UNIVERSIDADE DO ESTADO DE SANTA CATARINA – UDESC
CENTRO DE EDUCAÇÃO SUPERIOR DO OESTE - CEO
PROGRAMA DE PÓS-GRADUAÇÃO EM ZOOTECNIA

DISSERTAÇÃO DE MESTRADO

IDENTIFICAÇÃO DE GENES E POLIMORFISMOS ASSOCIADOS COM A MANIFESTAÇÃO DE HÉRNIA UMBILICAL EM SUÍNOS UTILIZANDO SEQUENCIAMENTO DE NOVA GERAÇÃO

IGOR RICARDO SAVOLDI

CHAPECÓ, 2020.

IGOR RICARDO SAVOLDI

IDENTIFICAÇÃO DE GENES E POLIMORFISMOS ASSOCIADOS COM A MANIFESTAÇÃO DE HÉRNIA UMBILICAL EM SUÍNOS UTILIZANDO SEQUENCIAMENTO DE NOVA GERAÇÃO

Dissertação apresentada ao Curso de Mestrado do Programa de Pós-Graduação em Zootecnia, Área de Concentração Ciência e Produção Animal, da Universidade do Estado de Santa Catarina (UDESC), como requisito parcial para obtenção de grau de **Mestre em Zootecnia**

Orientadora: Mônica Corrêa Ledur
Coorientadora: Adriana Mércia Guaratini Ibelli
Coorientador: Diego de Córdova Cucco

Chapecó, SC, Brasil
2020

Ficha catalográfica elaborada pelo programa de geração automática da

Biblioteca Setorial do CEO/UDESC,

com os dados fornecidos pelo(a) autor(a)

Savoldi, Igor Ricardo
IDENTIFICAÇÃO DE GENES E POLIMORFISMOS ASSOCIADOS COM A MANIFESTAÇÃO DE HÉRNIA UMBILICAL EM SUÍNOS UTILIZANDO SEQUENCIAMENTO DE NOVA GERAÇÃO / Igor Ricardo Savoldi. -- 2020.
134 p.

Orientadora: Mônica Corrêa Ledur
Coorientadora: Adriana Mércia Guaratini Ibelli
Coorientador: Diego de Córdova Cucco
Dissertação (mestrado) -- Universidade do Estado de Santa Catarina, Centro de Educação Superior do Oeste, Programa de Pós-Graduação em Zootecnia, Chapecó, 2020.

1. Genes. 2. defeitos congênitos. 3. transcriptoma. 4. exoma. 5. GWAS. I. Ledur, Mônica Corrêa . II. Ibelli, Adriana Mércia Guaratini. Cucco, Diego de Córdova. III. Universidade do Estado de Santa Catarina, Centro de Educação Superior do Oeste, Programa de Pós-Graduação em Zootecnia. IV. Titulo.

Universidade do Estado de Santa Catarina
UDESC Oeste
Programa de Pós-Graduação em Zootecnia

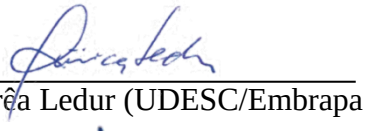
A Comissão Examinadora, abaixo assinada, aprova
a Dissertação de Mestrado

IDENTIFICAÇÃO DE GENES E POLIMORFISMOS ASSOCIADOS COM A MANIFESTAÇÃO DE HÉRNIA UMBILICAL EM SUÍNOS UTILIZANDO SEQUENCIAMENTO DE NOVA GERAÇÃO

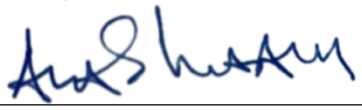
Elaborada por
IGOR RICARDO SAVOLDI

como requisito parcial para obtenção do grau de
Mestre em Zootecnia

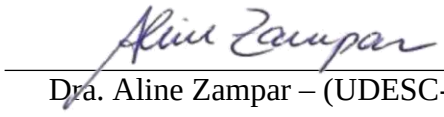
Comissão Examinadora:



Dra. Mônica Corrêa Ledur (UDESC/Embrapa Suínos e Aves)



Dra. Ana Silvia Alves Meira Tavares Moura (FMVZ/Unesp)



Dra. Aline Zampar – (UDESC-Oeste)

Chapecó, 30 de Julho de 2020.

AGRADECIMENTOS

Primeiro agradeço a Deus, pela força, pela persistência e principalmente pela oportunidade de sempre continuar crescendo e evoluindo como profissional.

Inicio os agradecimentos à minha família e aos meus pais, que sempre primaram pela minha educação. Obrigado Ricardo J. Savoldi e Solange C. M. Savoldi por, além de me oferecerem a oportunidade de estudar, sempre estarem presentes e sou muito feliz por isso.

Agradeço a Universidade do Estado de Santa Catarina (UDESC-Oeste) e ao programa de Pós-graduação que proporcionou totais condições para o desenvolvimento da dissertação de mestrado.

Agradeço também a todos os professores que tive a oportunidade de conhecer durante o mestrado. Em especial aos professores Diego Cucco e Aline Zampar que me ensinaram muito mais que conteúdos em sala de aula, me mostraram que o caráter e a dedicação com a pesquisa são essenciais para que possamos crescer cada dia mais. Vocês dois são ótimos profissionais e excelentes professores!!

E claro, aos meus colegas de turma do mestrado. Colegas especiais! Sou grato por ter conhecido vocês, agradeço pela recepção de vocês, pelas risadas e pela convivência tão maravilhosa, vocês me ajudaram muito na UDESC e por isso espero que nossa amizade siga em frente além da UDESC. Muito obrigado!

Agradeço as agências de fomento CNPq e CAPES que financiaram esta pesquisa, e agradeço ao Programa de Monitoria da Pós-graduação (PROMOP) da UDESC que me contemplou com uma bolsa de estudos durante o mestrado. Gostaria de dizer que as agências de fomento e as bolsas de estudos são extremamente importantes para a ciência brasileira e sem elas, este projeto e muitos outros não poderiam ser realizados. Portanto sou grato e friso a importância das agências financiadoras.

Agradeço a Embrapa Suínos e Aves pela oportunidade e disponibilidade de estar desenvolvendo a dissertação de mestrado, local onde em 2015 pude ingressar realizando meu estágio de IC e então comecei a despertar meu interesse pela pesquisa. Agradeço a todos os profissionais que lá estão que de alguma forma me ajudaram muito nesse período.

Gostaria de agradecer todos os colegas do laboratório de Genética da Embrapa, pois quando iniciei meu estágio em 2015 todos me acolheram muito bem e foi onde conheci

muitas pessoas importantes para mim, que se tornaram grandes amigos que tem a minha admiração.

São tantos os colegas que eu conheci que me ajudaram (e que ainda ajudam) nesta longa jornada até o mestrado, quero agradecer a cada um deles que eu pude conhecer de diversos lugares onde eu pude aprender muito com cada um de vocês.

Agradeço a Raíra e a Kamila que foram minhas colegas de estágio durante bastante tempo, Mayla, Willian, Haniel, Essamai, Neide e Jéssica colegas e amigos que me ajudaram e me ensinaram muito. Carol, uma pessoa incrível, uma amiga especial, mais recentemente Débora, Fernanda e Ágata que pude conhecer e que me ajudaram muito, além das nossas conversas e risadas diárias, eu admiro muito vocês e quero que a nossa amizade construída seja longa.

Quero agradecer o Maurício, Alexandre e Michele por terem me ajudado, vocês me ensinaram muitas coisas durante o mestrado são pessoas que eu guardo com muito carinho.

Gostaria de agradecer as minhas orientadoras Dra. Adriana, Dra. Jane e Dra. Mônica. Sou grato por tudo, eu tenho muita admiração por vocês, agradeço pela oportunidade de fazer parte da equipe de vocês, uma equipe ótima, sou grato por tudo o que eu aprendi com vocês. Agradeço todo o incentivo, paciência e compreensão que vocês tiveram, eu tenho muito carinho por vocês, um profundo respeito e admiração!! Tenho muito orgulho de ter feito parte dessa equipe. Gratidão eterna a vocês!

E a todos o meu MUITO OBRIGADO!

RESUMO

Dissertação de Mestrado
Programa de Pós-Graduação em Zootecnia
Universidade do Estado de Santa Catarina

IDENTIFICAÇÃO DE GENES E POLIMORFISMOS ASSOCIADOS COM A MANIFESTAÇÃO DE HÉRNIA UMBILICAL EM SUÍNOS UTILIZANDO SEQUENCIAMENTO DE NOVA GERAÇÃO

AUTOR: Igor Ricardo Savoldi
ORIENTADORA: Mônica Corrêa Ledur
Chapecó, 30 de Julho de 2020

A hérnia umbilical (HU) é uma condição que afeta a produção de suínos, reduzindo o bem-estar e o desempenho dos animais. As potenciais causas que podem levar à HU são os distúrbios associados ao metabolismo do colágeno, estrutura muscular e reparo do tecido conjuntivo, levando à uma maior fragilidade e flacidez da região umbilical, fazendo com que a abertura umbilical não feche corretamente e que os intestinos projetem-se pela parede abdominal formando o saco herniário. Contudo, os componentes genéticos envolvidos com a UH são pouco compreendidos. Portanto, o objetivo deste estudo foi identificar genes e polimorfismos associados com a manifestação de hérnia umbilical em suínos utilizando sequenciamento de nova geração. Para isso, 10 leitoas com 90 dias de idade, 5 herniadas e 5 não-herniadas foram selecionadas. A extração de DNA foi realizada com o Purelink Genomic Mini Kit (Invitrogen), as bibliotecas exômicas foram preparadas usando o kit SeqCap EZ Library SR (1.0) (NimbleGen / Roche) e o sequenciamento foi realizado no equipamento Illumina HiSeq 2500 (2x100pb). Além disso, também foram utilizados dados de transcriptoma do anel umbilical disponíveis no banco de dados SRA para identificação de variantes. As leituras exômicas e transcriptômicas foram submetidas ao controle de qualidade utilizando a ferramenta Trimmomatic e mapeadas contra o genoma suíno de referência (Sscrofa11.1) utilizando o software BWA-MEM para o sequenciamento exômico e o STAR para o transcriptômico. A identificação dos polimorfismos foi realizada usando o Genome Analysis Toolkit (GATK) e o Variant Effect Predictor (VEP) do Ensembl foi usado para anotar e predizer o efeito das variantes. Essas variantes foram também comparadas com resultados de um estudo de associação genômica (GWAS). No exoma foram identificadas 209 variantes, 5 localizadas em regiões intergênicas e 204 localizadas em 72 genes. No transcriptoma foram identificadas 81 variantes localizadas em 42 genes. Dessas, 29 variantes do tipo SNP (Polimorfismo de Nucleotídeo Único) foram classificadas através do VEP como variantes *missense*. Comparando as três metodologias, obtivemos 79 variantes concordantes entre o exoma e o transcriptoma, localizadas em 23 genes. Além disso, 8 genes também foram identificados pela análise de GWAS. Portanto, destacamos as variantes nos genes *FYN, SPN, ITGB2, MYLPF, MYH13, MYH8, MYH2, MYH3, MYO19, VIM, ROCK2, RABGEF1* e *UBASH3B* como fortes candidatas no desenvolvimento de UH em suínos. Destaca-se também a importância da combinação das três abordagens para identificar genes e polimorfismos potencialmente envolvidos no desencadeamento da UH.

Palavras-chave: Genes, defeitos congênitos, transcriptoma, exoma, GWAS, suínos.

ABSTRACT

Master's Dissertation
Programa de Pós-Graduação em Zootecnia
Universidade do Estado de Santa Catarina

IDENTIFICATION OF GENES AND POLYMORPHISMS ASSOCIATED WITH THE MANIFESTATION OF UMBILICAL HERNIA IN PIGS USING NEXT GENERATION SEQUENCING

AUTHOR: Igor Ricardo Savoldi
ADVISER: Mônica Corrêa Ledur
Chapecó, 30 July 2020

Umbilical hernia (UH) is a condition that affects pig production, which reduces welfare and performance. The potential causes that can lead to UH are disorders associated with collagen metabolism, muscle structure and connective tissue repair, leading to a weakness and flabbiness in the umbilical region. In consequence, umbilical opening do not close properly and the intestines project through the abdominal wall, forming the herniary sac. However, the genetic components involved with UH are poorly understood. Therefore, the aim of this study was to identify genes and polymorphisms associated with the manifestation of umbilical hernia in pigs using next generation sequencing. Therefore, 10 gilts with 90 days of age, 5 herniated and 5 non-herniated were selected. DNA extraction was performed with the Purelink Genomic Mini Kit (Invitrogen), the exomic libraries were prepared using the SeqCap EZ Library SR (1.0) kit (NimbleGen / Roche) and the sequencing was performed in the Illumina HiSeq 2500 (2x100pb) equipment. Moreover, transcriptome data available in the SRA database were also used to identify variants. The exomic and transcriptomic sequencies were subjected to quality control using the Trimmomatic tool and mapped against the swine reference genome (Sscrofa11.1) using the BWA-MEM software for the exomic and the STAR for the transcriptomic sequencing. Polymorphisms were identified using the Genome Analysis Toolkit (GATK). The Ensembl's Variant Effect Predictor (VEP) was used to annotate and predict the effect of the variants. These variants were also compared with results from a genome-wide association study (GWAS). In the exome, 209 variants were found, 5 located in intergenic regions and 204 located in 72 genes. In the transcriptome, 81 variants were identified, which were located in 42 genes. From those, 29 single nucleotide polymorphisms (SNP) were classified with the VEP tool as missense variants. Comparing the three methodologies, we obtained 79 concordant variants between the exome and the transcriptome, located in 23 genes. Moreover, 8 genes were also identified by GWAS analysis. Therefore, we highlight the variants in the *FYN, SPN*, *ITGB2*, *MYLPF*, *MYH13*, *MYH8*, *MYH2*, *MYH3*, *MYO19*, *VIM*, *ROCK2*, *RABGEF1* and *UBASH3B* genes as strong candidates in the development of UH in pigs. We also highlight the importance of combining the three approaches to identify genes and polymorphisms potentially involved with the onset of UH.
Keywords: Genes, congenital defects, transcriptomics, exome, GWAS, swine.

SUMÁRIO

1 CAPÍTULO I

REVISÃO DE LITERATURA

1.1 SUINOCULTURA

A suinocultura exerce um papel de destaque mundial na produção de carne devido ao investimento tecnológico que foi implementado nos modelos de sistemas de criação dos suínos. Este avanço proporcionou um crescimento de 45,7% na produção mundial de carne suína nos últimos 24 anos, passando de 78,2 milhões de toneladas em 1995 para 116,4 milhões de toneladas de carne em 2019 (ABCS, 2020). O continente asiático detém a maior produção de carne suína do mundo (66,2 milhões de toneladas) e a Europa é o segundo maior produtor de suínos com 28,4 milhões de toneladas, seguida das Américas, com 22,2 milhões de toneladas (ABPA, 2020).

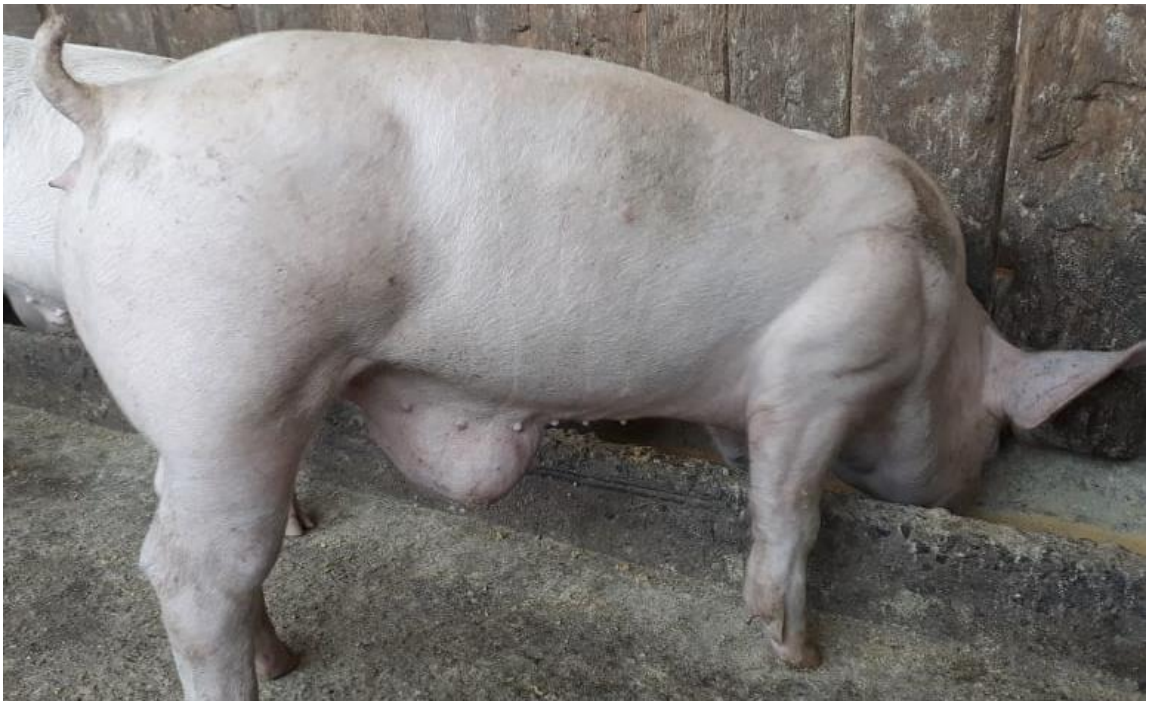
A suinocultura brasileira passou pelo processo de intensificação e inovação tecnológica no decorrer dos últimos anos, de modo que tal avanço proporcionou um aumento significativo na criação e produção de suínos. Em 2006, a produção de suínos era de 2,94 milhões de toneladas, aumentando para 3,75 milhões de toneladas em 2017 (USDA, 2018). Portanto, em 11 anos houve um crescimento significativo na produção, avanço que exemplifica claramente a evolução tecnológica do setor nas áreas de genética, nutrição, sanidade e manejo. Desta forma, os suínos puderam ser abatidos com maior peso, menor quantidade de gordura e melhor conversão alimentar (ABCS, 2018). Este avanço propiciou ao Brasil um lugar de destaque na suinocultura mundial, ocupando a quarta posição de maior produtor, com aproximadamente 3,98 milhões de toneladas, sendo 19% desta destinada à exportação, conferindo ao país a posição de quarto maior exportador mundial. Os outros 81% da produção são destinados ao mercado interno (ABPA, 2020).

Contudo, um dos maiores desafios impostos na suinocultura moderna é a melhoria de alguns aspectos produtivos para torná-la mais eficiente, como: oferta e procura de insumos, ausência de investimento, abandono da atividade, perdas produtivas, sanidade e outros. Dentre eles, os problemas congênitos, como as hérnias escrotais e umbilicais, estão entre os principais responsáveis pelas perdas produtivas (PETERSEN *et al.*, 2008).

1.2 HÉRNIA UMBILICAL

A hérnia umbilical (Figura 1) é caracterizada pela migração de parte das vísceras abdominais do intestino através do anel umbilical, formando um acúmulo de volume na área ventral do abdômen (JAY GROSFELD *et al.*, 2006). Isso ocorre devido ao enfraquecimento dos músculos ao redor do umbigo, fazendo com que a abertura umbilical não feche corretamente e que os intestinos se projetem pela parede abdominal (STRAW; BATES; MAY, 2009), formando o saco herniário. Além disso, distúrbios associados ao metabolismo do colágeno e a estrutura muscular lisa podem comprometer a reparação de tecido conectivo pelos fibroblastos (BENDAVID, 2004) e a flacidez nessa região inguinal pode evoluir para hérnia (BEUERMANN *et al.*, 2009; FRANZ *et al.*, 2001).

Figura 1: Leitoa comercial com aproximadamente 90 dias de idade acometida pela hérnia umbilical.



Fonte: Arquivo Pessoal.

As hérnias afetam negativamente a conversão alimentar e ganho de peso, aumentando a contaminação de carcaça devido à ruptura do intestino, podendo ocasionar a morte dos suínos. Desta forma, a ocorrência das hérnias reduz o bem-estar animal (GRINDFLEK *et*

al., 2006) e causa perdas econômicas significativas. Straw, Bates e May (2009) observaram correlação entre a mortalidade e o tamanho do saco herniário. Para avaliação do tamanho das hérnias sugeriu-se uma escala de escores entre pequeno, médio e grande. As taxas de mortalidade foram 4,0%, 3,1% e 8,3%, respectivamente (STRAW; BATES; MAY, 2009). Leitões acometidos por hérnias umbilicais apresentam menor ganho de peso antes do desmame em relação a leitões normais (SEARCY-BERNAL, GARDER e HIRD, 1994). No frigorífico, as carcaças com hérnia têm menor valorização do que as demais (STRAW; BATES; MAY, 2009), sendo esta desvalorização muito significativa para as indústrias de carne. No Brasil, Zanchin (2015) avaliou 235 suínos comerciais suspeitos de apresentar HU no frigorífico, sendo que 153 carcaças foram confirmadas com a presença de HU. Destas, 100 (65.4%) foram classificadas como não exportáveis (destinadas ao mercado interno), 48 (31,4%) foram destinadas para tratamento térmico que inclui conservas e embutidos cozidos e 5 (3,3%) foram para a graxaria (farinha de carne), evidenciando, dessa forma, grandes perdas produtivas e econômicas tanto para a indústria como para produtor.

Na Dinamarca, verificou-se que a ocorrência de hérnia umbilical foi de aproximadamente 6,6%, mostrando a alta incidência desta condição em suínos destinados à produção (PETERSEN *et al.*, 2008). Em plantéis de suínos em Ontário (Canadá), foram observadas incidências variando entre 0,4% e 1,5% (KEENSLISIDE, 2006), refletindo o grande impacto que esse problema causa na suinocultura. É importante ressaltar que no Brasil apenas um estudo foi identificado relatando a incidência de hérnia umbilical nos rebanhos suínos. Poltronieri *et al*. (2015), avaliando a incidência de hérnias escrotais/inguinais e umbilicais durante o desmame de leitões aos 21 dias em uma granja de suínos comerciais, verificaram que do total de 4.000 leitões nascidos no período avaliado, 60 apresentaram hérnias escrotais/inguinais e 20 apresentaram hérnias umbilicais, representando, respectivamente, 1.5% e 0.5% do total de leitões nascidos. No Brasil, 32 milhões de suínos são abatidos anualmente. Se considerarmos que destes, 0,5% forem acometidos com hérnia umbilical, estima-se que 160 mil suínos são comercializados com HU por ano. Leitões herniados mantidos na granja têm um valor de mercado 50% menor que animais sadios. Considerando a suinocultura brasileira, isto geraria um prejuízo superior a R$ 20 milhões de reais por ano, sem considerar os possíveis óbitos. Essas

estimativas não consideram ainda as perdas econômicas no desempenho e as causadas por problemas de bem-estar animal.

Sevillano *et al.*, (2015) estimaram a ocorrência da hérnia inguinal/escrotal em duas populações de suínos das raças Large White (LW) e Landrace (LD), observando que a raça LW possui um maior índice de ocorrência (0,42%) em comparação com a raça LD (0,34%). Para a hérnia umbilical, estima-se que a prevalência desta anomalia em linhas puras esteja entre 1,2 e 2,1%, conforme relatado para as raças Duroc e Large White, respectivamente (SEARCY-BERNAL, GARDER e HIRD, 1994).

1.3 FATORES PREDISPONENTES À HERNIAÇÃO

Para compreender os mecanismos biológicos e genéticos do desenvolvimento da hérnia umbilical, grande parte dos estudos é realizada com a espécie humana, extrapolando-se, portanto, as descobertas descritas em humanos para as outras espécies animais semelhantes como, por exemplo, os suínos (MEIER *et al.*, 2001). De uma forma geral, a hérnia umbilical está associada a fatores genéticos e ambientais, podendo, portanto, ser classificada em congênita ou adquirida (COATS; HELIKSON; BURD, 2000). A adquirida é desencadeada após o parto em consequência da fraqueza na cicatrização do umbigo dos recém-nascidos ou também por diversos fatores como trauma, lesão e infecções do umbigo (PETERSEN *et al*., 2008). A congênita é observada durante a fase embrionária devido a um defeito no desenvolvimento dos músculos na parede abdominal ocasionando a protrusão de parte do intestino para fora do abdômen (BRANDT, 2008). Este tipo de falha pode estar associado a fatores genéticos (FALL *et al.*, 2006; FRANZ, 2008).

Brandt (2008) define que todos os recém-nascidos apresentam um pequeno defeito no umbigo ao nascimento, através do qual os vasos umbilicais atravessam (BRANDT, 2008; FALL *et al.*, 2006). Posteriormente, o fechamento do anel umbilical é espontâneo. Sendo assim, este tipo de processo é o único que é programado geneticamente para fechar (FALL *et al.*, 2006). Contudo, a base molecular para o fechamento do anel umbilical é pouco conhecida, porém, existem evidências de que mecanismos genéticos afetam o fechamento do anel umbilical (COATS; HELIKSON; BURD, 2000; MEREI, 2006). Em humanos, crianças afro-americanas e africanas apresentam uma incidência muito maior de

hérnias umbilicais (MEIER *et al.*, 2001). Além disso, alguns distúrbios já foram associados com o aumento da ocorrência de hérnia umbilical em crianças, incluindo a síndrome do gigantismo (Beckwith Wiedemann - SBW) e a síndrome de Down (BRANDT, 2008; FALL *et al.*, 2006).

Geralmente, as hérnias abdominais ocorrem devido a distúrbios dos tecidos ao redor do umbigo (FRANZ *et al.*, 2001), incluindo perda funcional muscular de sustentação, falha ao fechamento do anel umbilical através de uma falha no tecido colagenoso ou por falha da cicatrização do tecido ao passar por um procedimento cirúrgico (BRANDT, 2008). O metabolismo anormal do colágeno já foi diretamente associado com o desenvolvimento da hérnia, sendo que moléculas imaturas de colágeno foram relacionadas ao processo de herniação (HUTSON *et al.*, 2015), indicando uma base genética para a formação da hérnia.

A hidroxilação desordenada das fibras do colágeno foi associada ao aumento da atividade proteolítica desse tecido causando fraqueza estrutural (FALL *et al.*, 2006). Um gene que está diretamente associado ao colágeno é o gene da *matrix metaloprotease* (*MMP)* que atua diretamente na formação dos fibroblastos e sua superexpressão já foi associada ao desenvolvimento de hérnia inguinal em humanos (BELLÓN *et al.*, 2001). Em pacientes com hérnia inguinal, os níveis de expressão dos genes *MMP13* e *MMP2* estavam elevados indicando que a alta expressão da matrix metaloprotease está relacionada com o desenvolvimento da hérnia inguinal (BELLÓN *et al.*, 2001; ZHENG *et al.*, 2002). Além disso, a diminuição nos níveis de RNAm do colágeno tipo I e III na pele de pacientes com hérnia incisional mostrou uma baixa presença deste tipo celular no saco herniário indicando uma visível falha na cicatrização deste tecido no local da hérnia (BELLÓN *et al.*, 2001; FRANZ, 2008).

Processos inflamatórios também já foram associados com o desenvolvimento das hérnias, onde há o envolvimento de várias células pró-inflamatórias e células de defesa como macrófagos e linfócitos (KATSUMI *et al.*, 2005). Quando um tecido está lesionado ou mais susceptível ao dano, este perde sua força de sustentação no que corresponde a fase de inflamação (BRANDT, 2008; FRANZ, 2008). Portanto, uma fase prolongada de inflamação é vista pelo organismo como um componente estranho o qual inibe a cicatrização do tecido bloqueando a síntese de colágeno, a contração tecidual e a

cicatrização do local, tornando-o propício para o desenvolvimento de hérnias (FRANZ, 2008).

1.4 COMPONENTES GENÉTICOS ASSOCIADOS ÀS HÉRNIAS

Sabe-se que a ocorrência de hérnias em suínos está sob forte influência do ambiente e do manejo dos animais, além de estar associada com a linhagem (MANALAYSAY *et al.*, 2017). Estimativas de herdabilidade para hérnia escrotal/inguinal de 0,31 (SEVILLANO *et al*., 2015) e para hérnia umbilical de 0,25 (SEARCY-BERNAL, GARDNER E HIRD, 1994; THALLER; DEMPFLE; HOESCHELE, 1996) foram observadas, evidenciando a influência genética no surgimento destes defeitos.

Muitos estudos na área genômica são voltados para a identificação de regiões de QTL (*Quantitative trait loci*), que são caracterizadas como regiões possivelmente responsáveis pela expressão de características fenotípicas de interesse (LI *et al.*, 2019).

As hérnias inguinais são provocadas pela fraqueza da região inguinal e abdominal, atrelados a falhas na formação dos tecidos no período embrionário (BRANDT, 2008). Além disso, distúrbios ligados aos fibroblastos poderiam comprometer a reparação do tecido conectivo (FRANZ, 2008) comprometendo a região inguinal, podendo se desenvolver para hérnia escrotal (BEUERMANN *et al.*, 2009). Grindeflek *et al.* (2006), utilizando a metodologia de análise de ligação, identificaram regiões genômicas associadas com a formação da hérnia inguinal nos cromossomos SSC 3, 6, 7, 12 e 15 dos suínos. Nessas regiões foram localizados alguns genes como o *COL9A1, ESR1, BAX, HOXB9* e o *HOXB5*, que podem ser importantes para o desenvolvimento desta anomalia em suínos. Segundo Gatphayak *et al.* (2007), polimorfismos encontrados nos genes *INSL3* e *BAX* dos suínos podem estar diretamente associados ao desenvolvimento da hérnia inguinal/escrotal, uma vez que estes dois genes são importantes para a formação e diferenciação do gubernáculo.

Já a hérnia umbilical é diagnosticada pela protrusão do conteúdo abdominal formando o saco herniário na região umbilical. Geralmente, nesses casos, as musculaturas ao redor do umbigo estão enfraquecidas, permitindo que a abertura umbilical não feche adequadamente e propiciando a formação da hérnia umbilical (SEARCY-BERNAL, GARDNER E HIRD, 1994). No entanto, a manifestação clínica da doença é variável e a confirmação da anomalia logo nos primeiros dias de vida do suíno pode ser confusa, pois

geralmente aparece entre 9 e 14 semanas de idade (PETERSEN *et al.*, 2008). Com isso, há uma maior necessidade de elaboração de estudos aplicados ao melhoramento genético e formas de redução do problema. Ding *et al.* (2009) identificaram 11 regiões associadas a hérnia umbilical nos cromossomos SSC1, 2, 3, 6, 7, 8, 10 e 11, indicando o *locus SWR1928* no SSC7 e o *SW830* no SSC10 como os microssatélites mais relacionados a esta anomalia. Atualmente, 55 QTLs relacionados à hérnia umbilical já foram identificados e inseridos na base de dados QTLdb (https://www.animalgenome.org/cgi-bin/QTLdb/SS/traitmap?trait_ID=596) (SOUZA et al., 2020), auxiliando para o melhor entendimento dos mecanismos genéticos que controlam a manifestação desse defeito nos suínos.

É indiscutível que os SNPs são a maior fonte de variação genética dos indivíduos em termos de números (LANDER, 2011). No entanto, a descoberta das CNVs (do inglês, copy number variation) abriu uma nova possibilidade para entendimento da variação fenotípica, evolução e susceptibilidade às doenças. CNVs são definidas como alterações genômicas com pelo menos 1 kb de extensão, que envolve ganhos ou perdas genéticas em comparação com um genoma de referência (IAFRATE *et al.*, 2004; LEVY *et al.*, 2007; SCHERER *et al.*, 2007; SEBAT *et al.*, 2004). Para o genoma humano, as CNVs compreendem aproximadamente 12% do genoma do indivíduo (REDON *et al.*, 2006). Numericamente, a estimativa da variação individual humana pode chegar a mais de 350 milhões de pares de base, colocando definitivamente as CNVs como uma importante fonte da variação genética (REDON *et al.*, 2006; WOODWARK; BATEMAN, 2011).

Novas metodologias como RNA-seq (ROMANO *et al.*, 2020), sequenciamento do exoma (GUIATTI *et al.*, 2015), genotipagem e GWAS (*Genome-wide association studies*) (SEVILLANO *et al.*, 2015) são ferramentas importantes que cada vez mais estão sendo utilizadas nas pesquisas voltadas para a produção animal. Além disso, estas metodologias inovadoras também estão sendo implementadas nos programas de melhoramento genético por meio da seleção genômica dos suínos. Contudo, poucos trabalhos vêm sendo desenvolvidos com o intuito de elucidar melhor quais são os componentes genéticos associados com o desenvolvimento da hérnia umbilical em suínos e compreender suas vias de atuação em relação a esse problema.

1.5 METODOLOGIAS GENÔMICAS E APLICAÇÃO DO SEQUENCIAMENTO DE NOVA GERAÇÃO

O sequenciamento do DNA ou a determinação da sequência das bases nitrogenadas iniciou-se com a técnica descrita por Sanger (MAJEWSKI *et al.*, 2011) e aprimorada posteriormente com o método de eletroforese por capilar. A partir de 2005, as novas plataformas de sequenciamento surgiram, trazendo consigo uma nova tecnologia, o Sequenciamento de Nova Geração, do inglês: *Next generation sequencing* (NGS) (MAJEWSKI *et al.*, 2011; SAWYER *et al.*, 2016). Essas novas plataformas, representadas pelas empresas *Ilumina, Ion Torrent da Life Technologies, PacBio* e *MinION*, utilizam métodos diferentes de sequenciamento (METZKER, 2010). Portanto, o NGS é capaz de fornecer a leitura de milhares de pares de base em uma única reação (corrida), possibilitando o sequenciamento de vários genes em paralelo, o exoma inteiro ou o genoma de um organismo em uma única corrida (MAJEWSKI *et al.*, 2011; METZKER, 2010).

