An intronic insertion in *KPL2* results in aberrant splicing and causes the immotile short-tail sperm defect in the pig

Anu Sironen*†, Bo Thomsen‡, Magnus Andersson§, Virpi Ahola¶, and Johanna Vilkki*

*MTT Agrifood Research Finland, Animal Production Research, Animal Breeding, FIN-31600, Jokioinen, Finland; ‡Department of Genetics and Biotechnology, Danish Institute of Agricultural Sciences, P.O. Box 50, DK-8830 Tjele, Denmark; §Department of Clinical Veterinary Sciences, Saari Unit, Faculty of Veterinary Medicine, University of Helsinki, FIN-04920, Saarentaus, Finland; and ¶MTT Agrifood Research Finland, Food Research, FIN-31600, Jokioinen, Finland

Edited by Ryuzo Yanagimachi, University of Hawaii, Honolulu, HI, and approved February 3, 2006 (received for review July 25, 2005)

The immotile short-tail sperm defect is an autosomal recessive disease within the Finnish Yorkshire pig population. This disease specifically affects the axoneme structure of sperm flagella, whereas cilia in other tissues appear unaffected. Recently, the disease locus was mapped to a 3-cM region on porcine chromosome 16. To facilitate identification of candidate genes, we constructed a porcine-human comparative map, which anchored the disease locus to a region on human chromosome 5p13.2 containing eight annotated genes. Sequence analysis of a candidate gene *KPL2* revealed the presence of an inserted retrotransposon within an intron. The insertion affects splicing of the *KPL2* transcript in two ways; it either causes skipping of the upstream exon, or causes the inclusion of an intronic sequence as well as part of the insertion in the transcript. Both changes alter the reading frame leading to premature termination of translation. Further work revealed that the aberrantly spliced exon is expressed predominantly in testicular tissue, which explains the tissue-specificity of the immotile short-tail sperm defect. These findings show that the *KPL2* gene is important for correct axoneme development and provide insight into abnormal sperm development and infertility disorders.

cilia | retrotransposon | spermatogenesis

Cilia and flagella play important roles in many physiological processes, including cellular and fluid movement, sensory perception, and development. The biogenesis and maintenance of cilia depend on the intraflagellar transport system, which is required for the assembly and elongation of cilia by transporting ciliary precursors to their site of incorporation (1). The internal cytoskeletal structure of cilia, flagella, basal bodies, and centrioles, called the axoneme, is highly conserved among eukaryotic cells and consists of ≈250 polypeptides (2, 3). The axoneme structure of most motile cilia and flagella consists of nine outer doublet microtubules surrounding a central pair of singlet microtubules. Projecting from the doublet microtubules are an inner and an outer row of dyneins, which are ATP-dependent motor proteins. Neighboring peripheral doublet microtubules are linked to each other by the elastic protein nexin, which are also connected to the inner singlets by radial spokes. Axonemal bending, which provides the force for cilia movement, is generated by transient interactions of dyneins and doublet microtubules that cause sliding between pairs of outer microtubules. The central pair and radial spokes complex selectively interact with subsets of dynein arms to regulate the sliding movements of microtubules (4, 5). Furthermore, primary cilia are usually immotile and contain a "9 + 0" axoneme that lacks the central pair singlets, the radial spokes, and the dynein complex. Primary cilia are ubiquitous organelles in most vertebrates (6).

Mutations in proteins that function in basal bodies, intraflagellar transport system machinery, axonemes, ciliary matrix, and ciliary membrane can lead to cilia related diseases in the human such as polycystic kidney disease, retinal dystrophy, neurosensory impairment, Bardet-Biedl syndrome, or primary ciliary dyskinesia (PCD) (7–10). PCD is a genetically heterogeneous group of disorders with axonemal abnormalities affecting one in 16,000 individuals (11). PCD is characterized by the complete absence of or occurrence of defective cilia and flagella. Structural defects have been observed in several axoneme components, including outer and inner dynein arms, radial spokes, nexin links, and microtubules. Thus far, only mutations in genes *DNAI1*, *DNAH5*, and *DNAH11* encoding for proteins of the outer dynein arms have been identified (11–13). The common clinical manifestations of PCD are situs inversus, bronchiectasis, chronic sinusitis, and male sterility.

The first case of the immotile short-tail sperm (ISTS) defect in pigs was detected in Finnish Yorkshire (Large White) boars in 1987 and, to date, 82 boars are known to be affected. The ISTS phenotype is characterized as lowered sperm counts, short sperm tails, and axonemal abnormalities. Electron microscopic examination of flagella cross-sections has revealed that typically one or both of the central microtubules are missing, and often there are less than nine doublets and the subunits of the doublets are broken apart. However, dynein arms appear normal in cross-sections. Approximately 5% of spermatozoa of affected boars have flagella of normal length, but none are motile (14). The disorder appears to be specific to sperm tail development, because no effects on the structure of cilia in the respiratory or female reproductive tract have been observed (14). The ISTS defect provides an ideal opportunity to analyze the function of a gene affecting cilia and sperm tail development. Homozygosity mapping and haplotype analysis has located the ISTS associated gene to porcine chromosome 16 within a 3-cM region proximal to SW419 (15). In the present study, we fine-mapped the causative mutation to an interval corresponding to 1.158 kbp on human chromosome 5 containing eight annotated genes. We show that, in one of these genes, *KPL2*, a retrotransposon within an intron in homozygous affected boars leads to aberrant splicing in testicular tissue. These findings are consistent with earlier studies reporting that *KPL2* is expressed predominantly in ciliated tissues and at specific stages of sperm cell development in the rat (16).

