A 63-bp insertion in exon 2 of the porcine *KIF21A* gene is associated with arthrogryposis multiplex congenita

Zih-Hua Fang^{1*}, Adéla Nosková¹, Danang Crysnanto¹, Stefan Neuenschwander², Peter Vögel², Hubert Pausch¹

¹ Animal Genomics, Institute of Agricultural Science, D-USYS, ETH Zürich, 8092 Zürich, Switzerland

² Animal Genetics unit, Institute of Agricultural Science, D-USYS, ETH Zürich, 8092 Zürich, Switzerland

* corresponding author: zih-hua.fang@usys.ethz.ch

Running title: Fine-mapping of porcine arthrogryposis multiplex congenita

Summary

Arthrogryposis multiplex congenita (AMC) is a recessively inherited fatal disease detected almost 20 years ago in the Swiss Large White pig population. A diagnostic marker test enabled the identification of carrier animals, but the underlying causal mutation remains unknown. To identify the mutation underlying AMC, we collected whole-genome genotyping and sequencing data for 11 affected piglets and 23 healthy pigs. Haplotype-based case-control association testing using 47,829 SNPs confirmed that AMC maps to SSC5 ($P = 9.4 \times 10^{-13}$). Subsequent autozygosity mapping revealed a common 6.06 Mb region (from 66,757,970 to 72,815,151 bp) of extended homozygosity in 11 piglets affected by AMC. We detected a 63-bp insertion in the second exon of *KIF21A* gene encoding Kinesin Family Member 21A using whole-genome sequences of a carrier boar, two of its affected and two heterozygous piglets. This insertion was compatible with the recessive inheritance of AMC. The 63-bp insertion likely represents a loss-of-function allele because it is predicted to introduce a premature stop codon in *KIF21A* gene (p.Val41_Phe42insTer) that truncates 1,614 amino acids (~ 97%) from the protein. Lack of KIF21A protein is lethal in mice, thus providing additional evidence that a loss-of function allele of *KIF21A* might cause fatal AMC in pigs. We found that this deleterious allele still segregates at low frequency in the Swiss Large White pig

population. The unambiguous detection of carrier animals can now facilitate the eradication of the deleterious allele from the population.

Keywords (3 to 10)

pig, mutation, monogenic, recessive, rare disease, genetic disorder, whole genome sequencing, homozygosity mapping, genome-wide association

Small effective population size and the wide-spread use of a few breeding animals may result in the frequent phenotypic manifestation of recessive alleles in livestock populations. Microarray-based genotyping in affected and unaffected individuals facilitates rapid mapping of recessive genetic defects (Charlier *et al.* 2008). The analysis of whole-genome sequence variant genotypes of either affected or carrier animals has proven to be very efficient in detecting causal variants for recessive traits (<https://omia.org/home/>). Yet uncovering causal variants remains challenging when they reside in genomic regions with either assembly problems or gaps in reference sequences (Pausch *et al.* 2015). Diagnostic haplotype- or microsatellite-based testing may be used to detect carrier animals and avoid at-risk matings even when the molecular basis of genetic defects is unknown. However, the presence of either recombinant or ancestral haplotypes may prevent unambiguous identification of carrier animals when causal mutations are not genotyped (e.g. Jung *et al.* 2014).

Arthrogryposis multiplex congenita (AMC) is a fatal autosomal recessive disease that was first described almost 20 years ago in the Swiss Large White pig population (Genini *et al.* 2004). Affected piglets die either shortly before or during birth and show various congenital abnormalities including malformed and contracted joints of fore- and hindlimbs, brachygnathia inferior and scoliosis (Genini *et al.* 2004, Haubitz *et al.* 2012). Using linkage analysis, a 2.32 Mb region (1.86 Mb based on the Sscrofa 11.1 assembly) of SSC5 has been associated with AMC (Haubitz *et al.* 2012). Microsatellite testing was applied to exclude carrier animals from breeding and remove the deleterious allele from the population. However, a causal mutation had not been detected so far. Here, we attempt to eventually identify the causal mutation for AMC using whole-genome genotyping and sequencing data.

