Molecular characterization and exclusion of porcine GUSB as a candidate gene for congenital hernia inguinalis/scrotalis

Julia Beck†1, Kirsten Bornemann-Kolatzki†1, Christoph Knorr1, Helge Taeubert2 and Bertram Brenig*1

Address: 1Institute of Veterinary Medicine, Georg-August-University Goettingen, Burckhardtweg 2, 37077 Goettingen, Germany and 2Institute of Animal and Genetics, Georg-August-University Goettingen, Albrecht-Thaer-Weg 3, 37075 Goettingen, Germany

Email: Julia Beck - jbeck1@gwdg.de; Kirsten Bornemann-Kolatzki - kbornem@gmx.de; Christoph Knorr - cknorr@gwdg.de; Helge Taeubert - htaeube@gwdg.de; Bertram Brenig* - bbrenig@gwdg.de

* Corresponding author †Equal contributors

Abstract

Background: Inguinal hernias are usually caused by a congenital defect, which occurs as a weakness of the inguinal canal. Porcine β-glucuronidase gene (GUSB) was chosen as functional candidate gene because of its involvement in degradation of hyaluronan within gubernacular tissue during descent of testes. Since a genome-wide linkage analysis approach has shown evidence that two regions on porcine chromosome 3 (SSC 3) are involved in the inheritance of hernia inguinalis/scrotalis in German pig breeds, GUSB also attained status as a positional candidate gene by its localization within a hernia-associated chromosomal region.

Results: A contig spanning 17,157 bp, which contains the entire GUSB, was assembled. Comparative sequence analyses were conducted for the GUSB gene locus. Single nucleotide polymorphisms (SNPs) located within the coding region of GUSB were genotyped in 512 animals. Results of transmission disequilibrium test (TDT) for two out of a total of five detected SNPs gave no significant association with the outcome of hernia in pigs.

Conclusion: On the basis of our studies we are able to exclude the two analyzed SNPs within the porcine GUSB gene as causative for the transmission of inguinal hernia.

Background

Inguinal and scrotal hernias are usually caused by a congenital defect, which occurs as a weakness of the inguinal canal. About 1% of all litters in German pig breeds are affected [1]. In several studies the estimated values of heritability for inguinal hernias in pigs range from $h^2 = 0.2$ [2] to $h^2 > 0.6$ [3,4]. The inguinal/scrotal hernia condition in pigs and in humans is affected by both environmental and genetic factors, with the involvement of multiple genes and likely incomplete penetrance, although the mode of inheritance has not been clarified so far and is still controversially discussed in literature [3,5-7]. In humans a high segregation ratio suggesting an autosomal dominant inheritance with incomplete penetrance and sex influence was reported. Additionally, the authors of this study noted a preferential paternal transmission of the disease gene and suggested a possible role of genomic imprinting in the aetiology of the hernia condition in man [8]. Anatomical factors such as an abnormal wide inguinal canal or a not obliterated processus vaginalis are regarded as predis-

Received: 19 December 2005
Accepted: 28 April 2006

BMC Veterinary Research 2006, 2:14 doi:10.1186/1746-6148-2-14

This article is available from: http://www.biomedcentral.com/1746-6148/2/14
© 2006 Beck et al; licensee BioMed Central Ltd.
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/2.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.
posing for inguinal and scrotal hernia. Thus a predisposition takes place during and shortly after testicular descent. The **gubernaculum** or **genitofemoral ligament** is emerging as the key anatomical structure in the control of testicular descent. The first phase of testicular migration [9] is characterized by rapid growth of the **gubernaculum**, which serves to dilate the inguinal canal prior to descent of the testes. This is accomplished through a rapid swelling of the **gubernaculum** due to formation of hydrated space, mediated mainly by the formation and deposition of hyaluronan (HA) within the **gubernacular** tissue [10,11].

During the second phase of testicular descent the **gubernaculum** involutes, presumably due to the removal of HA, thus allowing the testes to descend into the scrotum [12]. Enzymes that act in the biodegradation of HA feature the second phase of testicular descent, these are hyaluronidase, **β**-hexosaminidase and **β**-glucuronidase.