Results

Fine Mapping. Fine mapping was initiated by isolating porcine BAC clones containing the markers SW2411, SW419, and S0006.

Conflict of interest statement: No conflicts declared.

This paper was submitted directly (Track II) to the PNAS office.

Abbreviations: PCD, primary ciliary dyskinesia; ISTS, immotile short-tail sperm; qPCR, quantitative PCR.

Data deposition: The sequences reported in this paper have been deposited in the GenBank database (accession nos. DQ119847 for the Sus scrofa *KPL2* mRNA sequence and DQ092447 for the Sus scrofa *KPL2* partial genomic sequence).

†To whom correspondence should be addressed. E-mail: anu.sironen@mtt.fi.

End sequences of BAC clones were used to construct a partial contig of the disease-associated region by chromosome walking (data not presented). BLASTN searches with BAC-end sequences from the contig against the human NCBI database located the area between SW419 and S0006 within a 2-Mbp region on human chromosome 5p13.2. Porcine ESTs corresponding to genes in this region were sequenced and SNPs were found in the genes *AMACR* and *RAI14*. A recombination in two affected boars between the ISTS defect and *RAI14* reduced the disease-associated region to 1,158 kbp in the human map, which harbors eight annotated genes (*RAI14*, *FLJ25439*, *RAD1*, *BRIX*, *LOC134218*, *AGXT2*, *PRLR*, and *FLJ23577*). Intriguingly, the hypothetical human protein *FLJ23577* has a 79.4% identity to the rat *KPL2* protein sequence, which is highly expressed in testicular seminiferous tubules, and is therefore an attractive candidate gene (16). The ability to amplify a *KPL2*-specific product by PCR using the isolated BAC-clones as a template showed that the *KPL2* gene is indeed located in close proximity to the disease-linked marker SW419. Expression of *KPL2* in porcine testis was confirmed by PCR amplification of testicular cDNA with *KPL2*-specific primers. The cDNA of *KPL2* was sequenced from a normal and an affected boar. This analysis identified 10 SNPs and also revealed that exon 30 was absent in the *KPL2* transcript of the affected boar. Furthermore, genotyping of additional animals showed that all 10 SNPs were homozygous in several normal boars as well as in affected boars, excluding the possibility that these SNPs are disease-causing mutations. Mutation-associated exon skipping is known to be the underlying cause for an increasing number of diseases (17, 18), which prompted us to investigate the mechanism leading to the absence of exon 30 in affected animals in greater detail.

Mutation Detection. The *KPL2* intron 29 (3,305 bp) and the first 1,500 bp of intron 30 were sequenced by using genomic DNA of a normal and an affected boar. Three SNPs were found, all of which were heterozygous in affected boars. During this sequencing, some primer pairs designed to amplify the beginning of intron 30 were found not to produce a product using DNA from affected boars. Therefore, we used PCR amplification to scan for possible insertions or deletions by positioning overlapping primer pairs in intron 30 (Fig. 1*A*). Most primer pairs resulted in identical PCR fragments in normal and affected boars. However, certain primer combinations generated a product only in normal animals. The inability to generate a PCR fragment suggested the presence of a large insertion in intron 30 in affected animals. To corroborate this, we performed a Southern blot analysis using a probe that spans the junction between exon 30 and intron 30 (Fig. 1*B*). The data indicate that a large segment of ≈9,000 bp has been inserted into intron 30 in affected boars. Finally, long-range PCR was used to amplify a 9-kbp fragment, verifying the presence of a large insertion in intron 30 (Fig. 1*B*). Sequencing of the fragment is currently ongoing. Preliminary sequence information was used to search GenBank, which indicated that one end of the sequence was homologous to a porcine genomic sequence containing the PERV-A retrovirus (GenBank accession no. AY160111, identities 542/544), and to the 5′ genomic sequence of a porcine endogenous retrovirus clone PERV-A (GenBank accession no. AJ304824, identities 387/387). The sequence of the other end of the insertion showed 86% identity over >900 bp to LINE-1 elements, which are abundant retrotransposons in mammals (19). These data indicate that a retrotransposition event has disrupted the intron, although the exact structure of the retrotransposon remains unclear.

To generate a PCR fragment (KPL2i) diagnostic of the presence of the insertion, we designed a reverse primer within the insertion and used this primer in combination with a forward primer within exon 30 to genotype normal, carrier, and ISTS-affected pigs. Another fragment (KPL2n) amplified with a forward primer at the end of exon 30, and a reverse primer in the intron 30 downstream of the insertion site was used as a marker for unaffected chromosomes. This PCR-based assay showed that the insertion was homozygous only in ISTS affected boars and heterozygous in carrier pigs, whereas no product for KPL2i was observed in samples of normal individuals from the Yorkshire (n = 10), Duroc, Hampshire, and Landrace breeds (four individuals from each breed). Thus, these data show that the presence of the 9,000-bp insertion in intron 30 is associated with the ISTS defect.

