REFERENCE ONLY

UNIVERSITY OF LONDON THESIS

Degree: PhD | Year: 2005 | Name of Author: HOSKINS, B. E.

COPYRIGHT
This is a thesis accepted for a Higher Degree of the University of London. It is an unpublished typescript and the copyright is held by the author. All persons consulting the thesis must read and abide by the Copyright Declaration below.

COPYRIGHT DECLARATION
I recognise that the copyright of the above-described thesis rests with the author and that no quotation from it or information derived from it may be published without the prior written consent of the author.

LOANS
Theses may not be lent to individuals, but the Senate House Library may lend a copy to approved libraries within the United Kingdom, for consultation solely on the premises of those libraries. Application should be made to: Inter-Library Loans, Senate House Library, Senate House, Malet Street, London WC1E 7HU.

REPRODUCTION
University of London theses may not be reproduced without explicit written permission from the Senate House Library. Enquiries should be addressed to the Theses Section of the Library. Regulations concerning reproduction vary according to the date of acceptance of the thesis and are listed below as guidelines.

A. Before 1962. Permission granted only upon the prior written consent of the author. (The Senate House Library will provide addresses where possible).

B. 1962 - 1974. In many cases the author has agreed to permit copying upon completion of a Copyright Declaration.

C. 1975 - 1988. Most theses may be copied upon completion of a Copyright Declaration.

D. 1989 onwards. Most theses may be copied.

This thesis comes within category D.

☑ This copy has been deposited in the Library of UCL

☐ This copy has been deposited in the Senate House Library, Senate House, Malet Street, London WC1E 7HU.

The molecular and functional genetics of Bardet-Biedl syndrome

Bethan Elinor Hoskins

Molecular Medicine Unit
Institute of Child Health, London

A thesis submitted for the degree of Doctor of Philosophy to
the University of London

September, 2004

UMI Number: U591727

All rights reserved

INFORMATION TO ALL USERS
The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript
and there are missing pages, these will be noted. Also, if material had to be removed,
a note will indicate the deletion.

UMI U591727
Published by ProQuest LLC 2013. Copyright in the Dissertation held by the Author.
Microform Edition © ProQuest LLC.
All rights reserved. This work is protected against
unauthorized copying under Title 17, United States Code.

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

Abstract

Bardet-Biedl Syndrome (BBS) is a rare genetic disease comprising obesity, retinal dystrophy, polydactyly and renal abnormalities. When this study began there were six known loci for BBS, only one of which, *BBS6*, had been identified. The initial aims of the project were to find *BBS* genes by screening of candidate genes within the known intervals and performing a genome wide screen using previously unlinked patients.

During the time of this project three of the known *BBS* loci (*BBS1*, *2* and *4*) have been identified and a further two novel genes (*BBS7* and *8*) have also been cloned. BBS was previously thought to be a recessive disease, but through mutation screening of all *BBS* genes in our cohort of 160 patients we have found evidence for complex inheritance involving the requirement for three mutations; a homozygous mutation in one *BBS* gene and a further heterozygous mutation at a second *BBS* locus, to manifest disease in some families.

As a quick and cheap alternative to sequencing for mutation detection, I developed the technique of multiplex capillary heteroduplex analysis (MCHA) which is now in routine use for screening new BBS cases. To determine the possible function of the BBS4 protein, a yeast-two-hybrid screen was undertaken to identify interactors of BBS4. Pericentriolar Material 1 (PCM1), one of the potential interactors, also co-localises with the protein at the centriolar satellites of centrosomes and basal bodies of primary cilia. The BBS8 protein, which shares homology with BBS4 also localises to the basal body and, from immunohistochemistry experiments using the BBS8 antibody, has been shown to be expressed in ciliated tissues such as kidney, retina, brain and spermatids, implicating basal body dysfunction in the aetiology of BBS.

During the course of my studies, we have progressed from a handful of genetic loci to the determination of the putative protein function which may pave the way to the design of future therapies.

Table of contents

List of figures

List of tables

List of appendices

Acknowledgements

Firstly I would like to thank everyone, in particular my parents and brother, who have helped me throughout the last four years and my friends Helen, Kate and Nahid for being so supportive in the last few months.

I would also like to thank all the past and present members of the MMU, especially all the members of 'Team BBS': Susan Darling, Jo Hill, Helen May-Simera, Amy Norman, Alison Ross and Sonia Sibold. I would also like to thank all the BBS patients and their families, particularly those that I have met through the LMBBS society, for their continued support and enthusiastic involvement.

As part of my PhD was spent at Johns Hopkins University, I would also like to thank UCL for awarding me a Bogue fellowship, and everyone in the Katsanis lab: Nico, Steve Ansley, Jose Badano, Jantje Gerdes, Carmen Leitch, Stephen Murhpy and Tanya Teslovich. Also from my time at Johns Hopkins, Lori Kotch, Andy McCallion, the Pandey lab and the Holt family.

Finally I would like to thank my supervisors Pete Scambler, Adrian Woolf and especially Phil Beales.

List of Abbreviations

3AT	3-aminotriazole
ABC	avidin-biotin complex
AD	activation domain
ADPKD	autosomal dominant polycystic kidney disease
APS	ammonium persulphate
ARF	ADP-ribosylation factor
ARL	ARF-like
ARPKD	autosomal recessive polycystic kidney disease
ASP	affected sib pair analysis
ATRA	all trans retinoic acid
BBS	Bardet-Biedl syndrome
BMI	body mass index
bp	base pair
BSA	bovine serum albumin
CC	connecting cilium
CCM	chemical cleavage of mismatches
cM	centiMorgan
CC-RCC	clear cell renal cell carcinoma
CRF	chronic renal failure
c.SNP	coding single nucleotide polymorphism
CVS	chorionic villus sampling
DAB	3,3-diaminobenzidine tetrahydrochloride
DBD	DNA-binding domain
dbEST	database of expressed sequence tags
DGGE	denaturing gradient gel electrophoresis
DHPLC	denaturing high-pressure liquid chromatography
DMSO	dimethyl sulfoxide

dsDNA	double-stranded DNA
E	embryonic stage (mouse)
EABA	endogenous avidin binding activity
ER	estrogen receptor
ERG	electroretinogram
ESRD	end-stage renal disease
EST	expressed sequence tag
FABB	flagellar apparatus-basal body
FFM	fat-free mass
FM	fat mass
GDB	genome database
HBD	homozygosity by descent
HDAC	histone deacetylase complex
HPSF	high purity salt free
HRP	horseradish peroxidase
IBD	identical by descent
IC	intermediate chain
IFT	intraflagellar transport
Ig	immunoglobulin
IS	inner segments (retina)
ISH	*in situ* hybridisation
LPA	linear polyacrylamide
5mC	methylcytosine
MCHA	multiplex capillary heteroduplex analysis
MCS	multiple cloning site
MKKS	McKusick-Kaufman syndrome
N-CoR	nuclear corepressor
NHR	nuclear hormone receptor
NIDDM	non-insulin dependent diabetes mellitus
NPHP	nephronophthisis
NRIP1	nuclear receptor interacting protein 1

OD	optical density
OE	olfactory epithelium
ONL	outer nuclear layer
OPL	outer plexiform layer
ORF	open reading frame
OS	outer segments (retina)
OSN	olfactory sensory neurons
PCM1	pericentriolar material 1
PCR	polymerase chain reaction
PEG	polyethylene glycol-400
PFA	paraformaldehyde
pI	isoelectric point
PTT	protein truncation test
PWS	Prader-Willi syndrome
RAR	retinoic acid receptor
RD	repression domain
RMR	resting metabolic rate
RNAi	RNA interference
RP	retinitis pigmentosa
RPE	retinal pigment epithelium
RRM	RNA recognition motif
RT-PCR	reverse transcriptase PCR
RXR	retinoid X receptor
SD	synthetic dropout
SHARP	SMRT/HDAC associated repressor protein
SID	SMRT interaction domain
SIFT	sorting intolerant from tolerant
SKIP	Ski-interacting protein
SLS	Senior-Löken syndrome
SMRT	silencing mediator for retinoid acid and thyroid hormone receptor
SNP	single nucleotide polymorphism

SRA	steroid receptor RNA activator
SSCP	single stranded conformational polymorphism
TD	touchdown PCR program
TE	telencephalon
TPR	tetratricopeptide repeat
UAS	upstream activating sequence
UPD	uni-parental disomy
UTR	untranslated region
VDR	vitamin D receptor
WHO	world health organisation
wt	wild-type

Chapter 1 Introduction

Bardet-Biedl syndrome (BBS) is a rare, multisystem genetic disorder displaying both a highly variable phenotype and considerable genetic heterogeneity. Despite first being reported almost a century ago, the underlying genetic or biochemical defect leading to the manifestation of the features of BBS was not known. It is only within the last five years, with the recent cloning of several *BBS* genes that it has been possible to propose a likely mechanism that, when defective, results in BBS.

1.1 History

The occurrence of retinitis pigmentosa (RP), hypogenitalism and mental retardation was first reported almost 150 years ago by Laurence and Moon (1866). Over half a century later, and quite independently, George Bardet (1920) reported the case of two siblings suffering from RP, polydactyly, hypogenitalism and obesity, followed in 1922 by a short case report from Arthur Biedl (1922) on siblings with RP, polydactyly, hypogenitalism, obesity and mental retardation. Initially the condition was known as Bardet-Biedl syndrome, however, in 1925, on the recommendation of Solis-Cohen and Weiss (1925), the name Laurence-Moon-Bardet-Biedl or Laurence-Moon-Biedl syndrome was commonly used. It was only following a review of the literature by Ammann (1970) that essential differences between the cases reported by Laurence and Moon, and those reported by Bardet and Biedl were highlighted, prompting a split in the nomenclature. The four siblings reported by Laurence and Moon developed a spastic paraparesis and there was no mention of obesity or polydactyly in these patients.

In 1969 Klein and Ammann (1969), following a study of 57 BBS cases in Switzerland, published the first clinical and genetic population-based survey of BBS patients. Twenty-six of the patients (45.6%) were classified as 'complete' cases manifesting the five cardinal features of the disease: RP, obesity, polydactyly, hypogenitalism and mental retardation. The

remaining patients were classified as 'incomplete' (19.3%; one or two cardinal features missing), 'abortive' (10.5%; only one or two cardinal features present), 'atypical' (8.8%; the ocular phenotype was something other than RP such as optic atrophy, myopia or anophthalmia) or undetermined type (15.8%).

Initial diagnostic criteria for BBS were set out in 1982 by Schachat and Maumenee (1982). They recommended that four of the five cardinal signs reported by Klein and Ammann (1969) be present for a diagnosis of BBS to be made. They did propose however, that retinal degeneration be an essential requirement for diagnosis and suggested that in its absence, a diagnosis of Prader-Willi syndrome (PWS) should be considered. As the renal component of BBS was rarely noted before 1980, it was not until several years after it was initially reported (Bauman and Hogan, 1973) that it was suggested as a cardinal feature of the disease in reports by Churchill *et al.* (1981) and Green *et al.* (1989). In the latter report, a study of 32 patients on the Island of Newfoundland, 21 patients were studied in detail to establish the frequency of both structural and functional renal abnormalities. All patients showed some form of abnormality, the most common structural features being abnormal calyces (in 95% of patients) and the persistence of fetal lobulation (95%), and a defect in the ability to concentrate urine being the most common functional anomaly (14/17 tested; 82%). Of the 28 patients that underwent an ophthalmic examination, 100% showed signs of retinal dystrophy and all but one of these patients were registered blind. Although this high incidence of retinal disease is likely a true reflection of the retinal involvement of the syndrome, particularly in this geographically isolated population, it may also be influenced by an ascertainment bias as the patients in this study were primarily recruited through an Ocular Genetics Clinic. Fifty-eight percent of patients (18/31) had polydactyly, but when other digital abnormalities such as syndactyly and brachydactyly were also included, 90% of patients were found to have dysmorphic extremities. Despite all patients showing signs of inappropriate behaviour and shallow affect, less than half of patients (41%) scored within the mentally retarded range on verbal and performance tests. Taking all of their results into account, Green *et al.* (1989) suggested that the cardinal manifestations of BBS should be considered to be retinal dystrophy, dysmorphic extremities, obesity, hypogenitalism in males and renal disease.

1.2 Diagnostic Criteria

The latest diagnostic criteria to be published are the first to incorporate the many secondary features that are often also seen in BBS cases. Beales *et al.* (1999) carried out a study of 109 patients, the largest survey of living BBS patients ever conducted, to establish the frequency of both the cardinal manifestations of the disease, and the less common secondary features. As observed in other studies, the majority of patients (102/109, 93%) had retinal dystrophy. The slightly lower frequency of retinal disease in this report in comparison to that of Green *et al.* (1989) is due to the fact that seven of the patients, all below the age of eight, did not have retinal dystrophy at the time of study. Since the publishing of this population survey, all patients have now developed RP resulting in a frequency of 100% amongst UK BBS patients (P. Beales, Personal communication). Obesity was common in patients; 72% of post-pubertal subjects were overweight according to the WHO body mass index classification ($BMI>25kg/m^2$, where BMI = weight in kg ÷ (height in m)^2) and 52% were obese ($BMI>30kg/m^2$). Renal defects were reported in only 26 patients but, as only half of the patients had had any kind of radiological investigation, it is possible that some patients had a non-symptomatic structural renal defect of which they would have been unaware. A variety of secondary features were noted amongst this population of BBS patients including developmental delay, speech disorders, dental anomalies, ataxia or poor coordination, diabetes mellitus and behavioural problems. Beales *et al.* (1999) therefore recommended that the diagnostic criteria be modified to incorporate these findings and suggested that for a diagnosis of BBS to be made there should be the presence of either four primary features or three primary features and two secondary features (for a list of primary and secondary features, see Table 1. 1). From this study the mean age at diagnosis was found to be nine years of age, despite some parents first noticing abnormalities in their children at the age of three. A period of six years between the initial development of clinical signs and confirmed diagnosis is relatively long but is probably due to the late development of some features of the disease, such as functional renal abnormalities and retinal dystrophy, and a possible reluctance of clinicians to make a diagnosis of BBS in the absence of a retinal phenotype.

Diagnostic criteria: Four primary features *or* three primary and two secondary features
Primary features
Rod-cone dystrophy (atypical retinitis pigmentosa)
Polydactyly
Obesity
Learning difficulties
Hypogonadism in males
Renal anomalies
Secondary features
Dental crowding/hypodontia/small roots/high arched palate
Brachydactyly/syndactyly
Speech disorders/delay
Strabismus/cataracts/astigmatism
Polyuria/polydipsia (nephrogenic diabetes insipidus)
Ataxia/poor coordination/imbalance
Left ventricular hypertrophy/congenital heart disease
Mild spasticity (especially lower limbs)
Diabetes mellitus
Developmental delay
Hepatic fibrosis

Table 1. 1: Current diagnostic criteria for BBS, published by Beales *et al.* (1999)

1.3 Demography and prevalence

BBS is a rare disease that is found throughout the world. The prevalence rate in the U.S.A and Northern Europe is between 1 in 100,000 and 1 in 160,000 (Beales et al., 1997; Croft et al., 1995; Klein and Ammann, 1969). However, high incidences of BBS are found in two specific populations: the Bedouin tribes of the Middle East and on the island of Newfoundland. The prevalence of BBS in the mixed Arab population is estimated to be 1 in 36,000, but a much higher incidence is seen in the Bedouin, particularly in Kuwait, where cases of BBS are as frequent as 1 in 13,500 live births (Farag and Teebi, 1989). The Bedouin are considered to be the most homogeneous Arab population due to the continued practice of consanguineous marriage, with first-cousin marriages particularly common; consanguinity rates as high as 54.3% have been reported in Kuwait where over half the population is Bedouin (Al-Awadi et al., 1985). Newfoundland, located just off the coast of Canada, is a large island with a population of only 550,000 and a prevalence of BBS of 1 in 17,500 (Parfrey et al., 2002). The island was settled in the late eighteenth and early nineteenth centuries by fishermen from the South-West coast of England and the South-East coast of Ireland. It has been estimated that 90% of the current population are descended from the original 20,000 to 30,000 settlers (Bear et al., 1988; Young et al., 1999a). In addition to the geographic isolation of the island, and the corresponding lack of immigration, a poor road network has resulted in the majority of families living in small communities of less than 2,500 with a high level of kinship and possible consanguinity inferred from the marriage of individuals with shared surnames

1.4 Primary Features

1.4.1 Retinal dystrophy

Rod-cone dystrophy (atypical retinitis pigmentosa) is a frequent finding in BBS patients (See Figure 1. 1a). The retinal degeneration commonly begins with night-blindness, first

noted by parents at a mean age of 8.5 years, and progresses more rapidly than typical RP resulting in the patient being registered legally blind at a mean age of 15.5 years in the UK BBS population (Beales et al., 1999). Results from the Swiss BBS study support this finding of a more rapid rate of retinal degeneration; Klein and Ammann (1969) reported that 86.4% of patients were registered blind by the age of 30 in contrast to blindness occurring between the ages of 40 to 50 as is the case in typical RP (predominantly autosomal dominant RP). As with most aspects of the BBS phenotype, there is a degree of both inter- and intrafamilial variation with respect to the age at which vision begins to deteriorate. In a study on the intrafamilial variation of the BBS phenotype, Riise *et al.* (1997) found that amongst the 11 families studied, first signs of night-blindness developed at a mean age of 4.6 years with a maximum intrafamilial variation of +/- 6 years. There was a much larger degree of variation seen in the age at which first signs of daytime visual impairment developed; the mean age was 6.6 years with a maximum variation of +/-14 years. There was also a noticeable difference in the pattern of visual loss; in one family with three affected sibs, one sib developed nightblindness followed by daytime visual impairment (as is the case in most BBS patients), a second developed the symptoms in the reverse order and the third developed night and daytime visual impairment simultaneously. The significance of these results however is hard to interpret as the sample size is very small and nightblindness is a largely subjective feature, which is difficult to measure.

In addition to retinal degeneration, other ocular abnormalities have been reported including strabismus, nystagmus, astigmatism, macular oedema and degeneration, cataracts, myopia and optic atrophy (Beales et al., 1999; Green et al., 1989; Klein and Ammann, 1969). Some of these features were observed at a significant frequency and are now considered to be secondary features of BBS (Beales et al., 1999) and may be present before the vision begins to deteriorate, allowing an earlier diagnosis (Musarella, 2001). An electroretinogram (ERG) is the investigation of choice to study the retina of patients. Although it can be hard to perform on infants, it may be altered as young as 14 months of age and can therefore be used, in conjunction with support and advice, as a means of preparing the child for an adult life without sight.

1.4.2 Obesity

Obesity is a common finding in BBS but, as with the other features of the disease, it is not universal. The extent of obesity varies between different studies due to different measurement criteria. Klein and Ammann (1969) reported obesity in 96% of their patients, classifying the obese as those that were above the 50^{th} percentile for weight when plotted on weight distribution charts. Green *et al.* (1989) also plotted patient weights on percentile charts, in this case height-adjusted, but described only patients above the 90^{th} percentile as overweight and therefore reported a slightly lower incidence of obesity in Newfoundland patients (88%). The WHO classification system for obesity is based on the calculation of the body mass index (BMI) that is an effective and non-invasive way of assessing obesity, as the height of the individual is incorporated into the calculation. Using this system, 72% of patients in the Beales *et al.* study were overweight (BMI>25kg/m^2) and 52% were obese (BMI>30kg/m^2). In this study it was also found that male patients were significantly shorter than the general population (mean height of male patients, 1.73m, compared to mean of general population, 1.76m), although female patients were also shorter than the general population, this difference was not significant (mean height of female patients, 1.62m, general population, 1.63m).

There are several genetic diseases including BBS, Prader-Willi (PWS) and Alström syndromes that involve the association of childhood obesity with hypogonadism, short stature and mental retardation. A child presenting with excess weight, normal intelligence and above average height for their age is more likely to be a case of simple isolated obesity, it is therefore essential to determine the child's height when considering the diagnosis of a genetic condition (Kopelman, 1994). The aetiology of the obesity seen in BBS is not known but it usually develops by the end of the first year and continues into adulthood where it is predominantly localised to the trunk and proximal limbs (Figure 1.1b, Beales et al., 2004). In contrast to PWS patients, who have a substantially lower resting metabolic rate (RMR) in comparison to controls, a study of the energy metabolism in BBS revealed that there were no significant differences in RMR, fat mass (FM) or fat-free mass (FFM) between a group of 20 BBS patients and 20 control subjects matched for age, gender and BMI (Grace et al., 2003). There was however, preliminary evidence in this study for a reduction in physical activity in

the BBS group relative to controls, which is unsurprising considering the physical handicaps associated with BBS. There was also evidence for an increased energy intake amongst BBS patients.

In addition to the renal component of the disease, obesity is one of the largest causes of morbidity and mortality in patients and is also associated with a number of co-morbidities. In adults a BMI>28kg/m^2 is linked to an increased risk of several conditions, including ischaemic heart disease, stroke, diabetes, cancer and renal failure. Childhood obesity can lead to a low self-esteem, bullying, early puberty and hypertension. In addition, an increased risk of death from cardiovascular disease has been observed in individuals whose BMI was above the 75th percentile for their age as an adolescent, whether or not the obesity persisted into adulthood (Kiess et al., 2001).

1.4.3 Renal abnormalities

The renal abnormalities in BBS can be structural; such as renal cysts, calyceal clubbing and fetal lobulation, functional; such as an inability to concentrate the urine, or both (Figure 1.1c, Beales et al., 1999; Harnett et al., 1988; Parfrey et al., 2002). Although not considered to be a cardinal manifestation of BBS until the 1980s (Churchill et al., 1981; Green et al., 1989), Klein and Ammann (1969) did note renal abnormalities in some patients, including cystic kidneys in three of the 57 patients in their study (5.26%). A study on the spectrum of renal disease amongst the BBS population on Newfoundland in 1988 (Harnett et al., 1988) found that although the majority of patients had a structural and/or functional abnormality, only 15% of patients suffered from a marked symptomatic renal impairment. Hypertension and frequent urinary tract infections are also common in patients and must be well managed in order to preserve renal function for as long as possible.

The kidneys of a BBS patient can show an altered appearance on ultrasound very early in life and, in some cases, prenatally (Cassart et al., 2004). Dippell and Varlam (1998) performed serial renal ultrasound examinations from birth on a group of seven patients from five

different families to study the changes in the renal appearance of patients in the first few months and years of life. At birth six patients had a renal volume in relation to body weight above the 97th percentile and all seven patients exhibited marked hyperechogenicity of the parenchyma. Over the initial 12 months, the kidneys lost their growth potential and by the end of the first year were below the 10th percentile for renal volume. In two patients this loss of renal mass continued until at least five years of age. The finding of large hyperechogenic kidneys, particularly in association with polydactyly, on a prenatal ultrasound scan can indicate an affected fetus and be used as a means of prenatal diagnosis in an at-risk pregnancy (Dar et al., 2001; Gershoni-Baruch et al., 1992); however, as neither renal abnormalities nor polydactyly are universal features of the disease, the absence of both features does not guarantee a healthy fetus.

Reports of the percentage of BBS patients to develop end-stage renal disease (ESRD) vary; Beales *et al.* (1999) reported ESRD in only 5% of patients (6/109, four of these six patients were children), Harnett *et al.* (1988) reported ESRD in 15% of patients (3/20) and O'Dea *et al.* (1996) estimate that 25% of patients will have developed ESRD by the age of 48 years. In patients with ESRD, hospital-based haemodialysis is the preferred treatment as the obesity, reduced manual dexterity and low IQ of BBS patients make home peritoneal dialysis a difficult option (Collins et al., 1994). There have been reports of successful renal transplant in several patients (Beales et al., 2004; Collins et al., 1994; Norden et al., 1991); however the renal transplant can lead to complications caused by other features of BBS. Both patients reported by Collins *et al.* (1994) had stable renal function post-transplant but were unable to control their weight due to hyperphagia, often seen in BBS patients, being exacerbated by the steroids used after transplantation. The transplant patient reported by Devarajan (1995) also suffered with morbid obesity post-transplant, despite not having suffered with it on initial presentation, and, having contracted three urinary tract infections in the six weeks after transplant, was found to be suffering from complex genital malformations that had not been detected previously. Considering these potential complications it is important to ensure that a thorough examination, including genitourinary anatomy, is performed prior to a transplant being given and that minimal steroids are used to maintain the transplanted organ. A new generation of anti-rejection drugs that are now available should help to elevate some of the problems experienced with the use of high levels of steroids.

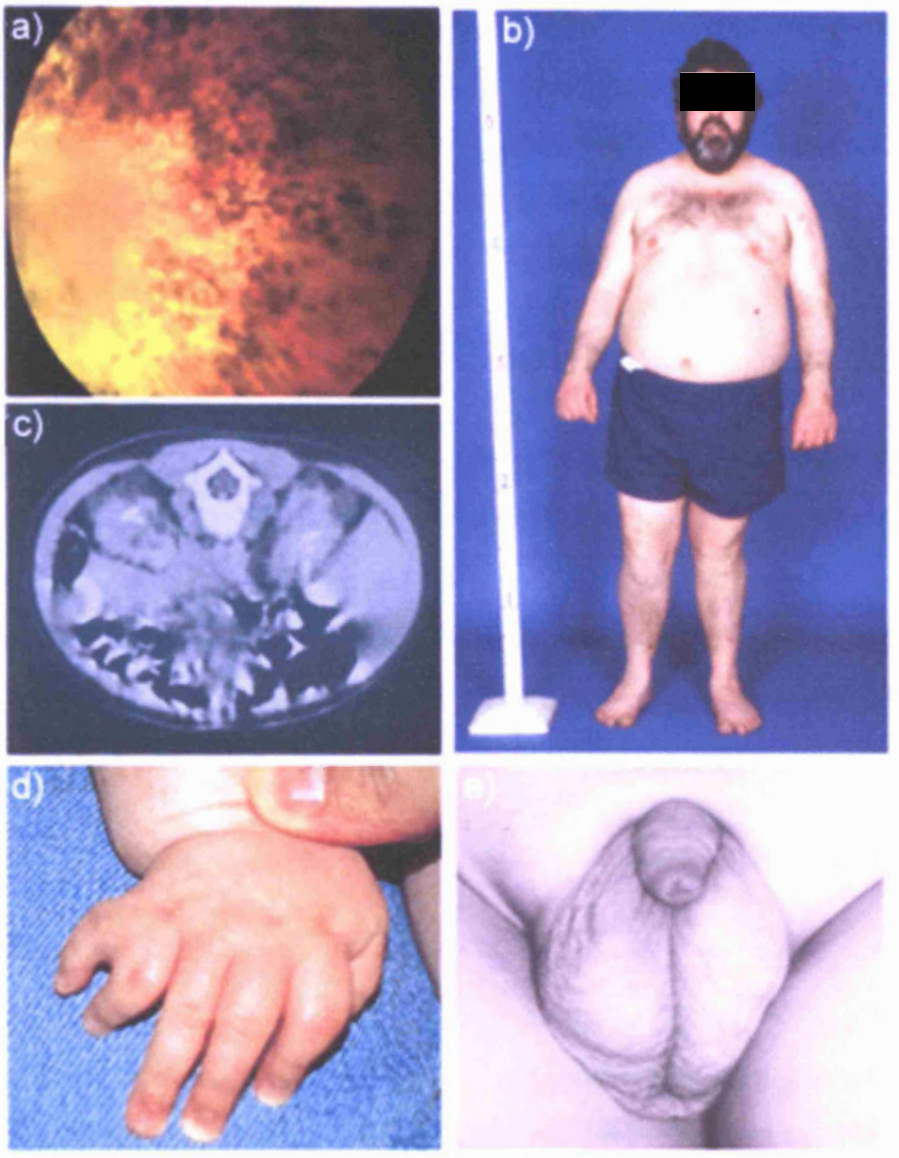


Figure 1. 1: Primary features of BBS. a) Fundoscopy from a 32 year old patient showing rod-cone dystrophy. b) Truncal obesity. c) Abdominal CT scan of an 11-year-old patient, showing irregular shaped kidneys with multiple cysts. d) Postaxial polydactyly of the hand. e) Hypogonadism.

1.4.4 Polydactyly

Polydactyly of one or more limbs was seen in 69% of patients in both the Swiss and UK studies (Beales et al., 1999; Klein and Ammann, 1969). The extra digit can vary from a simple skin tag to a fully formed digit (Figure 1.1d, Beales et al., 2004) and is almost always postaxial, although there have been occasional reports of preaxial polydactyly involving a bifid thumb (Klein and Ammann, 1969; Manouvrier-Hanu et al., 1999). In the UK study, 21% of patients exhibited polydactyly of all four limbs and polydactyly involving only the feet (21%) was almost three times more common than polydactyly involving only the hands (8%), findings similar to those of Green *et al* (1989) and Klein and Ammann (1969). There is extensive variation in the number and position of extra digits. Of the 34 patients in the UK study that did not have polydactyly, almost half of these (16/34) had affected sibs with polydactyly and in a case of monozygotic male twins, one twin had polydactyly of three limbs whereas the other had no polydactyly (Beales et al., 1999). Carmi *et al.* (1995a) have suggested that there is a genotype-phenotype relationship involving the extent of polydactyly; 11 of 12 *BBS3* linked patients from their study had an extra digit on all four limbs, while the polydactyly observed in patients linked to *BBS4* was predominantly restricted to the hands. These results however may not be universal; only a single family was studied for each genetic locus and later studies involving a larger number of families have not supported this observation (Riise et al., 1997).

Other limb abnormalities include brachydactyly of the hands and feet (Figure 1. 2a), partial syndactyly (particularly involving the 2nd and 3rd toes), fifth finger clinodactyly and a large 'sandal' gap between the 1st and 2nd toes (Beales et al., 1999; Green et al., 1989; Klein and Ammann, 1969).

1.4.5 Genital abnormalities

Hypogenitalism is more frequently reported in BBS males than females but despite the easier diagnosis in males, it is not thought that this accounts for all of the excess (Klein and

mentally retarded and reported observing 'mild feeble-mindedness' (mild mental retardation) in just over half of both male and female patients. Sixty-two percent of UK BBS patients were considered to be mentally retarded and half of these had received special education of some form (Beales et al., 1999). Despite a difference in the frequency of mental retardation in their studies, Beales *et al.* (1999) and Green *et al.* (1989) both reported altered behaviour in the majority of patients including volatile outbursts, emotional immaturity and shallow affect. A more detailed study of the behavioural phenotype and IQ of 21 BBS children (ages ranged from three to 18 years) conducted by Barnett *et al.* (2002) confirmed the presence of atypical behavioural traits in young patients and mild to moderate mental retardation in the majority of cases (11/21 (52.4%) of children had a Full Scale IQ in the mental retardation range). Cases of parents reporting internalising problems such as withdrawal, anxiety and depressed mood were common amongst this study group and the preference for routines in many patients was reported to have a negative impact on family life. Two of the 21 patients in this study were in the mild to moderate range and a further two were in the severe range for autism.

Figure 1. 2: Secondary characteristics of BBS. a) Brachydactyly of the hands (note also the scarring from the removal of extra digits). b) High-arched palate. c) Examples of the characteristic face associated with BBS.

1.5 Secondary features

In addition to a detailed study of the cardinal features of BBS, Beales *et al.* (1999) also described several secondary features of the disease, seen at varying frequencies within their patient cohort. These secondary features included:

Motor defects

Forty-three patients (39.4%) showed signs of ataxia and poor coordination. Young patients appear clumsy and adults tend to have poor balance.

Developmental delay

Half of patients were developmentally delayed and late in reaching milestones. Thirty-four male patients (31.2%) were also late in passing through puberty.

Dental Anomalies

Dental anomalies, including dental crowding, small teeth and enamel hypoplasia are common in BBS (Borgstrom et al., 1996; Kobrin et al., 1990). A high-arched palate is also present in the majority of patients (Figure 1.2b), Beales et al., 1999).

Speech deficits

Speech therapy was required in 54% of patients to correct a vocal or speech defect. The voice is high-pitched with a breathy quality with oral and palatal movements appearing uncoordinated.