Dessa forma, a partir do sequenciamento dos genomas por NGS foi possível identificar um número maior de SNPs nas diferentes espécies (OZAKI *et al.*, 2002), o que permitiu a construção de painéis de alta densidade de SNPs para genotipagem (SEVILLANO *et al.*, 2015), possibilitando estudos de associação global do genoma (GWAS) também em animais de produção (MCCARTHY *et al.*, 2008). Além disso, estudos do transcriptoma e do exoma também se tornaram possível em outras espécies devido a redução do custo do sequenciamento (METZKER, 2010). Essas três metodologias serão abordadas a seguir.

1.5.1 ESTUDOS DE ASSOCIAÇÃO GLOBAL DO GENOMA (GWAS)

Os estudos de associação global do genoma (GWAS) permitem a associação de milhares de *loci* genômicos com características complexas permitindo compreender melhor a relação entre genótipo e fenótipo dos indivíduos (MCCARTHY *et al.*, 2008). Esses estudos vêm sendo aplicados em diferentes espécies animais e vegetais buscando compreender melhor os mecanismos genéticos atuantes na expressão do fenótipo (BARBAN *et al.*, 2016; SANTANA *et al.*, 2015; WOLFE *et al.*, 2016). Dessa forma,

busca-se identificar regiões de maior efeito sobre a característica de interesse para, posteriormente, realizar uma investigação aprofundada das funções biológicas dos genes e das regiões funcionais, visando compreender melhor a influência genética sobre a expressão do fenótipo (XIANG *et al.*, 2015).

Neste tipo de análise são utilizados os SNPs que são a forma de variação mais frequente no genoma (OZAKI *et al.*, 2002), mas CNVs também podem ser utilizadas (LEE *et al.*, 2012; WANG *et al.*, 2015). Os SNPs, por exemplo, estabelecem localizações no genoma conhecidas como marcadores moleculares que contribuem para a variabilidade genética das populações (LLINARES-LÓPEZ *et al.*, 2015). O objetivo da busca da associação entre os alelos ou genótipos com o fenótipo é determinar se um alelo associa-se com determinada doença ou característica produtiva na população de interesse (ROSELLI *et al.*, 2018) e esta associação pode indicar a relação causal direta ou indireta com o desenvolvimento de uma determinada doença (SANNA *et al.*, 2011).

A metodologia foi desenvolvida inicialmente através de estudos epidemiológicos humanos (MCCARTHY *et al.*, 2008), porém, vem sendo aplicada em diferentes espécies como, por exemplo, humanos (BARBAN *et al.*, 2016), bovinos (SANTANA *et al.*, 2015), plantas (WOLFE *et al.*, 2016), aves (PERTILLE *et al.*, 2017; MOREIRA *et al.*, 2019) e suínos (GRINDFLEK *et al.*, 2018; SEVILLANO *et al.*, 2015). Por meio de GWAS foi possível identificar SNPs para várias doenças complexas, incluindo câncer de mamas (EASTON *et al.*, 2007; GHOUSSAINI *et al.*, 2012), Parkinson (LI *et al.*, 2012), artrite reumatoide (KURREEMAN *et al.*, 2012; PLENGE *et al.*, 2007) e diabetes tipo 1 e 2 (EASTON *et al.*, 2007; LU *et al.*, 2012; SANDHOLM *et al.*, 2012; ZEGGINI *et al.*, 2008) em humanos.

Em suínos, os estudos de associação global do genoma têm permitido também a identificação de regiões genômicas e genes associados com a manifestação das hérnias. Genes candidatos presentes em regiões de QTLs previamente descritas vêm sendo associados com o desenvolvimento da hérnia inguinal/escrotal. Entre eles, o *INSL3* (*insulin like* 3), que está relacionado ao crescimento e diferenciação do gubernáculo mediando a descida testicular intra-abdominal (BURKHARDT *et al.*, 1994; RICHARD, 2007), o *MIS* (*Mikimopine synthase*), que provoca a regressão dos ductos de Muller no embrião masculino e atua na diferenciação de células foliculares (REY *et al.*, 2003) e o *CGRP*

(Calcitonina), que é responsável pela regeneração do tecido nervoso após lesão (WOOLLEY *et al.*, 2017).

Sevillano *et al.* (2015) identificaram, através da análise de GWAS, SNPs relacionados ao desenvolvimento da hérnia escrotal/inguinal em duas raças de suínos. Para a raça Large White, 10 SNPs foram identificados em cinco regiões de QTLs localizados nos cromossomos 3, 5, 7, 8 e 13, sendo que o SNP mais significativo em cada região de QTL explicou entre 1,22% e 1,60% da variância total da incidência destas hérnias na linhagem. Para a raça Landrace, 22 SNPs foram significativos em cinco regiões de QTL localizadas nos cromossomos 1, 2, 4, 10 e 13. O SNP mais significativo em cada região de QTL explicou entre 1,15% e 1,46% da variância total da incidência da hérnia escrotal/inguinal na população Landrace (SEVILLANO *et al.*, 2015). Este estudo relatou a identificação de novas regiões associadas ao desenvolvimento da hérnia escrotal/inguinal bem como a sugestão dos genes *RHOA, LEF1 e EGF* como candidatos à suceptibilidade a essa anomalia em suínos, o que permitiu avançar no conhecimento de regiões genômicas associadas à manifestação da hérnia escrotal/inguinal.

Em outro estudo de GWAS, Liao *et al.* (2015) identificaram dois *loci* sugestivos para predisposição à hérnia umbilical em suínos, um no SSC2 (rs81358018) e outro no SSC17 (rs81479278), em uma população de suínos das raça Duroc. Esses autores sugerem o gene *SRC* como candidato posicional e funcional à suceptibilidade a essa anomalia. Long *et al.* (2016), avaliando as raças de suínos Duroc, Landrace e Yorkshire, identificaram oito regiões de CNVs (do inglês, copy number variation) associadas ao desenvolvimento de hérnias umbilicais, incluindo uma CNV no *locus* do gene *NUGGC (Nuclear GTPase Germinal Center Associated),* sendo que todos os animais afetados compartilharam um haplótipo do gene *NUGGC.* Dessa forma, os autores concluiram que as CNVs encontradas no gene *NUGGC* contribuem para a ocorrência de hérnias umbilicais em suínos.

Grindflek *et al*. (2018) identificaram, através de GWAS, 126 SNPs no SSC14 entre a posição 47,16Mb e 58,91Mb associados ao desenvolvimento da hérnia umbilical na raça Landrace explicando cerca de 8,6% da variância fenotípica para essa condição. Ainda, nesta região de QTL, os autores identificaram os genes *Leukemia inhibitory factor* (*LIF*) e *Oncostatin M* (*OSM*) como candidatos para essa anomalia. Além disso, Fernandes *et al.* (2018) identificaram cinco SNPs associados com hérnia umbilical em suínos comerciais

nos cromossomos 4, 6, 11 e 13, sendo que alguns genes candidatos foram identificados nessas regiões, como o *TBX15 (T-box 15), WARS2 (tryptophanyl-tRNA synthetase 2), LIPI (lipase I)* e *RBM11 (RNA Binding Motif Protein 11).* Mais recentemente, Li et al. (2019) detectaram um SNP no gene *CAPN9* (*Calpain 9*) do cromossomo 14 do suíno significativamente associado à hérnia umbilical. Além disso, mais dois SNPs significativos nos cromossomos SSC9 e 16 foram identificados. Neste mesmo estudo, uma mutação no exon 10 do gene *CAPN9* foi identificada nos animais afetados com hérnia umbilical a qual não existia em animais sem hérnias.

1.5.2 SEQUENCIAMENTO DO TRANSCRIPTOMA

O transcriptoma é o conjunto completo de transcritos de uma célula, podendo proporcionar conhecimento dos processos biológicos que estão ocorrendo em uma célula em determinado momento (LI *et al.*, 2012). A regulação do transcriptoma é fundamental para processos fisiológicos, patológicos e de desenvolvimento, já que o perfil de expressão do mRNA (RNA mensageiro) determinará as características das células ou dos tecidos em seu estado específico (MARIONI *et al.*, 2008). Os perfis de um transcriptoma em resposta a estímulos biológicos e fisiológicos fornecem informações e respostas valiosas para a interpretação dos elementos funcionais do genoma revelando mecanismos moleculares das células e para a compreensão do desenvolvimento de processos biológicos fundamentais (FU *et al.*, 2009).

O sequenciamento do RNA (RNA-seq) é uma abordagem desenvolvida recentemente para quantificar e mapear genes de um transcriptoma, sendo uma ferramenta poderosa de análise de expressão gênica (TRAPNELL *et al.*, 2013). Após a extração total do RNA, seleciona-se o RNAm, que é fragmentado e convertido em uma biblioteca de cDNA, que representa de maneira abrangente a composição do transcriptoma (WANG *et al.*, 2011). Após o sequenciamento, as sequências oriundas dos transcritos expressos em um determinado tecido podem ser alinhadas contra um genoma de referência (WANG; GERSTEIN; SNYDER, 2009), permitindo descobrir novos genes, quantificar a expressão gênica, montar e anotar novos transcritos, quantificar sítios de *splicing* e identificar mutações e polimorfismos de interesse (TRAPNELL *et al.*, 2013; YU *et al.*, 2012).

A utilização da técnica de RNA-seq como ferramenta para análise de expressão gênica já foi descrita em diferentes espécies como, por exemplo, avaliando níveis de RNAm de derivados de globulina em bovinos e equinos (CORREIA *et al.*, 2018), regiões de *splicing* em tecido muscular cardíaco de adolescentes (STARK; GRZELAK; HADFIELD, 2019), regulação da espermatogênese em camundongos (JI *et al.*, 2019), doenças do sistema imunológico em peixes (SUDHAGAR; KUMAR; EL-MATBOULI, 2018) e problemas locomotores em aves (PEIXOTO *et al.*, 2019).

Além disso, alguns estudos utilizando RNA-seq em suínos já foram desenvolvidos. Romano *et al.* (2020) identificaram os mecanismos biológicos e genéticos envolvidos na ocorrência e manifestação da hérnia escrotal em suínos utilizando análises do transcriptoma. Dos 13.498 genes expressos no tecido do anel inguinal, 703 foram diferencialmente expressos, sendo que 209 estavam mais expressos e 494 menos expressos nos suínos herniados em relação aos não-herniados. Os genes *MYBPC1*, *BOK*, *SLC25A4*, *SLC8A3*, *DES*, *TPM2*, *MAP1CL3C* e *FGF1* foram destacados como fortes candidatos no desenvolvimento da hérnia escrotal. Já, Souza *et al.* (2020) avaliaram o perfil de expressão gênica no anel umbilical de suínos normais e acometidos pela hérnia umbilical utilizando RNA-seq. Foram identificados 13.216 genes expressos no tecido herniário. Destes, 145 foram menos expressos e 85 mais expressos nos suínos herniados em relação aos não-herniados. Além disso, as vias biológicas mais importantes identificadas foram as do sistema imune, desenvolvimento anatômico, adesão celular e matriz extracelular, sendo os genes *ACER2*, *SLC2A6*, *PTGS1*, *LGALS3*, *KANK3* e *FOS* destacados como importantes candidatos envolvidos no desenvolvimento da hérnia umbilical em suínos.

1.5.3 SEQUENCIAMENTO DO EXOMA

O sequenciamento do exoma é uma metodologia genômica importante para caraterização do genoma funcional que sequencia todos os exons de genes codificadores de proteínas, que representa entre 1 e 2% do genoma, dependendo da espécie (WARR *et al.*, 2015). Na espécie humana, os exons correspondem a menos de 2% do genoma e é nesta porção do DNA que se concentram parte das mutações com potencial patogênico que podem ser determinadas geneticamente (MAJEWSKI *et al.*, 2011; SAWYER *et al.*, 2016). Além disso, em humanos, aproximadamente 85% das mutações causadoras de doenças

conhecidas podem ser encontradas na região codificadora ou nos locais de *splicing* dos genes codificadores de proteínas (BARBAN *et al.*, 2016; CHOI *et al.*, 2009; ZHAO *et al.*, 2016).

O DNA é constituído por nucleotídeos que são ligados uns aos outros pelas ligações fosfodiester, dando origem a formação da longa fita (ALBERTS *et al.*, 2017; ZAHA, FERREIRA e PASSAGLIA, 2014). Esse DNA constitui os genes, que são divididos em exons e introns. Os exons são as sequências (regiões) codificadoras de proteínas que são intercaladas por sequências (regiões) não codificadoras chamadas de íntrons (ALBERTS *et al.*, 2017; ZAHA, FERREIRA e PASSAGLIA, 2014). Para o sequenciamento do exoma, a prioridade é sequenciar a parte codificadora do DNA (METZKER, 2010), ou seja, os exons. Portanto, é necessária a retirada das regiões não codificadoras, como os íntrons, regiões regulatórias e regiões intergênicas (ZAHA, FERREIRA e PASSAGLIA, 2014). Para isso, a depleção dos íntrons é realizada através de kits específicos que contem sondas que capturam somente a parte codificadora do genoma, permitindo o sequenciamento das regiões funcionais. Esta metodologia de sequenciamento permite identificar variantes funcionais localizadas no genoma e predizer as alterações estruturais na formação das proteínas (WARR *et al.*, 2015). De Ligt *et al.* (2012) avaliaram pacientes com déficit intelectual grave e, a partir do sequenciamento do exoma dessas amostras, obtiveram 79 polimorfismos associados a este problema. Destes, 16 eram sinônimos e 63 não sinônimos, afetando genes envolvidos com déficit intelectual grave. Han *et al.* (2019) sequenciaram o exoma completo de um grupo de pessoas com transtorno bipolar e identificaram diversos genes associados a esse distúrbio, incluindo um SNP no gene *KMT2C (Lysine Methyltransferase 2C)* que regula a metilação do gene *H3K4 (histona H3 lisina 4*) envolvido na remodelação da cromatina.

O sequenciamento total do exoma (do inglês, *whole-exome sequencing* - WES) já provou ser um excelente método de detecção de câncer metastático, identificando várias alterações frequentes em genes supressores de tumores conhecidos, como *ATM, RB1, TP53, CDKN2A/B, PTEN* e *PPP2R2A* (ADALSTEINSSON *et al.*, 2017). Além disso, a metodologia de análise do exoma permite identificar padrões de alteração de CNVs entre pacientes doentes e saudáveis.

O estudo do exoma tem acelerado a descoberta de variantes codificadoras de proteínas atuantes nas características fenotípicas em humanos e camundongos (BADEMCI *et al.*, 2014). Robert *et al.* (2014) validaram esta metodologia em suínos, em que o sequenciamento do exoma mostrou-se ser uma ótima ferramenta para identificar variações em regiões codificadoras associadas a características de produção, função embrionária e neonatal. Em animais cujo objetivo é a produção, há poucos trabalhos utilizando informações sobre o sequenciamento do exoma, relacionando essas descobertas com os problemas enfrentados no sistema atual de produção. Guiatti *et al.* (2015) realizaram uma comparação entre o exoma e o genoma de referência do suíno Sscrofa10.2 e observaram um total de 232.530 SNPs, dos quais 20,6% foram mapeados em regiões exômicas e 49,5% em regiões intrônicas. Porém, com suínos ainda há a necessidade de desenvolver estudos que consigam associar a genética com o desenvolvimento de determinadas anomalias. Dessa forma, esta metodologia pode ser usada para o estudo de fatores desencadeadores de hérnias que são anomalias ligadas ao desenvolvimento dos suínos.

2 OBJETIVOS

2.1 Objetivo geral

Identificar genes e polimorfismos associados com a manifestação de hérnia umbilical em suínos utilizando o sequenciamento de nova geração.

2.1.1 Objetivos específicos

Identificar polimorfismos (SNPs, inserções e deleções - InDels) e genes presentes no exoma e no transcriptoma do anel umbilical de suínos normais e afetados com HU.

Identificar os polimorfismos (SNPs e InDels) e genes associados à hérnia umbilical a partir do sequenciamento do exoma e do transcriptoma.

Comparar os polimorfismos e genes identificados entre o exoma e o transcriptoma associados à hérnia umbilical.

Comparar os polimorfismos e genes encontrados pelas metodologias do exoma e do transcriptoma com resultados de estudos de associação global do genoma (GWAS).

Caracterizar as estruturas proteicas resultantes das variantes funcionais identificadas entre o exoma e o transcriptoma.

Avançar no conhecimento dos mecanismos moleculares que atuam na manifestação da hérnia umbilical em suínos.

3 CAPÍTULO II

MANUSCRITO

Os resultados desta dissertação são apresentados na forma de um manuscrito, com sua formatação de acordo com as orientações da revista BMC Genomics.

3.1 **MANUSCRITO I**

A joint analysis using exome, transcriptome and GWAS data identifies candidate polymorphisms and genes involved with umbilical hernia in pigs

Autores: Igor Ricardo Savoldi, Adriana Mércia Guaratini Ibelli, Mauricio Egídio Cantão, Jane de Oliveira Peixoto, Michele Porto Pires, Marcos Antônio Zanella Mores, Essamai Brizola Lagos, Jader Silva Lopes, Ricardo Zanella, Mônica Corrêa Ledur

De acordo com normas para publicação em: BMC Genomics

A joint analysis using exome, transcriptome and GWAS data identifies candidate polymorphisms and genes involved with umbilical hernia in pigs

Igor Ricardo Savoldi[1], Adriana Mércia Guaratini Ibelli[2,3], Mauricio Egídio Cantão[2], Jane de Oliveira Peixoto[2,3], Michele Porto Pires[4], Marcos Antônio Zanella Mores[2], Essamai Brizola Lagos[5], Jader Silva Lopes[6], Ricardo Zanella[7,8], Mônica Corrêa Ledur[1,2*]

Igor Ricardo Savoldi[1]
E-mail: Igorsaavoldii@gmail.com

Adriana Mércia Guaratini Ibelli[2,3]
E-mail: adriana.ibelli@embrapa.br

Mauricio Egídio Cantão[2]
E-mail: mauricio.cantao@embrapa.br

Jane de Oliveira Peixoto[2]
E-mail: jane.peixoto@embrapa.br

Michele Porto Pires[4]
E-mail: micheleportopires@gmail.com

Marcos Antônio Zanella Mores[2]
E-mail: marcos.mores@embrapa.br

Essamai Brizola Lagos[5]
E-mail: essamay.brlagos@gmail.com

Jader Silva Lopes[6]
Email: jader.lopes@brf-br.com

Ricardo Zanella[7,8]
E-mail: ricardozanella@upf.br

Mônica Corrêa Ledur[1,2*]
E-mail: monica.ledur@embrapa.br

*Corresponding author

Mônica Corrêa Ledur

E-mail: monica.ledur@embrapa.br

Embrapa Suínos e Aves,

Rodovia BR-153, Km 110, Distrito de Tamanduá, Caixa Postal: 321, Concórdia, Santa Catarina 89715-899, Brazil.

[1]Programa de Pós-Graduação em Zootecnia, Universidade do Estado de Santa Catarina, UDESC-Oeste, Chapecó, SC, Brazil.
[2]Embrapa Suínos e Aves, Concórdia, Santa Catarina, Brazil.
[3]Programa de Pós-Graduação em Ciências Veterinárias, Universidade Estadual do Centro-Oeste, Guarapuava, Paraná, Brazil
[4]Instituto Catarinense de Sanidade Agropecuária, ICASA, Brasil.
[5]Programa de Pós-Graduação em Zootecnia, Universidade Estadual de Ponta Grossa, UEPG, PR, Brazil.
[6]BRF S.A, Curitiba, Paraná, Brazil.
[7]Universidade de Passo Fundo, Passo Fundo, RS, Brazil
[8]Programa de Mestrado em BioExperimentação, UPF, Passo Fundo, RS, Brazil

Abstract

Background: Umbilical Hernia (UH) is characterized by the passage of part of the intestine through the umbilical canal forming the herniary sac. There are several potential causes that can lead to the umbilical hernia such as bacterial infections, management conditions and genetic factors. Since the genetic components involved with UH are poorly understood, this study aimed to identify polymorphisms and genes associated with the manifestation of umbilical hernia in pigs using next generation sequencing.

Results: In the exome sequencing, a total of 459,728 polymorphisms were identified in all 10 samples evaluated and, when comparing normal and UH-affected pigs, 209 variants located in 72 genes were identified. In the transcriptome, a total of 160,555 polymorphisms were found and, when comparing normal and UH-affected pigs, 81 variants were identified, located in 42 genes. Out of those, 29 variants were of SNPs type classified through the VEP as missense variants. Comparing the two methodologies, we obtained 79 concordant variants between the exome and transcriptome analyses, which were located in 23 genes. Moreover, 8 genes were also shared with the GWAS data, totaling 91 identified genes

distributed in 48 biological processes (BP). Among the BP involved with UH it is possible to highlight muscle contraction, immune response, and cell-matrix adhesion.

Conclusions: We have generated the first exome sequencing related to normal and umbilical hernia-affected pigs, which allowed us to identify several variants involved with this disorder. Moreover, comparing these variants with the results from RNA-Seq and GWAS, it was possible to identify variants and genes that had not been previously related to the development of umbilical hernia in pigs. We highlight the variants in the *FYN, SPN, ITGB2, MYLPF, MYH13, MYH8, MYH2, MYH3, MYO19, VIM, ROCK2, RABGEF1* and *UBASH3B* genes as strong candidates in the development of UH in pigs, as well as the importance of combining the three approaches to identify candidate genes and polymorphisms involved with UH.

Keywords: Candidate genes, congenital defects, transcriptomics, GWAS, swine.

Background

Umbilical hernia (UH) is a condition that negatively affects pigs, being considered the most common congenital defect in this species and is characterized by the abdominal content protruding through the umbilical ring [1]. In addition to economic losses caused by reduced performance, UH results in welfare concerns to the modern pig industry [1,2,3]. This condition occurs due to weakened support of muscles around the umbilical ring or umbilicus of the animal [4], causing the non-closing of the umbilical area properly and, consequently, the intestines protrude through the abdominal wall to form the herniary sac. Moreover, collagen metabolism can be involved with the UH development [5].

The etiology of UH is likely to be multifactorial, affected by genetic and environmental factors, such as physical injury, obesity, inappropriate removal of the umbilical cord and infections [6]. It has been reported that the occurrence of UH ranges from 0.40 to 2.25%, affecting mainly 9 to 14 week-old pigs [1,7]. Moreover, animals under the same management conditions may be affected or not by hernias [8] and the heritability

estimate of 0.25 for UH in pigs [9] indicate that the development of umbilical hernia is controlled by a genetic component.

Few studies have already been developed seeking to understand the UH genetic inheritance. Ding et al. [10] observed significant linkage between markers and scrotal/inguinal and umbilical hernia in pigs on 12 different chromosomes. Liao et al. [8] identified two suggestive loci predisposing to umbilical hernia on SSC2 and SSC17 in a Duroc population. In a genome-wide association study (GWAS) with commercial pigs, Fernandes at al. [11] identified five SNPs associated with umbilical hernia: one in SSC4 (rs334706328), two in SSC6 (rs80813241, rs81337222), one in SSC13 (rs337360700) and the other with unknown position in the pig genome. Moreover, Grindflek et al. [12], studying the Norwegian Landrace pigs, identified a highly significant Quantitative Trait Loci (QTL) for umbilical hernia, detected between 48 and 51 Mb on SSC14.

Even though some studies have been performed, they suggested that this disorder is complex and affected by multiple genes and causal variants. Thus, further studies are required to identify additional susceptibility loci and causative genes for UH in pigs using different strategies. Therefore, to clarify the genetic basis of swine umbilical hernia, this study aimed to identify polymorphisms and genes associated with UH in pigs through the whole-exomic sequencing, and additional transcriptome and GWAS data analyses.

Methods

Animals and sample collection

A total of 10 Landrace purebred females (with approximately 90 days of age) was used in a case-control design. These gilts were selected from the same nucleus farm with high sanitary status, located in Santa Catarina State, south of Brazil. From those, 5 were affected by umbilical hernia and 5 were healthy selected from families with no history of umbilical hernia. The animals were transported to the Embrapa Swine and Poultry National Research Center to be necropsied and to confirm the presence or absence of UH, as described by Souza et al. [13]. The euthanasia was performed by electrocution for 10 seconds following the procedure approved by the Embrapa Swine and Poultry National Research Center

Ethical Committee of Animal Use (CEUA) under the protocol number 011/2014. Ear tissue samples were collected and stored at -20 ºC until DNA extraction.

DNA isolation

Genomic DNA was extracted from 70-100 mg of ear tissue using Purelink Genomic DNA Mini kit (Thermo Fisher Scientific, Waltham, MA, USA). Briefly, tissue digestion was performed adding 200µL of Genomic Digestion Solution Buffer and 20µL Proteinase K for 4 hours at 55 °C. The samples were centrifuged at 14,000 rpm for three minutes at room temperature, and 20µL RNAse, 200µL Pure Link Lysis/Binding Buffer and 200µL 100% alcohol were added. The solution was pipetted into the silica column with the washing steps performed with 500µL Wash Buffer 1 and 2 centrifuged at 12000 rpm per 1 minute to bind DNA to the silica column. Finally, the DNA was eluted in 50µL Elution Buffer solution. The concentration and quality of samples were measured in a Biodrop spectrophotometer (Biodrop, England, UK) and in a 1,5% agarose gel electrophoresis. Only DNA samples showing the 260/280 ratio between 1.8 and 2.0 were used for further analyses.

Exomic capture and sequencing

To prepare the Next Generation Sequencing (NGS) libraries, the SeqCap EZ Exome Probes v3.0 kit were used. The DNA fragmentation was performed using the Bioruptor® equipment (Diagenode, Denville, NJ, USA) following the recommendations of the protocols. The gDNA was fragmented to an insertion size of approximately 150 bp, generating dsDNA fragments with 3' or 5' overhangs to index the transport adapters for sequencing. Samples with a concentration of 10 ng/µl diluted in TE (10 mM Tris, 1 mM EDTA, pH 7.5 - 8.0) were pipetted into 0.1 ml Bioruptor® Microtubes. The sonication was performed in three stops of seven minutes each containing 15 cycles (30”/ 30” on / off time) submerged in water at 4 °C.

After fragmentation, the clean-up of fragmented DNA was performed using purification beads (SPB), followed by the blunt end repair (ERP3). After the final repair, the library size was selected using SPB beads. Next, the adapters were ligated and samples were prepared for the probe hybridization. Finally, the enrichment of the DNA fragments was performed, as this process selects and enriches the DNA fragments that have adapters

at the ends and amplifies the amount of DNA in the library. Furthermore, sequencing was performed on Illumina's HiSeq 2500 at the ESALQ/USP Functional Genomics Center (São Paulo/Brazil) using a paired-end 100 bp library.

Sequencing analysis and annotation

The FASTQ files were submitted to quality control (QC) analysis using the Trimmomatic tool [14) to remove low-quality sequences (PHRED ≤ 20). The remaining reads were mapped against the *Sus scrofa* reference genome (Sscrofa11.1) using the BWA-MEM [15]. Variant calling (SNP and InDel) was performed with GATK tool v.3.6 following the general guidelines for whole-exome sequencing (WES) [16]. The variant effect predictor (VEP) tool available in the Ensembl 98 [17] was used to annotate and identify the effects and consequences of all variants that differed between normal and UH-affected pigs. For this analysis, the data resulting from the GATK were imputed in the VCF format. The list of variants was submitted to the VEP tool from Ensembl 98 using its standard criteria, in which additional identifiers for genes (gene symbol, transcript version, and protein), transcripts and variants were used (transcript biotype, exon and intron numbers, phenotypes and Upstream/Downstream distance 5,000 bp), co-located variants and frequency data.

Additionally, the SIFT score (Sorting Intolerant From Tolerant) [17,18], available in the VEP tool, was used to identify the potential impact of amino acid substitutions on protein structure and function, which can, consequently, alter the phenotype. The use of this tool implies in a better prediction of the effect of non-synonymous coding variants [19]. The SIFT score is given for each variant by which one can predict whether the variant can affect the protein function or not. A SIFT score ranging from 0.00 to 0.05 classifies a variant as deleterious and from >0.05 to 1.00 as tolerated. However, SIFT also assigns values classified as tolerated low confidence. These values are assigned to SNPs located in conserved regions of the genome, which could indicate that this substitution may be at a position that had no time to evolve and it is conserved by chance and, hence, predicted to be deleterious [17, 19].

Transcriptome variants involved with UH

To verify the concordance of variants found in the DNA using the exomic sequencing from those found in the transcriptome between healthy and herniated pigs, samples from the Bioproject PRJNA445856 were used. This transcriptomic data was previously described by Souza et al. [13], which was generated with the same animals used for the exome sequencing. The sequences generated in the transcriptome analysis were submitted to quality control using the Trimmomatic tool [14] to remove low-quality sequences (PHRED $\leq$ 20) and mapped against the reference genome *Sus scrofa* (NCBI Sus scrofa 11.1 – Ensembl 98) using the STAR tool [20]. The identification of different variants between both groups was performed using the GATK tool v. 3.6, following the Guide of Best Practices for using GATK [16]. Subsequently, the data obtained was submitted to the VEP Ensembl 98 program for annotation and prediction of variants using the same input criteria as those used for the exome analysis.

Genome-Wide Association Study (GWAS)

Here, GWAS results obtained from a previous study from our group (Lagos, in preparation) were used in an integrated analysis with the exomic and transcriptomic data. Briefly, a total of 325 commercial pigs (92 herniated and 233 non-herniated) were genotyped using the GGP Porcine Bead Chip 50 K (Neogen, Lincoln, NE, USA). The Blossoc [21] program, which consists in a multiple marker approach, was used to identify SNPs associated with UH. A total of 735 genes located in regions associated and suggestively associated with UH was prospected in commercial pigs and, in the current study, this set of genes was used to compare the results found (identified polymorphisms and genes) using the other two methods.

Functional analysis with the three approaches

To evaluate the functions of the common identified genes with the exome, transcriptome, and GWAS approaches, the DAVID 6.7 database (https://david.ncifcrf.gov) [22] was used to classify the Gene Ontology (GO) categories of cellular component, biological process (BP) and molecular function. Afterwards, the BP enriched with genes through the DAVID were grouped using REViGO (http://revigo.irb.hr) [22] for better visualization. Interactions

between genes were predicted with the NetworkAnalyst program (https://www.networkanalyst.ca) [23]. Furthermore, it was verified whether the genes found in our study were in QTL regions previously mapped for umbilical hernia in pigs or not using the Pig QTLdb from the Animal Genome Database (http://www.animalgenome.org/QTLdb/app).

Results

Whole-exomic analysis

Using the GATK for variant discovery, a total of 459,728 variants (SNPs and InDels) were identified in all 10 samples evaluated. The highest number of variants was found in the swine chromosome (SSC) 6, 2 and 1, respectively (Fig. 1).

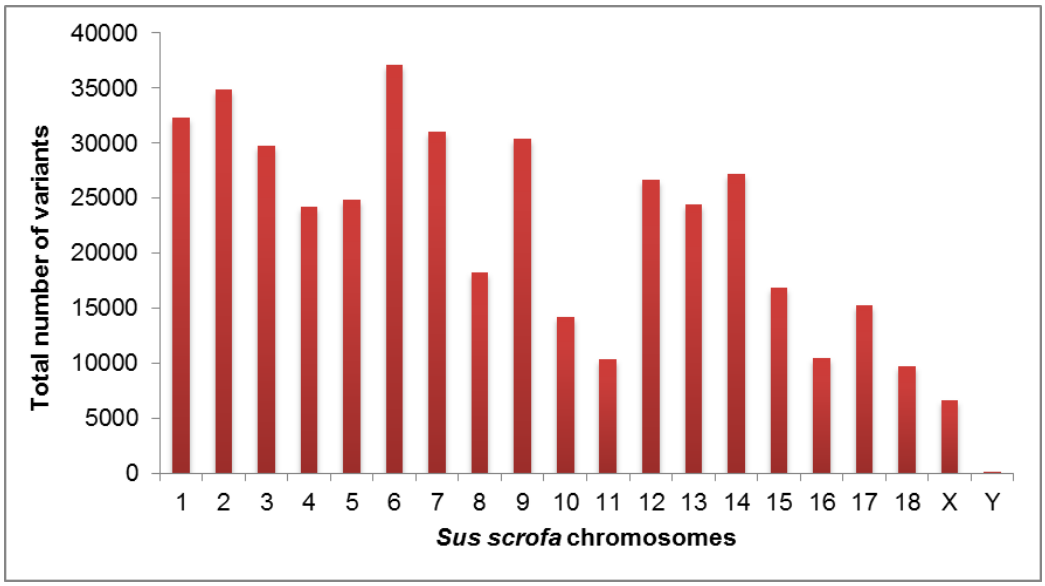


Figure 1. Total number of variants detected in each chromosome using the pig exome sequencing.

Considering only the variants between the two groups (healthy and UH-affected pigs), 209 polymorphisms were identified (Table 1). From those, 158 have already been described in pigs, and the other 41 are new polymorphisms that were firstly described here. From the 209 variants, using the VEP tool, 5 were classified in intergenic regions and 204

were classified in gene regions according to their coding consequence, where 61% were grouped as synonymous, 38% as missense and 1% as frameshift variant.