Our hypothesis is that the swelling of the **gubernaculum** exceeds through accumulation of HA due to a diminished degradation. The inguinal ring is thus extended in an unphysiological way and may remain open predisposing the male pigs for inguinal and scrotal hernia. The limitation of HA degradation may be based on mutations in one or several of the genes coding for the HA degrading enzymes.

A genome-wide linkage scan was conducted to identify chromosomal regions involved in the inheritance of hernia in pigs [13,14]. **GUSB** first considered as functional candidate was mapped within a region on porcine chromosome 3 (SSC3) [15] that was found to be associated with congenital inguinal and scrotal hernia. Therefore, the gene was considered as a positional candidate in addition to its functional relevance.

Here we report the isolation and complete genomic characterization of the porcine **GUSB** gene, its status as a positional candidate gene for congenital inguinal and scrotal hernia in combination with results obtained from linkage and association analysis of SSC3. Further, an association study of two SNPs within the **GUSB** coding sequence was conducted to elucidate their relevance for the phenotype hernia inguinalis/scrotalis.

### Results and discussion

**Linkage and association analysis on SSC3**

| Markers typed on SSC3 | cM (singlepoint) | NPL in singlepoint/p-valuea | cM (multipoint)b | NPL in multipoint/p-valueb | TDT-Score/p-valuea |
|-----------------------|------------------|----------------------------|------------------|---------------------------|--------------------|
| SW274                 | 0                | -0.03 / 0.5091             | 0                | 1.19 / 0.1170             | 4.90 / 0.555       |
| SW2021                | 12.4             | 1.370 / 0.0873             | 12.4             | 2.49 / 0.0074             | 16.55 /           |
| SW2429                | 17.2             | 1.580 / 0.0587             | 17.2             | 2.14 / 0.0177             | 13.30 /           |
| SW833                 | 17.3             | 0.530 / 0.2969             | 17.3             | 2.24 / 0.0137             | 11.55 /           |
| SW72                  | 17.8             | **2.290 / 0.0125**         | 17.8             | 2.92 / 0.0022             | 11.40 /           |
| SE51018               | 19               | 1.490 / 0.0694             | 19               | 2.58 / 0.0057             | 9.72 / 0.373      |
| S0701                 | 31.46            | 0.460 / 0.3188             | 31.46            | 1.37 / 0.0866             | 3.33 / 0.188      |
| S0206                 | 42.3             | 0.260 / 0.3949             | 42.3             | 0.92 / 0.1740             | 4.69 / 0.584      |
| SW2408                | 94.2             | 1.480 / 0.0707             | 94.2             | 1.88 / 0.0318             | 9.21 / 0.602      |
| S0002                 | 102.2            | **2.310 / 0.0119**         | 102.2            | 2.38 / 0.0097             | 19.00 /           |
| S0397                 | 109              | 0.050 / 0.4758             | 109              | 1.49 / 0.0697             | 10.85 /           |
| SW1327                | 109.6            | 1.730 / 0.0441             | 109.6            | 1.62 / 0.0543             | 14.03 /           |
| SW349                 | 112.6            | 0.380 / 0.0852             | 112.6            | 1.73 / 0.0434             | 13.64 /           |

---

*a Significant results are printed in bold letters  
*b NPL scores were calculated using the multipoint algorithm of Allegro (version 1.0). The first value in each cell refers to the exact marker position, the second value refers to an additional theoretical position, which are set in the middle between adjacent markers.
their position on the chromosome and the location of the significant nonparametric LOD scores (NPL) obtained in single- and multipoint analyses. The information content was 0.62 on average. Two regions on chromosome 3 (region I: between 12.4 cM and 19 cM; region II between 98.2 cM and 102.2 cM) showed significant NPL. The singlepoint analysis revealed a significant NPL for markers SW72 (region I, p = 0.0125) and S0002 (region II, p = 0.0119), while multipoint linkage analysis gave significant NPL values for the five region I markers SW2021 (p = 0.0074), SW2429 (p = 0.0177), SW833 (p = 0.0137), SW72 (p = 0.0022), SE51018 (p = 0.0057), and for the two region II markers SW2408 (p = 0.0318) and S0002 (p = 0.0097).