Hearing loss

Twenty-one percent of patients suffered with hearing loss as a result of chronic glue ear, but in the majority of cases this had resolved by puberty. Three other patients (2.8%) had an unexplained sensorineural hearing loss and five patients (8.8%) with a progressive nerve deafness were also noted in the Swiss study (Klein and Ammann, 1969).

Diabetes mellitus

Six percent of UK patients had non-insulin dependent diabetes mellitus (NIDDM) (Beales et al., 1999) . Klein and Ammann (1969) and Green *et al.* (1989) also reported NIDDM in 8.8% and 45% of patients respectively. The much higher frequency within the Newfoundland cohort is likely to be the most accurate estimate of NIDDM in BBS as a fasting glucose tolerance test was carried out on all patients as part of the study; in both the Swiss and UK studies, as only a minority of patients had undergone such testing, it is possible that a number of patients had undetected glucose intolerance or NIDDM.

Hypertension

Hypertension has been noted in several studies, although at very different frequencies. Klein and Ammann (1969) and Beales *et al.* (1999) reported low frequencies of 1.8% and 8% respectively, whereas Green *et al.* (1989) report hypertension in 62% of patients. This discrepancy may be due to the increased incidence of renal disease within the Newfoundland cohort and may also be influenced by different measurement techniques.

Characteristic face

As is the case with some other genetic syndromes there may be a characteristic face associated with BBS as proposed by Beales *et al.* (1999) and supported by Lorda-Sanchez *et al.* (2001); many patients have deep set eyes, hypertelorism, a long philtrum, thin upper lip and a prominent forehead with early frontal balding in male patients (Figure 1. 2c).

1.6 Consanguinity rates

BBS is a rare autosomal recessive disease with considerably higher frequencies in the Bedouin of Kuwait and on the island of Newfoundland, where consanguinity rates are high. In her review of previously published cases of BBS, Bell (1958) reported a consanguinity rate of 39% amongst parents of patients, slightly higher than that reported on Newfoundland (Green et al., 1989). A surprisingly high rate of consanguinity was observed in the Swiss BBS study; consanguinity was reported in 20 of the 38 sibships (52.6%) described by Klein

and Ammann (1969). This high rate is explained by the distribution of many of their families in isolated areas where consanguinity is likely to be much higher than the more populated regions of Switzerland. As expected due to the outbred nature of the British population, a low rate of consanguinity (8.0%) was seen in the UK BBS cohort (Beales et al., 1999).

1.7 Heterozygous effects

Partial disease manifestations of BBS have been reported in some obligate heterozygous relatives of BBS patients. Through their study of 109 UK BBS patients, Beales *et al.* (2000) also observed an increased incidence of both clear cell renal cell carcinoma (CC-RCC) in parents, and renal malformations in unaffected sibs of BBS patients. In a study of 180 parents, three (1.67%) were found to have CC-RCC. None of the parents fell into a high-risk category for renal carcinoma such as the obese, smokers and those on chronic dialysis and all three parents were below 55 years of age at which point the risk of renal carcinoma rises sharply. The general population risk of developing CC-RCC below the age of 55 is 1 in 1041; therefore, the relative cumulative risk to parents of BBS patients is 17 times higher than that of the general population. In the same study, the incidence of unilateral renal agenesis (2/123) was 20 times higher in siblings of patients than in the general population. An increase in the incidence of obesity in BBS heterozygote parents has also been reported. A study by Croft and Swift (1990) on a single extended family showed an increased incidence of obesity, renal disease, hypertension and NIDDM in relatives of four BBS affected sibs. The authors suggested that this high frequency of features of BBS is a result of partial disease manifestations in the heterozygous state, but also acknowledge that the hypertension and diabetes in these individuals may be present merely as a consequence of their obesity. In a later study by the same group (Croft et al., 1995), the proportion of severely overweight fathers (26.7%) amongst the 34 parents studied was found to be significantly higher than in comparably aged American males (8.9%). However, this higher incidence of obesity in fathers was not observed in the considerably larger study of the UK BBS population by Beales *et al* (1999). An increased incidence of hypertension was also seen in the Croft *et al.* (1995) study but was not significant. As obesity was the only factor to

be more prevalent in both studies, Croft *et al.* (1995) concluded that the cases of hypertension, NIDDM and renal disease observed in the initial study must have been linked to something other than the BBS gene segregating in that family.

A high incidence of rare recessive diseases in isolated populations is often explained by a founder effect with genetic drift but in certain cases other factors, in particular selection, must also play a part in maintaining these diseases at high frequencies in isolated populations (Zlotogora, 1998). The high prevalence of BBS on the island of Newfoundland can not be the result of a single founder effect due to the occurrence of both genetic and allelic heterogeneity among BBS patients on the island (See *1.9.1 Mapping of BBS1-5*). It has therefore been proposed that one reason for the high frequency of BBS is past selection for heterozygous individuals based on their ability to store fat under the adverse conditions often seen in the harsh winters on Newfoundland. To determine if the *BBS1* gene (the high frequency of which is likely due to a founder effect) is linked to obesity on the island, 200 obese and 200 non-obese subjects were screened for the presence of heterozygous *BBS1* mutations (Fan et al., 2004b). Although three of the 200 obese subjects (1.5%) were found to be heterozygous for the M390R mutation (the most common mutation in *BBS1* that accounts for approximately 80% of cases of *BBS1*), the mutation was also seen in heterozygous form in three subjects from the non-obese cohort, indicating that this mutation is not linked to obesity in the general population of Newfoundland. In the same study the BMI of obligate heterozygous carriers of *BBS1* mutations were studied and compared to the BMI of non-carriers; no significant difference in BMI was found between carriers and non-carriers.

1.8 Related syndromes

The combination of the late onset of some of the features of BBS, such as renal disease and loss of vision, and the existence of other genetic syndromes with similar cardinal manifestations can lead to confusion amongst clinicians and the possibility of misdiagnosis.

Differential diagnoses include:

- Laurence-Moon syndrome – As described earlier (see *1.1 History*) there has been confusion for many years as to whether the conditions described by Laurence and Moon (1866), and those described by Bardet (1920) and Biedl (1922) are distinct entities or are in fact the same disease. At the current time the conditions are considered to be separate based on the presence of spastic paraparesis and the absence of obesity and polydactyly in Laurence-Moon patients.

- McKusick-Kaufman syndrome (MKKS) – MKKS occurs in both sexes but is more common in females. It is usually diagnosed in infancy and is characterised by the presence of postaxial polydactyly, hydrometrocolpos and, in 10-20% of patients, congenital heart defects (David et al., 1999). In affected males the features are limited to polydactyly and occasionally genital anomalies such as micropenis or hypospadias. Cases of BBS that present at birth or in infancy with polydactyly and hydrometrocolpos may be misdiagnosed as MKKS if no follow-up to assess for the later development of additional features of BBS is performed. To avoid misdiagnosis, Slavotinek and Biesecker (2000) recommended that a diagnosis of MKKS not be made in a sporadic female infant with hydrometrocolpos and polydactyly until five years of age, following monitoring of the infant for development of further features of BBS.

- Alström syndrome – Alström syndrome shares several features with BBS including truncal obesity, atypical retinitis pigmentosa and renal dysplasia but can be distinguished from BBS by the presence of a sensorineural hearing loss in 70% of patients and NIDDM in most patients (Beales et al., 2004). Other secondary features of Alström syndrome not reported in BBS include acanthosis nigricans, hypertriglyceridemia and skeletal abnormalities (Dyer et al., 1994).

- Biemond II syndrome – Another syndrome involving the features of obesity, polydactyly, hypogenitalism and mental retardation is the Biemond II syndrome. The substitution of iris coloboma for retinitis pigmentosa and the additional feature of

dwarfism in this condition should ensure that there is little confusion between the diagnosis of Biemond II syndrome and BBS (Schachat and Maumenee, 1982).

1.9 The Genetics of BBS

BBS is an autosomal recessive, single gene disorder. It was initially expected that mutations at a single locus would account for all cases of BBS but early linkage studies demonstrated the existence of genetic heterogeneity that has lead to the mapping, and subsequent cloning, of several *BBS* genes.

1.9.1 Mapping of *BBS1-5*

The first *BBS* gene was mapped in 1993 by Kwitek-Black *et al.* (1993). Two large inbred Bedouin kindreds from the Negev region in Israel were used in the linkage study; pedigree 1 contained nine affected members and pedigree 2, twelve. Both families were therefore large enough to be used independently for linkage analysis. The initial stage of the linkage study was to exclude loci that are known to be associated with heritable retinal disease as RP is one of the most common cardinal features of BBS (see *1.4.1 Retinal dystrophy*). Loci that were studied included RP loci on chromosomes 7 and 8, Usher's syndrome loci on chromosomes 1, 11 and 14, the rhodopsin locus on chromosome 3 and the Best's disease locus on chromosome 11. All loci were excluded in both BBS families and a genome wide screen was carried out in pedigree 1 with linkage found to chromosome 16q21. All affected individuals in the pedigree were homozygous for marker D16S408 and a lod score of 4.2 at a recombination fraction of 0 ($\theta = 0$) confirmed statistically significant linkage to this area. Individuals in pedigree 2 were then haplotyped to test for linkage to 16q21 in the second kindred. The locus was found to be excluded in this pedigree, giving evidence for locus heterogeneity in BBS. As it was registered second, this locus was designated *BBS2*.

In 1994 the second BBS locus, *BBS1*, was mapped by Leppert *et al.* (1994) using 31 outbred North American pedigrees. Twenty-nine of the pedigrees were of North European descent and three, Hispanic. A genome screen was carried out and a second *BBS* locus was mapped, with 17 of the 31 families showing linkage to 11q13 with a lod score of 4.31. The 14 families that were excluded from *BBS1* were also excluded from the previously mapped *BBS2* locus giving evidence for the existence of at least a third locus. No combination or severity of traits were identified that could distinguish the *BBS1* linked pedigrees from those that were unlinked, indicating that linkage to a particular locus could not be inferred by phenotype.

BBS3 was also mapped in 1994 to chromosome 3p12-13 by Sheffield *et al.* (1994) using conventional linkage analysis in a large inbred Bedouin pedigree. The family used in this linkage study was also used to test the pooled DNA approach to mapping recessive diseases in inbred populations. This approach to gene mapping works on the assumption that in isolated inbred populations most, or all, of the affected individuals will share a common chromosomal region that is associated with the disease due to identity by descent (IBD) from a common ancestor (Figure 1. 3). Equal amounts of DNA from all the affected members of the pedigree are pooled and used as a template for PCR using polymorphic markers. Two control pools, one containing unaffected siblings and the other containing the parents of the patients, are also generated and amplified with the markers. Markers that are not linked to the disease locus will have multiple alleles in the patient DNA pool, similar to the mixed alleles of the control pools. Markers that are linked to the disease locus however, will show a shift towards a single allele, that is likely to be associated with the disease, and will differ from the control pools. If a shift in allele frequencies between the patient and control pools is seen for markers, the individual members of the pedigree are genotyped for the markers, and those that flank them, and a lod score calculated. This DNA pooling strategy was also used to successfully map *BBS4* to chromosome 15q23 (Carmi et al., 1995b).

In 1999, Woods *et al.* (1999) carried out a survey of 17 Newfoundland families to investigate the high prevalence of BBS seen on the island and to establish whether, as anticipated, it is the result of a single gene founder effect. All families were genotyped using markers spanning the critical intervals of the four known loci and lod scores calculated. Unexpectedly, linkage to several loci was found indicating that the high frequency of BBS on

the island is not due to a founder effect. Three of the families were assigned to *BBS1*, one to *BBS2*, one to *BBS3*, six were excluded from all four loci and linkage results for the remaining six pedigrees were inconclusive and they could therefore not be assigned to any known locus. The family that was assigned to *BBS3* was the first family of North European descent to be linked to this locus, which had previously only been reported in a single consanguineous Bedouin family. One of the six families that were excluded from all loci was used to perform a genome-wide screen for the fifth *BBS* locus, the first locus to be mapped in a large inbred family of European ancestry. The DNA pooling approach described by Sheffield *et al.* (1994) was used, and linkage to chromosome 2q31 was found, with all affected members of the pedigree being homozygous by descent (HBD) for an ancestral haplotype spanning 13 cM (Young et al., 1999a).

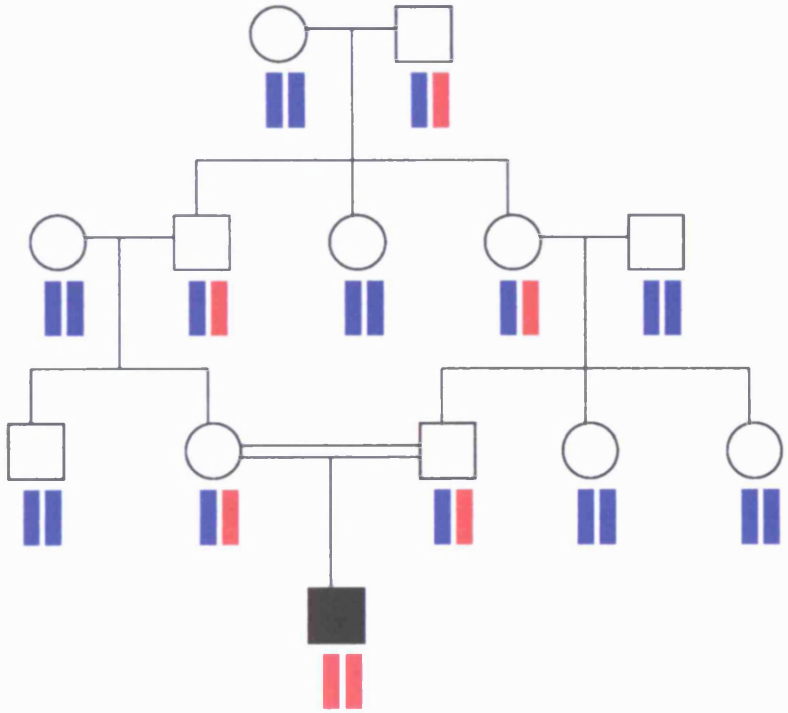


Figure 1. 3: Identity-by-descent (IBD). The affected individual is homozygous, or identical by descent, for the disease-associated haplotype (red), inherited from his great-grandfather.

1.9.2 Cloning of *BBS* genes

In the last four years considerable progress has been made in the cloning of the genes responsible for BBS. A candidate gene approach, for the identification of known loci, and an alternative homology-based approach, to identify novel genes through their homology to known *BBS* genes, have both been successful.

1.9.2.1 Positional Cloning of *BBS6*

The *BBS6* gene was the last *BBS* gene to be mapped to a chromosomal locus using conventional linkage analysis, and also the first *BBS* gene to be cloned. Following the identification of several Newfoundland families that were unlinked to *BBS1-5* (Woods et al., 1999), a genome screen to map a sixth *BBS* locus was carried out using the DNA pooling method in a single unlinked consanguineous pedigree with two affected individuals (Katsanis et al., 2000). A 3:1 reduction of alleles was seen from the control pool to the affected pool with marker D20S189 on chromosome 20p12. The remaining four pedigrees were also genotyped for D20S189 with a cumulative lod score of 3.92, indicating a new *BBS* locus on 20p12. To delineate the *BBS6* critical interval further, all Newfoundland pedigrees were genotyped for 20p markers; a 1.9cM critical interval was identified between markers D20S185 and D20S189. Evidence of a founder effect within *BBS6* families was found as the haplotype of one of the disease-carrying chromosomes in pedigrees NF-B1 and NF-B5 matched that of pedigrees NF-B3 and NF-B4.

Shortly before the mapping of *BBS6*, the gene that, when mutated, causes MKKS was cloned (Stone et al., 2000). Given the clinical overlap between MKKS and BBS (See *1.8 Related syndromes*), and the concordant mapping position, it was hypothesized that *MKKS* was a candidate gene for *BBS6* (Katsanis et al., 2000; Slavotinek et al., 2000). To test this hypothesis, mutation screening of all coding exons and splice junctions of the *MKKS* gene was carried out in the five Newfoundland families. Several coding region alterations were found. The affected individuals in pedigrees NF-B3 and NF-B4 were homozygous for a 1bp deletion (c.280delT) that results in a frameshift and premature termination of the protein at

codon 103 (F94fsX103). This deletion was also seen in heterozygous form in NF-B1 and NF-B5, supporting the haplotype inferred prediction of a common ancestral chromosome in these four pedigrees. The other disease-associated allele in NF-B1 was a complex frameshift mutation, c.429_430delCT + 433_435delAG (D143fsX157), and in NF-B5, was a c.830T>C transition (L277P) that was not seen in 234 control chromosomes. The identification of mutations in *MKKS* in five Newfoundland families prompted mutational analysis of this gene in 37 families of North American or European origin. Mutations were found in two of these pedigrees. In pedigree AR237 a homozygous c.110A>G (Y37C) mutation was found to segregate with the disease, this mutation had previously been reported in heterozygous form in an MKKS patient (Stone et al., 2000). Although *MKKS* mutations account for only 5-7% of BBS cases in North America and Europe, they account for 34% of BBS seen on Newfoundland, making *BBS6* the second most prevalent locus on the island after *BBS1*. MKKS is tenfold less prevalent than BBS and has not been reported in the Newfoundland population, suggesting that most mutations in *MKKS* result in BBS whereas milder, hypomorphic alleles result in MKKS. Slavotinek *et al.* (2000) also found *MKKS* mutations in four unrelated BBS probands providing an independent confirmation that *MKKS* was the first gene to be associated with BBS. The MKKS protein shows similarity to type II chaperonins that are responsible for folding a range of proteins. Alterations that change the shape of the molecule or truncations of the protein are likely to affect or abolish the ability of the MKKS protein to fold target peptides.

In order to delineate the critical intervals for the *BBS* loci, and determine the involvement of *BBS6*, Beales *et al.* (2001) conducted genetic and mutational analyses of a cohort of 163 pedigrees from diverse ethnic backgrounds. The critical interval of *BBS1* had previously been refined to a genetic distance of 2.9cM by linkage and haplotype analysis in 91 pedigrees (Katsanis et al., 1999), but intervals for some of the other loci remained large. The *BBS2* interval was substantially reduced from the18cM interval previously published by Kwitek-Black *et al.*(1993) to 2cM, by haplotype analysis of eight linked pedigrees. The *BBS3* interval was also reduced from 6cM to 2cM by analysis of two pedigrees that were consistent with linkage to 3p13. The inclusion of four Turkish pedigrees and one Pakistani pedigree in the haplotype analysis of *BBS4* meant that it was possible to refine the 2cM interval to a distance of 1.3cM. It was not possible to reduce the *BBS5* critical interval of 13cM, although

with 3 pedigrees showing linkage to 2q31 it provided independent confirmation of this locus that had previously only been described in a single Newfoundland pedigree. Strong evidence for a seventh *BBS* locus was provided by this study, as 14% of pedigrees were unlinked to any of the six known loci.

All 163 pedigrees were also screened for mutations in *MKKS*, regardless of any haplotype inferred chromosomal assignment. In eight pedigrees, missense alterations were found that segregated with the disease and were not present in a minimum of 188 control chromosomes. In seven of these pedigrees, despite good coverage of the open reading frame (ORF) of the gene, and splice junctions, only one mutation was identified. One of the heterozygous alterations, c.724 G>T (A242S) in pedigree B14, has also been reported in conjunction with an H84Y (c.250C>T) mutation in an MKKS patient (Stone et al., 2000), suggesting that it is a pathogenic mutation. Haplotype analysis of B14 showed that both the affected and unaffected sibs had identical 20p12 haplotypes, but that the affected sib was also homozygous across the *BBS2* interval (Figure 1. 4). The A242S (c.724G>T) allele was seen once in 142 Newfoundland controls, but not in 188 North American controls. These findings suggested that either the A242S is a very rare polymorphism or that BBS may not follow the classical model of recessive inheritance.

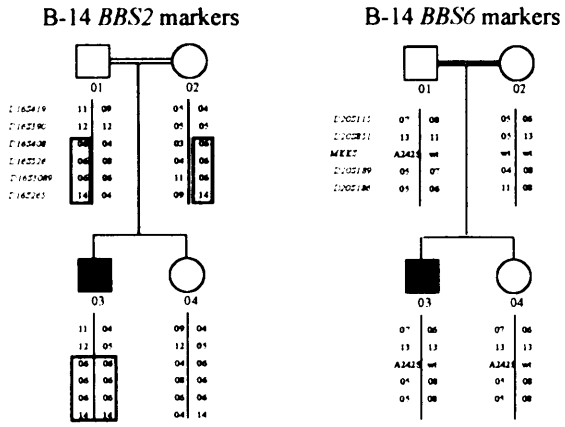


Figure 1. 4: Haplotype analysis of pedigree B14. Affected (03) and unaffected (04) offspring have identical haplotypes across the *BBS6* interval, but only 03 is IBD for the *BBS2* locus (Beales et al., 2001).

1.9.2.2 Identification of the second *BBS* gene; *BBS2*

The second *BBS* gene to be identified was, *BBS2* (Nishimura et al., 2001). *BBS2* is a novel gene with 17 coding exons and a predicted protein product of 721 amino acids. The gene is widely expressed, with strongest expression in the fetal and adult kidney, and adult heart, skeletal muscle, thyroid, spinal cord and adrenal gland. Homozygous mutations in *BBS2* were found in two large inbred Bedouin families, one of which was used to map the locus. The families had different haplotypes around the *BBS2* locus suggesting that they harboured different mutations. In one family a 1bp deletion in exon 8 (c.940delA) was found to segregate with the disease; this frameshift causes truncation of the protein at codon 324 (I314fsX324). The deletion was not seen in 96 controls. Two sequence variants, both homozygous and present *in cis*, were found to segregate with the disease in the second family. A c.224T>G (V75G) non-conservative substitution that was thought to be the disease-causing mutation in this family, and a c.367A>G (I123V) change that, based on its conservative nature, was assumed not to be pathogenic, although the possibility that both alterations are required for disease cannot be ruled out. The identification of mutations in *BBS2* in these two linked families prompted the screening of 18 unrelated probands of which three (17%) were found to harbour *BBS2* mutations including two nonsense mutations in exon 8, c.814C>T and c.823C>T (R272X and R275X), and a 1bp insertion in exon 10, c.1206insA (R403fsX408) (Nishimura et al., 2001).

1.9.2.3 Identification of *BBS4*

The *BBS4* gene, cloned by Mykytyn *et al* (2001), comprises 16 exons encoding a 519 amino acid protein. Like *BBS2*, *BBS4* is a novel gene that is ubiquitously expressed with the highest levels of expression in the kidney. Mutations in *BBS4* were found in five consanguineous pedigrees that showed evidence of linkage to chromosome 15q. The Bedouin kindred used to map the *BBS4* locus harboured a homozygous missense mutation, c.884G>C, in exon 12 (R295P), which was not present in 48 Bedouin controls. In an Italian pedigree, exons 3 and 4 failed to amplify by PCR in affected patients suggesting a partial gene deletion. Primers were designed to amplify across the deleted region and sequencing of the junction fragment

revealed a complete deletion of exons 3 and 4 (IVS2_IVS5), resulting in a loss of 48 codons. The same 6kb deletion was also found in an Israeli-Arab family. The two families did not share haplotypes suggesting that the mutation occurred independently in the two families, but the possibility of a distantly related ancestor cannot be ruled out. The deletion breakpoints are within *Alu* repeat elements in introns 2 and 4, it is therefore possible that the deletion is caused by unequal homologous recombination between the *Alu* repeats. In two other European families, mutations in the splice donor site of exon 4 (c.220+1G>C) and the splice acceptor site of exon 7 (c.406-2A>C) were found. In a single non-consanguineous BBS family that was screened for *BBS4* mutations, a heterozygous 2bp insertion, c.585-586insTG, was found in exon 8 (V195fsX209) but a second mutation in this family was not identified.

1.9.2.4 Cloning of the most common *BBS* gene; *BBS1*

The initial critical interval for the *BBS1* gene, mapped in 1994 (Leppert et al., 1994), was large at 26cM. Several studies attempted to narrow the interval and, through a combination of haplotype analysis in 91 pedigrees and linkage disequilibrium in the Newfoundland BBS population, a reduction of the interval to just 1cM was published (Katsanis et al., 1999; Young et al., 1999b). Haplotype analysis by Mykytyn *et al.* (2002) in several extended families generated an alternative critical interval distal to, and not overlapping with, either the Young *et al.* (1999b) interval or the slightly larger Katsanis *et al.* (1999) interval. Mutation screening of several candidate genes within this distal interval in a cohort of patients resulted in the identification of sequence alterations in a novel gene with 17 exons, now known as *BBS1*, which segregated with disease. Mutations identified in affected individuals included a homozygous c.1645G>T nonsense mutation in exon 16 (E549X) in a consanguineous Puerto Rican pedigree, a one base deletion in exon 10, c.851delA, resulting in a frameshift (Y284fsX288) and a common c.1169T>G missense mutation in exon 12 (M390R, accounts for ~80% of cases of *BBS1*) seen in homozygous form in 16 pedigrees and in heterozygous form in a further six pedigrees. Results of these mutation analyses confirmed that *BBS1* is the most common *BBS* gene, as predicted by linkage analysis. A study of the BBS1 protein sequence did not provide any clues as to the function of the protein but it does

show similarity to the BBS2 protein (BBS1 and BBS2 are 23% identical and 40% similar over an area of 192 amino acids).

1.9.2.5 Identification of a seventh *BBS* locus; *BBS7*

In contrast to the conventional approach to gene identification of linkage mapping and positional cloning by candidate gene screening, the fifth *BBS* gene to be cloned was identified through its similarity to the *BBS2* gene (Badano et al., 2003a). Fragments of the BBS2 peptide sequence were used to screen the conceptual translation of the human EST database (dbEST); two genes with modest similarity to *BBS2* were identified. One of the genes was independently found to be *BBS1* (Mykytyn et al., 2002) and in the other, mutations including a 4bp deletion, c.709-714delAAGA (K237fsX296), and two missense mutations, c.968A>G and c.662C>T (H323R and T221I), were found to segregate with disease, resulting in this novel gene being named *BBS7*, and designated as the seventh genetic locus to be associated with BBS (Badano et al., 2003a).

1.9.3 Further *BBS* loci

Despite the recent cloning of five *BBS* genes (*BBS1*, *2*, *4*, *6* and *7*), a large portion of patients, possibly as many as half (Katsanis et al., 2001b), remain unlinked to any of these known loci indicating the existence of additional *BBS* loci within the human genome. The extensive genetic heterogeneity seen in BBS and the relatively small contribution of some of the known genes to the syndrome may make identification of additional *BBS* loci by conventional methods difficult. A greater knowledge of the function of the known BBS proteins, the proteins with which they interact and how mutations in their respective genes can lead to an identical phenotype will likely assist the identification of new *BBS* genes by more functional approaches, as was successfully done in the case of *BBS7* (Badano et al., 2003a).

1.10 Aims of the project

Shortly before this project began, the first *BBS* gene, *BBS6*, was cloned. Initial aims of this project were:

- To identify the known loci by screening of candidate genes within the *BBS1-5* critical intervals.
- To map novel *BBS* loci through a genome-wide homozygosity screen in unlinked consanguineous pedigrees of Indian, Pakistani and Turkish origin.

Within the first year of the project the *BBS2* gene was cloned by another group, followed shortly afterwards by the cloning of *BBS4*, *1* and *7*. Cloning of these *BBS* genes resulted in the following additional aims of the project:

- Mutation analysis of all cloned *BBS* genes by direct sequencing in our patient cohort.
- Development of a quicker and more cost-effective mutation detection technique for screening of new BBS cases.
- Identification of interactors of the BBS4 protein by yeast-two-hybrid analysis.
- Cloning of new *BBS* genes using a novel homology based approach, followed by expression analysis of the gene products.

Chapter 2 Materials and Methods

2.1 Materials

2.1.1 Reagents

All reagents were of AnalaR grade and obtained from Sigma Aldrich or BDH unless otherwise stated. All solutions were made using Milli-Q purified water and autoclaved where appropriate.

2.1.1.1 Stock solutions

TBE	89mM Tris borate, 2.5mM EDTA (pH 8.3)
TE	10mM Tris-HCl, 1mM EDTA (pH 8.0)

2.1.1.2 Gel loading buffers

Sucrose loading dye	1.28M sucrose, bromophenol blue
1kb DNA ladder	0.1μg/μl 1kb ladder (Invitrogen) in TE
Genotyping loading buffer	three parts deionised formamide, one part GS-500 TAMRA size standard (Applied Biosystems), one part blue dextran (50mg/ml)/EDTA (50mM)

2.1.1.3 Yeast media

YPAD broth	2% bacto-peptone, 1% bacto-yeast extract, after autoclaving, 2% filter-sterilized glucose and 0.004% Adenine sulfate added

1xTE/LiAc	0.01M Tris, 0.1M lithium acetate
1xTE/LiAc/PEG	1xTE/LiAc in 50% polyethylene glycol-400 (PEG)
Synthetic dropout broth (SD)	0.67% bacto-yeast nitrogen base without amino acids, after autoclaving, dropout mix added
Dropout mix	20mg/l Adenine sulfate, 20mg/l Uracil, 20mg/l L-Tryptophan, 20mg/l L-Histidine, 20mg/l L-Methionine, 60mg/l L-Leucine and 30mg/l L-Lysine (30mg/l)
SD agar	SD broth with 2% bacto-agar added prior to autoclaving
Rescue solution	2% Triton X-100, 1% SDS, 0.1M sodium chloride, 0.01M Tris, 0.001M EDTA

2.1.1.4 Immunohistochemistry solutions

Carnoy's fluid	six parts absolute alcohol, three parts chloroform, one part acetic acid
Peroxidase block	3% hydrogen peroxide in 100% methanol
Block	10% goat serum, 1% BSA in PBS
PBT	0.1% Triton X-100 in PBS (Gibco-BRL Life Technologies)

2.1.2 DNA samples

A diagnosis of BBS in all patients was made by a clinician, based on published diagnostic criteria (Beales et al., 1999; Green et al., 1989). Where possible, samples from unaffected siblings and parents were also taken. For all samples, informed consent was obtained from either the individual, or in the case of a child, from a parent. DNA extraction from blood was either performed using the PUREGENE DNA Purification Kit (Gentra, see *2.2.1.1 DNA extraction*), or was carried out by the DNA Diagnostic Lab of Great Ormond Street Hospital. DNA samples from affected individuals were aliquoted at ~30ng/µl in a 96-well plate format for PCR amplification. DNA samples from NPHP patients were provided by Friedhelm Hildebrandt (University of Michigan), diluted to ~20ng/µl.

2.1.3 Primer design

Primers were designed to amplify the coding region of *BBS1*, *2*, *4*, *6*, *7* and *8* using the Primer v3 program (www.genome.wi.mit.edu/cgi-bin/primer/primer3.cgi). Amplicons were designed to include the exon and 50-100bp of flanking intronic sequence to ensure complete coverage of both splice junctions. Amplicon lengths ranged from 213-543bp. For *BBS2*, *4*, *7* and *8* each amplicon was designed to contain a single exon. In the case of *BBS1* certain exons and surrounding introns were sufficiently short to incorporate two exons within a single amplicon; exons 2 and 3 were contained within a single amplicon, as were exons 4 and 5. In contrast, exons 3 and 6 of *BBS6* were too large to be amplified in a single amplicon; four overlapping amplicons (3a, 3b, 3c and 3d) were designed to amplify exon 3 and two (6a and 6b) were designed for exon 6.

Primers were then tested for primer dimer formation or amplification of non-specific products using Amplify v.2.1 software (Bill Engels, University of Wisconsin, http://engels.genetics.wisc.edu/amplify/amplify). To increase the stability of some primers, M13 tags (Forward-TGTAAAACGACGGCCAGT, reverse-CAGGAAACAGCTATGACC) were added. Primers were manufactured and purified using a standard high purity salt free (HSPF) procedure (MWG Biotech, Sigma Genosys). To adapt primer pairs for MCHA, the forward primer of each primer pair was 5' end-labelled with either 6-FAM, HEX or TET (MWG Biotech, Sigma Genosys). For information on primer sequences, amplicon size, PCR conditions and colour of MCHA label, see Appendices 1-6.

