Cell, Vol. 117, 541–552, May 14, 2004, Copyright ©2004 by Cell Press

Comparative Genomics Identifies a Flagellar and Basal Body Proteome that Includes the *BBS5* Human Disease Gene

Jin Billy Li,[1] Jantje M. Gerdes,[2] Courtney J. Haycraft,[4] Yanli Fan,[5] Tanya M. Teslovich,[2] Helen May-Simera,[6] Haitao Li,[7] Oliver E. Blacque,[5] Linya Li,[1] Carmen C. Leitch,[2] Richard Allan Lewis,[8] Jane S. Green,[9] Patrick S. Parfrey,[9] Michel R. Leroux,[5] William S. Davidson,[5] Philip L. Beales,[6] Lisa M. Guay-Woodford,[7] Bradley K. Yoder,[4] Gary D. Stormo,[1] Nicholas Katsanis,[2,3] and Susan K. Dutcher[1,*]

[1]Department of Genetics
Washington University School of Medicine
St. Louis, Missouri 63110
[2]McKusick-Nathans Institute of Genetic Medicine
[3]Wilmer Eye Institute
Johns Hopkins University
Baltimore, Maryland 21205
[4]Department of Cell Biology
University of Alabama at Birmingham
Birmingham, Alabama 35294
[5]Department of Molecular Biology and Biochemistry
Simon Fraser University
Burnaby, British Columbia
Canada BC V5A 1S6
[6]Molecular Medicine Unit
Institute of Child Health
University College London
London WC1N 1EH
United Kingdom
[7]Division of Genetic and Translational Medicine
Departments of Medicine, Pediatrics, and Genetics
University of Alabama at Birmingham
Birmingham, Alabama 35294
[8]Departments of Molecular and Human Genetics, Ophthalmology, Pediatrics, and Medicine
Baylor College of Medicine
Houston, Texas 77030
[9]Departments of Medical Genetics and Clinical Epidemiology
Memorial University
St. John's, Newfoundland
Canada

Summary

Cilia and flagella are microtubule-based structures nucleated by modified centrioles termed basal bodies. These biochemically complex organelles have more than 250 and 150 polypeptides, respectively. To identify the proteins involved in ciliary and basal body biogenesis and function, we undertook a comparative genomics approach that subtracted the nonflagellated proteome of *Arabidopsis* from the shared proteome of the ciliated/flagellated organisms *Chlamydomonas* and human. We identified 688 genes that are present exclusively in organisms with flagella and basal bodies and validated these data through a series of in silico, in vitro, and in vivo studies. We then applied this resource to the study of human ciliation disorders and have identified *BBS5*, a novel gene for Bardet-Biedl syndrome. We show that this novel protein localizes to basal bodies in mouse and *C. elegans*, is under the regulatory control of *daf-19*, and is necessary for the generation of both cilia and flagella.

*Correspondence: dutcher@genetics.wustl.edu

Introduction

Cilia and flagella are central to many developmental processes in a diverse range of organisms, and disruption of their structure or function can have profound phenotypic consequences (Pazour and Witman, 2003). Among vertebrates, the primary cilium is virtually ubiquitous, comprising a "9+0" axoneme that consists of nine outer doublet microtubules. Centrioles can interconvert to basal bodies by assembling a transition zone, which is essential for the assembly of cilia or flagella. The transition zone serves as a docking site for intraflagellar transport (IFT) proteins and their motors, kinesin II (Cole et al., 1998; Piperno and Mead, 1997). Basal bodies can also convert to centrioles that serve as the microtubule-organizing center (MTOC) in animal and algal cells. However, land plants, fungi, and diatoms do not contain centrioles, which suggests that these organisms may have evolved other structures to organize microtubules (Pickett-Heaps, 1971). In organisms that retain centrioles, these structures are essential and play key roles (Bobinnec et al., 1999). Their removal suggests that centrioles are important to complete cytokinesis and are essential for the transition from G1 to S phase of the cell cycle (Hinchcliffe et al., 2001; Khodjakov and Rieder, 2001). Missegregation or overduplication of centrioles is associated with genomic instability (Meraldi and Nigg, 2001). Notably, neoplastic cells have increased numbers of centrioles and abnormal centriole architecture, which is similar to the morphologies observed in mutant Chlamydomonas strains (Dutcher, 2003a; Lingle and Salisbury, 1999).

Cilia and basal bodies have been implicated directly in several developmental processes including left-right asymmetry, heart development, maintenance of the renal epithelium, respiratory function, electrolyte balance in the cerebrospinal fluid, and reproductive fecundity. Nodal cilia in early development are fundamental to initiate the molecular cascade that ultimately leads to the visceral asymmetry in vertebrates (Essner et al., 2002; Tabin and Vogan, 2003; Takeda et al., 1999). A lack or dysfunction of sensory cilia in the epithelium of renal tubules underlies polycystic kidney disease, although the exact process of cyst formation is not understood (Haycraft et al., 2001; Pazour et al., 2000). In the retina, successful translocation of several molecules from the cell body to the outer segment of the photoreceptor depends on transport along a modified cilium; defective transportation of certain proteins results in the retinal

degeneration associated with retinitis pigmentosa (Marszalek and Goldstein, 2000). Although the exact role has not been elucidated, sensory cilia present on neuronal cells have somatostatin receptors embedded in their ciliary membrane (Handel et al., 1999). Defective basal bodies have been implicated in the pathogenesis of the pleiotropic Bardet-Biedl syndrome (BBS). This multisystemic disorder is characterized primarily by retinal dystrophy, obesity, polydactyly, renal and genital malformations, and learning disabilities. Six genes associated with BBS (*BBS1, BBS2, BBS4, BBS6/MKKS, BBS7*, and *BBS8*) have been identified (Ansley et al., 2003; Badano et al., 2003a; Katsanis et al., 2000; Mykytyn et al., 2001, 2002; Nishimura et al., 2001; Slavotinek et al., 2000).

Centrioles and flagella are biochemically complex organelles. The flagellar axoneme of *Chlamydomonas* contains over 250 polypeptides (Dutcher, 1995a), whereas the surrounding flagellar membrane and matrix together contain over 50 polypeptides (Witman et al., 1972; our unpublished data). Preparations enriched in *Chlamydomonas* basal bodies show over 150 polypeptides (Dutcher, 1995b).

We undertook a comparative genomics approach to identify the eukaryotic flagellar/ciliary and basal body proteome. Our strategy rests on two observations. First, *Chlamydomonas* flagellar and basal body proteins show high similarity to human orthologs. Second, organisms that have lost flagella and basal bodies through evolution have also lost most of the genes needed for building and regulating these organelles (Cole, 2003; Dutcher, 2004; Perrone et al., 2003). Cumulatively, our data provide a comprehensive description of the conserved flagellar and basal body proteome and a rich platform to understand the function of these cellular structures.

Results

The recently released genome sequence of *Chlamydomonas reinhardtii* has approximately 95 Mb of sequence and encodes 19,832 predicted proteins (http://shake.jgi-psf.org/chlre2/chlre2.home.html). These predicted proteins were compared to the 32,035 predicted proteins of the human genome using WU-BLASTP (Lopez et al., 2003) with a cutoff E value of 10^{-10}. This comparison yielded a dataset of 4,348 genes (Figure 1), which were then compared to the 28,786 predicted proteins of *Arabidopsis*. All proteins with a match in *Arabidopsis* were removed from consideration, which leaves only the proteins both present in organisms with flagella and basal bodies and absent from an organism that lacks flagella and basal bodies. This procedure was repeated using the genomes of *Caenorhabditis elegans*, *Ciona intestinalis*, *Drosophila melanogaster*, and *Mus musculus*. The comparisons with each organism yield between 400 and 688 proteins (Table 1). Using human, mouse, and *Ciona* in combination yielded 504 proteins, which suggests that many of the same genes are found in all of the organisms tested (Table 1). The datasets generated with *Caenorhabditis elegans* and *Drosophila melanogaster* are smaller (400 and 481, respectively) than the human, mouse, or *Ciona* sets. We will refer to the *Chlamydomonas* + Human − *Arabidopsis* collection of proteins as the flagellar apparatus-basal body (FABB) proteome

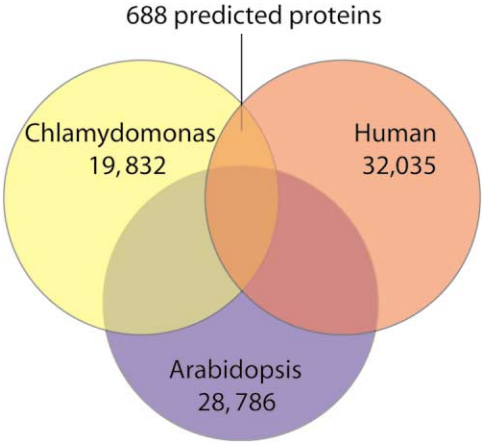

Figure 1. 688 Proteins in the Flagllear Apparatus Basal Body (FABB) Proteome

Venn diagram of the comparative approach to enrich for flagellar and basal body proteins. The proteome of *Chlamydomonas* (yellow) was compared to the proteome of human (red) by WU-BLASTP to find all matches with a cutoff E value of 10^{-10}. These 4,348 matches were then compared to the proteome of Arapidopsis (blue) by WU-BLASTP to remove all matches with a cutoff E value of 10^{-10}. The 688 proteins remaining are designated the FABB proteome (in light orange).

(Supplemental Table S1 at http://www.cell.com/cgi/content/full/117/4/541/DC1).

