Deletion of an intronic polypyrimidine tract of porcine *DNAH17* perturbs splicing and causes defective sperm flagella

Adéla Nosková*[,1], Maya Hiltpold*[,1], Fredi Janett[†], Thomas Echtermann[‡], Zih-Hua Fang*, Xaver Sidler[‡], Christin Selige[§], Andreas Hofer[§], Stefan Neuenschwander**, Hubert Pausch*

*Animal Genomics, Institute of Agricultural Sciences, ETH Zürich, 8315 Lindau, Switzerland

[†]Clinic of Reproductive Medicine, Vetsuisse Faculty, University of Zurich, 8057 Zurich, Switzerland

[‡]Division of Swine Medicine, Vetsuisse Faculty, University of Zurich, 8057 Zurich, Switzerland

[§]SUISAG, 6204 Sempach, Switzerland

**Animal Genetics, Institute of Agricultural Science, ETH Zürich, 8092 Zürich, Switzerland

[1]authors contributed equally

AN: adela.Noskova@usys.ethz.ch
MH: maya.hiltpold@usys.ethz.ch
FJ: fredi.janett@uzh.ch
TE: techtermann@vetclinics.uzh.ch
ZHF: zih-hua.fang@usys.ethz.ch
XS: xsidler@vetclinics.uzh.ch
CS: cse@suisag.ch
AH: aho@suisag.ch
SN: stefan.neuenschwander@usys.ethz.ch
HP: hubert.pausch@usys.ethz.ch

Data availability:

- ***DNA and RNA sequencing*** data are available at the European Nucleotide Archive (ENA) of the EMBL at BioProject
  - PRJEB38156 (https://www.ebi.ac.uk/ena/browser/view/PRJEB38156)
  - PRJEB37956 (https://www.ebi.ac.uk/ena/browser/view/PRJEB37956)
  - PRJEB39374 (https://www.ebi.ac.uk/ena/browser/view/PRJEB39374)
  - PRJNA622908 (https://www.ebi.ac.uk/ena/browser/view/PRJNA622908)
  - PRJNA506525 (https://www.ebi.ac.uk/ena/browser/view/PRJNA506525)
- ***Whole-genome genotypes*** in the form of plink binary files and **phased haplotypes** of 6 affected and 100 unaffected Large White boars are available at zenodo: https://doi.org/10.5281/zenodo.4081475
- **Genotypes of candidate causal variants** for 5 affected and 82 unaffected boars are available in a VCF file at zenodo: https://doi.org/10.5281/zenodo.4081475

Running title: *DNAH17* exon skipping in infertile boars

Key words: Sperm flagella defect, impaired male fertility, polypyrimidine tract, exon skipping, genetic disorder

Corresponding author

Prof. Dr. Hubert Pausch
ETH Zurich
Animal Genomics
Institute of Agricultural Sciences (IAS)
EHB E21
Eschikon 27
8315 Lindau, Switzerland
Phone +41 52 354 92 34
hubert.pausch@usys.ethz.ch
www.ag.ethz.ch

ABSTRACT

Artificial insemination in pig (*Sus scrofa domesticus*) breeding involves the evaluation of the semen quality of breeding boars. Ejaculates that fulfill predefined quality requirements are processed, diluted and used for inseminations. Within short time, eight Swiss Large White boars producing immotile sperm that had multiple morphological abnormalities of the sperm flagella were noticed. The eight boars were inbred on a common ancestor suggesting that the novel sperm flagella defect is a recessive trait. Transmission electron microscopy cross-sections revealed that the immotile sperm had disorganized flagellar axonemes. Haplotype-based association testing involving microarray-derived genotypes at 41,094 SNPs of six affected and 100 fertile boars yielded strong association ($P=4.22 \times 10^{-15}$) at chromosome 12. Autozygosity mapping enabled us to pinpoint the causal mutation on a 1.11 Mb haplotype located between 3,473,632 and 4,587,759 bp. The haplotype carries an intronic 13-bp deletion (Chr12:3,556,401-3,556,414 bp) that is compatible with recessive inheritance. The 13-bp deletion excises the polypyrimidine tract upstream exon 56 of *DNAH17* (XM_021066525.1:c.8510-17_8510-5del) encoding dynein axonemal heavy chain 17. Transcriptome analysis of the testis of two affected boars revealed that the loss of the polypyrimidine tract causes exon skipping which results in the in-frame loss of 89 amino acids from DNAH17. Disruption of DNAH17 impairs the assembly of the flagellar axoneme and manifests in multiple morphological abnormalities of the sperm flagella. Direct gene testing may now be implemented to monitor the defective allele in the Swiss Large White population and prevent the frequent manifestation of a sterilizing sperm tail disorder in breeding boars.

INTRODUCTION

Artificial insemination is the most frequent method of breeding in pigs. The semen of breeding boars is collected once or twice per week at semen collection centers. All ejaculates are macro- and microscopically evaluated. Traits that are routinely measured in all ejaculates include ejaculate volume, sperm concentration, motility, and morphology. Only ejaculates that meet the quality requirements for artificial insemination are processed, diluted and used for inseminations (Colenbrander et al., 1993),(Broekhuijse et al., 2011),(Holt et al., 1997). Semen quality and insemination

success vary within and between boars due to environmental, permanent environmental and genetic effects (Marques et al., 2017).

Access to longitudinal data on standardized semen traits for large cohorts of males is unique to livestock populations (Thibier and Wagner, 2002),(Knox, 2016). Boars suitable for breeding are selected based on semen quality and genomic predictions derived from dense SNP microarray-derived genotypes. Dense genotypes and repeated semen trait measurements for thousands of males are prerequisites to unravel the genetic architecture of male fertility. The comprehensive genetic analysis of male fertility is challenging in most species other than livestock. For instance, the analysis of male fertility in humans typically relies on correlated proxy phenotypes such as family size and birth rate (Kosova et al., 2012), small cohorts of males (Sato et al., 2018),(Sato et al., 2020), and observations from a single ejaculate per individual (Rahban et al., 2019).

Due to the widespread use of artificial insemination in livestock populations, the semen from individual male breeding animals may be used for thousands of inseminations. Boar semen quality is positively correlated with insemination success and litter size (Holt et al., 1997). Semen traits that are measured at semen collection centers have medium heritability (Marques et al., 2017). Thus, they are well suited response variables in genome-wide association studies for male reproductive performance (Diniz et al., 2014),(Hiltpold et al., 2020). Genome-wide association studies on male fertility carried out in livestock discovered novel phenotype-genotype associations that improved our biological understanding of mammalian fertilization (Pausch et al., 2014),(Lamas-Toranzo et al., 2020, p. 95),(Noda et al., 2020),(Hiltpold et al., 2020),(Ni et al., 2020, p. 19),(Noskova et al., 2020).

Ejaculates from hundreds of boars are evaluated every year at semen collection centers as a service to the pig breeding industry. This rich resource of semen quality records facilitates to investigate sporadically occurring sperm defects using case-control association testing (Sironen et al., 2006),(Sironen et al., 2011),(Noskova et al., 2020). Once causative variants have been identified, direct gene tests and genome-based mating strategies may be implemented to avoid the birth of infertile males. Moreover, the discovery of genes that harbor pathogenic alleles that compromise male fertility is important to enhance the diagnostic yield of genetic testing also in species other than livestock (Xavier et al., 2020).

128 Here, we investigate an autosomal recessive sperm tail defect of Swiss Large White
129 boars. Using genome-wide association testing, we map the disorder to porcine
130 chromosome 12. The analysis of genome-wide DNA and RNA sequencing data of
131 affected boars revealed that the loss of an intronic polypyrimidine tract of porcine
132 *DNAH17* is causal for the morphological abnormalities of the sperm flagella.

133

134

135 **MATERIAL AND METHODS**

136 **Ethics approval and consent to participate**

137 Semen samples of five breeding boars were collected by laboratory technicians at an
138 approved semen collection center as part of their regular service to the Swiss pig
139 breeding industry. Semen samples of three boars were flushed from the epididymis
140 post mortem. Testes were collected after regular slaughter at an approved
141 slaughterhouse. Our study was approved by the veterinary office of the Canton of
142 Zurich (animal experimentation permit ZH 070/20).

143 **Consent for publication**

144 SUISAG, the Swiss pig breeding and competence center provided written consent to
145 the analyses performed and agreed to publish results and data.

146 **Animals**

147 Eight Swiss Large White boars with a sperm tail defect were considered in our study
148 (**Table 1**). Five of them were noticed at the semen collection center of SUISAG
149 because their ejaculates contained immotile spermatozoa that had multiple
150 morphological abnormalities of the flagella. Because the five boars were healthy and
151 pedigree analysis indicated that they were inbred on a common ancestor, recessive
152 inheritance of the sperm tail defect was suspected. A mating between two suspected
153 carrier animals was performed in the field to assess phenotypic manifestations in
154 their offspring. The pregnant sow was purchased and maintained at the research
155 barn of the Division of Swine Medicine, Vetsuisse Faculty, University of Zurich. The
156 sow gave birth to a litter with eleven piglets (eight females, three males). One of the
157 male piglets (Boar_1246) died at the age of 200 days due to hemorrhagic bowel
158 syndrome (Grahofer et al., 2017). The other two male boars (Boar_1249,
159 Boar_1254) were slaughtered at the age of 17 months at a regular slaughterhouse.

160 All three male piglets from the mating of suspected carrier animals expressed the
161 sperm tail defect (see below). Pedigree records were analyzed using the PYPEDAL
162 software package (Cole, 2007).

163

164 **Table 1: Experimental design of the study**

Boar	AI center	Research station	GWAS cohort	PCR-based genotyping	Whole-genome sequence data	RNA sequencing data	Macro- and microscopic sperm analyses	Eosin-nigrosin staining	TEM sections
Boar_78	X		X		SAMEA6813549		X		
Boar_79	X		X		SAMEA6813550		X		
Boar_80	X		X	X	SAMEA6813551	SAMEA6832284	X		
Boar_81	X		X		SAMEA6813552	SAMEA6832285	X		
Boar_82	X		X		SAMEA6813553		X		
Boar_1246[a]		X	X				X		
Boar_1249		X		X			X	X	X
Boar_1254		X		X			X	X	X

165 [a] died at the age of 200 days due to hemorrhagic bowel syndrome; sperm were flushed from the
166 epididymis post mortem.
167 Details on the eight boars with a sperm tail defect. Genotypes and haplotypes for the GWAS cohort
168 are available at https://doi.org/10.5281/zenodo.4081475. DNA and RNA sequencing data from the
169 boars are available at the European Nucleotide Archive (ENA) of the EMBL at the BioProject
170 PRJEB38156 under the accession numbers listed in the table.

171

172 **Phenotypes**

173 Semen samples of five breeding boars were macro- and microscopically evaluated
174 by laboratory technicians as part of the routine service of SUISAG to the Swiss pig
175 breeding industry. Spermatozoa from three boars (Boar_1246, Boar_1249,
176 Boar_1254) were flushed from the epididymis post mortem. Semen samples from a
177 fertile boar were provided by SUISAG.

