A premature stop mutation in the porcine myostatin gene is a candidate causative variant for a recessive leg weakness syndrome and affects muscle depth

A premature stop mutation in the porcine myostatin gene is a candidate causative variant for a recessive leg weakness syndrome and affects muscle depth

O. Matika[1], D. Robledo[1], R. Pong-Wong[1], J. A. Woolliams[1], S. C. Bishop[1], V. Riggio[1], A. E. Hoste[2], G. A. Walling[2], A. L. Archibald[1] & R. D. Houston[1]

[1] The Roslin Institute and Royal (Dick) School of Veterinary Studies, The University of Edinburgh, Midlothian, EH25 9RG, United Kingdom
Oswald.matika@roslin.ed.ac.uk (Corresponding Author)
[2] JSR Genetics, Southburn, Driffield, East Yorkshire, YO25 9ED, United Kingdom

Summary

Lameness in piglets is a major animal welfare and economic problem in pig production. Following observation of a high incidence of a leg weakness syndrome in a commercial pig lines, a variance components approach was used to assess the genetic basis of the condition. The results suggested a monogenic recessive mode of inheritance, and homozygosity mapping was used to identify a region associated with the leg weakness syndrome on SSC15. Whole genome resequencing of cases and controls identified an outstanding candidate mutation in this region which results in a premature stop codon within exon 3 of the porcine *MSTN* locus. Myostatin (MSTN) is a transforming growth factor-β family member that is a critical regulator of skeletal muscle development. Mutations in the *MSTN* gene lead to muscle hypertrophy and are responsible for the 'double muscling' phenotype observed in several mammalian species, and is a common target for gene editing experiments in farm animals. The candidate causal mutation in MSTN was in Hardy-Weinberg equilibrium at birth, but significantly distorted amongst animals still in the herd at 110 kg weight. In heterozygous form, the *MSTN* mutation was associated with a major increase in muscle depth and decrease in fat depth, explaining 31 and 18 % of the genetic variation respectively. *MSTN* ablation by gene editing in pigs is associated with problems of low piglet survival and lameness. Thus, in the current population, it is likely that this *MSTN* mutation was deleterious for piglet survival, but was maintained due to selection for increased muscle associated with heterozygous animals. The association between this *MSTN* mutation and fitness traits (leg weakness, survival) has clear implications for the potential use of gene editing of the porcine *MSTN* locus for increased meat production, and provide a plausible explanation for the lack of disrupting *MSTN* mutations in pigs despite intense selection for lean growth and their relatively high frequency in other species.

Keywords: pigs, leg weakness, MSTN, growth traits

Introduction

Leg weakness is a major cause of lameness in piglets which has major negative animal welfare and economic impact in pig production. While leg weakness is a heterogeneous condition, notably high heritability estimates in certain pig breeds (up to 0.61) suggest a strong underlying genetic basis in some cases (Jørgensen & Vestergaard, 1990). Further,

genetic correlations between leg weakness traits and other production traits (such as growth and muscle depth) have previously been found (Jørgensen & Vestergaard, 1990). In the current study, the genetic basis of a leg weakness syndrome was characterised in a commercial line of Large White pigs with a high incidence of this syndrome. Variance component analyses were used to assess genetic parameters and the mode of inheritance for the syndrome. Homozygosity mapping was applied to identify possible regions associated with the phenotype, after which whole genome resequencing of affected and control animals was used to identify a candidate causative mutation for further testing.

Material and methods

Phenotype description

The data used were from a Large White commercial line of piglets from a genetic nucleus unit reared under standard commercial conditions. The data comprised 19,006 piglets phenotyped since the leg weakness syndrome was first noticed in 2007. Since the condition seemed to affect muscle and tendons, resulting in the piglet not being able to straighten the legs to stand, the animals were then visually classified as normal or affected (0/1). A subset of affected piglets were video recorded to characterise the features of the syndrome. The full pedigree comprised 26,908 animals over seven generations with 262 sires mated to 1,583 dams. These data were used to characterise the mode of inheritance and genetic parameter estimates. All samples were collected on a commercial nucleus farm as part of normal husbandry and management procedures in the nucleus flock and complied with conventional UK red tractor farm assurance standards (https://assurance.redtractor.org.uk/) where sick or injured livestock that do not respond to treatment are promptly and humanely euthanized by a trained and competent stockperson.

