Man

[a] Marom et al., 2017

[b] Rappaport et al., 2017

[c] Sivaram et al., 2009

[d] Tripathy et al., 2018

[e] Cheng et al., 1995; Mu et al., 1996

[f] Doty et al., 2002; He et al., 2010; Lin et al., 2010

[g] Gregersen et al., 2019

[h] Sampson and Hewitt, 2003

[i] Nijman et al., 2005

Figure legends

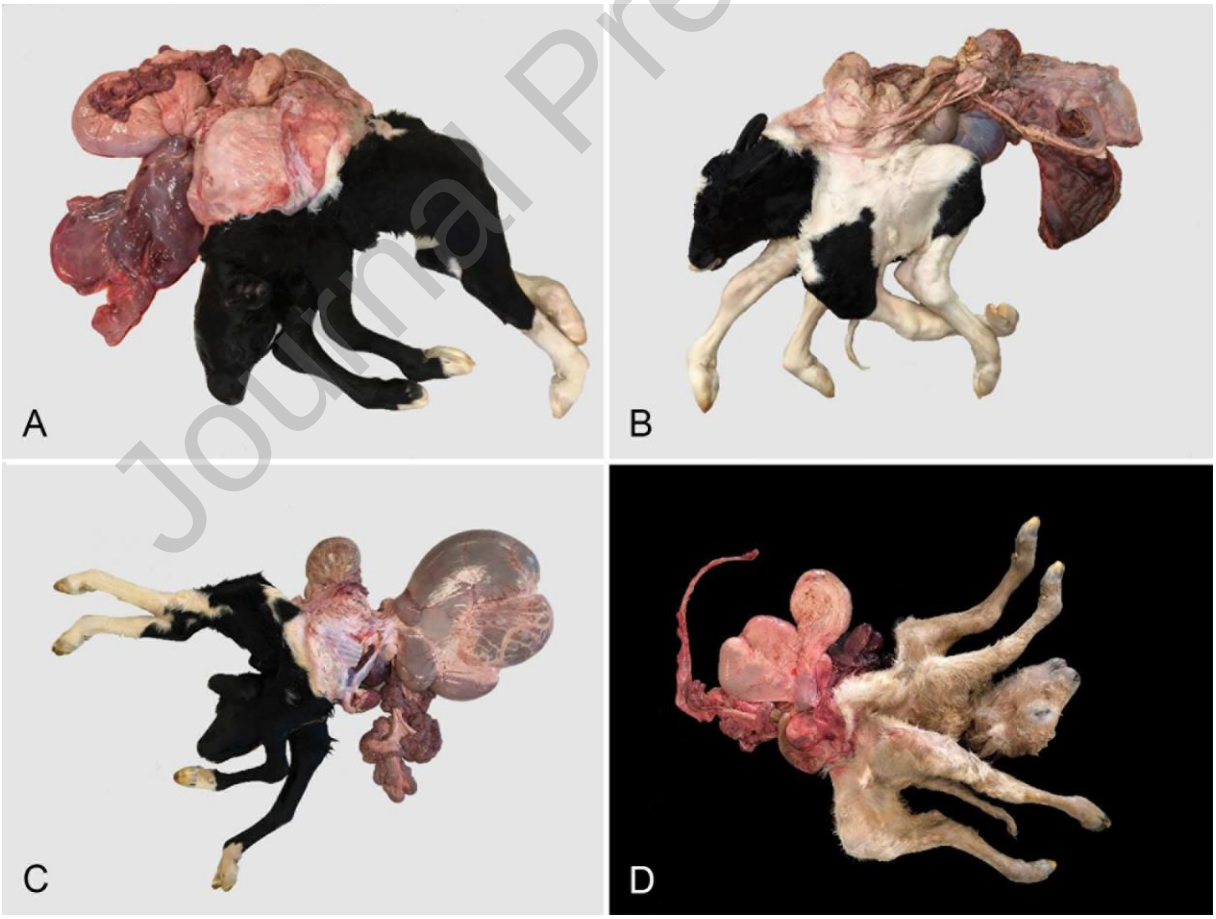

Fig.1. Gross morphology of four schistosoma reflexum cases.