Initially, Genini *et al.*, 2004 mapped AMC to a 4.61 Mb interval on SSC5 between two microsatellite markers SW152 (69,619,907 bp, Sscrofa 11.1) and SW904 (74,230,251 bp, Sscrofa 11.1) using linkage analysis. Using a larger dataset and additional markers enabled refining the AMC-associated region to a 1.86 Mb interval between two SNP markers ALGA0032767 (70,454,750 bp, Sscrofa 11.1) and DRGA0006010 (72,315,123 bp, Sscrofa 11.1; Haubitz *et al.* 2012). To confirm the genomic region associated with AMC, we genotyped cases and controls using Illumina PorcineSNP60 Genotyping Bead chips (see data availability). To do so, we isolated genomic DNA of 11 affected piglets and 23 healthy pigs for genotyping at 62,543 SNPs using either frozen EDTA-stabilized blood or frozen tail biopsies. The physical positions of the SNPs were based on the Sscrofa11.1 assembly of the porcine genome (Warr *et al.* 2019). Quality control was carried out using PLINK (version 1.9; Chang *et al.* 2015). We retained 47,829 autosomal SNPs that were genotyped in more than 90% of the samples, had MAF greater than 0.01 and did not deviate significantly from Hardy-Weinberg proportions (P -value greater than 1×10^{-4}). To impute sporadically missing genotypes and infer haplotypes, we applied 25 iterations of the phasing algorithm implemented in Beagle 5.0 (Browning *et al.* 2018) assuming an effective population size of 200 and all other parameters set to default values. Subsequently, we applied a sliding window-based approach to compare haplotype frequency in cases and controls as described in Venhoranta *et al.* 2014 and Pausch *et al.* 2016. Specifically, we shifted a sliding window consisting of 20 contiguous SNPs along the chromosomes in steps of 5 SNPs. Within each window, haplotypes with frequency greater than 1% were tested for association with AMC using Fisher's exact test of allelic association.

The strongest association signal resulted from six adjacent haplotype windows with identical P values ($P = 9.4 \times 10^{-13}$) that were located on SSC5 between 69,601,646 and 71,330,075 bp (Fig.1a). The top haplotypes partly overlap the AMC-associated interval defined by Haubitz *et al.* (2012). The top haplotypes are in the center of a 6.06 Mb segment (from 66,757,970 to 72,815,151 bp) of extended homozygosity that was identical by descent in the 11 piglets affected by AMC. This segment was not detected in the homozygous state in 23 healthy pigs (Fig. 1b). According to the Ensembl (version 98) annotation of the porcine genome, the segment encompasses 67 genes. No haplotypes located outside the 6.06 Mb segment of extended homozygosity were associated with AMC at a Bonferroni-corrected significance threshold of 4.0×10^{-7} .

Next, we sequenced a carrier boar, two of its affected and two heterozygous piglets to an average of 13.2-fold read depth using 2×150 bp paired-end reads (see data availability). Illumina TruSeq PCR-free libraries with insert sizes of 350 bp were prepared and sequenced with an Illumina NovaSeq6000 instrument. We used the fastp software (Chen *et al.* 2018) to remove adapter sequences and reads that had phred-scaled quality less than 15 for more than 15% of the bases. Subsequently, the filtered reads were aligned to the Sscrofa 11.1 assembly of the porcine genome using the BWA-MEM algorithm implemented in the BWA software (Li 2013). The Picard tools software suite (Picard Toolkit, 2019) and Sambamba (Tarasov *et al.* 2015) were applied to mark duplicates and sort the alignments by coordinates, respectively. The variant detection (SNPs and Indels) in the five sequenced animals was performed together with 93 pigs for which sequence data were also available