178 Sperm concentration, total sperm count and sperm motility were determined with an
179 IVOS II CASA system (Hamilton Thorne Inc., Beverly, U.S.A.) using Leja 2-chamber
180 slides (Leja, Nieuw-Vennep, the Netherlands). For morphological examination,
181 semen was fixed in buffered formol saline solution (Na2HPO4 4.93g, KH2PO4
182 2.54g, 38% formaldehyde 125ml, NaCl 5.41g, distilled water q.s. 1000ml) and
183 smears prepared (Hancock, 1957). At least 200 spermatozoa were subsequently

184 evaluated by phase contrast microscopy using oil immersion (Olympus BX50,
185 UplanF1 100×/1.30, Olympus, Wallisellen, Switzerland).

186 Assessment of sperm viability was performed using the eosin-nigrosin staining
187 method (Blom, 1950). In stained slides at least 200 spermatozoa were evaluated
188 under oil immersion on a light microscope (Olympus BX50, UplanF1 100×/1.30,
189 Olympus, Wallisellen).

190 In order to prepare sperm flushed from the epididymis for transmission electron
191 microscopy (TEM), the samples were washed twice in phosphate buffered saline
192 (PBS) and subsequently centrifuged at 300g. A semen sample from an unaffected
193 boar was first centrifuged at 300g to increase the sperm concentration and then
194 washed in PBS. The pellet was fixed with equal volume of 6% glutaraldehyde,
195 resuspended gently, and centrifuged at 6000g. After removing the supernatant, the
196 sperm were fixed for a second time with 3% glutaraldehyde, and finally pelleted at
197 6000g. Pellets were washed three times in PBS, post-fixed in 1% osmium tetroxide,
198 washed in ddH2O, stained in 1% uranyl acetate, dehydrated in graded series of
199 ethanol (25, 50, 75, 90, and 100%), and embedded in Epoxy resin through
200 increasing concentrations (25, 50, 75, and 100%) using PELCO Biowave+ tissue
201 processor, and then cured at 60 °C for 3 days. Embedded blocks were sectioned
202 using Leica FC6 microtome and a DIATOME diamond knife with 45° angle into
203 60 nm sections and mounted on Quantifoil copper grids with formvar carbon films.
204 Sections were post-stained with 2% uranyl acetate followed by lead citrate. Grids
205 were imaged using FEI Morgagni 268 electron microscope operated at 100 kV at
206 20 k magnification.

207 **Genotypes and haplotype inference**

208 Genotypes of 9955 Large White boars (including six boars with the sperm tail defect)
209 and sows were provided by SUISAG. All pigs were genotyped using Illumina
210 PorcineSNP60 Bead chips that comprised between 62,163 and 68,528 SNPs. We
211 considered only autosomal SNPs. Physical positions of the SNPs corresponded to
212 the Sscrofa11.1-assembly of the porcine genome (Warr et al., 2019). Quality control
213 on the genotypes was carried using the PLINK (version 1.9) software (Chang et al.,
214 2015). Animals and SNPs with more than 10% missing genotypes were excluded
215 from subsequent analyses. We removed SNPs with minor allele frequency (MAF)
216 less than 0.005 and SNPs for which the observed genotype distribution deviated

217 significantly (P < 0.00001) from Hardy-Weinberg proportions. After quality control,
218 our dataset contained 9848 pigs and genotypes at 43,254 autosomal SNPs.
219 Sporadically missing genotypes were imputed and haplotypes were inferred using 12
220 iterations of the phasing algorithm implemented in the BEAGLE (version 5.0) software
221 (Browning et al., 2018), assuming an effective population size of 500 while all other
222 parameters were set to default values.

223 **Genome-wide association testing**

224 Six boars that produced sperm with multiple morphological abnormalities of the
225 flagella were considered as case group for a genome-wide association study. The
226 control group consisted of 100 randomly selected boars that produced normal sperm
227 and were fertile, i.e., each of them sired at least one litter in the Swiss Large White
228 breeding unit.

229 We tested the association between the affection status of the boars and 41,094
230 autosomal SNPs for which the frequency of the minor allele was greater than 5%.
231 Single marker-based case-control association testing was performed using Fisher's
232 exact tests of allelic association as implemented in the PLINK (version 1.9) software
233 (Chang et al., 2015).

234 A linear haplotype-based association model was implemented as described in
235 (Venhoranta et al., 2014) and (Hiltpold et al., 2020). We shifted a sliding window
236 consisting of 25 contiguous SNPs corresponding to an average haplotype length of
237 1.28 ± 0.65 Mb along the chromosomes in steps of 5 SNPs. Within each of the 8583
238 sliding windows considered, we tested the association of between 1 and 10
239 haplotypes that had frequency greater than 5% with the sperm tail defect using the
240 linear model (Hiltpold et al., 2020) $y = \mu + \sum_{j=1}^{30} a_j PC_j + bHT + e$, where y is a vector
241 of phenotypes (1-fertile, 2-infertile), μ is the intercept, PC_j are the top 30 principal
242 components of a genomic relationship matrix built based on genotypes of 43,254
243 autosomal SNPs using the PLINK software, a and b are effects of the principal
244 components and the haplotype (HT) tested, respectively, and e is a vector of
245 residuals that are assumed to be normally distributed. In total, we tested 47,685
246 haplotypes for association with the sperm tail defect.

247 SNPs and haplotypes that exceeded the Bonferroni-corrected significance threshold
248 (P_{SNP}=1.22 x 10^{-6}, P_{HAPLO}=1.05 x 10^{-6}) were considered as significantly associated.

249 The inflation factors of the test statistics were calculated using the *estlambda()*-
250 function of the GENABEL R package (Aulchenko et al., 2007).

251 **Analysis of positional candidate genes**

252 The top association signal revealed a 1.11 Mb segment of extended homozygosity at
253 chromosome 12. Genes located within the segment of homozygosity were obtained
254 from the Refseq annotation of the porcine genome (version 106, available at
255 ftp://ftp.ncbi.nlm.nih.gov/genomes/refseq/vertebrate_mammalian/Sus_scrofa/annotat
256 ion_releases/106/GCF_000003025.6_Sscrofa11.1/GCF_000003025.6_Sscrofa11.1_
257 genomic.gff.gz). Using publicly available RNA sequencing data (BioProject
258 PRJNA506525, sample accession numbers SAMN10462191 – SAMN10462197)
259 from (Robic et al., 2019), we quantified mRNA abundance of the positional candidate
260 genes in porcine testes. To this end, we downloaded between 47 and 58 million 2 x
261 150 bp paired-end sequencing reads that were generated from RNA extracted from
262 testis tissue of seven pubertal boars from European Piétrain and Landrace
263 populations. The RNA sequencing reads were pseudo-aligned to an index of the
264 porcine transcriptome (Refseq version 106) and transcript abundance was quantified
265 using the KALLISTO software (Bray et al., 2016). We used the R package TXIMPORT
266 (Soneson et al., 2015) to aggregate transcript abundances to the gene level.

267 **Whole-genome sequencing and sequence variant genotyping**

268 We sequenced five Swiss Large White boars that produced sperm that had multiple
269 morphological abnormalities of the sperm flagella using 2 ×150 bp paired-end reads.
270 Illumina TruSeq DNA PCR-free libraries with 400 bp insert sizes were prepared and
271 sequenced at an Illumina NovaSeq6000 instrument. We used the FASTP software
272 (Chen et al., 2018) to remove adapter sequences, poly-G tails and reads that had
273 phred-scaled quality less than 15 for more than 15% of the bases. Between
274 81,145,807 and 114,616,836 filtered read pairs were aligned to the Sscrofa11.1
275 assembly of the porcine genome using the MEM algorithm of the BWA software (Li,
276 2013) with default parameter settings. We marked duplicates using the PICARD tools
277 software suite (https://github.com/broadinstitute/picard) and sorted the alignments by
278 coordinates using SAMBAMBA (Tarasov et al., 2015). The number of reads covering a
279 position were extracted from the alignments using the MOSDEPTH software (version
280 0.2.2) with default parameter settings (Pedersen and Quinlan, 2018). The realized

281 sequencing read depth of the five boars was 10.9-fold and it ranged from 8.9- to
282 12.5-fold.

283 Sequence variants (SNPs and Indels) of the five sequenced boars were genotyped
284 together with 93 pigs from different breeds for which whole-genome sequence data
285 were available from our in-house database using the multi-sample variant calling
286 approach of the GENOME ANALYSIS TOOLKIT (GATK, version 4.1.0; (Depristo et al.,
287 2011)). We filtered the sequence variants by hard-filtering according to best practice
288 guidelines of the GATK. A detailed description of our reference-guided variant
289 discovery and filtration approach is described in (Crysnanto et al., 2019). Functional
290 consequences of polymorphic sites were predicted according to the Refseq
291 annotation (version 106) of the porcine genome using the VARIANT EFFECT PREDICTOR
292 software from Ensembl (McLaren et al., 2016) along with the SpliceRegion.pm plugin
293 (https://github.com/Ensembl/VEP_plugins/blob/release/101/SpliceRegion.pm). We
294 considered 82 pigs with known pedigree as a control group for the present study.

295 **Identification of candidate causal variants**

296 We considered 14,806 SNPs and 3,856 Indels that were detected within the 1.11 Mb
297 segment (between 3,473,632 and 4,587,759 bp) of extended homozygosity at
298 chromosome 12 as positional candidate causal variants. To identify variants
299 compatible with recessive inheritance of the sperm tail disorder, we applied a filtering
300 strategy that takes into account flawed genotypes and the under-calling of
301 heterozygous genotypes due to relatively low sequencing coverage. Specifically, we
302 screened for alleles that had following frequency:

303 - $\geq$ 0.8 in five affected boars (at least 8 out of 10 alleles),
304 - $\leq$ 0.05 in 82 control pigs from different breeds (less than 8 alleles).

305 This filtration resulted in only one compatible variant which was a 13-bp intronic
306 deletion.

307 **Whole transcriptome sequencing and read alignment**

308 Testes were collected immediately after slaughter at an approved slaughterhouse
309 from two boars (Boar_80, Boar_81) that were homozygous for the 13-bp deletion
310 and produced sperm with multiple morphological abnormalities of the flagella. Tissue
311 samples were frozen in liquid nitrogen and stored at -80°C until RNA extraction.
312 Paired-end RNA libraries (2 x 150 bp) were prepared using total RNA purified from

313 testis tissue using the Illumina TruSeq RNA Sample Preparation Kit (Illumina, San
314 Diego, CA, USA). The libraries were sequenced at an Illumina NovaSeq6000
315 instrument yielding 49,198,010 and 70,451,148 reads. Quality control on the raw
316 RNA sequencing reads was performed using the FASTP software using default
317 parameter settings (see above). The filtered read pairs (48,889,296 and 69,988,376)
318 were aligned to the Sscrofa11.1 reference sequence and the Refseq gene
319 annotation (version 106) using the splice-aware read alignment tool STAR (version
320 2.7.3a) (Dobin et al., 2013). The number of RNA sequencing reads that covered a
321 position was counted using the MOSDEPTH (see above).