The transmission disequilibrium test (TDT) was applied to all markers on SSC3 the resulting p-values are given in Table 1. TDT confirmed the results of the nonparametric linkage analysis within region I for the markers SW2429, SW833, SW72 (p ≤ 0.05) and within region II for marker S0002 (p ≤ 0.1). A significant TDT score (p ≤ 0.1) was also found for marker SW1327 located at position 109.6 cM, about 6 cM distal of region II.

Physical distances between GUSB and microsatellite markers

GUSB was mapped by analyses of a whole-genome radiation hybrid panel [16] and fluorescence in situ hybridization, the most significantly linked markers were SW251 and SW72 [15]. Twopoint analyses gave a distance of 61 cR to SW72 and one of 54 cR to SW251. Hawken and colleagues estimated a conversion ratio (kb to cR) for the panel of 50.4 kb/cR on SSC3 [17]. Therefore, the distances in cR correspond to a distance of 3.07 Mb between GUSB locus and SW72 and one of 2.72 Mb between GUSB locus and SW251 at 42.3 cM [18].

These results confirmed GUSB as a positional candidate gene for congenital inguinal and scrotal hernia in addition to its previously proposed functional relevance.

Table 2: Comparative sequence statistics

|          | Human | Pig   | Mouse |
|----------|-------|-------|-------|
| Cytogenetic localization | HSA7q11.21 | SSC3p14-p16 | MMU5G1.3 |
| Sequence considered (bp) | 25000 | 17121 | 16449 |
| Length of gene within sequence considered (bp) | 21287 | 14436 | 14148 |
| GC content (%) | 51.8 | 54.7 | 48.9 |
| CpG islands (position within sequence considered: length (bp)) | 3730 – 4379; 650 | 1510 – 2653; 1144 | 2239 – 2565; 327 |
| Repetitive sequence total (%) | 57.6 | 20.2 | 34.2 |
| SINE (%) | 51.7 | 14.7 | 31.9 |
| Alu/B1 (%) | 50.6 | -- | 9.3 |
| B2-B4 (%) | -- | -- | 22.1 |
| MiR (%) | 1.1 | 0.0 | 0.0 |
| IDs (%) | -- | 0.5 | 0.5 |
| LINE (%) | 0.5 | 4.5 | 0.0 |
| LTR (%) | 4.9 | 0.0 | 0.0 |
| Others (%) | 0.3 | 1.0 | 2.3 |

The transmission disequilibrium test (TDT) was applied to all markers on SSC3 the resulting p-values are given in Table 1. TDT confirmed the results of the nonparametric linkage analysis within region I for the markers SW2429, SW833, SW72 (p ≤ 0.05) and within region II for marker S0002 (p ≤ 0.1). A significant TDT score (p ≤ 0.1) was also found for marker SW1327 located at position 109.6 cM, about 6 cM distal of region II.

Structural characterization of clone TAIGP714N09100Q

Pulsed field electrophoresis after NotI digestion of the isolated PAC clone TAIGP714N09100Q revealed an insert size of 80 kb. Subsequent partial sequence characterization confirmed that the clone contained the whole porcine functional GUSB gene as well as at least one exon of the so far uncharacterized porcine VKORC1L1 (vitamin-K epoxide reductase paralog) gene.

Characterization of genomic sequence and comparative sequence analyses

A contig spanning 17,157 bp, which contained the entire GUSB gene and 2.7 kb of its 5’ and 3’ flanking regions was assembled [GenBank:DQ095863] by a combination of subcloning, primerwalking and sequencing of long range PCR-products. The porcine GUSB gene spans over 14,436 bp (from start to stop codon) and consists of 12 exons ranging in length from 85 bp to 213 bp, which are separated by introns ranging in length from 84 bp to 3,928 bp. While the length of the exons is highly conserved between humans and pig, the size of the introns varies considerably, thus accounting for the larger size (21,521 bp) of the human GUSB compared to its porcine ortholog.