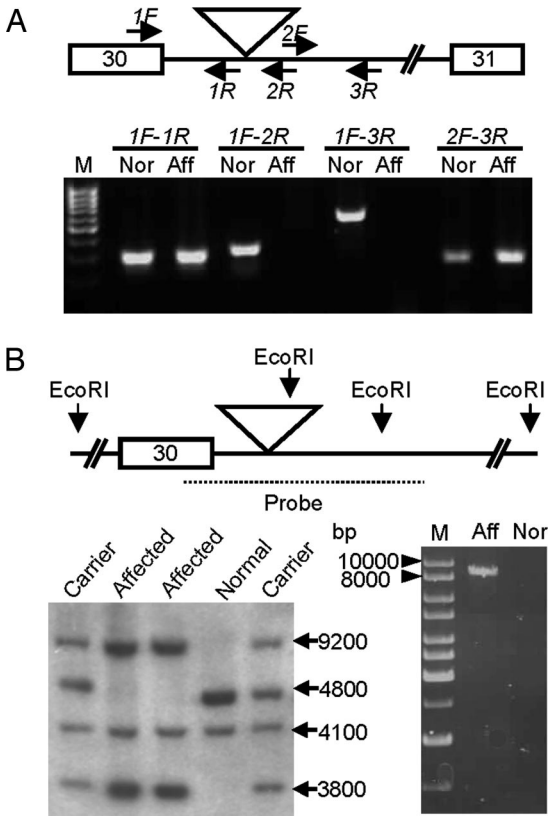

Fig. 1. Characterization of an insertion in intron 30 of the *KPL2* gene. (*A*) Identification of the insertion site in individuals affected with ISTS by PCR analysis. Primer pairs on both sides of the insertion resulted in similar PCR fragments for normal (Nor) and affected boars (Aff), whereas primer pairs over the insertion site only produced a PCR product for normal boars. (*B*) A schematic presentation of the genomic region containing the insertion in intron 30 of the porcine *KPL2* gene. Positioning of the EcoRI restriction sites and the insertion are indicated by arrows. Southern blot analysis of the region is shown in *Lower Left*, and long-range PCR in *Lower Right*. The data are consistent with the presence of a large insertion within the probed area in ISTS affected boars. Long-range PCR allowed a more precise estimate of the insertion size of ≈9,000 bp.

Characterization of Aberrant Splice Products. Subsequently, *KPL2* transcript splicing in various tissues was characterized. Semi-quantitative PCR amplification of cDNA across base pairs 3978–4466 (KPL2e29–36) produces a fragment of 487 bp in normal boars and a fragment of 257 bp in ISTS affected boars. A less abundant fragment of 998 bp was also detected in two affected boars (Fig. 2*A*). Sequencing of these fragments showed that exon 30 (230 bp) was missing in the shorter fragment, and that exon 30 was present in the longer fragment, but only in combination with part of the intron 30 (59 bp preceding the insertion) and the beginning of the insertion (452 bp). Importantly, the reading frame in both of these aberrantly spliced

GENETICS

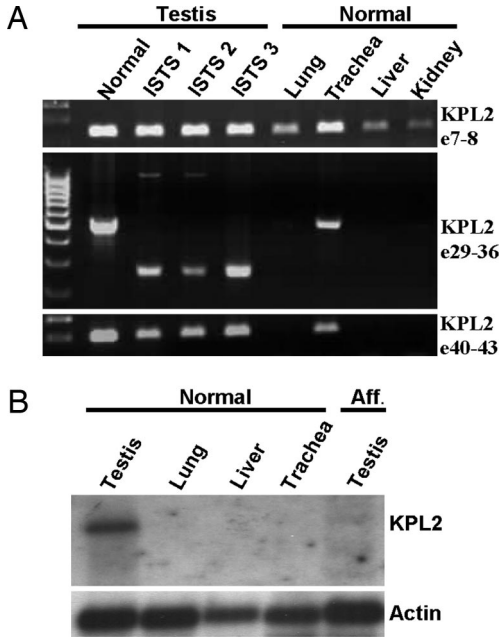

Fig. 2. Aberrant splicing of exon 30 in affected boars. (*A*) Analysis of expression of different parts of the *KPL2* gene in different tissues. RT-PCR was used to analyze the fragments KPL2e7–8 (exons 7–8), KPL2e29–36 (exons 29–30 and 36), and KPL2e40–43 (exons 40–43) in the testis of normal and ISTS affected boars, and in different tissues of a normal boar. The KPL2e29–36 fragment is mainly expressed in the testis and at lower levels in the trachea. Two splicing variants are expressed in ISTS-affected boars for KPL2e29–36. The shorter fragment (257 bp) is depleted of exon 30, and the longer fragment (998 bp) includes exon 30 together with part of intron 30 and the beginning of the insertion, as verified by sequencing. KPL2e40–43 is also only expressed in the testis and trachea. The KPL2e7–8 fragment was expressed in all tissues examined. (*B*) Northern blot analysis. The transcription level of *KPL2* exons 3–7 is moderate in normal testis and extremely low in the lung, liver, and trachea. Furthermore, *KPL2* expression is reduced in ISTS affected boars. Expression of *β-actin* was used as a control.

transcripts is disrupted, generating premature stop codons, and as a result truncation of the encoded protein (Fig. 3). This analysis also revealed that expression of KPL2e29–36 was tissue-specific, being expressed in the testis, and at a lower level in the trachea, but not in any of the other tissues examined. Likewise, another fragment containing exons 40, 41, and 43 (KPL2e40–43, base pairs 5191–5389) was also expressed specifically in the testis and trachea. In contrast, a fragment from *KPL2* exons 7–8 (KPL2e7–8, base pairs 956-1081) was found to be expressed in all tissues analyzed, but at lower levels in the lung, liver, and kidney, for both normal and affected boars.