Genetic parameter estimates

Initially the data were analysed with the leg weakness treated either as a continuous or as a binary trait fitting a logit link function using ASREML software (Gilmour *et al*., 2009). Models explored random effects due to the animal, sire, dam and their combinations. Other non-genetic random effects fitted were the permanent environmental effects due to the sow and litter effects. Environmental effects fitted included month and year of birth, and sow parity as fixed effects, with numbers born alive or dead included as covariates. The data were also explored using complex segregation analyses (Walling *et al*., 2002), implemented using a Gibbs sampler to formally investigate the major gene hypothesis.

Homozygosity mapping

Our hypothesis was that the leg weakness syndrome is caused by the effect of a single mutation with the deleterious allele exhibiting a recessive mode of inheritance. If this is true, then all affected individuals would be homozygous for the deleterious allele and the homozygosity status would extend to the area closely linked to the mutation. Hence, homozygosity mapping identifies a candidate region where all affected individuals are homozygous for the same haplotype (and the controls are not) (Charlier *et al.*, 2008). To test this hypothesis DNA samples were collected from 10 cases (putative homozygotes for the

causative allele) and 10 unaffected littermates (heterozygous or homozygous wild-type). Animals were genotyped using the Illumina PorcineSNP60 SNP chip (Ramos *et al*., 2009). After quality control, candidate regions were found by identifying regions with all affected individuals sharing the same runs of homozygosity status.

Whole genome resequencing

To discover candidate causative mutations, six 'cases' from the original piglets were pooled and used to define the homozygous segment and six separate putative heterozygous dams were sequenced on an Illumina HiSeq2500 platform by Edinburgh Genomics. The dams were sequenced at an average depth of ~10x, and the pool of piglets to a depth of ~20x. The DNA-sequencing output comprised ~1.3 billion paired-end reads. The piglets had an average of 48 million paired-end reads per sample, while the dams had on average 157 million paired-end reads per sample. Quality filtering and removal of residual adaptor sequences was conducted on read pairs using Trimmomatic v.0.32 (Bolger *et al.* 2014). Only reads where both pairs had a length greater than 32 bp post-filtering were retained, leaving a total of ~1.2 bn paired-end reads (92%). Whole genome resequencing was followed by alignment to the published pig genome assembly (Sscrofa10.2, GCF_000003025.5; (Groenen *et al.* 2012)) using Burrows-Wheeler Aligner with default parameters (Li & Durbin 2010). Variant calling was performed using Genome Analysis Toolkit (GATK) HaplotypeCaller after read recalibration (DePristo *et al.* 2011).

Validation of SNP

A kompetitive allele specific PCR (KASP) was designed and performed by LGC Genomics (UK) to enable genotyping of the *MSTN* mutation in large numbers of animals. A population of 686 piglets was genotyped at birth, of which 381 pigs had performance test data recorded at 110 kg (i.e. slaughter weight) and were used to validate the candidate causative SNP variant. Their pedigree comprised 1,107 animals with 178 sires mated to 439 dams.

The genotype patterns were assessed by testing Hardy-Weinberg equilibrium (HWE) at the two sampling time points (i.e. birth and 110kg), under the hypothesis that a causative mutation (or closely linked marker) should be in HWE at birth, but then significantly different from HWE due to an absence or near-absence of homozygous recessive genotypes in the performance test samples. An association analysis was conducted to investigate the effect of the mutation and the performance traits including days to 40 kg, days from 40-110 kg, weight at start of test period (average 85 days), weight at end of test period (average 138 days), muscle depth and fat depth. The fixed effects accounted for in the linear mixed model included year, sex, parity, age and SNP, with animal fitted as a random effect.