If the subtracted dataset were enriched for flagellar and basal body proteins, it should meet several criteria. First, it should contain many of the known genes. Second, the domain characteristics of the proteome should be consistent with the domains of the known flagellar and basal body proteins. Third, a substantial fraction of the genes should be upregulated after deflagellation. Fourth, ablation of the genes or messages should cause flagellar and basal body phenotypes. Finally, the proteome should contain novel genes involved in disorders of human ciliation. We have tested each of these predictions.

Validation with Known *Chlamydomonas* Flagellar and Basal Body Genes

We used 58 well-studied genes that encode flagellar proteins or components of the IFT complex in *Chlamydomonas* (Supplemental Table S2 on the *Cell* website). This set of test genes excludes tubulins, kinesins, and calcium binding proteins because these proteins have required roles in all eukaryotes. Thus, any gene retained in *Arabidopsis* for another function should not be found in the FABB proteome. Fifty-two of the fifty-eight known flagellar and IFT genes (90%) are present in the FABB proteome. Six genes that localize or are required for the assembly of the basal body of *Chlamydomonas* were examined, and four are in the FABB proteome (67%).

The similarity and identity of these 64 *Chlamydomonas* proteins were compared to the best matches from human, mouse, *Ciona*, *Drosophila*, *C. elegans*, and *Arabidopsis* (Supplemental Table S2 online). Overall, human and mouse matches show more significant E values with *Chlamydomonas* proteins than with *Drosophila* and

Table 1. Comparison of Flagellated and Nonflagellated Organism Proteomes

Organisms in Comparison	No. Genes	Percentage of Known *Chlamydomonas* Flagellar Genes (n = 41)	Percentage of Known *Chlamydomonas* IFT Genes (n = 17)	Percentage of Known *Chlamydomonas* Basal Body Genes (n = 6)
Chlamydomonas + Human − *Arabidopsis*	688	88	94	67
Chlamydomonas + Mouse (32,911) − *Arabidopsis*	671	86	94	67
Chlamydomonas + *Ciona* (15,852) − *Arabidopsis*	668	91	94	50
Chlamydomonas + *Drosophila* (18,498) − *Arabidopsis*	481	83	88	33
Chlamydomonas + *C. elegans* (22,221) − *Arabidopsis*	400	59	94	0
Chlamydomonas + Human + Mouse − *Arabidopsis*	611	86	94	67
Chlamydomonas + Human + Mouse + *Ciona* − *Arabidopsis*	504	86	94	50

C. elegans. Importantly, *C. elegans* lacks motile cilia and has simpler basal bodies with singlet microtubules. Proteins that are present in *Chlamydomonas* and *Drosophila* but absent from *C. elegans* may be proteins needed for motile flagella and more complex basal bodies. There are 326 proteins of the FABB proteome that are not found in *C. elegans* (Supplemental Table S1). Consistent with this prediction, this set includes the known radial spoke, central pair, and dynein arm proteins, which are needed for motility.

In addition to established *Chlamydomonas* basal body and flagellar genes, the FABB proteome contains several known mammalian disease genes (Supplemental Table S1). Five of the six genes that result in Bardet-Biedl syndrome in humans appear in the dataset (Ansley et al., 2003; Badano et al., 2003a; Mykytyn et al., 2001, 2002; Nishimura et al., 2001). Two genes that cause kidney disease in humans, NPHP4 (Mollet et al., 2002) and fibrocystin (Ward et al., 2002), are also present in the FABB proteome. Fibrocystin localizes to primary cilia of the kidney (Wang et al., 2004). Hydin, which causes hydrocephalus in mice and localizes to cilia, is present in the FABB proteome (Davy and Robinson, 2003). IFT172 and IFT88 are in the FABB proteome. Mice with mutations in these proteins show neural tube patterning defects that act through the hedgehog signaling pathway (Huangfu et al., 2003). IFT88 is also needed for proper kidney function (Pazour et al., 2002; Taulman et al., 2001). A tubby superfamily member (TUSP) is also present (Li et al., 2001).

Validation with the Distribution of Domains

The domain/motif distribution of the human orthologs in the FABB proteome was analyzed by the algorithm Pfam (Bateman et al., 2004) and compared to the motifs found in 785 randomly selected human proteins not present in the FABB proteome (Figure 2; Supplemental Table S3). Consistent with the predicted enrichment for flagellar and basal body proteins, FABB is essentially devoid of Zn-finger, ring finger, bHLH, and immunoglobulin motifs. By contrast, dynein heavy chain proteins, ion channels, and domains involved in cyclic nucleotide metabolism, which are associated with flagella, basal bodies, and ciliated neurons, are enriched (Anholt and Rivers, 1990; Pasquale and Goodenough, 1987; Wong et al., 2000). Intriguingly, we found a significant overrepresentation of tetratricopeptide repeat motifs (TPRs) in the FABB (14 to 1). TPR domains are found in several IFT and BBS proteins, which suggests that this domain might be a useful a priori determinant of flagellar or basal body function. We also observed significant enrichment for WD repeat-containing proteins (21 to 7) and IQ calmodulin binding motifs that may represent important domains for protein-protein interactions in the basal body and flagella.

Validation with Transcriptional Regulation of Flagellar Genes

Amputation of the flagella in *Chlamydomonas* is followed by transcription of flagellar genes. Consequently, a 3- to 12-fold upregulation of the mRNAs for structural components of the flagella, the IFT complex, and its molecular motors has been observed 30 min after deflagellation (Lefebvre et al., 1980). Therefore, the genes in the FABB proteome that are transcriptionally upregulated following deflagellation are likely to encode flagellar proteins. RSP3, α-tubulin, IFT57, and D2LIC serve as positive controls and δ-tubulin and ϵ-tubulin are genes that show no upregulation (Supplemental Table S4 online). We found 38% (n = 103) of the genes tested to be upregulated 3- to 13-fold (Figure 3; Supplemental Table S4) and to be likely to play roles in flagellar assembly and function. Twenty-two genes were also tested at 10 min and at 50 min after deflagellation. Fifteen of the genes that were not upregulated at 30 min also show no change at either 10 or 50 min. For the seven genes that were upregulated, the peak levels were at 30 min and not at 10 or 50 min.

To determine what fraction of genes not in the FABB proteome is upregulated by deflagellation, ten genes that are present in *Chlamydomonas*, human, and *Arabidopsis* were analyzed. This set should not be enriched for flagellar and basal body genes. Only one of the ten, 167087, is upregulated. 167087 is the homolog of SRP54 (signal recognition particle); this upregulation may reflect a need for increased protein sorting upon deflagellation.

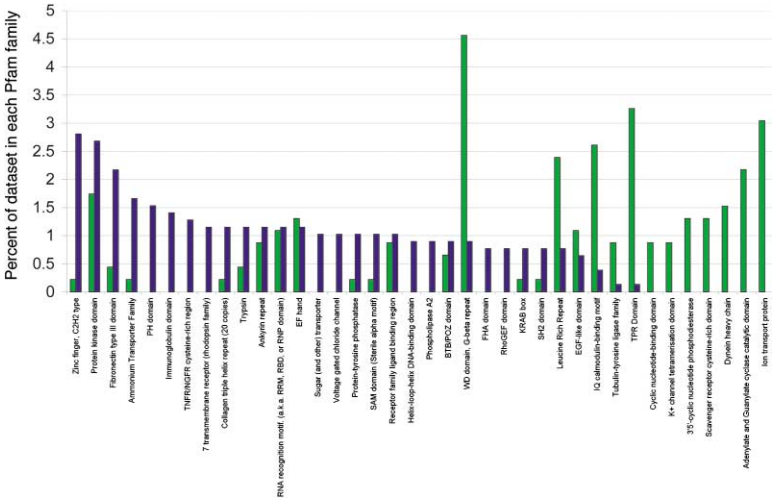

Figure 2. Some Pfam Domains Are Enriched in the FABB Proteome

Histogram of domains found in the human matches of the FABB (green; n = 461) compared to nonselected human proteins (blue; n = 785). The y axis indicates the percentage of genes containing the domains listed on the x axis. Many of the proteins (n = 227) had no Pfam domain. (See Supplemental Table S3 online for list and statistical evaluation using Fisher's exact p value).

Validation of Basal Body Genes by RNA Interference

RNA interference (RNAi) has been used successfully in *Chlamydomonas* to lower the levels of several mRNAs and thereby to produce mutant phenotypes (Huang and Beck, 2003; Pfannenschmid et al., 2003; Sineshchekov et al., 2002; Wu-Scharf et al., 2000). Constructs expressing short (150–200 bp) hairpins loops were transformed into wild-type *Chlamydomonas*. Multiple transformants were examined to generate an allelic series.