from our in-house database using the multi-sample variant calling approach implemented in the Genome Analysis Toolkit (GATK, version 4.1.0; Depristo *et al.* 2011). Subsequently, we filtered the detected sequence variants by hard-filtering according to GATK's best practice guidelines. The retained sequence variants were annotated using the Ensembl Variant Effect Predictor (VEP, release 98; McLaren *et al.* 2016). We also discovered and genotyped large structural variants (> 50 bp) including insertions, deletions, inversions, duplications and translocations using Pindel v0.2.5b9 (Ye *et al.* 2009) and Delly v0.7.8 (Rausch *et al.* 2012) with the default settings. However, we did not detect structural variants (> 50 bp) that were compatible with recessive inheritance of AMC. The analysis of sequencing depth along chromosome 5 in affected and unaffected pigs using mosdepth (Pedersen and Quinlan 2018) did not reveal large deletions or duplications segregating with the AMC-associated haplotype.

To detect candidate causal variants for AMC, we considered 75,110 sequence variants located within the 6.06 Mb segment of extended homozygosity on SSC5 (from 66,757,970 to 72,815,151 bp). We retained variants as candidate causal mutations if they were homozygous in two affected piglets, heterozygous in the carrier boar and two carrier piglets, and not homozygous in 93 AMC-unaffected pigs for which the haplotype status was unknown. This filtering identified 809 variants that were compatible with recessive inheritance (see data availability). Of these 809 variants, only two were predicted to alter protein-coding sequences: a 6-bp in-frame insertion (NC_010447.5:g71673414_71673414insCGCCCG, p.Ala56_Gly57dup) in *SLC2A13* gene encoding Solute Carrier Family 2 Member 13 that was predicted to have a moderate impact on the resulting protein, and a 63-bp insertion in *KIF21A*

gene encoding Kinesin Family Member 21A that was predicted to introduce premature termination of translation (NC_010447.5:g70964237_g70964238insAAGATAGAGGTTCTTTCCTCTAGACTTGAGGTTCTCCTGGTGTGACAGATGGTTCTTTCCTCT, p.Val41_Phe42insTer). Furthermore, 807 out of 809 variants were detected in the heterozygous state in at least five AMC-unaffected control pigs from our in-house database for which the haplotype status was unknown, indicating that they occur at least at low to moderate frequency in the population. Only two variants were never detected either in the heterozygous or homozygous state in control pigs: the 63-bp insertion in *KIF21A* gene and a variant (NC_010447.5:g71800313) in the 5'UTR of *LRKK2*, further suggesting that the 63-bp insertion is a plausible causal mutation for AMC. We confirmed the presence of the 63-bp insertion in *KIF21A* gene in 11 affected piglets using PCR and subsequent sanger sequencing of PCR products from one affected piglet and one homozygous wildtype pig (Table S1, Fig.1c).

Fig.1 Identification of a 63-bp insertion in exon 2 of porcine *KIF21A* gene.

(a). Manhattan plot representing the result of a haplotype-based genome-wide association study. Red color represents haplotypes with a P value less than 4.03×10^{-7} (Bonferroni corrected-significance threshold). (b). Homozygosity mapping in 11 cases and 23 controls using SNPs located on SSC5. Blue and pale blue represent homozygous genotypes (AA and BB), heterozygous genotypes are displayed in light grey. The red bar indicates a 6.06 Mb segment of extended homozygosity shared in 11 cases. (c). The 63-bp insertion is located in exon 2 (ENSSSCT000000006489) of *KIF21A* (transcript ID: ENSSSCT000000000843). Sanger sequencing chromatograms from one wild type (wt/wt) pig and one mutant (mt/mt) piglet.