Results and discussion

A high estimate of heritability (0.70 ± 0.16) was obtained fitting a logit transformed mixed linear sire model. In this model, low estimates of 0.17 ± 0.02 and 0.11 ± 0.02 were observed for permanent environmental effects due to the dam and litter respectively. There was no direct comparison for this particular leg weakness syndrome reported in literature (http://omia.angis.org.au/OMIA000585/9825/). The overall prevalence of leg weakness was 6.3%. The mean proportion of affected piglets, summing across all piglets in affected litters,

was 23% ± 0.7. The results from the complex Bayesian segregation analysis allowing for dominance, suggested almost all the variation could be explained by a single gene with virtually no polygenic or environmental variation. Results from the Bayesian segregation analysis gave a mean estimate of the additive effect of 0.50 ±0.001 and mean dominance effect of -0.50 ± 0.001, which is in agreement with the recessive gene model hypothesis.

Homozygosity mapping based on the Illumina PorcineSNP60 SNP genotypes of cases and controls revealed a region with the longest homozygous segment in all cases that was not homozygous for the same alleles in controls of ~8.5Mb containing 62 SNPs from the SNP array on pig chromosome (SSC)15. This was assumed to represent a selective sweep that is likely to contain the underlying causative mutation.

Whole genome resequencing followed by alignment to the published pig genome assembly resulted in the identification of 40 SNPs and 10 InDels within the homozygosity region, which co-segregated with the phenotype under the assumption of a single causative mutation. Functional annotation of these variants identified an outstanding candidate functional mutation in the third exon of the *MSTN* locus that resulted in t a stop codon at position 274 (p.Glu274*), which would results in a MSTN protein missing the last 101 amino acids, which is a highly conserved region across many animal species. This mutation had not previously been reported in pigs. No other polymorphisms with discernible function were identified.

The results of all 686 piglets sampled at birth in our population for the *MSTN* mutation showed that this mutation was in HWE at birth (p=0.79, q=0.21, χ^2=0.01, P > 0.05) but significantly deviated from HWE in the subset of animals present at slaughter weight (p=0.80, q=0.20, χ^2=23.55, P<0.05), with most of the homozygous mutant animals dying within a week and none present at slaughter.

Few pleiotropic major genes have been identified that control both production and fitness traits. Mutations in the myostatin gene (*MSTN*) have been reported to cause muscle hypertrophy and are responsible for the exceptional ‘double muscling’ phenotype observed in several mammalian species, most famously in Belgian Blue cattle (McPherron & Lee, 1997). MSTN is a member of the transforming growth factor beta (TGF-β) superfamily, is highly conserved across species, and is typically expressed in developing and mature skeletal muscle as a key regulator of muscle growth (McPherron & Lee, 1997). The structure of the gene comprises three exons and two introns in all livestock species studied, and mutations in exon 3 cause the double-muscled phenotype in cattle (e.g. McPherron & Lee, 1997). The causal relationship between the *MSTN* mutations and double-muscle phenotypes in cattle was initially inferred from the phenotypes of mice in which the *MSTN* gene had been knocked-out. More recent experiments in cattle and sheep in which the *MSTN* gene was disrupted by gene editing technology have supported this inference of causality (Proudfoot *et al.,* 2015; Wang *et al*., 2016). Whilst the homozygous *MSTN* double-muscle phenotype in cattle presents some challenges including a frequent requirement for caesarean deliveries, the *MSTN* mutation homozygotes are viable.

Gene editing has been used to disrupt the porcine myostatin gene, and an increase in muscle mass has been reported in edited pigs (Kang *et al*., 2014; Wang *et al*., 2015; Bi *et al*., 2016; Rao *et al*., 2016; Tanihara *et al*., 2016; Kang *et al*., 2017),. However, a number of problems, including leg weakness, have been reported for the pigs carrying these artificial *MSTN* mutations (Kang *et al*., 2014; Rao *et al*., 2016; Xing *et al*., 2017).