322 **Bioinformatic analysis**

323 Putative branch point sequences within *DNAH17* (XM_021066525.1) intron 55 were
324 predicted using the BPP algorithm (Zhang et al., 2017). Putative 3' splice acceptor
325 sites were predicted using the NNSPLICE software tool
326 (https://www.fruitfly.org/seq_tools/splice.html (Reese et al., 1997)). Domains of
327 porcine DNAH17 (protein-ID: A0A287AFU3) were retrieved from Uniprot.

328 **Genotyping of the 13-bp deletion**

329 DNA was isolated from either hair roots (living pigs) or spleen (slaughtered pigs)
330 using DNeasy blood and tissue kit (Qiagen). A PCR test was established consisting
331 of a FAM-labelled forward primer (Fw-*DNAH17*: 5'-TGAGCATCTTCTTGGCGAGG-
332 3') and a reverse primer (Re-*DNAH17*: 5'-GCTGTTGATCAGCACCAGGA-3') which
333 bind to the wild type and mutant alleles, as well as a wild type specific reverse primer
334 (Re-wt-*DNAH17*: 5'-TCGAGCTAGACGCGAGGG-3'). PCR fragments were analyzed
335 on a DNA analyzer 3130xl (Applied Bioscience).

336 **Availability of data**

337 Whole-genome sequence data of 87 pigs including five boars with a sperm tail defect
338 have been deposited at the European Nucleotide Archive (ENA) of the EMBL at
339 BioProject PRJEB38156, PRJEB37956, PRJEB39374 and PRJNA622908, under
340 sample accession numbers listed in **Supplemental file S1**. The genotypes of five
341 affected and 82 unaffected boars at 14,806 SNPs and 3,856 Indels that were
342 detected within the segment of autozygosity are available from Zenodo
343 (https://doi.org/10.5281/zenodo.4081475). RNA sequencing data of testicular tissue
344 samples of two boars homozygous for the 13-bp deletion have been deposited at the

345 European Nucleotide Archive (ENA) of the EMBL at the BioProject PRJEB38156
346 under sample accession numbers SAMEA6832284 and SAMEA6832285. RNA
347 sequencing data of testicular tissue samples of seven pubertal boars are accessible
348 via the European Nucleotide Archive (ENA) of the EMBL at the BioProject
349 PRJNA506525 under sample accession numbers SAMN10462191 –
350 SAMN10462197. The microarray-derived genotypes and haplotypes of cases and
351 controls, the eigenvectors considered to account for stratification, the R script use to
352 carry out haplotype-based association testing, and the results from the haplotype-
353 based association study are available from Zenodo
354 (https://doi.org/10.5281/zenodo.4081475).

355

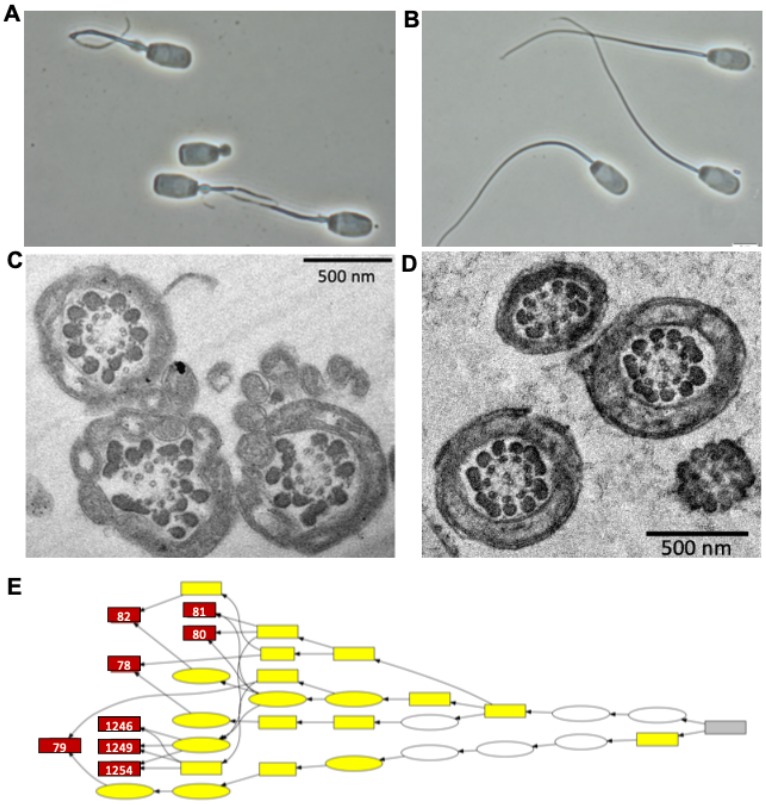
356

357 **RESULTS**

358 **Phenotypic manifestation of a sperm tail disorder in Large White boars**
359 Five Swiss Large White boars maintained at a semen collection center produced
360 ejaculates that contained spermatozoa with defective tails. Microscopic semen
361 analysis revealed multiple morphological abnormalities of the sperm flagella
362 including rudimentary, short, coiled and irregularly shaped tails (**Figure 1A**).
363 Proximal cytoplasmic droplets were frequently observed at the junction of sperm
364 head and tail. Progressively motile sperm were not detected in the ejaculates. Eosin-
365 nigrosin staining of semen flushed from the epididymis of two affected boars
366 indicated that 48 and 60% of the sperm were viable (**Supplemental file S2**).
367 Ejaculate volume and sperm concentration were normal. Apart from producing
368 sperm with defective tails, the boars were healthy and had no testicular
369 abnormalities. Due to the absence of motile sperm and the high degree of
370 abnormalities of the flagella, the boars were not suitable for breeding and were
371 slaughtered. The sperm tail defect of the boars is commonly referred to as
372 asthenoteratozoospermia.

373 Transmission electron microscopy cross-sections of sperm flagella from a fertile boar
374 revealed the typical axonemal arrangement of nine outer microtubule doublets
375 surrounding the central pair (9x2 + 2) (**Figure 1D**). In contrast, cross-sections of
376 sperm that were flushed from the epididymis of two affected boars revealed multiple
377 ultrastructural abnormalities including completely disorganized axonemes (**Figure**

378 1C). Central and outer microtubule doublets were absent in some flagella. Some
379 sperm had superfluous but disorganized axonemal structures. The cytoplasmic bags
380 that were detected at the junction between sperm head and tail using light
381 microscopy contained unassembled axonemal components (**Supplemental file S3**).



382

383 **Figure 1. Characteristics of a recessive sperm tail defect of Large White boars**. Representative
384 phase contrast microscopy images of sperm produced from affected (A) and fertile (B) boars.
385 Ejaculates of affected boars contained sperm with multiple morphological abnormalities of the flagella
386 including short, coiled, broken and bent tails. A proximal cytoplasmic droplet was detected between
387 the sperm head and tail at most of the spermatozoa from affected boars. Representative transmission
388 electron microscopy cross-sections of sperm flagella from affected (C) and fertile (D) boars. Cross
389 sections of the sperm flagella revealed the typical arrangement of nine outer microtubule doublets
390 surrounding the central pair (9×2 + 2) in fertile boars (D). Lack of central and outer microtubules
391 doublets and complete axonemal disorganization were frequently detected in cross-sections of the
392 sperm flagella in affected boars. Pedigree of eight boars that produced sperm with defective tails (D).
393 The pedigree contains only obligate mutation carriers. Red color indicates the eight affected boars.
394 The numbers within the boxes correspond to the IDs in Table 1. Boars 1246, 1249 and 1254 originate
395 from a mating between two suspected carrier animals. Ovals and boxes represent female and male
396 ancestors, respectively. The grey box indicates a common ancestor (born in 2005) on which all
397 affected boars were inbred. Shapes filled with yellow color represent individuals that were confirmed
398 to carry the top haplotype in heterozygous state using Illumina Porcine60K genotypes.

399 The boars that produced defective sperm were closely related (**Figure 1E**). Two
400 affected boars were fullsibs. The average relationship coefficient between the boars
401 was 0.28 and it ranged from 0.14 to 0.63. Their average coefficient of inbreeding was
402 0.087, which is slightly higher than in the fertile boars (F=0.06). All affected boars
403 were inbred on a boar born in 2005. This common ancestor was present in both the
404 paternal and maternal ancestry of significantly less (N=10, $P_{Fisher's\ exact}$=4.69 x 10^{-6})
405 fertile boars. Three sires of affected boars were used in artificial insemination. None
406 of their ejaculates (n=168) contained an anomalous amount of sperm with defective
407 flagella. These observations are compatible with an autosomal recessive inheritance
408 of the sperm tail defect.

409 In order to verify the preseumed recessive inheritance and monitor phenotypic
410 manifestations in homozygous boars, a mating between two suspected carrier
411 animals was performed in the field (**Figure 1E**). The sow gave birth to a litter with
412 eight female and three male piglets (Boar_1246, Boar_1249, Boar_1254). The three
413 male piglets were maintained at a research barn. Boar_1246 died at the age of 200
414 days due to hemorrhagic bowel syndrome. Testes and epididymis were collected
415 post mortem. Testicular abnormalities were not detected. Microscopic analysis of
416 sperm flushed from the epididymis revealed that the sperm flagella had multiple
417 morphological abnormalities. Boar_1249 and Boar_1254 were healthy and
418 developed normal at the research barn. At 17 months, both boars were slaughtered
419 at a regular slaughterhouse and testes and epididymis were collected. Microscopic
420 analysis of sperm flushed from the epididymis revealed that both boars produced
421 sperm with multiple morphological abnormalities of the flagella.

422 **Haplotype-based association testing maps the sperm tail defect to a 1.11 Mb**
423 **interval on porcine chromosome 12**

424 Six affected boars (**Table 1**) were genotyped with the Illumina PorcineSNP60
425 microarray. As a control group, we considered 100 randomly selected fertile boars of
426 the Swiss Large White breeding population that also had Illumina PorcineSNP60-
427 derived genotypes. Following quality control, genotypes at 41,094 autosomal SNPs
428 were used for association testing. Single marker-based case/control-association
429 testing using Fisher's exact test of allelic association revealed eleven SNPs located
430 on porcine chromosomes 11, 12 and 13 that exceeded the Bonferroni-corrected
431 significance threshold (**Figure 2A**). The strongest association signal (P=1.87 x 10^{-9})

resulted from SNP *ASGA0052524* that is located at 2,785,333 bp on porcine chromosome 12. An inflation factor of 1.55 indicated that the SNP-based association study was enriched for false positive association signals. Inspection of the top principal components of the genomic relationship matrix revealed clustering of the boars with the sperm tail defect (**Supplemental file S4**), suggesting that population stratification confounded the association analysis, thereby producing spurious association signals.