To investigate if the AMC-associated allele still segregates unnoticed in the Swiss Large White population, we screened genotypes (Porcine 60K SNP Bead chips) of 9,206 healthy pigs available from routine genomic breeding value estimation for the presence of the AMC-associated haplotype. This haplotype was in perfect LD with the 63-bp insertion in our AMC pedigree of 11 affected piglets and 23 unaffected pigs. The frequency of the AMC-associated haplotype was 1.61% in the Swiss Large White pig population. We detected the AMC-associated haplotype in 293 pigs in the heterozygous state and, surprisingly, in two healthy pigs in the homozygous state. We obtained biological materials (blood, ear tissue or hair samples) of 30 randomly selected heterozygous haplotype carriers and one homozygous haplotype carrier to isolate genomic DNA for testing the presence of the 63-bp insertion. Biological material was not available for the second homozygous haplotype carrier. PCR testing revealed that only 2 (out of 30) heterozygous haplotype carriers also carried the 63-bp insertion, indicating that either an ancestral or recombinant haplotype without the insertion also segregates in the population. The one homozygous haplotype carrier did not carry the 63-bp insertion in *KIF21A* gene. Assuming that 7% (2/30) of the haplotype carriers are indeed mutation carriers, the 63-bp insertion has an allele frequency of approximately 0.1% in the Swiss Large White population.

The *KIF21A* gene encodes Kinesin Family Member 21A, a microtubule-dependent motor protein that is predominantly expressed in the central nervous system and muscle (Marszalek *et al.* 1999, Desai *et al.* 2012). Kinesin Family Member 21A plays a crucial role in cell morphogenesis and neuronal development by regulating microtubule growth (van der Vaart *et al.* 2013). The premature stop codon in porcine *KIF21A* truncates 1,614 amino acids (~ 97%) from KIF21A (transcript ID:

ENSSSCT00000000843). If not degraded via nonsense-mediated mRNA decay, the truncated protein lacks a complete motor domain (Yamada *et al.* 2003). Mice lacking the motor domain of KIF21A die within 24h of birth (Cheng *et al.* 2014). Piglets homozygous for the 63-bp insertion were either stillborn or died shortly after birth and had congenital malformations, suggesting that a lack of the KIF21A motor domain is also lethal in the homozygous state in pigs. The multi-species comparison of KIF21A protein sequences using the constraint-based alignment tool (COBALT; Papadopoulos and Agarwala, 2007) shows a strong evolutionary conservation of the motor domain (position 9-371), further suggesting its importance in the physiological protein function (Fig. S1).

Heterozygous missense mutations in human *KIF21A* gene impair the function of the oculomotor nerve due to fibrosis of extraocular muscles (Yamada *et al.* 2003, Lu *et al.* 2008, Bianchi *et al.* 2016). Apart from eye movement disorders, people with pathogenic *KIF21A* mutation are healthy. Pigs carrying the AMC-associated haplotype harboring the 63-bp insertion in the heterozygous state were also healthy. However, their eye movement had not been investigated and described in detail. Fibrotic processes may possibly underlie the characteristically contracted joints observed in the piglets with AMC (Kalampokas *et al.* 2012). Interestingly, loss-of-function mutations of human *KIF21A* have not been observed in the homozygous state so far, suggesting that they might be lethal (Walsh and Engle 2010). We did not detect the 63-bp insertion in *KIF21A* in the homozygous state, suggesting that it represents a lethal allele. Furthermore, mutations in kinesin superfamily proteins (KIFs) cause spasticity of limbs (Reid *et al.* 2002, Dor *et al.* 2014, Duchesne *et al.*

2018), supporting our findings that a loss-of-function allele in porcine *KIF21A* gene is associated with AMC.

The identification of a plausible causal mutation for a recessive disease emerged almost 20 years ago evidences that current whole-genome genotyping and sequencing technologies may readily reveal plausible candidate causal mutations for unresolved genetic disorders. However, our results do not provide functional evidence that the putative loss-of-function allele in *KIF21A* is causal for AMC. Porcine AMC was investigated decades ago, and tissue samples of affected piglets had not been preserved once diagnostic marker testing indicated that the deleterious allele was eradicated from the population. Due to the lack of appropriate tissue samples, we could not investigate either mRNA or protein expression in affected piglets to validate the consequence of the 63-bp insertion in *KIF21A* gene. Nevertheless, our findings indicate that deleterious alleles of *KIF21A* warrant close scrutiny in unresolved cases of neurological and muscular disorders.