To determine if the genes in the FABB proteome that were not upregulated after deflagellation represented false positives or bona fide basal body candidates, genes for six predicted proteins (171684, 171242, 158402, 162018, 165684, and 169453) that do not show upregulation were tested. The phenotypes of four to nine transformants for each gene were examined. We observed flagellar phenotypes in the RNAi strains for five of the six genes (Table 2). These phenotypes ranged from slowly swimming, which resembled the defect observed in inner dynein arm mutant strains, to paralyzed flagella, to lack of flagella altogether. Previously studied

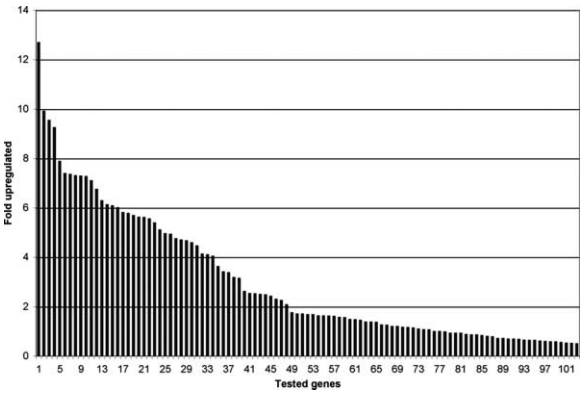

Figure 3. Many Genes of the FABB Proteome Show Upregulation

Histogram of the fold-upregulation of mRNAs for 103 genes assayed before and 30 min after deflagellation. Thirty-nine genes show greater than 3-fold upregulation; these are likely to be flagellar genes.

Chlamydomonas strains with defects in basal body assembly show a constellation of phenotypes (Dutcher, 2003a). Many of these mutant strains lack flagella or have abnormal numbers of flagella. Other phenotypes include incomplete basal body microtubules, disorganized cytoplasmic microtubules known as rootlet microtubules, and disorganized centrin fibers. Disorganized rootlet microtubules and centrin fibers result in misplaced cleavage furrows (Ehler et al., 1995). Three strains (171242, 158402, and 162018) show obvious defects in cleavage furrow placement (Table 2). The level of reduction in the specific mRNAs was quantified by real-time RT-PCR in duplicate experiments. Interestingly, differences in phenotypes among transformants are reflected by differences in the degree of mRNA reduction. For example, three independent 171242 transformants have different phenotypes. 171242-T1 cells completely lack flagella and have a defect in cleavage furrow placement; the mRNA is reduced 7.8-fold. 171242-T2 cells show a weaker phenotype in that they are only 59% aflagellate and do not have a cleavage furrow defect; the mRNA is reduced only 3.4-fold. Finally, 171242-T3 cells show the weakest phenotype and the mRNA is reduced only 1.4-fold. In general, the more severe phenotypes are observed in cells with greater reduction in mRNA for most of the different sets of transformants that show variable phenotypes (Table 2).

Only one of six proteins chosen for analysis has a clear ortholog. 171684, which shows no phenotype, is the homolog of echinoderm microtubule-associated protein. Perhaps, greater reductions in level of message result in lethality so that no transformants with a strong phenotype were identified. Protein 171242 has a cobalamin adenosyltransferase domain whose function in flagellar and basal bodies is unknown. Protein 162018 has a NIPSNAP1 domain, and its function is unclear. The remaining three proteins do not contain recognizable domains or motifs (Bateman et al., 2004).

Identifying Human Disease Genes with the FABB Proteome

Recent data suggest that several genes that cause Bardet-Biedl syndrome encode proteins involved in

Table 2. Phenotypes Observed with RNA Interference for Potential Basal Body Genes

	Phenotypes of Representative Transformants N = 200 Cells				
Chlamydomonas Protein	Name	Two Flagella (%)	Zero Flagella (%)	Other Phenotypes	Fold Reduction in mRNA (Ave. of 2 Samples)
171684 Echinoderm microtubule-associated protein	T1	95	5	None	1.7
171242 Cobalamin adenosyltransferase domain	T1	0	100	Cleavage furrow defect	7.8
	T2	41	59	Paralyzed flagella	3.4
	T3	75	25	Short, jerky motility	1.4
158402 No known Pfam domain	T3	0	100	Cleavage furrow defects	3.3
	T2	95	5	Slow smooth swimming	1.9
	T1	95	5	Slow, smooth swimming	1.3
162018 NIPSNAP1 domain	T2	0	100	Cleavage furrow defects	2.7
	T4	0	100	Aflagellate	2.0
	T1	33	67		1.3
	T3	20	80		1.2
165684 No known Pfam domain	T1	0	100		1.3
	T3	96	4	Slow, smooth swimming	1.4
	T2	96	4		0.9
169453 No known Pfam domain	T2	54	46	Paralyzed flagella	2.2
	T3	3	97		1.4
	T1	95	5	Paralyzed flagella	1.3

basal body/flagellar assembly and function (Ansley et al., 2003). The FABB proteome contains orthologs of all known *BBS* genes except for *BBS6*, which appears to have evolved specifically in the mammalian lineage. Therefore, we hypothesized that genes in our dataset might offer useful candidates for identifying novel *BBS* genes and would further validate our approach.

BBS5 maps to human chromosome 2q31 between *D2S156* and *D2S1238*; the region spans 14 Mb and contains approximately 230 predicted genes (Young et al., 1998). Two genes in this interval are present in the FABB proteome. They are FLJ11457 (NM_024753), which encodes IFT139 (Cole, 2003), and DKFZp762I194 (NM_152384), which encodes a well-conserved, novel protein of unknown function. The complete open reading frame and intron-exon junctions of FLJ11457 in a patient from Newfoundland family NFB9 that defined the *BBS5* locus (Young et al., 1998) was sequenced and yielded several polymorphisms, but no potentially pathogenic mutations. We then sequenced each of the 12 coding exons of DKFZp762I194 and identified a homozygous A→G transition at the +3 position of the exon 6 splice donor that segregates with the BBS phenotype in the extended family. The two available affected individuals are homozygous G/G, both parents are heterozygous A/G, and four unaffected sibs are either homozygous A/A or heterozygous A/G (Figure 4A). Because this allele segregates with the disease and it is absent from 100 Newfoundland control chromosomes, its effect in a lymphoblastoid cell line from one of the NFB9 patients was investigated. Amplification of exons 4 to 9 of DKFZp762I194 in a control lymphoblastoid cell line and subsequent cloning and sequencing of the PCR products revealed two RNA splice isoforms; one lacks the 63 bp exon 8 (Figure 4B). In contrast, patient NFB9 produces three RNA species. One species lacks exon 6, one species lacks exons 6 and 8, and the third uses a cryptic splice site (/GTA) in intron 5 (Figure 4B). All three messages result in a frameshift and premature termination in exon 7. Therefore, we conclude that this intronic mutation compromises the splicing of DKFZp762I194 and its protein product. Thus, DKFZp762I194 is a strong candidate for *BBS5*.

To investigate further the candidacy of this gene, we genotyped 24 Saudi Arabian consanguineous families with markers encompassing each of the known *BBS* loci. We determined that in one of these families, KK63, both affected sibs, but none of the unaffected individuals, were homozygous by descent (HBD) across the *BBS5* critical interval. The other seven *BBS* loci were excluded by haplotype analysis. Sequencing DKFZp762I194 in this family revealed a premature termination mutation $L_{142}X$ in exon 6 that is not present in 166 ethnically matched control chromosomes (Figure 4C). This mutation may lead to nonsense-mediated decay of the message. The combined data from these two families support the hypothesis that the candidate gene DKFZp762I194 is *BBS5*.

We next ascertained the contribution of *BBS5* mutations to the pool of BBS patients by screening a combined cohort of 259 independent BBS families from various ethnic backgrounds. We identified likely pathogenic mutations in two families. In consanguineous family PB108 of Turkish descent, we identified an insertion-deletion mutation, 263–271 indelGCTCTTA, which has the net effect of removing a single base and resulting in a premature termination codon (Figure 4D). In a second family, PB127 of Kurdish descent, the patient harbors a homozygous nonsense mutation $W_{59}X$, which by conceptual translation will lead to premature termination (Figure 4E). These data suggest that *BBS5* contributes approximately 2% to the BBS mutation pool; this estimate is consistent with the contribution of most other *BBS* loci (Ansley et al., 2003; Badano et al., 2003b; Beales et al., 2001; Katsanis et al., 2002).

BBS5 May Interact Genetically with *BBS1*

We have shown previously that some families segregate three mutations in two *BBS* genes and that the presence of the third mutation may be a modifier of either pene-

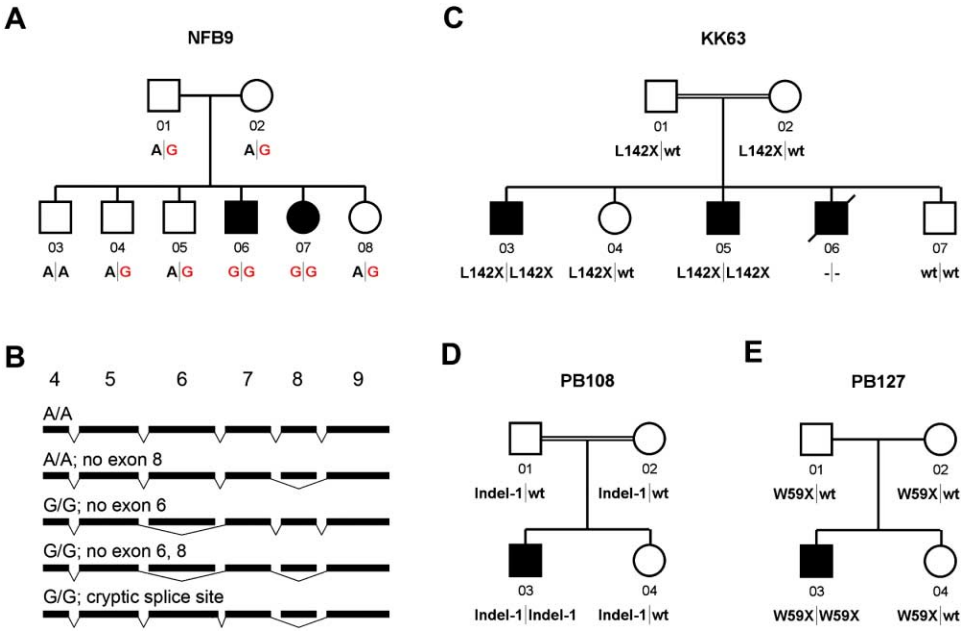

Figure 4. DKFZp762I194 Is *BBS5*

(A) Newfoundland family NFB9 segregates an A→G transition at the +3 position of the exon 6 splice donor; genotypes are shown below each individual.