A comparison of the porcine GUSB protein-encoding region (1,959 bp) with mammalian orthologs revealed nucleotide sequence identities of 88% with the feline
GUSB [GenBank:NM_001003191], 87% with the canine GUSB [GenBank:NM_001003191], 84% with the human GUSB [GenBank:NM_001003191], and 81% with the murine Gusb [GenBank:J02836]. The observed identities between the known mammalian GUSB coding sequences are consistent with the molecular phylogenetic relationships between the different species. PiPMaker program [19] was further used to align the porcine sequence DQ095836 to a human BAC clone [GenBank:AC073261] on chromosome 7, that contains the human GUSB gene. Based on this alignment (data not shown) a region of 25 kb from position 35,000 to position 60,000 of the human BAC clone was used for further comparative sequence analyses. The murine sequence was taken from GenBank entry J02836. The results of screening for CpG islands and repetitive elements are shown in Table 2. DQ095863 reveals a high GC content of 54.7% compared to the mean GC content recently estimated at 49.6% and 40.7% for porcine coding sequences and introns respectively [20]. The high GC content is consistent with the GC content of 51.8% of the human GUSB gene. Contrary to that the GC content in the murine Gusb sequence is below 50% although five CpG islands were detected. Like many other housekeeping genes, human, murine, and porcine GUSB genes contain a CpG island that spans the promoter region, exon 1, and parts of intron 1. Only this promoter-associated CpG island is present in the three species, although the human (650 bp) and the murine (327 bp) promoter-associated CpG islands are significantly shorter than their porcine counterpart (1144 bp). Moreover, the porcine sequence contains three CpG islands in total the mouse ortholog five regions, whereas human GUSB contains only two CpG islands. These differences in amount and length of CpG rich regions contribute to the difference in the total GC content between the human, murine and porcine sequences. The CpG islands annotated in Table 2 matched the commonly defined criteria for a CpG island: 1) regions of DNA of at least 200 bp in length; and 2) a CG content above 50% and a ratio of observed versus expected CpGs above 0.6 [21].

The human, murine and porcine GUSB sequences were also screened for repetitive elements using the RepeatMasker database. The total content of repetitive sequences in human GUSB accounts 57.6%. Contrary, the porcine GUSB contains only 20.2% repetitive elements. An intermediate repeat content of 34.2% was found in the murine Gusb gene. In all three species the vast majority among the repetitive elements are species-specific SINE repeats, with 51.7% in the human and 31.9% in the murine but only 14.7% in the porcine GUSB sequence. This discrepancy can be explained by the fact that pig-specific SINE repeats (PRES) [22] are less characterized than human-specific Alus or murine B1-B4 elements. It is, however, most likely that this discrepancy primarily reflects a true difference rather than a biased detection in the species-specific repeat content, since all human GUSB introns are longer than their porcine counterparts and should therefore contain more SINE repeats.

Furthermore, 97.9% of the human SINEs are Alu elements, which are specific for primates. Recently, the human GUSB gene has been described to show an extremely high percentage of Alu repeats compared to the human genome average of 6-12%. The authors postulated that the Alu elements serve a vital function that precedes pseudogene proliferation within the human genome. Therefore, the authors concluded that most of the human GUSB pseudogenes have arisen from involvement of Alu-Alu recombination [23]. As the porcine GUSB contains only few SINEs in its 5’ region and since we found no evidence for unprocessed GUSB pseudogenes within the pig genome during screening of the genomic PAC library our results support their hypothesis.

**SNP analysis**

After screening all of the 12 GUSB exons with the adjacent intron junctions a total of five SNPs located in exons 5, 6, and 7 as well as in introns 3 and 8 were discovered. At positions 1,044 (exon 6) and 1,227 (exon 7) each (num-bering refers to predicted CDS) thymine is substituted by

| Exon | Primers<sup>a</sup> | Position<sup>b</sup> | Test-Assay | Resulting fragments | Allele-Frequencies |
|------|---------------------|---------------------|------------|--------------------|-------------------|
| 5    | EX5outer.f/EX5outer.r/EX5innerG/EX5innerA | 805 | Tetra-primer | G: 606, 342 A: 606, 308 | G: 0.707 A: 0.293 |
| 6    | Ex6Seq.f/Ex6Seq.r   | 1044 | Hpol-PCR-RFLP | T: 132, 109 C: 241 | T: 0.651 C: 0.349 |
| 7    | Ex7Seq.f/Ex7Seq.r   | 1227 | Hhol-PCR-RFLP | T: 295 C: 123, 118 | T: 0.987 C: 0.013 |

<sup>a</sup> Primer sequences given in table 1  
<sup>b</sup> Numbering refers to CDS as outlined in GenBank:DQ095863
cytosine, but both substitutions do not lead to an exchange of the encoded amino acid. The transition of guanine to adenine at position 805 (exon 5), however, changes the encoded amino acid from valine to methionine. Additionally two SNPs in intron 3 (T → C) and in intron 8 (C → T) were detected, but both of them do not alter splice motifs.