2.1.4 Microsatellite markers

For the homozygosity screen, the ABI PRISM Linkage Mapping Set (ABI, Applied Biosystems) was primarily used. The set consisted of 358 dinucleotide repeat markers distributed throughout the genome with an average spacing of 10cM. The markers were organized into 28 panels based on expected allele size and colour of fluorescent label (FAM, HEX or TET). In order to achieve the desired density of markers, gaps caused by a failure of

markers from the ABI set to amplify were filled, where possible, using trinucleotide and tetranucleotide markers from the Single Chromosome Scan Human Screening Set (RG, Research Genetics, Inc.). If further markers were required to increase the density in an area of homozygosity, suitable markers in the area were identified using the Genome Database (GDB, www.gdb.org). The primers, one of which was fluorescently labelled, were manufactured according to standard methods (Sigma genosys, MWG biotech).

2.1.5 Yeast strain, bait and cDNA library

2.1.5.1 Yeast strain

The PJ69-4A strain of *Saccharomyces cerevisiae* (James et al., 1996) was used as the host strain for the two-hybrid screen. The strain consists of three reporter genes, each under the control of individual GAL4-inducible promoters (*GAL2-ADE2*, *GAL1-HIS3* and *GAL7-lacZ*). The genotype of the yeast is: MATa, *gal4*Δ, *gal80*Δ, *trp*1-901, *leu*2-3, *ura*3-52, *his*3-200, *GAL2-ADE2*, *LYS2::GAL1-HIS3*, *met2::GAL7-lacZ*.

2.1.5.2 pGBDU-BBS4 bait plasmid

Full-length BBS4 cDNA was cloned into the multiple cloning site (MCS) of pGBDU (James et al., 1996) using *Eco*RI and *Bam*HI sites, amplified and the plasmid extracted by Dr.Alison Ross.

2.1.5.3 cDNA library

The Human Kidney MATCHMAKER cDNA Library (Clontech) was used as the prey in the two-hybrid screen. The library consisted of ~3.5×10^6 independent clones, ranging in size

from 0.5-4.0kb. cDNA inserts were cloned into the *Eco*RI and *Xho*I sites of the MCS of the pACT2 vector. The vector contains a nutritional marker (*LEU2*) for selection in yeast.

2.1.6 BBS antibodies

BBS4 and BBS8 polyclonal antibodies were generated and purified by CovalAb. Computer prediction software was used to select likely immunogenic peptide sequences which were then synthesised and conjugated to a carrier protein such as keyhole limpet hemacyanin (KLH) or bovine serum alubmin (BSA) before being used to immunise a rabbit. Peptides sequences were - QFPVSTESQKPRQKK (BBS4) and GFLRPSTQSGRPGTME (BBS8). Diluted antisera for both antibodies were used in immunohistochemical staining.

2.2 Experimental procedures

2.2.1 General methods

2.2.1.1 DNA extraction from blood

Genomic DNA was extracted from 10ml of whole blood using the PUREGENE DNA Purification Kit (Gentra) according to manufacturers instructions. Following pelleting and lysis of white blood cells the sample was then treated with RNase A. The proteins were precipitated out and the supernatant mixed with 100% isopropanol to precipitate the DNA. Following washing of the DNA pellet with 70% ethanol, the DNA was resuspended in DNA hydration solution and stored at –20°C.

2.2.1.2 PCR reactions

Standard PCR reactions were carried out in a 25µl volume containing 1.5mM $MgCl_2$, 10mM Tris, 40mM NaCl, 0.25mM spermidine, 200µM of each dNTP (Amersham Biosciences), 5pmol of each primer, 1U *Taq* polymerase (Bioline) and ~30ng of genomic DNA. PCR reactions were carried out in a 96-well plate on a standard thermocycler (Eppendorf, Perkin Elmer, GRI). All primer pairs were initially tested on control genomic DNA using the Touchdown PCR program and optimised further (annealing temperature or magnesium concentration) where necessary. For difficult PCR reactions, the FAILSAFE PCR PreMix Selection Kit (Epicentre) was used. Primers were tested in a PCR reaction using the FAILSAFE Enzyme Mix with one of 12 different buffers (2x FAILSAFE PCR PreMix A-L). The reaction mix was comprised of two parts; a 12.5µl volume containing 5pmol of each primer, 1.25U FAILSAFE Enzyme Mix and ~30ng of genomic DNA, with 12.5µl of PreMix A-L. Following testing of a primer pair with each of the PreMixes, the PCR was repeated using a PreMix that gave a single strong PCR product on agarose gel electrophoresis.

PCR programs used were either:

1. A two-step touchdown program (TD) of 95°C for 5 minutes, followed by step 1: 10 cycles comprising 94°C for 30 seconds, an initial annealing temperature of 65°C for 30 seconds set to decrease by 1°C per cycle, an extension of 72°C for 30 seconds, followed by step 2: 25 cycles of 94°C for 30 seconds, 55°C for 30 seconds and 72°C for 30 seconds, followed by a final extension at 72°C for 10 minutes. Or,

2. A standard PCR program of 95°C for 2 minutes, followed by 30 cycles of 95°C for 30 seconds, an annealing temperature of either 55°C, 58°C, 61°C, 63°C or 71°C dependent on the fragment for 30 seconds, 72°C for 45 seconds, followed by a final extension step of 72°C for 10 minutes. Or,

3. The FAILSAFE PCR program: 94°C for 1 minute followed by, 30 cycles of 94°C for 30 seconds and 72°C for 1 minute.

2.2.1.3 Agarose gel electrophoresis

To estimate the yield of PCR reactions, a 5µl aliquot of each reaction was added to 2µl sucrose loading dye and run on a 1% agarose gel (1g electrophoresis grade agarose (Gibco-BRL Life Technologies) in 100ml of 1xTE) with ethidium bromide at a final concentration of 0.5µg/ml. To check that products were of the correct size, a 5µl aliquot of 1kb ladder was also run alongside the PCR products in a single well of the gel.

2.2.1.4 Restriction digests

Following detection of a sequence alteration in a patient sample, the 'Cut Map' function in the Sequencher v4.1 software (Genecodes) was used to determine if the change affected any restriction sites around the alteration. If the change either created or abolished a restriction site, a restriction digest using the corresponding enzyme was carried out on the patient

sample, a control DNA sample and also any available relatives of the patient, to confirm the presence of the change in the patient and correct segregation within the family. Digests were carried out in a 20µl volume containing 0.2µl bovine serum albumin (BSA), 20U enzyme (10u/µl, New England Biolabs), 2µl enzyme reaction buffer (New England Biolabs), 5.8µl water and 10µl PCR product. After gentle mixing, the reactions were then incubated at the optimal temperature for the enzyme (either 37°C or 55°C) overnight. 5µl of sucrose loading dye was added to the reaction and the whole volume run on a 2% low melting point agarose gel (2g NuSieve GTG low melting point agarose (FMC Bioproducts) in 100ml 1xTE with 0.5µg/ml ethidium bromide). Samples from the patient and any relatives were compared to the pattern of bands seen in the control sample to confirm presence of the sequence change.

2.2.2 Homozygosity mapping

2.2.2.1 Amplification of markers

To amplify markers, the True Allele PCR mix (Applied Biosystems) was used in a 7.5µl reaction volume consisting of 4.5µl True Allele PCR Mix, 2.5pmol of each primer, 1.5µl water and ~30ng genomic DNA. The True Allele PCR program was used; an initial denaturation at 95°C for 12 minutes followed by a two step cycling program, step 1: 10 cycles of 94°C for 15 seconds, 55°C for 15 seconds and 72°C for 30 seconds, then step 2: 20 cycles of 89°C for 15 seconds, 55°C for 15 seconds and 72°C for 30 seconds, followed by a final extension of 72°C for 10 minutes.

2.2.2.2 Pooling of markers

Markers were pooled according to the Panel Guide for the ABI PRISM marker set or combined such that there was a minimum difference of 20bp between the largest expected allele size for one marker and the smallest expected allele size for the following marker. For pooling, 2.5µl of HEX and TET labelled markers and 1.5µl of FAM labelled markers were

combined in a clean sample plate. Following pooling, a 1.5µl aliquot of the mix was then added to 3.5µl genotyping loading buffer.

2.2.2.3 Gel preparation for electrophoresis of samples

Markers were run on a 96-well 6% polyacrylamide gel on an ABI 377 sequencer (Applied Biosystems). Prior to pouring the gel, all apparatus including glass plates (36cm well-to-read distance), spacers and combs were thoroughly cleaned using warm water and a solution of Alconox detergent to remove all traces of dirt and residual fluorescence. The plates were then rinsed in distilled water and allowed to air dry before the apparatus was assembled. The glass plates, separated by 0.4mm thick spacers, were placed in the frame and locked into position. To make the gel, 10.8g urea and 3ml 6% Long Ranger Gel Solution (BMA) were combined in a 30ml solution of 1x TBE. The mix was then filter sterilized through a 150ml filter (Nalgene) using a vacuum pump. Just prior to pouring of the gel, 150µl freshly prepared 10% ammonium persulphate (APS) and 21µl TEMED (N,N,N',N'- tetramethylethylenediamine, Gibco-BRL life technologies) were added to the gel mix to accelerate polymerisation of the acrylamide. The mix was then injected between the glass plates using a 20ml syringe. Whilst the gel was injected, the glass plates were tapped to prevent air bubbles from forming in the gel. The straight edge of a plastic casting comb was inserted into the top of the gel and clamped in place during the polymerisation process (~45 minutes). When the gel was set, the comb was removed, the top of the gel rinsed out using distilled water and a paper 96-well comb (PE Biosystems) inserted.

2.2.2.4 Gel electrophoresis of markers

The gel apparatus was fitted inside the sequencer and both the upper and lower buffer chambers filled with 1xTBE. The 'Prerun' protocol was then carried out to warm the gel to the run temperature of 51°C. During the Prerun step, the samples were denatured at 95°C for three minutes and placed on ice. Following rinsing of the wells with buffer, the samples were loaded using a Kloehn multichannel loader (Anachem). The first 48 samples were loaded in

alternate lanes, the gel run for a further two minutes and the final 48 samples loaded. The gel was then run at 3kV for 3 hours.

2.2.2.5 Genotyping of markers

To analyse the results, the Gene Scan v.3.1.2 analysis software (Applied Biosystems) was first used to track the lanes of the gel and to size the alleles using the internal TAMRA size standard present in each lane. The Genotyper v.2.5 software (Applied Biosystems) was then used to view the sized fragments that were designated appropriate allele sizes for each marker based on the expected allele size range from the genome database (GDB, www.gdb.org).

2.2.3 Direct sequencing

2.2.3.1 PCR reaction clean-up

Following the PCR reaction, excess primers and dNTPs were removed using the exonuclease I (*Exo* I) and shrimp alkaline phosphatase (SAP) enzymes (New England biolabs). 2U SAP (1U/µl) and 10U *Exo* I (10U/µl) were added to the PCR product, held at 37°C for 15 minutes and then heated to 80°C for 10 minutes to denature the enzymes.

2.2.3.2 Cycle sequencing

Cycle sequencing was carried out using Big Dye Terminator v3.1 (Applied Biosystems) in a 15µl reaction volume; 5µl PCR product, 3µl 5X Big Dye sequencing buffer, 2µl BigDye cycle sequencing reaction mix, 5pmol primer and 4µl water. The cycle sequencing program used was: 95°C for 2 minutes, followed by 35 cycles of 95°C for 20 seconds, 50°C for 10 seconds and 60°C for 3 minutes.

2.2.3.3 Sequencing reaction clean up

To remove excess salt, sequence samples were passed through a sephadex (G-50) column. To make the column: sephadex was poured onto the loader and compacted using the scraper. A clean 96-well MultiScreen plate (Millipore) was then placed on top of the loader, the two parts (loader and plate) were turned over and the loader tapped to transfer the sephadex to the MultiScreen plate. Before use, the sephadex was rehydrated; 350µl of water was added to the sephadex and the plate stored at 4°C for three hours. Following rehydration of the sephadex, the plate was spun down at 910g for five minutes, the flow-through discarded, 150µl water added and spun for a further five minutes to rinse the column. Prior to passing the sequence samples through the sephadex column, 5µl of distilled water was added to each sample to bring the volume to 20µl. The sephadex plate was then transferred to a clean, skirted 96-well plate, the sample added and the plate spun for a final five minutes.

2.2.3.4 Capillary electrophoresis

After removal of salt from the sample by sephadex, the samples were run on the MegaBACE 1000 DNA Analyser (Amersham Biosciences), a capillary-based DNA fragment analyser composed of 96 40cm long acrylamide-coated capillaries with a 50µm internal diameter. Standard sequencing materials and conditions were used according to manufacturer's instructions. Approximately one hour before the run, the 3% Long Range linear polyacrylamide (LPA) matrix (Amersham Biosciences) was removed from the refrigerator to allow the matrix to reach room temperature and a buffer plate containing 150µl 1xLPA buffer (Amersham Biosciences) per well prepared. The 'Matrix Fill and Prerun' followed by the 'Inject Samples and Run' protocols were performed. The samples were injected at a voltage of 3kV for 40 seconds and run conditions of 9kV for 60-100minutes (dependent on fragment length) were used.

2.2.3.5 Sequence analysis

Sequence chromatograms were visualised using MegaBACE Sequence Analyzer v.3.0 software (Amersham Biosciences) and exported to ABD format. Chromatograms were then analysed using the Sequencer v.4.1 program (Genecodes). For each gene, a contig containing the published genomic sequence was created. All patient chromatograms were then aligned with the published sequence for analysis.

2.2.4 MCHA

2.2.4.1 Heteroduplex formation

To form heteroduplexes, 5µl (~300ng) of the patient PCR product was mixed with an equal quantity of the control DNA product and denatured at 95°C for 5 minutes with 2°C decrements every 2 minutes until the base holding temperature (25°C) was reached.

2.2.4.2 Sample plate preparation

2µl of the heteroduplexes were mixed with 0.5µl ET Rox size standard (Amersham Biosciences) and made up to 10µl with distilled water. When pooling of samples was carried out, 1.5µl aliquots of each of the heteroduplexes were pooled, size standard added and made up to 10µl with distilled water. In each run, a known wt homoduplex amplified from control DNA was included in a single well of the sample plate as a reference peak when analysing the patient samples.

2.2.4.3 Capillary electrophoresis

MCHA was carried out on the MegaBACE 1000 DNA Analyser (Amersham Biosciences). To remove excess salt from the samples the 'Preinject Samples' protocol was performed prior to injection of matrix into the capillaries. Following preinjection, 3% non-denaturing LPA matrix (Amersham Biosciences) was injected under high pressure for 200 seconds into the capillaries. A standard sample injection voltage of 3kV over 45 seconds was used. Each sample was run at 25°C at a voltage of 10kV for between 40 and 60 minutes dependent on fragment length.

2.2.4.4 MCHA Analysis

The results were analysed using Genetic Profiler v.2.0 software (Amersham Biosciences). Peaks were compared with the reference peak and any with an abnormal morphology (e.g. split or extra peaks) were reamplified using unlabelled primers, sequenced and analysed according to the direct sequencing protocol (see *2.2.3 Direct sequencing*).

2.2.5 Yeast two-hybrid

2.2.5.1 Autoactivation test

To test the ability of the DBD-BBS4 hybrid alone to activate the system, the pGBDU-*BBS4* plasmid construct was transformed into PJ69-4A using a small-scale transformation protocol. A single colony of PJ69-4A was used to inoculate a 5ml volume of YPAD broth. The culture was incubated, with gentle shaking, overnight at 30°C. The following day, the optical density (OD_{600}) of the culture was taken and a sufficient volume was used to inoculate a 60ml volume of YPAD broth to an OD_{600} of 0.1. The culture was again incubated at 30°C with gently shaking until the OD_{600} had reached between 0.5 and 0.7 (four to six hours of growth).

Following incubation, cells were pelleted at 1,500g for five minutes. The supernatant was then discarded, the cells resuspended in 20ml of distilled water and pelleted a second time. The cell pellet was finally resuspended in 300µl of 1xTE/LiAc and aliquoted into three sterile 1.5ml eppendorf tubes. 100ng of pGBDU-*BBS4* plasmid DNA and 50ng of denatured carrier DNA (Herring testes DNA, Clonetech) was added to each eppendorf, followed by 300µl 1xTE/LiAc/PEG. The contents of the tube was mixed gently by inversion and incubated at 30°C, with shaking, for 30 minutes. 70µl dimethyl sulfoxide (DMSO) was then added to each tube, mixed by inversion and incubated at 42°C for a further 15 minutes. Following a brief centrifugation (10 seconds at 15,800g), the supernatant was discarded and the pellet resuspended in 500µl of distilled water. The cell suspension from each tube was plated out onto a fresh SD-Ura agar plate and incubated at 30°C for two days.

Following two days growth, two colonies were picked from each of the three SD-Ura plates and streaked out in duplicate onto SD-Ura-Ade and SD-Ura-His+3AT plates and incubated for a further four to five days. The absence of growth on either the SD-Ura-Ade or the SD-Ura-His+3AT, indicated that there was no autoactivation of the system by the DBD-BBS4 hybrid protein. 3AT (3-aminotriazole) is an inhibitor of the histidine pathway and is added to the media to suppress leaky expression of *HIS3*.

2.2.5.2 Large scale transformation of library plasmids

Following the small-scale transformation of PJ69-4A with the pGBDU-*BBS4* bait plasmid to test for autoactivation, a library screen was carried out using a large-scale transformation protocol to introduce the library plasmids into the yeast. A 200ml volume of SD-Ura media was inoculated with a whole scraping of a streak from the SD-Ura plates used in the autoactivation test and incubated at 30°C with shaking for 24 hours. After the incubation, the OD_{600} of the culture was taken and an appropriate volume of the starter culture added to 1 litre of prewarmed (30°C) YPAD broth to give an OD_{600} of 0.1-0.2. The culture was again incubated at 30°C with gently shaking until the OD_{600} had reached 0.5-0.7. The culture was pelleted in 250ml bottles at 960g for five minutes. The cell pellet from one of the bottles was resuspended in 20ml of 1xTE/LiAc which was then used to resuspend the cell pellet from the

second bottle. The 20ml solution was pipetted into a sterile 1 litre conical flask and 10mg of denatured carrier DNA (Herring testes, Clontech) and 250µg of library plasmid DNA added. The contents of the conical flask was then mixed by swirling, 140ml of 1xTE/LiAc/PEG added and incubated at 30°C, with gentle shaking, for 30 minutes.

Following the incubation, 17.6ml of DMSO was added to the conical flask, mixed gently and aliquoted equally between ten 50ml falcon tubes. The cells were heat shocked at 42°C for six minutes and immediately transferred to a sterile 1 litre conical flask containing a 400ml volume of room temperature YPAD broth without glucose. The cells were pelleted at 960g in 250ml bottles, each pellet washed with 250ml YPAD without glucose and pelleted a second time. The pellets were resuspended in 1 litre of prewarmed YPAD and incubated in a 2 litre conical flask, with gentle shaking, at 30°C for 1 hour. The cells were then pelleted, resuspended in 500ml SD-Ura-Leu, pelleted a second time and finally resuspended in 1 litre of prewarmed SD-Ura-Leu. The culture was incubated, with gentle shaking, at 30°C for ~10 hours.

2.2.5.3 Transfection efficiency

The 1 litre culture was pelleted in 250ml bottles, washed twice with SD without glucose, pelleted again and resuspended in 10ml SD-Ura-Leu-Ade-His+3AT. 10µl of the cell suspension was removed and serially diluted (1:10, 1:100, 1:1,000, 1:10,000, 1:100,000 and 1:1,000,000) in six eppendorf tubes. Each of the dilutions was plated out onto a 10cm SD-Ura-Leu plate, using a sterile glass spreader to disperse the cells. The plates were incubated at 30°C for 3-4 days, at which point the number of colonies on each plate was counted.

2.2.5.4 First round of selection - Histidine

The remainder of the 10ml cell suspension was plated out in 550µl aliquots using a sterile glass spreader onto large (25cm) SD-Ura-Leu-His+3AT plates. The plates were then incubated at 30°C for 4-5 days.

2.2.5.5 Second round of selection - Adenine

After 4-5 days, the plates from the first round of selection were assessed for colony growth. Each of the positive colonies from the SD-Ura-Leu-His+3AT plates were streaked out onto individual grids of large, gridded SD-Ura-Leu-Ade plates using an inoculation loop.

2.2.5.6 Isolation of prey plasmid

Following the second round of selection, the prey plasmid was isolated to allow identification of the positive clone. To recover the pACT2 plasmid, a 5ml SD-Leu culture, inoculated with a scraping from the streak on the SD-Ura-Leu-Ade plate, was grown overnight at 30°C with shaking. The following day, the culture was centrifuged at 5,000g for 5 minutes, the pellet resuspended in the residual liquid and transferred to an eppendorf tube. 200µl of rescue solution, 100µl phenol:chloroform (1:1), followed by ~0.3g acid-washed glass beads (Sigma) was added and the contents of the tube mixed vigorously using a vortex for two minutes. The tubes were spun in a microcentrifuge at full speed (15,800g) for five minutes and the top phase (~200µl) applied to a QIAprep spin column (QIAGEN) in a 1.5ml collection tube. The column was spun at full speed for 1 minute and the flow-through discarded. 500µl of PB solution (QIAGEN) was added to the column, spun for 1 minute and the flow-through discarded. The column was then washed with 750µl of PE solution (QIAGEN), spun down, flow-through discarded and spun a second time to remove all traces of PE solution (QIAGEN) from the column. The column was transferred to a clean eppendorf tube for elution of the plasmid DNA. 50µl EB solution (QIAGEN) was added to the centre of the column, allowed to stand for 1 minute and then spun at full speed for 1 minute.

2.2.5.7 Amplification and identification of the prey cDNA sequence

To identify the clone, the insert was first amplified by PCR and sequenced. For amplification, a standard PCR reaction was carried out using ~2µg plasmid DNA and pACT2 primers (F - AAAGAGATCTGTATGGCTTAC, R - CAGTATCTACGATTCATAGATC).

A standard PCR program with an annealing temperature of 60°C was used (See *2.2.1.2 PCR reactions*). PCR reactions were cleaned-up and sequenced as described earlier (See *2.2.3 Direct sequencing*). The sequence text was then used to search the blastn database (www.ncbi.nlm.nih.gov/BLAST/Blast.cgi) to identify the cDNA insert.

2.2.6 Immunohistochemistry

2.2.6.1 Fixation of mouse tissues

Mouse eyes were removed from sacrificed wt female mice, washed in PBS (Gibco-BRL Life technologies) and fixed in 100% methanol, 2% paraformaldehyde (PFA), 4% PFA or Carnoy's fluid overnight at 4°C. Following fixation, the tissue was washed twice in PBS (10 minutes each) and then dehydrated through an ethanol series; 25% ethanol in PBS (60 minutes), 50% ethanol in PBS (60 minutes), 75% ethanol in PBS (60 minutes), 1:1 ratio 100% ethanol: Histoclear (RA Lamb) (30 minutes) and Histoclear (30 minutes). The tissue was then transferred to fresh Histoclear in a glass bottle for storage at room temperature.

2.2.6.2 Paraffin embedding and sectioning

Fixed tissues were embedded in paraffin following dehydration, for microtome sectioning. Tissues were transferred to Histoclear: wax (30 minutes), followed by three further incubations in wax. All wax incubations were carried out in a 55°C oven. In the final wax step, tissues were orientated appropriately and the wax allowed to cool and harden in the plastic cassette. When the paraffin block had hardened, the cassette was fitted into the microtome, 4 micron sections taken and placed on TESPA (3-aminopropyltriethoxysilane) coated glass slides. Slides were stored at room temperature.

2.2.6.3 Deparaffinisation of sections

Prior to staining, sections were deparaffinised in a Histoclear wash and ethanol series; Histoclear for 10 minutes, followed by 100% ethanol, 90% ethanol in PBS, 70% ethanol in PBS, 50% ethanol in PBS, 30% ethanol in PBS and water, each for five minutes.

2.2.6.4 Antigen retrieval

As the fixation and embedding processes can damage some antigens, an antigen retrieval step was also performed on some sections to determine whether it was required to improve the staining of BBS proteins. Following the deparaffinisation step, slides were placed, in a plastic slide holder, in a preheated solution of 0.01M citric acid and microwaved for 4, 7 or 10 minutes. The slides were then cooled slowly with running water.

2.2.6.5 Immunohistochemistry staining

Before application of the primary antibody to the sections, a series of blocking steps were performed. Firstly, endogeneous peroxidase activity within the sections was blocked using a solution of 3% hydrogen peroxide in methanol for 15 minutes, followed by two five minute rinses in water. A blocking solution containing 10% goat serum and 1% BSA in PBS was then applied to the sections to block non-specific binding of the secondary antibody and incubated at room temperature for 30 minutes. All antibody dilutions were made in block solution. The block was removed from the slides and the BBS4 or BBS8 antisera, diluted 1 in 500 or 1 in 1,000 respectively, were applied to the sections. The slides were placed in a humified box to prevent drying out of the sections and incubated at 4°C overnight. The next morning, the primary antibody was removed and the sides washed three times in PBT (five minutes each) and once in PBS. A 1 in 100 dilution of biotinylated goat anti-rabbit (DAKO) secondary antibody was then prepared and applied to the sections. The slides were returned to the humified box and incubated at room temperature for 30 minutes. During this incubation, the components of the ABComplex/HRP kit (DAKO) were mixed; reagent A

(avidin) and reagent B (biotinylated HRP) were both diluted 1 in 100 and allowed to stand at room temperature for 30 minutes. After removal of the secondary antibody, the slides were again washed three times in PBT and the ABComplex mixture applied, with a final incubation of 30 minutes at room temperature. Following the incubation, the slides were washed three times in PBT and then treated with a solution of DAB chromogen (1% DAB in water). The colour reaction was viewed under the microscope and, when a sufficient colour had developed (5-10 minutes), was quenched by plunging the slides into a trough of water. The slides were then allowed to air dry before mounting of the coverslip with DPX media (RA Lamb).

Chapter 3 Homozygosity and candidate gene screens

3.1 Introduction

The first step in identifying the gene responsible for a genetic condition is to localise the gene to a specific chromosomal region. In recessive conditions, this can be done either by linkage analysis in small and large nuclear pedigrees, or by homozygosity mapping in consanguineous pedigrees. Following the localisation of the disease gene and narrowing of the critical interval where possible, mutation screening of candidate transcripts is required to successfully identify the disease gene. Detection of pathogenic mutations within the ORF of a candidate gene or EST in pedigrees that where used to map the locus provides confirmation that the gene is a *BBS* gene.

3.1.1 Linkage analysis

According to Mendel's second law, the law of independent assortment, the segregation of alleles of one gene is independent of the segregation of alleles at another locus. This law applies to loci that are physically well separated, but is not true of genes that are in close proximity to each other. During meiosis (Figure 3. 1), homologous chromosomes pair up and undergo recombination, or crossing-over, which involves an exchange of chromosomal material between one chromatid of each homologous chromosome. Following recombination, two of the four gametes that are formed will contain original chromosomes and two will contain a modified chromosome that is made up of part of the original chromosome and part of the homologous chromosome. The proportion of recombinant chromosomes present after meiosis is referred to as the recombination fraction, θ, and is used as an estimate of genetic distance. As only two of the four chromatids are involved in recombination, θ has a maximum value of 0.50 for two loci that are far apart on the same

chromosome, or on different chromosomes, and a minimum of 0 for loci that are adjacent to one another.

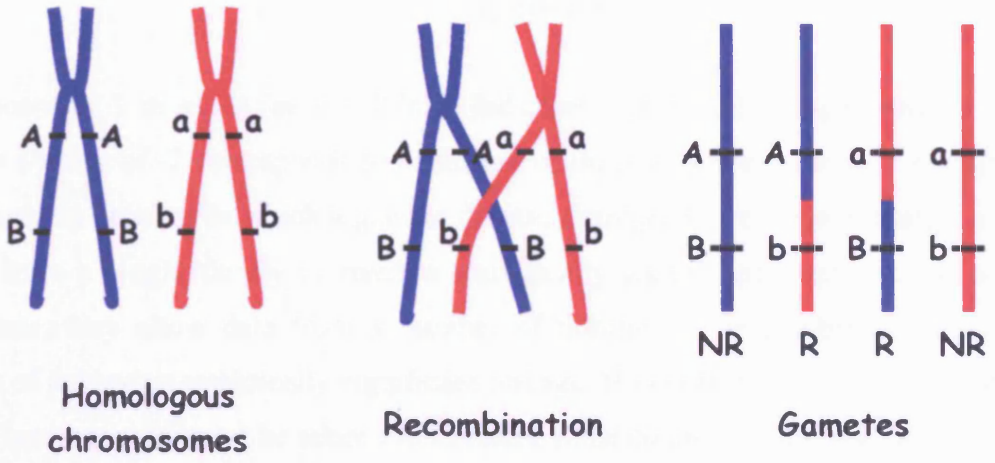


Figure 3. 1: The process of recombination. During meiosis recombination occurs at random along the length of the chromosome. In this example the individual is a double heterozygote for two loci on the same chromosome (AaBb). Following a recombination event, two non-recombinant (NR) gametes (AB and ab) and two recombinant (R) gametes (Ab and aB) will be formed.

Linkage analysis is the study of the inheritance of two or more characteristics, such as a disease phenotype and a polymorphic marker, to determine if the two co-segregate more often than would be expected under Mendel's second law. If two characteristics are found to have a recombination fraction of <0.5, they are considered to be genetically linked. The object of linkage analysis is therefore to estimate θ and determine whether it differs significantly from 0.5. In simple cases, it is possible to estimate the value of θ by counting the proportion of recombinant individuals in the total number of offspring, but linkage analysis in humans is rarely that simple. In disease gene mapping, as it is often not possible to count the number of recombinant and non-recombinant offspring, methods such as the maximum likelihood method, or lod score method, are commonly used (Morton, 1955). Computer programs are used to calculate the likelihood of a given pedigree under different recombination fractions, L (θ). This calculated likelihood is then tested against the likelihood of observing the given pedigree under the hypothesis that the two loci are unlinked, L (θ = 0.5).

In practice the common logarithm of the ratio of these two likelihoods is taken, giving the lod score:

$$Z(\theta) = \log_{10} \frac{L(\theta)}{L(\theta = 0.5)}$$

A lod score of 3 or more (or $P < 0.001$) indicates significant linkage between two loci, whereas a value of -2 corresponds to exclusion of linkage between the two. Pedigree sizes, particularly in families in which a genetic disease segregates, are often not large enough for results from a single family to reach a statistically significant level. As lod scores are logarithms, they allow data from a number of families to be combined, improving the chances of achieving statistically significant linkage. However, the possibility of non-allelic genetic heterogeneity must be taken into account when combining linkage data from several different families. To remove this problem from linkage calculations, large pedigrees, that are themselves big enough to generate a statistically significant lod score, are required as locus homogeneity can only reliably be assumed within a single family.

Alternative gene mapping approaches include the affected sib pair (ASP) method and homozygosity mapping in consanguineous pedigrees. ASP is based on the assumption that on average a sib pair will share 0, 1 or 2 parental haplotypes with the respective frequencies of 0.25, 0.5 and 0.25. In an affected sib pair, chromosomal regions associated with the disease are expected to be shared at frequencies that deviate from this ratio; if the disease is dominant, affected sibs will share at least one parental haplotype, and if the disease is recessive, they will share both haplotypes more frequently (Strachan and Read, 1999). As a genetic model does not have to be specified to use this method (nonparametric), it is very useful for studying a disease for which the inheritance pattern is unclear. The requirement of a large number of affected sib pairs for detection of significant linkage and the often large size of candidate regions identified through this method make it unsuitable for use in our patient cohort. Homozygosity mapping has been used successfully to map a number of *BBS* loci (Carmi et al., 1995b; Katsanis et al., 2000; Kwitek-Black et al., 1993; Sheffield et al., 1994; Young et al., 1999a).