From all the variants that differed between healthy and UH-affected pigs, 197 were located in the autosomes with the largest number mapped in the SSC3, 6 and 12 (Table 1). Also, two variants were mapped on the X chromosome. The SSC3 had the largest number of variants identified, with a total of 118 variants located in 34 genes (Table 1). We highlight some of them: RAB Guanine Nucleotide Exchange Factor 1 (*RABGEF1)*, sulfatase modifying factor 2 *(SUMF2),* Chaperonin containing TCP1 subunit 6A *(CCT6A),* zinc finger protein 713 *(ZNF713),* integrin subunit alpha M *(ITGAM),* Vitamin K epoxide reductase complex subunit *1 (VKORC1),* serine protease 53 *(PRSS53),* calcium release-activated calcium modulator 3 *(ORAI3),* sialophorin *(SPN),* rho associated coiled-coil containing protein kinase 2 *(ROCK2),* myosin light chain phosphorylation fast skeletal muscle *(MYLPF),* and ring finger protein 40 *(RNF40)* (Table 1). On SSC 6, 27 variants were identified, located in 8 genes, such as EPH receptor B2 (*EPHB2),* elongin A *(ELOA),* cannabinoid receptor 2 *(CNR2)* and Chloride Intracellular Channel 4 (*CLIC4)* (Table 1). On SSC12, 23 variants were identified in 12 genes, such as myosin XIX *(MYO19),* Acetyl-CoA carboxylase alpha *(ACACA),* myosin-13 *(MYH13),* myosin-8 *(MYH8),* myosin, heavy chain 2, skeletal muscle, adult *(MYH2),* myosin-3 *(MYH3)* and ADP-ribose/CDP-alcohol diphosphatase, manganese dependent *(ADPRM)* (Table 1).

Table 1. Total number of variants detected by analyzing the exome between normal and umbilical hernia-affected pigs.

SNP name	Location	Gene/Symbol
rs332203161	1:3427322	-
rs707913396	1:3427328	-
rs713738573	1:3427330	-
rs81505732	2:7592175	*SLC22A11*
rs333210122	2:7592186	*SLC22A11*
rs322282589	2:7592192	*SLC22A11*
new	2:7592249	*SLC22A11*
rs787087930	2:7592295	*SLC22A11*
rs334404716	2:7598086	*SLC22A11*
rs338734805	2:7598137	*SLC22A11*
rs792193084	2:7602943	*SLC22A11*
new	2:7602945	*SLC22A11*
new	3:15942680	*ENSSSCG00000007733*
rs81309268	3:15953792	*ENSSSCG00000007733*

new	3:16078557	*ENSSSCG00000007733*
rs713010025	3:16376678	*RABGEF1*
rs337976532	3:16382428	*KCTD7*
rs320729536	3:16382619	*KCTD7*
rs323726488	3:16383246	*KCTD7*
rs333208968	3:16383250	*KCTD7*
rs333661817	3:16383284	*KCTD7*
rs326053487	3:16383454	*KCTD7*
rs336316568	3:16383763	*KCTD7*
rs329707669	3:16383919	*KCTD7*
rs713060696	3:16383936	*KCTD7*
rs701844432	3:16383956	*KCTD7*
rs324583382	3:16384043	*KCTD7*
rs345204099	3:16384270	*KCTD7*
rs327324699	3:16384293	*KCTD7*
rs329831862	3:16384540	*KCTD7*
rs332144111	3:16384634	*KCTD7*
rs323662654	3:16384658	*KCTD7*
rs335859180	3:16385181	*KCTD7*
new	3:16385215	*KCTD7*
rs332268785	3:16385727	*KCTD7*
new	3:16844063 - 16844064	*SUMF2*
rs330731365	3:16844265	*CCT6A*
rs343913735	3:16844271	*SUMF2*
rs701592732	3:16844932	*CCT6A*
rs691819169	3:16844934	*SUMF2*
rs706237879	3:16844935	*CCT6A*
new	3:16847861	*SUMF2*
rs1108762720	3:16851119	*CCT6A*
rs1113974639	3:16852667	*ENSSSCG00000033141*
rs336023254	3:16854410	*CCT6A*
rs341960076	3:16855237	*CCT6A*
rs339277541	3;16958038	*ZNF713*
rs331020706	3:16958150	*ZNF713*
rs340816565	3:16958194	*ZNF713*
rs327975767	3:16958979	*ZNF713*
rs325589354	3:16959068	*ZNF713*
rs337698799	3:16959239	*ZNF713*
rs318863693	3:16959265	*ZNF713*
rs329457108	3:16959434	*ZNF713*
rs334498862	3:16959634	*ZNF713*
new	3:16959718	*ZNF713*
rs693699557	3:16959721	*ZNF713*
rs324279159	3:16960031	*ZNF713*
rs788766021	3:16960220 - 16960224	*ZNF713*

rs336887529	3:16960465	*ZNF713*
rs343256392	3:16960493	*ZNF713*
rs325887031	3:16960740	*ZNF713*
rs337918521	3:16960969	*ZNF713*
rs320802395	3:16961030	*ZNF713*
rs330288467	3:16961219	*ZNF713*
rs323115420	3:16964045	*ZNF713*
rs335343250	3:16964916	*ZNF713*
rs329911516	3:16970816	*ZNF713*
rs342012840	3:16971089	*ZNF713*
rs324205762	3;16971143	*ZNF713*
rs333780109	3:16971156	*ZNF713*
rs343615406	3:16971169	*ZNF713*
rs328383315	3:16971295	*ZNF713*
new	3:17001571 - 17001575	*SEPTIN14*
rs342352307	3:17001888	*SEPTIN14*
new	3:17026001	*SEPTIN14*
new	3:17033960	*LOC106509673*
new	3:17033997	*LOC106509673*
new	3:17034051	*LOC106509673*
new	3:17034234	*LOC106509673*
new	3:17034293	*LOC106509673*
new	3:17034405	*LOC106509673*
rs702384560	3:17034426	*LOC106509673*
rs709083818	3:17034534	*LOC106509673*
rs324169719	3:17034589	*LOC106509673*
rs330223831	3:17034658	*LOC106509673*
new	3:17243996	*ITGAM*
rs318763073	3:17243999	*ITGAM*
rs344858642	3:17244108	*ITGAM*
rs326942919	3:17246305	*ITGAM*
rs323402259	3:17246405	*ITGAM*
rs327289001	3:17254444	*ITGAM*
rs327947675	3:17399455	*PRSS53*
rs337670844	3:17399477	*ZNF646*
rs328098289	3:17430228	*STX4*
rs341530377	3:17457840	*HSD3B7*
rs344714132	3:17458203	*STX1B*
rs344526901	3:17458409	*HSD3B7*
rs334776843	3:17458467	*STX1B*
rs322669402	3:17459844	*STX1B*
rs793318116	3:17460656	*HSD3B7*
rs327942559	3:17466860	*SETD1A*
rs339276563	3:17468216	*HSD3B7*
rs330957838	3:17468302	*SETD1A*
rs340370195	3:17479170	*SETD1A*

rs345676220	3:17491423	*ORAI3*
rs332536957	3:17497571	*ORAI3*
rs344892486	3:17610468	*ZNF629*
rs345908614	3:17613140	*ZNF629*
rs335540465	3:17613531	*ZNF629*
new	3:17617319 - 17607323	*ZNF629*
rs1107804156	3:17618533	*RNF40*
rs789266896	3:17628688	*ZNF629*
rs319144424	3:17631027	*ZNF629*
rs696120779	3:17636757	*PHKG2*
rs792624385	3:17639370	*RNF40*
rs324849698	3:17639613	*RNF40*
rs343172093	3:17639636	*RNF40*
rs706476605	3:17639641	*RNF40*
rs706021010	3:17639643	*RNF40*
rs690340579	3:17639901	*RNF40*
rs335008558	3:17640590	*RNF40*
rs340945745	3:17767030	*ZNF689*
new	3:17979825	*SPN*
new	3:17979826	*MYLPF*
new	3:17979827	*SPN*
rs328258515	3:101613325	*GALM*
new	3:125492648 - 125492649	*ROCK2*
new	4:89321959	*DEDD*
rs319871479	4:128985819	-
rs329715150	5:17364441	*GRASP*
rs345481021	6:80824380	*EPHB2*
rs326115442	6:80824416	*EPHB2*
rs337528160	6:80824455	*EPHB2*
rs81389080	6:80826802	*EPHB2*
rs343091930	6:80836094	*EPHB2*
rs337803448	6:80836565	*EPHB2*
rs327572607	6:80837841	*EPHB2*
rs323316435	6:80840438	*EPHB2*
rs81389091	6:80842615	*EPHB2*
rs710496863	6:80842862	*EPHB2*
new	6:80843072 - 80843073	*EPHB2*
new	6:80843074 - 80843081	*EPHB2*
rs342283188	6:80843098	*EPHB2*
rs81389092	6:80843265	*EPHB2*
rs346223430	6:80843336	*EPHB2*
rs325312968	6:80843363	*EPHB2*
rs328228830	6:80843833	*EPHB2*

rs331678844	6:80843991	*EPHB2*
rs81389093	6:80844018	*EPHB2*
rs81233096	6:80897035	*TEX46*
rs328260148	6:80977371	*LUZP1*
rs344493496	6:80984204	*LUZP1*
rs329148574	6:81000551	*LUZP1*
rs325089032	6:81571496	*ELOA*
rs345825612	6:81681049	*FUCA1*
rs332149968	6:81681123	*CNR2*
rs321503722	6:82401042	*CLIC4*
rs322593756	7:96619078	*PAPLN*
new	7:113442605	*FBLN5*
rs696812713	8:6157581	*OTOP1*
rs712855168	8:6157582	*OTOP1*
rs328470251	9:1324980	*LOC100736607*
rs345798145	9:49630761	*UBASH3B*
rs332848504	10:8836715	*ENSSSCG00000038539*
rs706216562	10:15297251	*LOC102166270*
rs690508595	10:15297252	*LOC102166270*
rs698813384	10:15297253	*LOC102166270*
new	10:43353476	*CUBN*
new	12:19348500	*DUSP3*
rs335136145	12:38026003	*MYO19*
rs331463738	12:38129509	*DHRS11*
rs340781986	12:38624687	*ACACA*
rs324236192	12:38624714	*ACACA*
rs327539493	12:55096339	*MYH13*
rs324996684	12:55097665	*MYH13*
rs341831793	12:55099832	*MYH13*
rs346367824	12:55103125	*MYH13*
new	12:55146905	*MYH8*
new	12:55152605	*MYH8*
new	12:55164179	*MYH8*
new	12:55164180	*MYH8*
new	12:55164182	*MYH8*
new	12:55197434	*MYH2*
new	12:55197435	*MYH2*
new	12:55205557	*MYH2*
new	12:55263362	*MYH2*
new	12:55267170	*MYH2*
rs319222050	12:55356662	*MYH3*
rs337961174	12:55356663	*MYH3*
new	12:55372418 - 55372420	*MYH3*
new	12:55438152 - 55438155	*ADPRM*
rs342689686	13:12410437	*RARB*

rs330943524	14:108858300	*MMS19*
rs324529516	15:50299014	*UNC5D*
rs790346070	X:39631586	-
new	X:46911316	*ENSSSCG00000045777*

From those 209 total variants, 10 were classified as chromosome scaffold (additional file 1) and 12 were insertion and deletion (InDels) and were located in the following genes: potassium channel tetramerization domain containing 7 *(KCTD7)*, septin 14 *(SEPTIN14)*, zinc finger protein 629 *(ZNF629)*, solute carrier family 66 member 3 *(SLC66A3)*, fibulin 5 *(FBLN5)*, transmembrane protein 220 *(TMEM220)*, *ZNF713*, *RNF40*, *ROCK2*, *EPHB2*, *MYH3*, *ADPRM*, *SUMF2* and *ENSSSCG000000331*. The other 199 variants were classified as SNP and were located in several genes (Tables 1 and 2).

Moreover, according to the VEP, some variants were located in 2 or more genes, because the position of the genes in the genome depends on which strand (forward or reverse) it is located, with different (transcribed) consequences. In general, all variants were grouped into gene, flanking or intergenic regions. In the flanking regions, VEP subdivided the variants in upstream, with 26 (9%) variants, 5' UTR with 12 (4%) variants, downstream with 62 (20%) variants and 3' UTR with 58 (19%) variants. For the coding and intronic gene regions, 55 (18%) variants were grouped in the exons (21 missense, 26 synonymous and 1 frameshift variant) and 86 (28%) were located in the introns. Moreover, 5 (2%) variants were located in intergenic regions (Fig. 2).

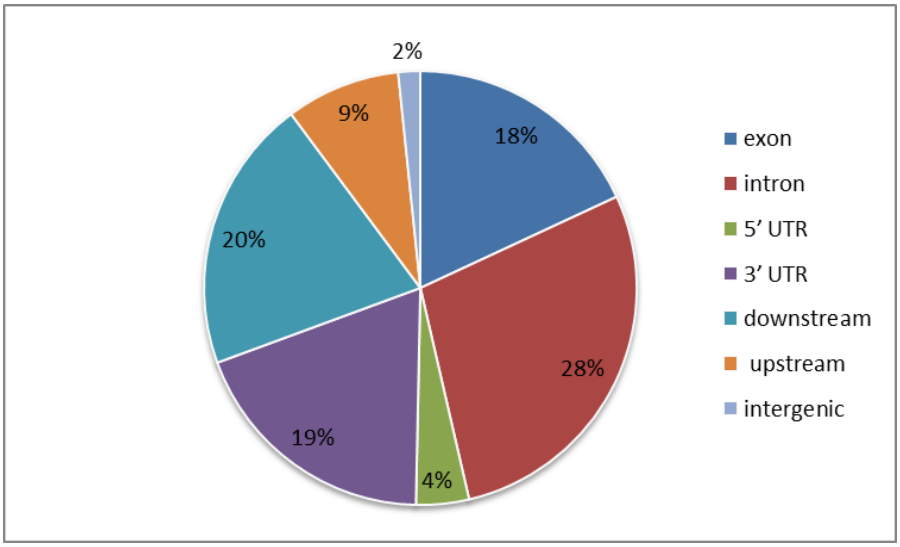


Figure 2. Gene position of 209 variants (SNPs and InDels) generated with the VEP tool when analyzing the pig exomic sequencies.

The 209 polymorphisms identified in the whole exome sequencing between healthy and UH-affected pigs were located in 72 different genes (Table 2). However, several polymorphisms were located in the same gene, for example, the *KCTD7* gene had 19 polymorphisms, most of which were located downstream and 3 'UTR, the *ZNF713* gene had 28 variants, some of which located in 3' UTR, and the *EPHB2* gene had 18 polymorphisms located in different gene positions (additional file 1).

Table 2. Position of the variants obtained by whole-exome sequencing located within 72 genes in the swine genome.

Variant position	**Gene symbol**
Exon	*MYO19, MYH8, MYH3, MYH2, SLC22A11, ZNF713, ITAGM, ZNF646, SETD1A, RNF40, ELOA, OTOP1, LOC106509673, SPN, HSD3B7, ZNF629, CCDC189, LOC100736607, STX4, TYWY* and *CCT6A*
Intron	*DUSP3, DHRS11, MYH13, MYH8, MYH3, MYH2, MMS19, UNC5D, SLC22A11, KCTD7, ZNF713, SEPTIN14, ITGAM, SETDA1, ZNF629, RNF40, CCDC189, MYLPF, DEDD, NIT1, GRASP, EPHB2, TEX46, LUZP1, LYPLA2, CNR2, FUCA1, PAPLN, FBLN5, LOC110259104, CCT6A, TYWY, CUBN, ENSSSCG00000038539, LOC102166270, RARB, ENSSSCG00000047703* and *ENSSSCG00000033141*
Upstream	*MYH3, HSD3B7, SETD1A, ORAI3, RNF40, PFDN2, CNR2, CLIC4, UBASH3B, LOC100514433, ENSSSCG00000049720, LOC106509673, ENSSSCG00000018553, ENSSSCG00000033141, SUMF2* and *ENSSSCG00000007735*
5' UTR	*ZNF713, BCKDK, ZNF629, RNF40, ZNF689, CNR2, TYWY* and *LOC106509673*
Downstream	*DHRS11, ADPRM, TMEM220, KCTD7, ZNF713, SEPTIN14, PRSS53, VKORC1, HSD3B7, STX1B, SETD1A, RNF40, ZNF629, PHKG2, SRSF7, GALM, ROCK2, SLC66A3, NIT1, EPHB2, ENSSSCG00000045777, LOC100514433, LOC110259104, ENSSSCG00000047341, ENSSSCG00000018553, ENSSSCG00000033141* and *ENSSSCG00000027233*
3' UTR	*DUSP3, DHRS11, MYO19, DHRS11, ADPRM, KCTD7, ZNF713, STX1B, HSD3B7, ORAI3, FBXL19, ZNF629, PHKG2, EPHB2* and *CCT6A*

Variants in the transcriptome analysis

When analyzing variants in the transcriptome of the same 10 samples, a total of 160,555 polymorphisms (SNP and InDel) were identified. The highest number of variants was found in chromosomes 6, 3 and 2, respectively (Fig. 3).

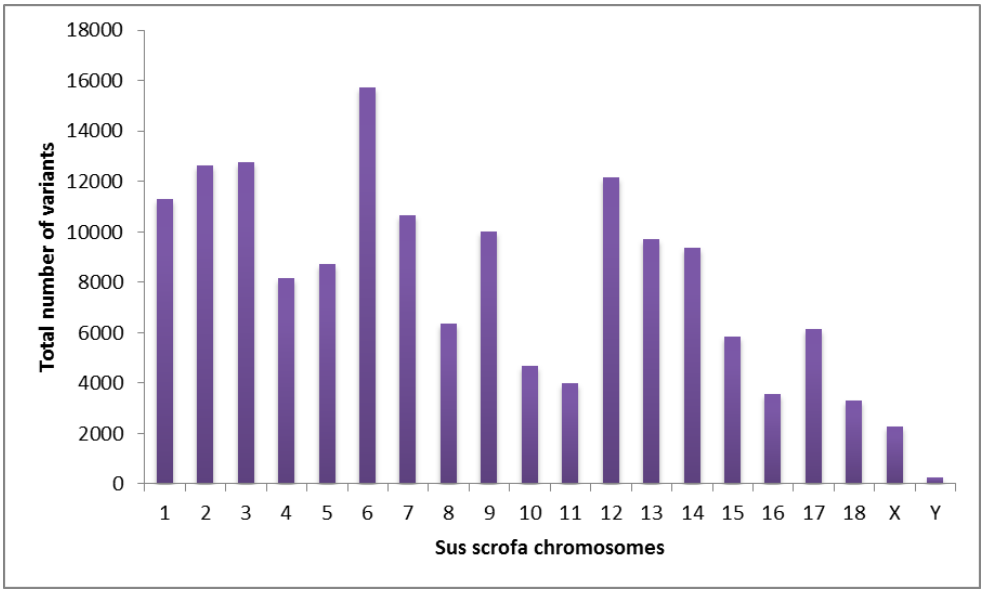


Figure 3. Total number of variants detected in each chromosome using the transcriptomics sequencing from normal and umbilical hernia-affected pigs.

To identify possible polymorphisms related to UH, the comparison between the two groups resulted in 81 different polymorphisms (SNPs and InDels) between herniated and non-herniated pigs (additional file 2). The 81 variants were located in coding regions, introns and flanking or regulatory regions. From those, 67 (82.7%) were existing variants and 14 (17.3%) were new, and they were classified according to their coding consequence, where 71% were synonymous, 22% missense and 7% were frameshift variants. Only 4 variants were InDels, located in the *SUMF2*, *CCT6A*, *ENSSSCG00000033141*, *ZNF629*, *RNF40* and *EPHB2* genes. Furthermore, similar to the exome results, some variants were classified according to the gene position in more than one different gene, where 28 (22%) variants were located in exons (comprising 11 synonymous, 6 missense and 2 frameshift variants), 9 (7%) in introns and 92 (71%) in flanking regions (including 33 (26%) downstream, 12 (9%) upstream, 7 (5%) 5'UTR and 40 (31%) 3'UTR variants) (Fig. 4).

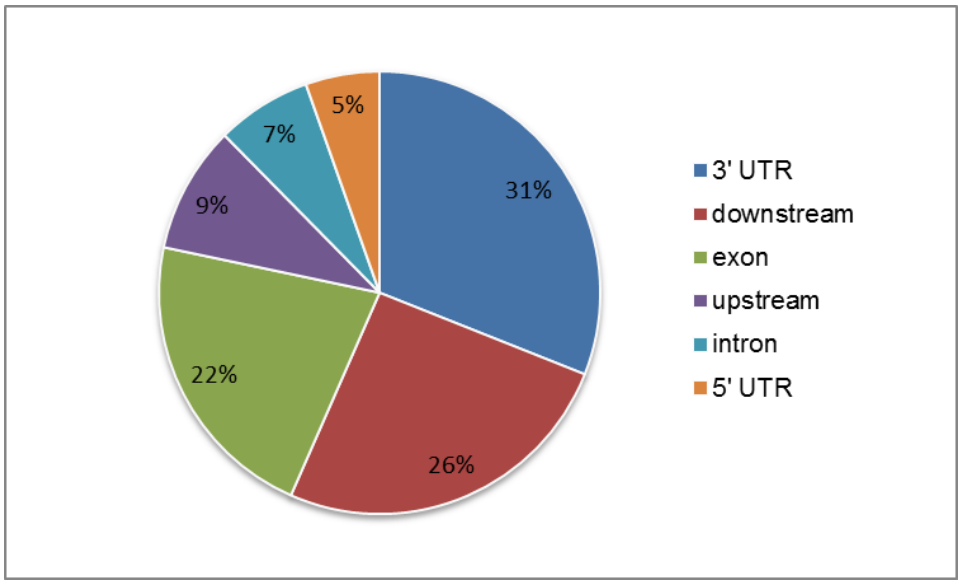


Figure 4. Classification of 81 variants (SNPs and InDels) generated with the VEP tool when analyzing pig transcriptomic sequences according to their position in genes.

The 81 polymorphisms were located in 42 genes (Table 3). Some genes harbored a high number of variants, such as the *KCTD7* gene, with 13 variants located in the 3 'UTR and downstream positions, whereas the *EPHB2* gene had 10 variants located in exons (synonyms and missense variants) and 3' UTR, and the *ITGAM* gene presented 4 exon variants (synonyms and missense variants).

Table 3. Position of the variants obtained by transcriptomics sequencing located within the 42 genes in the swine genome.

Variant position	**Gene symbol**
Exon	*NCOA7*, *SEC62*, *ZNF713*, *ITGAM*, *ZNF646*, *SETD1A*, *RNF40*, *ELOA*, *ACACA*, *HSD3B7*, *ORAI3*, *ZNF629*, *EPHB2*, *CCT6A*, *TYW1* and *ENSSSCG00000031037*
Intron	*VIM*, *DHRS11*, *TBC1D20*, *KCTD7*, *SLC5A2*, *ZNF629*, *CCT6A*, *LMO4*, and *ENSSSCG00000006928*
Downstream	*PHACTR2*, *DHRS11*, *ITGB2*, *CREB3L2*, *KCTD7*, *PRSS53*, *VKORC1*, *SETD1A*, *STX1B*, *BCL7C*, *RNF40*, *RPRD1A*, *ENSSSCG00000033141* and *ENSSSCG00000018553*
Upstream	*TRDMT1*, *HSD3B7*, *SETD1A*, *SUMF2*, *ENSSSCG00000033141* and *ENSSSCG00000018553*

3' UTR	*FYN, DHRS11, EOGT, ITGB2, CREB3L2, KCTD7, C16orF58, HSD3B7, ORAI3, BCL7C, ZNF629, IL18R1, IL1RL1, EPHB2, CLIC4, RPRD1A* and *CCT6A*
5' UTR	*ZNF713, BCKDK* and *ZNF629*

Variants with effects on proteins

The VEP tool in the Ensembl 98 uses the extension SIFT which is a value that predicts whether an amino acid substitution affects the function of the protein or not. The general status of the SIFT score obtained using VEP is summarized in Table 4.

From the exome analysis, according to the SIFT prediction, 21 variants of the SNP type were classified as missense, possibly altering the amino acid as a result from the substitution, and 5 variants were classified as deleterious in the following genes: solute carrier family 22 member 11 *(SLC22A11)* (0.03), *ZNF713* (0.02), *ITGAM* (0.01) and two in the *SPN* gene (0 and 0.05). Moreover, 13 SNPs were designated as tolerated with the SIFT score > 0.05 to 1.00 and 3 SNPs were classified as tolerated low confidence on the SET domain containing 1A, histone lysine methyltransferase *(SETD1A)* and olfactory receptor 7A17-like *(LOC106509673)* genes with SIFT scores ranging from 0.00 to 0.05 (Table 4).

When analyzing the RNA-Seq variants, 8 were classified as missense on chromosomes 3 and 6, located in 6 genes. There was only one deleterious variant according to the SIFT score located in the *ITGAM* gene and one classified as tolerated low confidence in the *SETD1A* gene. The other polymorphisms were classified as tolerated and were located in the *ZNF713*, zinc finger protein 646 (*ZNF646*), *RNF40* and *ELOA* genes (Table 4).

Table 4. Missense variants observed between normal and umbilical-hernia affected pigs with SIFT score calculated in the dbSNP database (Ensembl).

Approach	Existing variation	Location	Gene Symbol	SIFT
Exomic	rs792193084	2:7602943	*SLC22A11*	Tolerated (0.37)
	new	2:7602943	*SLC22A11*	Deleterious (0.03)
	rs337918521	3:16960969	*ZNF713*	Deleterious (0.02)
	rs323115420	3:16964045	*ZNF713*	Tolerated (0.65)
	new	3:17033960	*LOC106509673*	tolerated low confidence (0.07)
	new	3:17034293	*LOC106509673*	tolerated low confidence (0.8)
	rs327289001	3:17254444	*ITGAM*	Deleterious (0.01)
	rs337670844	3:17399477	*ZNF646*	Tolerated (0.1)
	rs330957838	3:17468302	*SETD1A*	tolerated low confidence (0.34)
	rs789266896	3:17628688	*RNF40*	Tolerated (0.62)
	new	3:17979825	*SPN*	Deleterious (0.01)
	new	3:17979826	*SPN*	Deleterious (0.05)
	rs325089032	6:81571496	*ELOA*	Tolerated (0.1)
	rs696812713	8:6157581	*OTOP1*	Tolerated (0.58)
	rs712855168	8:6157582	*OTOP1*	Tolerated (0.53)
	rs335136145	12:38026003	*MYO19*	Tolerated (0.36)
	new	12:55146905	*MYH8*	Tolerated (0.51)
	new	12:55164180	*MYH8*	Tolerated (0.1)
	new	12:55197434	*MYH2*	Tolerated (0.1)
	new	12:55267170	*MYH2*	Tolerated (0.1)
	rs319222050	12:55356662	*MYH3*	Tolerated (0.19)
Transcriptomics	rs323115420	3:16964045	*ZNF713*	Tolerated (0.65)
	rs327289001	3:17254444	*ITGAM*	Deleterious (0.01)
	rs337670844	3:17399477	*ZNF646*	Tolerated (0.1)
	rs330957838	3:17468302	*SETD1A*	tolerated low confidence (0.34)
	rs789266896	3:17628688	*RNF40*	Tolerated (0.60)
	rs325089032	6:81571496	*ELOA*	Tolerated (0.1)

SIFT score varies from 0 to 1; Deleterious: SIFT score <=0.05; Tolerated: SIFT score >0.05.

Polymorphisms and common genes identified with the exome, transcriptome and GWAS approaches

Comparing the exome with the umbilical ring transcriptome [13], 47 genes were unique in the exome, 19 genes were unique in the transcriptome and 23 genes were common to both approaches (Fig. 5). Moreover, 79 identical variants were found with both methodologies (additional file 3). From these, 12 variants were located in 3 genes of the zinc finger protein family *(ZNF629, ZNF646* and *ZNF713),* 11 variants in 5 genes of anti-inflammatory response and immune system *(ITGAM, RNF40, BCKDK, PRSS53* and *DHRS11),* 5 in the *SETD1A* gene, 10 variants in 3 genes of amino acid metabolism *(ACACA, HSD3B7* and *SUMF2),* 14 variants in 3 genes of blood pressure and size *(TYW1, CCT6A* and *EPHB2),* 1 polymorphism in the *ELOA* gene, 4 variants in 3 genes of transport of molecules *(VKORC1, CLIC4* and *STX1B),* 16 polymorphisms in 2 genes of calcium and potassium modulator *(ORAI3* and *KCTD7),* and 6 variants in 2 new genes *ENSSSCG00000018553* and *ENSSSCG00000033141* (additional file 3). Furthermore, when comparing the results from the exome/transcriptome with those of GWAS from Lagos, (in preparation), we have 8 genes in common (Fig. 5, Table 5).

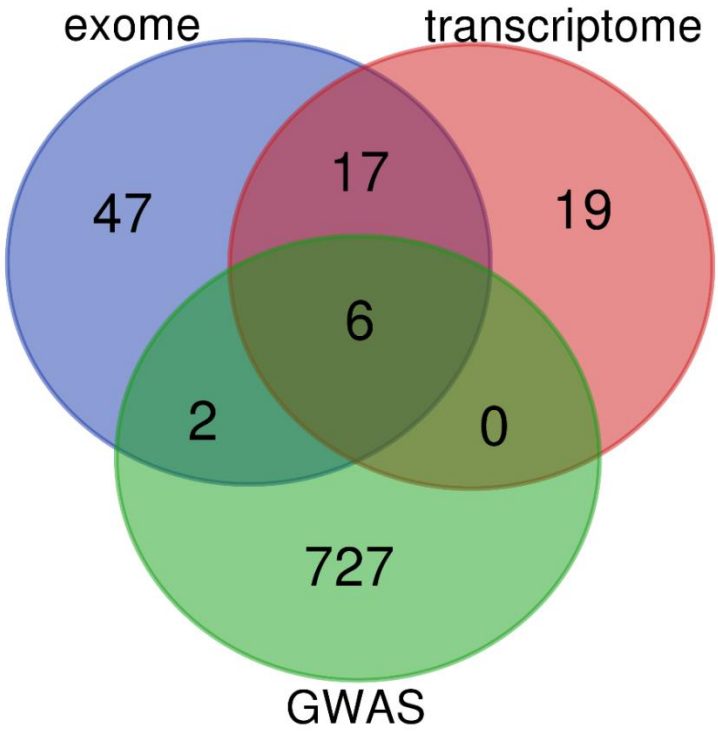


Figure 5. Venn diagram grouping all 72 genes obtained through the exome, 42 obtained from the transcriptome and 735 obtained from the results of a GWAS, indicating the number of common genes among the three methodologies.

Table 5. Common candidate genes for umbilical hernia identified with the three different methodologies

Methodology	**Gene Symbol**
Exomic/ Transcriptomics	*EPHB2, TYW1, SUMF2, ITGAM, BCKDK, PRSS53, ORAI3, ZNF629, SETD1A, RNF40, DHRS11, ACACA, ENSSSCG00000018553, CCT6A, STX1B, ELOA, ZNF646, ZNF713, HSD3B7, ENSSSCG00000033141, VKORC1* and *KCTD7*
GWAS	*ITGAM, BCKDK, PRSS53, STX4, STX1B, ZNF646, VKORC1* and *ENSSSCG00000049720*

Functional analysis and gene enrichment

The gene ontology analysis carried out with all 91 genes obtained through the analysis of the exome and transcriptome and the 8 common genes from the GWAS was performed in the DAVID database. Forty-eight biological processes (additional file 4) were found, which were summarized in 6 superclusters with the REViGO tool: protein localization to membrane, cardiac septum development, cellular response to chemical stimulus, locomotion, metabolism and organic substance metabolism (Fig. 6).

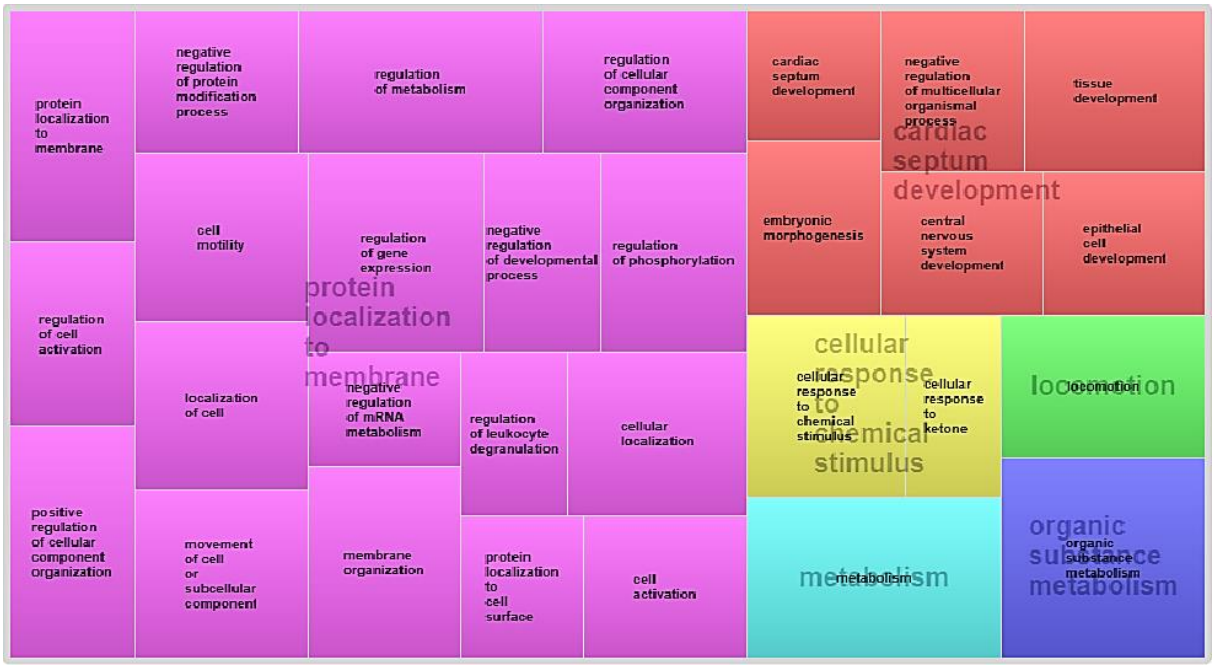


Figure 6. Superclusters of biological processes enriched by genes involved with umbilical hernia found in the exome, transcriptome and GWAS approaches. Different colors show different superclusters and the size of each box is determined by the uniqueness of the categories.