In conclusion, we report a plausible candidate causal variant for AMC in Swiss Large White pigs. A 63-bp insertion in *KIF21A* gene likely represents a loss-of-function allele that causes this fatal disease in the homozygous state.

Acknowledgements

We acknowledge SUISAG for providing genotype data and biological materials of healthy control animals. The sequencing of unaffected control animals was financially supported by SUISAG, Micarna SA and the ETHZ Foundation.

Availability of data

Whole-genome sequences are available at the Sequence Read Archive of the NCBI at the BioProject PRJNA622908 under sample accession numbers: SAMN14532191 and SAMN14532478 (2 piglets affected with AMC), SAMN14532792 (the carrier boar), and SAMN14532769 and SAMN14532790 (two heterozygous piglets). Whole-genome genotyping data of 11 affected and 23 healthy animals together with a VCF file containing 809 sequence variants compatible with recessive inheritance are available at <https://doi.org/10.5281/zenodo.3746178>.

References

- Bianchi S., Van Riel W.E., Kraatz S.H.W., Olieric N., Frey D., Katrukha E.A., Jaussi R., Missimer J., Grigoriev I., Olieric V., Benoit R.M., Steinmetz M.O., Akhmanova A., and Kammerer R.A. 2016. Structural basis for misregulation of kinesin KIF21A autoinhibition by CFEOM1 disease mutations. *Scientific Reports*, 6, 30668.
- Browning B.L., Zhou Y., and Browning S.R. 2018. A One-Penny Imputed Genome from Next-Generation Reference Panels. *American Journal of Human Genetics*, 103 (3), 338–348.
- Chang C.C., Chow C.C., Tellier L.C.A.M., Vattikuti S., Purcell S.M., and Lee J.J. 2015. Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience*, 4;7.
- Charlier C., Coppieters W., Rollin F., Desmecht D., Agerholm J.S., Cambisano N., Carta E., Dardano S., Dive M., Fasquelle C., Frennet J.C., Hanset R., Hubin X., Jorgensen C., Karim L., Kent M., Harvey K., Pearce B.R., Simon P., Tama N., Nie H., Vandeputte S., Lien S., Longeri M., Fredholm M., Harvey R.J., and Georges M. 2008. Highly effective SNP-based association mapping and management of recessive defects in livestock. *Nature Genetics*, 40 (4), 449–454.
- Chen S., Zhou Y., Chen Y., and Gu J. 2018. Fastp: An ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics*, 34 (17), i884–i890.
- Cheng L., Desai J., Miranda C.J., Duncan J.S., Qiu W., Nugent A.A., Kolpak A.L., Wu C.C., Drokhlyansky E., Delisle M.M., Chan W.M., Wei Y., Propst F., Samara L.R.P., Fritsch B., and Engle E.C. 2014. Human CFEOM1 mutations attenuate KIF21A autoinhibition and cause oculomotor axon stalling. *Neuron*, 82 (2), 334–349.
- Depristo M.A., Banks E., Poplin R., Garimella K. V., Maguire J.R., Hartl C., Philippakis A.A., Del Angel G., Rivas M.A., Hanna M., McKenna A., Fennell T.J., Kernysky A.M., Sivachenko A.Y., Cibulskis K., Gabriel S.B., Altshuler D., and Daly M.J. 2011. A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nature Genetics*, 43 (5), 491–501.
- Desai J., Velo M.P.R., Yamada K., Overman L.M., and Engle E.C. 2012. Spatiotemporal expression pattern of KIF21A during normal embryonic development and in congenital fibrosis of the extraocular muscles type 1 (CFEOM1). *Gene Expression Patterns*, 12 (5–6), 180–188.
- Dor T., Cinnamon Y., Raymond L., Shaag A., Bouslam N., Bouhouche A., Gaussen M., Meyer V., Durr A., Brice A., Benomar A., Stevanin G., Schuelke M., and Edvardson S. 2014. KIF1C mutations in two families with hereditary spastic paraparesis and cerebellar dysfunction. *Journal of Medical Genetics*, 51 (2), 137–142.
- Duchesne A., Vaiman A., Frah M., Floriot S., Legoueix-Rodriguez S., Desmazières A., Fritz S., Beauvallet C., Albaric O., Venot E., Bertaud M., Saintilan R., Guatteo R., Esquerré D., Branchu J., Fleming A., Brice A., Darios F., Vilotte J.L., Stevanin G., Boichard D., and El Hachimi K.H. 2018. Progressive ataxia of Charolais cattle highlights a role of KIF1C in sustainable myelination. *PLoS Genetics*, 14 (8), 1–25.
- Genini S., Malek M., Špilar Š., Nguyen T.T., Ménétrey F., Gebert S., Hagger C., Neuenschwander S., Kadarmideen H.N., Stranzinger G., and Vögeli P. 2004. Arthrogryposis multiplex congenita (AMC), a hereditary disease in swine, maps