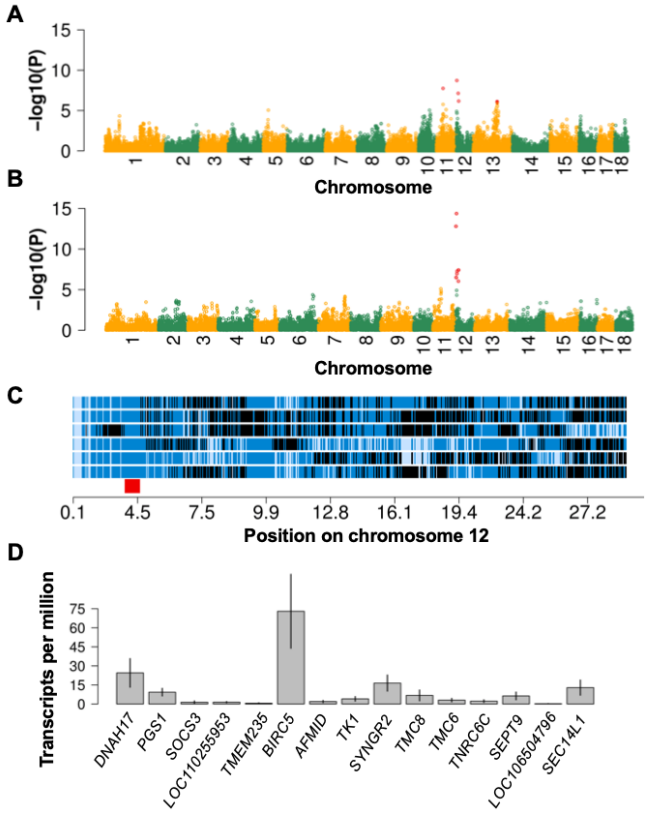


Figure 2. A 1.11 Mb segment on porcine chromosome 12 is associated with the sperm tail defect. Manhattan plot of a SNP- (A) and haplotype-based (B) genome-wide association study. Red color represents eleven SNPs and twenty-one haplotypes that exceeded the Bonferroni corrected-significance threshold. Homozygosity mapping in six boars that produced sperm with defective tails using SNPs located on porcine chromosome 12 (C). Blue and pale blue represent homozygous genotypes (AA and BB), heterozygous genotypes are displayed in black. The red bar indicates a 1.11 Mb segment of extended homozygosity that was shared in all six affected boars. mRNA abundance (quantified in transcripts per million) of 15 genes that are located within the segment of extended homozygosity in testicular tissue of pubertal boars (D). The vertical line represents the standard deviation estimated across seven samples.

451 To take population stratification into account, we repeated the association analysis
452 using a haplotype-based linear model that included 30 principal components of a
453 genomic relationship matrix as covariates. An inflation factor of 1.15 indicated that
454 the principal components-based control of population stratification was mostly
455 successful. Twenty-one haplotypes located between 94,345 and 9,256,260 bp on
456 chromosome 12 (SSC12) exceeded the Bonferroni-corrected significance threshold
457 (**Figure 2B**). The strongest associations ($P=4.22 \times 10^{-15}$) were detected for three
458 overlapping haplotypes located between 2,166,472 and 3,755,956 bp. The top
459 haplotype occurred at a frequency of 3.5% among the 100 fertile boars. Seven fertile
460 boars were heterozygous carriers of the top haplotype. The top haplotype was not
461 detected in the homozygous state in fertile boars. No haplotypes located on
462 chromosomes other than SSC12 were associated with the sperm tail defect.

463 The haplotype-based association study peaked at a 1.11 Mb segment of extended
464 homozygosity (between 3,473,632 and 4,587,759 bp) at porcine chromosome 12
465 that was identical by descent in all six affected boars, corroborating recessive
466 inheritance (**Figure 2C**). According to the Refseq annotation of the porcine genome,
467 the segment of extended autozygosity encompasses 15 genes (**Figure 2D**).

468 We suspected that the mutation causing the sperm tail defect affects a transcript that
469 is abundant in the testes. We found expression levels greater than 10 transcripts per
470 million for four genes (*DNAH17*, *BIRC5*, *SYNGR2*, *SEC14L1*) that were annotated
471 within the 1.11 Mb segment of extended homozygosity (**Figure 2D**). Among those,
472 only *DNAH17* encoding the dynein axonemal heavy chain 17, shows a highly testis-
473 biased expression in humans (https://gtexportal.org/home/), sheep
474 (http://biogps.org/dataset/BDS_00015/sheep-atlas/) (Clark et al., 2017) and cattle
475 (http://cattlegeneatlas.roslin.ed.ac.uk/) (L. Fang et al., 2020). Pathogenic alleles in
476 human and murine orthologs of *DNAH17* have been detected in males that produce
477 immotile sperm with multiple morphological abnormalities of the flagella (Whitfield et
478 al., 2019),(Zhang et al., 2020). Thus, we considered *DNAH17* as a compelling
479 positional and functional candidate gene for a sperm flagella defect in boars.

480 **An intronic 13-bp deletion is associated with the sperm tail defect**

481 We sequenced five affected boars to an average coverage of 10.9-fold using short
482 paired-end reads. We also considered 82 sequenced pigs from our in-house variant
483 database that were not affected by the sperm tail disorder. The analysis of

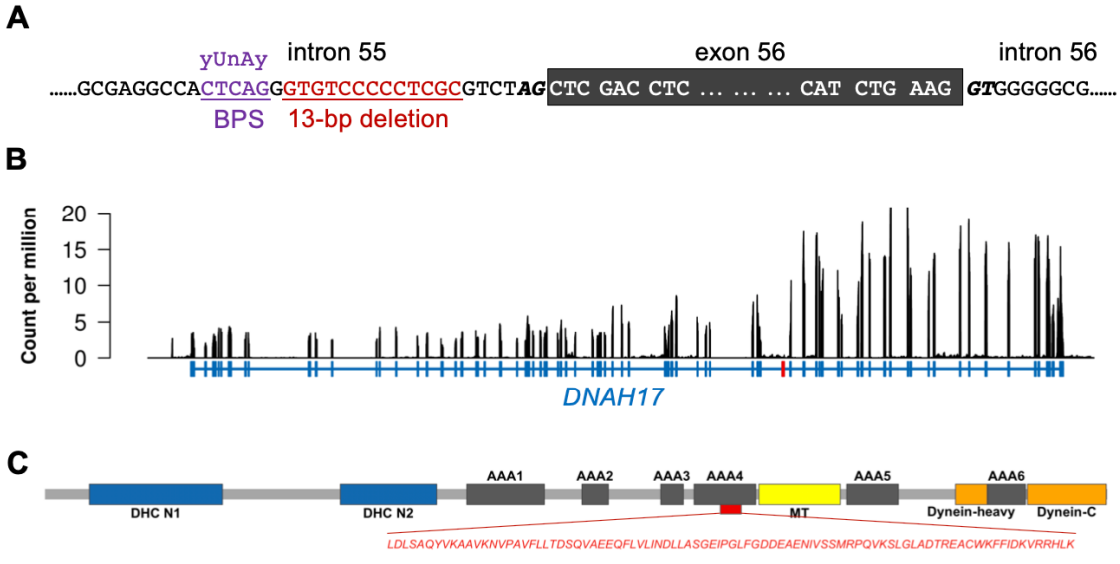
484 sequencing depth at the proximal region of SSC12 did not reveal large deletions or
485 duplications that segregate with the haplotype. To identify candidate causal
486 mutations, we considered 18,662 variants (14,806 SNPs and 3,856 Indels) that were
487 detected within the 1.11 Mb interval (between 3,473,632 and 4,587,759 bp) of
488 extended homozygosity. Of the 82 sequenced control animals, 65 also had array-
489 called genotypes. Three of the 65 pigs that had array- and sequence-called
490 genotypes carried the top haplotype in heterozygous state, yielding a haplotype
491 frequency of 2% in the sequenced control animals.

492 Of 18,662 positional candidate causal variants, a single variant was compatible with
493 recessive inheritance of the sperm tail defect: a 13-bp deletion at Chr12:3,556,401-
494 3,556,414 bp (XM_021066525.1:c.8510-17_8510-5del) located in an intron of
495 *DNAH17*. The 13-bp location resides within the top window from the haplotype-
496 based association study. Sequence variant genotyping using *GATK* indicated that
497 the five sequenced boars with the sperm tail defect were homozygous for the
498 deletion. Only the three boars from the control group, that carried the top haplotype
499 in the heterozygous state, also carried the 13-bp deletion in the heterozygous state.
500 The deletion did not occur in the homozygous state in animals from the control
501 group. According to the variant classification algorithm of the VEP software, the
502 intronic 13-bp deletion was annotated as "splice_region_variant & intron_variant". Its
503 impact on protein function was predicted to be low.

504 **RNA sequencing reveals that the 13-bp deletion causes skipping of DNAH17**
505 **exon 56**

506 Closer inspection of the 13-bp deletion revealed that it excises an intronic pyrimidine-
507 rich sequence containing a continuous stretch of eight pyrimidine bases between 5
508 and 17 bases upstream of a canonical 3' splice site of *DNAH17* exon 56 (**Figure**
509 **3A**). The VEP plugin SpliceRegion.pm indicated that the deletion coincides with a
510 putative polypyrimidine tract. A branch point consensus sequence (yUnAy) (Gao et
511 al., 2008) was predicted (z-score: 4.79) 2-6 nucleotides upstream the 13-bp deletion.
512 The pyrimidine content between the predicted branch point adenine and the splice
513 acceptor site at the 3' end of intron 55 is 68.4%. Considering that the polypyrimidine
514 tract is an important cis-acting element for spliceosome assembly in canonical "GT-
515 AG"-type acceptor splice sites (Coolidge et al., 1997), we suspected that the 13-bp
516 deletion perturbs pre-mRNA splicing. *In silico* splice site prediction using NNSPLICE

517 indicated (prediction score 0.46) that the 13-bp deletion likely prevents recognition of
518 the 3' splice acceptor site at the intron-exon boundary of exon 56, thus possibly
519 causing exon skipping.