To elucidate the interaction between genes and BPs, the 91 genes obtained through the three methodologies were used to build a network with the NetworkAnalyst tool using the KEGG database (Kyoto Encyclopedia of Genes and Genomes). A total of 10 BPs was found including 21 genes (Fig. 7). Immune response and muscle contraction were some of the most enriched BP, grouping genes that interact with natural killer cell activation and cell-matrix adhesion.

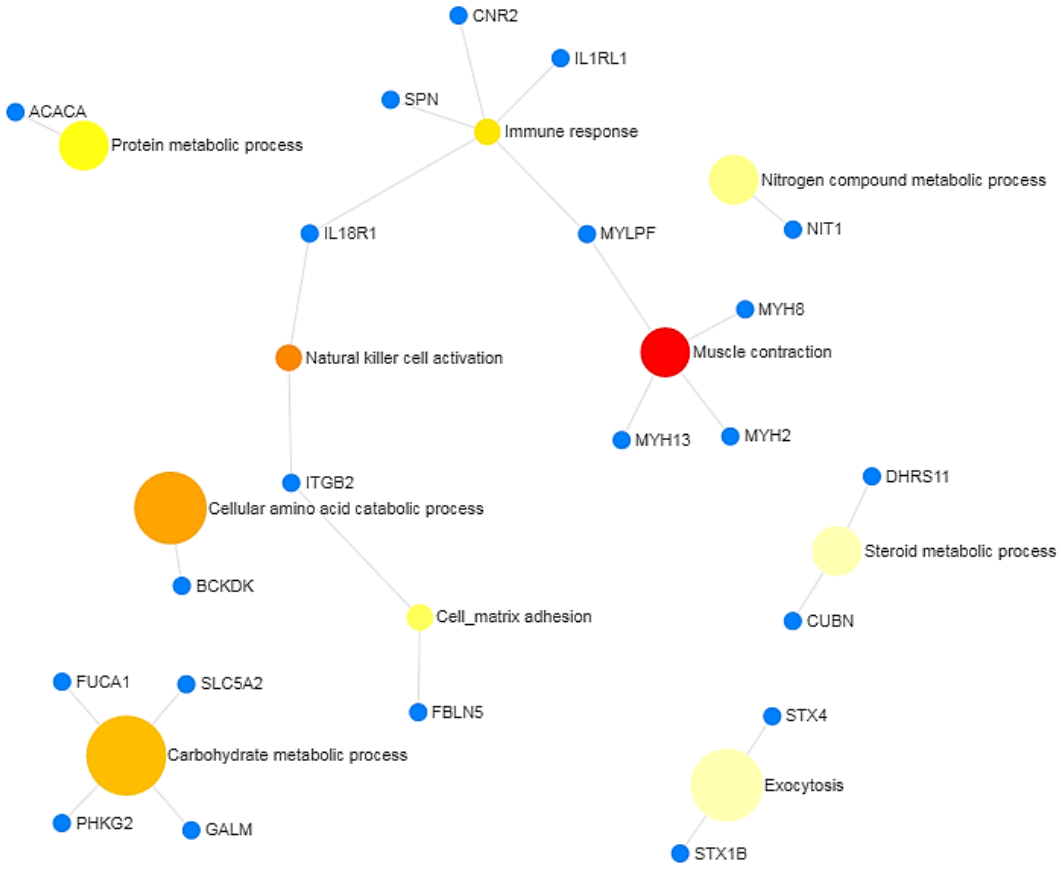


Figure 7. Bipartite gene network of genes and biological processes related to umbilical hernia in pigs built with the NetworkAnalyst tool using the KEGG database (Kyoto Encyclopedia of Genes and Genomes). The yellow and red nodes indicate the number of predicted genetic interactions, where yellow nodes indicate low and red nodes indicate high direct interaction with genes. Hub genes are highlighted in blue.

The central group, including the genes *MYLPF*, *MYH13*, *MYH2*, *MYH8* and *MYH3*, are mainly involved in the formation of myosin and differentiation of actin and myosin filaments (Fig. 7). When analyzing the exome and transcriptome, several variants were identified in SSC12, located in some genes of the myosin family, responsible for muscle contraction (Table 1). Among the main variants, the SNP rs335136145 identified in the *MYO19* gene and located in a missense exon region (additional file 1) can be highlighted. In the *MYH13* gene, two SNPs (rs341831793 and rs346367824) were identified in synonymous regions (additional file 1). In the *MYH8* gene, four new variants were identified and in the *MYH2* gene, five new variants were found (additional file 1).

Moreover, three new variants were identified in intronic regions of the *MYLPF* gene (additional file 2).

Another group of genes was related to immune response, natural killer cell activation and cell-matrix adhesion: Integrin subunit beta 2 (*ITGB2),* Interleukin 1 receptor-like 1 *(IL1RL1*), Interleukin 18 receptor 1 *(IL18R1*) *SPN, CNR2,* and *FBLN5* (Fig. 7). Several variants were found in the exome and transcriptome analyses on SSC3, 6, 7 and 13 (additional files 1 and 2). Among the main variants, three new variants were identified only in the exome, in the SPN gene located on SSC3 (additional file 1). The *CNR2* gene is located on SSC6 and two SNPs (rs345825612 and rs332149968) were identified in this gene (additional file 1). Another gene identified only in the exome was the *FBLN5,* located on SSC7, with a new InDel type variant identified in an intronic region of this gene (additional file 1). In the transcriptome, two new variants were identified in the *ITGB2* gene located on SSC13 (additional file 2). Moreover, a new polymorphism was identified in the regulatory region 3'UTR of the *IL1RL1* and *IL18R1* genes in SSC3 (additional file 2).

Also, a SNP (rs331463738) was identified in the 3 'UTR region of the *DHRS11* gene on SSC12, which is responsible for oxidoreductase activity and coenzyme binding, and a new variant located on SSC10 was identified in an intronic region of the Cubilin (*CUBN)* gene, which is responsible for steroids metabolic process (Fig. 7). Moreover, other important genes have also been found interacting with BPs: the carbohydrate metabolic process included the alpha-L-fucosidase (*FUCA1),* galactose mutarotase *(GALM),* phosphorylase kinase catalytic subunit gamma 2 *(PHKG2)* and solute carrier family 5 member 2 *(SLC5A2)* genes (Fig. 7). The protein metabolic process involved the *ACACA* gene. The cellular amino acid catabolic process included *BCKDK* and the exocytosis BP involved the syntaxin 4 (*STX4)* and *STX1B* genes (Fig. 7). In addition, we observed through the NetworkAnalyst tool the gene-gene interactions with the 91 genes identified (Fig. 8). Therefore, some of the genes, such as *MMS19, FYN, VKORC1* and *VIM* could be considered hubs, since they are involved in the expression of several other genes in the network (Fig. 8).

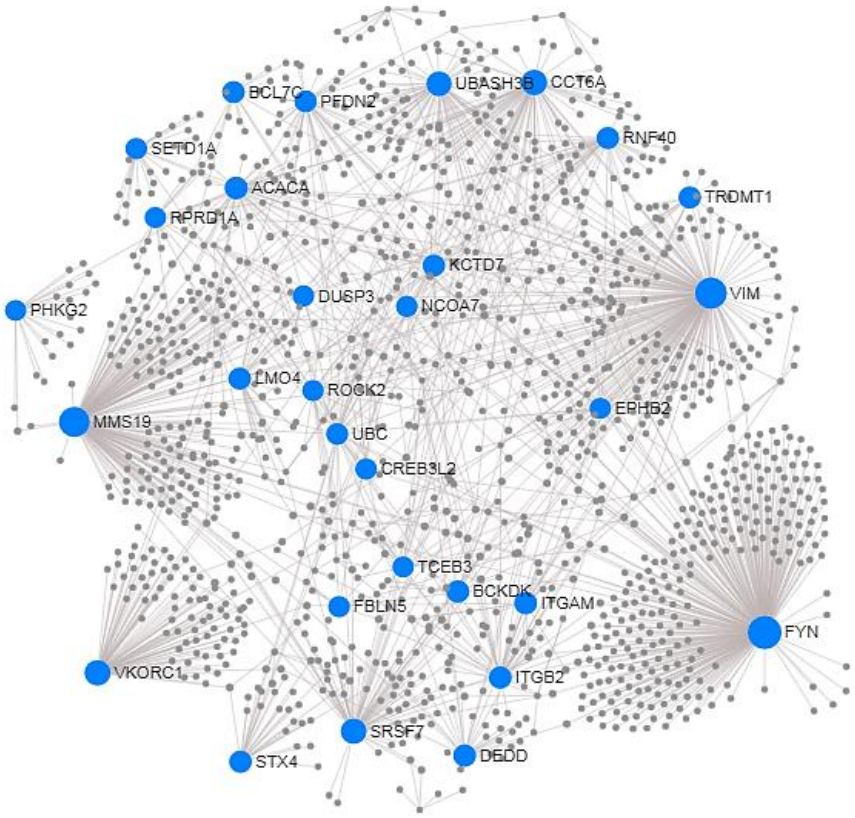


Figure 8. Gene enrichment analysis and biological processes related to umbilical hernia in pigs using the NetworkAnalyst tool. Nodes indicate the number of predicted gene interactions. Strong and large circles contain high number of genes. Blue circles are the genes identified in this study and grey circles are gene interactions.

Discussion

The knowledge regarding genes that act in the UH formation is still scarce. Understanding the role and function of the variants in genes is extremely important since they can alter their function and expression and, as a consequence, can directly influence the formation of some anomalies, such as UH [4,24]. Despite the importance of this problem in pig farming, the knowledge in this field is still limited, although some previous studies using different methodologies pointed out some candidate genes and important biological processes related to umbilical hernia in pigs [11,12,13]. To fill this gap, we have generated the first WES of healthy and UH-affected pigs. Moreover, this is the first study integrating three methodologies: WES, RNA-Seq and GWAS to identify putative variants and genes involved with the development of UH in Landrace pigs.

Several genetic and environmental factors contribute to the formation and development of UH [4,5]. Some environmental factors like infection of the umbilical cord during birth [2] contribute to the development of hernias. Furthermore, it is known that the weakness of muscle tissue around the umbilical area interferes in the closing of the umbilical canal allowing the intestinal loops to project through the abdominal wall [4,25]. However, the genetic control of UH is still unknown. A hereditary cause has been suggested by [7], who carried out a progeny test in purebred pigs and showed that the chance of finding a pig with umbilical hernia is different between breeds. Some studies [12,26,27] revealed that the development of UH is a polygenic trait, which justifies that just a few candidate genes have already been reported for UH [6]. Therefore, in the current study, polymorphisms and genes related to the formation of UH were identified. Using the three approaches, exome sequencing, RNA-Seq and GWAS, allowed us to identify different genes and variants related to UH with a great confidence. A total of 209 variants were identified differing between normal and UH-affected pigs from the exome sequencing (additional file 1), which were located in 72 genes (Table 1). In the transcriptome, 81 variants were identified (additional file 2) in 42 genes (Table 2), with 79 of these variants being concordant with the exome, comprising 23 genes common to both approaches. In addition, 8 genes identified as involved with UH in both the exome and transcriptome from a Landrace purebred were also found suggestively associated to UH in a GWAS carried out in a commercial pig population. To better clarify the UH etiology, the identification of genes and biological processes involved with this disorder is essential. From the 48 BP found in our study (Fig. 6, Fig. 7), muscle contraction, immune response, and cell-matrix adhesion are highlighted. Altogether, our results allowed a close observation of the relationship of the genes with these BP, as further discussed.

Muscle Contraction related to Umbilical Hernia

It is suggested that the muscle tissue plays an important role in the development of umbilical hernia [2]. In children, it has been reported that the muscles of the umbilical region influence the development of the umbilicus [29]. Xu et al. [30] deduced by histologically examining human fetuses with 8 to 40 weeks of age that muscle contraction probably plays a critical role in closing the umbilical ring after birth, tracing a strong

correlation between the umbilicus and the abdominal wall. In our study, several polymorphisms (additional file 1, additional file 2) were identified in genes related to the muscle contraction BP using both the exome and transcriptome methodologies (Fig. 6, Fig. 7). The *MYLPF, MYH13, MYH2, MYH3, MYH8, MYO19* and *ROCK2* genes are responsible for the formation of myosin, actin filaments and protein kinase, skeletal muscle development and muscle contraction [31, 32, 33, 34, 35].

The *MYLPF* had three new mutations found in its 1/7 intron region (additional file 1). *MYLPF* acts in the calcium ion binding, metal ion binding and structural constituent of muscle [36]. Wang et al. [37] produced healthy and knockout mice for *MYLPF* and observed that the lack of *MYLPF* transcription causes a complete lack of skeletal muscle during development. Therefore, three new variants in introns can interfere in the alternative splicing process of this gene [37], which can act in the generation of different transcripts of this gene, influencing the healthy formation of muscle tissue [36] This happens because the myosin molecule without its subunit-containing two heavy chains can be degraded by the cellular mechanism to remove incorrectly folded proteins during embryogenesis [36]. Another possibility is that the myosin heavy chains were unable to form thick filaments; therefore, the lack of thick filaments would prevent healthy myofibrillogenesis leading to functional loss of cell and muscle tissue [37].

Several variants have been identified in the myosin family: four polymorphisms were identified in regions of exon (synonymous variant) and intron of the *MYH13* gene (additional file 1). This gene is a fundamental component for the microfilament motor activity and actin-binding [32]. Three variants were found in the *MYH3* gene, two located in exons of the type missense and synonymous variant and a new deletion located in the region upstream of the gene (additional file 1). *MYH3* function includes nucleotide binding, motor activity and protein binding [33]. In the *MYH2* gene, another member of the myosin family, five new variants located in regions of exons (missense and synonymous variant) and introns were identified (additional file 1). The *MYH2* is responsible for the generation of mechanical force in eukaryotic cells and skeletal muscle contraction [32]. Moreover, in the *MYH8*, we identified four new polymorphisms also in regions of exon (missense and synonymous variant) and intron (additional file 1). This gene is predominantly expressed in fetal skeletal muscle [33].

In the mammalian genome, myosin is composed by 16 genes, encoding proteins expressed in muscle and non-muscle tissues [32]. In our results, we identified variants in five myosin genes. Xu et al. [30] identified several genes of the myosin family in skeletal muscle in humans and observed that changes in the expression of this family of genes interfere with muscle contraction. Here, the results indicate that the variants found in myosin genes can be strong candidates to trigger UH in pigs, because the musculature, in particular muscular contraction, is extremely important to prevent the passage of the intestinal loops through the umbilical ring causing UH [28]. Furthermore, genes from the myosin family, the 1/3 myosin light chain skeletal muscle isoform (*MYL1)* and myosin light chain 3 (*MYL3)* have been described as candidate genes to the development of scrotal hernia in pigs [38], emphasizing the importance of muscle contraction in the development of hernias.

Moreover, a variant was identified in an exomic region (missense) of the *MYO19* gene (additional file 1), which is also responsible for ATP binding and actin-binding [30]. Finally, on SSC3, a new deletion in a downstream regulatory region of the *ROCK2* gene was identified (additional file 1). This gene is involved in regulation of smooth muscle contraction, actin cytoskeleton organization, stress fiber and focal adhesion formation [35]. *ROCK2* is a key regulator of the actin cytoskeleton that acts in the formation of the actin/myosin filaments [22]. Human studies indicate that *ROCK2* promotes cancer growth, in addition to degrading MMP2 [39]. In mouse, the lack of this gene can cause cardiac hypertrophy [40]. Furthermore, we identified a SNP (rs330195537) in the *VIM* gene (additional file 1), involved with the formation of the intermediate class III filaments found in several non-epithelial cells that make up the actin microtubules and microfilaments, which compose the cytoskeleton [41] and that is associated with *ROCK2*. Therefore, this new mutation in the *ROCK2* gene associated with *VIM* could impair the formation of the actin/myosin filaments, preventing the complete formation of muscle tissue around the navel, leaving this region flaccid and susceptible to the formation of UH.

Cell-matrix adhesion

The cell adhesion BP is related to the formation of UH since it is responsible for cell connections, cell adhesion, tissue development and maintenance, cell differentiation, migration and communication [42]. Some genes identified in the current study were classified in BPs in pigs, such as cell adhesion (Fig. 7), cell motility (GO: 0048870), cell activation (GO: 0001775) and cell migration (GO: 0016477) (Additional file 4), for example, the ubiquitin associated and SH3 domain containing B (*UBASH3B*), LIM domain only 4 (*LMO4*), *RABGEF1*, *FYN*, *ITGB2*, *ROCK2*, *FBLN5* and *EPHB2* (additional file 1, additional file 2).

Most of these genes have functions that can be related to the herniation process. The *RABGEF1* gene acts on cell recycling, also acting as an ubiquitin ligase by similarity [43]. The *UBASH3B* gene can inhibit the endocytosis of the epidermal growth factor receptor (EGFR) and promote the accumulation of activated target receptors, T cells and EGFR on the cell surface [44]. The *EPHB2* is involved in several cellular processes, including cell motility, division and differentiation [45]. The *LMO4* gene encodes a protein that plays a role as a regulator of transcription [46]. The *FBLN5* promotes the adhesion of endothelial cells through the interaction of integrins (*ITGB2/ITGAM)* [47]. The *FBLN5* and *ITGB2/ITGAM* play a very important role in the cell adhesion process promoting cell binding. Mutations in these genes were associated with delayed growth and development of the tissues and, also, the Fibulin can be regulated by growth factors [48].

In the current study, the *RABGEF1* presented one SNP (rs713010025) and the *UBASH3B* another SNP (rs345798145), both located in upstream regulatory regions. These genes are involved with increasing levels of epithelial growth factors (EGF) and transforming growth factor (TGF) [49]. Studies have indicated that when there is injury, epithelial cells, macrophages and fibroblasts produce growth factors such as EGF and TGF to heal the injury [50, 51]. Further, in the injury site, there is an increase in the rates of healing and regeneration of the composite fibrous tissue by Fibulin [52,53]. However, when these growth factors are not balanced, the problem becomes chronic [54]. Probably, those SNPs in the *RABGEF1* and *UBASH3B* genes could modulate the expression of growth factors, since these variants are located in upstream regulatory regions. Therefore, they may trigger an immune disorder in the umbilical ring tissue, favoring the defense cells to attack the tissue itself, thus, favoring the occurrence of umbilical hernias in pigs.

Immune Response

In our study, several genes were enriched in biological processes related to the immune system (additional file 4). Also, the DAVID database indicated some significant PBs: regulation of immune response, with the main genes *ITGAM, ITGB2* and *FYN*, regulation of leukocyte degranulation (*STX4, SPN*) and regulation of myeloid leukocyte mediated immunity (*IL1RL1, IL18R1*) (Fig. 6, Fig. 7).

The *ITGAM* and *ITGB2* genes are involved in several receptor interactions of monocytes, macrophages and granulocytes [53]. Integrin *ITGAM/ITGB2* is also a fibrinogen receptor and regulates the migration of neutrophils [55]. The *FYN* gene is responsible for regulating cell growth and survival, cell adhesion, integrin-mediated signaling and cytoskeletal remodeling [56], while *STX4* acts on the coupling of transport vesicles [57]. *SPN* is predominant on the surface of the leukocyte cell, directly influencing T cells, regulating their proliferation, differentiation and migration [58]. The *IL1RL1/IL18R1* are receptors that are induced by pro-inflammatory stimuli, helping in the connection between tissue and auxiliary T cells [59].

Umbilical disorders, including omphalophlebitis, present a significant challenge to the health and well-being of a newborn [60]. Omphalophlebitis is an inflammation or infection of the umbilical region [61], which is the main cause of abscesses [62]. In pig farming, this condition can be caused by mismanagement of cutting or cleaning the umbilical cord [63,64]. The omphalophlebitis can develop in animals with compromised immune systems and concomitant health problems [65]. Several polymorphisms in genes related to the immune system were identified in our study, pointing to an ongoing inflammation process. Therefore, the *FYN*, *SPN*, *IL1RL1/IL18R1*, *STX4*, *ITGAM* and *ITGB2* genes might be strong candidates for the development of UH in pigs.

Genes located in QTL regions

Some studies using different approaches have already been developed seeking to identify QTLs for umbilical hernia in pigs [10, 12, 13]. We identified through the exome analysis that the papillin proteoglycan like sulfated glycoprotein (*PAPLN*) gene was located in a QTL region already described in the literature for umbilical hernias in pigs [10].

The PAPLN is a component of the extracellular matrix (66), widely studied in humans, and it has been related to liver diseases and growth [66, 67]. It is one of the main glycoproteins in the extracellular matrix [67, 68]. Suppression of Papillin in Drosophilas has already been associated with embryonic death during embryogenesis due to disorders and abnormalities in muscle formation and Malpighi tubules malformation [68]. Moreover, in humans, the Papillin is known to have an important role in the modulation of metalloproteins during organogenesis, acting directly in the differentiation of the three germ leaflets [67]. This indicates that this gene can interfere in the differentiation of ectoderm, mesoderm and endoderm for the formation of organs. However, there is not much information about the performance of this gene in pigs and this is the first time that this gene has been related to the development of UH.

Variants and candidate genes that cause UH

Combining the results from the three methodologies (Fig. 5) and those from the VEP tool using the SNPdatabase, we evaluated the consequences of the polymorphisms found when they were predicted to affect the production of amino acids. Twenty-one variants in the exome and six variants in the transcriptome were classified with the SIFT score (Table 3). Polymorphisms in the regulatory and coding regions of the genome may be implicated in the development of diseases and congenital problems. Generally, non-synonymous SNPs, such as missense type variants, lead to amino acid changes in protein products, in which they represent approximately half of the known genetic problems responsible for human hereditary diseases [69]. The *ITGAM, SPN, MYH8, MYH2, MYH3* and *MYO19* genes were identified in the three methodologies. Some variants located in these genes were identified in the DNA (genomic) and were also expressed in the mRNA of the umbilical ring tissue from normal and UH-affected pigs. As previously discussed, integrin and sialophorin are important components of the immune system and cell adhesion [55,58]. Therefore, this shows that some variants, mainly of the missense type, that alter proteins are strong candidates as factors predisposing the occurrence of UH. These variants and many others have been identified in many genes present in several BP; some of them already discussed above are key genes for triggering UH. These results show that the variants found in our study are fundamental pieces for understanding the etiology of the UH.

Our results have shown that many of these genes were associated with the development of UH for the first time. The polymorphisms identified in these genes are strong candidates for the development of UH in pigs. However, further studies are required to validate the function of these groups of polymorphisms and genes in the development of UH. Nevertheless, the combination of the three methodologies used greatly improved the reliability of our results, providing the discovery of variants possibly involved with the onset of UH and the paths to understand the umbilical hernia development.

Conclusions

We have generated the first exome sequencing related to normal and umbilical hernia-affected pigs, which allowed us to identify several variants involved with this disorder. Moreover, comparing these variants with the results from RNA-Seq and GWAS, it was possible to identify variants and genes that had not been previously related to the development of umbilical hernia in pigs. Muscle contraction, immune response and cell-matrix adhesion were the main active biological processes related to the umbilical hernia occurrence. Therefore, the variants found in the *FYN, SPN, ITGB2, MYLPF, MYH13, MYH8, MYH2, MYH3, MYO19, VIM, ROCK2, RABGEF1* and *UBASH3B* genes can be highlighted as strong candidates to the development of UH in pigs. These results contribute to better understand the genetic mechanisms involved with the occurrence of UH in pigs and possibly in other mammals, including humans.

Abbreviations

UH: umbilical hernia
BP: biological processes
CEUA: Ethical Committee for Animal Use
VEP: variant effect predictor
QTL: Quantitative Trait Loci
GWAS: Genome-Wide Association Study
SNP: single nucleotide polymorphism
InDel: insertion and deletion

QC: quality control
GO: Gene Ontology
SIFT: Sorting Intolerant From Tolerant
NGS: Next Generation Sequencing
WES: Whole-Exome Sequencing

Declarations

Ethics approval: All animal procedures were performed in accordance with the Ethics Committee on Animal Utilization from the Embrapa Swine and Poultry National Research Center under protocol number 011/2014.

Consent for publication: Not applicable

Availability of data and material: The datasets used or analyzed during the current study are available from the corresponding author on reasonable request. The transcriptome sequences are available in the SRA database with BioProject number PRJNA352962.

Competing interests: The authors declare that they have no competing interests.

Funding: This study was funded by project #476146/2013-5 from the National Council of Scientific and Technological Development (CNPq) from the Brazilian Government. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Authors' contributions: AMGI, JOP, RZ and MCL conceived and designed the experiment. JSL provided the animals used in this study. IRS, AMGI, JOP, MCL and MAZM were responsible for the data collection. MEC and AMGI performed the exome and transcriptome analysis. EBL and MPP performed the genome-wide association analysis. IRS and AMGI performed the variant analyses. AMGI, IRS and EBL performed the functional analyses of the genes. AMGI, IRS, JOP, MPP and MCL interpreted the

results and evaluated the conclusions. IRS wrote the manuscript. All authors reviewed, edited and approved the final manuscript.

Acknowledgements: The authors are grateful to Alexandre L. Tessmann for technical assistance. IRS was sponsored by a PROMOP/Udesc scholarship. MPP was a post-doctoral ICASA/SC fellow at the Embrapa Swine and Poultry National Research Center. MCL is recipient of a productivity fellowship from CNPq. We thank the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior-Brasil (CAPES-Finance Code 001) for the free access to the journals used in the literature review.

Additional Files

Additional file 1: Total number of variants (SNPs and InDels) identified in the whole-exome sequencing that differed between normal and umbilical hernia-affected pigs and its annotation and consequence predicted with the VEP tool.

Additional file 2: Total number of variants (SNPs and InDels) identified in the transcriptome analysis that differed between normal and umbilical hernia-affected pigs its annotation and consequence predicted with the VEP tool.

Additional file 3: Polymorphisms, genes and regions associated to umbilical hernia in pigs annotated in the exome, transcriptome and GWAS analyses.

Additional file 4: GO biological process enriched based on the candidate genes identified for umbilical hernia through the analyses of the whole exome, RNA sequencing and GWAS data.

References

1. Petersen, H. H., Nielsen, E. O., Hassing, A. G., Ersbøll, A. K. & Nielsen, J. P. Prevalence of clinical signs of disease in Danish finisher pigs. **Vet. Rec**. 162, 377–382, 2008.
2. Straw, B., Bates, R. & May, G. Anatomical abnormalities in a group of finishing pigs: Prevalence and pig performance. **Journal Swine Health Production**. 17, 28–31, 2009.

3. Yun, J., Olkkola, S., Hänninen, M.-L., Oliviero, C. & Heinonen, M. The effects of amoxicillin treatment of newborn piglets on the prevalence of hernias and abscesses, growth and ampicillin resistance of intestinal coliform bacteria in weaned pigs. **PLoS One** 12, e0172150, 2017.
4. Brandt, M. L. Pediatric Hernias. **Surgical Clinics of North America** 88, 27–43, 2008.
5. Franz, M. G. The Biology of Hernia Formation. **Surgical Clinics of North America** 88, 1–15, 2008.
6. Rutten-Ramos, S. C. & John, D. Association between umbilical hernias and genetic line in a swine multiplication herd and methods to differentiate the role of sire in the incidence of umbilical hernias in offspring. **Journal Swine Health Production**. 14, 317–322, 2006.
7. Searcy-Bernal R., Gardner, I. A. Hird, D.W. Effects of and factors associated with umbilical hérnias in a swine herd. **Journal of the American Veterinary Medinal Association.** vol. 4, no.10, p. 1660–1664, 1994.
8. Liao, X. J. *et al.* Susceptibility Loci for Umbilical Hernia in Swine Detected by Genome-Wide Association. **Genetika.** 51, 1163–1170, 2015.
9. Thaller, G., Dempfle, L. & Hoeschele, I. Investigation of the inheritance of birth defects in swine by complex segregation analysis. **Journal Animal Breeding Genetics**. 113, 77–92, 1996.
10. DING, N. et al. A Genome-Wide Scan Reveals Candidate Susceptibility Loci for Pig Hernias in an Intercross Between White Duroc and Erhualian. **Journal of Animal Science.** V. 87 p. 2469-2474, 2009.
11. Fernandes, L. T. *et al.* Novel putative candidate genes associated with umbilical hernia in pigs. **Proceedings of the World Congress on Genetics Applied to Livestock Production** 11, 743–749, 2018.
12. Grindflek, E., Hansen, M. H. S., Lien, S. & Van Son, M. Genome-wide association study reveals a QTL and strong candidate genes for umbilical hernia in pigs on SSC14. **BMC Genomics** 19, 412, 2018.
13. Souza, M. R. *et al.* Transcriptome analysis identifies genes involved with the development of umbilical hernias in pigs. **PLoS One** 15, 2020.

14. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: A flexible trimmer for Illumina sequence data. **Bioinformatics** 30, 2114–2120, 2014.
15. Heng, L. & Durbin, R. Fast and accurate short read alignment with Burrows–Wheeler transform . **Bioniformatics** 25, 1754–1760, 2009.
16. McKenna, A. *et al.* The genome analysis toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. **Genome Research**. 20, 1297–1303, 2010.
17. McLaren, W. *et al.* Deriving the Consequences of Genomic Variants With the Ensembl API and SNP Effect Predictor - **Nature**. 26, 2069–2070, 2010.
18. Vaser, R., Adusumalli, S., Ngak Leng, S., Sikic, M. & Ng, P. C. SIFT missense predictions for genomes. **Nature Protocls.** 11, 1073–1081, 2015.
19. Lopes, M. C. *et al.* A Combined Functional Annotation Score for Non-Synonymous Variants. **Human Hereditary**. 73, 47–51, 2012.
20. Dobin, A. *et al.* STAR: ultrafast universal RNA-seq aligner. **Bioinformatics** 29, 15–21, 2012.
21. Mailund, T., Besenbacher, S. & Schierup, M. H. Whole genome association mapping by incompatibilities and local perfect phylogenies. **BMC Bioinformatics** 7, 454, 2006.
22. Huang, D. W. Brad, T. S. & Lempicki, R. A. Bioinformatics Enrichment Tools: Paths Toward the Comprehensive Functional Analysis of Large Gene Lists **Nucleic Acids Research.** 37, 1–13, 2009.
23. Zhou, G. *et al.* NetworkAnalyst 3.0: a visual analytics platform for comprehensive gene expression profiling and meta-analysis. **Nucleic Acids Research**. 47, W234–W241, 2019.
24. Franz, M. G. The biology of hernias and the abdominal wall. **Hernia** 10, 462–471, 2006.
25. Fall, I. *et al.* Strangulated umbilical hernias in children. **Pediatric Surgery International.** 22, 233–235, 2006.
26. Sevillano, C. A. *et al.* Genome-wide association study using deregressed breeding values for cryptorchidism and scrotal/inguinal hernia in two pig lines. **Genetics Selection and Evolution**. 47, 18, 2015.

27. Grindflek, E. *et al.* Genome-wide linkage analysis of inguinal hernia in pigs using affected sib pairs. **BMC Genetics**. 7, 25, 2006.
28. Franz, M. G. *et al.* Fascial incisions heal faster than skin: A new model of abdominal wall repair. **Surgery** 129, 203–208, 2001.
29. Meier, D. E. *et al.* Incidence of umbilical hernia in african children: Redefinition of 'normal' and reevaluation of indications for repair. **World Journal Surgery**. 25, 645–648, 2001.
30. Xu, D. *et al.* Umbilicus and the rectus sheath: a study using human fetuses. *Surg.* **Radiology Anatomy.** 42, 461–471, 2020.
31. Wang, Y. *et al.* Fast skeletal muscle regulatory light chain is required for fast and slow skeletal muscle development. **FASEB Journal**. 21, 2205–2214, 2007.
32. Mascarello, F. *et al.* Expression and identification of 10 sarcomeric MyHC isoforms in human skeletal muscles of different embryological origin. Diversity and similarity in mammalian species. **Annals of Anatomy.** 207, 9–20, 2016.
33. Agarwal, M. *et al.* Myosin heavy chain-embryonic regulates skeletal muscle differentiation during mammalian development. **Development.** 147, 2020.
34. Wu, L. *et al.* Integrative transcriptomics and proteomic analysis of extraocular muscles from patients with thyroid-associated ophthalmopathy. **Experience Eye Research**. 193, 107962, 2020.
35. Shimizu, Y. *et al.* ROCK-I regulates closure of the eyelids and ventral body wall by inducing assembly of actomyosin bundles. **Journal Cell Biology.** 168, 941–953, 2005.
36. Kondo, T. *et al.* Characterization of myosin II regulatory light chain isoforms in HeLa cells. **Cytoskeleton.** v. 72, 609-620, 2015.
37. Wang, Y. *et al.* Fast skeletal muscle regulatory light chain is required for fast and slow skeletal muscle development. **FASEB Journal**. v. 21, 2205-2214, 2007.
38. Romano, G. de S. *et al.* Inguinal ring RNA sequencing reveals downregulation of muscular genes related to scrotal hernia in pigs. **Genes.** 11, 2020.
39. Zanin-Zhorov, A. & Waksal, S. D. ROCKing cytokine secretion balance in human T cells. **Cytokine** 72, 224–225, 2015.
40. Okamoto, R. *et al.* FHL2 prevents cardiac hypertrophy in mice with cardiac-specific

deletion of ROCK2. **FASEB Journal.** 27, 1439–1449, 2013.