***KPL2* Expression Levels in Various Tissues.** Testis tissue of three normal and three ISTS-affected boars were analyzed for relative expression of KPL2e7–8 and KPL2e29–30 using quantitative PCR (qPCR). Expression of *KPL2* was also determined in samples of lung, trachea, and liver of three normal and three ISTS-affected boars. In normal boars, the expression of the fragment KPL2e7–8 was 2.3-, 6-, and 12-fold lower in the trachea, lung, and liver, respectively, relative to the expression in the testis (Fig. 4*A*). In addition, affected boars showed a down-regulated KPL2e7–8 expression, primarily in the testis (3.8-fold), possibly caused by nonsense-mediated RNA decay of the mutated transcript, whereas the trachea, lung, and liver appeared less affected (Fig. 4*A*). In normal boars, the expression of KPL2e29–30 fragment was highest in the testis and ≈4.3-fold lower in the trachea, whereas no transcription was detected in the liver or lung. In affected boars, the expression is decreased 15-fold in the testis and ≈3-fold in the trachea (Fig. 4*B*), compared with normal individuals. The qPCR data were supported by Northern blot analysis, confirming that *KPL2* is transcribed predominantly in the testis and that the expression pattern is markedly altered in ISTS boars (Fig. 2*B*). Taken together, these data show that the *KPL2* gene is differentially expressed in healthy animals. Thus, the region coding for the C-terminal part of the protein appears to be expressed only in the testis and trachea (tissues with motile cilia), whereas the region coding for the N-terminal part is expressed in all of the tissues analyzed (including the lung, liver, and kidney). Lung tissue consists of bronchioles and alveoli, but only bronchioles contain motile cilia. The lower number of motile cilia in lung tissue relative to the testis and trachea would account for the lack of expression of the KPL2e29–30 fragment in lung tissue samples. Furthermore, the intronic 9,000-bp insertion in affected animals generates a C-terminally truncated version of the protein in the testis and trachea, and also leads to a decreased level of *KPL2* transcripts.

```
NORMAL   : GGCCAGTGTGGAAAAATTAACTGAGGTAGCTCGCTATCACATTGAAACATCTACCAAAATTCAGAATGAGCTTTATTTGGATCAAGAAGACTTC
ISTS1    : GGCCAG----------------------------------------------------------------------------------------
ISTS2    : GGCCAGTGTGGAAAAATTAACTGAGGTAGCTCGCTATCACATTGAAACATCTACCAAAATTCAGAATGAGCTTTATTTGGATCAAGAAGACTTC
NORMALp  : --A--S--V--E--K--L--T--E--V--A--R--Y--H--I--E--T--S--T--K--I--Q--N--E--L--Y--L--D--Q--E--D--F-
ISTS1p   : --A----------------------------------------------------------------------------------------
ISTS2p   : --A--S--V--E--K--L--T--E--V--A--R--Y--H--I--E--T--S--T--K--I--Q--N--E--L--Y--L--D--Q--E--D--F-

NORMAL   : TTCATTAACGGCAATATCAAGGTCTTTCCAGACCCGCCTCCTCCTGTTCGTCCTCCACCTGTAGAAAAGGAGGAGAATGGCACCCTGACCATCG
ISTS1    : ----------------------------------------------------------------------------------------
ISTS2    : TTCATTAACGGCAATATCAAGGTCTTTCCAGACCCGCCTCCTCCTGTTCGTCCTCCACCTGTAGAAAAGGAGGAGAATGGCACCCTGACCATCG
NORMALp  : -F--I--N--G--N--I--K--V--F--P--D--P--P--P--P--V--R--P--P--P--V--E--K--E--E--N--G--T--L--T--I--
ISTS1p   : ----------------------------------------------------------------------------------------
ISTS2p   : -F--I--N--G--N--I--K--V--F--P--D--P--P--P--P--V--R--P--P--P--V--E--K--E--E--N--G--T--L--T--I--

NORMAL   : AGCAGCTTGACAACCTTCGAGATCAGTTCTTAGATATAGCGCCTAAAGGCGTCATAGGAAATAAAGCATTTGCTGACATTCTGCTGGACTTGGT
ISTS1    : ----------------------------------------GCGTCATAGGAAATAAAGCATTTGCTGACATTCTGCTGGACTTGGT
ISTS2    : AGCAGCTTGACAACCTTCGAGATCAGTTCTTAGATATAGCGCCTAAAGTCCATCTCATGTTCGCATCATACAACATAGACCTGAATTAACTCAA
NORMALp  : E--Q--L--D--N--L--R--D--Q--F--L--D--I--A--P--K--G--V--I--G--N--K--A--F--A--D--I--L--L--D--L--V
ISTS1p   : ----------------------------------------------R--R--H--R--K--*--S--I--C--*--H--S--A--G--L--G
ISTS2p   : E--Q--L--D--N--L--R--D--Q--F--L--D--I--A--P--K--V--H--L--M--F--A--S--Y--N--I--D--L--N--*--L--K
```

Fig. 3. Aberrant splice products of ISTS-affected boars. The DNA and protein sequences are shown for the end of exon 29, exon 30, and the beginning of exon 31 of the *KPL2* gene in normal and ISTS affected boars. Both skipping of exon 30 in ISTS affected individuals (ISTS1) and partial inclusion of intron 30 and the insertion (ISTS2) disrupt the reading frame and produce several translation stop codons (gray shading), leading to premature termination of translation.