3.1.2 Homozygosity mapping in BBS

In a consanguineous pedigree, an alternative to the lod score method for gene mapping can also be used. Individuals with a rare recessive disease are likely to be IBD for markers linked to the disease locus due to inheritance of the disease-associated haplotype from a recent common ancestor (Figure 1. 3). A genome screen, using polymorphic markers spaced throughout the genome, followed by haplotype analysis to identify haplotypes common to all affected individuals, can therefore be used to identify disease loci. This technique was used to map a number of the *BBS* loci, both in Bedouin pedigrees (*BBS2, 3* and *4* (Carmi et al., 1995b; Kwitek-Black et al., 1993; Sheffield et al., 1994)) and also in Newfoundland pedigrees (*BBS5* and *6* (Katsanis et al., 2000; Young et al., 1999a)). For a summary on the mapping of the individual loci, see *1.9.1 Mapping of BBS1-5.*

Shortly after the cloning of *BBS6*, Beales *et al.* (2000) performed haplotype analysis of a cohort of 163 pedigrees to estimate the contribution of each of the known loci to all cases of BBS and to refine the critical intervals of *BBS1-5*. All known loci were accounted for within the cohort, but 14% of pedigrees were found to be unlinked to *BBS1-6*. A further 28% were unable to be assigned to the known loci due to small pedigree size and/or uninformative markers in some of the critical intervals; a proportion of these families may also be unlinked. These results provide strong evidence for additional *BBS* loci within the genome.

3.1.3 Positional cloning of mapped *BBS* genes

The process of identifying a disease phenotype, mapping of the locus and cloning of the gene was initially referred to as 'reverse genetics' (Orkin, 1986; Ruddle, 1984). These methods are in fact part of a classical genetics approach to disease gene identification. The transgenic or 'knock-out' method, in which mutations are introduced into a novel gene that is then reintroduced into a model organism such as the mouse, is an example of true reverse genetics (McKusick, 2002). For this reason, in 1992, Francis Collins recommended that the phenotype to locus to gene approach be known as positional cloning (Collins, 1992).

Mapping of a disease locus allows the search for the causative gene to be concentrated on a particular chromosomal region. Depending on the size of the interval and the density of genes, a large number of genes may exist within the critical interval. As screening of all genes within an interval would be a costly and time-consuming process, it is necessary to prioritise the candidate genes for mutation screening. This can be done using a number of different criteria (Strachan and Read, 1999):

- *Appropriate expression pattern* – A good candidate gene should have an expression pattern consistent with the disease phenotype. Expression of the gene does not have to be restricted to the tissues or organs that are affected by the disease (such as the retina, kidney, developing limb bud and gonads in the case of BBS), but must be expressed in the relevant areas. Analysis of the *BBS1* critical interval for possible candidate genes revealed that the *ROM1* gene, which is expressed in rod outer segments, was located within the interval. However, as the pattern of *ROM1* expression is very specific and restricted to only the retina, the gene was not considered to be a suitable candidate for BBS (Leppert et al., 1994).

- *Appropriate function* – If the function of a gene within a critical interval is known, it can be used to determine if the gene is a good candidate for the disease under study. The association of a gene with a similar disease phenotype also indicates that the gene is a good candidate. Due to both the clinical and mapping overlap between MKKS and BBS, the *MKKS* gene was proposed as a candidate for BBS. Detection of pathogenic mutations in BBS patients, by two independent groups, provided confirmation that the *MKKS* gene was *BBS6* (Katsanis et al., 2000; Slavotinek et al., 2000).

- *Homology to a known gene* – A high degree of similarity between a candidate gene and a gene that is already known to be associated with the disease is strong evidence for the candidate gene also being associated with the same disease. Based on its homology at the protein level to BBS2, the novel transcript encoding BBS1 was considered to be a strong candidate; pathogenic mutations in a number of families confirmed that this gene was *BBS1* (Mykytyn et al., 2002).

- *Homology to a gene in a model organism* – If a mouse model of a disease exists, for which the causative gene has been identified, it is highly likely that the orthologous gene accounts for the disease in humans.

3.1.4 Conclusions

Previous linkage studies to map *BBS1-6* highlighted the extensive genetic heterogeneity in BBS and also provided evidence that further loci for the disease must exist within the human genome. In order to understand more about the molecular mechanisms involved in this pleiotropic disorder and how, if at all, the high degree of heterogeneity at the genetic level contributes to the extensive variation in the clinical phenotype, it is essential to identify the genes responsible for the disease.

3.2 Materials and methods

3.2.1 DNA samples

Patients from 17 consanguineous pedigrees were genotyped in the genome-wide homozygosity screen. Ten of the pedigrees contained only a single affected offspring, the remaining seven contained two or more affected sibs. In two pedigrees with multiple affected sibs (PB011 and PB016), due to poor DNA quality for some patient samples, DNA from only one affected was included in the sample plate. If an area of homozygosity was identified in either PB011 or PB016, where possible, the remaining affected offspring were also genotyped for the relevant markers. All pedigrees were of Indian, Kurdish, Pakistani or Turkish origin. See Table 3. 1 for a summary of the pedigrees used. DNA from the 23 patient samples was extracted and aliquoted as described in *2.2.1.1 DNA extraction from blood.*

Pedigree	No. of affecteds	Origin
PB011*	2	Kurdish
PB012	3	Pakistani
PB016*	3	Kurdish
PB018	2	Kurdish
PB026	1	Indian
PB030	2	Pakistani
PB031	2	Pakistani
PB035	2	Pakistani
PB036	1	Turkish
PB040	1	Turkish
PB041	1	Turkish
PB042	1	Kurdish
PB049	1	Turkish
PB050	1	Pakistani
PB051	1	Pakistani
PB052	1	Turkish
PB053	1	Turkish

Table 3. 1: Summary of pedigrees used in the genome screen. * Indicates a pedigree with multiple affected sibs, but only one affected offspring was included in the DNA sample plate.

3.2.2 Microsatellite markers

Amplification, pooling and gel electrophoresis of the markers used in the homozygosity screen was performed as described (See *2.2.2 Homozygosity mapping*), followed by analysis using the Gene Scan v.3.1.2 and Genotyper v.2.5 programs (Applied Biosystems) (See *2.2.2.5 Genotyping of markers*)

3.2.3 Candidate gene identification and screening

Following identification of a region of homozygosity common to a number of pedigrees, the area was analysed for the presence of possible candidate genes using the previously described criteria (See *3.1.3 Positional cloning of mapped BBS genes).* In order to determine exon-intron boundaries, the genomic sequence was first identified by screening of the nr human genome database with the cDNA using the blastn function of BLAST (www.ncbi.nlm.nih.gov/BLAST/Blast.cgi). The genomic and cDNA sequences were then aligned using the bl2seq function (www.ncbi.nlm.nih.gov/blast/bl2seq/bl2.html). Primers were designed to amplify the ORF of the candidate gene as described (See *2.1.3 Primer design*). Affected individuals from a pedigree that demonstrated IBD to the area were then screened for pathogenic mutations by direct sequencing (See *2.2.3 Direct sequencing*).

The same approach was also used to identify and screen candidate genes within the critical intervals of the known *BBS* loci (*BBS1-5*).

3.2.4 Candidate genes screened

Two candidate genes located in a region of homozygosity identified through the genome screen (14q23.3), were sequenced in pedigrees with haplotypes consistent with linkage to this area:

- *HSPA2* (heat shock 70kDa protein 2) – HSPA2, a heat shock protein with chaperone activity, was considered to be a good candidate for BBS based on its similarity to BBS6.

- *SKIP* (SKI interacting protein) – SKIP is a nuclear matrix-associated coactivator that can bind to retinoid receptors to enhance retinoic acid and oestrogen mediated gene expression. This gene was considered to be a candidate for BBS as it was found to interact with BBS4 on yeast-two-hybrid analysis (See *7.4.3.3 Ski-interacting protein (SKIP)*).

Two candidates for the mapped *BBS* genes were also screened:

- *STIP1* (stress-induced-phosphoprotein 1) – *STIP1*, encoding an Hsp70/Hsp90-organizing protein, was screened as a candidate for *BBS1* based on its similarity to BBS6.

- *TLE3* (transducin-like enhancer of split 3) – Transducin is a G-protein found specifically in rod outer segments where it mediates the activation of cyclic GTP-specific phosphodiesterase by rhodopsin. Based on its similarity to transducin, *TLE3* was considered to be a good candidate for *BBS4*.

3.3 Results

3.3.1 Homozygosity screen

Three possible regions of homozygosity, shared between a number of families, were found during the genome-wide homozygosity screen.

3.3.1.1 Chromosome 4q25-31.2

Genotyping of chromosome 4q markers from the ABI marker set revealed two consecutive markers, D4S402 and D4S1575, that were homozygous in four pedigrees (Table 3. 2). Marker D4S402 was homozygous in seven of the eight affecteds from these pedigrees (results were not achieved for this marker in individual PB016.2). Marker D4S1575 was homozygous in both sibs from pedigree PB031, but was homozygous in only one sib in pedigrees PB016, PB018 and PB030. Genotyping of additional markers from the RG set revealed homozygosity in seven of the eight patients for marker D4S2394, located between the two ABI markers. Two flanking markers, D4S2623 and D4S1644, were heterozygous in seven and five patients respectively. The minimum stretch of homozygosity shared by all individuals from the four pedigrees is between markers D4S2623 and D4S2394, a length of ~19.4Mb.

3.3.1.2 Chromosome 8q23.3-24.13

As several of the ABI chromosome 8q markers in this region did not amplify, RG markers were also used. A potential region of homozygosity was identified across three consecutive markers, D8S1142, D8S592 and D8S514 in five pedigrees (Table 3. 3). Markers identified through GDB (D8S1694, D8S522 and D8S1802) were also genotyped in order to increase the density of markers across this interval. Heterozygosity of one or more of these additional markers in most individuals, restricts the area of homozygosity in this region to between markers D8S1694 and D8S522, a ~1.2Mb distance.

3.3.1.3 Chromosome 14q22.3-24.3

Homozygosity of ABI markers in five pedigrees lead to genotyping of further markers in this area of chromosome 14q (Table 3. 4). Although the area of homozygosity stretches for ~23.0Mb between D14S276 and D24S74 in two pedigrees (PB031 and PB041), the minimum region of homozygosity, shared between all five pedigrees, is restricted to a ~5.6Mb distance between D14S63 and D14S588. The *HSPA2* gene was found to be located within this interval, very close to D14S63 (0.3Mb distal to the marker), and was considered to be a good candidate for BBS as the protein is a heat shock protein with chaperone activity. *SKIP*, located within the wider interval of homozygosity present in four of the pedigrees (PB031, PB041, PB049 and PB053), was also screened as the SKIP protein was found to interact with BBS4 (See *7.4.3.3 Ski-interacting protein (SKIP)*).

3.3.2 Candidate gene screen

No pathogenic mutations were identified in *HSPA2*, *SKIP*, *STIP1* or *TLE3*.

Patient ID	D4S2623 (RG)	D4S402 (ABI)	D4S2394 (RG)	D4S1575 (ABI)	D4S1644 (RG)
PB016.2	223/231	-	209	291/293	192/196
PB016.3	223	129	209	293	192/196
PB018.1	211/223	111	206/209	291/293	192/200
PB018.2	211/223	111	209	291	192/200
PB030.1	203/207	105	209	289/293	192/200
PB030.2	203/207	105	209	293	196
PB031.1	211/227	111	209	293	196
PB031.2	211/227	111	209	293	196

Table 3. 2: Genotyping results for chromosome 4q25-31.21 markers. Single allele sizes in red indicate homozygosity for the corresponding allele. - Represents no results.

Patient ID	D8S1471 (RG)	D8S1142 (RG)	D8S1694 (GDB)	D8S592 (RG)	D8S522 (GDB)	D8S1802 (GDB)	D8S514 (ABI)
PB026.1	277/297	354	246/254	161	217/219	275	216/218
PB031.1	289/297	354	252	165	213	279	216/218
PB031.2	289/237	354	252	165	213	279	216/218
PB035.1	293/297	362	192/200	157	213	279/281	216
PB035.2	297	354	192	-	211/213	281	216
PB036.1	281/289	354	254/256	169	219	275/279	216
PB041.1	285/289	350/362	196/200	161	219/223	275/277	218

Table 3. 3: Haplotypes for chromosome 8q23.3-24.13 markers. An initial interval of homozygosity spanning three markers (D8S1142, D8S592 and D8S514) was found to not be continuous when the density of markers in this region was increased.

Patient ID	D14S276 (ABI)	D14S592 (RG)	D14S63 (ABI)	D14S588 (RG)	D14S258 (ABI)	D14S74 (ABI)
PB031.1	235/239	247	180	121	199	295/299
PB031.2	235/239	247	180	121	199	295/299
PB040.1	237	250	180	117/121	193/199	-
PB041.1	237	247	184	121	189	299
PB049.1	239/243	241/247	182/186	125	195	301/305
PB053.1	239	235/247	180	119/121	191	289

Table 3. 4: Chromosome 14q22.3-24.3 haplotypes. Homozygosity across a 5.3Mb interval between markers D14S592 and D14S63 in five pedigrees lead to screening of the candidate gene *HSPA2*. No mutations were identified.

3.4 Discussion

3.4.1 Homozygosity screen

Evidence for the existence of additional *BBS* loci has been provided in a number of recent studies on the contribution of the known loci to BBS (Beales et al., 2001; Fauser et al., 2003; Katsanis, 2004). Using a subset of consanguineous pedigrees from our cohort that had either been excluded from linkage to the known loci in a previous study (Beales et al., 2001), or were too small to perform linkage analysis, a genome-wide homozygosity screen was performed to map a novel *BBS* locus. Patients from 17 different pedigrees of Indian, Pakistani or Turkish origin were studied using markers spaced across the whole genome. No significant regions of homozygosity shared amongst all pedigrees were identified, indicating that a single novel *BBS* locus does not account for all currently unmapped cases of BBS in pedigrees of Middle Eastern and Asian origin. Taking into account the high degree of genetic heterogeneity observed in previous linkage screens and the relatively small contribution of the majority of the loci to disease (Katsanis, 2004), it is most likely that a number of rare loci are responsible for disease in these pedigrees.

Three regions of homozygosity requiring further study were identified on the genome-wide screen. In one of these regions, following genotyping of additional markers, the area of IBD was found to only be continuous in one of the five pedigrees initially thought to exhibit IBD across the region. Suitable candidate genes for mutation screening were only identified in one of the regions; no pathogenic mutations were detected.

3.4.1.1 Chromosome 4q25-31.2

An area of homozygosity on chromosome 4q was present in four pedigrees, each of which contained two affecteds. Genotyping results for the ABI marker set in all pedigrees revealed that marker D4S402 was homozygous in seven patients from four different pedigrees (PB016, PB018, PB030 and PB031, see Table 3. 2). The distal marker, D4S1575, was also homozygous in both affected sibs from pedigree PB031, but in the remaining three pedigrees,

one sib was homozygous for the marker and the other, heterozygous. Genotyping of additional markers from the RG set confirmed that the IBD region did not extend beyond D4S1575 in these three pedigrees, as five of the six patients were heterozygous for marker D4S1644. Heterozygosity for marker D4S2394, located between D4S402 and D4S1575, in individual PB018.2, suggested that this marker is the distal boundary of the IBD region shared by all four pedigrees. It was also possible to define the proximal boundary of IBD using RG markers; marker D4S2623 was homozygous in only a single patient (PB016.2).

Although extending further in three pedigrees, the shared region of IBD on chromosome 4q lies between markers D4S2623 and D4S2394 (4q25-28.2), a distance of ~19.4Mb. Analysis of known genes in this area at the time of the genome screen did not reveal any candidate genes for BBS that demonstrated an expression pattern consistent with the clinical phenotype and/or similarity to *BBS6*, the only cloned *BBS* gene at the time. After completion of the genome screen, the *BBS7* gene was cloned on 4q27 based on its homology to BBS2 (Badano et al., 2003a). As *BBS7* is located within this region of IBD, pedigrees PB016, PB018, PB030 and PB031 were screened for mutations in the *BBS7* ORF. No coding sequence alterations were found in any patients from the four pedigrees.

There are a number of possible explanations for these results:

- *BBS7* mutations that were not detected by sequencing may exist in these pedigrees. Although this is possible, it is unlikely that all four pedigrees carry a regulatory element mutation or large deletion or rearrangement in *BBS7*.
- A second *BBS* gene is located in this area, in close proximity to *BBS7*. To date there has been no evidence of clustering of *BBS* genes in the human genome, but as more novel *BBS* genes are cloned, incidences of clustering may yet be found.
- The identification of a number of homozygous markers in the same chromosomal location as a recently cloned *BBS* gene could also be coincidental and *BBS7* may be the only *BBS* gene on 4q27.

Genotyping of additional markers in this area and a further analysis of known genes and ESTs in the area would be required to determine if this region does contain a novel *BBS* gene.

3.4.1.2 Chromosome 8q23.3-24.13

Evidence for a small region of homozygosity on chromosome 8q was also found in five pedigrees, although results from additional markers did not support the initial evidence for IBD (See Table 3. 3). Homozygosity of two or all of the consecutive markers D8S1142, D8S592 and D8S514 in five pedigrees (PB026, PB031, PB035, PB036 and PB041) lead to the generation of additional markers identified through GDB. Although results from these markers confirmed a ~10Mb stretch of IBD in PB031 (from D8S1471 to D8S514), heterozygosity was observed for one or more of the GDB markers in each of the other pedigrees.

These results suggest therefore that the area of homozygosity common to all five pedigrees is restricted to only a 1.2Mb interval between D8S1694 and D8S522 (8q24.11-8q24.12). This may correspond to a novel *BBS* locus but, in some pedigrees (PB026, PB035 and PB041), the interval represents homozygosity at only a single marker, D8S592. As the heterozygosity of this marker is relatively low (0.67), it is possible that the observation of homozygosity for this marker in five of 17 pedigrees is coincidental. Genotyping of further markers with higher heterozygosities is required to establish whether this is in fact a true *BBS* locus.

3.4.1.3 Chromosome 14q22.3-24.3

Haplotypes from chromosome 14q revealed a number of homozygous markers in six individuals from five different families (Table 3. 4). The distal boundary of the region was defined by heterozygosity for the marker D14S588 in two patients (PB040.1 and PB053.1), and the proximal boundary by marker D14S63 that was heterozygous in PB049.1. Analysis of this ~5.6Mb interval for suitable candidate genes, lead to the identification of *HSPA2*, a gene encoding a heat shock protein (HSPA2), located 0.3Mb distal to D14S63. Mutation

screening of the gene in pedigrees exhibiting linkage to this area did not reveal any pathogenic coding mutations. The *SKIP* gene, located just proximal to D14S74 in the extended region of IBD that is shared between four of the pedigrees (PB031, PB041, PB049 and PB053), was also sequenced. Sequence analysis of the ORF in these pedigrees did not result in the detection of any mutations.

3.4.1.4 Marker density

It is possible that novel *BBS* loci have been missed in this screen due to a lack of complete coverage of the genome. Despite an average spacing of 10cM between loci in the marker set used for the screen (ABI PRISM, Applied Biosystems), some areas of the genome were not well covered. Gaps between markers were greater than 20cM in 11 cases and in four of these, the distance was greater than 25cM. On chromosome 9q32 there was only a single marker in an interval of 41.7cM and on chromosome 8q24, the neighbouring markers to D8S258 were a distance of 19.1cM (proximal) and 24cM (distal) away.

3.4.2 Candidate gene screens

It was initially expected that the cloning of the first *BBS* gene would aid the identification of the remaining mapped genes. As BBS6 shows similarity at the amino acid level to the chaperonin family (Stone et al., 2000), genes within the critical intervals of *BBS1-5* that also demonstrated this similarity were considered to be good candidates (Sheffield et al., 2001). In addition to screening of two candidate genes identified through the homozygosity screen, genes were also screened as possible candidates for *BBS4* (*TLE3*) and *BBS1* (*STIP1*). Since the candidate gene screen, both *BBS4* and *BBS1* have been cloned (Mykytyn et al., 2001; Mykytyn et al., 2002) and were found to be novel transcripts of unknown function. As *BBS2*, *7* and *8* have now also been cloned (Ansley et al., 2003; Badano et al., 2003a; Nishimura et al., 2001) and more information about the probable function of some of the BBS proteins is known, new criteria when selecting candidate genes for BBS is available. Although expression levels within the affected organs such as retina and kidney are high, all the

recently cloned *BBS* genes were found to be widely expressed and have not shown any similarity to chaperones (Ansley et al., 2003; Badano et al., 2003a; Mykytyn et al., 2001; Mykytyn et al., 2002; Nishimura et al., 2001).

3.4.3 Mutations identified in unlinked pedigrees

To assess the contribution of the cloned *BBS* genes to our cohort of patients, all patients, including those from the homozygosity screen, were screened for mutations in *BBS1*, *2*, *4*, *7* and *8* by direct sequencing (See *Chapter 4 Mutation screening of known BBS genes and identification of a novel gene BBS8*). Of the 17 consanguineous pedigrees from the genome screen, one (PB026) was found to carry a homozygous 2bp deletion in *BBS2*, a second (PB012) defined the *BBS8* locus and in a third (PB053), a single mutation in *BBS1* was identified (See Table 4. 1). The remaining 14 pedigrees do not carry mutations in any of the known *BBS* genes.

3.4.4 Summary

No major novel *BBS* loci were detected in our cohort of consanguineous pedigrees. Three regions of IBD were found to be shared amongst a small number of families but further work is required with additional highly informative markers to establish whether these regions represent novel loci, or are merely chance associations of homozygosity for consecutive markers in different pedigrees. Further information on the structure and function of the BBS proteins will allow more accurate selection criteria to be used when looking for candidate genes for BBS.

Chapter 4 Mutation screening of known *BBS* genes and identification of a novel gene, *BBS8*

4.1 Introduction

Since the identification of pathogenic mutations in the newly cloned *MKKS* gene (Stone et al., 2000) in BBS patients, defining the *BBS6* gene (Katsanis et al., 2000; Slavotinek et al., 2000), considerable progress has been made in the identification of *BBS* genes; three more of the known *BBS* genes (*BBS1*, *2* and *4* (Mykytyn et al., 2001; Mykytyn et al., 2002; Nishimura et al., 2001)) have now been identified and a novel, unmapped, gene (*BBS7* (Badano et al., 2003a)) has also recently been cloned (See Appendices 1-6 for schematics of *BBS1*, *2*, *4*, *6*, *7* and *8*) . *BBS1*, *2* and *4* were identified using a conventional positional cloning approach involving mutation screening of candidate genes and ESTs within the critical intervals of the respective loci. The presence of pathogenic coding mutations in the consanguineous pedigrees that were used to map the loci and in additional unrelated families (in the case of *BBS1* and *2*) confirmed that mutations in these novel genes result in BBS. In contrast, the *BBS1* gene was a strong candidate owing to its location within the critical interval and its shared homology with the previously identified BBS2 protein (BBS1 and BBS2 are 40% similar over a 192-amino acid region, (Mykytyn et al., 2002)). *BBS7* was also suggested as a candidate gene for BBS based on its similarity to known BBS proteins (BBS7 is 42.5% similar to BBS2 across a 252-amino acid region which partly overlaps with the region of shared homology between BBS2 and BBS1); two missense mutations and one frameshift mutation in a cohort of BBS patients confirmed *BBS7* as a novel *BBS* gene (Badano et al., 2003a).

4.1.1 Spectrum of mutations reported in *BBS* genes

Several types of mutations including missense, nonsense, splice junction, frameshift and whole exon deletions have been reported in *BBS* genes (Mykytyn et al., 2001; Mykytyn et al., 2003; Mykytyn et al., 2002; Nishimura et al., 2001). To date no whole gene deletions have been reported in any of the known genes. In addition, no nonsense mutations have been reported in either *BBS4* or *7*, but as these loci account for only a small percentage of cases of BBS (the combined contribution of *BBS4* and *7* is estimated to be <5% (Katsanis, 2004)) this is likely to be purely a reflection of the small number of mutations reported in these genes; frameshift and missense mutations have been reported in all *BBS* genes (Badano et al., 2003a; Katsanis et al., 2000; Mykytyn et al., 2001; Mykytyn et al., 2003; Nishimura et al., 2001).

A 6kb deletion of exons 3 and 4 (IVS2_IVS5) has been reported in *BBS4* in two pedigrees, one Italian and the other Israeli-Arab (Mykytyn et al., 2001). Although the deletion breakpoints in the two families were identical, haplotype analysis of the pedigrees using microsatellite markers surrounding the *BBS4* gene indicated that the mutations had occurred independently. As the breakpoints are located in *Alu* repeat sequences in introns 2 and 4 it is likely that this deletion occurred through unequal homologous recombination between *Alu* sequences. Deletions or duplications involving *Alu* sequences have been reported in multiple genetic conditions and may account for as much as 0.3% of all human genetic disease (Deininger and Batzer, 1999). Analysis of the *BBS4* sequence by Mykytyn *et al.* (2001) identified 40 *Alu* sequences throughout the gene, with a particularly high frequency in intron 2 (12 *Alu* sequences were identified in this intron alone); the authors therefore suggest that *BBS4* may be predisposed to *Alu*-associated mutations and that further mutations of this type are likely to be identified in future mutation screens of the gene. High numbers of *Alu* repeats have not been reported in other *BBS* genes.

In addition to a variety of types of mutation reported, the distribution of mutations throughout the ORF of the *BBS1*, *2*, *4* and *6* genes appears to be relatively uniform, with little or no indication of mutation hotspots within the genes. Two nonsense mutations in exon 8 of

BBS2 (R272X and R275X) were however reported by Nishimura *et al.* (2001), and none of the five mutations reported in this study were located in the C-terminal third of the protein. It is possible that the close proximity of the nonsense mutations and high incidence of mutations towards the N-terminus is a true representation of the distribution of mutations within the *BBS2* gene, but it may also be coincidental and influenced by the small number of mutations detected in the study. In contrast to the other *BBS* genes, *BBS7* does appear to show clustering of mutations towards the middle of the protein, but again the sample size is very small (Badano et al., 2003a). Only three mutations have been reported in *BBS7* to date; all three are located within a 112-amino acid stretch in the centre of the 621-amino acid protein, in a region of shared homology with BBS2.

With the exception of M390R, a common missense mutation in *BBS1* (M390R accounts for 80% of cases of *BBS1* and over 30% of all cases of BBS (Mykytyn et al., 2003)), the majority of *BBS* mutations are either private (reported in a single family) or segregate with disease in only a small number of pedigrees. The M390R mutation has primarily been reported in patients of North European origin but Mykytyn *et al.* (2002) has also reported the mutation in affected members of Puerto Rican BBS families in both compound heterozygote (in association with other *BBS1* mutations) and homozygote individuals. Haplotype analysis of the M390R mutation suggests that it is an ancient mutation as it is present on a single haplotype both within and across populations (Mykytyn et al., 2003).

4.1.2 *TTC8* as a candidate gene for BBS

Despite the recent success in the cloning of *BBS* genes, mutations in the five known genes do not account for all cases of BBS. The percentage of patients that remain unlinked to any of the known loci has been reported to range from 14-42% (Badano et al., 2003a; Beales et al., 2001; Katsanis et al., 2001b) and in one particular study, only 43% of patients were accounted for by mutations in *BBS1*, *2*, *4* and *6* (Fauser et al., 2003). With so many patients unaccounted for by the known loci, further *BBS* loci must be present within the human genome. It is possible that there is only one further locus that accounts for all currently

unlinked cases of BBS but, as no significant regions of homozygosity were detected in the genome-wide homozygosity screen (See *3.3.1 Homozygosity screen*), this is unlikely. The existence of a number of additional, less common, loci is therefore more likely. In the absence of extended consanguineous pedigrees, that are themselves large enough to generate statistically significant linkage, mapping of novel, rare *BBS* loci is very difficult. Following the successful cloning of *BBS1* and *BBS7* based on their homology to BBS2, this alternative method is an effective way of cloning new *BBS* genes, particularly in a cohort of patients from small, non-consanguineous pedigrees.

In a similar approach to that used by Badano *et al.* (2003a) to clone the *BBS7* gene, fragments of the BBS4 protein sequence were used to screen the conceptual translation of the human genome and dbEST. As BBS4 contains at least ten tetratricopeptide repeats (TPRs) that are involved in protein-protein interactions, several TPR containing proteins were identified. One particularly good match, the hypothetical protein TTC8, showed a stretch of homology that included three consecutive TPR domains and, although not located within a mapped *BBS* locus (the *TTC8* gene maps to chromosome 14q32.11), was considered to be a strong candidate for BBS based on this degree of homology to BBS4 (Ansley et al., 2003).

4.1.3 Conclusions

To determine the contribution of the known loci to our patient cohort and potentially identify novel mutations within the known *BBS* genes, 120 affected individuals were screened for mutations in *BBS1*, *2*, *4* and *7* by direct sequencing. The same patient cohort was also screened for mutations within the ORF of *TTC8* to establish whether this was a novel *BBS* gene.

4.2 Methods

4.2.1 Patient Cohort

A cohort of 120 affected individuals was screened for mutations in *BBS1*, *2*, *4* and *7*. A large number of this patient cohort had previously been screened for mutations in *BBS6* during an earlier study (Beales et al., 2001). Patients recruited after that study were screened for *BBS6* mutations using the MCHA technique (See *6.3.2 Identification of alterations in new cases of BBS and in NPHP patients*). All patients were screened for the four genes regardless of any haplotype inferred chromosomal assignment from previous linkage studies. Patient DNA samples were either received directly from clinicians or were extracted from blood samples as described earlier (See *2.2.1.1 DNA extraction*).

4.2.2 PCR and cycle sequencing reactions

All PCR, sequencing reactions and sequence analysis were carried out in a 96-well format as described earlier (See *2.2.1.2 PCR reactions* and *2.2.3 Direct sequencing*). Following identification of a mutation in a patient, all available members of the family were screened to check the segregation pattern of the change. Where possible mutations were also confirmed using a restriction digest with the relevant enzyme (See *2.2.1.4 Restriction digests*); M390R (*BBS1*) – *Bsp*HI, Y24X (*BBS2*) – *Bfa*I and R275X (*BBS2*) – *Bsl*I.

4.3 Results

4.3.1 *BBS1* mutations

At least one mutant *BBS1* allele was identified in 30 of the 120 pedigrees (25.0%) screened by direct sequencing, confirming that *BBS1* is the most common *BBS* gene (Beales et al., 2003). Despite complete coverage of the ORF and splice junctions of the gene, single mutant alleles were detected in four of the 30 pedigrees (13.3%). Both previously published (missense, frameshift and nonsense) and novel (missense) mutations were identified in this patient cohort (Table 4. 1 and Appendix 1)

4.3.1.1 Previously published mutations

As described in previous mutation screens of *BBS1* (Mykytyn et al., 2001; Mykytyn et al., 2003), a common missense mutation (M390R) was the predominant *BBS1* mutation within our cohort (Table 4. 1). The majority of M390R mutations were found to segregate with disease in a recessive manner within small, outbred pedigrees of European descent. The M390R mutation, caused by a T>G transversion in exon 12, was seen in homozygous form in eleven isolated cases of BBS, for whom DNA from family members was not available, and also in six pedigrees (17/30, 56.7% of all *BBS1* mutations, (Beales et al., 2003)). Despite being unaffected, the fathers of two of these pedigrees (PB006 and PB029) were found to be homozygous for the M390R mutation. M390R alleles were also seen in compound heterozygous form in association with another *BBS1* mutation in four pedigrees (13.3%) and in heterozygous form without a second *BBS1* mutation in three pedigrees (10.0%). In one Turkish pedigree (PB086), three affected individuals in one branch of the family (individuals 07, 08 and 09) were found to carry a compound heterozygous mutation involving M390R and a c.1318C>T nonsense mutation in exon 13 (R440X, Table 4. 1). The affected individual in the other branch of the pedigree (01) carried only the R440X mutation, in homozygous form (Figure 4. 1). Interestingly, despite the fact that the parents of 01 (individuals 05 and 06) are first cousins, both R440X alleles in 01 have not been inherited from a common

ancestor. A heterozygous R440X mutation was also identified by MCHA in an isolated new case of BBS (F820, See *6.3.2 Identification of alterations in new cases of BBS and in NPHP patients*). A second compound heterozygous mutation involving M390R was found in association with a 1bp deletion in exon 10 (Y284fsX288) in a single English pedigree (PB008) comprising male monozygotic twins and an affected sister (Table 4. 1). The same frameshift mutation was also seen in homozygous form in a consanguineous Libyan pedigree (PB110). In three cases (PB074, PB097 and PB125), affected individuals were found to be heterozygous for the M390R mutation but no second pathogenic mutation within the *BBS1* ORF was identified.