41. Borra, P. S. *et al.* Structural and computational investigations of VIM-7: Insights into the substrate specificity of VIM metallo-β-lactamases. **Journal Molecular Biology.** 411, 174–189, 2011.
42. Johnson, A. *et al.* **Biologia Molecular da Célula**. 6° edição, 2017.
43. Marichal, T. *et al.* Guanine nucleotide exchange factor RABGEF1 regulates keratinocyte-intrinsic signaling to maintain skin homeostasis. **Journal Clinical Investigation.** 126, 4497–4515, 2016.
44. Krupina, K. *et al.* Ubiquitin Receptor Protein UBASH3B Drives Aurora B Recruitment to Mitotic Microtubules. **Development Cell** 36, 63–78, 2016.
45. Cissé, M. *et al.* Reversing EphB2 depletion rescues cognitive functions in Alzheimer model. **Nature** 469, 47–52, 2011.
46. Barban, N. *et al.* Genome-wide analysis identifies 12 loci influencing human reproductive behavior. **Nature Genetics**. 48, 1–7, 2016.
47. Lee, Y. & Schiemann, W. P. Fibulin-5 (FBLN5) initiates EMT and induces MMPs expression via β1 integrin/EGFR/ERK pathway. **Cancer Research** 73, A8–A15, 2013.
48. Loeys, B. *et al.* Homozygosity for a missense mutation in fibulin-5 (FBLN5) results in a severe form of cutis laxa. **Human Molecular Genetics.** 11, 2113–2118, 2002.
49. Okuda, S., Languino, L. R., Ruoslahti, E. & Border, W. A. Elevated expression of transforming growth factor-β and proteoglycan production in experimental glomerulonephritis. Possible role in expansion of the mesangial extracellular matrix. **Journal Clinical Investigation.** 86, 453–462, 1990.
50. Polosa, R. *et al.* Expression of c-erbB receptors and ligands in human nasal epithelium. **The Journal of Allergy and Clinical Immunology.** 106, 1124–1131, 2000.
51. Strutz, F. *et al.* Role of basic fibroblast growth factor-2 in epithelial-mesenchymal transformation. **Kidney International**. 61, 1714–1728, 2002.
52. Boyer, B., Vallés, A. M. & Edme, N. Induction and regulation of epithelial-mesenchymal transitions. **Biochemical Pharmacology**. 60, 1091–1099, 2000.
53. Blanco, F. J. *et al.* Genome-Wide Transcriptional and Functional Analysis of

Endoglin Isoforms in the Human Promonocytic Cell Line U937. **Journal Cell Physiology**. 230, 947–958, 2015.

54. Zavadil, J. *et al.* Genetic programs of epithelial cell plasticity directed by transforming growth factor-β. **Proceedings of the National Academy of Sciences.** 98, 6686–6691, 2001.
55. Yuferov, V. *et al.* Oxycodone Self-Administration Induces Alterations in Expression of Integrin, Semaphorin and Ephrin Genes in the Mouse Striatum. **Frontiers Psychiatry.** 9, 257, 2018.
56. Short, B. Fyn regeneration: A Src family kinase and several other signaling pathways help repair zebrafish fins after wounding. **Journal Cell Biology.** 199, 189–189, 2012.
57. Tian, W., Ma, C., Liu, Y. & Xu, T. An efficient co-expression and purification system for the complex of Stx4 and C-terminal domain of Synip. **Biochemical Biophysical Research Community.** 371, 366–370, 2008.
58. Kovacheva, M., Zepp, M., Berger, S. & Berger, M. R. Bone sialoprotein is an essential target in breast cancer skeletal metastasis. **Cancer Research** 74, 2614–2614, 2014.
59. Reijmerink, N. E. *et al.* Association of IL1RL1, IL18R1, and IL18RAP gene cluster polymorphisms with asthma and atopy. **Journal of Allergy and Clinical Immunology,** v. 122, 2008.
60. Agarwal, N. *et al.* Appendicitis in paraumbilical hernia mimicking strangulation: a case report and review of the literature. **Hernia.** 17, 531–532, 2013.
61. Brown, S. B. F., MacDuff, E. & O'Dwyer, P. J. Abdominal wall fibromatosis associated with previous laparoscopic hernia repair. **Hernia.** 17, 669–672, 2013.
62. Yau, K. K., Siu, W. T. & Chan, K. L. Strangulated Appendix Epiploica in Paraumbilical Hernia: Preoperative Diagnosis and Laparoscopic Treatment. **Surgical laparoscopy endoscopy & percutaneous technique.** 16, 49–51, 2006.
63. Murphey, M. D. *et al.* Musculoskeletal Fibromatoses: Radiologic-Pathologic Correlation. **RadioGraphics**. 29, 2143–2176, 2009.
64. Jacob, B. P. *et al.* Tissue ingrowth and bowel adhesion formation in an animal comparative study: Polypropylene versus Proceed versus Parietex Composite.

Surgery Endoscopy Other Interv Techniques. 21, 629–633, 2007.

65. Elleau, C. Pathologie ombilicale en période néonatale. **Journal pédiatrie puériculture.** 28, 161–165, 2015.
66. Hotta, K. *et al.* Identification of core gene networks and hub genes associated with progression of non-alcoholic fatty liver disease by RNA sequencing. **Hepatology Research**. 47, 1445–1458, 2017.
67. Roselli, C. *et al.* Multi-ethnic genome-wide association study for atrial fibrillation. **Nature Genetics.** 50, 1225–1233, 2018.
68. Kramerova, I. A. *et al.* Papilin in development; a pericellular protein with a homology to the ADAMTS metalloproteinases . **Development.** 127, 5475–5485, 2000.
69. Pauline C. N & Steven H. SIFT: predicting amino acid changes that affect protein function . **Nucleic Acids Research**. 31, 3812–3814, 2003.

Additional file 1: Total number of variants (SNPs and InDels) identified in the whole-exome sequencing that differed between normal and umbilical hernia-affected pigs and its annotation and consequence predicted with the VEP tool.

SNP name	SSC	start	end	Allele (ref)	Allele (alt)	Gene	Symbol	Gene name	Consequence (transcript)
rs332203161	1	3427322	3427322	T	C	-	new		intergenic variant
rs707913396	1	3427328	3427328	T	C	-	new		intergenic variant
rs713738573	1	3427330	3427330	T	G	-	new		intergenic variant
rs81505732	2	7592175	7592175	T	C	ENSSSCG00000020831	SLC22A11	solute carrier family 22 member 11	intron variant
rs333210122	2	7592186	7592186	A	G	ENSSSCG00000020831	SLC22A11	solute carrier family 22 member 11	intron variant
rs322282589	2	7592192	7592192	T	C	ENSSSCG00000020831	SLC22A11	solute carrier family 22 member 11	intron variant
new	2	7592249	7592249	G	A	ENSSSCG00000020831	SLC22A11	solute carrier family 22 member 11	intron variant
rs787087930	2	7592295	7592295	C	G	ENSSSCG00000020831	SLC22A11	solute carrier family 22 member 11	intron variant
rs334404716	2	7598086	7598086	G	A	ENSSSCG00000020831	SLC22A11	solute carrier family 22 member 11	intron variant
rs338734805	2	7598137	7598137	G	T	ENSSSCG00000020831	SLC22A11	solute carrier family 22 member 11	intron variant
rs338734805	2	7598137	7598137	G	C	ENSSSCG00000020831	SLC22A11	solute carrier family 22 member 11	intron variant
new	2	7602943	7602943	C	G	ENSSSCG00000020831	SLC22A11	solute carrier family 22 member 11	missense variant
rs792193084	2	7602943	7602943	C	T	ENSSSCG00000020831	SLC22A11	solute carrier family 22 member 11	missense variant
new	2	7602945	7602945	G	A	ENSSSCG00000020831	SLC22A11	solute carrier family 22 member 11	frameshift variant
new	3	1594268	15942680	G	C	ENSSSCG00000007	new	tRNA-yW synthesizing	intron variant

		0				733		protein 1 homolog	
rs81309268	3	1595379 2	15953792	C	A	ENSSSCG00000007 733	new	tRNA-yW synthesizing protein 1 homolog	5 ' UTR variant
rs81309268	3	1595379 2	15953792	C	A	ENSSSCG00000007 733	new	tRNA-yW synthesizing protein 1 homolog	intron variant
new	3	1607855 7	16078557	G	A	ENSSSCG00000007 733	new	tRNA-yW synthesizing protein 1 homolog	synonymous variant
rs713010025	3	1637667 8	16376678	G	T	ENSSSCG00000007 735	RABGEF1	RAB guanine nucleotide exchange factor 1	upstream gene variant
rs337976532	3	1638242 8	16382428	A	G	ENSSSCG00000040 985	KCTD7	potassium channel tetramerization domain containing 7	downstream gene variant
rs320729536	3	1638261 9	16382619	T	C	ENSSSCG00000040 985	KCTD7	potassium channel tetramerization domain containing 7	3 ' UTR variant
rs320729536	3	1638261 9	16382619	T	C	ENSSSCG00000040 985	KCTD7	potassium channel tetramerization domain containing 7	downstream gene variant
rs323726488	3	1638324 6	16383246	C	T	ENSSSCG00000040 985	KCTD7	potassium channel tetramerization domain containing 7	3 ' UTR variant
rs323726488	3	1638324 6	16383246	C	T	ENSSSCG00000040 985	KCTD7	potassium channel tetramerization domain containing 7	downstream gene variant
rs333208968	3	1638325 0	16383250	G	A	ENSSSCG00000040 985	KCTD7	potassium channel tetramerization domain containing 7	3 ' UTR variant
rs333208968	3	1638325 0	16383250	G	A	ENSSSCG00000040 985	KCTD7	potassium channel tetramerization domain containing 7	downstream gene variant

rs333661817	3	16383284	16383284	C	T	ENSSSCG00000040985	KCTD7	potassium channel tetramerization domain containing 7	downstream gene variant
rs333661817	3	16383284	16383284	C	T	ENSSSCG00000040985	KCTD7	potassium channel tetramerization domain containing 7	3 ' UTR variant
rs326053487	3	16383454	16383454	C	G	ENSSSCG00000040985	KCTD7	potassium channel tetramerization domain containing 7	3 ' UTR variant
rs326053487	3	16383454	16383454	C	G	ENSSSCG00000040985	KCTD7	potassium channel tetramerization domain containing 7	downstream gene variant
rs336316568	3	16383763	16383763	A	G	ENSSSCG00000040985	KCTD7	potassium channel tetramerization domain containing 7	3 ' UTR variant
rs336316568	3	16383763	16383763	A	G	ENSSSCG00000040985	KCTD7	potassium channel tetramerization domain containing 7	downstream gene variant
rs329707669	3	16383919	16383919	G	A	ENSSSCG00000040985	KCTD7	potassium channel tetramerization domain containing 7	3 ' UTR variant
rs329707669	3	16383919	16383919	G	A	ENSSSCG00000040985	KCTD7	potassium channel tetramerization domain containing 7	downstream gene variant
rs713060696	3	16383936	16383936	G	A	ENSSSCG00000040985	KCTD7	potassium channel tetramerization domain containing 7	downstream gene variant
rs713060696	3	16383936	16383936	G	A	ENSSSCG00000040985	KCTD7	potassium channel tetramerization domain containing 7	3 ' UTR variant

rs701844432	3	16383956	16383956	A	G	ENSSSCG00000040985	KCTD7	potassium channel tetramerization domain containing 7	downstream gene variant
rs324583382	3	16384043	16384043	C	T	ENSSSCG00000040985	KCTD7	potassium channel tetramerization domain containing 7	3 ' UTR variant
rs324583382	3	16384043	16384043	C	T	ENSSSCG00000040985	KCTD7	potassium channel tetramerization domain containing 7	downstream gene variant
rs345204099	3	16384270	16384270	C	T	ENSSSCG00000040985	KCTD7	potassium channel tetramerization domain containing 7	3 ' UTR variant
rs345204099	3	16384270	16384270	C	T	ENSSSCG00000040985	KCTD7	potassium channel tetramerization domain containing 7	downstream gene variant
rs327324699	3	16384293	16384293	C	T	ENSSSCG00000040985	KCTD7	potassium channel tetramerization domain containing 7	3 ' UTR variant
rs327324699	3	16384293	16384293	C	T	ENSSSCG00000040985	KCTD7	potassium channel tetramerization domain containing 7	downstream gene variant
rs329831862	3	16384540	16384540	A	G	ENSSSCG00000040985	KCTD7	potassium channel tetramerization domain containing 7	downstream gene variant
rs329831862	3	16384540	16384540	A	G	ENSSSCG00000040985	KCTD7	potassium channel tetramerization domain containing 7	3 ' UTR variant
rs332144111	3	16384634	16384634	C	G	ENSSSCG00000040985	KCTD7	potassium channel tetramerization domain containing 7	3 ' UTR variant

rs332144111	3	16384634	16384634	C	G	ENSSSCG00000040985	KCTD7	potassium channel tetramerization domain containing 7	downstream gene variant
rs323662654	3	16384658	16384658	T	C	ENSSSCG00000040985	KCTD7	potassium channel tetramerization domain containing 7	3 ' UTR variant
rs323662654	3	16384658	16384658	T	C	ENSSSCG00000040985	KCTD7	potassium channel tetramerization domain containing 7	downstream gene variant
rs335859180	3	16385181	16385181	G	C	ENSSSCG00000040985	KCTD7	potassium channel tetramerization domain containing 7	downstream gene variant
rs335859180	3	16385181	16385181	G	C	ENSSSCG00000040985	KCTD7	potassium channel tetramerization domain containing 7	intron variant
rs335859180	3	16385181	16385181	G	C	ENSSSCG00000040985	KCTD7	potassium channel tetramerization domain containing 7	3 ' UTR variant
new	3	16385215	16385215	G	CC	ENSSSCG00000040985	KCTD7	potassium channel tetramerization domain containing 7	downstream gene variant
new	3	16385215	16385215	G	CC	ENSSSCG00000040985	KCTD7	potassium channel tetramerization domain containing 7	3 ' UTR variant
new	3	16385215	16385215	G	CC	ENSSSCG00000040985	KCTD7	potassium channel tetramerization domain containing 7	intron variant
rs332268785	3	16385727	16385727	G	A	ENSSSCG00000040985	KCTD7	potassium channel tetramerization domain containing 7	intron variant

rs332268785	3	16385727	16385727	G	A	ENSSSCG00000040985	KCTD7	potassium channel tetramerization domain containing 7	3 ' UTR variant
new	3	16844063	16844064	GC	-	ENSSSCG00000007745	SUMF2	sulfatase modifying factor 2	upstream gene variant
new	3	16844063	16844064	GC	-	ENSSSCG00000033141	new	new	downstream gene variant
new	3	16844063	16844064	GC	-	ENSSSCG00000020808	CCT6A	chaperonin containing TCP1 subunit 6A	3 ' UTR variant
rs330731365	3	16844265	16844265	G	A	ENSSSCG00000020808	CCT6A	chaperonin containing TCP1 subunit 6A	3 ' UTR variant
rs330731365	3	16844265	16844265	G	A	ENSSSCG00000033141	new	new	downstream gene variant
rs330731365	3	16844265	16844265	G	A	ENSSSCG00000007745	SUMF2	sulfatase modifying factor 2	upstream gene variant
rs343913735	3	16844271	16844271	A	G	ENSSSCG00000020808	CCT6A	chaperonin containing TCP1 subunit 6A	3 ' UTR variant
rs343913735	3	16844271	16844271	A	G	ENSSSCG00000033141	new	new	downstream gene variant
rs343913735	3	16844271	16844271	A	G	ENSSSCG00000007745	SUMF2	sulfatase modifying factor 2	upstream gene variant
rs701592732	3	16844932	16844932	T	C	ENSSSCG00000020808	CCT6A	chaperonin containing TCP1 subunit 6A	3 ' UTR variant
rs701592732	3	16844932	16844932	T	C	ENSSSCG00000020808	CCT6A	chaperonin containing TCP1 subunit 6A	intron variant
rs701592732	3	16844932	16844932	T	C	ENSSSCG00000007745	SUMF2	sulfatase modifying factor 2	upstream gene variant
rs701592732	3	16844932	16844932	T	C	ENSSSCG00000033141	new	new	downstream gene variant
rs691819169	3	16844934	16844934	A	G	ENSSSCG00000020808	CCT6A	chaperonin containing TCP1 subunit 6A	3 ' UTR variant

rs691819169	3	16844934	16844934	A	G	ENSSSCG00000033141	new	new	downstream gene variant
rs691819169	3	16844934	16844934	A	G	ENSSSCG00000007745	SUMF2	sulfatase modifying factor 2	upstream gene variant
rs691819169	3	16844934	16844934	A	G	ENSSSCG00000020808	CCT6A	chaperonin containing TCP1 subunit 6A	intron variant
rs706237879	3	16844935	16844935	A	G	ENSSSCG00000020808	CCT6A	chaperonin containing TCP1 subunit 6A	3 ' UTR variant
rs706237879	3	16844935	16844935	A	G	ENSSSCG00000033141	new	new	downstream gene variant
rs706237879	3	16844935	16844935	A	G	ENSSSCG00000020808	CCT6A	chaperonin containing TCP1 subunit 6A	intron variant
rs706237879	3	16844935	16844935	A	G	ENSSSCG00000007745	SUMF2	sulfatase modifying factor 2	upstream gene variant
new	3	16847861	16847861	T	C	ENSSSCG00000033141	new	new	non coding transcript exon variant
new	3	16847861	16847861	T	C	ENSSSCG00000020808	CCT6A	chaperonin containing TCP1 subunit 6A	intron variant
new	3	16847861	16847861	T	C	ENSSSCG00000007745	SUMF2	sulfatase modifying factor 2	upstream gene variant
rs1108762720	3	16851119	16851119	A	G	ENSSSCG00000020808	CCT6A	chaperonin containing TCP1 subunit 6A	synonymous variant
rs1108762720	3	16851119	16851119	A	G	ENSSSCG00000033141	new	new	upstream gene variant
rs1108762720	3	16851119	16851119	A	G	ENSSSCG00000018553	new	new	downstream gene variant
rs1113974639	3	16852667	16852667	C	T	ENSSSCG00000018553	new	new	downstream gene variant
rs1113974639	3	16852667	16852667	C	T	ENSSSCG00000020808	CCT6A	chaperonin containing TCP1 subunit 6A	intron variant

rs1113974639	3	16852667	16852667	C	T	ENSSSCG00000033141	new	new	upstream gene variant
rs336023254	3	16854410	16854410	G	A	ENSSSCG00000018553	new	new	downstream gene variant
rs336023254	3	16854410	16854410	G	A	ENSSSCG00000020808	CCT6A	chaperonin containing TCP1 subunit 6A	intron variant
rs341960076	3	16855237	16855237	A	C	ENSSSCG00000020808	CCT6A	chaperonin containing TCP1 subunit 6A	intron variant
rs341960076	3	16855237	16855237	A	C	ENSSSCG00000020808	CCT6A	chaperonin containing TCP1 subunit 6A	synonymous variant
rs341960076	3	16855237	16855237	A	C	ENSSSCG00000018553	new	new	upstream gene variant
rs339277541	3	16958038	16958038	T	C	ENSSSCG00000029029	ZNF713	zinc finger protein 713	downstream gene variant
rs331020706	3	16958150	16958150	G	A	ENSSSCG00000029029	ZNF713	zinc finger protein 713	downstream gene variant
rs340816565	3	16958194	16958194	T	C	ENSSSCG00000029029	ZNF713	zinc finger protein 713	downstream gene variant
rs327975767	3	16958979	16958979	A	G	ENSSSCG00000029029	ZNF713	zinc finger protein 713	3 ' UTR variant
rs325589354	3	16959068	16959068	C	T	ENSSSCG00000029029	ZNF713	zinc finger protein 713	3 ' UTR variant
rs337698799	3	16959239	16959239	A	G	ENSSSCG00000029029	ZNF713	zinc finger protein 713	3 ' UTR variant
rs318863693	3	16959265	16959265	T	C	ENSSSCG00000029029	ZNF713	zinc finger protein 713	3 ' UTR variant
rs329457108	3	16959434	16959434	G	A	ENSSSCG00000029029	ZNF713	zinc finger protein 713	3 ' UTR variant
rs334498862	3	16959634	16959634	T	C	ENSSSCG00000029029	ZNF713	zinc finger protein 713	3 ' UTR variant
new	3	16959719	16959719	CT	-	ENSSSCG00000029	ZNF713	zinc finger protein 713	3 ' UTR variant

		8				029			
rs693699557	3	16959721	16959721	G	C	ENSSSCG00000029029	ZNF713	zinc finger protein 713	3 ' UTR variant
rs324279159	3	16960031	16960031	A	C	ENSSSCG00000029029	ZNF713	zinc finger protein 713	3 ' UTR variant
rs788766021	3	16960220	16960224	AAACT	-	ENSSSCG00000029029	ZNF713	zinc finger protein 713	3 ' UTR variant
rs336887529	3	16960465	16960465	T	G	ENSSSCG00000029029	ZNF713	zinc finger protein 713	3 ' UTR variant
rs343256392	3	16960493	16960493	A	G	ENSSSCG00000029029	ZNF713	zinc finger protein 713	3 ' UTR variant
rs325887031	3	16960740	16960740	G	A	ENSSSCG00000029029	ZNF713	zinc finger protein 713	3 ' UTR variant
rs337918521	3	16960969	16960969	C	A	ENSSSCG00000029029	ZNF713	zinc finger protein 713	missense variant
rs320802395	3	16961030	16961030	G	A	ENSSSCG00000029029	ZNF713	zinc finger protein 713	synonymous variant
rs330288467	3	16961219	16961219	T	C	ENSSSCG00000029029	ZNF713	zinc finger protein 713	synonymous variant
rs323115420	3	16964045	16964045	T	A	ENSSSCG00000029029	ZNF713	zinc finger protein 713	missense variant
rs335343250	3	16964916	16964916	G	C	ENSSSCG00000029029	ZNF713	zinc finger protein 713	intron variant
rs329911516	3	16970816	16970816	A	G	ENSSSCG00000029029	ZNF713	zinc finger protein 713	intron variant
rs342141649	3	16970819	16970819	C	T	ENSSSCG00000029029	ZNF713	zinc finger protein 713	intron variant
rs342012840	3	16971089	16971089	T	C	ENSSSCG00000029029	ZNF713	zinc finger protein 713	5 ' UTR variant
rs324205762	3	16971143	16971143	T	C	ENSSSCG00000029029	ZNF713	zinc finger protein 713	5 ' UTR variant

rs333780109	3	16971156	16971156	T	C	ENSSSCG00000029029	ZNF713	zinc finger protein 713	5 ' UTR variant
rs343615406	3	16971169	16971169	T	C	ENSSSCG00000029029	ZNF713	zinc finger protein 713	5 ' UTR variant
rs328383315	3	16971295	16971295	T	A	ENSSSCG00000029029	ZNF713	zinc finger protein 713	intron variant
new	3	17001571	17001575	ATTTG	-	ENSSSCG00000007751	SEPTIN14	septin 14	intron variant
rs342352307	3	17001888	17001888	T	C	ENSSSCG00000007751	SEPTIN14	septin 14	intron variant
new	3	17026001	17026001	T	C	ENSSSCG00000007751	SEPTIN14	septin 14	downstream gene variant
new	3	17033960	17033960	T	C	ENSSSCG00000036685	LOC106509673	olfactory receptor 7A17-like	missense variant
new	3	17033997	17033997	G	A	ENSSSCG00000036685	LOC106509673	olfactory receptor 7A17-like	synonymous variant
new	3	17034051	17034051	A	T	ENSSSCG00000036685	LOC106509673	olfactory receptor 7A17-like	synonymous variant
new	3	17034234	17034234	A	G	ENSSSCG00000036685	LOC106509673	olfactory receptor 7A17-like	synonymous variant
new	3	17034293	17034293	T	C	ENSSSCG00000036685	LOC106509673	olfactory receptor 7A17-like	missense variant
new	3	17034405	17034405	C	T	ENSSSCG00000036685	LOC106509673	olfactory receptor 7A17-like	synonymous variant
rs702384560	3	17034426	17034426	G	A	ENSSSCG00000036685	LOC106509673	olfactory receptor 7A17-like	synonymous variant
rs709083818	3	17034534	17034534	A	G	ENSSSCG00000036685	LOC106509673	olfactory receptor 7A17-like	synonymous variant
rs324169719	3	17034589	17034589	C	T	ENSSSCG00000036685	LOC106509673	olfactory receptor 7A17-like	5 ' UTR variant
rs330223831	3	1703465	17034658	C	T	ENSSSCG00000036	LOC106509	olfactory receptor	upstream gene

		8				685	673	7A17-like	variant
new	3	17243996	17243996	A	C	ENSSSCG00000007754	ITGAM	integrin subunit alpha M	intron variant
rs318763073	3	17243999	17243999	C	G	ENSSSCG00000007754	ITGAM	integrin subunit alpha M	intron variant
rs344858642	3	17244108	17244108	C	T	ENSSSCG00000007754	ITGAM	integrin subunit alpha M	synonymous variant
rs326942919	3	17246305	17246305	G	A	ENSSSCG00000007754	ITGAM	integrin subunit alpha M	synonymous variant
rs323402259	3	17246405	17246405	C	T	ENSSSCG00000007754	ITGAM	integrin subunit alpha M	intron variant
rs327289001	3	17254444	17254444	C	T	ENSSSCG00000007754	ITGAM	integrin subunit alpha M	missense variant
rs327947675	3	17399455	17399455	A	G	ENSSSCG00000007765	PRSS53	serine protease 53	downstream gene variant
rs327947675	3	17399455	17399455	A	G	ENSSSCG00000035364	VKORC1	vitamin K epoxide reductase complex subunit 1	downstream gene variant
rs327947675	3	17399455	17399455	A	G	ENSSSCG00000026817	ZNF646	zinc finger protein 646	synonymous variant
rs327947675	3	17399455	17399455	A	G	ENSSSCG00000007763	BCKDK	branched chain keto acid dehydrogenase kinase	5 ' UTR variant
rs337670844	3	17399477	17399477	C	T	ENSSSCG00000007763	BCKDK	branched chain keto acid dehydrogenase kinase	5 ' UTR variant
rs337670844	3	17399477	17399477	C	T	ENSSSCG00000007765	PRSS53	serine protease 53	downstream gene variant
rs337670844	3	17399477	17399477	C	T	ENSSSCG00000035364	VKORC1	vitamin K epoxide reductase complex subunit 1	downstream gene variant

rs337670844	3	17399477	17399477	C	T	ENSSSCG00000026817	ZNF646	zinc finger protein 646	missense variant
rs328098289	3	17430228	17430228	T	C	ENSSSCG00000049720	new	new	upstream gene variant
rs328098289	3	17430228	17430228	T	C	ENSSSCG00000007768	STX4	syntaxin 4	synonymous variant
rs341530377	3	17457840	17457840	G	A	ENSSSCG00000021238	STX1B	syntaxin 1B	3 ' UTR variant
rs341530377	3	17457840	17457840	G	A	ENSSSCG00000032369	HSD3B7	hydroxy-delta-5-steroid dehydrogenase, 3 beta- and steroid delta-isomerase 7	downstream gene variant
rs341530377	3	17457840	17457840	G	A	ENSSSCG00000021238	STX1B	syntaxin 1B	downstream gene variant
rs344714132	3	17458203	17458203	G	C	ENSSSCG00000032369	HSD3B7	hydroxy-delta-5-steroid dehydrogenase, 3 beta- and steroid delta-isomerase 7	downstream gene variant
rs344714132	3	17458203	17458203	G	C	ENSSSCG00000021238	STX1B	syntaxin 1B	downstream gene variant
rs344714132	3	17458203	17458203	G	C	ENSSSCG00000021238	STX1B	syntaxin 1B	3 ' UTR variant
rs344526901	3	17458409	17458409	C	T	ENSSSCG00000021238	STX1B	syntaxin 1B	3 ' UTR variant
rs344526901	3	17458409	17458409	C	T	ENSSSCG00000032369	HSD3B7	hydroxy-delta-5-steroid dehydrogenase, 3 beta- and steroid delta-isomerase 7	downstream gene variant
rs344526901	3	17458409	17458409	C	T	ENSSSCG00000021238	STX1B		downstream gene variant

rs334776843	3	17458467	17458467	C	T	ENSSSCG00000032369	HSD3B7	hydroxy-delta-5-steroid dehydrogenase, 3 beta- and steroid delta-isomerase 7	downstream gene variant
rs334776843	3	17458467	17458467	C	T	ENSSSCG00000021238	STX1B	syntaxin 1B	downstream gene variant
rs334776843	3	17458467	17458467	C	T	ENSSSCG00000021238	STX1B	syntaxin 1B	3 ' UTR variant
rs322669402	3	17459844	17459844	A	G	ENSSSCG00000032369	HSD3B7	hydroxy-delta-5-steroid dehydrogenase, 3 beta- and steroid delta-isomerase 7	3 ' UTR variant
rs322669402	3	17459844	17459844	A	G	ENSSSCG00000021238	STX1B	syntaxin 1B	downstream gene variant
rs322669402	3	17459844	17459844	A	G	ENSSSCG00000007782	SETD1A	SET domain containing 1A, histone lysine methyltransferase	downstream gene variant
rs793318116	3	17460656	17460656	C	T	ENSSSCG00000007782	SETD1A	SET domain containing 1A, histone lysine methyltransferase	downstream gene variant
rs793318116	3	17460656	17460656	C	T	ENSSSCG00000021238	STX1B	syntaxin 1B	downstream gene variant
rs793318116	3	17460656	17460656	C	T	ENSSSCG00000032369	HSD3B7	hydroxy-delta-5-steroid dehydrogenase, 3 beta- and steroid delta-isomerase 7	synonymous variant
rs336707684	3	17466842	17466842	A	G	ENSSSCG00000007782	SETD1A	SET domain containing 1A, histone lysine	intron variant

								methyltransferase	
rs336707684	3	17466842	17466842	A	G	ENSSSCG00000032369	HSD3B7	hydroxy-delta-5-steroid dehydrogenase, 3 beta- and steroid delta-isomerase 7	upstream gene variant
rs327942559	3	17466860	17466860	T	C	ENSSSCG00000007782	SETD1A	SET domain containing 1A, histone lysine methyltransferase	intron variant
rs327942559	3	17466860	17466860	T	C	ENSSSCG00000032369	HSD3B7	hydroxy-delta-5-steroid dehydrogenase, 3 beta- and steroid delta-isomerase 7	upstream gene variant
rs339276563	3	17468216	17468216	T	C	ENSSSCG00000007782	SETD1A	SET domain containing 1A, histone lysine methyltransferase	synonymous variant
rs339276563	3	17468216	17468216	T	C	ENSSSCG00000032369	HSD3B7	hydroxy-delta-5-steroid dehydrogenase, 3 beta- and steroid delta-isomerase 7	upstream gene variant
rs330957838	3	17468302	17468302	T	C	ENSSSCG00000032369	HSD3B7	hydroxy-delta-5-steroid dehydrogenase, 3 beta- and steroid delta-isomerase 7	upstream gene variant
rs330957838	3	17468302	17468302	T	C	ENSSSCG00000007782	SETD1A	SET domain containing 1A, histone lysine methyltransferase	missense variant

rs340370195	3	17479170	17479170	A	G	ENSSSCG00000007782	SETD1A	SET domain containing 1A, histone lysine methyltransferase	intron variant
rs345676220	3	17491423	17491423	G	A	ENSSSCG00000007782	SETD1A	SET domain containing 1A, histone lysine methyltransferase	upstream gene variant
rs345676220	3	17491423	17491423	G	A	ENSSSCG00000007770	ORAI3	ORAI calcium release-activated calcium modulator 3	3 ' UTR variant
rs332536957	3	17497571	17497571	T	G	ENSSSCG00000028186	FBXL19	F-box and leucine rich repeat protein 19	3 ' UTR variant
rs332536957	3	17497571	17497571	T	G	ENSSSCG00000007770	ORAI3	ORAI calcium release-activated calcium modulator 3	upstream gene variant
rs344892486	3	17610468	17610468	T	C	ENSSSCG00000007780	ZNF629	zinc finger protein 629	5 ' UTR variant
rs345908614	3	17613140	17613140	A	G	ENSSSCG00000007780	ZNF629	zinc finger protein 629	intron variant
rs335540465	3	17613531	17613531	T	C	ENSSSCG00000007780	ZNF629	zinc finger protein 629	synonymous variant
new	3	17617323	17617323	CAGAA	-	ENSSSCG00000007786	RNF40	ring finger protein 40	downstream gene variant
new	3	17617323	17617323	CAGAA	-	ENSSSCG00000007780	ZNF629	zinc finger protein 629	3 ' UTR variant
rs1107804156	3	17618533	17618533	G	T	ENSSSCG00000007780	ZNF629	zinc finger protein 629	intron variant
rs1107804156	3	17618533	17618533	G	T	ENSSSCG00000007786	RNF40	ring finger protein 40	downstream gene variant
rs789266896	3	1762868	17628688	T	G	ENSSSCG00000007	ZNF629	zinc finger protein 629	3 ' UTR variant