(B) Diagram of exons 4 to 9 that were amplified by RT-PCR and sequenced. Unaffected (A/A) individuals have an mRNA with exons 4–9 and an alternatively spliced mRNA that lacks exon 8. Affected members (G/G) of family NFB9 have three predominant mRNA species. One lacks exon 6, one lacks exons 6 and 8, and one starts at a cryptic splice site in intron 5.

(C, D, and E) Segregation of mutations in three families; each mutation is predicted to result in premature termination. Families KK63 (Saudi) and PB108 (Turkish) are consanguineous; consanguineity is also suspected for the Kurdish family PB127.

trance (Katsanis et al., 2001) or expressivity (Badano et al., 2003b). We queried whether *BBS5* also participates in complex inheritance. We identified four outbred Caucasian pedigrees that contain single, heterozygous mutations in *BBS5*. Two families, AR618 and AR711, harbor a single $R_{207}H$ mutation in exon 8. The $R_{207}H$ variant was not detected in 288 control chromosomes and is conserved in mammals and fish. The degree of conservation suggests that H207 may have a modifying effect on the BBS phenotype. However, neither patient has mutations in the known *BBS* loci. Two other families, AR291 and AR381, have an $N_{184}S$ variant in exon 7. One of the families harboring the $N_{184}S$ allele is homozygous for the $M_{390}R$ *BBS1* mutation, the most common disease-causing *BBS1* allele (Figure 5A) (Beales et al., 2003; Mykytyn et al., 2001, 2002). Importantly, the N184 residue lies in a predicted repeat domain, DM16, which is conserved in both copies of the domain in ten species from Trypanosome to *Homo sapiens* (Figure 5B). Consistent with a potentially modifying but not causal contribution, we found the $N_{184}S$ allele in 1 of 388 control chromosomes.

The Search for X Boxes

In the nematode *C. elegans*, 60 of the 302 neurons in hermaphrodites have cilia and multiple ciliated neurons are located in the male tail (Ward et al., 1975; White et al., 1976; Wicks et al., 2000). Several genes needed for functional cilia have a 14 nucleotide sequence known as the X box, which is found in the first 150 bp upstream of the ATG initiation codon. This motif is thought to be the binding site for the RFX transcription factor DAF-19 (Schafer et al., 2003; Swoboda et al., 2000). We searched the *C. elegans* genes in the FABB proteome for this motif with the algorithm PATSER (Hertz and Stormo, 1999). Twenty-five of the 400 *C. elegans* genes contain the X box motif (Supplemental Table S1). They include five IFT genes and the dynein intermediate light chain as well as *BBS1, 2, 7*, and *8* homologs (Ansley et al., 2003; Haycraft et al., 2001; Perrone et al., 2003). Ten of these genes encode unknown proteins as well; one of these is *BBS5*. When 500 bp upstream of each of the 22,221 *C. elegans* genes is searched, 221 genes have this motif (data not shown). Thus, the FABB proteome is enriched for *C. elegans* genes containing the X box motif by about 6-fold.

Expression and Localization of BBS-5

The presence of an X box in the promoter of *bbs-5* suggests that it will be expressed specifically in ciliated cells and that it will be regulated by DAF-19. To test this possibility, a transcriptional reporter construct that consists of 500 bp of upstream sequence and the first two exons of *bbs-5* was fused in-frame to the *gfp* coding region and used to generate transgenic worms. In agreement with a conserved role for BBS-5 in basal bodies at the base of cilia, GFP fluorescence was detected in all ciliated sensory neurons including the amphids, labial neurons, phasmids, and the sensory rays of the male tail (Figures 6A–6C and data not shown). No GFP was evident in other cell types, suggesting that *bbs-5* expression is restricted to ciliated cells.

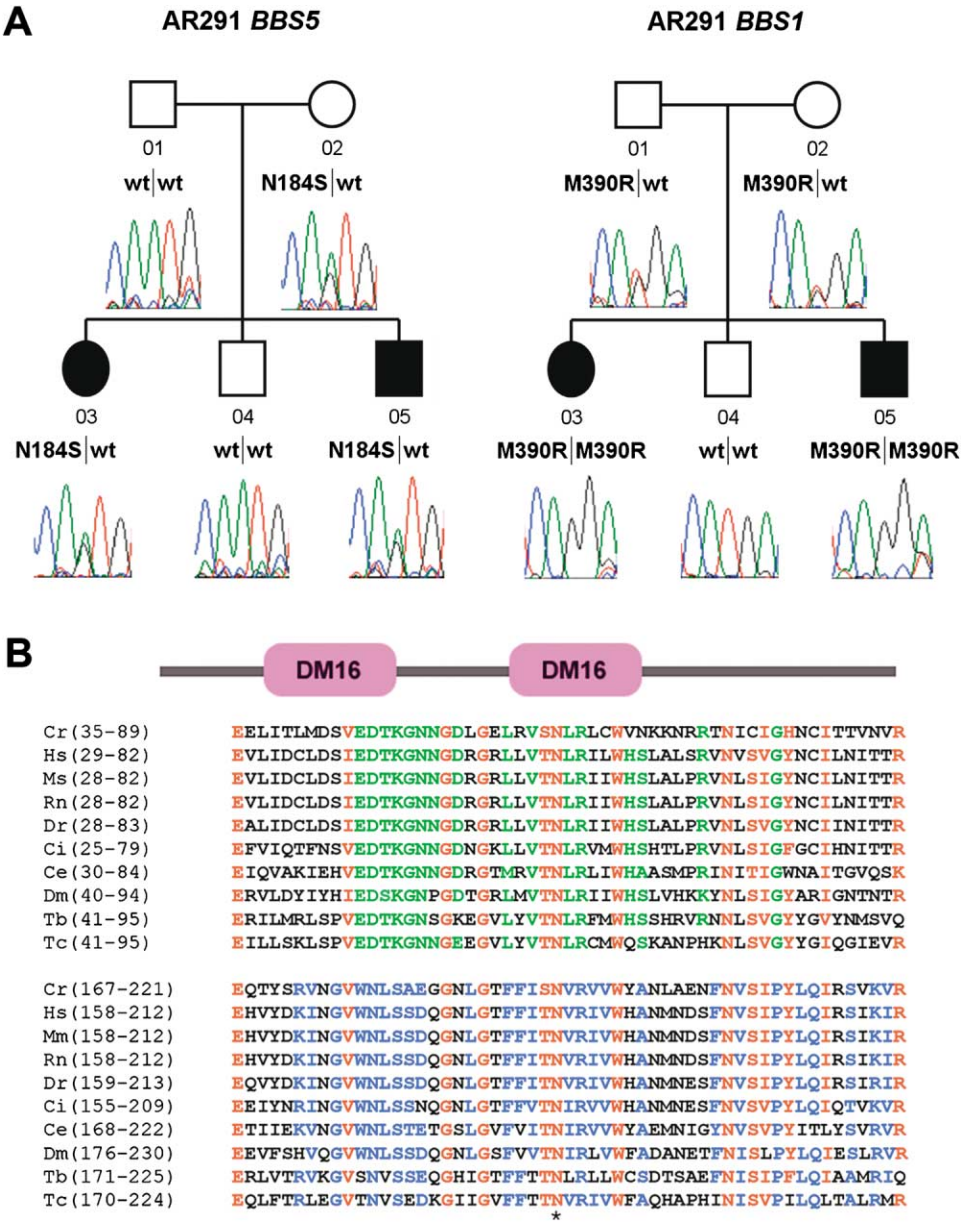

Figure 5. *BBS5* May Interact Genetically with *BBS1*

(A) Both patients from North American family AR291 are homozygous for the $M_{390}R$ mutation in *BBS1* and are heterozygous for the $N_{184}S$ mutation in *BBS5*.

(B) Schematic of the positions of the DM16 repeat domains of BBS5 and alignment of both DM16 domains contained in each BBS5 ortholog across ten species. Red depicts similarity of at least 90% in both domains, green depicts similarity of at least 80% in domain 1, and blue depicts similarity of at least 80% in domain 2. The N184 residue involved in family AR291 is conserved (asterisk). Cr = *Chlamydomonas reinhardtii*; Hs = *Homo sapiens*; Mm = *Mus musculus*; Rn = *Rattus norvegicus*; Dr = *Danio rerio*; Ci = *Ciona intestinalis*; Ce = *Caenorhabditis elegans*; Dm = *Drosophila melanogaster*; Tb = *Trypanosoma brucei*; Tc = *Trypanosoma cruzi*. Amino acid positions within each sequence are given in parentheses.