520

521 **Figure 3: A 13-bp deletion perturbs splicing of *DNAH17***. Genomic sequence at the intron-exon
522 boundaries of exon 56 of porcine *DNAH17* (XM_021066525.1) (A). The grey background indicates
523 the sequence of exon 56. Red color highlights 13 nucleotides that are deleted in boars with the sperm
524 tail defect. Splice acceptor and donor sites are represented in bold-italic type. Violet color indicates
525 the predicted (z-score: 4.79) branch point sequence (BPS). Expression of DNAH17 (in counts per
526 million) in testes of two Swiss Large White boars (SAMEA6832284, SAMEA6832285) homozygous
527 for the 13-bp deletion (B). The number of reads covering a position was divided by the total number of
528 reads (in million) mapped to transcripts, and averaged across both boars. Blue color indicates the
529 exon-intron structure of porcine *DNAH17*. Exon 56 is displayed in red color. DNAH17 domains
530 according to Uniprot (protein-ID: A0A287AFU3) (C). Different colored boxes represent conserved
531 domains in the stem of dynein heavy chains (DHC), the microtubule-binding region (MT) and six
532 ATPase domains (ATPases associated with diverse cellular activities, AAA1-AAA6). Skipping of exon
533 56 truncates 89 amino acids (red color) from AAA4.
534

535 Exon 56 of porcine *DNAH17* is regularly expressed in testis tissue of pubertal boars
536 (**Supplemental file S5**). In order to examine if the loss of the polypyrimidine tract
537 upstream exon 56 perturbs splicing, we sequenced RNA purified from testis tissue of
538 two boars that produced defective sperm and were homozygous for the 13-bp
539 deletion. We aligned 49 and 70 million paired-end (2x150 nt) RNA sequencing reads
540 to the porcine transcriptome to quantify transcript and exon abundance. An average
541 of 49 *DNAH17* transcripts per million (TPM) were detected in the two boars

542 homozygous for the 13-bp deletion. However, no reads aligned to the sequence of
543 exon 56 corroborating that the 13-bp deletion prevents recognition of the 3' splice
544 site, thereby causing exon skipping (**Figure 3B**).

545 The skipping of exon 56 causes an in-frame loss of 89 amino acids (residues 2836-
546 2924) representing 2% of porcine DNAH17 (XP_020922184.1). DNAH17 contains
547 six evolutionarily conserved "ATPase associated with a variety of cellular activities
548 (AAA)"-domains that are linked together as an asymmetric hexameric ring required
549 to power the beating movement of sperm flagella (Gleave et al., 2014),(Snider et al.,
550 2008). The 89 amino acids are truncated from AAA4 of DNAH17 (**Figure 3C**).

551

552

553 **DISCUSSION**

554 Eight Swiss Large White boars produced sperm with multiple morphological
555 abnormalities of the flagella. Considering that all sperm were immotile, the boars
556 were judged to be infertile and not used for breeding. Because a majority of the
557 sperm were viable, fertilization might be possible using intracytoplasmic sperm
558 injection (ICSI) (Nijs et al., 1996),(Ortega et al., 2011). However, success rates after
559 ICSI may vary considerably in individuals with asthenoteratozoospermia
560 (Wambergue et al., 2016),(Whitfield et al., 2019). In any case, the application of ICSI
561 in livestock populations seems unwarranted if the underlying male-factor infertility
562 follows recessive inheritance, because a defective allele would be transmitted to the
563 offspring.

564 Apart from producing sperm with defective tails, the boars were healthy. Diseases
565 and environmental stressors were less likely to be causal for the defective sperm
566 flagella (Jurewicz et al., 2009), because the boars were maintained at a semen
567 collection center with many other boars that produced normal semen. Idiopathic
568 infertility and poor semen quality in healthy males may result from pathogenic alleles
569 in genes that are specifically expressed in the male reproductive tract (Pausch et al.,
570 2014),(Iso-Touru et al., 2019),(Pausch et al., 2016b),(Noskova et al., 2020). The
571 eight affected boars were inbred on a common ancestor supporting the hypothesis of
572 a shared genetic etiology. While the presence of common ancestors in the pedigrees
573 of all affected boars is compatible with recessive inheritance, it does not rule out

574 other modes of inheritance (Bourneuf et al., 2017). Considering that all affected but
575 only few fertile boars were inbred on the common ancestor, a recessive mode of
576 inheritance was likely.

577 The morphological and ultrastructural flagella defects of the eight Swiss Large White
578 boars are similar to those observed in Finnish Yorkshire boars that are homozygous
579 for a loss-of-function allele in *KPL2* (a.k.a. *SPEF2*) (Andersson et al., 2000),(Sironen
580 et al., 2006). Although cross-breeding between Finnish Yorkshire and Swiss Large
581 White pigs had not been documented recently, the defective allele could be located
582 on an ancient shared haplotype (Pausch et al., 2016a),(Schwarzenbacher et al.,
583 2016). However, the genomic region on chromosome 16 encompassing *SPEF2* is
584 not associated with the sperm defect of the Swiss Large White boars, indicating
585 genetic heterogeneity.

586 SNPs on chromosomes 11, 12 and 13 were significantly associated with the sperm
587 disorder using Fisher's exact test of allelic association. This result was puzzling
588 because an oligogenic inheritance of the sperm flagella defect was unlikely.
589 Phenotypic misclassification and genetic heterogeneity could lead to an inconclusive
590 association study (Manchia et al., 2013). However, the morphological abnormalities
591 of the sperm flagella were strikingly similar in the ejaculates of eight boars.
592 Moreover, pedigree analysis suggested monogenic recessive inheritance. Principal
593 components analysis and a genomic inflation factor of 1.55 indicated that population
594 stratification confounded our SNP-based association study (Price et al., 2010).
595 Although Fisher's exact tests have been widely applied to map binary traits (Balding,
596 2006), these tests are prone to type I errors when cases and controls differ in their
597 ancestries. Following the approach from (Hiltpold et al., 2020), we fitted the top
598 principal components of a genomic relationship matrix as covariates in a haplotype-
599 based association model to take population stratification into account. A strong
600 association signal at chromosome 12 remained while all other association signals
601 disappeared. In agreement with previous studies (Price et al., 2006),(Pausch et al.,
602 2011),(Hiltpold et al., 2020) and evidenced by a low inflation factor, the principal
603 components-based correction successfully eliminated spurious signals that arose
604 due to population stratification. Considering that the most significantly associated
605 haplotype also encompasses the causal variant for the sperm flagella defect, our
606 findings corroborate that genome-wide haplotype-based association testing offers a

607 powerful approach to identify trait-associated regions in stratified mapping cohorts
608 (Kadri et al., 2014),(Hiltpold et al., 2020),(Pausch et al., 2016a).

609 The haplotype associated with multiple morphological abnormalities of the sperm
610 flagella carries a 13-bp deletion in an intron of *DNAH17* encoding dynein axonemal
611 heavy chain 17. Our results show that DNAH17 is abundant in porcine testis tissue.
612 Dynein axonemal heavy chains are required for the axonemal assembly and the
613 beating movement of the flagella (Inaba, 2003),(Ben Khelifa et al., 2014),(Tu et al.,
614 2019),(Sironen et al., 2020),(Liu et al., 2020),(Whitfield et al., 2019),(Zhang et al.,
615 2020, p. 17),(Sha et al., 2020). Pathogenic alleles in human and murine *DNAH17*
616 manifest in sperm with an abnormal mitochondrial sheath and cytoplasmic droplets,
617 as well as axonemal disorganization including absence of the central pair and
618 missing peripheral microtubule doublets (Whitfield et al., 2019), (Zhang et al., 2020,
619 p. 17),(Sha et al., 2020). Boars that are homozygous for a 13-bp deletion in intron 55
620 of *DNAH17* also produce sperm with multiple morphological and ultrastructural
621 abnormalities of the flagella. Although annotated as a low impact variant, the intronic
622 13-bp deletion perturbs pre-mRNA splicing because it disrupts the polypyrimidine
623 tract in intron 55 of *DNAH17*. The mutant DNAH17 lacks part of the AAA4 domain
624 which likely compromises axonemal assembly and the beating movement of the
625 sperm flagellum in the homozygous state. The assembly of the flagellar axoneme is
626 less disorganized when other domains of DNAH17 are affected by pathogenic
627 variants (Whitfield et al., 2019),(Zhang et al., 2020),(Sha et al., 2020), likely because
628 an intact AAA4 domain is essential to physiological DNAH17 function (Snider et al.,
629 2008). Thus, our findings suggest that the intronic 13-bp deletion is a loss-of-function
630 allele.

631 We previously identified mutations at splice donor sites and nearby splice acceptor
632 sites that manifest in male reproductive disorders due to aberrant splicing (Iso-Touru
633 et al., 2019),(Hiltpold et al., 2020). Our present study reveals an intronic deletion
634 from 5 to 17 nucleotides upstream a canonical 3' splice site that perturbs splicing as
635 the most likely causal variant for a novel porcine sperm flagella defect. The 13-bp
636 deletion compromises splicing because it excises a polypyrimidine tract upstream
637 *DNAH17* exon 56. Phenotypic consequences arising from polypyrimidine tract
638 mutations had rarely been described so far (Lefévre et al., 2002),(Sartelet et al.,
639 2015),(Abramowicz and Gos, 2018). To the best of our knowledge, our study reports

640 the first phenotype-genotype association for a mutation affecting an intronic
641 polypyrimidine tract in pigs. We provide evidence that the loss of the polypyrimidine
642 tract in porcine *DNAH17* intron 55 prevents recognition of the 3' splice site which
643 leads to the skipping of exon 56, thus causing defective sperm flagella. Sequence
644 motifs that govern the assembly of the spliceosome may be more distant to 3' splice
645 sites than the polypyrimidine tract in porcine *DNAH17* intron 55 (Zhang et al., 2017).
646 However, standard sequence variant annotation tools are largely blind to putative
647 consequences arising from mutations within intronic motifs. Thus, the systematic
648 characterization of branch points and polypyrimidine tracts seems warranted to
649 refine the functional classification of intronic variants.

650 Our findings enable the monitoring of the 13-bp deletion and unambiguous
651 identification of carrier and homozygous animals using direct gene testing. While
652 homozygosity for the 13-bp deletion is easily recognized in artificial insemination
653 boars using microscopic semen analysis, it may remain undetected in sows and
654 natural service boars. However, affected natural service boars will be noticed after
655 few matings due to low fertility. In agreement with previous findings on loss-of-
656 function alleles in human and murine *DNAH17* (Whitfield et al., 2019),(Zhang et al.,
657 2020),(Sha et al., 2020), homozygous mutation carriers were healthy. Considering
658 that the frequency of the 13-bp deletion is low and the majority of sows is
659 inseminated artificially, economic losses due to unsuccessful breeding with
660 homozygous boars are negligible in the Swiss Large White population.

661 The number of recessive conditions detected in livestock populations increases
662 steadily (Nicholas and Hobbs, 2014). Low effective population size, founder effects
663 and inbreeding favor the manifestation of recessive conditions. However, recent
664 advances in genome-wide genotyping and sequencing offer powerful tools for rapid
665 detection of causal variants underpinning inherited conditions (Bourneuf et al., 2017).
666 The phenotypic and economic consequences of the 13-bp deletion in *DNAH17* are
667 less detrimental than mutations that manifest in malformed and non-viable piglets
668 (Fang et al., 2020). However, the 13-bp deletion becomes a threat to the breeding
669 population if it attains higher frequency (Pausch et al., 2014),(Hiltpold et al., 2020).
670 The 13-bp deletion in *DNAH17* adds to a catalogue of sequence variants that cause
671 male-factor infertility in pigs (Sironen et al., 2011),(Sironen et al., 2006),(Noskova et
672 al., 2020). It remains an open question how an increasing number of undesired

673 recessive alleles may be considered appropriately in livestock populations (Cole,
674 2015),(Upperman et al., 2019). Although the newly identified 13-bp deletion does
675 neither compromise the welfare of homozygous animals nor result in huge economic
676 losses, it's frequency should be kept at a low level to avoid the birth of homozygous
677 boars that are infertile.