4.3.1.2 Novel missense mutations

In addition to the identification of previously reported *BBS1* mutations within our patient cohort, novel mutations, all missense, were also identified. In a single Pakistani consanguineous pedigree (PB034), a c.442G>A substitution in exon 5 (D148N) was present in homozygous form in the two affected offspring and in heterozygous form in both parents (Table 4. 1). The aspartic acid at position 148 of the BBS1 protein is conserved in *Mus musculus*, *Rattus norvegicus*, *Pan troglodytes*, *Canis familiaris* and *Anopheles gambiae*. Two other novel missense mutations were found in compound heterozygous individuals in association with M390R; a c.157A>G substitution in exon 3 (K53E) was identified in an English pedigree (PB017) with two affected offspring, and a contiguous c.1508-1509TT>AC substitution (L503H) was identified in a single individual (PB077, Table 4. 1). A heterozygous c.104A>G missense mutation in exon 2 (H35R) was also identified in a consanguineous pedigree (PB053); in this pedigree no second mutation was identified (Table 4. 1). Like D148, the histidine residue at position 35 is conserved in *Mus musculus*, *Rattus norvegicus*, *Pan troglodytes*, *Canis familiaris* and *Anopheles gambiae*.

See Appendix 1 for a schematic showing the distribution of the known and novel *BBS1* mutations identified in the patient cohort.

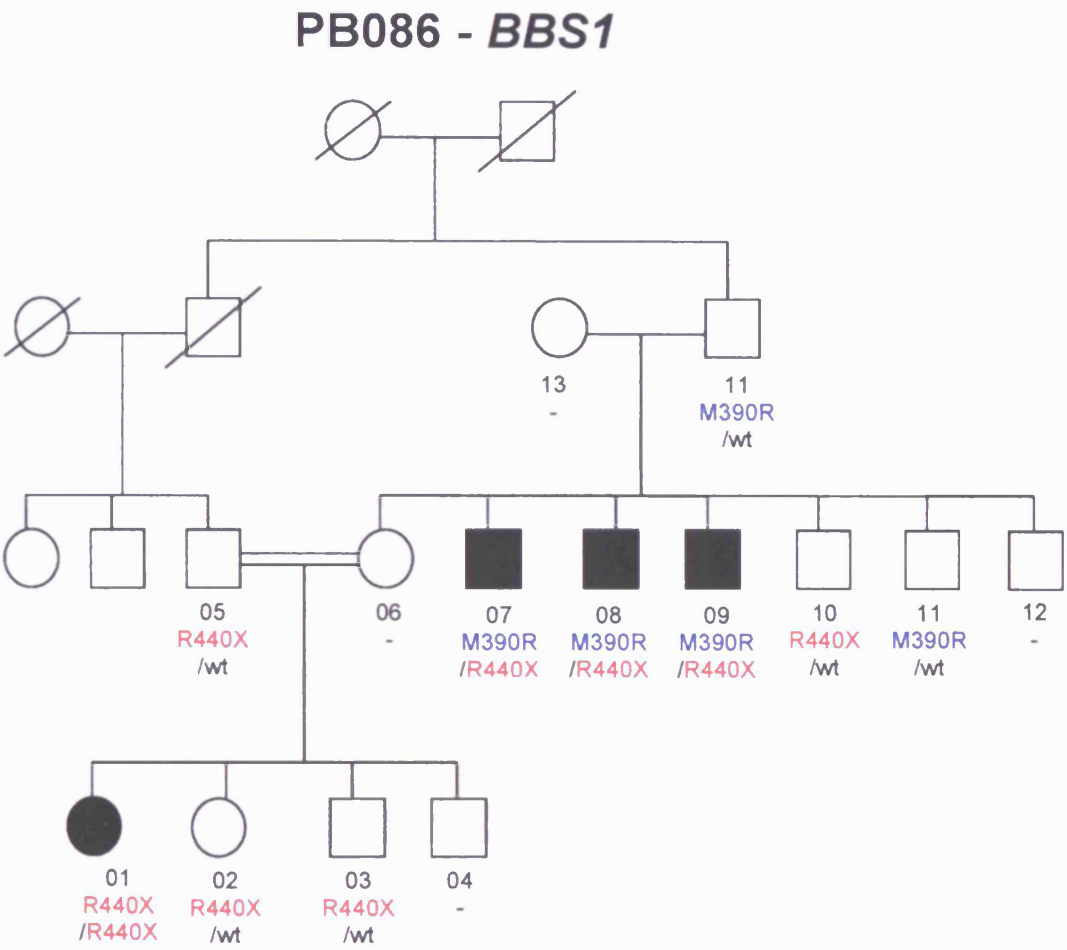


Figure 4. 1: Pedigree PB086. A consanguineous pedigree in which a homozygous nonsense mutation (R440X) segregates with disease in one branch of the family and in the other the heterozygous R440X mutation is associated with M390R.

4.3.2 *BBS2* mutations

Recessive mutations, including a deletion, nonsense and missense mutations, in *BBS2* were identified in six individuals (5.0%) from the patient cohort (Table 4. 1 and Appendix 2).

4.3.2.1 Previously reported mutations

Only a single known *BBS2* mutation was detected in our patient cohort. A C>T substitution in exon 8 (R275X), resulting in premature termination of the protein, was seen in homozygous form in one individual (PB065) and also in two affected sisters from a small English pedigree (PB005). This same mutation was also seen in association with heterozygous novel mutations in two compound heterozygous individuals (Table 4. 1).

4.3.2.2 Novel mutations

Three novel mutations were identified during the mutation screen of *BBS2* (Table 4. 1). A homozygous c.511-512delTT deletion in exon 4 (D170fsX171) was identified in the affected individual in a consanguineous Indian pedigree (PB026, Table 4. 1). A c.72C>G nonsense mutation in exon 1 (Y24X) was identified in homozygous form in a single family (PB0020) and also in heterozygous form, in combination with R275X, in a small outbred family (PB087, Table 4. 1). In this second family the Y24X mutation appeared to segregate with genital malformations in some of the otherwise unaffected relatives of the proband, indicating possible partial disease manifestation in heterozygous Y24X carriers in this family (Figure 4. 2). A second compound heterozygous mutation involving R275X was observed in an Irish ped. a maternally inherited heterozygous c.522T>A missense mutation (D174E). In this pedigree, the D174E mutation was discovered by MCHA (See *6.3.2 Identification of alterations in new cases of BBS and in NPHP patients*) due to an altered peak profile in exon 4 of *BBS2* in the patient sample Sequence analysis of the exon revealed the cause of the

MCHA result to be a T>A substitution which, on sequencing of the parent samples, was found to have been inherited from the mother.

See Appendix 2 for the position of the *BBS2* mutations in the gene.

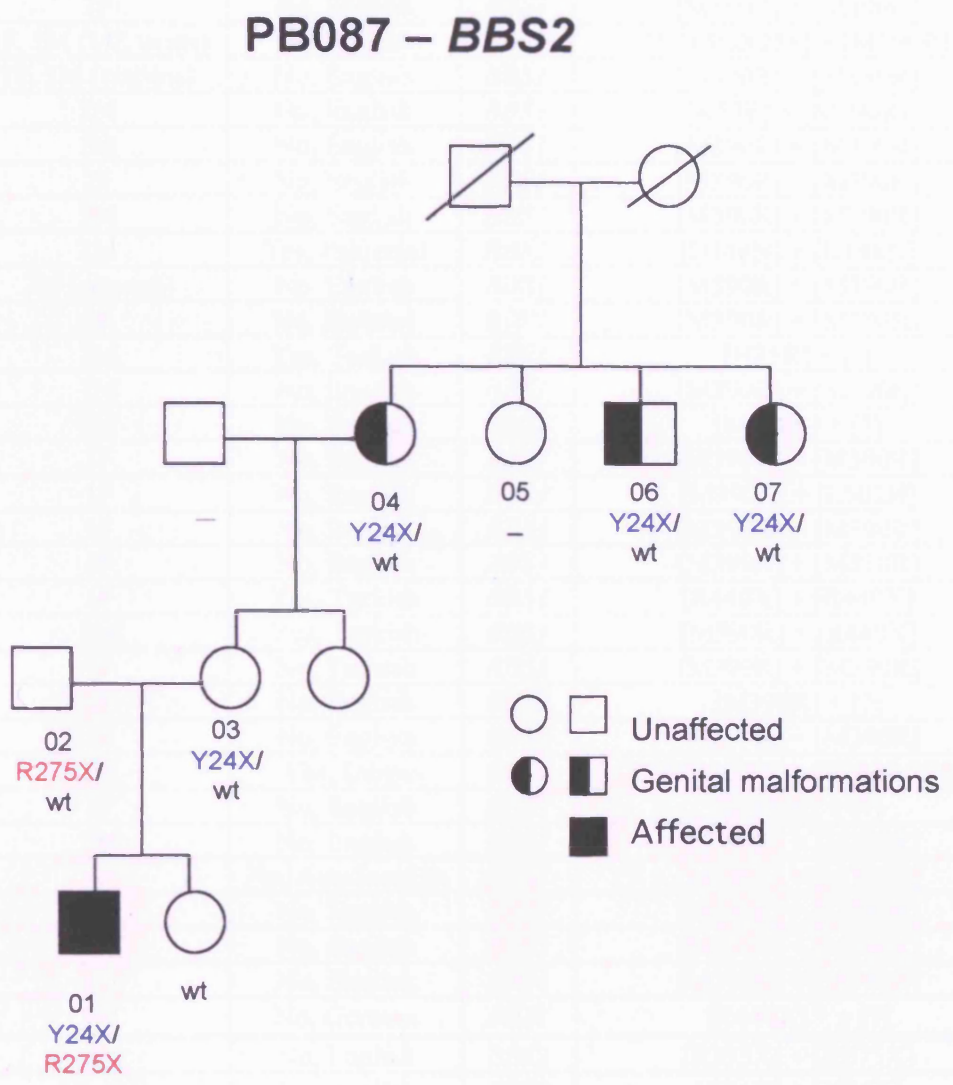


Figure 4. 2: Pedigree PB087. An English pedigree in which three relatives (04, 06 and 07) of the patient (01) have genital malformations in association with a heterozygous Y24X mutation.

Pedigree	Number of affecteds	Consanguineous, origin	Gene	Mutations
PB003	1F, 1M	No, English	*BBS1*	[M390R] + [M390R]
PB006	2F	No, English	*BBS1*	[M390R] + [M390R]
PB008	1F, 2M (MZ twins)	No, English	*BBS1*	[Y284fsX288] + [M390R]
PB013	1F, 1M (cousins)	No, English	*BBS1*	[M390R] + [M390R]
PB017	2M	No, English	*BBS1*	[K53E] + [M390R]
PB027	2M	No, English	*BBS1*	[M390R] + [M390R]
PB028	2F	No, English	*BBS1*	[M390R] + [M390R]
PB029	2M	No, English	*BBS1*	[M390R] + [M390R]
PB034	2M	Yes, Pakistani	*BBS1*	[D148N] + [D148N]
PB037	2M (cousins)	No, English	*BBS1*	[M390R] + [M390R]
PB038	1F	No, Holland	*BBS1*	[M390R] + [M390R]
PB053	1M	Yes, Turkish	*BBS1*	[H35R] + [?]
PB073	1M	No, English	*BBS1*	[M390R] + [M390R]
PB074	1M	No, English	*BBS1*	[M390R] + [?]
PB076	1F	No, English	*BBS1*	[M390R] + [M390R]
PB077	1F	No, English	*BBS1*	[M390R] + [L503H]
PB079	1F	No, English	*BBS1*	[M390R] + [M390R]
PB080	1F	No, English	*BBS1*	[M390R] + [M390R]
PB086a	1F	Yes, Turkish	*BBS1*	[R440X] + [R440X]
PB086b	3M	Yes, Turkish	*BBS1*	[M390R] + [R440X]
PB096	1F	No, Turkish	*BBS1*	[M390R] + [M390R]
PB097	1F	No, English	*BBS1*	[M390R] + [?]
PB100	1F	No, English	*BBS1*	[M390R] + [M390R]
PB110	1M	Yes, Libya	*BBS1*	[Y284fsX288] + [Y284fsX288]
PB125	1F	No, English	*BBS1*	[M390R] + [?]
PB139	1M	No, English	*BBS1*	[M390R] + [M390R]
PB150	1F	No, Aus/Scottish	*BBS1*	[M390R] + [M390R]
PB154	1M	No, English	*BBS1*	[M390R] + [M390R]
PB174	1F	No, English	*BBS1*	[M390R] + [M390R]
PB180	1M	No, English	*BBS1*	[M390R] + [M390R]
F820	1M	No, German	*BBS1*	[R440X] * + [?]
PB005	2F	No, English	*BBS2*	[R275X] + [R275X]
PB020	1F, 1M	No, English	*BBS2*	[Y24X] + [Y24X]
PB026	1F	Yes, Indian	*BBS2*	[D170fsX171] + [D170fsX171]
PB065	1F	No, Irish	*BBS2*	[R275X] + [R275X]
PB069	1M	No, Irish	*BBS2*	[D174E] * + [R275X]
PB087	1M	No, English	*BBS2*	[Y24X] + [R275X]
F523	1F	No, Turkish	*BBS6*	[I339V] * + [?]
PB054	1F	Yes, Turkish	*BBS7*	[K237fsX296] + [K237fsX296]
PB012	3M	Yes, Pakistani	*BBS8*	[IVS10delTGC] + [IVS10delTGC]

Table 4. 1: Recessive mutations identified in BBS patients. F - females, M - males, * - mutations that were identified using MCHA, all others were identified by direct sequencing. [?] indicates pedigrees in which a single mutant allele was identified.

4.3.3 *BBS4* mutations

No mutations in *BBS4* were found to segregate with disease in any individuals from the patient cohort. Sequence alterations that did not conform to a classic recessive model of inheritance were however discovered; these alterations will be discussed in *Chapter 5 Complex inheritance in BBS and a related syndrome.*

4.3.4 *BBS7* mutations

Only a single *BBS7* mutation was detected in our patient cohort. A c.711-714delGAGA deletion in exon 7 (K237fsX296) of the gene was identified in a consanguineous Turkish pedigree (PB054, Table 4. 1 and Appendix 5). This mutation, which results in a frameshift identical to the one reported by Badano *et al.* (2003a), is located within the 112-amino acid region of the BBS7 protein where all known *BBS7* mutations have been reported and that is shared exclusively between BBS7 and BBS2.

See Appendix 5 for a schematic showing the position of the frameshift mutation in the *BBS7* gene.

4.3.5 *TTC8* mutations

Sequencing of the ORF of *TTC8* revealed a 3bp deletion resulting in the abolition of the splice donor site of exon 10 (IVS10+2-4delTGC) in a consanguineous Pakistani pedigree (PB012, Table 4. 1 and Appendix 6), identifying *TTC8* as an eighth locus for BBS. The three affected members of the pedigree were all homozygous for the mutation; single copies of the mutation were present in the parents (obligate heterozygotes) and also in an unaffected sister. A possible consequence of this mutation is the skipping of exon 10, resulting in an in-frame deletion of 111bp. However, RT-PCR analysis of cDNA from cultured renal tubular cells from the patients using primers located in exons 9 and 13 of *BBS8*, suggest that the mutant

BBS8 transcript is degraded by nonsense mediated decay as no product was detectable in the patient samples, but was present in both their unaffected sister and an unrelated positive control (Ansley et al., 2003). Identification of mutations in *TTC8* that segregate in a recessive manner with BBS has resulted in this gene being referred to as *BBS8*, although it is still denoted *TTC8* by the nomenclature committee (Ansley et al., 2003).

See Appendix 6 for a schematic of *BBS8* including the position of the IVS10+2-4delTGC mutation.

4.4 Discussion

Mutation screening of the known *BBS* genes *BBS1*, *2*, *4* and *7*, and the novel gene *BBS8*, revealed that recessive mutations in these genes account for a relatively small portion (31.7%) of the cases of BBS within our patient cohort (38 of 120 patients screened had recessive mutations in *BBS1*, *2*, *7* or *8*).

4.4.1 Mutations identified in *BBS1*

Mutational analysis of the *BBS1* gene in our cohort supported previous studies confirming that it is the gene most frequently associated with BBS, and that a common missense mutation (M390R) accounts for the majority of cases of *BBS1* (Mykytyn et al., 2001; Mykytyn et al., 2003). Thirty patients (25.0% of pedigrees) had at least one mutation in *BBS1* including the previously reported M390R, Y284fsX288 and R440X mutations, and the four novel missense mutations H35R, K53E, D418N and L503H. The expected contribution of this locus to BBS was calculated by earlier linkage studies and haplotype analysis to be 36-56% (Bruford et al., 1997; Katsanis et al., 1999). The lower frequency of mutations seen amongst our patients, which was supported by a parallel screen conducted by our collaborators (at least one *BBS1* mutation was identified in 37 of 147 (25.2%) pedigrees (Beales et al., 2003)) may be explained by several different possibilities. A subset of mutations that would have been missed through our mutation screening approach may exist within our patient cohort. Mutations that would be missed by direct sequencing of the ORF and splice junctions include large deletions or insertions, heterozygous exonic deletions and also regulatory element or cryptic splice site mutations. This is a possibility, and may explain the absence of a second mutant *BBS1* allele in some pedigrees from our cohort (a single *BBS1* mutant allele was found in 16.7% (5/30) of families with *BBS1* mutations), but it is unlikely that undetected mutations in our cohort are able to account for such a large difference between the observed and expected contributions. A second possibility is that mutations exist within either additional exons of the *BBS1* gene, that were not detected on initial cloning of the gene, or in a second *BBS* locus located in close proximity to the *BBS1* gene on 11q13. The position of the *BBS1* gene outside both the Young *et al.* (1999b) and the

wider Katsanis *et al.* (1999) *BBS1* critical intervals raises the question as to whether there may be a second *BBS* locus proximal to *BBS1*. One pedigree from our cohort (PB010) and two from our collaborator's (AR37 and AR603) that showed linkage to the critical interval on 11q13 published by Katsanis *et al.* (1999), do not have coding mutations in *BBS1* (Beales et al., 2003). The occurrence of pedigrees that do not segregate *BBS1* mutations and are linked to the region proximal to the *BBS1* gene suggest that a second *BBS* gene may exist in this region on chromosome 11, a possibility that Mykytyn *et al.* (2001) do not exclude. Finally, it is also possible that initial calculations of the contribution of *BBS1* were an over estimation. Mykytyn *et al.* (2003), in their larger study of 129 pedigrees, report *BBS1* mutations in 32.0% (41/129) of their cohort, indicating that between a quarter and a third of all BBS cases can be accounted for by mutations in *BBS1*.

4.4.1.1 Contribution of M390R to *BBS1*

The contribution of M390R to *BBS1* in our cohort was similar to that of Mykytyn *et al.* (2003); of pedigrees from our cohort containing at least one M390R allele, 73.1% (19/26) were homozygotes in comparison to 69.2% (27/39) of the cohort studied by Mykytyn *et al.* A large difference was found however in the number of pedigrees in the two study groups containing a single M390R allele with no evidence for a second *BBS1* mutant allele segregating with disease. In two of the 12 (16.6%) M390R heterozygotes studied by Mykytyn *et al.* (2003), no second *BBS1* mutation was identified. The frequency of only a single M390R allele observed in a pedigree was higher in our cohort; of seven M390R heterozygotes, three (42.9%) did not have a second *BBS1* mutation. In addition, Mykytyn *et al.* (2003) did not report any incidences of unaffected relatives that were M390R homozygotes, an occurrence that was observed twice within our cohort (in both cases in a pedigree containing two affected offspring which were homozygous for M390R, the unaffected father was also found to be an M390R homozygote). Cases of patients with single mutant alleles and unaffected family members that carry two mutant alleles in a *BBS* gene may be a reflection of the complex inheritance which is seen in some BBS pedigrees (See *Chapter 5 Complex inheritance in BBS and a related syndrome* for a further discussion of complex inheritance in BBS).

4.4.1.2 Analysis of novel mutations

In addition to the detection of known *BBS1* mutations amongst pedigrees, four novel missense mutations (H35R, K53E, D148N and L503H) were also identified in our cohort. In order to determine the potential pathogenicity of the novel alterations, they were assessed using the following criteria:

- Segregation of the alteration with disease in the pedigree.
- Absence of the alteration in at least 200 matched control chromosomes.
- Whether or not the substitution is at a conserved residue.
- Analysis using the SIFT (Sorting Intolerant From Tolerant) program (http://blocks.fhcrc.org/sift/SIFT.html) to predict the effect of the substitution on protein function.

The novel missense mutations identified in *BBS1* all segregated with disease in the pedigrees in which they were found, were not seen in control samples and were also all at residues that were conserved in a number of species including mouse, rat, dog, chimp and mosquito. The SIFT program (Ng and Henikoff, 2002), which uses sequence homology between related proteins to predict whether an amino acid substitution will affect protein function, was used to determine whether the substitutions found in BBS1 were likely to be deleterious. Using the program, the human BBS1 sequence was aligned with sequences from *C. elegans* (Accession number - AL132876), *A. gambiae* (AAAB01008807), *D. melanogaster* (AE003560) and *P. pygmaeus* (CR858140). The leucine at amino acid position 503 was found to be highly conserved among the homologous proteins; it is therefore predicted that the L503H substitution (seen in a single pedigree from the cohort, in compound heterozygous form with M390R) has a deleterious effect on the BBS1 protein. Despite in some cases resulting in a change of group, for example, the substitution of lysine for glutamic acid (K53E) results in the exchange of a basic amino acid for an acidic one, the remaining substitutions were predicted to be tolerated at the positions at which they were found within the protein. As little is known about the function of the BBS1 protein, it is possible that these substitutions, although not deleterious, may have a more subtle effect on the interaction between BBS1 and its interactors, or may lie in an as yet undetermined functional domain. In addition to the

absence of all novel missense mutations in a group of control individuals, further support for the pathogenicity of the D148N mutation was achieved by the observation of a second family segregating the mutation in the cohort of our collaborator (Beales et al., 2003).

4.4.1.3 Frequency and distribution of mutations in the *BBS1* ORF

To date 30 different mutations (nine missense, eight nonsense, six frameshift, five splice site and two in-frame deletions) have been reported in *BBS1* (Beales et al., 2003; Fauser et al., 2003; Mykytyn et al., 2003; Mykytyn et al., 2002). Despite this number and range of mutations, with the exception of M390R, each mutation has been reported in a very small number of families; three Y284fsX288 alleles, one homozygous and one heterozygous pedigree, were detected within our cohort and also within the cohort of Mykytyn *et al.* (2003; 2002), and only two homozygous D148N pedigrees were identified, one in each of our cohort and that of our collaborator (Beales et al., 2003).

The mutations found in our cohort are distributed relatively evenly throughout the ORF of the *BBS1* gene, although no mutations in exons 6 to 9 were identified in our patients (Appendix 1). Only two mutations, an in-frame deletion (I200_T201del) and a missense mutation (c.699G>A, E234K) both in exon 8, have been identified in this region of the gene (Beales et al., 2003; Mykytyn et al., 2003), suggesting that this region may be part of an important functional domain and therefore less tolerant of mutations.

4.4.2 Mutations identified in *BBS2*

Although the second most common locus in BBS, the contribution of *BBS2* is considerably lower than that of *BBS1* (*BBS2* is estimated to account for only 8-16% of all cases of BBS (Katsanis, 2004; Katsanis et al., 2001b)) and therefore fewer mutations have been reported in *BBS2*. A slightly lower than expected frequency of *BBS2* mutations were found in our cohort; recessive mutations were identified in 5.0% of our pedigrees (6/120). Of the four different mutations present in the cohort, only one mutation (R275X) had been previously reported

(Nishimura et al., 2001) and was the most common *BBS2* mutation present in our patients. Affected individuals in two cases were homozygous for the nonsense mutation and in two other pedigrees, patients were found to be compound heterozygotes for R275X and a second *BBS2* mutation (Y24X or D174E). Identification of a homozygous R275X mutation in one of these pedigrees (PB005) confirmed haplotype analysis conducted in a previous study indicating linkage of this pedigree to the *BBS2* locus (Beales et al., 2001).

4.4.2.1 Novel *BBS2* mutations

A novel deletion, nonsense and missense mutation were all identified in our cohort. The known nonsense mutation R275X was present in two pedigrees in association with novel mutations. In one family, a maternally inherited D174E missense mutation was present with a paternally inherited R275X mutation. The D174E substitution was not predicted to be deleterious using the SIFT program but, was not present in at least 200 control chromosomes (screened by collaborators) and segregated with disease in the family, suggesting that it is likely a pathogenic alteration. In the second pedigree, an R275X allele was present in combination with an additional nonsense mutation, Y24X. In this pedigree (PB087, Figure 4. 2) there was a single child that was suspected of having BBS based on the presence of obesity, polydactyly, speech delay and hypospadias (as the child was only 18 months of age on initial presentation, key diagnostic features such as RP had yet developed). In an attempt to confirm the diagnosis of BBS, DNA from the child was screened for mutations in *BBS1*, *2*, *4* and *6* by direct sequencing. The identification of two nonsense mutations in *BBS2* (the paternally inherited R275X, and the maternally inherited Y24X) confirmed the diagnosis and provided the ability to offer a prenatal diagnosis on a second pregnancy in the family. Mutational analysis of a fetal sample taken by chorionic villus sampling (CVS), determined that the fetus did not carry either of the mutations in *BBS2* and a healthy child was born at term. In this family there was also evidence for the co-segregation of genital malformations with the Y24X mutation in unaffected relatives. The maternal grandmother and two of her three siblings, all of whom were heterozygous for the Y24X mutation, were found to have genital malformations. The Y24X mutation was also present in homozygous form in one other pedigree (PB020) of our cohort in which the affected individual also has genital

malformations. The observation of an association between the Y24X mutation and genital malformations raises questions as to whether heterozygous relatives of BBS patients have an increased risk of genital malformations relative to the general population, similar to the increased incidence of renal malformations and CC-RCC seen in some unaffected carrier relatives ((Beales et al., 2000) See *1.7 Heterozygous effects*). It is possible that the link between genital malformation and a single *BBS2* mutant allele in this pedigree is a general one between mutations in *BBS2* and genital malformations, or may be a more specific relationship restricted to the Y24X mutation itself. Functional studies are required to determine the effect that this nonsense mutation has on the BBS2 protein product; one possible outcome is translational reinitiation after the Y24X termination codon, leading to an N-terminally truncated protein (Kozak, 2001). Further studies of relatives of patients with *BBS2* mutations, in particular Y24X, for the presence of genital malformations and mutational analyses of isolated cases of genital malformation in the general population would be required to determine if this is a common occurrence or is particular to this family. It was previously assumed that, due to the genital malformations seen in MKKS patients, BBS patients exhibiting genital abnormalities were most likely to have mutations in *BBS6* (*MKKS*), these results suggest that this is not the case and that the presence or absence of genital malformations in a BBS patient is not locus-specific. *BBS* mutations may also confer a susceptibility to genital malformations on the general population.

4.4.3 Recessive mutations in *BBS4*

Mutations in *BBS4* account for fewer than 3% of BBS cases (Katsanis, 2004) and only five different recessive mutations (a deletion of exons 3 and 4, one missense and three splice site mutations have been reported to date (Katsanis et al., 2002; Mykytyn et al., 2001)). Despite the expectation of further *Alu*-associated mutations being identified within *BBS4*, due to the high number of *Alu* repeats present within the *BBS4* genomic sequence, the deletion of exons 3 and 4 (IVS2_IVS5) remains the only *Alu*-associated mutation reported in the gene. No recessive *BBS4* mutations were identified in our cohort.

4.4.4 A novel deletion in *BBS7*

A single *BBS7* frameshift mutation (K237fsX296) was found in a Turkish pedigree from our cohort. Although the deletion in our patient results in an identical frameshift to that identified by Badano *et al.* (2003a) in a Saudi pedigree, the deleted bases are not identical; in the Saudi pedigree, the 4bp deletion involves bases 709-712 (c.709-712delAAGA), whereas the Turkish deletion encompasses bases 711-714 (c.711-714delGAGA). It is possible therefore that this region of the *BBS7* gene is unstable and prone to deletions.

4.4.5 Identification of a novel *BBS* gene, *BBS8*

The identification of a splice site mutation in the *TTC8* gene (now known as *BBS8*) in a consanguineous BBS pedigree from our cohort proves that this novel transcript is a *BBS* gene. The detection of a 6bp deletion (E187_Y188del) in two unrelated Saudi pedigrees from the cohort of our collaborator provides independent confirmation that this is an eighth *BBS* locus (Ansley et al., 2003). The structure and function of the BBS8 protein is the first of the known BBS proteins to provide clues as to a possible disease mechanism underlying BBS. The protein, which shares homology with BBS4, contains a several TPR motifs lying towards the C-terminus and shows significant similarity to the prokaryotic *pilF* domain which is involved in twitching mobility and type IV pilus assembly (Appendix 6). This feature of the protein suggests that BBS may be caused by a defect in the function of cilia or flagella. Interestingly, one of the three affected individuals with a splice site mutation (IVS10+2-4delTGC) in *BBS8* has *situs inversus* in addition to BBS. *Situs inversus* is a defect of left-right asymmetry, known to be caused by dysfunction of the nodal cilia. The presence of *situs inversus* in one of three affected individuals indicates a randomisation of left-right asymmetry caused by this mutation. The function of BBS8 and the involvement of cilia in BBS will be discussed further in *Chapter 8 Expression analysis of BBS4 and BBS8 in mouse tissues.*

4.4.6 Summary

Mutational analysis of the known *BBS* genes and a newly cloned gene in our patient cohort has demonstrated the extensive genetic heterogeneity observed in BBS. Recessive mutations in five of the *BBS* genes (*BBS1*, *2*, *4*, *7* and *8*) account for ~30% of the 120 cases screened in this study. With the exception of *BBS1*, each of the other loci accounts for less than 10% of cases. Taking into account the rarity of the majority of the *BBS* loci, it is possible that a large number of additional loci exist amongst the remaining cases of BBS that are not yet accounted for by mutations in known genes. It is only with the cloning of the latest *BBS* gene (*BBS8*) that a potential mechanism that, when defective, results in the pleiotropic phenotype of BBS has been identified.

Chapter 5 Complex inheritance in BBS and a related syndrome

5.1 Introduction

BBS was traditionally considered to be an autosomal recessive disorder based on segregation patterns observed within pedigrees and populations. It was also initially expected that mutations at a single genetic locus would account for all cases of BBS but genetic heterogeneity was apparent from the initial linkage study to map *BBS1* and has steadily grown, resulting in the identification of eight known *BBS* loci (*BBS1-8*) with evidence for additional loci within the human genome (See *1.9 The genetics of BBS* for a summary of the mapping and cloning of *BBS1-7* and for a discussion of the cloning of the most recent *BBS* gene). Mutation analysis in a large cohort of patients has provided evidence for a possible non-Mendelian inheritance pattern in BBS involving mutations at more than one locus in some families (Beales et al., 2003).