To confirm a role for DAF-19 in regulating *bbs-5* transcription, the expression of *bbs-5* in wild-type and *daf-19* mutant worms was analyzed in vivo and by Northern blot analysis. *daf-19(m86)* transgenic lines carrying the *bbs-5::gfp* construct as an extrachromosomal array were mated to *daf-19(+)* worms. The resulting progeny were assayed for DAF-19 function based on the inability of *daf-19(m86)* worms to stain with DiO, a lipophilic fluorescent dye normally absorbed by a subset of ciliated sensory neurons (Figures 6F and 6G) (Fujiwara et al., 1999). Similar to the results reported for *osm-5*, a known DAF-19 transcriptional target, expression of the *bbs-5::gfp* reporter was markedly reduced but not abolished in the absence of DAF-19 (compare Figures 6D and 6E) (Haycraft et al., 2001). Northern blot analysis detected a 1.2 kb transcript that agrees with the predicted mRNA in WormBase (http://www.wormbase.org). The amount of the 1.2 kb transcript was reduced in *daf-*

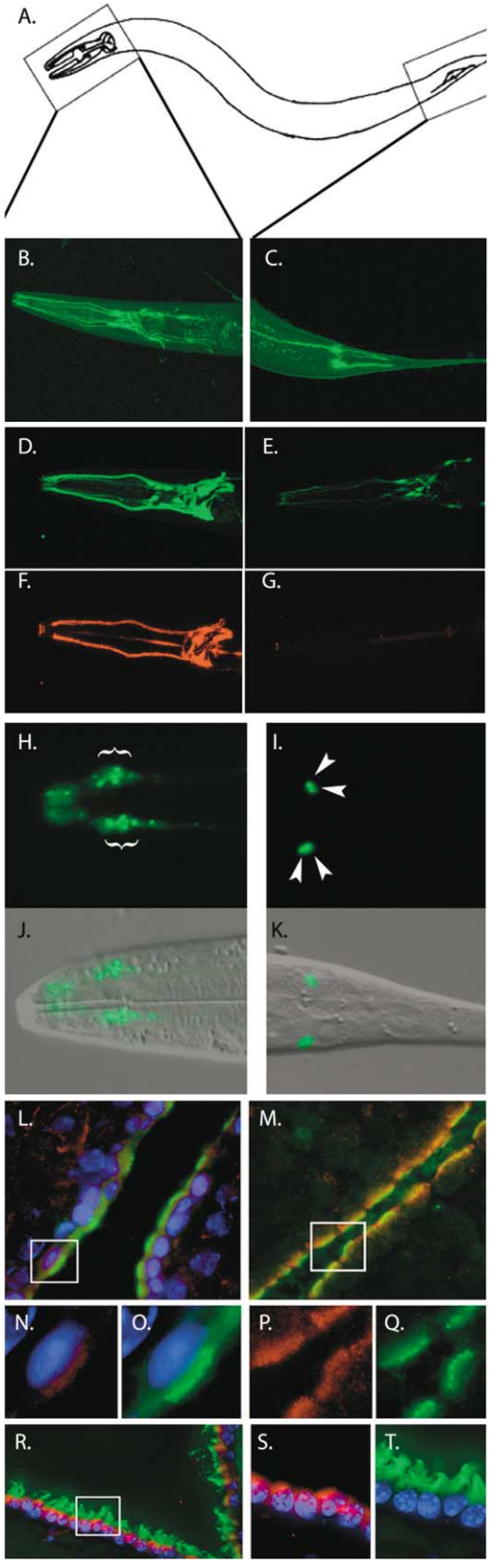

Figure 6. BBS-5 Expressed in Ciliated Cells

(A) Schematic of the adult hermaphrodite. The boxed areas indicate the regions shown in (B) and (C).

19 mutant worms, which confirms the in vivo expression analysis (data not shown).

Analysis of a *C. elegans* strain expressing a translational *bbs-5::GFP* transgene revealed that the GFP-tagged BBS-5 protein localizes specifically to the base of cilia in the ciliated head and tail neurons. BBS-5::GFP staining pattern is observed at the base of the cilia in the amphid neurons in the worm head and at the base of the left and right phasmid neurons in the tail (Figures 6H and 6I).

To explore the function of BBS5 further, a rabbit anti-peptide antibody was generated to the predicted mouse BBS5 protein. On immunoblots, the antibody recognizes two proteins with apparent molecular weights of 37 and 40 kDa in extracts from cultured IMCD cells and whole mouse (E10.5) embryos (Supplemental Figure S1). These sizes are consistent with the proteins predicted by the conceptual translation of the mouse cDNA and alternatively spliced isoform (Figure 4B). Importantly, neither of these products is detected on immunoblots probed with the preimmune sera.

In agreement with the localization of the BBS-5::GFP in *C. elegans*, prominent immunofluorescent staining was detected at the apical surface of the multiciliated ependymal cells lining the ventricles of the brain (Figures 6L–6T). The localization to basal bodies was confirmed by colocalization of BBS5 with γ-tubulin, a centrosomal and basal body marker, and the localization of BBS5 just beneath the cilia as detected by antibodies to acetylated α-tubulin. The domain of BBS5 staining was also found to extend beneath that of γ-tubulin signal, indicating that BBS5 is also localized to regions surrounding the basal bodies. Faint staining was detected in the cilium, although it was difficult to confirm that this signal was significantly different from that of the background.

RNA Interference of the *BBS5* Homolog in *Chlamydomonas*

Multiple *Chlamydomonas* transformants were obtained with a hairpin construct for *BBS5* (AY581824). Among nine transformants, four completely lack flagella and show a weak cleavage furrow defect. The remaining five

(B and C) Confocal images showing bbs-5::gfp expression in the amphid and labial neurons and phasmids of the adult hermaphrodite. Anterior is toward the left in (A)–(K).

(D–G) Confocal images of bbs-5::gfp expression (D and E) and DiO uptake (F and G) in wild-type (D and F) and mutant (E and G) worms. Expression of bbs-5::gfp is markedly reduced in *daf-19(m86)* mutant worms (compare D and E). Similar results were observed with four independent transgenic strains.

(H–K) Images showing the GFP fluorescence patterns from strains carrying a translational (BBS-5::GFP) transgene; shown are the GFP staining patterns in the amphid/labial neuron (H) and phasmid neuron (I) basal bodies, along with DIC-merged images (J and K).

(L–T) Confocal images of the ependymal cells of the lining of the mouse ventricles. (L, N, and O) Antibodies to γ-tubulin are in green, preimmune serum is red, and DNA is blue. (M, N, and Q) Antibodies to γ-tubulin are in green, antibodies to mouse BBS5 are in red, and DNA is in blue. The boxes in (L) and (M) indicate the regions enlarged in (N), (O), (P), and (Q). The box in (R) indicates the region enlarged in (S) and (T). Antibodies to BBS5 are in red, antibodies to acetylated α-tubulin are in green, and DNA is in blue (R, S, and T). BBS5 is found at the base of the cilia.

transformants are partially aflagellate. The severity of phenotype correlates with the fold-reduction in mRNA levels. In the aflagellate strains with the cleavage furrow defect, the mRNA was reduced 2.2- to 3.1-fold. In contrast, the cells that show a partial flagellar assembly defect have a 1.5- to 1.7-fold reduction. These results strongly support a role for Bbs5 in flagellar and basal body function and assembly.

Discussion

The study of model organisms has elucidated many critically important conserved processes that include the cell division cycle, apoptosis, and development. We have recently recognized the importance of cilia in diverse human diseases. To understand and to treat cilia-based diseases, we must understand the assembly and the function of flagella/cilia and basal bodies (Pazour and Rosenbaum, 2002). *Chlamydomonas* and *C. elegans* have been opportune model organisms to study flagella/cilia and centrioles/basal bodies. In this work, we have exploited the observation that *Chlamydomonas* and many metazoans have flagella/cilia and basal bodies while *Arabidopsis* lacks these structures to identify by comparative genomics a set of 688 proteins that is enriched for components of the flagella and basal body. Several methods have demonstrated that the dataset is highly enriched in components of the flagella and centriole/basal body. Our FABB proteome contains 88% of established *Chlamydomonas* flagellar and basal body genes. Of the eight known genes not in the FABB proteome, five are absent from the human proteome. One was present (IFT20) but, perhaps due to its small size, its E value was not significant. Two of the missing proteins (Oda11 and Pf16) have significant matches to human homologs, but missed the imposed cutoff E value of 10^{-10} (Supplemental Table S2). This reflects a potential caveat of our methodology; these are two proteins that could be included if the criteria were adjusted to reflect the difference rather than the absolute E values in the flagellated and nonflagellated comparisons. PF16 is the sperm-associated antigen 6 (SPAG6) in mammals and is clearly a flagellar component. Including these two proteins would raise our identification rate to 91%.

Amputation of *Chlamydomonas* flagella induces transcription of flagellar genes (Lefebvre et al., 1980). Thirty-eight percent of the 103 genes tested show at least a 3-fold increase in mRNA levels following deflagellation and about two-thirds of these genes encode unknown proteins. We suggest that these genes are components of the flagellar axoneme, membrane, or IFT particles. Among the upregulated genes that were characterized in mammals are hydin, NYD-SP28, which is involved in testis development, a leucine-rich testis-specific protein, a testis-specific AMY-1 binding protein, and KPL2, which is induced during ciliogenesis in the trachea.

Nineteen of the genes tested showed an increase in mRNA levels between 1.5- and 3-fold. These include *BBS1, 2, 4, 5, 7*, and *8* as well as tubulin tyrosine ligase and a testis-expressed gene. This intermediate induction level would be consistent with these genes encoding components needed in the transition zone of *Chlamydomonas*. The distal portions of the transition zone are lost upon deflagellation, and modest increases in transcription may be necessary for its reassembly (Dutcher, 2004). This interpretation agrees well with the localization of BBS-5 in *C. elegans* and in BBS5 mouse (Figure 6). However, without immunoelectron microscopy, it is difficult to determine if the proteins are localized to the basal body or to the transition zone.