678

679

680 **ACKNOWLEDGEMENTS**

681 We acknowledge financial support from SUISAG, Micarna SA and the ETH Zürich
682 Foundation. We thank Dr. Cecilia Bebeacua (ScopeM) for support in transmission
683 electron microscopy. We are thankful for the excellent technical support provided by
684 the ETH Zürich technology platforms FGCZ (https://fgcz.ch/), ScopeM
685 (https://scopem.ethz.ch/), and Genetic Diversity Centre (https://gdc.ethz.ch/),
686 respectively, for sequencing, microscopy and DNA fragment analysis.

687

688

689 **REFERENCES**

690 Abramowicz, A., Gos, M., 2018. Splicing mutations in human genetic disorders: examples,
691 detection, and confirmation. J Appl Genet 59, 253–268.
692 https://doi.org/10.1007/s13353-018-0444-7
693 Andersson, M., Peltoniemi, O., Makinen, A., Sukura, A., Rodriguez-Martinez, H., 2000. The
694 Hereditary 'Short Tail' Sperm Defect - A New Reproductive Problem in Yorkshire
695 Boars. Reproduction in Domestic Animals 35, 59–63. https://doi.org/10.1046/j.1439-
696 0531.2000.00196.x
697 Aulchenko, Y.S., Ripke, S., Isaacs, A., van Duijn, C.M., 2007. GenABEL: an R library for
698 genome-wide association analysis. Bioinformatics 23, 1294–1296.
699 https://doi.org/10.1093/bioinformatics/btm108
700 Balding, D.J., 2006. A tutorial on statistical methods for population association studies.
701 Nature Reviews Genetics 7, 781–791. https://doi.org/10.1038/nrg1916
702 Ben Khelifa, M., Coutton, C., Zouari, R., Karaouzène, T., Rendu, J., Bidart, M., Yassine, S.,
703 Pierre, V., Delaroche, J., Hennebicq, S., Grunwald, D., Escalier, D., Pernet-Gallay, K.,
704 Jouk, P.-S., Thierry-Mieg, N., Touré, A., Arnoult, C., Ray, P.F., 2014. Mutations in
705 DNAH1, which Encodes an Inner Arm Heavy Chain Dynein, Lead to Male Infertility
706 from Multiple Morphological Abnormalities of the Sperm Flagella. American Journal
707 of Human Genetics 94, 95. https://doi.org/10.1016/j.ajhg.2013.11.017
708 Blom, E., 1950. A one-minute live-dead sperm stain by means of eosin-nigrosin. Fertility and
709 Sterility 1, 176–177.

710 Bourneuf, E., Otz, P., Pausch, H., Jagannathan, V., Michot, P., Grohs, C., Piton, G.,
711 Ammermüller, S., Deloche, M.-C., Fritz, S., Leclerc, H., Péchoux, C., Boukadiri, A.,
712 Hozé, C., Saintilan, R., Créchet, F., Mosca, M., Segelke, D., Guillaume, F., Bouet, S.,
713 Baur, A., Vasilescu, A., Genestout, L., Thomas, A., Allais-Bonnet, A., Rocha, D., Colle,
714 M.-A., Klopp, C., Esquerré, D., Wurmser, C., Flisikowski, K., Schwarzenbacher, H.,
715 Burgstaller, J., Brügmann, M., Dietschi, E., Rudolph, N., Freick, M., Barbey, S., Fayolle,
716 G., Danchin-Burge, C., Schibler, L., Bed'Hom, B., Hayes, B.J., Daetwyler, H.D., Fries, R.,
717 Boichard, D., Pin, D., Drögemüller, C., Capitan, A., 2017. Rapid Discovery of De Novo
718 Deleterious Mutations in Cattle Enhances the Value of Livestock as Model Species.
719 Scientific Reports 7, 11466. https://doi.org/10.1038/s41598-017-11523-3
720 Bray, N.L., Pimentel, H., Melsted, P., Pachter, L., 2016. Near-optimal probabilistic RNA-seq
721 quantification. Nat Biotechnol 34, 525–527. https://doi.org/10.1038/nbt.3519
722 Broekhuijse, M.L.W.J., Feitsma, H., Gadella, B.M., 2011. Field data analysis of boar semen
723 quality. Reprod. Domest. Anim. 46 Suppl 2, 59–63. https://doi.org/10.1111/j.1439-
724 0531.2011.01861.x
725 Browning, B.L., Zhou, Y., Browning, S.R., 2018. A One-Penny Imputed Genome from Next-
726 Generation Reference Panels. American Journal of Human Genetics 103, 338–348.
727 https://doi.org/10.1016/j.ajhg.2018.07.015
728 Chang, C.C., Chow, C.C., Tellier, L.C., Vattikuti, S., Purcell, S.M., Lee, J.J., 2015. Second-
729 generation PLINK: rising to the challenge of larger and richer datasets. Gigascience 4,
730 7. https://doi.org/10.1186/s13742-015-0047-8
731 Chen, S., Zhou, Y., Chen, Y., Gu, J., 2018. Fastp: An ultra-fast all-in-one FASTQ preprocessor.
732 Bioinformatics 34, i884–i890. https://doi.org/10.1093/bioinformatics/bty560
733 Clark, E.L., Bush, S.J., McCulloch, M.E.B., Farquhar, I.L., Young, R., Lefevre, L., Pridans, C.,
734 Tsang, H.G., Wu, C., Afrasiabi, C., Watson, M., Whitelaw, C.B., Freeman, T.C.,
735 Summers, K.M., Archibald, A.L., Hume, D.A., 2017. A high resolution atlas of gene
736 expression in the domestic sheep (Ovis aries). PLOS Genetics 13, e1006997.
737 https://doi.org/10.1371/journal.pgen.1006997
738 Cole, J.B., 2015. A simple strategy for managing many recessive disorders in a dairy cattle
739 breeding program. Genet Sel Evol 47. https://doi.org/10.1186/s12711-015-0174-9
740 Cole, J.B., 2007. PyPedal: A computer program for pedigree analysis. Computers and
741 electronics in agriculture 57:107-113.
742 Colenbrander, B., Feitsma, H., Grooten, H.J., 1993. Optimizing semen production for
743 artificial insemination in swine. J. Reprod. Fertil. Suppl. 48, 207–215.
744 Coolidge, C.J., Seely, R.J., Patton, J.G., 1997. Functional analysis of the polypyrimidine tract
745 in pre-mRNA splicing. Nucleic Acids Res 25, 888–896.
746 Crysnanto, D., Wurmser, C., Pausch, H., 2019. Accurate sequence variant genotyping in
747 cattle using variation-aware genome graphs. Genetics Selection Evolution 51, 21.
748 https://doi.org/10.1186/s12711-019-0462-x
749 Depristo, M.A., Banks, E., Poplin, R., Garimella, K. V., Maguire, J.R., Hartl, C., Philippakis,
750 A.A., Del Angel, G., Rivas, M.A., Hanna, M., McKenna, A., Fennell, T.J., Kernytsky,
751 A.M., Sivachenko, A.Y., Cibulskis, K., Gabriel, S.B., Altshuler, D., Daly, M.J., 2011. A
752 framework for variation discovery and genotyping using next-generation DNA
753 sequencing data. Nature Genetics 43, 491–501. https://doi.org/10.1038/ng.806
754 Diniz, D.B., Lopes, M.S., Broekhuijse, M.L.W.J., Lopes, P.S., Harlizius, B., Guimarães, S.E.F.,
755 Duijvesteijn, N., Knol, E.F., Silva, F.F., 2014. A genome-wide association study reveals

a novel candidate gene for sperm motility in pigs. Animal Reproduction Science 151, 201–207. https://doi.org/10.1016/j.anireprosci.2014.10.014

Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., Gingeras, T.R., 2013. STAR: ultrafast universal RNA-seq aligner. Bioinformatics 29, 15–21. https://doi.org/10.1093/bioinformatics/bts635

Fang, L., Cai, W., Liu, S., Canela-Xandri, O., Gao, Y., Jiang, J., Rawlik, K., Li, B., Schroeder, S.G., Rosen, B.D., Li, C., Sonstegard, T.S., Alexander, L.J., Tassell, C.P.V., VanRaden, P.M., Cole, J.B., Yu, Y., Zhang, S., Tenesa, A., Ma, L., Liu, G.E., 2020. Comprehensive analyses of 723 transcriptomes enhance genetic and biological interpretations for complex traits in cattle. Genome Res 30,790-801. https://doi.org/10.1101/gr.250704.119

Fang, Z.-H., Nosková, A., Crysnanto, D., Neuenschwander, S., Vögeli, P., Pausch, H., 2020. A 63-bp insertion in exon 2 of the porcine KIF21A gene is associated with arthrogryposis multiplex congenita. Animal Genetics 51, 820–823. https://doi.org/10.1111/age.12984

Gao, K., Masuda, A., Matsuura, T., Ohno, K., 2008. Human branch point consensus sequence is yUnAy. Nucleic Acids Res 36, 2257–2267. https://doi.org/10.1093/nar/gkn073

Gleave, E.S., Schmidt, H., Carter, A.P., 2014. A structural analysis of the AAA+ domains in Saccharomyces cerevisiae cytoplasmic dynein. J Struct Biol 186, 367–375. https://doi.org/10.1016/j.jsb.2014.03.019

Grahofer, A., Gurtner, C., Nathues, H., 2017. Haemorrhagic bowel syndrome in fattening pigs. Porcine Health Manag 3. https://doi.org/10.1186/s40813-017-0074-1

Hancock, J.L., 1957. The morphology of boar spermatozoa. J R Microsc Soc 76, 84–97. https://doi.org/10.1111/j.1365-2818.1956.tb00443.x

Hiltpold, M., Niu, G., Kadri, N.K., Crysnanto, D., Fang, Z.-H., Spengeler, M., Schmitz-Hsu, F., Fuerst, C., Schwarzenbacher, H., Seefried, F.R., Seehusen, F., Witschi, U., Schnieke, A., Fries, R., Bollwein, H., Flisikowski, K., Pausch, H., 2020. Activation of cryptic splicing in bovine WDR19 is associated with reduced semen quality and male fertility. PLOS Genetics 16, e1008804. https://doi.org/10.1371/journal.pgen.1008804

Holt, C., Holt, W.V., Moore, H.D., Reed, H.C., Curnock, R.M., 1997. Objectively measured boar sperm motility parameters correlate with the outcomes of on-farm inseminations: results of two fertility trials. J. Androl. 18, 312–323.