The 512 animals of the half-sib pedigree were genotyped for the three SNPs in exons 5, 6, and 7. Test assays used for genotyping and resulting allele frequencies for each of the SNPs are given in Table 3. The SNP in exon 7 was excluded from further analyses, because of the uneven allele frequency distribution with a minor allele frequency of 0.0013.

**Association analysis for GUSB-SNPs**

TDT was applied to the SNPs in exons 5 and 6 to test simultaneously for association and linkage with congenital inguinal hernia. TDT analyses gave p-values of 0.229 and 0.195 for the SNPs in exon 5 and exon 6, respectively. Therefore, no significant association between the allelic variants and the disease outcome was detected.

**Conclusion**

All twelve exons of the porcine GUSB gene were screened for SNPs to evaluate a possible contribution of the gene in the inheritance of hernia. No structural mutation that significantly alters structure of the protein has been detected. In order to evaluate if so far undetected causative mutations that lie in linkage disequilibrium with one of the detected SNPs may contribute to the disease outcome a straightforward SNP-based association analyses was conducted. No association was found for the two tested SNPs. Nevertheless it is well recognized that especially for complex diseases common association analysis of candidate genes within all disease-linked chromosomal regions should shed further light on the contribution of single genes to the outcome of inguinal and scrotal hernia in pigs.

**Methods**

**Animals and sampling**

Blood, tissue and sperm specimen were collected on pig farms selected by different German pig breeding organisations. Before sampling, farmers reported affected animals and the hernia phenotype was proven with palpation done by veterinarians. Several additional data were recorded like position of hernia (left-sided or right-sided), number of the respective litter, date of birth. Collection of affected animals including their non-affected parents is still continuing and up to now a total of 1710 animals have been collected for the DNA repository. The affected piglets are crossbreds of the founder breeds Pietrain, Large White, Hampshire, and German Landrace. For this study, 275 affected piglets and their non-affected parents (84 boars and 153 sows) were taken. The pedigree consisted therefore of 84 boar families, each of them contains at least one affected sib pair which, if available, were completed with affected half sibs. Unaffected sibs of herniated pigs have not been collected. The size of a family ranged from a minimum of four individuals (an affected sib pair plus parents) up to a complex design with 22 individuals (five affected sib pairs and eight affected half sibs including the parents) for the largest family.

**Linkage analyses**

The pedigree was genotyped for 13 microsatellite markers distributed evenly across chromosome 3. Markers and the genetic distances were taken from the published MARC U.S. Department of Agriculture map [25]. Primers used for genotyping of the 13 microsatellites on SSC3 are given in Table 4. PCR was done using puReTaq Ready-To-Go
Table 5: Primers used for GUSB analysis

### a) Primers used for primerwalking

| Primer name         | Sequence of forward primer (5’–3’) | Strand orientation | Location in GUSB | Template DNA | Template DNA | T<sub>Ann</sub>(°C) |
|---------------------|------------------------------------|--------------------|------------------|--------------|--------------|----------------------|
| GAPIntron10.f       | ggggctgagtgttagacaatgca            | plus               | Intron 10        | PAC clone    |              | 53                   |
| GAP5prime.f         | cattaacggaagattgctttggt            | plus               | Intron 1         | PAC clone    |              | 53                   |
| 3-Strich.f          | gcagagagagctgactggatgtg            | plus               | Exon 12          | PAC clone    |              | 50                   |