The remaining 45 tested genes do not show significant changes in mRNA levels following deflagellation. These may represent basal body genes. Alternatively, they may be noise in our dataset. RNAi treatment of five of the six non-upregulated strains show flagellar defects; several show cleavage furrow defects. The severity of the phenotype correlates with the reduction in mRNA observed after transformation with RNAi constructs. These results suggest that this class is enriched for basal body genes.

One potential drawback to our dataset is the low abundance of coiled-coil proteins. These proteins have low information content since the structure may be more important than the exact sequence of the protein. Other proteins possibly missing from the FABB proteome include proteins that have functions in processes other than the flagella and basal bodies. Although the newer members of the tubulin superfamily, δ-, ϵ-, and η-tubulin clearly play a role in basal body function and assembly (Dutcher, 2003b), they are subtracted from our dataset by α-, β-, and γ-tubulin just as are kinesin- and calmodulin-like proteins. We may miss other proteins that are members of families with diverse roles. Finally, we will miss any proteins whose genes are evolving rapidly, which may include *BBS6* or cystin.

Taken together, our comparative genomics approach has produced a robust protein dataset that represents a significant portion of the *Chlamydomonas* proteome necessary for flagellar and basal body biogenesis and function. The core observation was the strong conservation of flagellar and basal body proteins. It is possible that a similar approach could be applied to other organelles and processes conserved in some phyla, but lost in others, to compartmentalize proteomes into functional subsets. The success of such approaches will depend on the extent of functional conservation of the process under study, the fidelity of the subtraction, particularly with respect to the choice of target genomes, and the ability to validate the subtracted protein fractions.

Notably, nearly half the proteins in our FABB set have no predicted function, and over a quarter of those lack recognizable motifs. This highlights the ability of proteome subtractions to assign initial function at a genome-wide level. Given the disparities between the rate of genomic data production and functional characterization of gene products, these types of approaches might provide the means to expedite annotation and accelerate the acquisition of biological tools. Finally, our FABB proteome will potentially serve as a useful template for the identification of genes involved in human ciliation disorders. The *BBS5* critical region spanned a 14 Mb interval containing in excess of 200 genes; the FABB proteome enabled us to clone *BBS5* by interrogating only two transcripts. Given that the involvement of cilia in a range of disorders is becoming more evident (Pazour and Rosenbaum, 2002), we predict that several genes involved in retinal degeneration, nephropathies, and pleiotropic syndromes with phenotypic overlap to BBS

will be included in the FABB proteome. Our collection will not only facilitate the genetic characterization of human disease phenotypes, but also provide initial information into the cellular basis of dysfunction.

Experimental Procedures

Searching the Proteomes

We used WU-BLASTP (Lopez et al., 2003) with $E \leq 10^{-10}$ to align all of the proteins predicted by GreenGenie (Li et al., 2003) in *Chlamydomonas* to all predicted proteins in several organisms with and without cilia.

Chlamydomonas Growth Conditions and Strains

Strain CC-124 was used in all transformation experiments. For RNA isolation and transformation, cells were grown in Sager and Granick II medium with stirring at 25°C to a density of 8×10^6 cells/ml. Transformation was performed as described previously (Perrone et al., 2003).

Quantitative Real-Time RT-PCR

We employed real-time RT-PCR to quantify the change in gene expression level after *Chlamydomonas* cells were deflagellated by pH shock (Piperno et al., 1977). Seventy-six genes had an EST to confirm the exon-intron junction; 27 genes had a well-conserved human homolog. Real-time RT-PCR was performed in 40 cycles of 95°C for 15 s and 60°C for 1 min using the ABI PRISM 7000 Sequence Detection System. The set of primers for each gene preferentially span an intron and sizes range from 50 to 150 bp. Oligonucleotides were obtained from Integrated DNA Technologies, Inc. (Coralville, Iowa). The SYBR Green PCR master mix (Applied Biosystems, Foster City, California) and uracil DNA glycosidase (Invitrogen, Carlsbad, California) were used according to the manufacturer's guidelines.

RNA Interference

Short hairpin constructs were constructed from a gene fragment containing exon and intron sequence and a fragment containing the same exonic sequence with added restriction enzyme sites. PCR fragments from genomic DNA were cloned into the TOPO TA-vector (Invitrogen) and confirmed by restriction enzyme digestion. The fragments were ligated together with T4 DNA ligase according to the manufacturer's guidelines (New England Biolabs, Boston, Massachusetts). The construct was cloned into the plasmid STL2, which is a derivative of pHisCrGFP (Fuhrmann et al., 1999). The resulting plasmid was cotransformed with pSI103, a plasmid carrying the aphVIII gene that confers resistance to paromomycin (Sizova et al., 2001). The presence of RNAi plasmid in transformants was verifed by colony PCR with primers to the sequence adjacent to the cloning site on pSTL2. The flagellar and cleavage furrow phenotypes of the transformants were examined by phase microscopy. Duplicate quantitative real-time RT-PCR was used to measure changes in the mRNA levels in transformants.

C. elegans Analysis

Growth and culture of all *C. elegans* strains was carried out using standard techniques (Brenner, 1974). RNA was isolated from mixed stage *daf-19(+)* and *daf-19(m86)* strains by standard protocols (Haycraft et al., 2001). Both *daf-19(+)* and *daf-19(m86)* strains also carry the *daf-12(sa204)* allele to suppress the Daf-c phenotype of *daf-19(m86)*. *daf-12(sa204); daf-19(m86)* adult hermaphrodites were transformed using standard techniques (Mello et al., 1991). Four independent transgenic lines were generated and analyzed. To obtain *daf-19(+)* worms for expression analysis, transgenic *daf-12(sa204); daf-19(m86)* males were mated to *daf-12(sa204)* hermaphrodites. Cross progeny were identified based on the presence of the extrachromosomal array and allowed to self fertilize. The subsequent generation was screened for rescue of the *daf-19(m86)* dye-filling defective phenotype by incubation in the hydrophobic dye DiO (Molecular Probes, Inc., Eugene, Oregon) as previously described (Fujiwara et al., 1999).

A promoter fusion was constructed with 500 bp upstream of the predicted translational start site of R01H10.6 through the second coding exon and was cloned into the GFP expression vector pPD95.75 (gift of A. Fire). The resulting construct was sequenced to ensure the correct translational reading frame and X box sequences. The entire *bbs-5* gene, including 496 bp of upstream promoter sequence, was fused in-frame with the coding region of *gfp* and the *unc-54* 3′ UTR. Transgenic lines expressing the *bbs-5::gfp* transgene were generated as described previously (Ansley et al., 2003).

Bardet-Biedl Syndrome Patients and Analysis

Individuals with Bardet-Biedl syndrome were ascertained by RAL, PBL, JSG, or PSP. Diagnosis was based on previously established criteria. For all kindreds, interviews were conducted, and genealogical records were constructed. With informed consent, all available medical records were obtained and reviewed. Appropriate individuals were genotyped for published markers from each *BBS* locus, and haplotypes constructed as described (Ansley et al., 2003). Primer sequences and amplification conditions are available upon request for the exons sequenced.

Generation of Antibodies to Mouse BBS5 Protein

Antisera against mammalian BBS5 were generated by injection of a peptide (DSIEDTKGNNGDRGC) corresponding to amino acids 35 through 48 of the mouse homolog into New Zealand White rabbits according to the standard protocols of Sigma-Genosys (Woodlands, Texas). Immunoblot analysis of mouse embryos (embryonic day 10.5) or cultured renal inner medullary collecting duct (IMCD; Stanton, 1997) cells were performed as described with a 1:500 dilution (Yoder et al., 2002). The horseradish peroxidase signal was detected using SuperSignal West Dura Extended Duration Substrate (Pierce, Rockford, Illinois) according to the manufacturer's instructions. For immunofluorescence, brain tissue was embedded in O.C.T. freezing compound (Sakura Finetek U.S.A., Inc., Torrence, California), cut into 4–6 μm sections, and fixed in 4% paraformaldehyde. Antibodies to acetylated α-tubulin (diluted 1:3000; Sigma, St. Louis, Missouri) or γ-tubulin (diluted 1:1000; Sigma) were used. Secondary antibodies were obtained from Jackson ImmunoResearch (West Grove, Pennsylvania) and nuclei were stained using Hoechst No. 33528 (Sigma). All images were captured under identical settings using a Nikon TE200 eclipse equipped with a Roper Scientific CoolSnap HQ CCD camera and MetaMorph imaging software (Universal Imaging, Downington, Pennsylvania).

Acknowledgments

Diego Martinez and Dan Rokhsar of DOE/JGI made much of this work possible and facilitated obtaining the *Chlamydomonas* genomic sequence. We thank the BBS families for their continued participation and the Research Department of the King Khaled Eye Specialist Hospital, Riyadh, KSA, and particularly Dr. John Cavender, former Director of Research. We thank N. Morrissette, J. Esparza, J. Keller, and A. Lippa for useful comments on the comparative genomics as well as S. Ansley, J. Badano, J. Kloss, C. Li, and A. Mah for technical assistance. J.B.L. thanks Dr. Robert Thach for his enthusiastic encouragement. S.K.D. is supported by NIH (GM-32843). J.B.L. was supported by the Monsanto Fellowship at Washington University. N.K. is supported by NIH (HD04260) and the March of Dimes. J.M.G. is supported by a fellowship from the German Academic Exchange Service. G.D.S. is supported by NIH (GM28755). B.K.Y. is supported by NIH (DK065655 and DK-62758). L.M.G.-W. is supported by NIH (DK055534). M.R.L. is supported by the HSFBC&Y and holds scholarship awards from the CIHR and MSFHR.