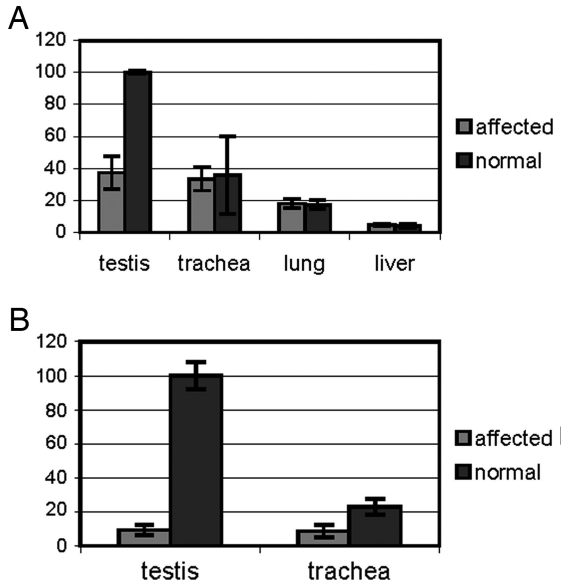

Fig. 4. Expression of different *KPL2* exons in various tissues of normal and ISTS affected boars relative to normal testis expression (100%). (*A*) The qPCR analysis of an mRNA fragment spanning *KPL2* exons 7 and 8. (*B*) The qPCR analysis of an mRNA fragment spanning *KPL2* exons 29 and 30. Amplification by qPCR was performed in triplicate on 50-ng cDNA samples of testicular, tracheal, lung, and liver tissues from normal and ISTS-affected boars. Mean values are presented and error bars indicate ± SD ($n = 3$; PCR analysis for each of three cDNA samples).

Sequence Alignments. The porcine *KPL2* nucleotide sequence from the present study (GenBank accession no. DQ119847) was translated into an amino acid sequence and aligned with the available full-length protein sequences from the NCBI database for the human *KPL2* isoform 1 (GenBank accession no. NP_079143) and the rat *KPL2* (GenBank accession no. NP_072142) using CLUSTALW (20). The pairwise sequence alignment scores were: human–pig 80%, human–rat 73%, and pig–rat 70%. Protein sequence lengths were 1,822, 1,812, and 1,744 aa, for the human, pig, and rat, respectively.

The NCBI LocusLink was used to identify exon/intron boundaries. *KPL2* in the human contains 43 exons spanning 197,525 bp, whereas the rat gene includes 40 exons spanning 170,943 bp, with exon numbering for the rat starting at exon 2 in the human gene. Exon numbering differs between species, such that the exon skipped in ISTS boars corresponds to exon 29 in the rat and exon 30 in the human. No published protein sequence contains all exons in the human, and exon differences are known to exist between the human and rat. Full-length protein sequences for the rat, human (isoform 1), and pig (testis cDNA), which are used for the alignment reported in Fig. 5, which is published as supporting information on the PNAS web site, lack human exons 31–35.

A complete protein–protein alignment between the human, rat, and pig showed that human exons 4 and 19 were missing in the rat sequence, and residues 1204–1226 (four imperfect copies of the sequence QAKKEKE) in exon 27 were absent in the human and pig sequence compared with that of the rat. The *KPL2* protein sequence contained several highly conserved exons (Fig. 5), including the porcine exon 30. A BLAST search with exon 30 also resulted in partial *KPL2* sequences for the dog, monkey, mouse, and chicken. The protein sequence for exon 30 was highly conserved across all mammalian species, with an alignment score varying mostly between 81 and 96%, being 89% between the human and pig, but only 53% between the chicken and pig.

Discussion

We present strong evidence that the presence of a retrotransposon in the *KPL2* gene causes the infertility phenotype of ISTS boars. The data demonstrate that processing of the transcript generated from the *KPL2* allele harboring the insertion produces two abnormal splice products. Thus, the majority of the *KPL2* transcripts in affected boars lacked exon 30, whereas a minor fraction retained exon 30, but also included intronic as well as retroelement sequences. Notably, the reading frames of both of these abnormally spliced transcripts are disrupted, generating premature translation stop codons, which truncate the protein at residues 1403 and 1487, respectively, from a total of 1,812 amino acids. The result is an entirely different and shorter C terminus, which is most likely the cause for the loss of function. Furthermore, the amount of *KPL2* transcripts were reduced in affected testicular tissue, possibly as a result of mRNA degradation by nonsense-mediated decay of transcripts containing premature termination codons.