### b) Primers used for long range PCR

| Primer name | Sequence of primer (5’–3’) | Sequence of reverse primer (5’–3’) | Amplified part of gene | Length of product | Template DNA | T<sub>Ann</sub>(°C) |
|-------------|-----------------------------|-----------------------------------|------------------------|-------------------|--------------|----------------------|
| AuEx8.f/AuEx9.r | caggctgcttactaatcaatgaaagcttggg          | ggttcaagactggagtggcg           | Intron 8              | 3260             | 55            | 55                   |
| AuEx9.f/AuEx10.r | cttctgctgctgtcgttctctcctccctcccctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctcta...
PCR beads (Amersham Pharmacia Biotech, Freiburg, Germany), reactions contained 100 ng genomic DNA and 10-50 pmol of each primer. For each reaction one of the primers was end-labelled with 6-FAM, JOE, NED or HEX and analysis of PCR-fragments was performed on an ABI Prism 3100 Genetic Analyzer (Applied Biosystems, Weiterstadt, Germany). Results of fragment analysis were evaluated using Genescan version 3.7 and Genotyper version 3.6 (Applied Biosystems, Weiterstadt, Germany).

Information content of markers, singlepoint and multipoint nonparametric linkage scores (NPL) were calculated using the program package Allegro 1.0 [26]. NPL \( \geq 1.96 \) and p-value \( \leq 0.05 \) were considered as significant.

**TDT analyses**
Transmission disequilibrium test (TDT) [27] was conducted using the program GASSOC [28]. A p-value of 0.05 was considered as significance limit.

**PAC clone isolation and characterization**
Isolation of a porcine PAC clone containing the GUSB gene from the library TAIGP714Q [29] was performed as described elsewhere [15]. The insert-size of the PAC clone (TAIGP714N09100Q) was determined by cleavage with NotI and separation of the fragments by pulsed field gel electrophoresis (CHEF-DR II; Bio-Rad, Munich, Germany) on a 0.8% agarose gel.

**Subcloning and sequencing**
The PAC clone was digested by appropriate restriction endonucleases, and the positive fragments were subcloned into pGEM-4Z (Stratagene, Heidelberg, Germany) poly linker after southern blotting and hybridization with a GUSB gene specific probe. Recombinant plasmids were used to transform *E. coli* XL1-Blue (Promega, Mannheim, Germany). If necessary the obtained fragments were further subcloned after enzymatic digestion. Purified plasmid DNA were bidirectionally sequenced using the ABI Prism BigDye Terminator Cycle Sequencing Kit (Applied Biosystems, Weiterstadt, Germany) and cloned in the pGEM-T Vector System (Promega, Mannheim, Germany) for sequencing and further subcloning. All sequencing reactions were analyzed using an ABI PRISM 3100 Genetic Analyzer (Applied Biosystems, Weiterstadt, Germany). Sequence data were evaluated and overlapping contigs were generated using the software Sequencher™ 4.1 (GeneCodes, Ann Arbor, USA). Contigs were generated out of at least two independent subclones and/or PCR-products. The consensus sequence is build of plus and minus strand information.

**Bioinformatical sequence analysis**
BLAST searches were conducted to identify orthologous sequence regions between species [30]. Repetitive sequences were identified, masked, and annotated with RepeatMasker [31] CpG islands were detected with CpG plot [32] using a 200-bp window and a minimum length of 200 bp. The threshold values for CpG islands were CpGobs/CpGexp > 0.6 and GC > 0.5.

**PCR conditions**
All applied PCR reactions were done in compliance with the following set-up unless otherwise indicated: 50 ng of template DNA, 200 \( \mu \)M of each dNTP, 0.4 \( \mu \)M of each primer, 1.5 mM MgCl\(_2\), and 1.5 U of Taq-polymerase in 1x buffer as recommended by the manufacturer (QIAGEN, Hilden, Germany) in a total reaction volume of 25 \( \mu \)l.

**SNP discovery**
To screen the GUSB exons and adjacent intron sequences (Table 5c) for the existence of SNPs, a direct sequencing approach was conducted. DNA from 15 unrelated herniated piglets from the pedigree described before and from 10 unrelated artificial insemination boars of different breeds (7x Pietrain, 2x German Landrace and 1x Large White) without any known herniated progeny was used in the comparative sequencing approach. The primer pairs specified in Table 5c were generated to amplify each exon separately. PCR amplifications were done with an initial denaturation step at 94°C for 4 min, followed by 30 cycles with 30 sec at 95°C, 30 sec at the primer specific annealing temperature and 30 sec at 72°C, followed by a final elongation for 4 min at 72°C. PCR products were sequenced directly as described above. Multiple sequence comparisons to evaluate SNPs were done with Sequencher 4.1 (GeneCodes, Ann Arbor, USA).