		8				780			
rs789266896	3	17628688	17628688	T	G	ENSSSCG00000007786	RNF40	ring finger protein 40	missense variant
rs319144424	3	17631027	17631027	C	A	ENSSSCG00000007780	ZNF629	zinc finger protein 629	downstream gene variant
rs319144424	3	17631027	17631027	C	A	ENSSSCG00000007786	RNF40	ring finger protein 40	intron variant
rs696120779	3	17636757	17636757	A	C	ENSSSCG00000007786	RNF40	ring finger protein 40	5 ' UTR variant
rs696120779	3	17636757	17636757	A	C	ENSSSCG00000007788	PHKG2	phosphorylase kinase catalytic subunit gamma 2	downstream gene variant
rs696120779	3	17636757	17636757	A	C	ENSSSCG00000007778	CCDC189	coiled-coil domain containing 189	intron variant
rs792624385	3	17639370	17639370	G	A	ENSSSCG00000007788	PHKG2	phosphorylase kinase catalytic subunit gamma 2	downstream gene variant
rs792624385	3	17639370	17639370	G	A	ENSSSCG00000007778	CCDC189	coiled-coil domain containing 189	intron variant
rs792624385	3	17639370	17639370	G	A	ENSSSCG00000007786	RNF40	ring finger protein 40	intron variant
rs324849698	3	17639613	17639613	T	C	ENSSSCG00000007788	PHKG2	phosphorylase kinase catalytic subunit gamma 2	downstream gene variant
rs324849698	3	17639613	17639613	T	C	ENSSSCG00000007786	RNF40	ring finger protein 40	intron variant
rs324849698	3	17639613	17639613	T	C	ENSSSCG00000007778	CCDC189	coiled-coil domain containing 189	intron variant
rs343172093	3	17639636	17639636	A	G	ENSSSCG00000007788	PHKG2	phosphorylase kinase catalytic subunit gamma 2	downstream gene variant

rs343172093	3	17639636	17639636	A	G	ENSSSCG00000007778	CCDC189	coiled-coil domain containing 189	intron variant
rs343172093	3	17639636	17639636	A	G	ENSSSCG00000007786	RNF40	ring finger protein 40	intron variant
rs706476605	3	17639641	17639641	A	C	ENSSSCG00000007788	PHKG2	phosphorylase kinase catalytic subunit gamma 2	downstream gene variant
rs706476605	3	17639641	17639641	A	C	ENSSSCG00000007778	CCDC189	coiled-coil domain containing 189	intron variant
rs706476605	3	17639641	17639641	A	C	ENSSSCG00000007786	RNF40	ring finger protein 40	intron variant
rs706021010	3	17639643	17639643	C	T	ENSSSCG00000007778	CCDC189	coiled-coil domain containing 189	intron variant
rs706021010	3	17639643	17639643	C	T	ENSSSCG00000007786	RNF40	ring finger protein 40	intron variant
rs706021010	3	17639643	17639643	C	T	ENSSSCG00000007788	PHKG2	phosphorylase kinase catalytic subunit gamma 2	downstream gene variant
rs690340579	3	17639901	17639901	A	G	ENSSSCG00000007788	PHKG2	phosphorylase kinase catalytic subunit gamma 2	downstream gene variant
rs690340579	3	17639901	17639901	A	G	ENSSSCG00000007778	CCDC189	coiled-coil domain containing 189	synonymous variant
rs690340579	3	17639901	17639901	A	G	ENSSSCG00000007786	RNF40	ring finger protein 40	upstream gene variant
rs335008558	3	17640590	17640590	T	C	ENSSSCG00000007788	PHKG2	phosphorylase kinase catalytic subunit gamma 2	3 ' UTR variant
rs335008558	3	17640590	17640590	T	C	ENSSSCG00000007786	RNF40	ring finger protein 40	upstream gene variant
rs335008558	3	1764059	17640590	T	C	ENSSSCG00000007	CCDC189	coiled-coil domain	synonymous variant

		0				778		containing 189	
rs340945745	3	17767030	17767030	T	C	ENSSSCG00000047341	new	new	downstream gene variant
rs340945745	3	17767030	17767030	T	C	ENSSSCG00000033496	ZNF689	zinc finger protein 689	5 ' UTR variant
new	3	17979825	17979825	A	G	ENSSSCG00000035256	SPN	sialophorin	missense variant
new	3	17979825	17979825	A	G	ENSSSCG00000007799	MYLPF	myosin light chain, phosphorylatable, fast skeletal muscle	intron variant
new	3	17979826	17979826	A	T	ENSSSCG00000007799	MYLPF	myosin light chain, phosphorylatable, fast skeletal muscle	intron variant
new	3	17979826	17979826	A	T	ENSSSCG00000035256	SPN	sialophorin	missense variant
new	3	17979827	17979827	T	C	ENSSSCG00000007799	MYLPF	myosin light chain, phosphorylatable, fast skeletal muscle	intron variant
new	3	17979827	17979827	T	C	ENSSSCG00000035256	SPN	sialophorin	synonymous variant
rs328258515	3	101613325	101613325	T	C	ENSSSCG00000008484	SRSF7	serine and arginine rich splicing factor 7	downstream gene variant
rs328258515	3	101613325	101613325	T	C	ENSSSCG00000008486	GALM	galactose mutarotase	downstream gene variant
new	3	125492648	125492649	TA	-	ENSSSCG00000008629	ROCK2	Rho associated coiled-coil containing protein kinase 2	downstream gene variant
new	3	125492648	125492649	TA	-	ENSSSCG00000008631	SLC66A3	solute carrier family 66 member 3	downstream gene variant
new	4	89321959	89321959	T	A	ENSSSCG00000006365	DEDD	death effector domain containing	intron variant

new	4	89321959	89321959	T	A	ENSSSCG00000006366	PFDN2	prefoldin subunit 2	upstream gene variant
new	4	89321959	89321959	T	A	ENSSSCG00000037793	NIT1	nitrilase 1	intron variant
new	4	89321959	89321959	T	A	ENSSSCG00000037793	NIT1	nitrilase 1	downstream gene variant
rs319871479	4	128985819	128985819	T	C	-	new		intergenic variant
rs329715150	5	17364441	17364441	G	A	ENSSSCG00000000234	GRASP	general receptor for phosphoinositides 1 associated scaffold protein	intron variant
rs345481021	6	80824380	80824380	T	G	ENSSSCG00000003527	EPHB2	EPH receptor B2	synonymous variant
rs326115442	6	80824416	80824416	T	C	ENSSSCG00000003527	EPHB2	EPH receptor B2	synonymous variant
rs337528160	6	80824455	80824455	A	G	ENSSSCG00000003527	EPHB2	EPH receptor B2	intron variant
rs81389080	6	80826802	80826802	C	T	ENSSSCG00000003527	EPHB2	EPH receptor B2	intron variant
rs343091930	6	80836094	80836094	T	C	ENSSSCG00000003527	EPHB2	EPH receptor B2	intron variant
rs337803448	6	80836565	80836565	C	T	ENSSSCG00000003527	EPHB2	EPH receptor B2	intron variant
rs327572607	6	80837841	80837841	T	C	ENSSSCG00000003527	EPHB2	EPH receptor B2	synonymous variant
rs323316435	6	80840438	80840438	G	A	ENSSSCG00000003527	EPHB2	EPH receptor B2	intron variant
rs81389091	6	80842615	80842615	T	C	ENSSSCG00000003527	EPHB2	EPH receptor B2	3 ' UTR variant
rs710496863	6	8084286	80842862	G	A	ENSSSCG00000003	EPHB2	EPH receptor B2	3 ' UTR variant

		2				527			
new	6	80843072	80843073	CA	-	ENSSSCG00000003527	EPHB2	EPH receptor B2	3 ' UTR variant
new	6	80843074	80843081	GGCGGGCA	-	ENSSSCG00000003527	EPHB2	EPH receptor B2	3 ' UTR variant
rs342283188	6	80843098	80843098	T	G	ENSSSCG00000003527	EPHB2	EPH receptor B2	3 ' UTR variant
rs81389092	6	80843265	80843265	A	G	ENSSSCG00000003527	EPHB2	EPH receptor B2	3 ' UTR variant
rs346223430	6	80843336	80843336	C	T	ENSSSCG00000003527	EPHB2	EPH receptor B2	3 ' UTR variant
rs325312968	6	80843363	80843363	A	G	ENSSSCG00000003527	EPHB2	EPH receptor B2	3 ' UTR variant
rs328228830	6	80843833	80843833	T	C	ENSSSCG00000003527	EPHB2	EPH receptor B2	downstream gene variant
rs331678844	6	80843991	80843991	G	T	ENSSSCG00000003527	EPHB2	EPH receptor B2	downstream gene variant
rs81389093	6	80844018	80844018	A	G	ENSSSCG00000003527	EPHB2	EPH receptor B2	downstream gene variant
rs81233096	6	80897035	80897035	A	C	ENSSSCG00000003530	TEX46	testis expressed 46	intron variant
rs328260148	6	80977371	80977371	T	C	ENSSSCG00000003532	LUZP1	leucine zipper protein 1	intron variant
rs344493496	6	80984204	80984204	A	G	ENSSSCG00000003532	LUZP1	leucine zipper protein 1	intron variant
rs329148574	6	81000551	81000551	T	C	ENSSSCG00000003532	LUZP1	leucine zipper protein 1	intron variant
rs325089032	6	81571496	81571496	G	A	ENSSSCG00000025440	ELOA	elongin A	missense variant
rs345825612	6	8168104	81681049	C	T	ENSSSCG00000027	CNR2	Interleukin 18 receptor	5 ' UTR variant

		9				849		1	
rs345825612	6	81681049	81681049	C	T	ENSSSCG00000027849	CNR2	cannabinoid receptor 2	intron variant
rs345825612	6	81681049	81681049	C	T	ENSSSCG00000027659	FUCA1	alpha-L-fucosidase	intron variant
rs345825612	6	81681049	81681049	C	T	ENSSSCG00000023191	LYPLA2	lysophospholipase 2	intron variant
rs332149968	6	81681123	81681123	C	T	ENSSSCG00000027849	CNR2	cannabinoid receptor 2	intron variant
rs332149968	6	81681123	81681123	C	T	ENSSSCG00000027659	FUCA1	alpha-L-fucosidase	intron variant
rs332149968	6	81681123	81681123	C	T	ENSSSCG00000027849	CNR2	cannabinoid receptor 2	upstream gene variant
rs332149968	6	81681123	81681123	C	T	ENSSSCG00000023191	LYPLA2	lysophospholipase 2	intron variant
rs321503722	6	82401042	82401042	A	C	ENSSSCG00000038994	CLIC4	chloride intracellular channel 4	upstream gene variant
rs322593756	7	96619078	96619078	C	T	ENSSSCG00000002341	PAPLN	papilin, proteoglycan like sulfated glycoprotein	intron variant
new	7	113442605	113442605	G	CA	ENSSSCG00000002444	FBLN5	fibulin 5	intron variant
rs696812713	8	6157581	6157581	T	C	ENSSSCG00000008731	OTOP1	otopetrin 1	missense variant
rs712855168	8	6157582	6157582	A	C	ENSSSCG00000008731	OTOP1	otopetrin 1	missense variant
rs328470251	9	1324980	1324980	C	T	ENSSSCG00000042620	LOC100736607	olfactory receptor 10A3	synonymous variant
rs345798145	9	49630761	49630761	C	T	ENSSSCG00000015136	UBASH3B	ubiquitin associated and SH3 domain containing B	upstream gene variant

rs332848504	10	8836715	8836715	G	A	ENSSSCG00000038539	new	new	non coding transcript exon variant
rs706216562	10	15297251	15297251	A	T	ENSSSCG00000031683	LOC102166270	new	non coding transcript exon variant
rs706216562	10	15297251	15297251	A	T	ENSSSCG00000031683	LOC102166270	new	splice region variant
rs690508595	10	15297252	15297252	C	T	ENSSSCG00000031683	LOC102166270	new	splice region variant
rs690508595	10	15297252	15297252	C	T	ENSSSCG00000031683	LOC102166270	new	non coding transcript exon variant
rs698813384	10	15297253	15297253	A	G	ENSSSCG00000031683	LOC102166270	new	non coding transcript exon variant
rs698813384	10	15297253	15297253	A	G	ENSSSCG00000031683	LOC102166270	new	splice region variant
new	10	43353476	43353476	C	T	ENSSSCG00000011030	CUBN	cubilin	intron variant
new	12	19348500	19348500	T	G	ENSSSCG00000020744	DUSP3	dual specificity phosphatase 3	3 ' UTR variant
new	12	19348500	19348500	T	G	ENSSSCG00000020744	DUSP3	dual specificity phosphatase 3	intron variant
rs335136145	12	38026003	38026003	G	A	ENSSSCG00000023373	ZNHIT3	zinc finger HIT-type containing 3	3 ' UTR variant
rs335136145	12	38026003	38026003	G	A	ENSSSCG00000017682	MYO19	myosin XIX	3 ' UTR variant
rs335136145	12	38026003	38026003	G	A	ENSSSCG00000017682	MYO19	myosin XIX	missense variant
rs335136145	12	3802600	38026003	G	A	ENSSSCG00000027	new	new	downstream gene

		3				233			variant
rs331463738	12	38129509	38129509	T	C	ENSSSCG00000017690	DHRS11	dehydrogenase/reductase 11	3 ' UTR variant
rs331463738	12	38129509	38129509	T	C	ENSSSCG00000017690	DHRS11	dehydrogenase/reductase 11	downstream gene variant
rs331463738	12	38129509	38129509	T	C	ENSSSCG00000017690	DHRS11	dehydrogenase/reductase 11	intron variant
rs340781986	12	38624687	38624687	G	A	ENSSSCG00000017694	ACACA	acetyl-CoA carboxylase alpha	synonymous variant
rs324236192	12	38624714	38624714	G	A	ENSSSCG00000017694	ACACA	acetyl-CoA carboxylase alpha	synonymous variant
rs327539493	12	55096339	55096339	C	G	ENSSSCG00000024416	MYH13	myosin-13	intron variant
rs324996684	12	55097665	55097665	A	C	ENSSSCG00000024416	MYH13	myosin-13	synonymous variant
rs341831793	12	55099832	55099832	T	A	ENSSSCG00000024416	MYH13	myosin-13	intron variant
rs346367824	12	55103125	55103125	A	C	ENSSSCG00000024416	MYH13	myosin-13	intron variant
new	12	55146905	55146905	T	G	ENSSSCG00000018005	MYH8	myosin-8	missense variant
new	12	55152605	55152605	A	C	ENSSSCG00000018005	MYH8	myosin-8	intron variant
new	12	55164179	55164179	G	A	ENSSSCG00000018005	MYH8	myosin-8	synonymous variant
new	12	55164179	55164179	G	A	ENSSSCG00000018005	MYH8	myosin-8	intron variant
new	12	55164180	55164180	C	G	ENSSSCG00000018005	MYH8	myosin-8	intron variant
new	12	55164180	55164180	C	G	ENSSSCG00000018005	MYH8	myosin-8	missense variant

new	12	55164182	55164182	C	G	ENSSSCG00000018005	MYH8	myosin-8	intron variant
new	12	55164182	55164182	C	G	ENSSSCG00000018005	MYH8	myosin-8	synonymous variant
new	12	55197434	55197434	T	G	ENSSSCG00000029441	MYH2	myosin, heavy chain 2, skeletal muscle, adult	intron variant
new	12	55197434	55197434	T	G	ENSSSCG00000029441	MYH2	myosin, heavy chain 2, skeletal muscle, adult	missense variant
new	12	55197435	55197435	T	C	ENSSSCG00000029441	MYH2	myosin, heavy chain 2, skeletal muscle, adult	synonymous variant
new	12	55197435	55197435	T	C	ENSSSCG00000029441	MYH2	myosin, heavy chain 2, skeletal muscle, adult	intron variant
new	12	55205557	55205557	T	C	ENSSSCG00000029441	MYH2	myosin, heavy chain 2, skeletal muscle, adult	synonymous variant
new	12	55205557	55205557	T	C	ENSSSCG00000029441	MYH2	myosin, heavy chain 2, skeletal muscle, adult	intron variant
new	12	55263362	55263362	T	C	ENSSSCG00000029441	MYH2	myosin, heavy chain 2, skeletal muscle, adult	intron variant
new	12	55267170	55267170	G	C	ENSSSCG00000029441	MYH2	myosin, heavy chain 2, skeletal muscle, adult	intron variant
new	12	55267170	55267170	G	C	ENSSSCG00000029441	MYH2	myosin, heavy chain 2, skeletal muscle, adult	missense variant
rs319222050	12	55356662	55356662	G	T	ENSSSCG00000018007	MYH3	myosin-3	missense variant
rs337961174	12	55356663	55356663	C	T	ENSSSCG00000018007	MYH3	myosin-3	synonymous variant
new	12	55372418	55372420	CGG	-	ENSSSCG00000018007	MYH3	myosin-3	upstream gene variant
new	12	5537241	55372420	CGG	-	ENSSSCG00000018	MYH3	myosin-3	intron variant

		8				007			
new	12	55438152	55438155	ATCG	-	ENSSSCG00000018009	ADPRM	ADP-ribose/CDP-alcohol diphosphatase, manganese dependent	downstream gene variant
new	12	55438152	55438155	ATCG	-	ENSSSCG00000036454	TMEM220	transmembrane protein 220	downstream gene variant
new	12	55438152	55438155	ATCG	-	ENSSSCG00000018009	ADPRM	ADP-ribose/CDP-alcohol diphosphatase, manganese dependent	3 ' UTR variant
rs342689686	13	12410437	12410437	T	C	ENSSSCG00000011212	RARB	retinoic acid receptor beta	intron variant
rs330943524	14	108858300	108858300	A	C	ENSSSCG00000010520	MMS19	MMS19 homolog, cytosolic iron-sulfur assembly component	intron variant
rs324529516	15	50299014	50299014	C	T	ENSSSCG00000015830	UNC5D	unc-5 netrin receptor D	intron variant
rs790346070	X	39631586	39631586	A	G	-	new		intergenic variant
new	X	46911316	46911316	C	T	ENSSSCG00000045777	new	new	downstream gene variant

Additional file 2: Total number of variants (SNPs and InDels) identified in the transcriptome analysis that differed between normal and umbilical hernia-affected pigs its annotation and consequence predicted with the VEP tool.

SNP name	SSC	start	end	Allele (ref)	Allele (alt)	Gene	SYMBOL	gene name	Consequence (transcript)
new	1	21441491	21441491	C	T	ENSSSCG00000004132	PHACTR2	phosphatase and actin regulator 2	downstream gene variant
new	1	21441491	21441491	C	T	ENSSSCG00000004132	PHACTR2	phosphatase and actin regulator 2	downstream gene variant
new	1	21441491	21441491	C	T	ENSSSCG00000004132	PHACTR2	phosphatase and actin regulator 2	downstream gene variant
new	1	21441491	21441491	C	T	ENSSSCG00000004132	PHACTR2	phosphatase and actin regulator 2	downstream gene variant
new	1	21441491	21441491	C	T	ENSSSCG00000004132	PHACTR2	phosphatase and actin regulator 2	downstream gene variant
new	1	21441491	21441491	C	T	ENSSSCG00000004132	PHACTR2	phosphatase and actin regulator 2	downstream gene variant
new	1	37122230	37122230	C	T	ENSSSCG00000004222	NCOA7	Nuclear receptor coactivator 7	frameshift variant
new	1	37122230	37122230	C	T	ENSSSCG00000004222	NCOA7	Nuclear receptor coactivator 7	frameshift variant
new	1	37122230	37122230	C	T	ENSSSCG00000004222	NCOA7	Nuclear receptor coactivator 7	frameshift variant
new	1	77585583	77585583	T	A	ENSSSCG00000004421	FYN	FYN proto-oncogene, Src family tyrosine kinase	3 ' UTR variant
new	1	77585583	77585583	T	A	ENSSSCG00000004421	FYN	FYN proto-oncogene, Src family tyrosine kinase	3 ' UTR variant
new	1	77585583	77585583	T	A	ENSSSCG00000004421	FYN	FYN proto-oncogene, Src family tyrosine	3 ' UTR variant

								kinase	
new	3	16078557	16078557	G	A	ENSSSCG00000007733	TYW1	TRNA-yW synthesizing protein 1 homolog	synonymous variant
new	3	16078557	16078557	G	A	ENSSSCG00000007733	TYW1	TRNA-yW synthesizing protein 1 homolog	synonymous variant
new	3	16078557	16078557	G	A	ENSSSCG00000007733	TYW1	TRNA-yW synthesizing protein 1 homolog	synonymous variant
new	3	16078557	16078557	G	A	ENSSSCG00000007733	TYW1	TRNA-yW synthesizing protein 1 homolog	synonymous variant
new	3	16078557	16078557	G	A	ENSSSCG00000007733	TYW1	TRNA-yW synthesizing protein 1 homolog	synonymous variant
rs320729536	3	16382619	16382619	T	C	ENSSSCG00000040985	KCTD7	potassium channel tetramerization domain containing 7	downstream gene variant
rs320729536	3	16382619	16382619	T	C	ENSSSCG00000040985	KCTD7	potassium channel tetramerization domain containing 7	3 ' UTR variant
rs320729536	3	16382619	16382619	T	C	ENSSSCG00000040985	KCTD7	potassium channel tetramerization domain containing 7	downstream gene variant
rs320729536	3	16382619	16382619	T	C	ENSSSCG00000040985	KCTD7	potassium channel tetramerization domain containing 7	downstream gene variant
rs323726488	3	16383246	16383246	C	T	ENSSSCG00000040985	KCTD7	potassium channel tetramerization domain containing 7	downstream gene variant

rs323726488	3	16383246	16383246	C	T	ENSSSCG00000040985	KCTD7	potassium channel tetramerization domain containing 7	3 ' UTR variant
rs323726488	3	16383246	16383246	C	T	ENSSSCG00000040985	KCTD7	potassium channel tetramerization domain containing 7	downstream gene variant
rs323726488	3	16383246	16383246	C	T	ENSSSCG00000040985	KCTD7	potassium channel tetramerization domain containing 7	downstream gene variant
rs333208968	3	16383250	16383250	G	A	ENSSSCG00000040985	KCTD7	potassium channel tetramerization domain containing 7	downstream gene variant
rs333208968	3	16383250	16383250	G	A	ENSSSCG00000040985	KCTD7	potassium channel tetramerization domain containing 7	3 ' UTR variant
rs333208968	3	16383250	16383250	G	A	ENSSSCG00000040985	KCTD7	potassium channel tetramerization domain containing 7	downstream gene variant
rs333208968	3	16383250	16383250	G	A	ENSSSCG00000040985	KCTD7	potassium channel tetramerization domain containing 7	downstream gene variant
rs333661817	3	16383284	16383284	C	T	ENSSSCG00000040985	KCTD7	potassium channel tetramerization domain containing 7	downstream gene variant
rs333661817	3	16383284	16383284	C	T	ENSSSCG00000040985	KCTD7	potassium channel tetramerization domain containing 7	3 ' UTR variant
rs333661817	3	16383284	16383284	C	T	ENSSSCG00000040985	KCTD7	potassium channel tetramerization domain containing 7	downstream gene variant
rs33366181	3	16383284	163832	C	T	ENSSSCG000000409	KCTD7	potassium channel	downstream gene

7			84			85		tetramerization domain containing 7	variant
rs326053487	3	16383454	16383454	C	G	ENSSSCG00000040985	KCTD7	potassium channel tetramerization domain containing 7	downstream gene variant
rs326053487	3	16383454	16383454	C	G	ENSSSCG00000040985	KCTD7	potassium channel tetramerization domain containing 7	3 ' UTR variant
rs326053487	3	16383454	16383454	C	G	ENSSSCG00000040985	KCTD7	potassium channel tetramerization domain containing 7	downstream gene variant
rs326053487	3	16383454	16383454	C	G	ENSSSCG00000040985	KCTD7	potassium channel tetramerization domain containing 7	downstream gene variant
rs336364520	3	16383470	16383470	G	A	ENSSSCG00000040985	KCTD7	potassium channel tetramerization domain containing 7	downstream gene variant
rs336364520	3	16383470	16383470	G	A	ENSSSCG00000040985	KCTD7	potassium channel tetramerization domain containing 7	3 ' UTR variant
rs336364520	3	16383470	16383470	G	A	ENSSSCG00000040985	KCTD7	potassium channel tetramerization domain containing 7	downstream gene variant
rs336364520	3	16383470	16383470	G	A	ENSSSCG00000040985	KCTD7	potassium channel tetramerization domain containing 7	downstream gene variant
rs329707669	3	16383919	16383919	G	A	ENSSSCG00000040985	KCTD7	potassium channel tetramerization domain containing 7	downstream gene variant
rs329707669	3	16383919	16383919	G	A	ENSSSCG00000040985	KCTD7	potassium channel tetramerization domain	3 ' UTR variant

								containing 7	
rs329707669	3	16383919	16383919	G	A	ENSSSCG00000040985	KCTD7	potassium channel tetramerization domain containing 7	downstream gene variant
rs329707669	3	16383919	16383919	G	A	ENSSSCG00000040985	KCTD7	potassium channel tetramerization domain containing 7	downstream gene variant
rs713060696	3	16383936	16383936	G	A	ENSSSCG00000040985	KCTD7	potassium channel tetramerization domain containing 7	downstream gene variant
rs713060696	3	16383936	16383936	G	A	ENSSSCG00000040985	KCTD7	potassium channel tetramerization domain containing 7	3 ' UTR variant
rs713060696	3	16383936	16383936	G	A	ENSSSCG00000040985	KCTD7	potassium channel tetramerization domain containing 7	downstream gene variant
rs713060696	3	16383936	16383936	G	A	ENSSSCG00000040985	KCTD7	potassium channel tetramerization domain containing 7	downstream gene variant
rs701844432	3	16383956	16383956	A	G	ENSSSCG00000040985	KCTD7	potassium channel tetramerization domain containing 7	downstream gene variant
rs701844432	3	16383956	16383956	A	G	ENSSSCG00000040985	KCTD7	potassium channel tetramerization domain containing 7	3 ' UTR variant
rs701844432	3	16383956	16383956	A	G	ENSSSCG00000040985	KCTD7	potassium channel tetramerization domain containing 7	downstream gene variant
rs701844432	3	16383956	16383956	A	G	ENSSSCG00000040985	KCTD7	potassium channel tetramerization domain containing 7	downstream gene variant

rs324583382	3	16384043	16384043	C	T	ENSSSCG00000040985	KCTD7	potassium channel tetramerization domain containing 7	downstream gene variant
rs324583382	3	16384043	16384043	C	T	ENSSSCG00000040985	KCTD7	potassium channel tetramerization domain containing 7	3 ' UTR variant
rs324583382	3	16384043	16384043	C	T	ENSSSCG00000040985	KCTD7	potassium channel tetramerization domain containing 7	downstream gene variant
rs324583382	3	16384043	16384043	C	T	ENSSSCG00000040985	KCTD7	potassium channel tetramerization domain containing 7	downstream gene variant
rs345204099	3	16384270	16384270	C	T	ENSSSCG00000040985	KCTD7	potassium channel tetramerization domain containing 7	downstream gene variant
rs345204099	3	16384270	16384270	C	T	ENSSSCG00000040985	KCTD7	potassium channel tetramerization domain containing 7	3 ' UTR variant
rs345204099	3	16384270	16384270	C	T	ENSSSCG00000040985	KCTD7	potassium channel tetramerization domain containing 7	downstream gene variant
rs345204099	3	16384270	16384270	C	T	ENSSSCG00000040985	KCTD7	potassium channel tetramerization domain containing 7	downstream gene variant
rs323662654	3	16384658	16384658	T	C	ENSSSCG00000040985	KCTD7	potassium channel tetramerization domain containing 7	downstream gene variant
rs323662654	3	16384658	16384658	T	C	ENSSSCG00000040985	KCTD7	potassium channel tetramerization domain containing 7	3 ' UTR variant
rs32366265	3	16384658	163846	T	C	ENSSSCG000000409	KCTD7	potassium channel	downstream gene

4			58			85		tetramerization domain containing 7	variant
rs323662654	3	16384658	16384658	T	C	ENSSSCG00000040985	KCTD7	potassium channel tetramerization domain containing 7	downstream gene variant
rs332268785	3	16385727	16385727	G	A	ENSSSCG00000040985	KCTD7	potassium channel tetramerization domain containing 7	intron variant
rs332268785	3	16385727	16385727	G	A	ENSSSCG00000040985	KCTD7	potassium channel tetramerization domain containing 7	3 ' UTR variant
rs332268785	3	16385727	16385727	G	A	ENSSSCG00000040985	KCTD7	potassium channel tetramerization domain containing 7	intron variant
rs332268785	3	16385727	16385727	G	A	ENSSSCG00000040985	KCTD7	potassium channel tetramerization domain containing 7	intron variant
rs694561033	3	16385959	16385959	G	C	ENSSSCG00000040985	KCTD7	potassium channel tetramerization domain containing 7	intron variant
rs694561033	3	16385959	16385959	G	C	ENSSSCG00000040985	KCTD7	potassium channel tetramerization domain containing 7	3 ' UTR variant
rs694561033	3	16385959	16385959	G	C	ENSSSCG00000040985	KCTD7	potassium channel tetramerization domain containing 7	intron variant
rs694561033	3	16385959	16385959	G	C	ENSSSCG00000040985	KCTD7	potassium channel tetramerization domain containing 7	intron variant
new	3	16844063	16844064	CG	-	ENSSSCG00000007745	SUMF2	Sulfatase modifying factor 2	upstream gene variant

new	3	16844063	16844064	CG	-	ENSSSCG00000020808	CCT6A	Chaperonin containing TCP1 subunit 6A	3 ' UTR variant
new	3	16844063	16844064	CG	-	ENSSSCG00000033141	new	new	downstream gene variant
new	3	16844063	16844064	CG	-	ENSSSCG00000020808	CCT6A	Chaperonin containing TCP1 subunit 6A	3 ' UTR variant
new	3	16844063	16844064	CG	-	ENSSSCG00000020808	CCT6A	Chaperonin containing TCP1 subunit 6A	3 ' UTR variant
new	3	16844063	16844064	CG	-	ENSSSCG00000007745	SUMF2	Sulfatase modifying factor 2	upstream gene variant
new	3	16844063	16844064	CG	-	ENSSSCG00000007745	SUMF2	Sulfatase modifying factor 2	upstream gene variant
rs330731365	3	16844265	16844265	G	A	ENSSSCG00000007745	SUMF2	Sulfatase modifying factor 2	upstream gene variant
rs330731365	3	16844265	16844265	G	A	ENSSSCG00000020808	CCT6A	Chaperonin containing TCP1 subunit 6A	3 ' UTR variant
rs330731365	3	16844265	16844265	G	A	ENSSSCG00000033141	new	new	downstream gene variant
rs330731365	3	16844265	16844265	G	A	ENSSSCG00000020808	CCT6A	Chaperonin containing TCP1 subunit 6A	3 ' UTR variant
rs330731365	3	16844265	16844265	G	A	ENSSSCG00000020808	CCT6A	Chaperonin containing TCP1 subunit 6A	3 ' UTR variant
rs330731365	3	16844265	16844265	G	A	ENSSSCG00000007745	SUMF2	Sulfatase modifying factor 2	upstream gene variant
rs330731365	3	16844265	16844265	G	A	ENSSSCG00000007745	SUMF2	Sulfatase modifying factor 2	upstream gene variant
rs343913735	3	16844271	16844271	A	G	ENSSSCG00000007745	SUMF2	Sulfatase modifying factor 2	upstream gene variant
rs343913735	3	16844271	16844271	A	G	ENSSSCG00000020808	CCT6A	Chaperonin containing TCP1 subunit 6A	3 ' UTR variant
rs34391373	3	16844271	168442	A	G	ENSSSCG000000331	new	new	downstream gene