In the rat, *KPL2* is expressed in tissues containing cilia-like structures such as the lung, trachea, testis, brain, and at lower levels in the kidney and spleen, whereas no expression is detected in the heart or liver, suggesting a role for this gene in ciliogenesis. Consistent with this suggestion, *KPL2* expression is closely correlated with ciliated cell differentiation in cultures of primary tracheal epithelial cells. Likewise, the spatio-temporal expression pattern in the seminiferous tubules at specific stages during sperm cell development supports *KPL2* as having a central role in the differentiation of axoneme-containing cells (16). The present study confirmed the expression pattern reported for rats with a high level expression in the testis, followed by an intermediate level in the trachea, and much lower expression in the lung, kidney, and liver. This suggests that the function of *KPL2* is not confined to motile cilia with a 9 + 2 axoneme structure, but that it may also play a role in immotile primary cilia. Within this context, it is important to note that several posttranscriptional regulatory pathways operate in combination with gene expression levels to determine the proteomic profile of cells. One mechanism is alternative splicing, which can generate a range of protein isoforms by the inclusion or skipping of exons, often in a cell-type or developmental stage-specific manner. In the human, two isoforms of *KPL2* have been identified (GenBank accession nos. NP_079143 and NP_653323), and several sequences with varying exon content from different tissues have been deposited in GenBank. Our results indicate that, in the pig, the primary *KPL2* transcript undergoes tissue-specific splicing to include exon 30 and presumably some other 3′end exons only in the testis and trachea. The intronic insertion leads to skipping of exon 30 and consequently causes termination of translation, which explains the sperm tail defects in ISTS boars. However, no respiratory dysfunction has been observed in ISTS pigs, and microscopic examination of tracheal cilia revealed no apparent effect on the axonemal structure (data not shown). It is unclear whether this is related to the 4- to 5-fold higher expression level in the testis relative to the trachea, which may indicate a more crucial role of at least one variant of *KPL2* in sperm tail development than in cilia differentiation. Alternatively, the phenotype could be dependent interactions of *KPL2* with other proteins, which are only expressed in the testis. However, it is also possible that symptoms take longer to develop in the trachea, and would not be detected in affected boars, which are usually slaughtered at a young age once infertility has been diagnosed. The clinical features of patients suffering form PCD also vary significantly (21). Typically males with immotile spermatozoa also have defective cilia in other tissues. However, cases have been reported where patients have immotile sperm, yet the structure and motility of other examined cilia are normal,

or in patients with no history of respiratory tract disorders (22–24).

Comparative sequence alignment of *KPL2* showed a high degree of cross-species conservation. Based on the human *KPL2* annotation (GenBank accession no. NP_079143), we performed sequence scans against protein domain databases to predict functional domains. This revealed the presence of a domain of unknown function called DUF1042 in the N terminus, which classifies *KPL2* together with other proteins implicated in flagella function such as the human *SPATA4* protein (spermatogenesis associate 4, GenBank accession no. NP_653245), the mouse sperm flagella protein *Spef1* (GenBank accession no. AY860964), and *CPC1* (central pair complex 1, GenBank accession no. AAT40991) of the unicellular organism *Chlamydomonas reinhardtii*. In addition, the N terminus contains a calponin homology domain, indicating potential actin binding activity. An adenylate kinase (ADK) domain and an ATP/GTP binding site (P-loop) are located centrally, where a number of other domains have been identified including two potential bipartite nuclear localization signals and a calcium-binding EF-hand motif, indicating that *KPL2* activity is modulated by calcium. The region containing the EF-hand is missing in ISTS boars, which may, to some extent, explain the sperm-specific structural changes (the positions, *E* values, and databases used are outlined in Table 1, which is published as supporting information on the PNAS web site). *KPL2* and *CPC1* share >40% similarity and both harbor several of the functional domains, including DUF1042, ADK, and EF-hands, which strongly suggest that they serve similar functions (25). Mutations in *CPC1* disrupt the assembly of the central pair microtubule-associated complex and alter flagellar beat frequency (25). However, the *KPL2* mutation produces a complex and more severe phenotype with fully disrupted central pair microtubules, as well as outer doublet defects, which may suggest different or additional roles of *KPL2* in the assembly of the axoneme. Furthermore, flagella in *CPC1* mutants can still beat, whereas all sperm flagella are immotile in ISTS boars, although ciliated cells other than spermatozoa are motile, for example in the ductuli efferentes of the testis. Mutagenesis studies in *Chlamydomonas* have revealed several genes (e.g., *pf18*, *pf19*, *pf20*, *pf6*, *pf16*, and *pf15*) coding for the central apparatus proteins that are essential for flagella motility, because their inactivation causes the paralysis of flagella (26, 27). Mutations in *pf16* (*spag6*) and *pf20* have also been shown to affect spermatogenesis in the mouse, where *spag6* interacts with *pf20.* Mutated *spag6* is known to cause infertility and truncated flagella (28, 29). These gene products have been localized to the central pair complex and are only expressed in the testis. These findings support the hypothesis that the *KPL2* protein may also be part of, or interact with, the central pair complex. However, the expression of *KPL2* in tissues devoid of motile cilia, such as the kidney and liver, suggests a wider role for at least one of the putative *KPL2* isoforms.

In conclusion, *KPL2* appears to be expressed in all cilia containing tissues, but presumably as different splice variants. The isoform containing the exon 30 encoded domain appears to play an important role in the correct assembly and function of the axoneme primarily in spermatozoa. Localization of the *KPL2* protein at the cellular level will elucidate the role of this gene in sperm tail development, and further studies are required to reveal the importance of different *KPL2* variants in other tissues. Mutations in various parts of *KPL2* are likely to produce different phenotypes, and thus the gene may be involved in multiple types of ciliary defects. Analysis of the function of *KPL2* may therefore provide a general insight into cilia malformations and abnormal development.

Materials and Methods

Pig Genomic Library Screening and Comparative Mapping. For fine mapping, BAC-clones (PigE BAC) from MCR geneservice (www.geneservice.co.uk/home; ref. 30) were picked up by PCR screening with markers SW2411, SW419, and S0006 located within the disease-associated region on porcine chromosome 16. The selected BAC clones were isolated with the Qiagen plasmid midi kit protocol in accordance with the manufacturer's recommendations. The ends of extracted BAC clones were sequenced and compared to the human sequence database (www.ncbi.nlm.nih.gov) to map the area between markers SW419 and S0006 on the human map.

PCR Amplification and DNA Sequencing. PCR amplification using BAC pools or pig genomic DNA as a template was performed with Dynazyme DNA polymerase (Finnzymes) according to the instructions from the supplier. Long-range PCR was performed by using Dynazyme EXT polymerase (Finnzymes) or a Long PCR Enzyme mix (Fermentas).