**SNP genotyping by PCR-RFLP**
For SNP genotyping PCR-RFLP tests were developed for the two SNPs located in exon 6 and 7. RFLP reactions were performed on 10 \( \mu \)l of PCR-sample using 1.5 U of the respective restriction enzyme (Table 3) in a total reaction
Biozym, Hessisch Oldendorf, Germany) in a total reaction volume of 25 µl. The primers EX5outer.f and EX5outer.r were end-labelled with 6-FAM and analysis of PCR-fragments was performed on an ABI Prism 3100 Genetic Analyzer. Results of fragment analysis were evaluated using Genotyper version 3.6 (Applied Biosystems, Weiterstadt, Germany).

**Authors' contributions**

JB performed the molecular genetic analysis of GUSB, sequence alignment, comparative sequence analysis and wrote the manuscript. KBK carried out the genomescan. CK edited the manuscript and participated in the coordination, design of the study. HT performed the statistical analysis. BB was responsible for funding, project coordination and manuscript preparation as senior author.

**Acknowledgements**

The authors acknowledge financial support from the FBF e.V. to BB and CK and from DFG to KBK. The project was funded by a grant of the Erxleben Research & Innovation Council (ERIC-BR1959-2001-09). The authors thank A. Siebels for expert technical assistance in sequencing and V. Raupach and F. Maier for technical assistance in SNP genotyping. Further, the authors thank German Breeding organizations namely BHZP Lüneburg, Schweinezuchtverband Nord-West Muenster, GFS Ascheberg, and SWE Oldenburg for providing samples.