Inaba, K., 2003. Molecular Architecture of the Sperm Flagella: Molecules for Motility and Signaling. jzoo 20, 1043–1056. https://doi.org/10.2108/zsj.20.1043

Iso-Touru, T., Wurmser, C., Venhoranta, H., Hiltpold, M., Savolainen, T., Sironen, A., Fischer, K., Flisikowski, K., Fries, R., Vicente-Carrillo, A., Alvarez-Rodriguez, M., Nagy, S., Mutikainen, M., Peippo, J., Taponen, J., Sahana, G., Guldbrandtsen, B., Simonen, H., Rodriguez-Martinez, H., Andersson, M., Pausch, H., 2019. A splice donor variant in CCDC189 is associated with asthenospermia in Nordic Red dairy cattle. BMC Genomics 20, 286. https://doi.org/10.1186/s12864-019-5628-y

Jurewicz, J., Hanke, W., Radwan, M., Bonde, J.P., 2009. Environmental factors and semen quality. Int J Occup Med Environ Health 22, 305–329. https://doi.org/10.2478/v10001-009-0036-1

Kadri, N.K., Sahana, G., Charlier, C., Iso-Touru, T., Guldbrandtsen, B., Karim, L., Nielsen, U.S., Panitz, F., Aamand, G.P., Schulman, N., Georges, M., Vilkki, J., Lund, M.S., Druet, T., 2014. A 660-Kb Deletion with Antagonistic Effects on Fertility and Milk Production Segregates at High Frequency in Nordic Red Cattle: Additional Evidence for the

803 Common Occurrence of Balancing Selection in Livestock. PLOS Genetics 10,
804 e1004049. https://doi.org/10.1371/journal.pgen.1004049
805 Knox, R.V., 2016. Artificial insemination in pigs today. Theriogenology 85, 83–93.
806 https://doi.org/10.1016/j.theriogenology.2015.07.009
807 Kosova, G., Scott, N.M., Niederberger, C., Prins, G.S., Ober, C., 2012. Genome-wide
808 Association Study Identifies Candidate Genes for Male Fertility Traits in Humans. The
809 American Journal of Human Genetics 90, 950–961.
810 https://doi.org/10.1016/j.ajhg.2012.04.016
811 Lamas-Toranzo, I., Hamze, J.G., Bianchi, E., Fernández-Fuertes, B., Pérez-Cerezales, S.,
812 Laguna-Barraza, R., Fernández-González, R., Lonergan, P., Gutiérrez-Adán, A., Wright,
813 G.J., Jiménez-Movilla, M., Bermejo-Álvarez, P., 2020. TMEM95 is a sperm membrane
814 protein essential for mammalian fertilization. eLife 9, e53913.
815 https://doi.org/10.7554/eLife.53913
816 Lefévre, S.H., Chauveinc, L., Stoppa-Lyonnet, D., Michon, J., Lumbroso, L., Berthet, P.,
817 Frappaz, D., Dutrillaux, B., Chevillard, S., Malfoy, B., 2002. A T to C mutation in the
818 polypyrimidine tract of the exon 9 splicing site of the RB1 gene responsible for low
819 penetrance hereditary retinoblastoma. Journal of Medical Genetics 39, E21.
820 https://doi.org/10.1136/jmg.39.5.e21
821 Li, H., 2013. Aligning sequence reads, clone sequences and assembly contigs with BWA-
822 MEM, arXiv:13033997.
823 Liu, C., Miyata, H., Gao, Y., Sha, Y., Tang, S., Xu, Z., Whitfield, M., Patrat, C., Wu, H., Dulioust,
824 E., Tian, S., Shimada, K., Cong, J., Noda, T., Li, H., Morohoshi, A., Cazin, C., Kherraf, Z.-
825 E., Arnoult, C., Jin, L., He, X., Ray, P.F., Cao, Y., Touré, A., Zhang, F., Ikawa, M., 2020.
826 Bi-allelic DNAH8 Variants Lead to Multiple Morphological Abnormalities of the
827 Sperm Flagella and Primary Male Infertility. The American Journal of Human
828 Genetics 107, 330–341. https://doi.org/10.1016/j.ajhg.2020.06.004
829 Manchia, M., Cullis, J., Turecki, G., Rouleau, G.A., Uher, R., Alda, M., 2013. The Impact of
830 Phenotypic and Genetic Heterogeneity on Results of Genome Wide Association
831 Studies of Complex Diseases. PLoS One 8.
832 https://doi.org/10.1371/journal.pone.0076295
833 Marques, D.B.D., Lopes, M.S., Broekhuijse, M.L.W.J., Guimarães, S.E.F., Knol, E.F.,
834 Bastiaansen, J.W.M., Silva, F.F., Lopes, P.S., 2017. Genetic parameters for semen
835 quality and quantity traits in five pig lines. J. Anim. Sci. 95, 4251–4259.
836 https://doi.org/10.2527/jas2017.1683
837 McLaren, W., Gil, L., Hunt, S.E., Riat, H.S., Ritchie, G.R.S., Thormann, A., Flicek, P.,
838 Cunningham, F., 2016. The Ensembl Variant Effect Predictor. Genome Biology
839 17:122. https://doi.org/10.1186/s13059-016-0974-4
840 Ni, X., Wang, J., Lv, M., Liu, C., Zhong, Y., Tian, S., Wu, H., Cheng, H., Gao, Y., Tan, Q., Chen,
841 B., Li, Q., Song, B., Wei, Z., Zhou, P., He, X., Zhang, F., Cao, Y., 2020. A novel
842 homozygous mutation in WDR19 induces disorganization of microtubules in sperm
843 flagella and nonsyndromic asthenoteratospermia. J Assist Reprod Genet 37, 1431–
844 1439. https://doi.org/10.1007/s10815-020-01770-1
845 Nicholas, F.W., Hobbs, M., 2014. Mutation discovery for Mendelian traits in non-laboratory
846 animals: a review of achievements up to 2012. Anim. Genet. 45, 157–170.
847 https://doi.org/10.1111/age.12103
848 Nijs, M., Vanderzwalmen, P., Vandamme, B., Segal-Bertin, G., Lejeune, B., Segal, L., van
849 Roosendaal, E., Schoysman, R., 1996. Fertilizing ability of immotile spermatozoa after

850 intracytoplasmic sperm injection. Hum. Reprod. 11, 2180–2185.
851 https://doi.org/10.1093/oxfordjournals.humrep.a019073
852 Noda, T., Lu, Y., Fujihara, Y., Oura, S., Koyano, T., Kobayashi, S., Matzuk, M.M., Ikawa, M.,
853 2020. Sperm proteins SOF1, TMEM95, and SPACA6 are required for sperm−oocyte
854 fusion in mice. PNAS 117, 11493–11502. https://doi.org/10.1073/pnas.1922650117
855 Noskova, A., Wurmser, C., Crysnanto, D., Sironen, A., Uimari, P., Fries, R., Andersson, M.,
856 Pausch, H., 2020. Deletion of porcine BOLL is associated with defective acrosomes
857 and subfertility in Yorkshire boars. Animal Genetics, Advanced online,
858 https://doi.org/10.1111/age.12998
859 Ortega, C., Verheyen, G., Raick, D., Camus, M., Devroey, P., Tournaye, H., 2011. Absolute
860 asthenozoospermia and ICSI: what are the options? Hum Reprod Update 17, 684–
861 692. https://doi.org/10.1093/humupd/dmr018
862 Pausch, H., Ammermüller, S., Wurmser, C., Hamann, H., Tetens, J., Drögemüller, C., Fries, R.,
863 2016a. A nonsense mutation in the COL7A1 gene causes epidermolysis bullosa in
864 Vorderwald cattle. BMC Genet. 17, 149. https://doi.org/10.1186/s12863-016-0458-2
865 Pausch, H., Flisikowski, K., Jung, S., Emmerling, R., Edel, C., Götz, K.-U., Fries, R., 2011.
866 Genome-wide association study identifies two major loci affecting calving ease and
867 growth-related traits in cattle. Genetics 187, 289–297.
868 https://doi.org/10.1534/genetics.110.124057
869 Pausch, H., Kölle, S., Wurmser, C., Schwarzenbacher, H., Emmerling, R., Jansen, S.,
870 Trottmann, M., Fuerst, C., Götz, K.-U., Fries, R., 2014. A Nonsense Mutation in
871 TMEM95 Encoding a Nondescript Transmembrane Protein Causes Idiopathic Male
872 Subfertility in Cattle. PLOS Genetics 10, e1004044.
873 https://doi.org/10.1371/journal.pgen.1004044
874 Pausch, H., Venhoranta, H., Wurmser, C., Hakala, K., Iso-Touru, T., Sironen, A., Vingborg,
875 R.K., Lohi, H., Söderquist, L., Fries, R., Andersson, M., 2016b. A frameshift mutation
876 in ARMC3 is associated with a tail stump sperm defect in Swedish Red (Bos taurus)
877 cattle. BMC Genet. 17, 49. https://doi.org/10.1186/s12863-016-0356-7
878 Pedersen, B.S., Quinlan, A.R., 2018. Mosdepth: Quick coverage calculation for genomes and
879 exomes. Bioinformatics 34, 867–868. https://doi.org/10.1093/bioinformatics/btx699
880 Price, A.L., Patterson, N.J., Plenge, R.M., Weinblatt, M.E., Shadick, N.A., Reich, D., 2006.
881 Principal components analysis corrects for stratification in genome-wide association
882 studies. Nature Genetics 38, 904–909. https://doi.org/10.1038/ng1847
883 Price, A.L., Zaitlen, N.A., Reich, D., Patterson, N., 2010. New approaches to population
884 stratification in genome-wide association studies. Nature reviews. Genetics 11, 459.
885 https://doi.org/10.1038/nrg2813
886 Rahban, R., Priskorn, L., Senn, A., Stettler, E., Galli, F., Vargas, J., Van den Bergh, M., Fusconi,
887 A., Garlantezec, R., Jensen, T.K., Multigner, L., Skakkebaek, N.E., Germond, M.,
888 Jørgensen, N., Nef, S., NICER Working Group, 2019. Semen quality of young men in
889 Switzerland: a nationwide cross-sectional population-based study. Andrology 7, 818–
890 826. https://doi.org/10.1111/andr.12645
891 Reese, M.G., Eeckman, F.H., Kulp, D., Haussler, D., 1997. Improved splice site detection in
892 Genie. J. Comput. Biol. 4, 311–323. https://doi.org/10.1089/cmb.1997.4.311
893 Robic, A., Faraut, T., Djebali, S., Weikard, R., Feve, K., Maman, S., Kuehn, C., 2019. Analysis
894 of pig transcriptomes suggests a global regulation mechanism enabling temporary
895 bursts of circular RNAs. RNA Biology 16, 1190–1204.
896 https://doi.org/10.1080/15476286.2019.1621621