The PCR amplicons were purified by using ExoSAP-IT (Amersham Pharmacia), whereas PCR fragments were sequenced in both directions with the same primers used in the amplification procedures. The BAC-ends were sequenced with universal primers T7 and SP6. Sequencing was performed on MegaBace 500 capillary DNA sequencer (Amersham Pharmacia) using DYEnamic ET Terminator kits with Thermo Sequenase II DNA Polymerase (Amersham Pharmacia).

Gene Expression. For analysis of candidate gene expression, samples of testicular, liver, kidney, tracheal, and lung tissue from normal and ISTS affected boars were collected and stored in RNAlater buffer (Qiagen). Samples of testis, trachea, lung, and liver were available from three affected and three normal boars, but kidney samples were only available from one of the normal boars. Total RNA purification was performed with RNeasy Protect Mini and Midi kits (Qiagen). Total RNA was reverse transcribed (RT-PCR) with random primers and an RNA PCR kit (SuperScript, Invitrogen, and ImProm-II Reverse Transcription System, Promega) according to the manufacturer's instructions and amplified by using gene-specific primers. Control reactions were performed with a ribosomal 18S RNA, and this process was repeated using RNA isolated from three different animals where possible. Fragments KPL2e7–8 (126 bp), KPL2e29–36 (487 bp), and KPL2e40–43 (199 bp) were used to examine the expression of various components of *KPL2* in different tissues. The locations of the fragments are shown in Fig. 5, and the primers used for PCR amplification are listed in Table 2, which is published as supporting information on the PNAS web site.

Real-Time qPCR. qPCR was used to measure the relative RNA transcript levels of *KPL2* in the testis, trachea, lung, and liver. Concentrations of cDNAs were measured and tissue cDNA samples were diluted 1:100 before use. Two fragments from different regions of *KPL2* (KPL2e7–8 and KPL2e29–30) were analyzed by using ribosomal 18S RNA as an internal reference gene (a list of primers used is given in Table 2). The qPCR was performed with an ABI 7000 Sequence Detection System in 96-well microtiter plates using Absolute qPCR SYBR Green ROX Mix (VWR). Amplification by qPCR contained 12.5 μl of Absolute qPCR SYBR Green Mix, 50 ng of cDNA, and 70 nM of each primer in a final volume of 25 μl. Amplifications were initiated with a 15-min enzyme activation at 95°C followed by 40 cycles of denaturation at 95°C for 15 s, primer annealing at 60°C for 30 s, and extension at 72°C for 30 s. All samples were amplified in triplicate, and the mean value was used for further calculations. Each run comprised of the products of amplifica-

tion for three control and three test samples with two primer pairs (reference and target gene) and a negative control also analyzed in triplicate. A standard curve for each primer pair was produced by serially diluting a control cDNA. Quantities of specific mRNA in the sample were measured according to the corresponding gene-specific standard curve. Raw data were analyzed with the sequence detection software (Applied Biosystems) and relative quantitation was performed within Microsoft EXCEL applying the RELATIVE EXPRESSION SOFTWARE tool (31, 32). Ratios between the target and reference gene were calculated by using the mean of these measurements. Specificity of RT-PCR products was determined by gel electrophoresis, which resulted in a single product of the desired length (RibS18, 188 bp; KPL2e29–30, 388 bp; KPL2e7–8, 126 bp). In addition, a melting curve analysis was performed allowing single product-specific melting temperatures to be determined. No primer–dimer formations were generated during the application of 40 real-time PCR amplification cycles. Differences in amplification were corrected by quantifying samples relative to the corresponding standard curves using ribosomal 18S RNA as an internal reference gene.

Southern and Northern Blotting. Porcine genomic DNA was digested with EcoRI, EcoRV, and BamHI, and separated by electrophoresis on a 0.8% agarose gel in TBE buffer and transferred onto positively charged Hybond-NX membranes (Amersham Pharmacia). The membranes were hybridized with the DNA Probe1 (1,177 bp, Table 1) labeled with EasyTides [α-^{32}P]dCTP, 250 μCi (PerkinElmer). Hybridization and washing were carried out at 65°C according to standard protocols, and the membranes were exposed to x-ray films.

Total RNA was extracted from tissues by RNeasy Protect Mini and Midi kits (Qiagen). Poly(A)+ RNA was isolated by using the Dynabeads DIRECT mRNA kit (Dynal) following the manufacturer's instructions. Each RNA sample was denatured by boiling for 10 min and loaded onto a 1% agarose-formaldehyde gel. RNA was transferred to positively charged nylon membranes by using the NorthernMax kit (Ambion). The Probe2 (*KPL2* exons 3–7, 756 bp, Table 2) was radioactively labeled with the Nick Translation System (GIBCO/BRL) using [α-^{32}P]dCTP (Amersham Pharmacia).

We thank Dr. O. Manninen for assistance with Southern blotting analysis, Dr. P. Pakarinen for assistance with Northern blot analysis, and A. Virta for sequencing analysis. This work was financially supported by The Finnish Animal Breeding Association.