- to Chromosome 5 by linkage analysis. *Mammalian Genome*, 15 (11), 935–941.
- Haubitz M., Neuenschwander S., and Vögeli P. 2012. Porcine arthrogryposis multiplex congenita (AMC): New diagnostic test and narrowed candidate region. *Molecular and Cellular Probes*, 26 (6), 248–252.
- Jung S., Pausch H., Langenmayer M.C., Schwarzenbacher H., Majzoub-Altweck M., Gollnick N.S., and Fries R. 2014. A nonsense mutation in PLD4 is associated with a zinc deficiency-like syndrome in Fleckvieh cattle. *BMC Genomics*, 15:623.
- Kalampokas E., Kalampokas T., Sofoudis C., Deligeoroglou E., and Botsis D. 2012. Diagnosing Arthrogryposis Multiplex Congenita: A Review. *ISRN Obstetrics and Gynecology*, 2012, 1–6.
- Li H. 2013. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM, arXiv:13033997.
- Lu S., Zhao C., Zhao K., Li N., and Larsson C. 2008. Novel and recurrent KIF21A mutations in congenital fibrosis of the extraocular muscles type 1 and 3. *Archives of Ophthalmology*, 126 (3), 388–394.
- Marszalek J.R., Weiner J.A., Farlow S.J., Chun J., and Goldstein L.S.B. 1999. Novel dendritic kinesin sorting identified by different process targeting of two related kinesins: KIF21A and KIF21B. *Journal of Cell Biology*, 145 (3), 469–479.
- McLaren W., Gil L., Hunt S.E., Riat H.S., Ritchie G.R.S., Thormann A., Flicek P., and Cunningham F. 2016. The Ensembl Variant Effect Predictor. *Genome Biology*, 17:122.
- Papadopoulos J.S. and Agarwala R. 2007. COBALT: Constraint-based alignment tool for multiple protein sequences. *Bioinformatics*, 23 (9), 1073–1079.
- Pausch H., Ammermüller S., Wurmser C., Hamann H., Tetens J., Drögemüller C., and Fries R. 2016. A nonsense mutation in the COL7A1 gene causes epidermolysis bullosa in Vorderwald cattle. *BMC Genetics*, 17:149.
- Pausch H., Schwarzenbacher H., Burgstaller J., Flisikowski K., Wurmser C., Jansen S., Jung S., Schnieke A., Wittek T., and Fries R. 2015. Homozygous haplotype deficiency reveals deleterious mutations compromising reproductive and rearing success in cattle. *BMC Genomics*, 16:312.
- Pedersen B.S. and Quinlan A.R. 2018. Mosdepth: Quick coverage calculation for genomes and exomes. *Bioinformatics*, 34 (5), 867–868.
- Picard Toolkit 2019. *Broad Institute, GitHub Repository*.
- Rausch T., Zichner T., Schlattl A., Stütz A.M., Benes V., and Korbel J.O. 2012. DELLY: Structural variant discovery by integrated paired-end and split-read analysis. *Bioinformatics*, 28 (18), 333–339.
- Reid E., Kloos M., Ashley-Koch A., Hughes L., Bevan S., Svenson I.K., Graham F.L., Gaskell P.C., Dearlove A., Pericak-Vance M.A., Rubinsztein D.C., and Marchuk D.A. 2002. A kinesin heavy chain (KIF5A) mutation in hereditary spastic paraplegia (SPG10). *American Journal of Human Genetics*, 71 (5), 1189–1194.
- Tarasov A., Vilella A.J., Cuppen E., Nijman I.J., and Prins P. 2015. Sambamba: Fast processing of NGS alignment formats. *Bioinformatics*, 31 (12), 2032–2034.
- van der Vaart B., van Riel W.E., Doodhi H., Kevenaar J.T., Katrukha E.A., Gumy L., Bouchet B.P., Grigoriev I., Spangler S.A., Yu K. Lou, Wulf P.S., Wu J., Lansbergen G., van Battum E.Y., Pasterkamp R.J., Mimori-Kiyosue Y., Demmers J., Olieric N., Maly I. V., Hoogenraad C.C., and Akhmanova A. 2013. CFEOM1-associated kinesin KIF21A is a cortical microtubule growth inhibitor. *Developmental Cell*, 27 (2), 145–160.
- Venhoranta H., Pausch H., Flisikowski K., Wurmser C., Taponen J., Rautala H., Kind A., Schnieke A., Fries R., Lohi H., and Andersson M. 2014. In frame exon