897 Sartelet, A., Li, W., Pailhoux, E., Richard, C., Tamma, N., Karim, L., Fasquelle, C., Druet, T.,
898 Coppieters, W., Georges, M., Charlier, C., 2015. Genome-wide next-generation DNA
899 and RNA sequencing reveals a mutation that perturbs splicing of the
900 phosphatidylinositol glycan anchor biosynthesis class H gene (PIGH) and causes
901 arthrogryposis in Belgian Blue cattle. BMC Genomics 16, 316.
902 https://doi.org/10.1186/s12864-015-1528-y
903 Sato, Y., Tajima, A., Kiguchi, M., Kogusuri, S., Fujii, A., Sato, T., Nozawa, S., Yoshiike, M.,
904 Mieno, M., Kojo, K., Uchida, M., Tsuchiya, H., Yamasaki, K., Imoto, I., Iwamoto, T.,
905 2020. Genome-wide association study of semen volume, sperm concentration, testis
906 size, and plasma inhibin B levels. Journal of Human Genetics 65, 683–691.
907 https://doi.org/10.1038/s10038-020-0757-3
908 Sato, Y., Tajima, A., Sato, T., Nozawa, S., Yoshiike, M., Imoto, I., Yamauchi, A., Iwamoto, T.,
909 2018. Genome-wide association study identifies ERBB4 on 2q34 as a novel locus
910 associated with sperm motility in Japanese men. J Med Genet 55, 1–7.
911 https://doi.org/10.1136/jmedgenet-2017-104991
912 Schwarzenbacher, H., Burgstaller, J., Seefried, F.R., Wurmser, C., Hilbe, M., Jung, S., Fuerst,
913 C., Dinhopl, N., Weissenböck, H., Fuerst-Waltl, B., Dolezal, M., Winkler, R., Grueter,
914 O., Bleul, U., Wittek, T., Fries, R., Pausch, H., 2016. A missense mutation in TUBD1 is
915 associated with high juvenile mortality in Braunvieh and Fleckvieh cattle. BMC
916 Genomics 17, 400. https://doi.org/10.1186/s12864-016-2742-y
917 Sha, Y., Wei, X., Ding, L., Mei, L., Huang, X., Lin, S., Su, Z., Kong, L., Zhang, Y., Ji, Z., 2020.
918 DNAH17 is associated with asthenozoospermia and multiple morphological
919 abnormalities of sperm flagella. Annals of Human Genetics 84, 271–279.
920 https://doi.org/10.1111/ahg.12369
921 Sironen, A., Shoemark, A., Patel, M., Loebinger, M.R., Mitchison, H.M., 2020. Sperm defects
922 in primary ciliary dyskinesia and related causes of male infertility. Cell. Mol. Life Sci.
923 77, 2029–2048. https://doi.org/10.1007/s00018-019-03389-7
924 Sironen, A., Thomsen, B., Andersson, M., Ahola, V., Vilkki, J., 2006. An intronic insertion in
925 KPL2 results in aberrant splicing and causes the immotile short-tail sperm defect in
926 the pig. PNAS 103, 5006–5011. https://doi.org/10.1073/pnas.0506318103
927 Sironen, A., Uimari, P., Venhoranta, H., Andersson, M., Vilkki, J., 2011. An exonic insertion
928 within Tex14 gene causes spermatogenic arrest in pigs. BMC Genomics 12, 591.
929 https://doi.org/10.1186/1471-2164-12-591
930 Snider, J., Thibault, G., Houry, W.A., 2008. The AAA+ superfamily of functionally diverse
931 proteins. Genome Biol 9, 216. https://doi.org/10.1186/gb-2008-9-4-216
932 Soneson, C., Love, M.I., Robinson, M.D., 2015. Differential analyses for RNA-seq: transcript-
933 level estimates improve gene-level inferences. F1000Res 4, 1521.
934 https://doi.org/10.12688/f1000research.7563.2
935 Tarasov, A., Vilella, A.J., Cuppen, E., Nijman, I.J., Prins, P., 2015. Sambamba: Fast processing
936 of NGS alignment formats. Bioinformatics 31, 2032–2034.
937 https://doi.org/10.1093/bioinformatics/btv098
938 Thibier, M., Wagner, H.-G., 2002. World statistics for artificial insemination in cattle.
939 Livestock Production Science 74, 203–212. https://doi.org/10.1016/S0301-
940 6226(01)00291-3
941 Tu, C., Nie, H., Meng, L., Yuan, S., He, W., Luo, A., Li, H., Li, W., Du, J., Lu, G., Lin, G., Tan, Y.-
942 Q., 2019. Identification of DNAH6 mutations in infertile men with multiple

943 morphological abnormalities of the sperm flagella. Scientific Reports 9, 15864.
944 https://doi.org/10.1038/s41598-019-52436-7
945 Upperman, L.R., Kinghorn, B.P., MacNeil, M.D., Van Eenennaam, A.L., 2019. Management of
946 lethal recessive alleles in beef cattle through the use of mate selection software.
947 Genetics Selection Evolution 51, 36. https://doi.org/10.1186/s12711-019-0477-3
948 Venhoranta, H., Pausch, H., Flisikowski, K., Wurmser, C., Taponen, J., Rautala, H., Kind, A.,
949 Schnieke, A., Fries, R., Lohi, H., Andersson, M., 2014. In frame exon skipping in
950 UBE3B is associated with developmental disorders and increased mortality in cattle.
951 BMC Genomics 15:890, 1–9. https://doi.org/10.1186/1471-2164-15-890
952 Wambergue, C., Zouari, R., Fourati Ben Mustapha, S., Martinez, G., Devillard, F., Hennebicq,
953 S., Satre, V., Brouillet, S., Halouani, L., Marrakchi, O., Makni, M., Latrous, H., Kharouf,
954 M., Amblard, F., Arnoult, C., Ray, P.F., Coutton, C., 2016. Patients with multiple
955 morphological abnormalities of the sperm flagella due to DNAH1 mutations have a
956 good prognosis following intracytoplasmic sperm injection. Hum Reprod 31, 1164–
957 1172. https://doi.org/10.1093/humrep/dew083
958 Warr, A., Affara, N., Aken, B., Beiki, H., Bickhart, D.M., Billis, K., Chow, W., Eory, L., Finlayson,
959 H.A., Flicek, P., Girón, C.G., Griffin, D.K., Hall, R., Hannum, G., Hourlier, T., Howe, K.,
960 Hume, D.A., Izuogu, O., Kim, K., Koren, S., Liu, H., Manchanda, N., Martin, F.J.,
961 Nonneman, D.J., O'Connor, R.E., Phillippy, A.M., Rohrer, G.A., Rosen, B.D., Rund, L.A.,
962 Sargent, C.A., Schook, L.B., Schroeder, S.G., Schwartz, A.S., Skinner, B.M., Talbot, R.,
963 Tseng, E., Tuggle, C.K., Watson, M., Smith, T.P.L., Archibald, A.L., 2019. An improved
964 pig reference genome sequence to enable pig genetics and genomics research.
965 bioRxiv 668921. https://doi.org/10.1101/668921
966 Whitfield, M., Thomas, L., Bequignon, E., Schmitt, A., Stouvenel, L., Montantin, G., Tissier, S.,
967 Duquesnoy, P., Copin, B., Chantot, S., Dastot, F., Faucon, C., Barbotin, A.L., Loyens,
968 A., Siffroi, J.-P., Papon, J.-F., Escudier, E., Amselem, S., Mitchell, V., Touré, A.,
969 Legendre, M., 2019. Mutations in DNAH17, Encoding a Sperm-Specific Axonemal
970 Outer Dynein Arm Heavy Chain, Cause Isolated Male Infertility Due to
971 Asthenozoospermia. The American Journal of Human Genetics 105, 198–212.
972 https://doi.org/10.1016/j.ajhg.2019.04.015
973 Xavier, M.J., Salas-Huetos, A., Oud, M.S., Aston, K.I., Veltman, J.A., 2020. Disease gene
974 discovery in male infertility: past, present and future. Hum Genet.
975 https://doi.org/10.1007/s00439-020-02202-x
976 Zhang, B., Ma, H., Khan, T., Ma, A., Li, T., Zhang, H., Gao, J., Zhou, J., Li, Y., Yu, C., Bao, J., Ali,
977 A., Murtaza, G., Yin, H., Gao, Q., Jiang, X., Zhang, F., Liu, C., Khan, I., Zubair, M.,
978 Hussain, H.M.J., Khan, R., Yousaf, A., Yuan, L., Lu, Y., Xu, X., Wang, Y., Tao, Q., Hao,
979 Q., Fang, H., Cheng, H., Zhang, Y., Shi, Q., 2020. A DNAH17 missense variant causes
980 flagella destabilization and asthenozoospermia. J Exp Med 217.
981 https://doi.org/10.1084/jem.20182365
982 Zhang, Q., Fan, X., Wang, Y., Sun, M., Shao, J., Guo, D., 2017. BPP: a sequence-based
983 algorithm for branch point prediction. Bioinformatics 33, 3166–3172.
984 https://doi.org/10.1093/bioinformatics/btx401
985

986

987

SUPPLEMENTAL FILES

Supplemental file 1

Title: Accession numbers of 87 boars

Format: xlsx

Description: Accession numbers of the sequence data of five affected (asthenozoospermic) and 82 unaffected (control) boars. The third column indicates the haplotype status (0 – not carrier, 1 – heterozygous carrier, 2 – homozygous carrier) of 65 sequenced boars that also had Illumina PorcineSNP60 BeadChip-called genotypes.

Supplemental file 2

Format: jpg

Title: Sperm viability assessment by eosin-nigrosin staining

Description: Representative light microscopic image of sperm from an affected boar stained with eosin-nigrosin. The sperm were flushed post mortem from the epididymis of Boar_1249. Live sperm appear white in color. Dead sperm appear pink in color because they take up eosin due to loss of membrane integrity.

Supplemental file 3

Title: Disorganized flagellar axonemes in two affected boars

Format: png

Description: Cross-sections of sperm from two affected boars reveal multiple morphological and ultrastructural abnormalities of the flagella including the absence of outer or inner doublets, secondary flagella-like structures within the same membrane, and flagellar components in the cytoplasmic droplet-like structure.

Supplemental file 4

Title: Principal components analysis of cases and controls

Format: png

Description: Scatter plot of the top five eigenvectors of a genomic relationship matrix. Grey and red color indicates 100 fertile and six asthenoteratozoospermic boars from the Swiss Large White population. The raw data underlying the plot (genotypes, haplotypes, eigenvectors) are available at available at https://doi.org/10.5281/zenodo.4081475.

1022

1023 **Supplemental file 5**

1024 Title: Expression profile of porcine *DNAH17* in pubertal boars

1025 Format: png

1026 Description: Expression of *DNAH17* quantified using testis RNAseq alignments of
1027 four pubertal boars (SRR2564762, SRR8237161, SRR8237162, SRR8237165). The
1028 number of reads covering a position was extracted from coordinate sorted BAM files
1029 using the MOSDEPTH software, divided by the total number of reads mapped to
1030 transcripts. Blue color indicates the exon-intron structure of porcine *DNAH17*. Exon
1031 56 is displayed in red color. Although exon expression is somewhat variable between
1032 the samples, the expression of exon 56 is evident in all samples. Raw data
1033 underlying the plot are available at the European Nucleotide Archive (ENA) of the
1034 EMBL at the BioProject PRJNA506525.