The impact of the natural porcine *MSTN* mutation observed in the current study (in heterozygous form) on performance traits was assessed in 381 pigs using an association

analysis. A significant association of the *MSTN* genotype with muscle depth and fatness traits was observed, with *MSTN* explaining 18 and 31 % of the genetic variation in fat and muscle depth respectively ($P < 0.001$) when assuming additivity. The heterozygous animals had on average 5 mm increased muscle depth, and 1.7 mm decreased backfat depth. The selection index applied in this population significantly benefited animals with positive muscle depth EBVs and negative fat depth EBVs. Therefore, it is plausible that the allele frequency of the *MSTN* mutation has been increased by selective breeding due to its favourable effect on muscle depth and fatness. Conversely, its association with piglet leg weakness and survival in homozygous form may explain the lack of naturally occurring major *MSTN* mutations reported in domesticated pigs to date, despite intense selection for lean growth.

Conclusions

A leg weakness syndrome was found to be highly heritable and controlled by a single major gene with a recessive mode of inheritance. Homozygosity mapping identified a single region of approximately 8.5 Mb on SSC15. The single most plausible candidate causative mutation causes a premature stop codon in exon 3 of the myostatin locus. A survey of commercial pigs shows the mutation in HWE at birth and complete absence of the homozygous mutant genotype at full market weight in these data. Pigs that were heterozygous for the mutation exhibited highly significantly increased muscle depth, and reduced fatness compared to wild-type animals. This implies that the mutation underlies a major QTL for these production traits. The pleiotropic effects of this naturally occurring nonsense myostatin mutation may explain why such mutations have not previously been reported in pigs, and are consistent with data from *MSTN* gene edited animals that suggest developmental problems associated with some of the genetically engineered *MSTN* genotypes in pigs under commercial farm conditions.

Acknowledgements

This work was supported by the Biotechnology and Biological Sciences Research Council Institute Strategic Programme Grants (BBS/E/D/20211553, BBS/E/D/30002275). The presentation at WCGALP in Auckland was partly supported by Alan Robertson Fund for Animal Genetics (scholarship) awarded by the British Society of Animal Science.

List of References

Bi, Y., Z. Hua, X. Liu, W. Hua, H. Ren, H. Xiao, L. Zhang, L. Li, Z. Wang, G. Laible, Y. Wang, F. Dong & X. Zheng, 2016. Isozygous and selectable marker-free MSTN knockout cloned pigs generated by the combined use of CRISPR/Cas9 and Cre/LoxP. Sci Rep UK 6: 31729.

Bolger, A.M., M. Lohse & B. Usadel, 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30: 2114-2120.

Charlier, C., W. Coppieters, F. Rollin, D. Desmecht, J.S. Agerholm, N. Cambisano, E. Carta, S. Dardano, M. Dive, C. Fasquelle, J.C. Frennet, R. Hanset, X. Hubin, C. Jorgensen, L. Karim, M. Kent, K. Harvey, B.R. Pearce, P. Simon, N. Tama, H. Nie, S. Vandeputte, S. Lien, M. Longeri, M. Fredholm, R.J. Harvey & M. Georges, 2008. Highly effective SNP-

based association mapping and management of recessive defects in livestock. Nat. Genet. 40: 449-454.

DePristo, M.A., E. Banks, R. Poplin, K.V. Garimella, J.R. Maguire, C. Hartl, A.A. Philippakis, G. del Angel, M.A. Rivas, M. Hanna, A. McKenna, T.J. Fennell, A.M. Kernytsky, A.Y. Sivachenko, K. Cibulskis, S.B. Gabriel, D. Altshuler & M.J. Daly, 2011. A framework for variation discovery and genotyping using next-generation DNA sequencing data. Nat. Genet. 43: 491-498.

Gilmour, A.R., B.J Gogel B.R. Cullis & R. Thompson 2009. ASReml User Guide Release 3.0, VSN Int. Ltd Hempstead,HP1 1ES, UK.