Although being associated with different mutations they share a general appearance characterized by U-shaped dorsal to dorsolateral retroflexion of the spine and exposure of the viscera. A) Case 4, Holstein; B) Case 11, Holstein; C) Case 3, Holstein; D) Case 16, Limousin.

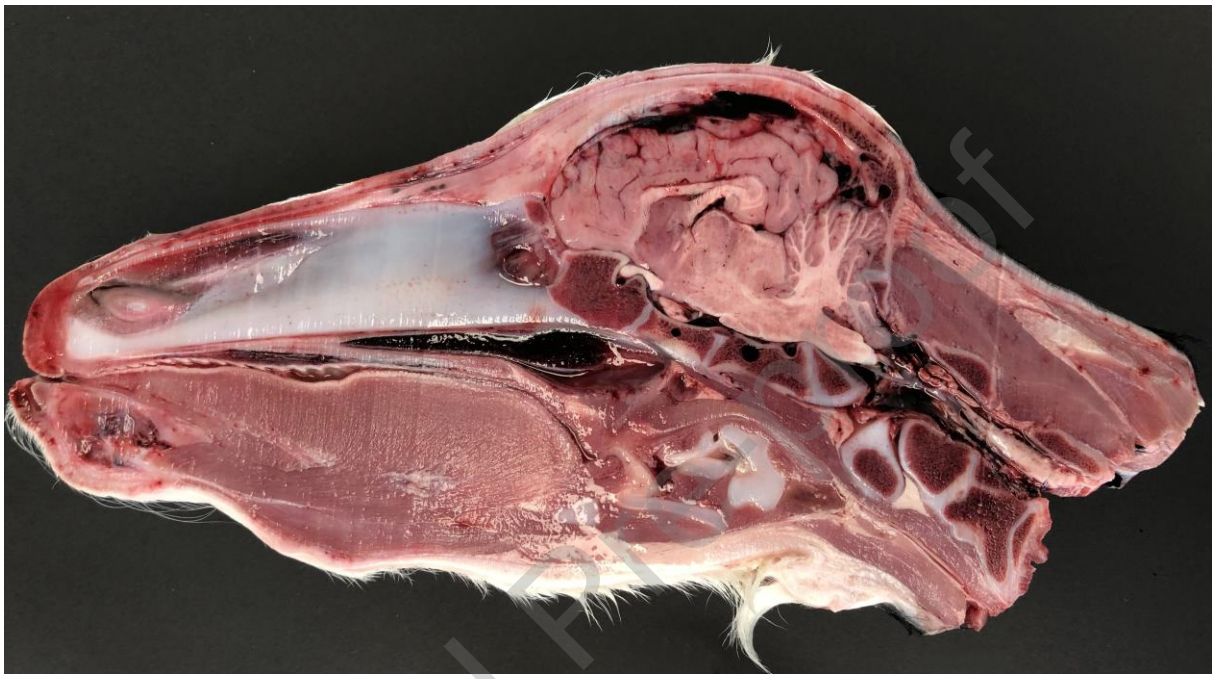

Fig. 2. Caudal displacement and elongation of the caudal aspect of the cerebellum (similar to Arnold-Chiari type 1 malformation) (arrows) in a Holstein calf (case 10). Median sagittal section through the head of a case of bovine schistosoma reflexum.

Highlights

- Schistosoma reflexum (SR), a lethal syndrome, was associated with dystocia in cattle
- Multiple independent de novo mutations were associated with the development of SR
- Genomic analysis showed that SR is not due to a single recessive mutation
- We highlighted the emerging potential of genomic precision diagnostics in livestock
- We provided insights into the function SR-related genes for translational medicine