Received: March 26, 2004
Revised: April 22, 2004
Accepted: April 23, 2004
Published: May 13, 2004

References

Anholt, R.R., and Rivers, A.M. (1990). Olfactory transduction: crosstalk between second-messenger systems. Biochemistry *29*, 4049–4054.

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). Basal body dysfunction is a likely cause of pleiotropic Bardet-Biedl syndrome. Nature *425*, 628–633.

Badano, J.L., Ansley, S.J., Leitch, C.C., Lewis, R.A., Lupski, J.R., and Katsanis, N. (2003a). Identification of a novel Bardet-Biedl syndrome protein, BBS7, that shares structural features with BBS1 and BBS2. Am. J. Hum. Genet. *72*, 650–658.

Badano, J.L., Kim, J.C., Hoskins, B.E., Lewis, R.A., Ansley, S.J., Cutler, D.J., Castellan, C., Beales, P.L., Leroux, M.R., and Katsanis, N. (2003b). Heterozygous mutations in BBS1, BBS2 and BBS6 have a potential epistatic effect on Bardet-Biedl patients with two mutations at a second BBS locus. Hum. Mol. Genet. *12*, 1651–1659.

Bateman, A., Coin, L., Durbin, R., Finn, R.D., Hollich, V., Griffiths-Jones, S., Khanna, A., Marshall, M., Moxon, S., Sonnhammer, E.L., et al. (2004). The Pfam protein families database. Nucleic Acids Res. *32 Database issue*, D138–D141.

Beales, P.L., Katsanis, N., Lewis, R.A., Ansley, S.J., Elcioglu, N., Raza, J., Woods, M.O., Green, J.S., Parfrey, P.S., Davidson, W.S., and Lupski, J.R. (2001). Genetic and mutational analyses of a large multiethnic Bardet-Biedl cohort reveal a minor involvement of BBS6 and delineate the critical intervals of other loci. Am. J. Hum. Genet. *68*, 606–616.

Beales, P.L., Badano, J.L., Ross, A.J., Ansley, S.J., Hoskins, B.E., Kirsten, B., Mein, C.A., Froguel, P., Scambler, P.J., Lewis, R.A., et al. (2003). Genetic interaction of BBS1 mutations with alleles at other BBS loci can result in non-Mendelian Bardet-Biedl syndrome. Am. J. Hum. Genet. *72*, 1187–1199.

Bobinnec, Y., Marcaillou, C., and Debec, A. (1999). Microtubule polyglutamylation in Drosophila melanogaster brain and testis. Eur. J. Cell Biol. *78*, 671–674.

Brenner, S. (1974). The genetics of Caenorhabditis elegans. Genetics *77*, 71–94.

Cole, D.G. (2003). The intraflagellar transport machinery of Chlamydomonas reinhardtii. Traffic *4*, 435–442.

Cole, D.G., Diener, D.R., Himelblau, A.L., Beech, P.L., Fuster, J.C., and Rosenbaum, J.L. (1998). Chlamydomonas kinesin-II-dependent intraflagellar transport (IFT): IFT particles contain proteins required for ciliary assembly in Caenorhabditis elegans sensory neurons. J. Cell Biol. *141*, 993–1008.

Davy, B.E., and Robinson, M.L. (2003). Congenital hydrocephalus in hy3 mice is caused by a frameshift mutation in Hydin, a large novel gene. Hum. Mol. Genet. *12*, 1163–1170.

Dutcher, S.K. (1995a). Flagellar assembly in two hundred and fifty easy-to-follow steps. Trends Genet. *11*, 398–404.

Dutcher, S.K. (1995b). Purification of basal bodies and basal body complexes from Chlamydomonas reinhardtii. Methods Cell Biol. *47*, 323–334.

Dutcher, S.K. (2003a). Elucidation of basal body and centriole functions in Chlamydomonas reinhardtii. Traffic *4*, 443–451.

Dutcher, S.K. (2003b). Long-lost relatives reappear: identification of new members of the tubulin superfamily. Curr. Opin. Microbiol. *6*, 634–640.

Dutcher, S., ed. (2004). Dissection of Basal Body and Centriole Function in the Unicellular Green Alga Chlamydomonas Reinhardtii (Weinheim: Wile-VCH).

Ehler, L.L., Holmes, J.A., and Dutcher, S.K. (1995). Loss of spatial control of the mitotic spindle apparatus in a Chlamydomonas reinhardtii mutant strain lacking basal bodies. Genetics *141*, 945–960.

Essner, J.J., Vogan, K.J., Wagner, M.K., Tabin, C.J., Yost, H.J., and Brueckner, M. (2002). Conserved function for embryonic nodal cilia. Nature *418*, 37–38.

Fuhrmann, M., Oertel, W., and Hegemann, P. (1999). A synthetic gene coding for the green fluorescent protein (GFP) is a versatile reporter in Chlamydomonas reinhardtii. Plant J. *19*, 353–361.

Fujiwara, M., Ishihara, T., and Katsura, I. (1999). A novel WD40 protein, CHE-2, acts cell-autonomously in the formation of C. elegans sensory cilia. Development *126*, 4839–4848.

Handel, M., Schulz, S., Stanarius, A., Schreff, M., Erdtmann-Vourliotis, M., Schmidt, H., Wolf, G., and Hollt, V. (1999). Selective targeting of somatostatin receptor 3 to neuronal cilia. Neuroscience *89*, 909–926.

Haycraft, C.J., Swoboda, P., Taulman, P.D., Thomas, J.H., and Yoder, B.K. (2001). The C. elegans homolog of the murine cystic kidney disease gene Tg737 functions in a ciliogenic pathway and is disrupted in osm-5 mutant worms. Development *128*, 1493–1505.

Hertz, G.Z. and Stormo, G.D. (1999). Identifying DNA and protein patterns with statistically significant alignments of multiple sequences. Bioinformatics. *15*, 563–577.

Hinchcliffe, E.H., Miller, F.J., Cham, M., Khodjakov, A., and Sluder, G. (2001). Requirement of a centrosomal activity for cell cycle progression through G1 into S phase. Science *291*, 1547–1550.

Huang, K., and Beck, C.F. (2003). Phototropin is the blue-light receptor that controls multiple steps in the sexual life cycle of the green alga Chlamydomonas reinhardtii. Proc. Natl. Acad. Sci. USA *100*, 6269–6274.

Huangfu, D., Liu, A., Rakeman, A.S., Murcia, N.S., Niswander, L., and Anderson, K.V. (2003). Hedgehog signalling in the mouse requires intraflagellar transport proteins. Nature *426*, 83–87.

Katsanis, N., Beales, P.L., Woods, M.O., Lewis, R.A., Green, J.S., Parfrey, P.S., Ansley, S.J., Davidson, W.S., and Lupski, J.R. (2000). Mutations in MKKS cause obesity, retinal dystrophy and renal malformations associated with Bardet-Biedl syndrome. Nat. Genet. *26*, 67–70.

Katsanis, N., Ansley, S.J., Badano, J.L., Eichers, E.R., Lewis, R.A., Hoskins, B.E., Scambler, P.J., Davidson, W.S., Beales, P.L., and Lupski, J.R. (2001). Triallelic inheritance in Bardet-Biedl syndrome, a Mendelian recessive disorder. Science *293*, 2256–2259.

Katsanis, N., Eichers, E.R., Ansley, S.J., Lewis, R.A., Kayserili, H., Hoskins, B.E., Scambler, P.J., Beales, P.L., and Lupski, J.R. (2002). BBS4 is a minor contributor to Bardet-Biedl syndrome and may also participate in triallelic inheritance. Am. J. Hum. Genet. *71*, 22–29.

Khodjakov, A., and Rieder, C.L. (2001). Centrosomes enhance the fidelity of cytokinesis in vertebrates and are required for cell cycle progression. J. Cell Biol. *153*, 237–242.

Lefebvre, P.A., Silflow, C.D., Wieben, E.D., and Rosenbaum, J.L. (1980). Increased levels of mRNAs for tubulin and other flagellar proteins after amputation or shortening of Chlamydomonas flagella. Cell *20*, 469–477.

Li, Q.Z., Wang, C.Y., Shi, J.D., Ruan, Q.G., Eckenrode, S., Davoodi-Semiromi, A., Kukar, T., Gu, Y., Lian, W., Wu, D., and She, J.X. (2001). Molecular cloning and characterization of the mouse and human TUSP gene, a novel member of the tubby superfamily. Gene *273*, 275–284.

Li, J.B., Lin, S., Jia, H., Wu, H., Roe, B.A., Kulp, D., Stormo, G.D., and Dutcher, S.K. (2003). Analysis of Chlamydomonas reinhardtii genome structure using large-scale sequencing of regions on linkage groups I and III. J. Eukaryot. Microbiol. *50*, 145–155.

Lingle, W.L., and Salisbury, J.L. (1999). Altered centrosome structure is associated with abnormal mitoses in human breast tumors. Am. J. Pathol. *155*, 1941–1951.

Lopez, R., Silventoinen, V., Robinson, S., Kibria, A., and Gish, W. (2003). WU-Blast2 server at the European Bioinformatics Institute. Nucleic Acids Res. *31*, 3795–3798.