Groenen, M.A.M., A.L. Archibald, H. Uenishi, C.K. Tuggle, Y. Takeuchi, M.F. Rothschild, C. Rogel-Gaillard, C. Park, D. Milan, H.-J. Megens, S. Li, D.M. Larkin, H. Kim, L.A.F. Frantz, M. Caccamo, H. Ahn, B.L. Aken, A. Anselmo, C. Anthon, L. Auvil, B. Badaoui, C.W. Beattie, C. Bendixen, D. Berman, F. Blecha, J. Blomberg, L. Bolund, M. Bosse, S. Botti, Z. Bujie, M. Bystrom, B. Capitanu, D. Carvalho-Silva, P. Chardon, C. Chen, R. Cheng, S.-H. Choi, W. Chow, R.C. Clark, C. Clee, R.P.M.A. Crooijmans, H.D. Dawson, P. Dehais, F. De Sapio, B. Dibbits, N. Drou, Z.-Q. Du, K. Eversole, J. Fadista, S. Fairley, T. Faraut, G.J. Faulkner, K.E. Fowler, M. Fredholm, E. Fritz, J.G.R. Gilbert, E. Giuffra, J. Gorodkin, D.K. Griffin, J.L. Harrow, A. Hayward, K. Howe, Z.-L. Hu, S.J. Humphray, T. Hunt, H. Hornshøj, J.T. Jeon, P. Jern, M. Jones, J. Jurka, H. Kanamori, R. Kapetanovic, J. Kim, J.-H. Kim, K.-W. Kim, T.-H. Kim, G. Larson, K. Lee, K.-T. Lee, R. Leggett, H.A. Lewin, Y. Li, W. Liu, J.E. Loveland, Y. Lu, J.K. Lunney, J. Ma, O. Madsen, K. Mann, , L. Matthews, S. McLaren, T. Morozumi, M.P. Murtaugh, J. Narayan, D.T. Nguyen, P. Ni, S.-J. Oh, S. Onteru, F. Panitz, E.-W. Park, H.-S. Park, G. Pascal, Y. Paudel, M. Perez-Enciso, R. Ramirez-Gonzalez, J.M. Reecy, S. Rodriguez-Zas, G.A. Rohrer, L. Rund, Y. Sang, K. Schachtschneider, J.G. Schraiber, J. Schwartz, L. Scobie, C. Scott, S. Searle, B. Servin, B.R. Southey, G. Sperber, P. Stadler, J.V. Sweedler, H. Tafer, B. Thomsen, R. Wali, J. Wang, J. Wang, S. White, X. Xu, M. Yerle, G. Zhang, J. Zhang, J. Zhang, S.-H. Zhao, J. Rogers, C. Churcher & L.B. Schook, 2012. Pig genomes provide insight into porcine demography and evolution. Nature 491: 393-398.

Jørgensen, B. & T. Vestergaard, 1990. Genetics of leg weakness in boars at the Danish pig breeding stations. Acta Agr. Scand. 40: 59-69.

Kang, J.-D., S. Kim, H.-Y. Zhu, L. Jin, Q. Guo, X.-C. Li, Y.-C. Zhang, X.-X. Xing, M.-F. Xuan, G.-L. Zhang, Q.-R. Luo, Y.S. Kim, C.-D. Cui, W.-X. Li, Z.-Y. Cui, J.-S. Kim & X.-J. Yin, 2017. Generation of cloned adult muscular pigs with myostatin gene mutation by genetic engineering. RSC Adv. 7: 12541-12549.

Kang, Q., Y. Hu, Y. Zou, W. HU, L. Li, F. Chang, Y. Li, D. Lu, Z. Sun, R. Zhang, Q. Li, Y. Dai & N. Li, 2014. Improving pig genetic resistance and muscle production through molecular biology**.** Proceedings, 10th World Congress of Genetics Applied to Livestock Production, Vancouver, BC, Canada. August 17-22, 2014.