**References**

1. Althoff W: Zur erbl. Abhängigkeit verschiedener Geburtsfehler beim Schwein. Bonn: Universitäet Bonn; 1985.
2. Althoff W, Mayer M, Richter L, Simon D: Zur erbl. Abhängigkeit der Geburtsfehler Brüche und Binnenhodigkeit beim Schwein. Züchtungskunde 1988, 60(4):319-329.
3. Mikami H, Fredeen HT: A genetic study of cryptorchidism and scrotal hernia in pigs. Can J Genet Cytol 1979, 21(1):9-19.
4. Thaller G, Dempfle L, Hoeschele I: Maximum likelihood analysis of rare binary traits under different modes of inheritance. Genetics 1996, 143(4):1819-1829.
5. Hurson R, Narpenstein D, Schoeneweis D, Leipold HW: Congenital defects in pigs. Vet Bull 1978, 48:645-674.
6. Thaller G: Komplxe Segregationsanalyse zum Nachweis von Hauptgenen bei dichotomen Merkmalen am Beispiel der Anomalien des Schweins. Muenchen: Technische Universitaet; 1992.
7. Warwick WL: A study of hernia in swine. Wisc Agric Exp Stat Bull 1926, 69:1-27.
8. Gong Y, Shao C, Sun Q, Chen B, Jiang Y, Guo C, Wei J, Guo Y: Genetic study of indirect inguinal hernia. J Med Genet 1994, 31(3):187-192.
9. Wensing CJ: Testicular descent in some domestic mammals. 3. Search for the factors that regulate the gubernacular reaction. Proc K Ned Akad Wet C 1973, 76(2):196-202.
10. Heyns CF, Human Hj, Werely CJ, De Klerk DP: The glycosaminoglycans of the gubernaculum during testicular descent in the fetus. J Urol 1990, 143(3):612-617.
11. Hurson JM, Hasthorpe S, Heyns CF: Anatomical and functional aspects of testicular descent and cryptorchidism. Endocr Rev 1997, 18(2):259-280.
12. Fentener van Vlissingen JM, Koch CA, Delpech B, Wensing CJ: Growth and differentiation of the gubernaculum testis during testicular descent in the pig: changes in the extracellular matrix, DNA content, and hyalurondase, beta-glucuronidase, and beta-N-acetylglucosaminidase activities. J Urol 1989, 142(3):837-843.
13. Bornemann-Kolatzki K: Durchführung eines Genomscans mit polymorphten DNA-Markern und Genom Mismatch Scanning (GMS) bei Sus scrofa zur Detektion Hernia inguinalis/
scrotalis assoziierter Genomregionen. Hannover/Goettingen: Tierarztlche Hochschule Hannover & Universit Konstanz; 2004.
14. Knorr C, Bornemann-Kolatzki K, Breig B. Genome-wide scan for porcine rupture identifies association with several chromosomal regions. In Plant & Animal Genome XIV: 2006 San Diego; 2006.
15. Beck J, Knorr C, Habermann F, Fries R, Breig B. Assignment of the beta-glucuronidase (GUSB) gene to porcine chromosome SSC3p16 →p14 by FISH and confirmation by hybrid panel analyses. Cytogenet Genome Res 2002, 97(3–4):277G.
16. Yerle M, Finton P, Robic A, Alfonso A, Palvaudea Y, Delcros C, Hawken R, Alexander L, Beattie CW, Schook LB, Alexander LJ. A first-generation porcine whole-genome radiation hybrid map. Mamm Genome 1999, 10(8):824-830.
17. Hawken RJ, Murtaugh J, Flickinger GH, Yerle M, Robic A, Milan D, Gellin J. Repeatmasker. BMC Genomics 2005, 6(1):70.
18. Schwartz S, Zhang Z, Frazer KA, Smit A, Riemer C, Bouck J, Gibbs R, Hardison R, Miller W. PipMaker – a web server for aligning two genomic DNA sequences. Genome Res 2000, 10(4):577-586.
19. Tomatsu S, Orii KO, Islam MR, Shah GN, Grubb JH, Suzuki Y, Orii T, Kondo N, Sly WS. Methylation patterns of the human beta-glucuronidase gene locus: boundaries of methylation and general implications for frequent point mutations. Cpg dinucleotides. Genomics 2002, 79(3):363-375.
20. Yu K, Gu CC, Xiong C, An P. Province MA: Global transmission disequilibrium tests based on haplotype sharing in multiple candidate genes. Genet Epidemiol 2005, 29(4):323-335.
21. Cross SH, Bird AP. CpG islands and genes. Curr Opin Genet Dev 1995, 5(3):309-314.
22. Breig B. Conserved nucleotide differences and subfamily structure of porcine short interspersed elements. Anim Genet 1999, 30(2):120-125.
23. Tomatsu S, Orii KO, Islam MR, Shah GN, Grubb JH, Suzuki Y, Orii T, Kondo N, Sly WS. Methylation patterns of the human beta-glucuronidase gene locus: boundaries of methylation and general implications for frequent point mutations. Cpg dinucleotides. Genomics 2002, 79(3):363-375.
24. Feng F, Wu CC, Xiong C, An P. Province MA: Global transmission disequilibrium tests based on haplotype sharing in multiple candidate genes. Genet Epidemiol 2005, 29(4):323-335.
25. Rohrer GA, Alexander LJ, Hu Z, Smith TP, Keele JW, Beattie CW. A comprehensive map of the porcine genome. Genome Res 1996, 6(5):371-391.
26. Kruglyak L, Daly MJ, Reeve-Daly MP, Lander ES: Parametric and nonparametric linkage analysis: a unified multipoint approach. Am J Hum Genet 1996, 58(6):1347-1363.
27. Spielman RS, McGinnis RE, Ewens WJ: Transmission test for linkage disequilibrium: the insulin gene region and insulin-independent diabetes mellitus (IDDM). Am J Hum Genet 1993, 52(3):506-516.
28. Schaid DJ: General score tests for associations of genetic markers with disease using cases and their parents. Genet Epidemiol 1996, 13(5):423-449.
29. Al-Bayati HK, Duscher S, Kollers S, Rettenberger G, Fries R, Breig B. Construction and characterization of a porcine PAC library covering 3.2 genome equivalents and cytogenetical assignment of six type I and type II loci. Mamm Genome 1999, 10(6):569-572.
30. Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 1997, 25(17):3836-3840.
31. Repeatmasker [http://www.repeatmasker.org]
32. CpG plot [http://www.ebi.ac.uk/emboss/cpgplot]
33. Ye S, Dhillon S, Ke X, Collins AR, Day IN: An efficient procedure for genotyping single nucleotide polymorphisms. Nucleic Acids Res 2001, 29(17):E88-E88.