Marszalek, J.R., and Goldstein, L.S. (2000). Understanding the functions of kinesin-II. Biochim. Biophys. Acta *1496*, 142–150.

Mello, C.C., Kramer, J.M., Stinchcomb, D., and Ambros, V. (1991). Efficient gene transfer in C.elegans: extrachromosomal maintenance and integration of transforming sequences. EMBO J. *10*, 3959–3970.

Meraldi, P., and Nigg, E.A. (2001). Centrosome cohesion is regulated by a balance of kinase and phosphatase activities. J. Cell Sci. *114*, 3749–3757.

Mollet, G., Salomon, R., Gribouval, O., Silbermann, F., Bacq, D., Landthaler, G., Milford, D., Nayir, A., Rizzoni, G., Antignac, C., and Saunier, S. (2002). The gene mutated in juvenile nephronophthisis type 4 encodes a novel protein that interacts with nephrocystin. Nat. Genet. *32*, 300–305.

Mykytyn, K., Braun, T., Carmi, R., Haider, N.B., Searby, C.C., Shastri,

M., Beck, G., Wright, A.F., Iannaccone, A., Elbedour, K., et al. (2001). Identification of the gene that, when mutated, causes the human obesity syndrome BBS4. Nat. Genet. *28*, 188–191.

Mykytyn, K., Nishimura, D.Y., Searby, C.C., Shastri, M., Yen, H.J., Beck, J.S., Braun, T., Streb, L.M., Cornier, A.S., Cox, G.F., et al. (2002). Identification of the gene (BBS1) most commonly involved in Bardet-Biedl syndrome, a complex human obesity syndrome. Nat. Genet. *31*, 435–438.

Nishimura, D.Y., Searby, C.C., Carmi, R., Elbedour, K., Van Maldergem, L., Fulton, A.B., Lam, B.L., Powell, B.R., Swiderski, R.E., Bugge, K.E., et al. (2001). Positional cloning of a novel gene on chromosome 16q causing Bardet-Biedl syndrome (BBS2). Hum. Mol. Genet. *10*, 865–874.

Pasquale, S.M., and Goodenough, U.W. (1987). Cyclic AMP functions as a primary sexual signal in gametes of Chlamydomonas reinhardtii. J. Cell Biol. *105*, 2279–2292.

Pazour, G.J., and Rosenbaum, J.L. (2002). Intraflagellar transport and cilia-dependent diseases. Trends Cell Biol. *12*, 551–555.

Pazour, G.J., and Witman, G.B. (2003). The vertebrate primary cilium is a sensory organelle. Curr. Opin. Cell Biol. *15*, 105–110.

Pazour, G.J., Dickert, B.L., Vucica, Y., Seeley, E.S., Rosenbaum, J.L., Witman, G.B., and Cole, D.G. (2000). Chlamydomonas IFT88 and its mouse homologue, polycystic kidney disease gene tg737, are required for assembly of cilia and flagella. J. Cell Biol. *151*, 709–718.

Pazour, G.J., San Agustin, J.T., Follit, J.A., Rosenbaum, J.L., and Witman, G.B. (2002). Polycystin-2 localizes to kidney cilia and the ciliary level is elevated in orpk mice with polycystic kidney disease. Curr. Biol. *12*, R378–R380.

Perrone, C.A., Tritschler, D., Taulman, P., Bower, R., Yoder, B.K., and Porter, M.E. (2003). A novel Dynein light intermediate chain colocalizes with the retrograde motor for intraflagellar transport at sites of axoneme assembly in chlamydomonas and Mammalian cells. Mol. Biol. Cell *14*, 2041–2056.

Pfannenschmid, F., Wimmer, V.C., Rios, R.M., Geimer, S., Krockel, U., Leiherer, A., Haller, K., Nemcova, Y., and Mages, W. (2003). Chlamydomonas DIP13 and human NA14: a new class of proteins associated with microtubule structures is involved in cell division. J. Cell Sci. *116*, 1449–1462.

Pickett-Heaps, J.D. (1971). The autonomy of the centriole: fact or fallacy? Cytobios *3*, 205–214.

Piperno, G., and Mead, K. (1997). Transport of a novel complex in the cytoplasmic matrix of Chlamydomonas flagella. Proc. Natl. Acad. Sci. USA *94*, 4457–4462.

Piperno, G., Huang, B., and Luck, D.J. (1977). Two-dimensional analysis of flagellar proteins from wild-type and paralyzed mutants of Chlamydomonas reinhardtii. Proc. Natl. Acad. Sci. USA *74*, 1600–1604.

Schafer, J.C., Haycraft, C.J., Thomas, J.H., Yoder, B.K., and Swoboda, P. (2003). XBX-1 encodes a dynein light intermediate chain required for retrograde intraflagellar transport and cilia assembly in Caenorhabditis elegans. Mol. Biol. Cell *14*, 2057–2070.

Sineshchekov, O.A., Jung, K.H., and Spudich, J.L. (2002). Two rhodopsins mediate phototaxis to low- and high-intensity light in Chlamydomonas reinhardtii. Proc. Natl. Acad. Sci. USA *99*, 8689–8694.

Sizova, I., Fuhrmann, M., and Hegemann, P. (2001). A Streptomyces rimosus aphVIII gene coding for a new type phosphotransferase provides stable antibiotic resistance to Chlamydomonas reinhardtii. Gene *277*, 221–229.

Slavotinek, A.M., Stone, E.M., Mykytyn, K., Heckenlively, J.R., Green, J.S., Heon, E., Musarella, M.A., Parfrey, P.S., Sheffield, V.C., and Biesecker, L.G. (2000). Mutations in MKKS cause Bardet-Biedl syndrome. Nat. Genet. *26*, 15–16.

Stanton, B.A. (1997). Cystic fibrosis transmembrane conductance regulator (CFTR) and renal function. Wien. Klin. Wochenschr. *109*, 457–464.

Swoboda, P., Adler, H.T., and Thomas, J.H. (2000). The RFX-type transcription factor DAF-19 regulates sensory neuron cilium formation in C. elegans. Mol. Cell *5*, 411–421.

Tabin, C.J., and Vogan, K.J. (2003). A two-cilia model for vertebrate left-right axis specification. Genes Dev. *17*, 1–6.

Takeda, S., Yonekawa, Y., Tanaka, Y., Okada, Y., Nonaka, S., and Hirokawa, N. (1999). Left-right asymmetry and kinesin superfamily protein KIF3A: new insights in determination of laterality and mesoderm induction by kif3A−/− mice analysis. J. Cell Biol. *145*, 825–836.

Taulman, P.D., Haycraft, C.J., Balkovetz, D.F., and Yoder, B.K. (2001). Polaris, a protein involved in left-right axis patterning, localizes to basal bodies and cilia. Mol. Biol. Cell *12*, 589–599.

Wang, S., Luo, Y., Wilson, P.D., Witman, G.B., and Zhou, J. (2004). The autosomal recessive polycystic kidney disease protein is localized to primary cilia, with concentration in the basal body area. J. Am. Soc. Nephrol. *15*, 592–602.

Ward, S., Thomson, N., White, J.G., and Brenner, S. (1975). Electron microscopical reconstruction of the anterior sensory anatomy of the nematode Caenorhabditis elegans. J. Comp. Neurol. *160*, 71–110.

Ward, C.J., Hogan, M.C., Rossetti, S., Walker, D., Sneddon, T., Wang, X., Kubly, V., Cunningham, J.M., Bacallao, R., Ishibashi, M., et al. (2002). The gene mutated in autosomal recessive polycystic kidney disease encodes a large, receptor-like protein. Nat. Genet. *30*, 259–269.

White, J.G., Southgate, E, Thomson, J.N., and Brenner, S. (1976). The structure of the ventral nerve cord of Caenorhabditis elegans. Philos. Trans. R. Soc. Lond. B Biol. Sci. *275*, 327–348.

Wicks, S.R., de Vries, C.J., van Luenen, H.G., and Plasterk, R.H. (2000). CHE-3, a cytosolic dynein heavy chain, is required for sensory cilia structure and function in Caenorhabditis elegans. Dev. Biol. *221*, 295–307.

Witman, G.B., Carlson, K., Berliner, J., and Rosenbaum, J.L. (1972). Chlamydomonas flagella. I. Isolation and electrophoretic analysis of microtubules, matrix, membranes, and mastigonemes. J. Cell Biol. *54*, 507–539.

Wong, S.T., Trinh, K., Hacker, B., Chan, G.C., Lowe, G., Gaggar, A., Xia, Z., Gold, G.H., and Storm, D.R. (2000). Disruption of the type III adenylyl cyclase gene leads to peripheral and behavioral anosmia in transgenic mice. Neuron *27*, 487–497.

Wu-Scharf, D., Jeong, B., Zhang, C., and Cerutti, H. (2000). Transgene and transposon silencing in Chlamydomonas reinhardtii by a DEAH-box RNA helicase. Science *290*, 1159–1162.

Yoder, B.K., Hou, X., and Guay-Woodford, L.M. (2002). The polycystic kidney disease proteins, polycystin-1, polycystin-2, polaris, and cystin, are co-localized in renal cilia. J. Am. Soc. Nephrol. *13*, 2508–2516.

Young, T.L., Woods, M.O., Parfrey, P.S., Green, J.S., O'Leary, E., Hefferton, D., and Davidson, W.S. (1998). Canadian Bardet-Biedl syndrome family reduces the critical region of BBS3 (3p) and presents with a variable phenotype. Am. J. Med. Genet. *78*, 461–467.