Li, H. & R. Durbin, 2010. Fast and accurate long-read alignment with Burrows-Wheeler transform. Bioinformatics 26: 589-595.

McPherron, A.C. & S.J. Lee, 1997. Double muscling in cattle due to mutations in the myostatin gene. P. Natl Acad Sci USA 94: 12457-12461.

Proudfoot, C., D.F. Carlson, R. Huddart, C.R. Long, J.H. Pryor, T.J. King, S.G. Lillico, A.J. Mileham, D.G. McLaren, C.B.A. Whitelaw & S.C. Fahrenkrug, 2015. Genome edited sheep and cattle. Transgenic Res 24: 147-153.

Ramos, A.M., R.P.M.A. Crooijmans, N.A. Affara, A.J. Amaral, A.L. Archibald, J.E. Beever, C. Bendixen, C. Churcher, R. Clark, P. Dehais, M.S. Hansen, J. Hedegaard, Z.-L. Hu, H.H. Kerstens, A.S. Law, H.J. Megens, D. Milan, D.J. Nonneman, G.A. Rohrer, M.F. Rothschild, T.P.L. Smith, R.D. Schnabel, C.P. Van Tassell, J.F. Taylor, R.T. Wiedmann, L.B. Schook. & M.A.M. Groenen, 2009. Design of a high density SNP genotyping assay in the pig using SNPs identified and characterized by next generation sequencing technology. *PLoS ONE* 4(8): e6524.

Rao, S., T. Fujimura, H. Matsunari, T. Sakuma, K. Nakano, M. Watanabe, Y. Asano, E. Kitagawa, T. Yamamoto & H. Nagashima, 2016. Efficient modification of the myostatin gene in porcine somatic cells and generation of knockout piglets. Mol. Reprod. Dev. 83: 61-70.

Tanihara, F., T. Takemoto, E. Kitagawa, S. Rao, L.T.K. Do, A. Onishi, Y. Yamashita, C. Kosugi, H. Suzuki, S. Sembon, S. Suzuki, M. Nakai, M. Hashimoto, A. Yasue, M. Matsuhisa, S. Noji, T. Fujimura, D. Fuchimoto & T. Otoi, 2016. Somatic cell reprogramming-free generation of genetically modified pigs. Sci. Adv. 2: e1600803.

Walling, G.A., S.C. Bishop, R. Pong-Wong, G. Gittus, A.J.F. Russel & S.M. Rhind, 2002. Detection of a major gene for litter size in Thoka Cheviot sheep using Bayesian segregation analyses. Anim. Sci. 75: 339–347.

Wang, K.K., H.S. Ouyang, Z.C. Xie, C.G. Yao, N.N. Guo, M.J. Li, H.P. Jiao & D.X. Pang, 2015. Efficient generation of myostatin mutations in pigs using the CRISPR/Cas9 system. Sci. Rep. UK 5: 16623.

Wang, X.L., Y.Y. Niu, J.K. Zhou, H.H. Yu, Q.F. Kou, A.M. Lei, X.E. Zhao, H.L. Yan, B. Cai, Q.Y. Shen, S.W. Zhou, H.J. Zhu, G.X. Zhou, W.Z. Niu, J.L. Hua, Y. Jiang, X.X. Huang, B.H. Ma & Y.L. Chen, 2016. Multiplex gene editing via CRISPR/Cas9 exhibits desirable muscle hypertrophy without detectable off-target effects in sheep. Sci. Rep. UK 6: 32271.

Xing, X.-X., M.-F. Xuan, L. Jin, Q. Guo, Z.-B. Luo, J.-X. Wang, Q.R. Luo, G.L. Zhang, C.-D. Cui, Z.-Y. Cui, J.-D. Kang & X.-J. Yin, 2017. Fibre-type distribution and expression of myosin heavy chain isoforms in newborn heterozygous myostatin-knockout pigs. Biotechnol. Lett. https://doi.org/10.1007/s10529-017-2422-2.