1. Blacque, O. E., Perens, E. A., Boroevich, K. A., Inglis, P. N., Li, C., Warner, A., Khattra, J., Holt, R. A., Ou, G. & Mah, A. K., *et al.* (2005) *Curr. Biol.* **15,** 935–941.
2. Inaba, K. (2003) *Zool. Sci.* **20,** 1043–1056.
3. Hackstein, J. H., Hochstenbach, R. & Pearson, P. L. (2000) *Trends Genet.* **16,** 565–572.
4. Omoto, C. K., Gibbons, I. R., Kamiya, R., Shingyoji, C., Takahashi, K. & Witman, G. B. (1999) *Mol. Biol. Cell* **10,** 1–4.
5. Porter, M. E. & Sale, W. S. (2000) *J. Cell Biol.* **151,** F37–F42.
6. Pazour, G. J. & Witman, G. B. (2003) *Curr. Opin. Cell Biol.* **15,** 105–110.
7. Ansley, S. J., Badano, J. L., Blacque, O. E., Hill, J., Hoskins, B. E., Leitch, C. C., Kim, J. C., Ross, A. J., Eichers, E. R. & Teslovich, T. M., *et al.* (2003) *Nature* **425,** 628–633.
8. Katsanis, N., Lupski, J. R. & Beales, P. L. (2001) *Hum. Mol. Genet.* **10,** 2293–2299.
9. Pazour, G. J. & Rosenbaum, J. L. (2002) *Trends Cell Biol.* **12,** 551–555.
10. Van's Gravesande, K. S. & Omran, H. (2005) *Ann. Med.* **37,** 439–449.
11. Pennarun, G., Escudier, E., Chapelin, C., Bridoux, A. M., Cacheux, V., Roger, G., Clement, A., Goossens, M., Amselem, S. & Duriez, B. (1999) *Am. J. Hum. Genet.* **65,** 1508–1519.
12. Olbrich, H., Haffner, K., Kispert, A., Volkel, A., Volz, A., Sasmaz, G., Reinhardt, R., Hennig, S., Lehrach, H. & Konietzko, N., *et al.* (2002) *Nat. Genet.* **30,** 143–144.
13. Bartoloni, L., Blouin, J. L., Pan, Y., Gehrig, C., Maiti, A. K., Scamuffa, N., Rossier, C., Jorissen, M., Armengot, M. & Meeks, M., *et al.* (2002) *Proc. Natl. Acad. Sci. USA* **99,** 10282–10286.
14. Andersson, M., Peltoniemi, O., Makinen, A., Sukura, A. & Rodriguez-Martinez, H. (2000) *Reprod. Domestic Anim.* **35,** 59.
15. Sironen, A. I., Andersson, M., Uimari, P. & Vilkki, J. (2002) *Mamm. Genome* **13,** 45–49.
16. Ostrowski, L. E., Andrews, K., Potdar, P., Matsuura, H., Jetten, A. & Nettesheim, P. (1999) *Am. J. Respir. Cell Mol. Biol.* **20,** 675–683.
17. Slaugenhaupt, S. A., Blumenfeld, A., Gill, S. P., Leyne, M., Mull, J., Cuajungco, M. P., Liebert, C. B., Chadwick, B., Idelson, M. & Reznik, L., *et al.* (2001) *Am. J. Hum. Genet.* **68,** 598–605.
18. Vuoristo, M. M., Pappas, J. G., Jansen, V. & Ala-Kokko, L. (2004) *Am. J. Med. Genet. A* **130,** 160–164.
19. Boissinot, S. & Furano, A. V. (2005) *Cytogenet. Genome Res.* **110,** 402–406.
20. Thompson, J. D., Higgins, D. G. & Gibson, T. J. (1994) *Nucleic Acids Res.* **22,** 4673–4680.
21. Chodhari, R., Mitchison, H. M. & Meeks, M. (2004) *Paediatr. Respir. Rev.* **5,** 69–76.
22. Afzelius, B. A. (2004) *J. Pathol.* **204,** 470–477.
23. Okada, H., Fujioka, H., Tatsumi, N., Fujisawa, M., Gohji, K., Arakawa, S., Kato, H., Kobayashi, S., Isojima, S. & Kamidono, S. (1999) *Hum. Reprod.* **14,** 110–113.
24. Neugebauer, D. C., Neuwinger, J., Jockenhovel, F. & Nieschlag, E. (1990) *Hum. Reprod.* **5,** 981–986.
25. Zhang, H. & Mitchell, D. R. (2004) *J. Cell Sci.* **117,** 4179–4188.
26. Adams, G. M., Huang, B., Piperno, G. & Luck, D. J. (1981) *J. Cell Biol.* **91,** 69–76.
27. Horowitz, E., Zhang, Z., Jones, B. H., Moss, S. B., Ho, C., Wood, J. R., Wang, X., Sammel, M. D. & Strauss, J. F., III (2005) *Mol. Hum. Reprod.* **11,** 307–317.
28. Zhang, Z., Kostetskii, I., Moss, S. B., Jones, B. H., Ho, C., Wang, H., Kishida, T., Gerton, G. L., Radice, G. L. & Strauss, J. F., III (2004) *Proc. Natl. Acad. Sci. USA* **101,** 12946–12951.
29. Sapiro, R., Kostetskii, I., Olds-Clarke, P., Gerton, G. L., Radice, G. L. & Strauss, J. F., III (2002) *Mol. Cell. Biol.* **22,** 6298–6305.
30. Anderson, S. I., Lopez-Corrales, N. L., Gorick, B. & Archibald, A. L. (2000) *Mamm. Genome* **11,** 811–814.
31. Pfaffl, M. W. (2001) *Nucleic Acids Res.* **29,** e45.
32. Pfaffl, M. W., Horgan, G. W. & Dempfle, L. (2002) *Nucleic Acids Res.* **30,** e36.