5.1.1 Initial evidence for complex inheritance in BBS

In 2000 the *BBS6* gene was cloned independently by two different groups (Katsanis et al., 2000; Slavotinek et al., 2000). In both screens, the majority of mutations in *BBS6* were found to be null alleles including frameshift mutations (D143fsX157 and F94fsX103) and a nonsense mutation (Y264X). It should be noted however, that there was an overlap in the patients studied by both groups (three Newfoundland pedigrees were shared between the cohorts). As different mutation nomenclature was used by each group, including different reference points in the same *MKKS* sequence designated as base 1 for the numbering of mutations, the overlap in patient samples was not initially apparent. All mutations were found to reside in exon 3, the largest and first coding exon of *BBS6*. This was in contrast to the less severe mutations within the gene that had been reported in MKKS patients (Stone et

al., 2000), suggesting that milder hypomorphic alleles result in the phenotype of MKKS, whereas truncating mutations result in the more severe phenotype of BBS (Katsanis et al., 2000). Slavotinek *et al.* (2000) were able to detect 100% (8/8 alleles) of mutant *BBS6* alleles within four different pedigrees, whereas in the slightly larger screen by Katsanis *et al.* (2000), only 93% (13/14) of mutant alleles were identified in seven pedigrees. Following the identification of causative *BBS6* mutations in a small number of families, a large-scale screen of 163 pedigrees was carried out by the same group to ascertain the contribution of the *BBS6* locus to all cases of BBS from their cohort (Beales et al., 2001). Mutations were detected in eight pedigrees. In contrast to previous screens, all but one of the nine mutations identified were missense mutations, all of which were absent from 188 control chromosomes. The majority of the mutations were located in exon 3 (6/9 mutations), but three missense mutations were also present in exon 6. In seven of the eight pedigrees (87.5%) containing mutations, only a single mutant allele was identified. There were a number of possible explanations for this high number of pedigrees with single mutations:

- *The single mutations may be polymorphisms rather than pathogenic mutations.*
  As the majority of mutations identified in this cohort were missense mutations it is possible that these sequence alterations are non-pathogenic polymorphisms. This is unlikely as one of the heterozygous mutations was a nonsense mutation (Q147X) and is therefore predicted to be severe, suggesting that the single mutations are likely to be pathogenic. Also one of the missense alterations (A242S) had previously been reported in an MKKS pedigree (Stone et al., 2000) and is therefore likely to be a disease-causing mutation. None of the single mutations were found in a sample of matched control chromosomes.

- *'Missing' mutations may exist in regulatory regions, cryptic splice sites or may be deletions of part of the gene*
  Although additional mutations that would have been missed by sequencing of the *BBS6* ORF may have existed in some families, not all cases of single mutant alleles detected in the screen could be explained in this way. In the consanguineous pedigree B14 (See Figure 1. 4), the A242S allele was identified in the affected individual (03) and was also found to be carried by his unaffected sister (04). Haplotype analysis

indicated that this pedigree was excluded from linkage to the *BBS6* locus as both offspring had identical haplotypes across the *BBS6* interval, ruling out the possibility of a second, unidentified, *BBS6* mutation existing within this pedigree in only the affected individual. Analysis of markers across the *BBS2* critical interval showed a stretch of IBD in only the affected individual, suggesting linkage to *BBS2*. In addition, two other pedigrees (AR-301 and AR259) with heterozygous *BBS6* mutations (one of which was the nonsense mutation Q147X present in pedigree AR-259) were consistent with linkage to *BBS2* (Beales et al., 2001).

- *Mutations at more than one locus may be both necessary and sufficient to cause BBS in some pedigrees.*

  The identification of a number of pedigrees with a single *BBS6* mutation which, in some cases, have been shown to be excluded from *BBS6* but show possible linkage to other *BBS* loci, suggests that a complex form of inheritance (oligogenic inheritance) involving multiple mutations at more than one locus may exist in BBS. This possibility is supported by evidence from a second consanguineous pedigree that was found to be IBD for both the *BBS2* and *BBS4* loci (Beales et al., 2001). Although this may be a chance occurrence of two regions of homozygosity at two *BBS* loci (in offspring of a first cousin union, the probability of any region being homozygous is 1 in 16), it may also indicate the presence of mutations at both loci within this pedigree.

5.1.2 Evidence of possible oligogenic inheritance in NPHP

In addition to BBS, evidence for oligogenic inheritance has been found recently in another recessive condition, nephronophthisis (NPHP), which shares similarities with BBS. NPHP is the most common cause of ESRD in children (Fanconi et al., 1951) and is characterised by clinical features including an inability to concentrate the urine, cyst formation and an increased echogenicity on renal ultrasound (Hildebrandt, 1999; Hildebrandt et al., 1997b; Waldherr et al., 1982), all features reported in BBS patients (See *1.4.3 Renal abnormalities* for a review of the renal phenotype of BBS patients). Extra-renal defects are also found in association with the NPHP renal phenotype; Joubert syndrome combines NPHP with

congenital hepatic fibrosis and cerebellar vermis aplasia (Boichis et al., 1973; Hildebrandt et al., 1998), Senior-Löken syndrome (SLS) is the association of the NPHP renal phenotype with RP (Løken et al., 1961; Senior et al., 1961) and developmental bone defects and *situs inversus* have both also been reported in patients with NPHP (Hildebrandt et al., 1992; Otto et al., 2003). Four *NPHP* genes have now been identified with a specific genotype-phenotype relationship associated with each locus; *NPHP1* (juvenile form) maps to 2q12.3 (Hildebrandt et al., 1997a), *INVS* (NPHP2, infantile) maps to 9q31 (Otto et al., 2003), *NPHP3* (adolescent) maps to 3q22 (Olbrich et al., 2003) and *NPHP4* (adolescent) maps to 1p36 (Otto et al., 2002). NPHP and BBS also share similarities at the cellular level; BBS4 and the recently identified BBS8 protein have both been shown to localise to the basal body of the cilia in ciliated tissues such as the renal primary cilia (Ansley et al., 2003; Kim et al., 2004), as have some of the NPHP proteins (Otto et al., 2003) (See *Chapter 8 Expression analysis of BBS4 and BBS8 in mouse tissues* for further discussion on the function of BBS8).

In addition to clinical and cellular similarities, NPHP shares similarities at the genetic level with BBS. Both conditions show extensive genetic heterogeneity with a number of cloned genes already identified in each condition (BBS – eight known genes, NPHP – four known genes), with evidence for further loci in both cases (~57% of BBS cases do not carry mutations in the known genes (Katsanis, 2004), 65% of NPHP do not have mutations in *NPHP1-4* (Hildebrandt, F. Personal communication)). Mutational analysis of known *NPHP* genes has also resulted in the identification of pedigrees in which single mutant alleles were found. During a screen of patients for mutations in *INVS*, the gene responsible for NPHP2, homozygous mutations were identified in six pedigrees and a single heterozygous mutation was identified in a seventh family (Otto et al., 2003). Mutations in *NPHP3* were found in nine pedigrees from the same patient cohort, but both mutant alleles were found in only three of these pedigrees (Olbrich et al., 2003). Analysis of the *NPHP4* gene by another group also resulted in the discovery of single alleles; two of five pedigrees with mutations in the gene had heterozygous mutations (Mollet et al., 2002).

5.1.3 Conclusions

The existence of families in which single mutant alleles occur in BBS, and also in NPHP, indicates that the inheritance pattern in these diseases may not be exclusively autosomal recessive as was previously thought. Mutations at more than one locus may be necessary to cause disease, or the phenotype resulting from mutations at one locus could be influenced by additional mutations at a second locus. To study whether non-Mendelian inheritance does exist in BBS, and the extent to which it occurs within our cohort of patients, mutational data for *BBS1*, *2*, *4*, *6*, *7* and *8* for all patients from the cohort was analysed. In view of the similarities between BBS and NPHP, the possibility of oligogenic inheritance involving the *BBS* and *NPHP* genes was also studied.

5.2 Methods

5.2.1 Mutational data

All sequence data generated from the direct sequencing of *BBS1*, *2*, *4*, *6*, *7* and *8* in the patient cohort as described in Chapter 4 was analysed using the Sequencher v.4.1 program (See *2.2.3.5 Sequence analysis*) to identify any mutations present in the *BBS* genes.

5.2.2 Restriction digests

Some mutations associated with oligogenic inheritance in BBS pedigrees were confirmed by restriction digest. All digests were performed as described in section *2.2.1.4 Restriction digests* using the following enzymes; R275X (*BBS2*) – *Bsl*I, T558I (*BBS2*) – *Ssp*I and A364E (*BBS4*) – *Mbo*I.

5.2.3 Microsatellite analysis of chromosome 11

In one pedigree from the cohort (PB056), both the affected mother and her affected daughter were found to be homozygous for the M390R mutation (*BBS1*). As the father of the affected child was unavailable, it was not possible to determine the segregation pattern of the mutation within the pedigree. To rule out the possibility of uni-parental disomy (UPD) of chromosome 11 from the mother or hemizygosity for part, or all, of the chromosome in the child, genotyping of both mother and daughter using microsatellite markers spaced evenly across the length of chromosome 11 was carried out. The markers used were: D11S1984, D11S2362, D11S1999, D11S1981, ATAE08, PYGM, D11S2371, D11S4960, D11S1979, D11S4952, D11S2000, D11S1986, D11S1998, D11S4464, D11S4463 and D11S2359. Amplification, pooling and gel electrophoresis of markers was carried as described in section *2.2.2 Homozygosity mapping.*

5.2.4 Mutational analysis of NPHP patients

To test the hypothesis that oligogenic inheritance may exist between *BBS* and *NPHP* genes, a cohort of 95 NPHP patients were studied through a collaboration with Prof. Friedhelm Hildebrandt at the University of Michigan. The patients had previously been screened for mutations in the known *NPHP* genes by direct sequencing (at the University of Michigan). Patients with mutations in *NPHP1*, *2*, *3* or *4*, and also patients that fulfilled diagnostic criteria for NPHP or associated syndromes (SLS, Joubert etc.) but did not carry mutations in the known genes were included in the cohort. MCHA was used to screen the cohort for mutations in *BBS1*, *2*, *4* and *6* (See *2.2.4 MCHA*).

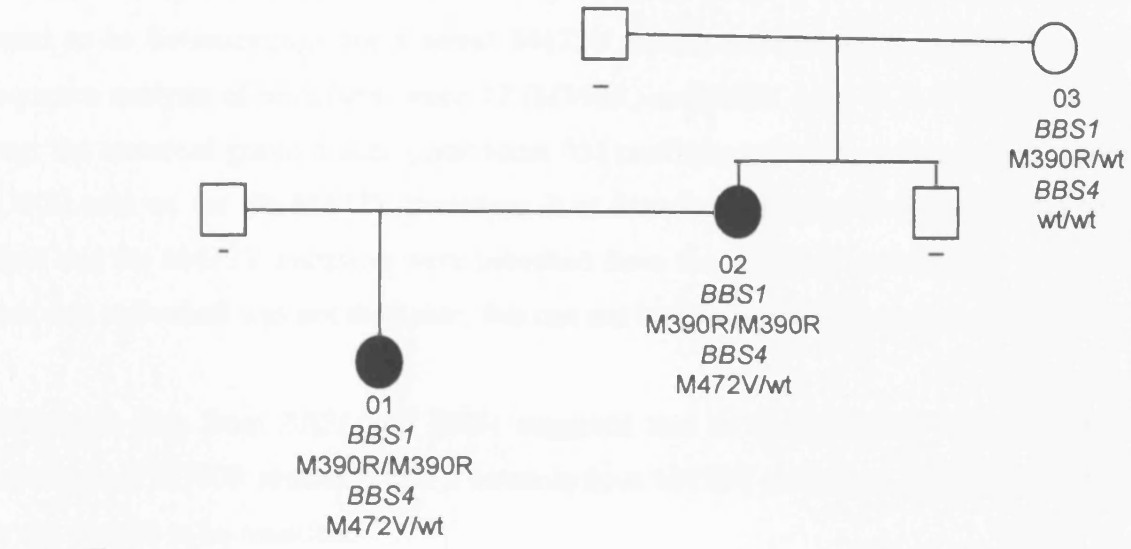
5.3 Results

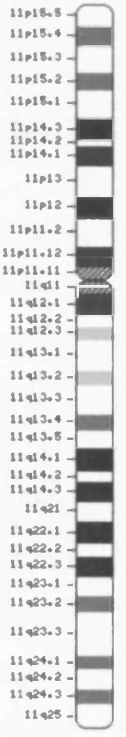
5.3.1 Evidence for oligogenic inheritance in BBS

Evidence for non-Mendelian inheritance in BBS was found in four pedigrees from our cohort.

5.3.1.1 Pedigree PB056

Pedigree PB056 (Figure 5. 1) consists of an affected mother (individual 02) and daughter (01). Sequence analysis revealed that both patients were homozygous for the M390R (*BBS1*) mutation. As the father of the affected child (01) from this pedigree was not available, it was not possible to confirm that the child had inherited a single copy of the M390R mutation from each parent. It was therefore necessary to rule out UPD of chromosome 11 from the mother by genotyping of microsatellite markers along the length of the chromosome in individuals 01 and 02. Microsatellite analysis indicated that UPD had not occurred as the affected child was found to have inherited one haplotype from her mother and another haplotype, different to that of the other maternal haplotype, which was presumed to have been inherited from her father (who must therefore carry at least one copy of the M390R mutation). Hemizygosity for all or part of the *BBS1* gene would have given an appearance identical to that of a homozygous T>G base substitution at codon 390 of *BBS1* (M390R) on sequence analysis. To rule out the possibility of hemizygosity across the exon 12 region of *BBS1*, coding SNPs (c.SNPs) within *BBS1* were also analysed; heterozygosity for both SNPs, either side of exon 12, supports the microsatellite analysis and suggests that the child has inherited two copies of the M390R mutation, one from each parent. The heterozygosity of the SNPs flanking the M390R mutation also suggest that the missense mutation has arisen independently on different haplotypes.





1984	160	168	160	180
2362	215	218	215	218
1999	107	119	107	119
1981	162	166	162	170
ATAE08	165	157	165	165
PYGM	**118**	**118**	118	124
SNP1	A	G	A	G
BBS1	**M390R**	**M390R**	**M390R**	**M390R**
SNP2	C	T	C	T
2371	212	200	212	196
4960	247	251	247	259
1979	**258**	**258**	**258**	**258**
4952	**146**	**146**	**146**	**146**
2000	**225**	**225**	225	207
1986	**176**	**176**	**176**	**176**
1998	150	154	150	154
4464	**237**	**237**	**237**	**237**
4463	128	124	128	120
2359	**216**	**216**	**216**	**216**

Figure 5. 1: Triallelic inheritance in pedigree PB056. Genotyping of microsatellites along the length of chromosome 11 excluded the possibility of UPD of a maternally inherited chromosome or hemizygosity of the chromosome in the region surrounding *BBS1* in individual 01.

This pedigree was screened for the presence of mutations in *BBS4* by MCHA (See *6.3.2 Identification of alterations in new cases of BBS and NPHP patients*). Both patients were found to be heterozygous for a novel M472V missense mutation in exon 15 of *BBS4*. Sequence analysis of both *BBS1* exon 12 (M390R) and *BBS4* exon 15 (M472V) in a sample from the maternal grandmother (individual 03) confirmed that she was a heterozygote for M390R and wt for the M472V mutation. It is therefore assumed that the second M390R allele and the M472V mutation were inherited from the maternal grandfather. As a sample from this individual was not available, this can not be confirmed by sequence analysis.

Mutational data from *BBS1* and *BBS4* suggests that in this pedigree three mutations (a homozygous M390R mutation and a heterozygous M472V mutation) appear to be required for the disease to be manifest.

5.3.1.2 Pedigree PB043

Pedigree PB043 (Figure 5. 2) is a consanguineous Turkish pedigree in which one of the five offspring is affected by BBS. In a previous study the affected individual (01) in this pedigree was shown to exhibit IBD across both the *BBS2* and *BBS4* critical intervals (Beales et al., 2001). Sequence analysis of *BBS2* revealed a homozygous c.1673C>T missense mutation (T558I) in the affected individual. This substitution creates an *Ssp*I site within the *BBS2* exon 14 fragment. Analysis of the others members of the pedigree by restriction digest indicated that the father (06) and one of the sibs (02) were heterozygous for the substitution. Despite being unaffected by disease, the mother (07) and the three remaining sibs (03, 04 and 05) were homozygous for the mutation. Sequence analysis of *BBS4* in the affected individual of the pedigree revealed the presence of a second homozygous *BBS* missense mutation, c.1091C>A (A364E). Again this substitution affected a restriction site; in this case the substitution creates a *Mbo*I site within the *BBS4* exon 13 fragment. By digest, both parents (06 and 07) and two of the sibs (02 and 04) were found to be heterozygous. The remaining two sibs (03 and 05) had not inherited the mutation from either parent and were wt for the change.

In this pedigree, owing to the presence of two homozygous missense mutations in the patient and three mutant alleles (homozygous T558I (*BBS2*) and heterozygous A364E (*BBS4*) mutations) in two asymptomatic individuals (04 and 07), four mutant *BBS* alleles would appear necessary for disease manifestation.

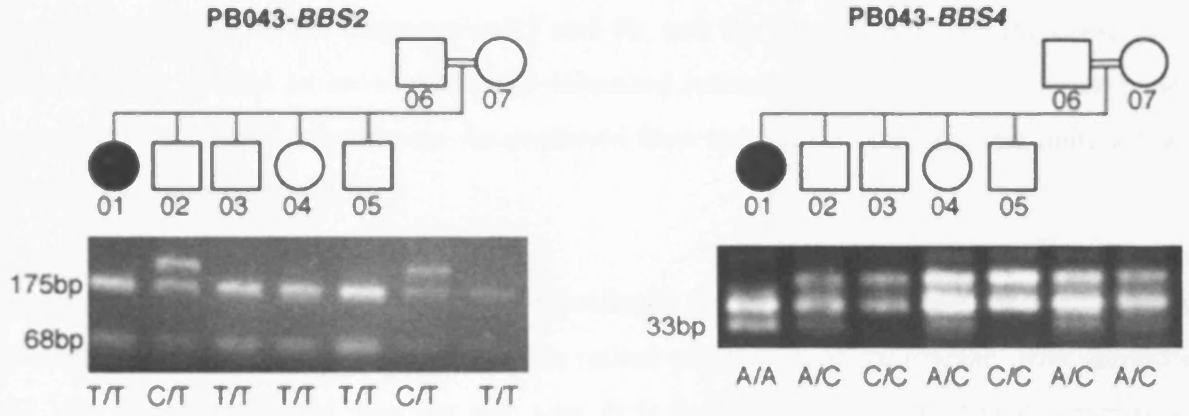


Figure 5. 2: Tetrallelic inheritance in pedigree PB043. The affected individual (01) has inherited two homozygous missense mutations; c.1673C>T (T558I) in *BBS2* and c.1091C>A (A364E) in *BBS4*. Segregation of both mutations in the pedigree was carried out by restriction digest. In the presence of the T558I mutation, 68 and 175bp fragments are generated following digestion with *Ssp*I. The A364E allele generates a 33bp fragment on digestion with *Mbo*I. Both a sib (04) and the mother (07) have inherited three mutant alleles but are asymptomatic (Katsanis et al., 2002).

5.3.1.1 Pedigree PB009

Pedigree PB009 (Figure 5. 3) contains three affected individuals. From the *BBS1* screen all three affected individuals (01, 02 and 03) were found to be homozygous for M390R (*BBS1*) and both parents (04 and 05), heterozygous. The three affected sibs in the pedigree were screened for mutations in *BBS2* by MCHA (See *6.3.2 Identification of alterations in new cases of BBS and NPHP patients*). On analysis of the exon 9 fragment in individuals 01 and 02, an altered peak profile was seen (See Figure 6. 5). Individual 03 did not have the same altered profile and, by comparison with the control fragment, was considered to be wt. Sequence analysis of the fragment in 01 and 02, and the parents, revealed the cause of the altered peak profile to be a paternally-inherited heterozygous c.1045T>G transversion, resulting in a L349W substitution. As predicted from the MCHA analysis, this mutation was not present in individual 03.

In this pedigree, the three affected individuals show a large degree of variability in phenotype, in particular with respect to the retinal component of the disease, with individual 03 less severely affected than her two sibs. It is therefore possible that in this family the presence of the third mutation, in *BBS2* (L349W), modifies the phenotype resulting from the homozygous M390R (*BBS1*) mutation.

	PB009-01	PB009-02	PB009-03
Sex	Female	Male	Female
Age at onset of nightblindness	17 years	13 years	32 years
Age at onset of RP	20 years	15 years	34 years
Obesity (BMI)	33	33	29
Behaviour	Disinhibited	Disinhibited	Normal

Table 5. 1 Phenotypic characteristics of affected individuals in pedigree PB009 (Badano et al., 2003b). Progression of retinal degeneration in individual 03 is significantly slower than in her two sibs (01 and 02).

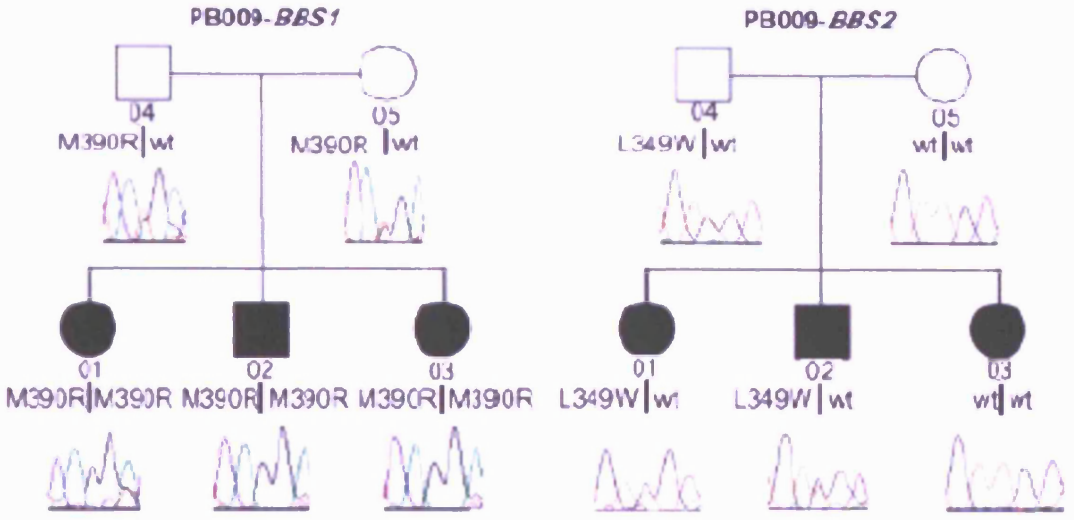


Figure 5. 3: Pedigree PB009. All affected offspring have inherited two copies of the M390R mutation in *BBS1*. A third mutant allele, a heterozygous L349W mutation in *BBS2*, has also been inherited by individuals 01 and 02 from their father (04). This additional allele was not present in individual 03, the least severely affected of the offspring (Badano et al., 2003b).

5.3.1.2 Pedigree PB061

Pedigree PB061 (Figure 5. 4) contains three offspring, two of which are affected by BBS. Sequencing of the ORF of *BBS2* identified a homozygous nonsense mutation (R275X) in both affected individuals (01 and 02). Confirmation of the mutation was carried out by digestion of the *BBS2* exon 8 fragment with *BslI*; the mutation was found to be present in both parents (04 and 05) and also the unaffected child (03) in heterozygous form.

Sequence analysis from individual 02 in the *BBS1* mutation screen identified an additional mutation in this individual. A heterozygous T>C mutation at the +2 position of the splice donor site of exon 15 (IVS15+2T>C) was present in this patient. Analysis of the other members of the pedigree revealed that the mutation had been inherited by both 02 and the unaffected sib (03) from the mother (05), who was heterozygous for the change. The mutation was not present in the second affected offspring (01), or the father (04).

Similar to pedigree PB009, the phenotype in this family is variable; of the two BBS sibs, 02 is the more severely affected. It would appear from this pedigree that the combination of the splice site mutation in *BBS1* with R275X (*BBS2*) results in a more severe phenotype than R275X alone.

	PB061-01	PB061-02
Sex	Female	Male
Development: sit	9 months	11 months
Development: walk	15 months	21 months
Development: speak	12 months	25 months
Learning	Normal	Severe delay
Obesity (BMI)	24 – no diet	25 – strict diet

Table 5. 2: Phenotypic differences between PB061-01 and PB061-02 (Badano et al., 2003b). In comparison to individual 01, individual 02 shows developmental delay and requires a strict diet to maintain a BMI of 25.

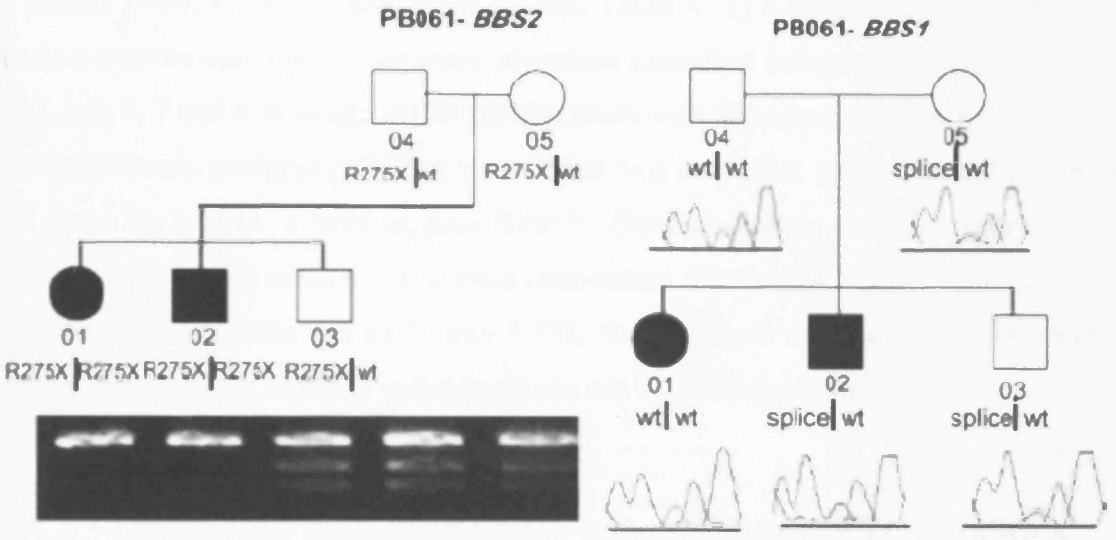


Figure 5. 4: A modifier mutation in PB061. The affected sibs (01 and 02) in the pedigree are both homozygous for the R275X (*BBS2*) nonsense mutation. The segregation of the mutation has been confirmed by restriction digest; the presence of the mutation in 01 and 02 abolishes a *Bsl*I site. A maternally inherited splice site mutation in *BBS1* (IVS15+2T>C) is also carried by individuals 02 and 03 (Badano et al., 2003b).

5.3.2 Examples of possible oligogenic inheritance in BBS

5.3.2.1 BBS patients with single mutations

In addition to pedigrees in which multiple *BBS* mutant alleles were present, patients with single heterozygous mutations were also identified in the cohort. In three cases of BBS from our cohort (PB074, PB097 and PB0125, See Table 4. 1) a heterozygous M390R (*BBS1*) mutation was the only coding sequence alteration identified following mutation screening of *BBS1*, *2*, *4*, *6*, *7* and *8*. A single H35R (*BBS1*) allele was also identified in a BBS patient from a consanguineous pedigree (PB053). In a further two cases that were screened for mutations in all genes by MCHA, a heterozygous R440X (*BBS1*) nonsense mutation (individual F820) and an I339V (*BBS6*) missense mutation (individual F523) were found. (Due to low DNA concentrations available for F820 and F523, the presence of a second *BBS1* and *BBS6* mutation, respectively, in these individuals can not be ruled out at this stage).

5.3.1.3 Unaffected individuals with homozygous *BBS1* mutations

Unaffected individuals with homozygous *BBS1* mutations were also identified in the cohort. In two cases (PB006 and PB029), the unaffected father of patients who were M390R (*BBS1*) homozygotes was also found to be homozygous for this common mutation.

5.3.2 Evidence for oligogenic inheritance between BBS and NPHP

Three cases of oligogenic inheritance involving mutations in both *NPHP* and *BBS* genes in patients affected with NPHP and RP (SLS) were identified from a screen of 95 NPHP patients.

5.3.2.1 Individual F712 II-1

Individual F712 II-1 is affected by SLS, developing ESRD at the age of 15 years. Mutation analysis of the NPHP genes in this individual revealed a homozygous gene deletion of *NPHP1* (the most common mutation reported in *NPHP1*, occurring in ~85% of patients (Otto et al., 2000). Screening of *BBS1*, *2*, *4* and *6* in this individual revealed the presence of a novel heterozygous c.147A>G missense mutation (K46R) in *BBS4*.

5.3.2.2 Individual F408 II-1

A second SLS patient was also found to be heterozygous for a novel missense mutation in *BBS4*, in this case a c.1322A>C substitution (K441T). This patient is also heterozygous for a missense mutation (L343F) within a novel gene predicted to cause NPHP, *NEK8*, and therefore may represent an example of digenic inheritance.

5.3.2.3 Individual F194 II-1

A novel c.67A>G heterozygous mutation (R23G) in *BBS6* was also identified in an SLS patient from the NPHP cohort. Mutation analysis of the *NPHP* genes in this patient identified a heterozygous nonsense mutation, R585X, in *NPHP1*.

Individual	Allele1 (gene)	Allele 2 (gene)	Allele 3 (gene)	Allele 4	Phenotype
PB056-01	M390R (*BBS1*)	M390R (*BBS1*)	M472V (*BBS4*)		BBS
PB056-02	M390R (*BBS1*)	M390R (*BBS1*)	M472V (*BBS4*)		BBS
PB056-03	M390R (*BBS1*)				Unaffected
PB043-01	T558I (*BBS2*)	T558I (*BBS2*)	A364E (*BBS4*)	A364E (BBS4)	BBS
PB043-02	T558I (*BBS2*)	A364E (*BBS4*)			Unaffected
PB043-03	T558I (*BBS2*)	T558I (*BBS2*)			Unaffected
PB043-04	T558I (*BBS2*)	T558I (*BBS2*)	A364E (*BBS4*)		Unaffected
PB043-05	T558I (*BBS2*)	T558I (*BBS2*)			Unaffected
PB043-06	T558I (*BBS2*)	A364E (*BBS4*)			Unaffected
PB043-07	T558I (*BBS2*)	T558I (*BBS2*)	A364E (*BBS4*)		Unaffected
PB009-01	M390R (*BBS1*)	L349W (*BBS2*)			Unaffected
PB009-02	M390R (*BBS1*)				Unaffected
PB009-03	M390R (*BBS1*)	M390R (*BBS1*)	L349W (*BBS2*)		Severe BBS
PB009-04	M390R (*BBS1*)	M390R (*BBS1*)	L349W (*BBS2*)		Severe BBS
PB009-05	M390R (*BBS1*)	M390R (*BBS1*)			Mild BBS
PB061-01	R275X (*BBS2*)				Unaffected
PB061-02	R275X (*BBS2*)	IVS15+2T>C (*BBS1*)			Unaffected
PB061-03	R275X (*BBS2*)	R275X (*BBS2*)	IVS15+2T>C (*BBS1*)		Severe BBS
PB061-04	R275X (*BBS2*)	R275X (*BBS2*)			Mild BBS
PB061-05	R275X (*BBS2*)	IVS15+2T>C (*BBS1*)			Unaffected
PB006-04	M390R (*BBS1*)	M390R (*BBS1*)			Unaffected
PB029-04	M390R (*BBS1*)	M390R (*BBS1*)			Unaffected
F712 II-1	Del (*NPHP1*)	Del (*NPHP1*)	K46R (*BBS4*)		SLS
F408 II-1	L343F (*NEK8*)	K441T (*BBS4*)			SLS
F194 II-1	R585X (*NPHP1*)	R23G (*BBS6*)			SLS

Table 5. 3: Mutations involved in complex inheritance in BBS and NPHP patients. Cases of potential triallelic and tetrallelic inheritance were identified in the patient cohort in addition to two cases where inheritance of a third mutant BBS allele modifies the phenotype. Missense mutations in *BBS4* and *BBS6* were also identified in SLS patients.