5			71			41			variant
rs343913735	3	16844271	16844271	A	G	ENSSSCG00000020808	CCT6A	Chaperonin containing TCP1 subunit 6A	3 ' UTR variant
rs343913735	3	16844271	16844271	A	G	ENSSSCG00000020808	CCT6A	Chaperonin containing TCP1 subunit 6A	3 ' UTR variant
rs343913735	3	16844271	16844271	A	G	ENSSSCG00000007745	SUMF2	Sulfatase modifying factor 2	upstream gene variant
rs343913735	3	16844271	16844271	A	G	ENSSSCG00000007745	SUMF2	Sulfatase modifying factor 2	upstream gene variant
rs1108762720	3	16851119	16851119	A	G	ENSSSCG00000018553	new	new	downstream gene variant
rs1108762720	3	16851119	16851119	A	G	ENSSSCG00000020808	CCT6A	Chaperonin containing TCP1 subunit 6A	synonymous variant
rs1108762720	3	16851119	16851119	A	G	ENSSSCG00000033141	new	new	upstream gene variant
rs1108762720	3	16851119	16851119	A	G	ENSSSCG00000020808	CCT6A	Chaperonin containing TCP1 subunit 6A	synonymous variant
rs1108762720	3	16851119	16851119	A	G	ENSSSCG00000020808	CCT6A	Chaperonin containing TCP1 subunit 6A	synonymous variant
rs341960076	3	16855237	16855237	A	C	ENSSSCG00000018553	new	new	upstream gene variant
rs341960076	3	16855237	16855237	A	C	ENSSSCG00000020808	CCT6A	Chaperonin containing TCP1 subunit 6A	synonymous variant
rs341960076	3	16855237	16855237	A	C	ENSSSCG00000020808	CCT6A	Chaperonin containing TCP1 subunit 6A	synonymous variant
rs341960076	3	16855237	16855237	A	C	ENSSSCG00000020808	CCT6A	Chaperonin containing TCP1 subunit 6A	intron variant
rs323115420	3	16964045	16964045	T	A	ENSSSCG00000029029	ZNF713	zinc finger protein 713	missense variant
rs342012840	3	16971089	16971089	T	C	ENSSSCG00000029029	ZNF713	zinc finger protein 713	5 ' UTR variant

rs324205762	3	16971143	16971143	T	C	ENSSSCG00000029029	ZNF713	zinc finger protein 713	5 ' UTR variant
rs333780109	3	16971156	16971156	T	C	ENSSSCG00000029029	ZNF713	zinc finger protein 713	5 ' UTR variant
rs343615406	3	16971169	16971169	T	C	ENSSSCG00000029029	ZNF713	zinc finger protein 713	5 ' UTR variant
new	3	17082026	17082026	C	T	ENSSSCG00000007753	C16orf58	Chromosome 3 C16orf58 homolog	3 ' UTR variant
new	3	17082026	17082026	C	T	ENSSSCG00000024058	SLC5A2	Solute carrier family 5 member 2	intron variant
new	3	17082026	17082026	C	T	ENSSSCG00000007753	C16orf58	Chromosome 3 C16orf58 homolog	3 ' UTR variant
new	3	17082026	17082026	C	T	ENSSSCG00000024058	SLC5A2	Solute carrier family 5 member 2	intron variant
new	3	17082026	17082026	C	T	ENSSSCG00000007753	C16orf58	Chromosome 3 C16orf58 homolog	3 ' UTR variant
rs344858642	3	17244108	17244108	C	T	ENSSSCG00000007754	ITGAM	integrin subunit alpha M	synonymous variant
rs344858642	3	17244108	17244108	C	T	ENSSSCG00000007754	ITGAM	integrin subunit alpha M	synonymous variant
rs344858642	3	17244108	17244108	C	T	ENSSSCG00000007754	ITGAM	integrin subunit alpha M	synonymous variant
rs344858642	3	17244108	17244108	C	T	ENSSSCG00000007754	ITGAM	integrin subunit alpha M	synonymous variant
rs344858642	3	17244108	17244108	C	T	ENSSSCG00000007754	ITGAM	integrin subunit alpha M	synonymous variant
rs344858642	3	17244108	17244108	C	T	ENSSSCG00000007754	ITGAM	integrin subunit alpha M	synonymous variant
rs344858642	3	17244108	17244108	C	T	ENSSSCG00000007754	ITGAM	integrin subunit alpha M	synonymous variant
rs32694291	3	17246305	172463	G	A	ENSSSCG000000077	ITGAM	integrin subunit alpha	synonymous variant

9			05			54		M	
rs326942919	3	17246305	17246305	G	A	ENSSSCG00000007754	ITGAM	integrin subunit alpha M	synonymous variant
rs326942919	3	17246305	17246305	G	A	ENSSSCG00000007754	ITGAM	integrin subunit alpha M	synonymous variant
rs326942919	3	17246305	17246305	G	A	ENSSSCG00000007754	ITGAM	integrin subunit alpha M	synonymous variant
rs326942919	3	17246305	17246305	G	A	ENSSSCG00000007754	ITGAM	integrin subunit alpha M	synonymous variant
rs326942919	3	17246305	17246305	G	A	ENSSSCG00000007754	ITGAM	integrin subunit alpha M	synonymous variant
rs326942919	3	17246305	17246305	G	A	ENSSSCG00000007754	ITGAM	integrin subunit alpha M	synonymous variant
rs327289001	3	17254444	17254444	C	T	ENSSSCG00000007754	ITGAM	integrin subunit alpha M	missense variant
rs327289001	3	17254444	17254444	C	T	ENSSSCG00000007754	ITGAM	integrin subunit alpha M	missense variant
rs327289001	3	17254444	17254444	C	T	ENSSSCG00000007754	ITGAM	integrin subunit alpha M	missense variant
rs327289001	3	17254444	17254444	C	T	ENSSSCG00000007754	ITGAM	integrin subunit alpha M	missense variant
rs327289001	3	17254444	17254444	C	T	ENSSSCG00000007754	ITGAM	integrin subunit alpha M	missense variant
rs327289001	3	17254444	17254444	C	T	ENSSSCG00000007754	ITGAM	integrin subunit alpha M	missense variant
rs327289001	3	17254444	17254444	C	T	ENSSSCG00000007754	ITGAM	integrin subunit alpha M	missense variant
rs327947675	3	17399455	17399455	A	G	ENSSSCG00000007765	PRSS53	serine protease 53	downstream gene variant
rs327947675	3	17399455	17399455	A	G	ENSSSCG00000026817	ZNF646	zinc finger protein 646	synonymous variant

rs327947675	3	17399455	17399455	A	G	ENSSSCG00000026817	ZNF646	zinc finger protein 646	synonymous variant
rs327947675	3	17399455	17399455	A	G	ENSSSCG00000035364	VKORC1	vitamin K epoxide reductase complex subunit 1	downstream gene variant
rs327947675	3	17399455	17399455	A	G	ENSSSCG00000007765	PRSS53	serine protease 53	downstream gene variant
rs327947675	3	17399455	17399455	A	G	ENSSSCG00000007763	BCKDK	Branched chain keto acid dehydrogenase kinase	5 ' UTR variant
rs327947675	3	17399455	17399455	A	G	ENSSSCG00000026817	ZNF646	zinc finger protein 646	synonymous variant
rs337670844	3	17399477	17399477	C	T	ENSSSCG00000007765	PRSS53	serine protease 53	downstream gene variant
rs337670844	3	17399477	17399477	C	T	ENSSSCG00000026817	ZNF646	zinc finger protein 646	missense variant
rs337670844	3	17399477	17399477	C	T	ENSSSCG00000026817	ZNF646	zinc finger protein 646	missense variant
rs337670844	3	17399477	17399477	C	T	ENSSSCG00000035364	VKORC1	vitamin K epoxide reductase complex subunit 1	downstream gene variant
rs337670844	3	17399477	17399477	C	T	ENSSSCG00000007765	PRSS53	serine protease 53	downstream gene variant
rs337670844	3	17399477	17399477	C	T	ENSSSCG00000007763	BCKDK	vitamin K epoxide reductase complex subunit 1	5 ' UTR variant
rs337670844	3	17399477	17399477	C	T	ENSSSCG00000026817	ZNF646	zinc finger protein 646	missense variant
rs322669402	3	17459844	17459844	A	G	ENSSSCG00000007782	SETD1A	SET domain containing 1A, histone lysine methyltransferase	downstream gene variant

rs322669402	3	17459844	17459844	A	G	ENSSSCG00000021238	STX1B	syntaxin 1B	downstream gene variant
rs322669402	3	17459844	17459844	A	G	ENSSSCG00000032369	HSD3B7	hydroxy-delta-5-steroid dehydrogenase, 3 beta- and steroid delta-isomerase 7	3 ' UTR variant
rs322669402	3	17459844	17459844	A	G	ENSSSCG00000021238	STX1B	syntaxin 1B	downstream gene variant
rs322669402	3	17459844	17459844	A	G	ENSSSCG00000021238	STX1B	syntaxin 1B	downstream gene variant
rs322669402	3	17459844	17459844	A	G	ENSSSCG00000032369	HSD3B7	hydroxy-delta-5-steroid dehydrogenase, 3 beta- and steroid delta-isomerase 7	3 ' UTR variant
rs340480028	3	17460060	17460060	G	C	ENSSSCG00000007782	SETD1A	SET domain containing 1A, histone lysine methyltransferase	downstream gene variant
rs340480028	3	17460060	17460060	G	C	ENSSSCG00000021238	STX1B	syntaxin 1B	downstream gene variant
rs340480028	3	17460060	17460060	G	C	ENSSSCG00000032369	HSD3B7	hydroxy-delta-5-steroid dehydrogenase, 3 beta- and steroid delta-isomerase 7	3 ' UTR variant
rs340480028	3	17460060	17460060	G	C	ENSSSCG00000021238	STX1B	syntaxin 1B	downstream gene variant
rs340480028	3	17460060	17460060	G	C	ENSSSCG00000021238	STX1B	syntaxin 1B	downstream gene variant
rs340480028	3	17460060	17460060	G	C	ENSSSCG00000032369	HSD3B7	hydroxy-delta-5-steroid dehydrogenase, 3 beta- and steroid delta-isomerase 7	3 ' UTR variant

rs793318116	3	17460656	17460656	C	T	ENSSSCG00000007782	SETD1A	SET domain containing 1A, histone lysine methyltransferase	downstream gene variant
rs793318116	3	17460656	17460656	C	T	ENSSSCG00000021238	STX1B	syntaxin 1B	downstream gene variant
rs793318116	3	17460656	17460656	C	T	ENSSSCG00000032369	HSD3B7	hydroxy-delta-5-steroid dehydrogenase, 3 beta- and steroid delta-isomerase 7	synonymous variant
rs793318116	3	17460656	17460656	C	T	ENSSSCG00000021238	STX1B	syntaxin 1B	downstream gene variant
rs793318116	3	17460656	17460656	C	T	ENSSSCG00000021238	STX1B	syntaxin 1B	downstream gene variant
rs793318116	3	17460656	17460656	C	T	ENSSSCG00000032369	HSD3B7	hydroxy-delta-5-steroid dehydrogenase, 3 beta- and steroid delta-isomerase 7	synonymous variant
rs339771716	3	17461694	17461694	C	T	ENSSSCG00000007782	SETD1A	SET domain containing 1A, histone lysine methyltransferase	downstream gene variant
rs339771716	3	17461694	17461694	C	T	ENSSSCG00000021238	STX1B	syntaxin 1B	downstream gene variant
rs339771716	3	17461694	17461694	C	T	ENSSSCG00000032369	HSD3B7	hydroxy-delta-5-steroid dehydrogenase, 3 beta- and steroid delta-isomerase 7	synonymous variant
rs339771716	3	17461694	17461694	C	T	ENSSSCG00000021238	STX1B	syntaxin 1B	downstream gene variant
rs339771716	3	17461694	17461694	C	T	ENSSSCG00000032369	HSD3B7	hydroxy-delta-5-steroid dehydrogenase, 3 beta- and steroid delta-	synonymous variant

								isomerase 7	
rs325937498	3	17466691	17466691	G	A	ENSSSCG00000007782	SETD1A	SET domain containing 1A, histone lysine methyltransferase	synonymous variant
rs325937498	3	17466691	17466691	G	A	ENSSSCG00000032369	HSD3B7	hydroxy-delta-5-steroid dehydrogenase, 3 beta- and steroid delta-isomerase 7	upstream gene variant
rs325937498	3	17466691	17466691	G	A	ENSSSCG00000032369	HSD3B7	hydroxy-delta-5-steroid dehydrogenase, 3 beta- and steroid delta-isomerase 7	upstream gene variant
rs343894209	3	17466745	17466745	C	T	ENSSSCG00000007782	SETD1A	SET domain containing 1A, histone lysine methyltransferase	synonymous variant
rs343894209	3	17466745	17466745	C	T	ENSSSCG00000032369	HSD3B7	hydroxy-delta-5-steroid dehydrogenase, 3 beta- and steroid delta-isomerase 7	upstream gene variant
rs343894209	3	17466745	17466745	C	T	ENSSSCG00000032369	HSD3B7	hydroxy-delta-5-steroid dehydrogenase, 3 beta- and steroid delta-isomerase 7	upstream gene variant
rs339276563	3	17468216	17468216	T	C	ENSSSCG00000007782	SETD1A	SET domain containing 1A, histone lysine methyltransferase	synonymous variant
rs339276563	3	17468216	17468216	T	C	ENSSSCG00000032369	HSD3B7	hydroxy-delta-5-steroid dehydrogenase, 3 beta- and steroid delta-isomerase 7	upstream gene variant
rs33927656	3	17468216	174682	T	C	ENSSSCG000000323	HSD3B7	hydroxy-delta-5-steroid	upstream gene variant

3			16			69		dehydrogenase, 3 beta- and steroid delta-isomerase 7	
rs330957838	3	17468302	17468302	T	C	ENSSSCG00000007782	SETD1A	SET domain containing 1A, histone lysine methyltransferase	missense variant
rs330957838	3	17468302	17468302	T	C	ENSSSCG00000032369	HSD3B7	hydroxy-delta-5-steroid dehydrogenase, 3 beta- and steroid delta-isomerase 7	upstream gene variant
rs345676220	3	17491423	17491423	G	A	ENSSSCG00000007770	ORAI3	ORAI calcium release-activated calcium modulator 3	3 ' UTR variant
rs345676220	3	17491423	17491423	G	A	ENSSSCG00000007782	SETD1A	SET domain containing 1A, histone lysine methyltransferase	upstream gene variant
rs324198007	3	17491763	17491763	C	A	ENSSSCG00000007770	ORAI3	ORAI calcium release-activated calcium modulator 3	synonymous variant
rs324198007	3	17491763	17491763	C	A	ENSSSCG00000007782	SETD1A	SET domain containing 1A, histone lysine methyltransferase	upstream gene variant
rs321032880	3	17584544	17584544	G	A	ENSSSCG00000007776	BCL7C	BAF chromatin remodeling complex subunit BCL7C	3 ' UTR variant
rs339255619	3	17588899	17588899	G	A	ENSSSCG00000007776	BCL7C	BAF chromatin remodeling complex subunit BCL7C	3 ' UTR variant
rs323862382	3	17589735	17589735	C	T	ENSSSCG00000007776	BCL7C	BAF chromatin remodeling complex subunit BCL7C	3 ' UTR variant

rs697381317	3	17591079	17591079	A	G	ENSSSCG00000007776	BCL7C	BAF chromatin remodeling complex subunit BCL7C	downstream gene variant
rs699677308	3	17591133	17591133	C	G	ENSSSCG00000007776	BCL7C	BAF chromatin remodeling complex subunit BCL7C	downstream gene variant
rs706065940	3	17591138	17591138	C	T	ENSSSCG00000007776	BCL7C	BAF chromatin remodeling complex subunit BCL7C	downstream gene variant
new	3	17591144	17591144	G	T	ENSSSCG00000007776	BCL7C	BAF chromatin remodeling complex subunit BCL7C	downstream gene variant
rs344892486	3	17610468	17610468	T	C	ENSSSCG00000007780	ZNF629	zinc finger protein 629	5 ' UTR variant
rs713239492	3	17612980	17612980	C	T	ENSSSCG00000007780	ZNF629	zinc finger protein 629	synonymous variant
rs335540465	3	17613531	17613531	T	C	ENSSSCG00000007780	ZNF629	zinc finger protein 629	synonymous variant
new	3	17617319	17617323	CAGAA	-	ENSSSCG00000007780	ZNF629	zinc finger protein 629	3 ' UTR variant
new	3	17617319	17617323	CAGAA	-	ENSSSCG00000007786	RNF40	ring finger protein 40	downstream gene variant
new	3	17617319	17617323	CAGAA	-	ENSSSCG00000007786	RNF40	ring finger protein 40	downstream gene variant
new	3	17617319	17617323	CAGAA	-	ENSSSCG00000007786	RNF40	ring finger protein 40	downstream gene variant
new	3	17617319	17617323	CAGAA	-	ENSSSCG00000007786	RNF40	ring finger protein 40	downstream gene variant
rs1107804156	3	17618533	17618533	G	T	ENSSSCG00000007780	ZNF629	zinc finger protein 629	intron variant
rs11078041	3	17618533	176185	G	T	ENSSSCG000000077	RNF40	ring finger protein 40	downstream gene

56			33			86			variant
rs1107804156	3	17618533	17618533	G	T	ENSSSCG00000007786	RNF40	ring finger protein 40	downstream gene variant
rs1107804156	3	17618533	17618533	G	T	ENSSSCG00000007786	RNF40	ring finger protein 40	downstream gene variant
rs1107804156	3	17618533	17618533	G	T	ENSSSCG00000007786	RNF40	ring finger protein 40	downstream gene variant
rs789266896	3	17628688	17628688	T	G	ENSSSCG00000007780	ZNF629	zinc finger protein 629	3 ' UTR variant
rs789266896	3	17628688	17628688	T	G	ENSSSCG00000007786	RNF40	ring finger protein 40	missense variant
rs789266896	3	17628688	17628688	T	G	ENSSSCG00000007786	RNF40	ring finger protein 40	missense variant
rs789266896	3	17628688	17628688	T	G	ENSSSCG00000007786	RNF40	ring finger protein 40	missense variant
rs789266896	3	17628688	17628688	T	G	ENSSSCG00000007786	RNF40	ring finger protein 40	missense variant
new	3	51803551	51803551	T	C	ENSSSCG00000008159	IL18R1	Interleukin 18 receptor 1	3 ' UTR variant
new	3	51803551	51803551	T	C	ENSSSCG00000032795	IL1RL1	Interleukin 1 receptor like 1	3 ' UTR variant
new	4	128903094	128903094	A	T	ENSSSCG00000047605		new	intron variant
new	4	128903094	128903094	A	T	ENSSSCG00000006928	LMO4	LIM domain only 4	intron variant
new	4	128903094	128903094	A	T	ENSSSCG00000006928	LMO4	LIM domain only 4	intron variant
rs345481021	6	80824380	80824380	T	G	ENSSSCG00000003527	EPHB2	EPH receptor B2	synonymous variant
rs345481021	6	80824380	80824380	T	G	ENSSSCG00000003527	EPHB2	EPH receptor B2	synonymous variant

rs345481021	6	80824380	80824380	T	G	ENSSSCG00000003527	EPHB2	EPH receptor B2	synonymous variant
rs345481021	6	80824380	80824380	T	G	ENSSSCG00000003527	EPHB2	EPH receptor B2	synonymous variant
rs326115442	6	80824416	80824416	T	C	ENSSSCG00000003527	EPHB2	EPH receptor B2	synonymous variant
rs326115442	6	80824416	80824416	T	C	ENSSSCG00000003527	EPHB2	EPH receptor B2	synonymous variant
rs326115442	6	80824416	80824416	T	C	ENSSSCG00000003527	EPHB2	EPH receptor B2	synonymous variant
rs326115442	6	80824416	80824416	T	C	ENSSSCG00000003527	EPHB2	EPH receptor B2	synonymous variant
rs327572607	6	80837841	80837841	T	C	ENSSSCG00000003527	EPHB2	EPH receptor B2	synonymous variant
rs327572607	6	80837841	80837841	T	C	ENSSSCG00000003527	EPHB2	EPH receptor B2	synonymous variant
rs327572607	6	80837841	80837841	T	C	ENSSSCG00000003527	EPHB2	EPH receptor B2	synonymous variant
rs327572607	6	80837841	80837841	T	C	ENSSSCG00000003527	EPHB2	EPH receptor B2	synonymous variant
rs81389091	6	80842615	80842615	T	C	ENSSSCG00000003527	EPHB2	EPH receptor B2	3 ' UTR variant
rs81389091	6	80842615	80842615	T	C	ENSSSCG00000003527	EPHB2	EPH receptor B2	3 ' UTR variant
rs81389091	6	80842615	80842615	T	C	ENSSSCG00000003527	EPHB2	EPH receptor B2	3 ' UTR variant
rs81389091	6	80842615	80842615	T	C	ENSSSCG00000003527	EPHB2	EPH receptor B2	3 ' UTR variant
new	6	80843072	80843073	CA	-	ENSSSCG00000003527	EPHB2	EPH receptor B2	3 ' UTR variant
	6	80843072	80843	CA	-	ENSSSCG00000003	EPHB2	EPH receptor B2	3 ' UTR variant

			73			27			
new	6	80843072	80843073	CA	-	ENSSSCG00000003527	EPHB2	EPH receptor B2	3 ' UTR variant
new	6	80843072	80843073	CA	-	ENSSSCG00000003527	EPHB2	EPH receptor B2	3 ' UTR variant
new	6	80843074	80843081	GGCGGGCA	-	ENSSSCG00000003527	EPHB2	EPH receptor B2	3 ' UTR variant
new	6	80843074	80843081	GGCGGGCA	-	ENSSSCG00000003527	EPHB2	EPH receptor B2	3 ' UTR variant
new	6	80843074	80843081	GGCGGGCA	-	ENSSSCG00000003527	EPHB2	EPH receptor B2	3 ' UTR variant
new	6	80843074	80843081	GGCGGGCA	-	ENSSSCG00000003527	EPHB2	EPH receptor B2	3 ' UTR variant
rs342283188	6	80843098	80843098	T	G	ENSSSCG00000003527	EPHB2	EPH receptor B2	3 ' UTR variant
rs342283188	6	80843098	80843098	T	G	ENSSSCG00000003527	EPHB2	EPH receptor B2	3 ' UTR variant
rs342283188	6	80843098	80843098	T	G	ENSSSCG00000003527	EPHB2	EPH receptor B2	3 ' UTR variant
rs342283188	6	80843098	80843098	T	G	ENSSSCG00000003527	EPHB2	EPH receptor B2	3 ' UTR variant
rs81389092	6	80843265	80843265	A	G	ENSSSCG00000003527	EPHB2	EPH receptor B2	3 ' UTR variant
rs81389092	6	80843265	80843265	A	G	ENSSSCG00000003527	EPHB2	EPH receptor B2	3 ' UTR variant
rs81389092	6	80843265	80843265	A	G	ENSSSCG00000003527	EPHB2	EPH receptor B2	3 ' UTR variant
rs81389092	6	80843265	80843265	A	G	ENSSSCG00000003527	EPHB2	EPH receptor B2	3 ' UTR variant
rs346223430	6	80843336	80843336	C	T	ENSSSCG00000003527	EPHB2	EPH receptor B2	3 ' UTR variant

rs346223430	6	80843336	80843336	C	T	ENSSSCG00000003527	EPHB2	EPH receptor B2	3 ' UTR variant
rs346223430	6	80843336	80843336	C	T	ENSSSCG00000003527	EPHB2	EPH receptor B2	3 ' UTR variant
rs346223430	6	80843336	80843336	C	T	ENSSSCG00000003527	EPHB2	EPH receptor B2	3 ' UTR variant
rs325089032	6	81571496	81571496	G	A	ENSSSCG00000025440	ELOA	elongin A	missense variant
rs325089032	6	81571496	81571496	G	A	ENSSSCG00000025440	ELOA	elongin A	missense variant
rs709055765	6	82481037	82481037	T	A	ENSSSCG00000038994	CLIC4	chloride intracellular channel 4	3 ' UTR variant
rs709055765	6	82481037	82481037	T	A	ENSSSCG00000038994	CLIC4	chloride intracellular channel 4	3 ' UTR variant
rs323015047	6	119763116	119763116	T	C	ENSSSCG00000027700	RPRD1A	Regulation of nuclear pre-mRNA domain containing 1A	3 ' UTR variant
rs323015047	6	119763116	119763116	T	C	ENSSSCG00000027700	RPRD1A	Regulation of nuclear pre-mRNA domain containing 1A	3 ' UTR variant
rs323015047	6	119763116	119763116	T	C	ENSSSCG00000027700	RPRD1A	Regulation of nuclear pre-mRNA domain containing 1A	downstream gene variant
rs330195537	10	43520964	43520964	G	C	ENSSSCG00000011033	VIM	Vimentin	intron variant
rs330195537	10	43520964	43520964	G	C	ENSSSCG00000021309	TRDMT1	TRNA aspartic acid methyltransferase 1	upstream gene variant
rs330195537	10	43520964	43520964	G	C	ENSSSCG00000011033	VIM	Vimentin	intron variant
rs330195537	10	43520964	43520964	G	C	ENSSSCG00000021309	TRDMT1	TRNA aspartic acid methyltransferase 1	upstream gene variant

rs330195537	10	43520964	43520964	G	C	ENSSSCG00000011033	VIM	Vimentin	intron variant
rs331463738	12	38129509	38129509	T	C	ENSSSCG00000017690	DHRS11	Mitochondrial rRNA methyltransferase 1	3 ' UTR variant
rs331463738	12	38129509	38129509	T	C	ENSSSCG00000017690	DHRS11	Mitochondrial rRNA methyltransferase 1	intron variant
rs331463738	12	38129509	38129509	T	C	ENSSSCG00000017690	DHRS11	Mitochondrial rRNA methyltransferase 1	3 ' UTR variant
rs331463738	12	38129509	38129509	T	C	ENSSSCG00000017690	DHRS11	Mitochondrial rRNA methyltransferase 1	downstream gene variant
rs340781986	12	38624687	38624687	G	A	ENSSSCG00000017694	ACACA	Acetyl-CoA carboxylase alpha	synonymous variant
rs340781986	12	38624687	38624687	G	A	ENSSSCG00000017694	ACACA	Acetyl-CoA carboxylase alpha	synonymous variant
rs340781986	12	38624687	38624687	G	A	ENSSSCG00000017694	ACACA	Acetyl-CoA carboxylase alpha	synonymous variant
rs340781986	12	38624687	38624687	G	A	ENSSSCG00000017694	ACACA	Acetyl-CoA carboxylase alpha	synonymous variant
rs324236192	12	38624714	38624714	G	A	ENSSSCG00000017694	ACACA	Acetyl-CoA carboxylase alpha	synonymous variant
rs324236192	12	38624714	38624714	G	A	ENSSSCG00000017694	ACACA	Acetyl-CoA carboxylase alpha	synonymous variant
rs324236192	12	38624714	38624714	G	A	ENSSSCG00000017694	ACACA	Acetyl-CoA carboxylase alpha	synonymous variant
rs324236192	12	38624714	38624714	G	A	ENSSSCG00000017694	ACACA	Acetyl-CoA carboxylase alpha	synonymous variant
rs710955781	13	50397619	50397619	C	T	ENSSSCG00000011504	EOGT	EGF domain specific O-linked N-acetylglucosamine transferase	3 ' UTR variant
rs71095578	13	50397619	503976	C	T	ENSSSCG000000115	EOGT	EGF domain specific	3 ' UTR variant

1			19			04		O-linked N-acetylglucosamine transferase	
rs710955781	13	50397619	50397619	C	T	ENSSSCG00000011504	EOGT	EGF domain specific O-linked N-acetylglucosamine transferase	3 ' UTR variant
new	13	108676707	108676707	G	A	ENSSSCG00000029608	SEC62	SEC62 homolog, preprotein translocation factor	frameshift variant
new	13	108676707	108676707	G	A	ENSSSCG00000029608	SEC62	SEC62 homolog, preprotein translocation factor	frameshift variant
new	13	108676707	108676707	G	A	ENSSSCG00000029608	SEC62	SEC62 homolog, preprotein translocation factor	frameshift variant
new	13	207513656	207513656	T	C	ENSSSCG00000025133	ITGB2	Integrin subunit beta 2	3 ' UTR variant
new	13	207513656	207513656	T	C	ENSSSCG00000025133	ITGB2	Integrin subunit beta 2	downstream gene variant
new	13	207513656	207513656	T	C	ENSSSCG00000025133	ITGB2	Integrin subunit beta 2	downstream gene variant
new	13	207513656	207513656	T	C	ENSSSCG00000025133	ITGB2	Integrin subunit beta 2	3 ' UTR variant
rs699968948	14	48800543	48800543	G	C	ENSSSCG00000031037	new	new	synonymous variant
rs1110474609	17	34777839	34777839	G	A	ENSSSCG00000036081	TBC1D20	TBC1 domain family member 20	intron variant
rs334463568	18	11660822	11660822	G	A	ENSSSCG00000016520	CREB3L2	CAMP responsive element binding protein 3 like 2	3 ' UTR variant

rs334463568	18	11660822	11660822	G	A	ENSSSCG00000016520	CREB3L2	CAMP responsive element binding protein 3 like 2	downstream gene variant

Additional file 3: Polymorphisms, genes and regions associated to umbilical hernia in pigs annotated in the exome, transcriptome and GWAS analyses.

Exome and Transcriptome

SNP name	Positions	Allele (ref / alt)	Symbol	Gene name	Consequence (transcript)
rs345481021	6:80824380	T/G	EPHB2	EPH receptor B2	synonymous variant
rs326115442	6:80824416	T/C	EPHB2	EPH receptor B2	synonymous variant
rs327572607	6:80837841	T/C	EPHB2	EPH receptor B2	synonymous variant
rs81389091	6:80842615	T/C	EPHB2	EPH receptor B2	3 ' UTR variant
rs342283188	6:80843098	T/G	EPHB2	EPH receptor B2	3 ' UTR variant
rs81389092	6:80843265	A/G	EPHB2	EPH receptor B2	3 ' UTR variant
rs346223430	6;80843336	C/T	EPHB2	EPH receptor B2	3' UTR variant / Downstream
new	6:80843072	CA/-	EPHB2	EPH receptor B2	3 ' UTR variant
new	6:80843074 80843081	GGCG GGCA/-	EPHB2	EPH receptor B2	3 ' UTR variant
new	3:16078557	G/A	TYW1	tRNA-yW synthesizing protein 1 homolog	synonymous variant
new	3:16844063	GC/-	SUMF2	sulfatase modifying factor 2	upstream gene variant
rs330731365	3:16844265	G/A	SUMF2	sulfatase modifying factor 2	upstream gene variant
rs343913735	3:16844271	A/G	SUMF2	sulfatase modifying factor 2	upstream gene variant
rs326942919	3:17246305	G/A	ITGAM	integrin subunit alpha M	synonymous variant

rs344858642	3:17244108	C/T	ITGAM	integrin subunit alpha M	synonymous variant
rs327289001	3:17254444	C/T	ITGAM	integrin subunit alpha M	missense variant
rs327947675	3:17399455	A/G	BCKDK	Branched chain keto acid dehydrogenase kinase	5 ' UTR variant
rs337670844	3:17399477	C/T	BCKDK	vitamin K epoxide reductase complex subunit 1	5 ' UTR variant
rs327947675	3:17399455	A/G	PRSS53	serine protease 53	downstream gene variant
rs337670844	3:17399477	C/T	PRSS53	serine protease 53	downstream gene variant
rs345676220	3:17491423	G/A	ORAI3	ORAI calcium release-activated calcium modulator 3	3 ' UTR variant
rs324198007	3:17491763	C/A	ORAI3	ORAI calcium release-activated calcium modulator 3	synonymous variant
rs344892486	3:17610468	T/C	ZNF629	zinc finger protein 629	5 ' UTR variant
rs335540465	3:17613531	T/C	ZNF629	zinc finger protein 629	synonymous variant
new	3:17617319 17617323	CAGA A/-	ZNF629	zinc finger protein 629	3 ' UTR variant
rs1107804156	3:17618533	G/T	ZNF629	zinc finger protein 629	intron variant
rs789266896	3:17628688	T/G	ZNF629	zinc finger protein 629	3 ' UTR variant
rs322669402	3:17459844	A/G	SETD1A	SET domain containing 1A, histone lysine methyltransferase	downstream gene variant
rs793318116	3:17460656	C/T	SETD1A	SET domain containing 1A, histone lysine methyltransferase	downstream gene variant
rs339276563	3:17468216	T/C	SETD1A	SET domain containing 1A, histone lysine	synonymous variant

				methyltransferase	
rs330957838	3:17468302	T/C	SETD1A	SET domain containing 1A, histone lysine methyltransferase	missense variant
rs345676220	3:17491423	G/A	SETD1A	SET domain containing 1A, histone lysine methyltransferase	upstream gene variant
new	3:17617319 17617323	CAGA A/-	RNF40	ring finger protein 40	downstream gene variant
rs1107804156	3:17618533	G/T	RNF40	ring finger protein 40	downstream gene variant
rs789266896	3:17628688	T/G	RNF40	ring finger protein 40	missense variant
rs331463738	12:38129509	T/C	DHRS11	Mitochondrial rRNA methyltransferase 1	3 ' UTR variant
rs340781986	12:38624687	G/A	ACACA	acetyl-CoA carboxylase alpha	synonymous variant
rs324236192	12:38624714	G/A	ACACA	acetyl-CoA carboxylase alpha	synonymous variant
rs1108762720	3:16851119	A/G	ENSSSCG00000018553	new	downstream gene variant
rs341960076	3:16855237	A/C	ENSSSCG00000018553	new	upstream gene variant
rs330731365	3:16844265	G/A	CCT6A	chaperonin containing TCP1 subunit 6A	3 ' UTR variant
rs343913735	3:16844271	A/G	CCT6A	chaperonin containing TCP1 subunit 6A	3 ' UTR variant
rs1108762720	3:16851119	A/G	CCT6A	chaperonin containing TCP1 subunit 6A	synonymous variant
rs341960076	3:16855237	A/C	CCT6A	chaperonin containing TCP1 subunit 6A	intron variant
rs339771716	3:17461694	C/T	STX1B	syntaxin 1B	downstream gene variant

rs322669402	3:17459844	A/G	STX1B	syntaxin 1B	downstream gene variant
rs325089032	6:81571496	G/A	ELOA	elongin A	missense variant
rs327947675	3:17399455	A/G	ZNF646	zinc finger protein 646	synonymous variant
rs337670844	3:17399477	C/T	ZNF646	zinc finger protein 646	missense variant
rs323115420	3:16964045	T/A	ZNF713	zinc finger protein 713	missense variant
rs342012840	3:16971089	T/C	ZNF713	zinc finger protein 713	5 ' UTR variant
rs324205762	3:16971143	T/C	ZNF713	zinc finger protein 713	5 ' UTR variant
rs333780109	3:16971156	T/C	ZNF713	zinc finger protein 713	5 ' UTR variant
rs343615406	3:16971169	T/C	ZNF713	zinc finger protein 713	5 ' UTR variant
rs322669402	3:17459844	A/G	HSD3B7	hydroxy-delta-5-steroid dehydrogenase, 3 beta- and steroid delta-isomerase 7	3 ' UTR variant
rs793318116	3:17460656	C/T	HSD3B7	hydroxy-delta-5-steroid dehydrogenase, 3 beta- and steroid delta-isomerase 7	synonymous variant
rs336707684	3:17466842	A/G	HSD3B7	hydroxy-delta-5-steroid dehydrogenase, 3 beta- and steroid delta-isomerase 7	upstream gene variant
rs339276563	3:17468216	T/C	HSD3B7	hydroxy-delta-5-steroid dehydrogenase, 3 beta- and steroid delta-isomerase 7	upstream gene variant
rs330957838	3:17468302	T/C	HSD3B7	hydroxy-delta-5-steroid dehydrogenase, 3 beta- and steroid delta-isomerase 7	upstream gene variant
new	3:1684406 16844064	CG/-	ENSSSCG000000	new	downstream gene variant

	3		33141		
rs330731365	3:16844265	G/A	ENSSSCG00000033141	new	downstream gene variant
rs343913735	3:16844271	A/G	ENSSSCG00000033141	new	downstream gene variant
rs1108762720	3:16851119	A/G	ENSSSCG00000033141	new	upstream gene variant
rs327947675	3:17399455	A/G	VKORC1	vitamin K epoxide reductase complex subunit 1	downstream gene variant
rs337670844	3:17399477	C/T	VKORC1	vitamin K epoxide reductase complex subunit 1	downstream gene variant
rs320729536	3:16382619	T/C	KCTD7	potassium channel tetramerization domain containing 7	3 ' UTR variant
rs323726488	3:16382619	T/C	KCTD7	potassium channel tetramerization domain containing 7	downstream gene variant
rs333208968	3:16383246	C/T	KCTD7	potassium channel tetramerization domain containing 7	3 ' UTR variant
rs333661817	3:16383246	C/T	KCTD7	potassium channel tetramerization domain containing 7	downstream gene variant
rs326053487	3:16383250	G/A	KCTD7	potassium channel tetramerization domain containing 7	3 ' UTR variant
rs329707669	3:16383284	C/T	KCTD7	potassium channel tetramerization domain containing 7	downstream gene variant

rs713060696	3:16383284	C/T	KCTD7	potassium channel tetramerization domain containing 7	3 ' UTR variant
rs701844432	3:16383454	C/G	KCTD7	potassium channel tetramerization domain containing 7	3 ' UTR variant
rs324583382	3:16383454	C/G	KCTD7	potassium channel tetramerization domain containing 7	downstream gene variant
rs345204099	3;16383763	A/G	KCTD7	potassium channel tetramerization domain containing 7	3 ' UTR variant
rs327324699	3:16383763	A/G	KCTD7	potassium channel tetramerization domain containing 7	downstream gene variant
rs323662654	3:16383936	G/A	KCTD7	potassium channel tetramerization domain containing 7	downstream gene variant
rs332268785	3:16384043	C/T	KCTD7	potassium channel tetramerization domain containing 7	3 ' UTR variant
new	3:16383956	A/G	KCTD7	potassium channel tetramerization domain containing 7	downstream gene variant

GWAS Region 11

Chromosome	**Positions**		**Symbol**	**Gene name**
3	17136161	17273657-1	ITGAM	integrin subunit alpha M
3	17375352	17402204-1	BCKDK	branched chain keto acid dehydrogenase kinase
3	17390437	17396852 1	PRSS53	serine protease 53

3	17416842	17431053-1	STX4	syntaxin 4
3	17439581	17459516 1	STX1B	syntaxin 1B
3	17384088	17406624-1	ZNF646	zinc finger protein 646
3	17386441	17396840 1	VKORC1	vitamin K epoxide reductase complex subunit 1
3	17431140	17439154 1	lncRNA	long-non-coding RNA

Additional file 4: GO biological process enriched based on the candidate genes identified for umbilical hernia through the analyses of the whole exome, RNA sequencing and GWAS data.