- skipping in UBE3B is associated with developmental disorders and increased mortality in cattle. *BMC Genomics*, 15:890 (1), 1–9.
- Walsh C.A. and Engle E.C. 2010. Allelic diversity in human developmental neurogenetics: Insights into biology and disease. *Neuron*, 68 (2), 245–253.
- Warr A., Affara N., Aken B., Beiki H., Bickhart D.M., Billis K., Chow W., Eory L., Finlayson H.A., Flicek P., Girón C.G., Griffin D.K., Hall R., Hannum G., Hourlier T., Howe K., Hume D.A., Izuogu O., Kim K., Koren S., Liu H., Manchanda N., Martin F.J., Nonneman D.J., O'Connor R.E., Phillippy A.M., Rohrer G.A., Rosen B.D., Rund L.A., Sargent C.A., Schook L.B., Schroeder S.G., Schwartz A.S., Skinner B.M., Talbot R., Tseng E., Tuggle C.K., Watson M., Smith T.P.L., and Archibald A.L. 2019. An improved pig reference genome sequence to enable pig genetics and genomics research. *bioRxiv*, 668921.
- Yamada K., Andrews C., Chan W.M., McKeown C.A., Magli A., De Berardinis T., Loewenstein A., Lazar M., O'Keefe M., Letson R., London A., Ruttum M., Matsumoto N., Saito N., Morris L., Del Monte M., Johnson R.H., Uyama E., Houtman W.A., De Vries B., Carlow T.J., Hart B.L., Krawiecki N., Shoffner J., Vogel M.C., Katowitz J., Goldstein S.M., Levin A. V., Sener E.C., Ozturk B.T., Akarsu A.N., Brodsky M.C., Hanisch F., Cruse R.P., Zubcov A.A., Robb R.M., Roggenkämper P., Gottlob I., Kowal L., Battu R., Traboulsi E.I., Franceschini P., Newlin A., Demer J.L., and Engle E.C. 2003. Heterozygous mutations of the kinesin KIF21A in congenital fibrosis of the extraocular muscles type 1 (CFEOM1). *Nature Genetics*, 35 (4), 318–321.
- Ye K., Schulz M.H., Long Q., Apweiler R., and Ning Z. 2009. Pindel: A pattern growth approach to detect break points of large deletions and medium sized insertions from paired-end short reads. *Bioinformatics*, 25 (21), 2865–2871.

Supporting information

Table S1 PCR primers and condition for genotyping and sanger sequencing the 63-bp insertion in porcine *KIF21A* gene

Fig. S1 Alignment of KIF21A protein sequences in multiple species using the constraint-based alignment tool (COBALT).