5.4 Discussion

Mutation screening of our cohort of BBS patients for mutations in all known *BBS* genes has lead to the identification of several cases of oligogenic inheritance involving mutations at more than one locus and/or a segregation pattern of mutations that does not conform to standard Mendelian laws of recessive inheritance. A study of a cohort of NPHP patients also identified incidences of oligogenic inheritance involving mutations in both *BBS* and *NPHP* genes resulting in an NPHP phenotype. See Table 5. 3 for a summary of mutations associated with complex inheritance identified in our cohort and Appendices 1 – 4 for the position of the mutations within the *BBS* genes.

5.4.1 Triallelic inheritance in BBS

Following the identification of pedigrees with single *BBS6* mutant alleles and evidence in some of the pedigrees for linkage to other BBS loci, questions of possible complex inheritance involving mutations at more than one *BBS* locus were raised (Beales et al., 2001). The cloning of *BBS2* in 2001 allowed this hypothesis to be tested by screening all pedigrees from the cohort for mutations in *BBS2*, regardless of any haplotype inferred chromosomal assignment or mutational data for the pedigree. Like the *BBS6* mutation screen, in addition to recessive *BBS2* mutations in six pedigrees (including three pedigrees, PB005, PB020 and PB026 from our own cohort, see *4.3.2 BBS2 mutations* for a discussion of the recessive mutations in these pedigrees), single *BBS2* mutant alleles were also detected in the cohort, in eight pedigrees (Katsanis et al., 2001a). To investigate the possibility of a second unidentified mutation (such as a large deletion or regulatory element mutation) in these pedigrees, linkage analysis was carried out using microsatellite markers around the locus to establish whether the pedigrees were linked to *BBS2*. Six of the eight pedigrees were large enough to perform linkage. One pedigree, in which a single nonsense allele (Y24X) was detected, was consistent with linkage to *BBS2* but, the remaining five pedigrees were all excluded from *BBS2* based on either sharing of only a single haplotype between affected sibs, or sharing of identical haplotypes between affected and unaffected sibs.

Analysis of all mutational data for *BBS2* and *BBS6* in the 163 patients from the cohort led to the identification of several pedigrees in which three mutant alleles (a homozygous mutation in *BBS2* with a heterozygous mutation in *BBS6*, or vice versa) segregated with disease (Katsanis et al., 2001a), resulting in a triallelic digenic inheritance pattern (Rivolta et al., 2002). In three pedigrees, homozygous or compound heterozygous *BBS2* mutations and a heterozygous mutation in *BBS6* were found to segregate with disease. In pedigree B14, one of the pedigrees that initially led to questions over the inheritance pattern of BBS (See Figure 1. 4), a homozygous Y24X (*BBS2*) nonsense mutation in the affected individual was detected in addition to the A242S heterozygous mutation carried by both the affected individual and his unaffected sib. In a second pedigree the affected individual was a compound heterozygote for frameshift (V158fsX170) and nonsense (R216X) *BBS2* mutations, with a third mutant allele in *BBS6* (C499S). In both these pedigrees it is possible that the third allele, a *BBS6* missense mutation in each case, is in fact a SNP rather than a pathogenic mutation. This is unlikely though as the A242S mutation has been reported in an MKKS patient (Stone et al., 2000) and neither the A242S or C499S substitutions were found in 384 matched control chromosomes (Katsanis et al., 2001a). Due to small pedigree sizes it is not possible to say that three mutations are required for disease in these pedigrees as there are no unaffected individuals that harbour two mutant alleles at the same locus. In the third pedigree (AR259, Figure 5. 5) however, it is possible to say that three mutant alleles are required to manifest disease. A compound heterozygous mutation involving two nonsense mutations (Y24X and Q59X) in *BBS2* was present in both the affected individual (03) and his unaffected sib (05). Analysis of *BBS6* mutation data in this pedigree revealed that only 03 carried a third nonsense mutation (Q147X).

In a fourth pedigree, a single N70S (*BBS2*) mutation was identified in combination with a homozygous Y37C (*BBS6*) mutation in both affected individuals. The pathogenicity of the third mutation in this case has been questioned as it has been suggested that it may be a SNP. The fact that the change was never observed in 384 control chromosomes, the asparagine at residue 70 is highly conserved and the substitution of a serine (or any other amino acid) at this position is predicted as intolerant using the SIFT analysis program (Ng and Henikoff, 2002), all suggest that the change may be pathogenic. Functional studies would be required

to determine whether this change is benign or disease-causing. These data suggest that, in a subset of BBS pedigrees, two mutant alleles at a single *BBS* locus are not always sufficient to cause disease.

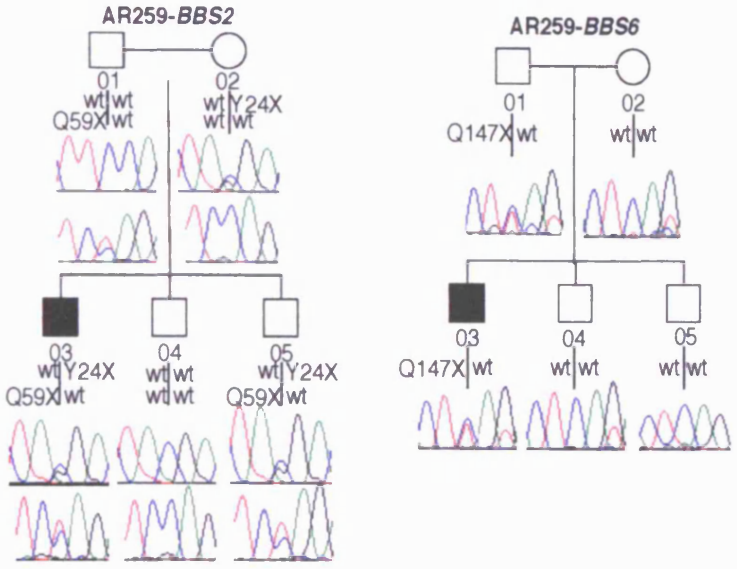


Figure 5. 5: Triallelic inheritance in AR259. In this pedigree three mutations are required for disease as the unaffected individual 05 is a compound heterozygote for two nonsense mutations in *BBS2* but does not carry the third nonsense allele in *BBS6* that has been inherited by his affected brother (03) (Katsanis et al., 2001a).

5.4.2 Prevalence of triallelic inheritance in BBS

Triallelic inheritance is not restricted to only *BBS2* and *BBS6*. Triallelism has been reported, to varying degrees, in all cloned *BBS* genes with the exception of *BBS8* (as only three pedigrees with mutations in *BBS8* have been reported at present, cases of triallelism involving this new gene may still be discovered). If triallelic inheritance occurs at a *BBS* locus, certain observations are expected to occur (Beales et al., 2003):

1. A homozygous mutation at the locus (locus A) with a single mutation at a second *BBS* locus (locus B) in an affected individual.

2. A single mutant allele at locus A with a homozygous mutation at locus B in an affected individual.
3. A single mutant allele at locus A with no evidence for additional mutations at any of the known loci in an affected individual.
4. Unaffected individuals with two mutant alleles (a homozygous or compound heterozygous mutation) at locus A.

Although the majority of cases of *BBS1* are accounted for by recessive mutations (for ~80% of *BBS1* mutations, two mutant alleles are sufficient to cause disease (Katsanis, 2004)), several pedigrees with *BBS1* mutations have been identified in which one or more of the expected observations of triallelic inheritance have been noted. Three of the expected observations were identified within pedigrees from our own cohort. In pedigree PB056 a homozygous *BBS1* mutation (M390R) was found in combination with a heterozygous missense mutation in *BBS4* (M472V) in both affected individuals. In this pedigree, as there are no unaffected individuals with two *BBS1* mutant alleles, it is not possible to say that three mutant alleles are required to manifest disease. However due to the high level of conservation of the methionine at codon 472 (it is conserved in several species including cow, chicken and xenopus) and the absence of the substitution from control chromosomes, it is unlikely that this change is a benign polymorphism (Beales et al., 2003).

No single *BBS1* mutations in combination with homozygous mutations at a second *BBS* locus were identified in our cohort, but were present in the cohort of our collaborator; a heterozygous *BBS1* missense mutation (E234K) was present in combination with a homozygous *BBS7* missense mutation (T211I) in one family and a single M390R allele was present in association with a complex *BBS2* mutation ([R315Q + IVS1+1G>C)] + [R315Q]) in a second family (Beales et al., 2003). In three pedigrees from our cohort, single M390R (*BBS1*) alleles were identified, with no additional sequence variations detected in either *BBS1* or any of the other known *BBS* genes in affected individuals. Also, only a single H35R (*BBS1*) allele was detected in a small consanguineous pedigree despite complete coverage of the *BBS1* ORF; no fu .her mutations in *BBS2*, *4*, *6*, *7* or *8* were identified in this pedigree. There may be a second *BBS1* mutation in these pedigrees which was not detected by direct sequencing, but it is also possible that the pedigrees are examples of triallelism with a

homozygous mutation in a novel *BBS* gene. (Single R440X (*BBS1*) and I339V (*BBS6*) alleles were also identified in two isolated cases of BBS. These individuals may represent cases of triallelism, but as it was not possible to achieve complete coverage of the ORF of the respective genes, the occurrence of a second mutation and a recessive pattern of inheritance in these patients can not be ruled out.)

5.4.3 Tetrallelic inheritance in a consanguineous pedigree

The results from mutation screens of the known *BBS* genes within our cohort, and that of our collaborator, have provided examples of segregation of three mutant alleles with disease in a subset of families (Badano et al., 2003a; Beales et al., 2003; Katsanis, 2004; Katsanis et al., 2001a). Mutational data from a consanguineous Turkish family from our cohort suggest that three alleles may not always be sufficient to cause BBS. The affected individual from pedigree PB043 is homozygous for two missense mutations; T558I in *BBS2* and A364E in *BBS4*. Restriction digest analysis in other members of the pedigree revealed inheritance of three mutant alleles (T558I homozygous mutation with a heterozygous A364E mutation) in an unaffected sib and the mother. As both mutations in this pedigree are missense, the pathogenicity of each is hard to predict. One explanation for the occurrence of unaffected family members segregating three *BBS* mutant alleles is that the T558I substitution in *BBS2* is a rare polymorphism and that the only disease causing sequence alteration in this family is the A364E allele in *BBS4*, which has a segregation pattern consist with standard recessive inheritance. The other explanation is that, in this pedigree, four mutant alleles (tetrallelic inheritance) are required to manifest disease. Neither the T558I nor the A364E substitutions were present in control chromosomes and both residues are at highly conserved positions within the respective proteins (Katsanis et al., 2002). A functional assay would be required to confirm whether it is only the A364E allele that accounts for disease in this pedigree or whether both missense mutations contribute to the BBS phenotype in the affected individual.

5.4.4 Mutations that modify the BBS phenotype

Initial findings of oligogenic inheritance in BBS suggested that, in pedigrees in which triallelism occurs, three mutations were required for disease. Individuals that have inherited only the homozygous or compound heterozygous mutation at one locus were therefore expected to be asymptomatic, as was the case in pedigree AR259 (See Figure 5. 5). Further screening of all BBS pedigrees for mutations in the known genes has resulted in the identification of two pedigrees from our cohort in which the third allele, rather than having an 'all-or-nothing' effect on the BBS phenotype, seems to modify the disease phenotype resulting from a homozygous mutation at a single *BBS* locus.

Pedigree PB009 contains three affected individuals that are all M390R (*BBS1*) homozygotes. The phenotype of the three sibs shows considerable variation, in particular in relation to the retinal features of the disease (See Table 5. 1). Individuals 01 and 02, aged 32 and 35 years respectively, each developed RP in the second decade of life (01 at 20 years of age, 02 at 15 years of age) and are now registered legally blind. In contrast, RP did not develop until the age of 34 years in individual 03, who, at the age of 37, stills enjoys good vision by day. Individuals 01 and 02 are also considered to be disinhibited and developed obesity early in infancy; both currently has a BMI of 33kg/m^2. Individual 03 does not share the altered behavioural phenotype of her sibs and, with a BMI of 29kg/m^2, is the least overweight of the three. A possible explanation for the more severe phenotype seen in individuals 01 and 02 is the inheritance of a third mutation, a paternally inherited L349W (*BBS2*) allele, present in these patients but not in their less severely affected sibling. The variability of the retinal phenotype in this pedigree is much greater than that seen in other pedigrees segregating one or two M390R alleles. Analysis of ten M390R pedigrees, each containing two or more affected sibs, that exhibit 'recessive' BBS (no evidence for a third mutant allele) revealed that the mean intrafamilial variation in the age of onset of RP was 2.3 years with a standard deviation of 1.7. In pedigree PB009, although the variation between individuals 01 and 02 was within the normal range (five years), the maximum variation in this pedigree of 19 years was well outside this range. The most likely explanation for this much higher degree of variation in this pedigree is the presence of the third allele in individuals 01 and 02. To

investigate how this additional mutation results in a more severe phenotype, functional studies of the effect of the L349W mutation on the BBS2 protein would be required. Using the SIFT analysis program (Ng and Henikoff, 2002), the substitution of a tryptophan for the leucine residue at codon 349 is predicted to be deleterious, suggesting that this allele does contribute to the severity of the phenotype (Badano et al., 2003b).

A modifier mutation at a second *BBS* locus was also identified in another pedigree from our cohort. In this pedigree (PB061), both patients were homozygous for the R275X (*BBS2*) nonsense mutation, which is also present in the unaffected sib in heterozygous form. Analysis of *BBS1* sequence data in this pedigree revealed the presence of a third allele, a splice site mutation (IVS15+2T>C) inherited from the mother, in individual 02. This third mutation was not present in the other affected individual (01) in the pedigree but was present in the unaffected offspring (03) who is therefore a double heterozygote for the R275X (*BBS2*) and IVS15+2T>C (*BBS1*) mutations. In this pedigree, similar to PB009, there is an apparent co-segregation of the third allele with a more severe BBS phenotype, in this case in relation to developmental delay and weight control (See Table 5. 2). Individual 01 exhibited a slight delay in reaching developmental milestones, but was within the normal range (sat at nine months, walked at 15 months, first words spoken at 12 months), she performed well at secondary school, requiring only minor support for her visual impairment, and has never suffered with obesity. In contrast, her brother (02) showed considerable developmental delay (sat at 11 months, walked at 21 months, first words spoken at two years of age), required support throughout schooling due to considerable learning difficulties and is only able to control his weight with a strict diet (Badano et al., 2003b). The most common outcome of donor splice site muations such as the IVS15+2T>C in this pedigree, is skipping of the exon containing the donor site mutation, in this case exon 15, corresponding to an in-frame deletion of 135bp (Attanasio et al., 2003; Brose et al., 2004).

5.4.5 Oligogenic inheritance in NPHP involving mutations in *NPHP* and *BBS* genes

Owing to the similarities between NPHP and BBS (including clinical features, protein function, degree of genetic heterogeneity and inheritance pattern, see *5.1.2 Evidence of possible oligogenic inheritance in NPHP*, for a comparison of NPHP and BBS), a cohort of 95 NPHP patients were screened for mutations in the *BBS* genes to investigate the possibility of oligogenic inheritance between the two disease genes families. Three SLS patients who had *NPHP* mutations (*NPHP1* and *NEK8*) with an additional heterozygous mutation in a *BBS* gene (*BBS4* or *6*) were identified. These results indicate that oligogenic inheritance may exist between the two syndromes. It is not known at this stage whether the *BBS* mutation is required to manifest NPHP and RP in these pedigrees or if the additional mutation modifies the phenotype caused by mutations at the *NPHP* locus. Further studies including segregation of all mutations within the pedigree and screening of additional NPHP patients for *BBS* mutations will be required to establish if the BBS mutation is essential and if there is a specific phenotype associated with mutations in both *NPHP* and *BBS* genes. Functional studies are also required to determine what effect the novel missense mutations (K46R and K441T in *BBS4* and R23G in *BBS6*) have on the BBS proteins; the substitutions may reside in important functional or interaction domains within BBS4 and BBS6 and may disrupt potential interactions between the NPHP and BBS proteins. A parallel study of BBS patients for mutations in the *NPHP* genes is currently underway.

5.4.6 Further examples of non-Mendelian inheritance in BBS

Mutation analysis of BBS patients by other groups has also provided evidence of a non-Mendelian pattern of inheritance in BBS. Following a study of 27 BBS and MKKS patients for mutations in the *BBS6* (*MKKS*) gene, Slavotinek *et al.* (2002) reported the identification of five BBS pedigrees in which only a single *BBS6* mutant allele was identified. To investigate the possibility of triallelic inheritance of *BBS6* and *BBS2* in these pedigrees, all patients with heterozygous *BBS6* mutations were screened for coding sequence alterations in

the *BBS2* gene. No pedigrees with mutations in both genes were found, however the authors confirm that the detection of only a single mutant allele in four pedigrees does suggest that inheritance of BBS may be more complex than first thought (Slavotinek et al., 2002).

A study of twenty-one patients of European origin by Fauser *et al.* (2003) resulted in the identification of pedigrees with mutations at more than one *BBS* locus. All patients were screened for the presence of sequence alterations in *BBS1*, *2*, *4* and *6*. Recessive inheritance of mutations in *BBS1* and *6* were identified in six patients (four had compound heterozygous mutations in *BBS1*, one was homozygous for M390R (*BBS1)* and the sixth pedigree had a homozygous mutation in *BBS6*). Mutations in *BBS2*, *4* and *6* were identified in a further three individuals, however these mutations did not fit with a recessive model of inheritance. A novel heterozygous D492N missense mutation (*BBS6*) was identified in one individual with no evidence for a second *BBS6* mutant allele. In two individuals heterozygous mutations in *BBS2* were found in combination with heterozygous mutations in *BBS4*. A nonsense mutation (R413X) in *BBS2* and a missense mutation (P503L) in *BBS4* were present in one individual, and in the other, missense mutations in both *BBS2* (R643H) and *BBS4* (K46R). The segregation pattern of mutations in two individuals from this cohort do not conform to the pattern observed in the triallelic pedigrees from our own cohort but do support a model of oligogenic inheritance in BBS in which mutations at more than one BBS locus are required in some pedigrees.

In 2003, Mykytyn *et al.* (2003) reported that there was no evidence for oligogenic inheritance involving *BBS1* in their cohort. To assess for oligogenic inheritance, 43 unrelated individuals with two *BBS1* mutations were screened for additional sequence alterations in *BBS2*, *4* and *6*. In total eight sequence variations were identified; I123V and A504V (*BBS2*), K46R, I70V and T354I (*BBS4*) and R517C and G532V (*BBS6*). All eight alterations were deemed to be SNPs by the authors based on the conservative nature of the change, the presence of the alteration in a sample of controls or a segregation pattern that does not fit with disease. Although applicable to recessive conditions, these criteria are not an appropriate means of assessing the pathogenicity of a sequence alteration within a gene associated with a complex form of inheritance. Given that alleles involved in complex inheritance are expected to have a higher carrier frequency in the general population than those associated with a recessive

condition, the observation of an allele in a sample of control chromosomes does not necessarily mean that the alteration is not pathogenic. As the authors do not state the number of control chromosomes that were screened or the frequency with which any of the eight alterations were observed, it is not possible to determine whether the alterations were present at a high frequency (expected if the alteration is a benign polymorphism) or at a low frequency which could be accounted for by an elevated carrier frequency.

In addition to a higher carrier frequency, oligogenic alleles will not conform to the segregation pattern of mutations involved in recessive inheritance and may therefore be present in unaffected individuals. As the authors do not fully elaborate on the segregation pattern of the alterations within the pedigrees in which they were detected, assessment of whether the segregation pattern is consistent with either triallelic inheritance or a modifier allele is not possible. Following the study by Mykytyn *et al.* (2003) the K46R allele in *BBS4* has been both identified in one of the NPHP patients from our NPHP-BBS oligogenicity study, and also reported in a BBS patient from the Fauser *et al.* (2003) study, suggesting that this mutation is likely to be associated with disease rather than a benign polymorphism.

Finally, in 16.7% (2/12) of families that were found to be heterozygous for the M390R allele in the Mykytyn *et al.* (2003) cohort, no second *BBS1* allele was reported. A single *BBS4* heterozygous mutation (V195fsX209) was also detected in a small consanguineous pedigree during the cloning of the gene by this group (Mykytyn et al., 2001). All of these factors are highly suggestive that more complex patterns of inheritance may in fact exist within this patient cohort.

5.4.7 Summary

Research into the genetics of BBS in recent years has uncovered a much higher degree of complexity than anticipated in what was initially thought to be a single gene recessive disorder. Eight *BBS* loci are now known, with evidence for at least one additional locus in the human genome and no genotype-phenotype relationship. Examples of triallelic, and

potentially tetra-allelic, inheritance in which multiple mutant alleles at two loci are required to manifest disease have been identified in a subset of families. The presence of a third mutation in certain pedigrees has also been found to co-segregate with a more severe BBS phenotype. The majority of known *BBS* genes have already been shown to participate in triallelic inheritance. The occurrence of single mutations in *BBS* genes with no additional mutations at the remaining known loci suggests that novel *BBS* genes may also play a part in this complex mode of inheritance. Examples of complex inheritance involving mutations in the *BBS* genes is not only restricted to the BBS phenotype. Heterozygous BBS mutations have also been detected in combination with mutations in the NPHP genes in nephronophthisis patients.

Chapter 6 Multiplex Capillary Heteroduplex Analysis (MCHA) as a mutation detection technique for BBS

6.1 Introduction

Direct sequencing is considered the gold standard for mutation detection but with a high cost per sample and the requirement for a strong, purified PCR fragment it is not always suitable for routine screening of a large number of samples. There were several reasons why an alternative to sequencing was required for screening our cohort of BBS patients. The occurrence of complex inheritance, involving mutations at more than one locus in some families, means that linkage analysis is no longer a reliable means of locus assignment and therefore all new cases of BBS must be screened for mutations in all known *BBS* genes. Almost 100 fragments are required to screen each patient for mutations in *BBS1*, *2*, *4*, *6*, *7* and *8* and with over half of patients unlinked to any of these known loci, mutation screening of a patient for mutations in known, and newly identified, *BBS* genes by sequencing is prohibitively expensive.

6.1.1 Mutation detection techniques

Mutation screening techniques can be classified as either specific, identifying defined changes within a short segment of DNA, or scanning, detecting uncharacterized alterations usually in longer DNA fragments. Specific techniques including primer extension, oligonucleotide ligation assay and allele-specific amplification are used to detect common pathogenic mutations in genes that account for the majority of cases of a given disease. This is not the case in BBS as several types of mutations including missense, nonsense, frameshift and whole-exon deletions have been described in association with disease, with no evidence for positional clustering of mutations or population specific genes or alleles. There are also

very few common mutations in *BBS* genes, the most common of which is the missense M390R in *BBS1* (Beales et al., 2003; Mykytyn et al., 2003; Mykytyn et al., 2002). The majority of mutations are therefore either private or segregate with disease in only a small number of families. A scanning method, that will detect known and novel mutations, is therefore more appropriate for screening new cases of BBS.

The characteristics of a good scanning technique combine high sensitivity, throughput, low cost with a quick and simple protocol (Cotton, 1997). The majority of scanning techniques utilise the properties of heteroduplex species that are formed due to the presence of a mismatch or small insertion/deletion in a double-stranded DNA (dsDNA) fragment. In comparison to a homoduplex fragment, heteroduplexes have an abnormal denaturing profile that is used in denaturing gradient gel electrophoresis (DGGE) and denaturing high performance liquid chromatography (DHPLC). The slower mobility of heteroduplexes in a non-denaturing gel is used in heteroduplex analysis (HA). The chemical cleavage of mismatches (CCM) method exploits the fact that mismatched bases in a heteroduplex are more sensitive to cleavage by chemicals.

All scanning methods have advantages and disadvantages making them suited to particular applications (See Table 6. 1). Single strand conformation polymorphism (SSCP), DHPLC and DGGE require considerable initial optimisation of primers, temperature, pH, and, in the case of DGGE, the denaturing gradient used. These factors mean that these techniques are most suited to screening a large number of patients for a small number of different fragments, the conditions for which have been individually optimised. The protein truncation test (PTT) only detects nonsense mutations and will therefore not pick up benign polymorphisms, a disadvantage of most scanning methods, but it will also miss potentially pathogenic missense mutations. The lack of optimisation required for individual fragments when using HA makes it a popular method, suitable for screening a cohort of patients for mutations in a number of different fragments.

Name	Features	Advantages	Disadvantages	Reference
Denaturing High Performance Liquid Chromatography (DHPLC)	Differential retention of heteroduplex and homoduplex DNA in a separation matrix under partially denaturing conditions.	High sensitivity in fragments from 150-700bp. Useful for screening large number of samples.	Requires extensive optimization. Surrounding sequence may affect detection ability.	Liu et al, 1998. Xiao and Oefner, 2001.
Denaturing Gradient Gel Electrophoresis (DGGE)	Differential migration of heteroduplex and homoduplex DNA fragments through a denaturing gradient gel.	High sensitivity and specificity. Fragments can be pooled to improve throughput.	Labour intensive design and testing for a range of primers for gene of interest.	Guldberg and Guttler, 1993. Hofstra et al, 2004.
Chemical Cleavage of Mismatches (CCM)	Mismatched bases are sensitive to cleavage by certain chemicals such as potassium permanganate.	Gives indication of position of mismatch.	Chemicals used are highly toxic and technique requires some skill to obtain good results.	Smooker and Cotton, 1993. Lambrinakos et al, 1999.
Protein Truncation Test (PTT)	Uses in vitro transcription/translation from amplified RNA to identify premature termination codons resulting in a truncated protein product.	Only functional mutation detection method.	Requires RNA from patient. Only detects nonsense mutations.	van der Luijt et al 1994. Den Dunnen and van Ommen, 1999.
Single Strand Conformation Polymorphism (SSCP)	Single stranded fragments of DNA that differ by a single base will have a different secondary structure and hence mobility in a non-denaturing gel.	Simple, non-toxic or radioactive protocol	Sensitivity can be influenced by electrophoresis conditions, position of mismatch and fragment length	Orita et al, 1989. Highsmith et al, 1999b.
Heteroduplex Analysis (HA)	Heteroduplexes will have a slower mobility than homoduplex DNA in a non denaturing gel.	Simple protocol, requires little optimization for different fragments.	Nature of mismatch affects detection ability. Optimal fragment length <700bp.	Highsmith et al, 1999a. Ganguly, 2002.

Table 6. 1: Comparison of commonly used mutation scanning techniques.

6.1.2 Heteroduplex species

The presence of a mutation in a dsDNA fragment can result in two different types of heteroduplexes depending on the nature of the mutation. In the case of a small insertion or deletion a 'bulge' is formed as a result of a number of unmatched bases, whereas in the case of base substitutions the corresponding mismatched bases result in a 'bubble' (Bhattacharyya and Lilley, 1989). Fragments containing a bulge have a large retardation through a gel, caused by bending of the DNA at the site of the bulge, making them easy to detect by HA. The larger the insertion/deletion, the greater the degree of bending and hence the greater the retardation of the fragment. The effect of a bubble on the mobility of a fragment is subtler, owing to less bending of the DNA, and can be influenced by several factors including the length of the fragment and the nature and position of the mismatch within the fragment (Highsmith et al., 1999a). In an individual with a heterozygous mutation, naturally occurring heteroduplexes will be present in the PCR product amplified from the patient DNA. In an individual with a homozygous mutation the patient amplified PCR product must be mixed with the PCR product amplified from control DNA, denatured and allowed to reanneal slowly to generate four species; a wild-type (wt) homoduplex, a mutant homoduplex and two heteroduplexes (Figure 6. 1).

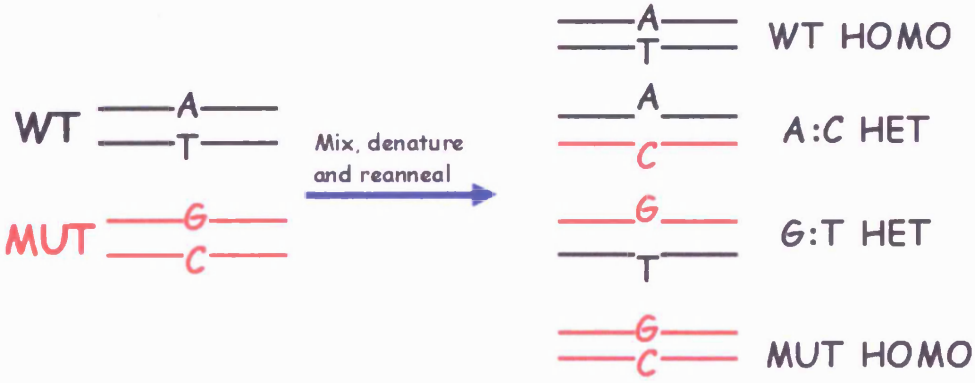


Figure 6. 1: Heteroduplex formation. In the above example, an amplified sample from an individual with a homozygous A>G substitution is mixed with a wt sample, denatured and allowed to reanneal slowly. Four species are then formed, a wt homoduplex, a mutant homoduplex and two heteroduplexes; one containing an A:C mismatch and the other, a G:T mismatch.

The occurrence of heteroduplex fragments with a retarded mobility in a polyacrylamide or agarose gel was first reported by Nagamine (1989) as a PCR artefact but has since been developed as a mutation detection technique. Traditional slab-gel HA is performed using standard gel electrophoresis apparatus with subsequent silver or ethidium bromide staining to visualise the bands. In 1998 Ganguly *et al.* modified the technique for use on a fluorescent platform; fragments labelled with either 6-FAM, HEX or NED were resolved using an ABI 377 sequencer. To improve the sensitivity and resolution of fluorescent HA further, Rozycka *et al.* (2000) adapted the technique for use on a single capillary fragment analyser (ABI 310) and were able to achieve high sensitivity in fragments of <350bp. Increases in throughput have been made possible by using a 96-well capillary sequencer, the MegaBACE 1000 (Amersham Biosciences), and pooling of several different sized fragments in each capillary. Following these modifications to standard HA this new technique has been named multiplex capillary heteroduplex analysis (MCHA).

6.2 Methods

6.2.1 Primer modification

Primers designed for the mutation screening of *BBS1*, *2*, *4*, *6* and *7* were adapted for use with the MCHA technique by the addition of a fluorescent label (FAM, HEX or TET) to the 5' of the forward primer of each pair. The colour of the label was selected to allow optimal pooling of fragments based on size and colour of label. See Appendices 7-12 for information on the size of the amplicons and the colour of label used.

6.2.2 Amplification and heteroduplex formation

All PCR reactions were carried out for both patient and control samples as described in section *2.2.1.2 PCR reactions*. Following amplification, patient samples were mixed with the control sample, denatured and allowed to reanneal slowly to generate heteroduplexes (See *2.2.4.1 Heteroduplex formation*).

6.2.3 Identification of base substitution or insertion/deletion

Following a positive MCHA result (additional peaks or altered peak profile in comparison to the control sample), the patient samples were re-amplified for the exon of interest using unlabelled primers and sequenced to identify the underlying alteration in the patient sample (See *2.2.3 Direct sequencing*).

6.3 Results

6.3.1 Sensitivity testing

To test the sensitivity of the MCHA method, known alterations including small insertion/deletions, missense and nonsense mutations in *BBS1* (D148N, Y284fsX288 and M390R), *BBS2* (D170fsX171 and R275X), *BBS4* (A364E) and *BBS7* (K237fsX296) and several single nucleotide polymorphisms (SNPs) in *BBS6* (P39P, I178I, c.985+16T>G, c.985+33G>C and c.1161+58A>G) were used. These alterations (with the exception of the SNPs) were selected as DNA from both homozygous and heterozygous individuals was available for each mutation, thus permitting us to test the efficiency of the heteroduplex formation step by comparison of the artificially created heteroduplexes (a homozygous sample mixed with wt) with the naturally occurring heteroduplexes (a heterozygous sample). In all cases the artificially created heteroduplexes gave an identical pattern to those of a heterozygous individual (See Figure 6. 2).