DAVID Bioprocesses	**Enriched genes**
GO:0072657 protein localization to membrane	RABGEF1, SEC62, STX4, ITGB2, ROCK2
GO:0050865 regulation of cell activation	UBASH3B, DUSP3, LMO4, RABGEF1, STX4
GO:0008152 metabolic process	DUSP3, RABGEF1, TYW1, FBXL19, ACACA, LMO4, ZNF629, VIM, FYN, FUCA1, RARB, UBASH3B, SRSF7, ELOA, MMS19, GALM, SLC66A3, EOGT, DEDD, RNF40, PHKG2, RPRD1A, ITGB2, STX1B, ZNF713, SLC22A11, LYPLA2, FBLN5, NCOA7, ROCK2
GO:0090150 establishment of protein localization to membrane	RABGEF1, SEC62, STX4, ROCK2
GO:0051130 positive regulation of cellular component organization	EPHB2, RNF40, STX4, PFDN2, FYN, ITGB2, ROCK2
GO:0003279 cardiac septum development	LMO4, LUZP1, RARB
GO:0031400 negative regulation of protein modification process	UBASH3B, DUSP3, RABGEF1, FYN, ROCK2

GO:0040011 locomotion	LMO4, EPHB2, RABGEF1, CNR2, STX4, FYN, ITGB2, ROCK2
GO:0050866 negative regulation of cell activation	UBASH3B, DUSP3, RABGEF1
GO:0019222 regulation of metabolic process	UBASH3B, DUSP3, SRSF7, RABGEF1, ELOA, MMS19, RNF40, RPRD1A, ITGB2, LMO4, ZNF713, ZNF629, VIM, FBLN5, FYN, RARB, NCOA7, ROCK2
GO:0051128 regulation of cellular component organization	LMO4, EPHB2, RABGEF1, RNF40, STX4, PFDN2, VIM, FYN, ITGB2, ROCK2
GO:0060255 regulation of macromolecule metabolic process	UBASH3B, SRSF7, DUSP3, RABGEF1, ELOA, MMS19, RNF40, RPRD1A, ITGB2, LMO4, ZNF713, ZNF629, VIM, FYN, RARB, NCOA7, ROCK2
GO:0031323 regulation of cellular metabolic process	UBASH3B, DUSP3, SRSF7, RABGEF1, ELOA, MMS19, RNF40, RPRD1A, ITGB2, LMO4, ZNF713, ZNF629, FBLN5, FYN, RARB, NCOA7, ROCK2
GO:0051129 negative regulation of cellular component organization	LMO4, EPHB2, RABGEF1, VIM, ROCK2
GO:0048598 embryonic morphogenesis	LMO4, EPHB2, LUZP1, RARB, ITGB2
GO:0051241 negative regulation of multicellular organismal process	UBASH3B, EPHB2, RABGEF1, VIM, RARB, ROCK2
GO:0051674 localization of cell	LMO4, RABGEF1, CNR2, STX4, FYN, ITGB2, ROCK2
GO:0048870 cell motility	LMO4, RABGEF1, CNR2, STX4, FYN, ITGB2, ROCK2
GO:0006928 movement of cell or subcellular component	LMO4, EPHB2, RABGEF1, CNR2, STX4, FYN, ITGB2, ROCK2
GO:0003205 cardiac chamber development	LMO4, LUZP1, RARB
GO:0010468 regulation of gene expression	SRSF7, LMO4, ELOA, ZNF629, MMS19, ZNF713, RNF40, VIM, FYN, RARB, RPRD1A, NCOA7, ROCK2
GO:0009888 tissue development	CREB3L2, LMO4, MYLPF, VIM, LUZP1, RARB, ITGB2, ROCK2

GO:0051093 negative regulation of developmental process	UBASH3B, EPHB2, VIM, RARB, ROCK2
GO:0044802 single-organism membrane organization	RABGEF1, SEC62, STX4, ITGB2, ROCK2
GO:0007417 central nervous system development	LMO4, EPHB2, VIM, FYN, RARB
GO:0042325 regulation of phosphorylation	UBASH3B, DUSP3, LMO4, RABGEF1, FYN, ITGB2, ROCK2
GO:0080090 regulation of primary metabolic process	UBASH3B, SRSF7, DUSP3, RABGEF1, ELOA, MMS19, RNF40, RPRD1A, ITGB2, LMO4, ZNF713, ZNF629, FYN, RARB, NCOA7, ROCK2
GO:0071704 organic substance metabolic process	DUSP3, RABGEF1, FBXL19, ACACA, LMO4, ZNF629, VIM, FYN, FUCA1, RARB, SRSF7, UBASH3B, ELOA, MMS19, GALM, SLC66A3, EOGT, RNF40, PHKG2, RPRD1A, ITGB2, STX1B, ZNF713, SLC22A11, LYPLA2, NCOA7, ROCK2
GO:0050686 negative regulation of mRNA processing	SRSF7, RNF40
GO:0045595 regulation of cell differentiation	UBASH3B, LMO4, EPHB2, VIM, FYN, RARB, ROCK2
GO:0016477 cell migration	RABGEF1, CNR2, STX4, FYN, ITGB2, ROCK2
GO:0070887 cellular response to chemical stimulus	CREB3L2, RABGEF1, CNR2, ACACA, STX4, FBLN5, FYN, ITGB2, ROCK2
GO:0002064 epithelial cell development	VIM, RARB, ROCK2
GO:1903312 negative regulation of mRNA metabolic process	SRSF7, RNF40
GO:0061024 membrane organization	RABGEF1, SEC62, STX4, ITGB2, ROCK2
GO:0002886 regulation of myeloid leukocyte mediated immunity	RABGEF1, STX4

GO:0043300 regulation of leukocyte degranulation	RABGEF1, STX4
GO:1901655 cellular response to ketone	ACACA, ROCK2
GO:0051641 cellular localization	CREB3L2, UBASH3B, RABGEF1, SEC62, STX4, FBLN5, FYN, ITGB2, ROCK2
GO:0034394 protein localization to cell surface	STX4, FBLN5
GO:0034613 cellular protein localization	RABGEF1, SEC62, STX4, FBLN5, FYN, ITGB2, ROCK2
GO:0070727 cellular macromolecule localization	RABGEF1, SEC62, STX4, FBLN5, FYN, ITGB2, ROCK2
GO:0001775 cell activation	UBASH3B, DUSP3, RABGEF1, STX4, FYN
GO:0010605 negative regulation of macromolecule metabolic process	SRSF7, UBASH3B, DUSP3, RABGEF1, RNF40, FYN, RARB, ROCK2
GO:0010563 negative regulation of phosphorus metabolic process	UBASH3B, DUSP3, RABGEF1, ROCK2
GO:0045936 negative regulation of phosphate metabolic process	UBASH3B, DUSP3, RABGEF1, ROCK2
GO:0032269 negative regulation of cellular protein metabolic process	UBASH3B, DUSP3, RABGEF1, FYN, ROCK2
GO:0050776 regulation of immune response	DUSP3, RABGEF1, STX4, FYN

4 CONSIDERAÇOES FINAIS

O estudo do exoma de suínos normais e afetados pela hérnia umbilical (HU) foi realizado visando identificar variantes (SNPs e InDels) e possíveis genes candidatos à formação e ao desenvolvimento da HU, proporcionando avanço no conhecimento científico relacionado ao surgimento da HU. Esses resultados contribuem para a compreensão dos mecanismos genéticos que envolvem a manifestação da HU em suínos e, possivelmente, em outras espécies de mamíferos, incluindo humanos.

Do total de polimorfismos identificados, 204 através do exoma e 81 através do transcriptoma, 79 foram iguais entre o exoma e o transcriptoma. Os 114 genes identificados (72 provenientes do exoma e 42 do transcriptoma) foram submetidos a uma base de dados para enriquecimento e análise dos processos biológicos (PB). Destacam-se os processos biológicos de contração muscular, resposta do sistema immune, ativação de células *Natural Killer* e adesão da matrix celular, pois apresentam afinidade com os fatores fisiológicos envolvidos na manifestação de hérnias. Além disso, os 79 polimorfismos em comum entre o exoma e o transcriptoma estão localizados em 23 genes.

Os resultados deste trabalho demonstram um grande número de genes e polimorfismos que podem causar e que estão envolvidos com o processo de herniação em suínos da raça Landrace. Com isso, este estudo confirma que a manifestação da HU ocorre devido a atuação de vários genes e polimorfismos. No entanto, são necessários mais estudos com o intuito de validar esses genes em outras populações de suínos e verificar se os polimorfismos identificados estão segregando nas próximas gerações, além de avaliar animais de outras raças e idades. Também se faz necessário investigar a atuação desses polimorfismos quanto a alteração nos aminoácidos e proteínas relacionadas à HU, para aprimorar a interpretação dos mecanismos de regulação gênica destes genes. O conhecimento dos fatores genéticos envolvidos na manifestação da hérnia umbilical é de grande importância para a cadeia produtiva de suínos, bem como para melhorar o bem-estar desses animais na produção, pois estes estudos favorecem o desenvolvimento de estratégias efetivas para reduzir o surgimento destes defeitos nos rebanhos.

5 REFERÊNCIAS

ABCS (Brasil). Associação Brasileira De Criadores De Suínos. **Evolução genética.** Disponível em: < http://abcs.gasoline-digital.com/wp-content/uploads/2020/06/Dados-Mercado-de-Su%C3%ADnos_2019-1.pdf>. Acesso em: 05 abr. 2020.

ABPA (Brasil). Associação Brasileira de Proteína Animal. **Histórico da suinocultura no Brasil.** 2020. Disponível em: <http://abpa-br.org/wp-content/uploads/2020/05/abpa_relatorio_anual_2020_portugues_web.pdf>. Acesso em: 10 ago 2020.

ALBERTS, B. J. *et al*. **Biologia molecular da célula**. 6°. ed. Porto Alegre: Artmed, 2017.

ADALSTEINSSON, V. A. *et al.* Scalable whole-exome sequencing of cell-free DNA reveals high concordance with metastatic tumors. **Nature Communications**, v. 8, n. 1, p. 1–13, 2017.

BADEMCI, G. *et al.* Identification of copy number variants through whole-exome sequencing in autosomal recessive nonsyndromic hearing loss. **Genetic Testing and Molecular Biomarkers**, v. 18, n. 9, p. 658–661, 2014.

BARBAN, N. *et al.* Genome-wide analysis identifies 12 loci influencing human reproductive behavior. **Nature Genetics**, v. 48, n. 12, p. 1–7, 2016.

BELLÓN, J. M. *et al.* Fibroblasts From the Transversalis Fascia of Young Patients With Direct Inguinal Hernias Show Constitutive MMP-2 Overexpression. **Annals of Surgery**, v. 233, n. 2, p. 287–291, 2001.

BENDAVID, R. **The Unified Theory of hernia formation.** Springer Paris, 2004.

BEUERMANN, C. *et al.* Tissue Calcium Content in Piglets with Inguinal or Scrotal Hernias or Cryptorchidism. **Journal of Comparative Pathology**, v. 140, n. 2–3, p. 182–186, 2009.

BRANDT, M, L. **Pediatric Hernias**. Surg Clin North Am, 2008.

BURKHARDT, E. *et al.* Structural organization of the porcine and human genes coding for a leydig cell-specific insulin-like peptide (ley i-l) and chromosomal localization of the human gene (INSl3). **Genomics**, v. 20, n. 1, p. 13–19, 1994.

CHOI, M. *et al.* Genetic diagnosis by whole exome capture and massively parallel DNA sequencing. **Proceedings of the National Academy of Sciences of the United States of America**, v. 106, n. 45, p. 19096–19101, 2009.

COATS, R. D. HELIKSON, M. A. & BURD, R. S. Presentation and management of epigastric hernias in children. **Journal of Pediatric Surgery**, v. 35, n. 12, p. 1754–1756, 2000.

CORREIA, C. N. *et al.* RNA Sequencing (RNA-Seq) Reveals Extremely Low Levels of Reticulocyte-Derived Globin Gene Transcripts in Peripheral Blood From Horses (Equus caballus) and Cattle (Bos taurus). **Frontiers in Genetics**, v. 9, n. AUG, p. 278, 2018.

DING, N. *et al.* A Genome-Wide Scan Reveals Candidate Susceptibility Loci for Pig Hernias in an Intercross Between White Duroc and Erhualian. **Journal of Animal Science.** V. 87 p. 2469-2474, 2009.

DE LIGT, J. *et al.* Diagnostic Exome Sequencing in Persons with Severe Intellectual Disability. **New England Journal of Medicine**, v. 367, n. 20, p. 1921–1929, 2012.

EASTON, D. F. *et al.* Genome-wide association study identifies novel breast cancer susceptibility loci. **Nature**, v. 447, n. 7148, p. 1087–1093, 2007.

FALL, I. *et al.* Strangulated umbilical hernias in children. **Pediatric Surgery International**, v. 22, n. 3, p. 233–235, 2006.

FERNANDES, L. T *et al.* Novel putative candidate genes associated with umbilical hernia in pigs. **In. Proceedings Of The World Congress On Genetics Applied To Livestock Production.** New Zeland, p. 743-749. 2018.

FRANZ, M. G. *et al.* Fascial incisions heal faster than skin: A new model of abdominal wall repair. **Surgery**, v. 129, n. 2, p. 203–208, 2001.

FRANZ, M. G. The Biology of Hernia Formation. **Surg Clin North Am**, v. 88, n. 1, p. 1–16, 2008. Disponível em: https://doi.org/10.1016/j.suc.2007.10.007

FU, X. *et al.* Estimating accuracy of RNA-Seq and microarrays with proteomics. **BMC Genomics**, v. 10, 2009.

GATPHAYAK, K. *et al.* Identification of porcine hernia inguinalis/scrotalis using single nucleotide polymorphism in INSL3 and BAX genes. **Tropical Agriculture**, v. 10 p. 1-5, 2007.

GHOUSSAINI, M. *et al.* Genome-wide association analysis identifies three new breast cancer susceptibility loci. **Nature Genetics**, v. 44, n. 3, p. 312–318, 2012.

GRINDFLEK, E. *et al.* Genome-wide linkage analysis of inguinal hernia in pigs using affected sib pairs. **BMC Genetics**, v. 7, n. 1, p. 25, 2006.

GRINDFLEK, E. *et al.* Genome-wide association study reveals a QTL and strong candidate genes for umbilical hernia in pigs on SSC14. **BMC Genomics**, v. 19, n. 1, p. 412, 2018.

GUIATTI, D. *et al.* Exome capture with heterologous enrichment in pig (Sus scrofa). **PLoS ONE**, v. 10, n. 10, 2015.

HAN, M. R. *et al.* Whole-exome sequencing identifies variants associated with structural MRI markers in patients with bipolar disorders. **Journal of Affective Disorders**, v. 249, p.

159–168, 2019.

HUTSON, J. M. *et al.* **Regulation of testicular descent**. Springer Verlag, 2015.

IAFRATE, A. J. *et al.* Detection of large-scale variation in the human genome. **Nature Genetics**, v. 36, n. 9, p. 949–951, 2004.

JAY G. *et al.* **Pediatric Surgery** . 6. ed. elsevier, 2006. v. 2*E-book*. Disponível em: https://www.elsevier.com/books/pediatric-surgery/9780323028424. Acesso em: 24 maio. 2020.

JI, J. *et al.* An association study using imputed whole-genome sequence data identifies novel significant loci for growth-related traits in a Duroc × Erhualian F 2 population. **Journal of Animal Breeding and Genetics**, v. 136, n. 3, p. 217–228, 2019.

KATSUMI, A. *et al.* Integrin activation and matrix binding mediate cellular responses to mechanical stretch. **Journal of Biological Chemistry**, v. 280, n. 17, p. 16546–16549, 2005.

KEENLISIDE, J. Belly and scrotal ruptures (umbilical and inguinal hernias). **8th Ann Swine Technol Workshop**, Red Deer, Alberta, Canada. 2006.

KURREEMAN, F. A. S. *et al.* Use of a multiethnic approach to identify rheumatoid-arthritis- susceptibility loci, 1p36 and 17q12. **American Journal of Human Genetics**, v. 90, n. 3, p. 524–532, 2012.

LANDER, E. S. Initial impact of the sequencing of the human genome. **Nature**, v.470, n.7333, p.187-97, Feb 10, 2011.

LEVY, S. *et al.* The Diploid Genome Sequence of an Individual Human. **PLoS Biology**, v. 5, n. 10, p. e254, 2007.

LEE, K. *et al.* Genome wide association studies (GWAS) and copy number variation (CNV) studies of the major psychoses: What have we learnt? **Neuroscience & Biobehavioral Reviews.** V. 36, n. 1, p. 556-571, 2012.

LIAO, X *et al.* susceptibility loci for umbilical hernia in swine detected by genome-wide association. **Genetics.** v. 51. p. 1163-1170, 2015.

LI, X. *et al*. Genome-wide association study identifies variants in the *CAPN9* gene associated with umbilical hernia in pigs. **Animal Genetics**, v. 50, n. 2, p.162-165, 2019.

LI, S. *et al.* GWAS identifies novel susceptibility loci on 6p21.32 and 21q21.3 for hepatocellular carcinoma in chronic hepatitis B virus carriers. **PLoS Genetics**, v. 8, n. 7, 2012.

LLINARES-LÓPEZ, F. *et al.* Genome-wide detection of intervals of genetic heterogeneity associated with complex traits. **Bioinformatics**, v. 15, n. 31, p. i240–i249, 2015.

LONG, Y. *et al.* A genome-wide association study of copy number variations with umbilical hernia in swine. **Animal Genetics**, v. 47, n. 3, p. 298–305, 2016.

LU, T. *et al.* Cinnamon extract improves fasting blood glucose and glycosylated hemoglobin level in Chinese patients with type 2 diabetes. **Nutrition Research**, v. 32, n. 6, p. 408–412, 2012.

MAJEWSKI, J. *et al.* What can exome sequencing do for you?. **Journal Medicine Genetics,** 2011.

MANALAYSAY, J. G. *et al.* Screening of BCL-2 associated X protein gene polymorphism associated with scrotal hernia in domesticated swine using polymerase chain reaction-restriction fragment length polymorphism. **Asian-Australasian Journal of Animal Sciences**, v. 30, n. 2, p. 262–266, 2017.

MARIONI, J. C. *et al.* RNA-seq: An assessment of technical reproducibility and comparison with gene expression arrays. **Genome Research**, v. 18, n. 9, p. 1509–1517, 2008.

MCCARTHY, M. I. *et al.* Genome-wide association studies for complex traits: Consensus, uncertainty and challenges. **Nature Review Genetics,** 2008.

MEIER, D. E. *et al.* Incidence of umbilical hernia in african children: Redefinition of "normal" and reevaluation of indications for repair. **World Journal of Surgery**, v. 25, n. 5, p. 645–648, 2001.

MEREI, J. M. Umbilical hernia repair in children: Is pressure dressing necessary. **Pediatric Surgery International**, v. 22, n. 5, p. 446–448, 2006.

METZKER, M. L. Sequencing technologies the next generation. **Nature Publishing Group,** 2010.

MOREIRA, G. *et al*. Genome-wide association scan for QTL and their positional candidate genes associated with internal organ traits in chickens. **BMC Genomics**, v. 20, p. 669, 2019.

OZAKI, K. *et al.* Functional SNPs in the lymphotoxin-α gene that are associated with susceptibility to myocardial infarction. **Nature Genetics**, v. 32, n. 4, p. 650–654, 2002.

PEIXOTO, J. O. *et al.* Proximal femoral head transcriptome reveals novel candidate genes related to epiphysiolysis in broiler chickens. **BMC Genomics**, v. 20, n. 1, p. 1031, 2019.

PETERSEN, H. H. *et al.* Prevalence of clinical signs of disease in Danish finisher pigs. **Veterinary Record**, v. 162, n. 12, p. 377–382, 2008.

PERTILLE, F. *et al.* Genome-wide association study for performance traits in chickens using genotype by sequencing approach. **Scientific Reports,** v. 7, p. 41748, 2017.

PLENGE, R. M. *et al.* Two independent alleles at 6q23 associated with risk of rheumatoid arthritis. **Nature Genetics**, v. 39, n. 12, p. 1477–1482, 2007.

POLTRONIERI, I. A, *et al.* Avaliação da ocorrência de hérnias escrotais, inguinais e umbilicais em propriedade do alto Uruguai gaúcho. ***In:*** **semana do conhecimento UPF-Universidade de Passo Fundo.** Disponível em: <http://semanadoconhecimento.upf.br/download/anais-2015/ciencias-agrarias/iuri-astolfi-poltronieri-avaliacao-da-ocorrencia-de-hernias.pdf> Acesso em: 13/09/2020.

REDON, R. *et al.* Global variation in copy number in the human genome. **Nature**, v. 444, n. 7118, p. 444–454, 2006.

REY, R. *et al.* AMH/MIS: What we know already about the gene, the protein and its regulation. *In*: 2003, **Molecular and Cellular Endocrinology**. : Elsevier Ireland Ltd, 2003. p. 21–31.

RICHARD, F. J. Regulation of meiotic maturation1. **Journal of Animal Science**, v. 85, n. suppl_13, p. E4–E6, 2007.

ROBERT, C. *et al.* Design and development of exome capture sequencing for the domestic pig (Sus scrofa). **BMC Genomics**, v. 15, n. 1, p. 550, 2014.

ROMANO, G. S. *et al.* Inguinal ring RNA sequencing reveals downregulation of muscular genes related to scrotal hernia in pigs. **Genes**, v. 11, n. 2, 2020.

ROSELLI, C. *et al.* Multi-ethnic genome-wide association study for atrial fibrillation. **Nature Genetics**, v. 50, n. 9, p. 1225–1233, 2018.

SANDHOLM, N. *et al.* New Susceptibility Loci Associated with Kidney Disease in Type 1 Diabetes. **PLOS Genetics**, v. 29, n. 9, p. 1002921, 2012.

SANNA, S. *et al.* Fine Mapping of Five Loci Associated with Low-Density Lipoprotein Cholesterol Detects Variants That Double the Explained Heritability. **PLoS Genetics**, v. 7, n. 7, p. e1002198, 2011.

SANTANA, M. H. A. *et al.* A genomewide association mapping study using ultrasound-scanned information identifies potential genomic regions and candidate genes affecting carcass traits in Nellore cattle. **Journal of Animal Breeding and Genetics**, v. 132, n. 6, p. 420–427, 2015.

SAWYER, S. L. *et al.* Utility of whole-exome sequencing for those near the end of the diagnostic odyssey: time to address gaps in care. **Clinical Genetics**, v. 89, n. 3, p. 275–284, 2016.

SCHERER, S. W. *et al.* Challenges and standards in integrating surveys of structural variation. **Nature Genetics**, v. 39, n. 7S, p. S7–S15, 2007.

SEARCY-BERNAL R., GARDNER I. A., & HIRD D.W. Effects of and factors associated with umbilical hérnias in a swine herd. **Journal of the American Veterinary Medinal Association**. vol. 4, no.10, p. 1660–1664. 1994.

SEBAT, J. *et al.* Large-scale copy number polymorphism in the human genome. **Science**, v. 305, n. 5683, p. 525–528, 2004.

SEVILLANO, C. A. *et al.* Genome-wide association study using deregressed breeding values for cryptorchidism and scrotal/inguinal hernia in two pig lines. **Genetics Selection Evolution**, v. 47, n. 1, p. 18, 2015.

SOUZA, M. R. *et al.* Transcriptome analysis identifies genes involved with the development of umbilical hernias in pigs. **PLoS ONE**, v. 15, n. 5, 2020.

STARK, R., GRZELAK, M., & HADFIELD, J. RNA sequencing: the teenage years. **Nature Publishing Group**, 2019.

STRAW, B. BATES, R. & MAY, G. Anatomical abnormalities in a group of finishing pigs: Prevalence and pig performance. **Journal of Swine Health and Production**, v. 17, n. 1, p. 28–31, 2009.

SUDHAGAR, A. KUMAR, G. & EL-MATBOULI, M. Transcriptome Analysis Based on RNA-Seq in Understanding Pathogenic Mechanisms of Diseases and the Immune System of Fish: A Comprehensive Review. **International Journal of Molecular Sciences**, v. 19, n. 1, p. 245, 2018.

THALLER, G. DEMPFLE, L. & HOESCHELE, I. Investigation of the inheritance of birth defects in swine by complex segregation analysis. **Journal of Animal Breeding and Genetics**, v. 113, n. 1–6, p. 77–92, 1996.

TRAPNELL, C. *et al.* Differential analysis of gene regulation at transcript resolution with RNA-seq. **Nature Biotechnology**, v. 31, n. 1, p. 46–53, 2013.

USDA (Estados Unidos). United States Department of Agriculture. Foreign Agricultural Service: **Pork Production - Selected Countries Summary**. 2018. Disponível em: <https://apps.fas.usda.gov/psdonline/app/index.html#/app/downloads>. Acesso em: 08 abr 2020.

WANG, L. *et al.* Copy number variation-based genome wide association study reveals additional variants contributing to meat quality in Swine. **Scientific Reports** 5 12535, 2015.

WANG, Y. *et al.* Evaluation of the coverage and depth of transcriptome by RNA-Seq in chickens. **BMC Bioinformatics**, v. 12, n. SUPPL. 10, p. S5, 2011..

WANG, Z. GERSTEIN, M. & SNYDER, M. RNA-Seq: A revolutionary tool for transcriptomics. **Nature Review Genetics,** 2009.

WARR, A. *et al.* Exome sequencing: Current and future perspectives. **G3: Genes, Genomes, Genetics**, v. 5, n. 8, p. 1543–1550, 2015.

WOLFE, M. D. *et al.* Genome-Wide Association and Prediction Reveals Genetic Architecture of Cassava Mosaic Disease Resistance and Prospects for Rapid Genetic Improvement. **The Plant Genome**, v. 9, n. 2, p.0110.0118, 2016.

WOODWARK, C. & BATEMAN, A. The characterisation of three types of genes that overlie copy number variable regions. **PLoS ONE**, v. 6, n. 5, 2011.

WOOLLEY, M. J. *et al.* Understanding the molecular functions of the second extracellular loop (ECL2) of the calcitonin gene-related peptide (CGRP) receptor using a comprehensive mutagenesis approach. **Molecular and Cellular Endocrinology**, v. 454, p. 39–49, 2017.

XIANG, T. *et al.* Imputation of genotypes in Danish purebred and two-way crossbred pigs using low-density panels. **Genetics Selection Evolution**, v. 47, n. 1, p. 54, 2015.

XUE Q. *et al.* A genome-wide association study in Han Chinese identifies multiple susceptibility loci for IgA nephropathy. **Nature Genetics**, v. 44, n. 2, p. 178–182, 2012.

ZAHA, A. FERREIRA, H.B. & PASSAGLIA, L.M.P. **Biologia molecular básica.** 5°. ed. Porto Alegre: Artmed, 2014.

ZANCHIN, F. **Diagnósticos diferenciais de hérnias umbilicais em suínos no abate.** Dissertação (mestrado em ciências veterinárias) – Faculdade de Veterinária, Universidade Federal do Rio Grande do Sul. Rio Grande do Sul. p. 55. 2015

ZEGGINI, E. *et al.* Meta-analysis of genome-wide association data and large-scale replication identifies additional susceptibility loci for type 2 diabetes. **Nature Genetics**, v. 40, n. 5, p. 638–645, 2008.

ZHAO, S. *et al.* Mutational landscape of uterine and ovarian carcinosarcomas implicates histone genes in epithelial-mesenchymal transition. **In: Proceedings of the National Academy of Sciences of the United States of America**, v. 113, n. 43, p. 12238–12243, 2016.

ZHENG, H. *et al.* Recurrent inguinal hernia: Disease of the collagen matrix? **World Journal of Surgery**, v. 26, n. 4, p. 401–408, 2002.

6 ANEXOS

Embrapa	Certificado de Conduta Ética	ETICA 1/1

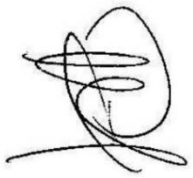
CERTIFICADO

Certificamos que o Protocolo nº**(000/AAAA): 011/2014**, sob título **"Identificação de genes e polimorfismos associados à formação de hérnias em suínos pela combinação do sequenciamento exômico total e do RNA",** sob responsabilidade de **Mônica Ledur** está de acordo com os Princípios Éticos na Experimentação Animal, adotados pelo Colégio Brasileiro de Experimentação Animal (COBEA), **TENDO SIDO CONSIDERADO APROVADO PELA** Comissão de Ética no Uso de Animais (CEUA/CNPSA) em reunião realizada em **07/ 11/ 2014**.

CERTIFICATE

We certify that the Protocol nº **(000/YYYY): 011/2014**, under the following title **"Identification of genes and polymorphisms associated with formation of hernias in swines combining RNA and whole exons sequencing."** is in agreement with the Ethical Principles in Animal Research adopted by Brazilian College of Animal Experimentation (COBEA) and was approved by the Embrapa Swines and Poultry Ethical Committee for Animals utilization in experimentation (CEUA/CNPSA) in **11/07/2014.**

Concórdia, **07/11/2014**.



Presidente CEUA/CNPSA

FQ4-053-05
Rev(03)