Eleven of the 12 known changes were successfully detected using the initial primer set and run conditions. The only change that was not detected was the M390R mutation in the *BBS1*x12 fragment. In an attempt to detect this change, in both heterozygous and homozygous individuals, the MCHA run temperature was increased by 5°C to 30°C. Despite this change, the sequence alteration remained undetected. When a plate of all known changes was run at both 25°C and 30°C, no difference in peak profile was observed, indicating that a difference of 5°C, under standard run conditions for all changes tested, has no effect on the sensitivity of the assay. As altering the run conditions did not improve the ability to detect this change, the forward primer for the *BBS1*x12 fragment was redesigned, increasing the size of the fragment (from 368bp to 458bp in length), decreasing the GC content slightly (from 54.5% to 52.5%) and changing the position of the mismatch within the fragment (from 0.59 to 0.67). Following the design of a new forward primer, and resulting alteration to the *BBS1*x12 fragment, the M390R mutation was detected in all known heterozygous and homozygous individuals, giving a sensitivity of 100% for all known changes (Figure 6. 3).

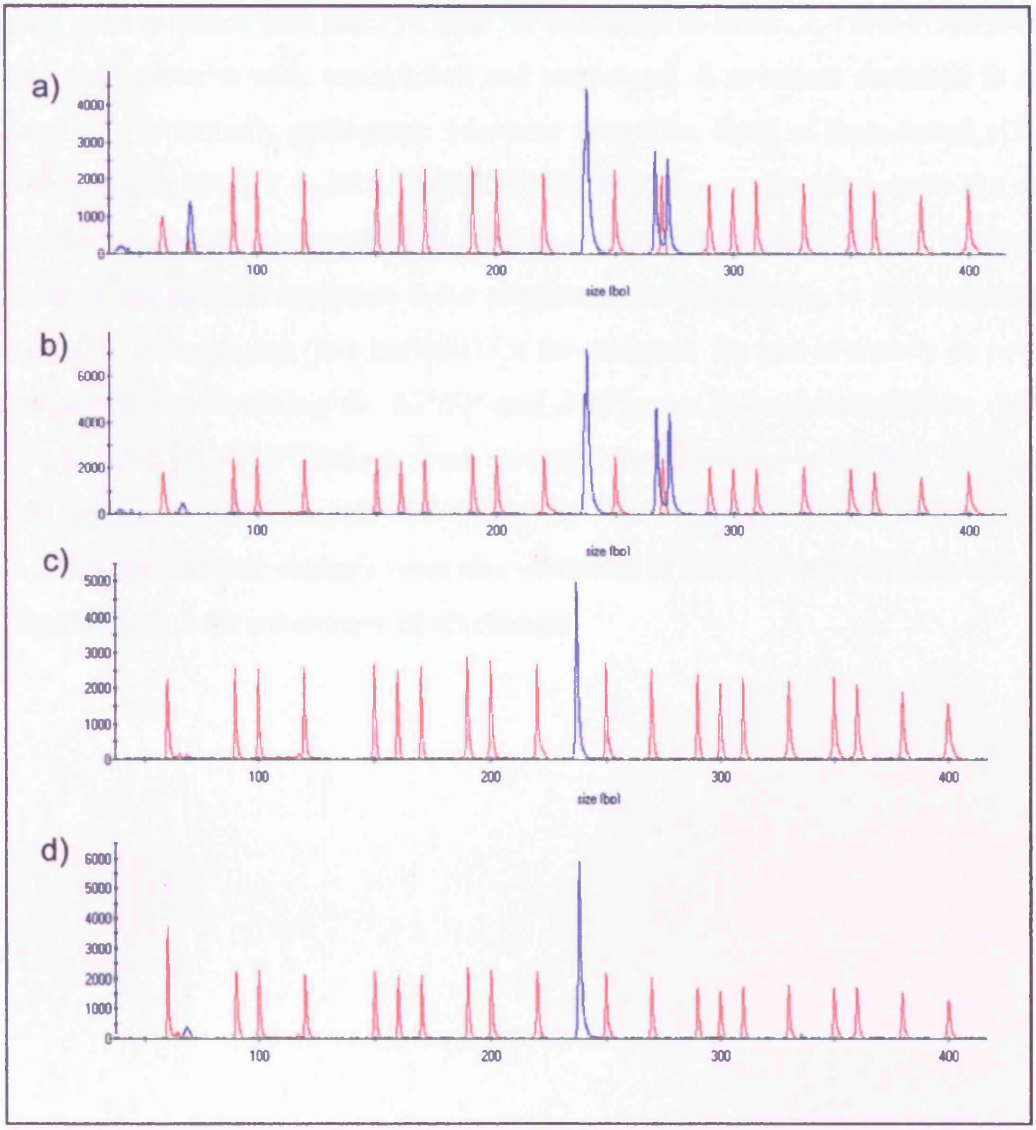


Figure 6. 2: Electropherograms showing a 2bp deletion in *BBS2*x4. a) A homozygous patient sample mixed with wt, giving a mutant homoduplex, a wt homoduplex and two heteroduplexes, b) A sample from a parent of the patient who is heterozygous for the mutation, c) Mutant homoduplex from the patient sample, d) wt homoduplex. ET-Rox size standard (red) was also added to the samples.

6.3.2 Identification of alterations in new cases of BBS and in NPHP patients

As a very high sensitivity was achieved for the assay, the MCHA technique was subsequently used to screen new cases of BBS for mutations in *BBS1*, *2*, *4* and *6*. All samples with altered peak patterns were reamplified and sequenced. A nonsense mutation in *BBS1* (R440X) and four potentially pathogenic missense mutations, three of them novel, (D174E and L349W in *BBS2*, M472V in *BBS4* and I339V in *BBS6*) were identified, each in a single family. In addition five patients with peak profiles similar to that of the pattern observed in the presence of the M390R mutation were confirmed on sequencing to be heterozygous (three patients) or homozygous (two patients) for the mutation. As part of a study on possible oligogenic inheritance involving the *NPHP* and *BBS* genes in nephronophthisis (NPHP) patients, a cohort of 95 NPHP patients were screened for mutations in the four *BBS* genes, three novel missense mutations (K46R and K441T in *BBS4* and R23G in *BBS6*) were identified. SNPs and intronic changes were also identified in *BBS1*, *2* and *4* in both cohorts of patients. See Table 6. 2 for a summary of all changes.

Amplicon	GC content of amplicon	Local GC content	Alteration	Result of change	Position of alteration (5'>3')	Hetero-duplexes
***BBS1* x5+6**	**59.8%**	**55.0%**	**c.442G>A**	**D148N**	**0.33 (138/417)**	**A:C T:G**
BBS1 x7	48.2%	55.0%	c.519-76C>T	Intronic	0.23 (88/381)	A:C T:G
***BBS1* x10**	**57.0%**	**60.0%**	**c.851delA**	**Y284fsX288**	**0.50 (214/432)**	
***BBS1* x12**	**52.5%**	**55.0%**	**c.1169T>G**	**M390R**	**0.67 (306/458)**	**G:A C:T**
BBS1 x13	56.0%	60.0%	c.1318C>T	R440X	0.61 (280/460)	A:C T:G
BBS1 x17	64.1%	60.0%	c.1782+7A>G	Intronic	0.70 (332/472)	A:C T:G
***BBS2* x4**	**39.7%**	**45.0%**	**c.511-512delTT**	**D170fsX171**	**0.51 (123/239)**	
BBS2 x4	39.7%	40.0%	c.522T>A	D174E	0.54 (130/239)	A:A T:T
BBS2 x5	37.5%	45.0%	c.612+12C>A	Intronic	0.79 (190/240)	G:A C:T
BBS2 x8	36.6%	25.0%	c.805-20A>G	Intronic	0.28 (98/349)	A:C T:G
***BBS2* x8**	**36.6%**	**55.0%**	**c.823C>T**	**R275X**	**0.39 (136/349)**	**A:C T:G**
BBS2 x9	50.0%	50.0%	c.1045T>G	L349W	0.56 (173/308)	G:A C:T
BBS4 x1	64.3%	60.0%	c.1-17C>A	Intronic	0.35 (120/343)	G:A C:T
BBS4 x3	38.4%	30.0%	c.147A>G	K46R	0.48 (164/343)	A:C T:G
BBS4 x5	44.4%	35.0%	c.332+27-28insA	Intronic	0.78 (326/417)	
BBS4 x13	46.0%	40.0%	c.1061T>C	T354I SNP	0.67 (214/318)	A:C T:G
***BBS4* x13**	**46.0%**	**55.0%**	**c.1091C>A**	**A364E**	**0.77 (244/318)**	**G:A C:T**
BBS4 x15	49.9%	40.0%	c.1249-33G>C	Intronic SNP	0.23 (108/473)	C:C G:G
BBS4 x15	49.9%	35.0%	c.1322A>C	K441T	0.46 (216/473)	G:A C:T
BBS4 x15	49.9%	55.0%	c.1414A>G	M472V	0.65 (308/473)	A:C T:G
BBS4 x16	42.6%	35.0%	c.1452-45C>T	Intronic SNP	0.30 (129/429)	A:C T:G
BBS6 x3a	39.6%	65.0%	c.67A>G	R23G	0.42 (209/492)	A:C T:G
***BBS6* x3a**	**39.6%**	**55.0%**	**c.117C>T**	**P39P SNP**	**0.53 (260/492)**	**A:C T:G**
***BBS6* x3c**	**39.5%**	**50.0%**	**c.534C>T**	**I178I SNP**	**0.23 (91/392)**	**A:C T:G**
***BBS6* x3d**	**40.5%**	**40.0%**	**c.985+16T>G**	**Intronic SNP**	**0.71 (298/421)**	**G:A C:T**
***BBS6* x3d**	**40.5%**	**30.0%**	**c.985+33G>C**	**Intronic SNP**	**0.75 (315/421)**	**C:C G:G**
BBS6 x4	34.9%	45.0%	c.1015A>G	I339V	0.28 (135/475)	A:C T:G
***BBS6* x4**	**34.9%**	**31.0%**	**c.1161+58A>G**	**Intronic SNP**	**0.71 (339/475)**	**A:C T:G**
***BBS7* x7**	**27.0%**	**40.0%**	**c.711-714del GAGA**	**K237fsX296**	**0.32 (151/468)**	

Table 6. 2: Table of changes studied using the MCHA method. Local GC content relates to the GC content of the 10bp either side of the change. Known changes that were used to test the sensitivity of the technique are in bold, all others are novel changes that were detected using the technique.

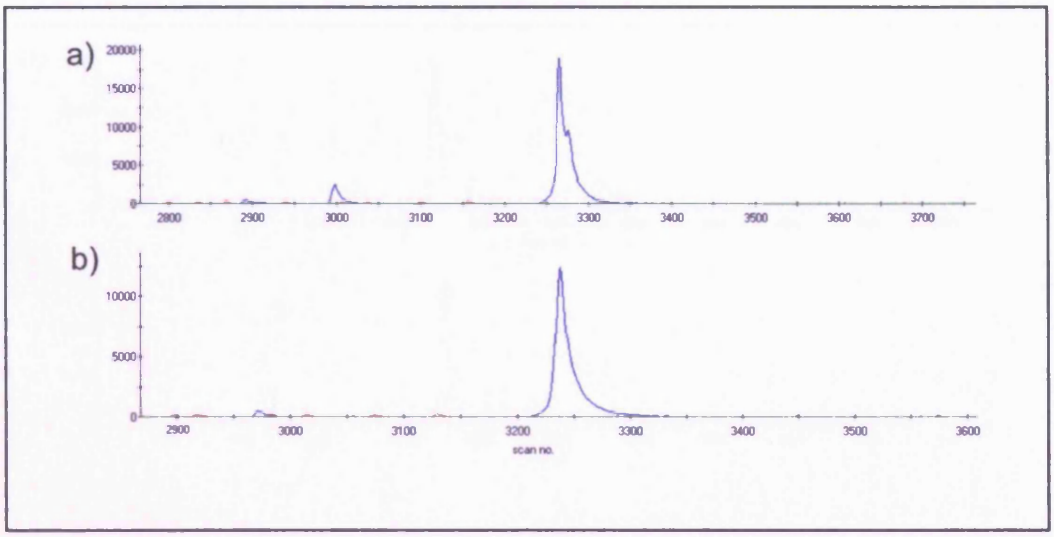


Figure 6. 3: Electropherograms showing the peak profile of the M390R (*BBS1*) mutation. a) A heterozygous individual with a G:A mismatch in the 458bp FAM labelled *BBS1*x12 fragment, b) a wt homoduplex.

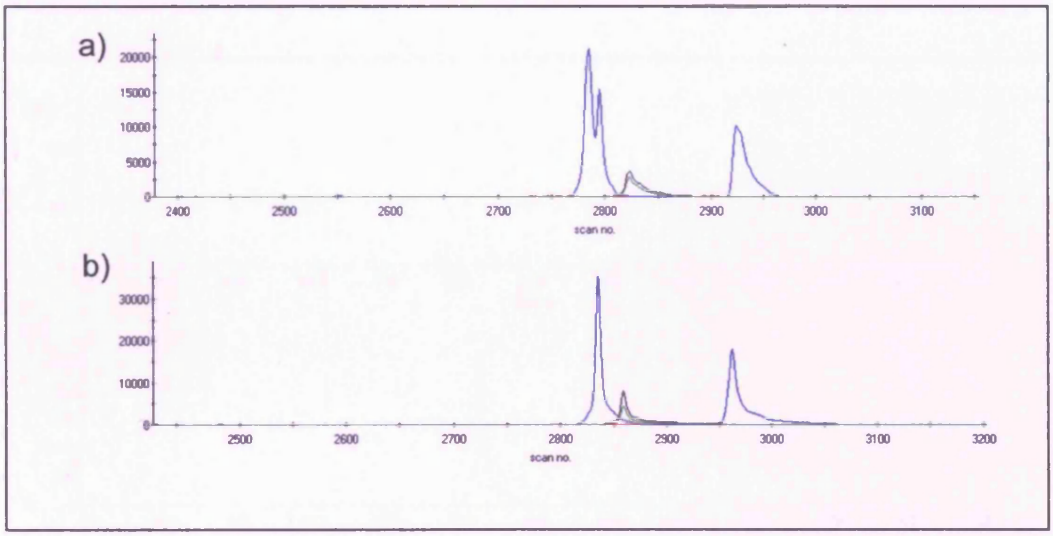


Figure 6. 4: Pooling of multiple fragments improves the throughput of MCHA. Three fragments (*BBS6*x3c, *BBS6*x3d and *BBS6*x3a) from the same individual are pooled in the same capillary. a) Results from an individual who is heterozygous for SNPs in *BBS6*x3c (I178I) and *BBS6*x3a (P39P), b) wt homoduplexes for each fragment.

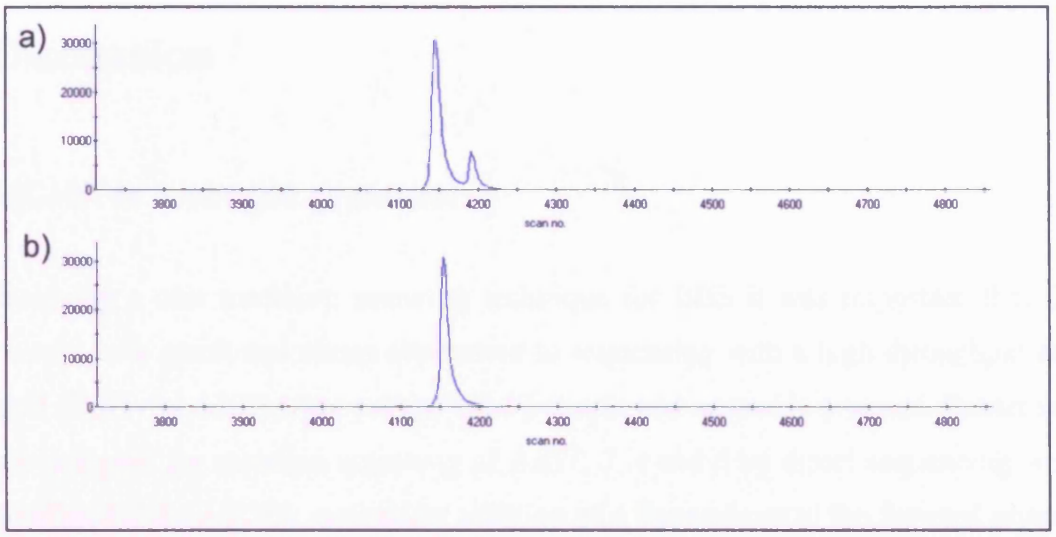


Figure 6. 5: MCHA is able to detect novel alterations. Peak profile of a novel T>G substitution (L349W) in *BBS2*x9, a 308bp FAM labelled PCR fragment. a) Sample from a heterozygous patient, b) a wt homoduplex.

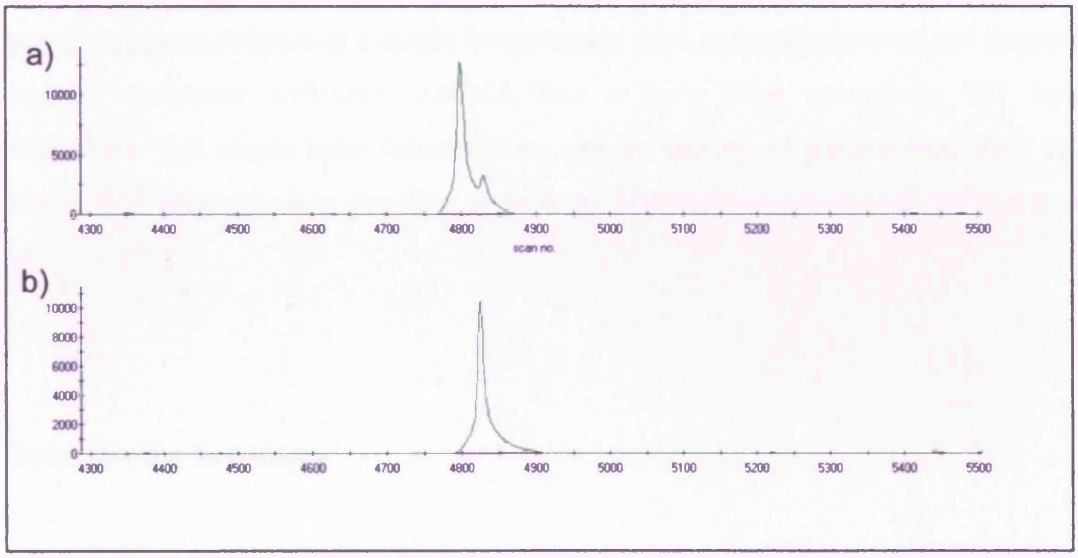


Figure 6. 6: A novel C>T intronic substitution in *BBS4*x16, a 429bp TET labelled fragment. a) Heterozygous patient sample, b) wt sample.

6.4 Discussion

6.4.1 MCHA is a simple protocol

When designing a new mutation scanning technique for BBS it was important that the method would be a quick and cheap alternative to sequencing with a high throughput and comparable sensitivity. MCHA has proved to be a simple and adaptable protocol. Primer sets previously designed for mutation screening of *BBS1*, *2*, *4* and *6* by direct sequencing were easily modified for the MCHA method by addition of a fluorophore to the forward primer. For each amplicon the colour of the fluorophore was selected based on the size of the fragment for optimal multiplexing. The ability to multiplex several different fragments in each capillary has greatly increased the throughput of HA on a capillary sequencer (Figure 6. 4). The genotyping protocol on the MegaBACE 1000 instrument was adapted by lowering of the run temperature to 25°C and use of a 3% non-denaturing LPA matrix with standard LPA buffer. Analysis of results by comparison of the patient electropherogram with a control DNA electropherogram containing a single homoduplex peak is straightforward and requires no additional computer software. MCHA has a very high sensitivity for both insertion/deletions and single base substitutions and by mixing of patient-amplified and control DNA PCR products it is possible to pick up both heterozygous and homozygous mutations.

6.4.2 Sensitivity testing

Conventional HA is known to have a high sensitivity for insertion/deletion mutations owing to the large effect that these alterations have on the mobility of a fragment through a gel matrix. The level of sensitivity for detecting single base substitutions is thought to be lower and influenced by a number of different factors relating to the fragment in which the mismatch is located such as the fragment length, the GC content and the nature of the mismatch. To test the sensitivity of the MCHA method, known alterations in five *BBS* genes (*BBS1*, *2*, *4*, *6* and *7*) including missense, nonsense and small insertions/deletions mutations,

and SNPs were used. Eleven of the 12 alterations (three deletions and eight base substitutions) tested were successfully detected in both heterozygous and homozygous individuals under standard run conditions using the initial primer set. The single alteration that was not detected was the M390R missense mutation, a T>G transversion, in *BBS1*x12. On analysis of the fragment it was not immediately clear why this alteration was not detected; the length and GC content of the fragment and the position of the mismatch within the fragment were all within the range in which mismatches had been detected in other fragments. A T>G transversion had also been successfully detected in another fragment (a further T>G mutation was also detected as a novel mutation during the screen of new BBS patients), suggesting that the nature of the mismatch was not the reason for non-detection. As modification of the run conditions by increasing the run temperature to 30°C still did not allow detection of this change, the forward primer of the pair was redesigned. Performing MCHA on PCR products generated from the new primer pair allowed detection of the mutation in all patients tested (Figure 6. 3); an overall sensitivity for the technique approaching 100%. The new forward primer had a modest effect on the properties of the *BBS1*x12 fragment; the greatest difference, an increased length by 92bp. In addition the GC content decreased slightly from 54.5% to 52.5% and the position of the mismatch within the fragment was changed from 0.59 to 0.67 (5' to 3').

6.4.3 High Specificity

In addition to a high sensitivity, MCHA also has a low false positive and negative rate. Situations that may lead to a false positive result include excess salt in the sample or a peak intensity outside the optimal range. Prior to the 'Matrix Fill and Prerun' step on the MegaBACE 1000, the 'Preinject Samples' protocol is performed to remove excess salt from the samples. If this step is omitted high salt levels can generate spurious peaks in the patient electropherogram. In my judgement, the optimal peak intensity for any sample is between 5,000 and 60,000U. Peaks with a low intensity resulting from weak PCR products often have a ragged appearance and a very high intensity can lead to splitting of the tip of the peak, both of which could result in a false positive result. Mixing of patient and control DNA PCR

products is an effective way of generating heteroduplexes in the case of a homozygous mutation but this step is also a source of potential false negative results. It is important to ensure that both the patient and control PCR products are of equal intensity, as failure or poor yield of either PCR will mean no heteroduplexes will be formed and mutations will be missed. To test the true false negative rate further known alterations must be tested.

6.4.4 Factors that influence sensitivity and peak shape

Several studies have tried to establish which mismatches are most easily detected by HA and what factors have the greatest effect on the ability to detect a mismatch. Ganguly (2002) studied a group of 12 fragments of identical size each containing a mismatch in the centre of the fragment and found that the different mismatches showed differing degrees of retardation when subjected to slab-gel HA, but that there was no relationship between the nature of the mismatch and the level of retardation. In a larger study, Highsmith *et al.*(1999a) generated a DNA toolbox consisting of several different fragments, each containing a specific mismatch, and varying in length, GC content and position of the mismatch. Each fragment was subjected to slab-gel HA, with ease of detection determined by the degree of separation between homoduplex and heteroduplex species. Fragment length (to 600bp), position of the mismatch and GC content were all found to have no effect on the ability to detect a mismatch. The nature of the mismatch was the main factor that influenced the degree of separation between duplexes, the separation order of the mismatches being: G:G/C:C > A:C/T:G = A:G/T:C > A:A/T:T.

The properties of all fragments in which alterations were detected using MCHA, either known mutations and SNPs (highlighted in bold) or novel alterations, are shown in Table 6. 2.

6.4.4.1 Fragment length

In our samples the fragment length appears to have little effect on the ability to detect mismatches, consistent with the findings of Highsmith *et al.* (1999a). The length of the fragments ranges from 239 to 492bp with little or no reduction in sensitivity at either end of the spectrum. In the longest fragment, *BBS6*x3a, two transitions were detected, an A>G at base 209 and a C>T at base 260. Although detected, these substitutions had very subtle effects on the shape of the *BBS6*x3a peak and were the hardest alterations to detect. This may be due to the effect of the bending of the DNA around the mismatch being obscured by the inherent flexibility of longer DNA fragments (Ganguly, 2002): *BBS6*x3c, a 392bp fragment of a similar GC content to *BBS6*x3a (GC content of 39.5% and 39.6% respectively), also contains a C>T transition which causes greater separation between the duplex peaks and hence easier detection (Figure 6. 4). Alternatively, the subtleness of the change to the peak could be a feature of the individual *BBS6*x3a fragment as all substitutions in the *BBS4*x15 fragment (a G>C transversion and, A>G and A>C transitions), only 19bp shorter in length, were easily detected.

6.4.4.2 Position of the mismatch

The position of the mismatch within the fragments ranges from 0.23 to 0.79 (as calculated from the 5' end of the forward primer). Slab-gel HA has a low sensitivity for mismatches within 50bp of the ends of a fragment (Rozycka et al., 2000); using MCHA it was possible to detect a C>A transversion located 50bp from the 3' end of the *BBS2*x5 fragment (240bp). When creating a set of oligos for mutation screening of a new *BBS* gene the primers are positioned such that there will be 50bp or more of intronic sequence flanking the exon on either side, or, where necessary, large exons will be divided into several overlapping fragments. Designing primers in this way ensures that no potentially pathogenic exonic alterations will be missed by a reduction in sensitivity in the final 50bp of a fragment.

6.4.4.3 GC content

The GC content of fragments containing single base substitutions ranges from 34.9% to 64.3% (the *BBS7*x7 fragment has a lower GC content, 27.0%, but contains a 4bp deletion). As with other factors, there does not appear to be a direct link between GC content of the fragment and ease of detection of a mismatch. The local GC content surrounding the alteration (calculated from the 10bp either side of the mismatch and shown in Table 6. 2) is likely to have a larger effect than the overall GC content. Ganguly (2002) observed a better resolution between duplexes when mismatches were located in an AT-rich sequence context and in earlier studies was unable to detect three substitutions in a GC-rich area of the *COL3A1* gene (Ganguly et al., 1993). There does appear to be a slight correlation between local GC content and ease of detection in our fragments: some of the alterations with the highest local GC contents (R23G and P39P in *BBS6*x3a and the intronic SNP in the *BBS1*x17 amplicon) do show only slight separation between duplexes, but others with a comparable local GC content show good separation. In addition the local GC content is the only factor that did not change on redesigning of primers for the *BBS1*x12 fragment, suggesting that it is not the only factor to affect the detection ability of a particular substitution.

6.4.4.4 Nature of the mismatch

The presence of anng a A:T match, a mutant homoduplex containing a G:C match and two heteroduplexes, one containing an A:C mismatch and the other a G:T mismatch will be formed (See Figure 6. 1). Of the 12 possible base substitutions, nine were detected in our fragments with varied frequencies (see Table 6. 3). It is likely that the three substitutions that were not detected (C>G, A>T and G>T) were simply not present in any of the fragments as the heteroduplex species that form as a result of these changes were successfully detected for other substitutions. For example, A:G and C:T heteroduplexes are created in the presence of four different substitutions: A>C, C>A, T>G and G>T. Seven base substitutions were detected due to the differential migration of G:A and C:T heteroduplexes in various sequence contexts and fragment lengths (see Table 6. 2) suggesting that had a G>T substitution been present in any of the fragments it would have been successfully detected.

Original nucleotide	Number of times substituted by				
	T	C	A	G	Total
T	-	**1**	**1**	**3**	5
C	6 (4)	-	**3**	**0**	9
A	**0**	**1**	-	7	8
G	**0**	**2**	1 (1)	-	3
Total	5	5	5	10	25

Table 6. 3: Table showing frequency of each substitution in studied fragments. Substitution numbers in bold represent transversions, those in plain text, transitions. Numbers in brackets correspond to CG>TG or CG>CA transitions resulting from deamination of 5-methylcytosine (5mC).

Of the 25 base substitutions studied the majority (60%) were transitions, consistent with the proportion of transitions present in an analysis of all mutations recorded in the Human Gene Mutation Database (HGMD) in 1997 (Krawczak et al., 1998). One reason for the increased occurrence of transitions in any set of alterations is the relatively high incidence of CG>TG and CG>CA substitutions caused by deamination of methylcytosine (5mC). In eukaryotic genomes 5mC occurs primarily within CpG dinucleotides and is susceptible to undergo deamination to form thymine leading to CG>TG transitions. CG>CA substitutions occur if, following deamination of 5mC in the antisense strand, there is miscorrection of G to A in the sense strand. The frequency of mutations at CpG dinucleotides varies between genes, in the *ADA* gene the rate is as high as 50% whereas in the *β-globin* and *HPRT* genes it is as low as 10% (Antonarakis et al., 2001). From analysis of the HGMD, Krawczak *et al.* (1998) found these mutations to account for 23.0% of all substitutions and 36.9% of transitions, consistent with this, 20% of substitutions and 33.3% of transitions within our fragments were accounted for by deamination of 5mC. Of the remaining transitions the majority were A>G substitutions (70%), with five of the seven substituted adenines preceded by another adenine. This high incidence of AA>AG transitions within our fragments may be as a result of a higher sensitivity of MCHA for detection of A>G substitutions, which, by coincidence, were most commonly preceded by adenine, or may be due to a more complex mutation mechanism specific to the *BBS* gene family.

6.4.5 Summary

Analysis of the alterations detected using MCHA shows that the technique has a high sensitivity for both insertion/deletions and base substitutions by differential migration of each of the four types of heteroduplex formed in the presence of a substitution (A:A/T:T, C:C/G:G, A:C/T:G and G:A/C:T). Rather than being determined by a single factor it is likely that the ability to detect a mismatch, and the shape of the peak resulting from such a mismatch, will be influenced by a variety of factors including fragment length, position and nature of the mismatch and the GC content of both the fragment and the area directly surrounding the mismatch. For this reason although it is not possible to predict the nature of a mismatch from the shape of the peak, individual mismatches do have characteristic shapes (Figures 6.6 and 6.7). MCHA also has a low false negative and positive rate and a quick and simple protocol making it an ideal mutation screening method for mutations in *BBS* genes.

Chapter 7 Functional analysis of BBS4

7.1 Introduction

The first steps in the functional analysis of a novel gene are often a study of the nucleotide or polypeptide sequence using database searches. Homology of all or part of the sequence to a protein of known function often provides initial clues about the functional properties of the protein. Motif searches can also be used to identify short sequences that indicate specific functions carried out by the protein. If no similarity to known polypeptides is found, it is also possible to ascertain a potential function for the protein through identification of its interactors by protein-protein interaction studies such as the yeast two-hybrid system.

7.1.1 The yeast two-hybrid system

The yeast two-hybrid system is a commonly used method to study protein-protein interactions in eukaryotes (Chien et al., 1991; Fields and Song, 1989). The system can be used either to test a direct interaction between two known proteins or to identify novel interactors of a protein of interest through a library screen. The method utilises properties of the GAL4 protein, a yeast transcriptional activator with a separate DNA binding (DBD) and activation domain (AD). The DBD binds to specific upstream activating sequences (UASs) within yeast promoters and transcription is activated by the acidic regions of the AD. In practice, to identify novel interactors, two plasmids are constructed that both encode hybrid proteins. The first consists of the DBD of GAL4 (residues 1-147), fused to the protein of interest (the bait), and the second, of the GAL4 AD (residues 768-881) fused to a protein (the prey) encoded by a clone from a library of cDNA fragments. Both constructs are introduced into a *Saccharomyces cerevisiae* yeast strain containing a reporter gene under the control of a GAL4-inducible promoter with a phenotype that is easy to score. If there is an interaction

between the bait and prey proteins, the two domains of the GAL4 protein will be in close proximity to one another, resulting in transcription of the reporter gene (Figure 7. 1). The system was initially tested using two yeast proteins, SNF1 and SNF4, which were known to interact. High levels of transcription of the reporter gene occurred only when both hybrid proteins were introduced into yeast (Fields and Song, 1989).

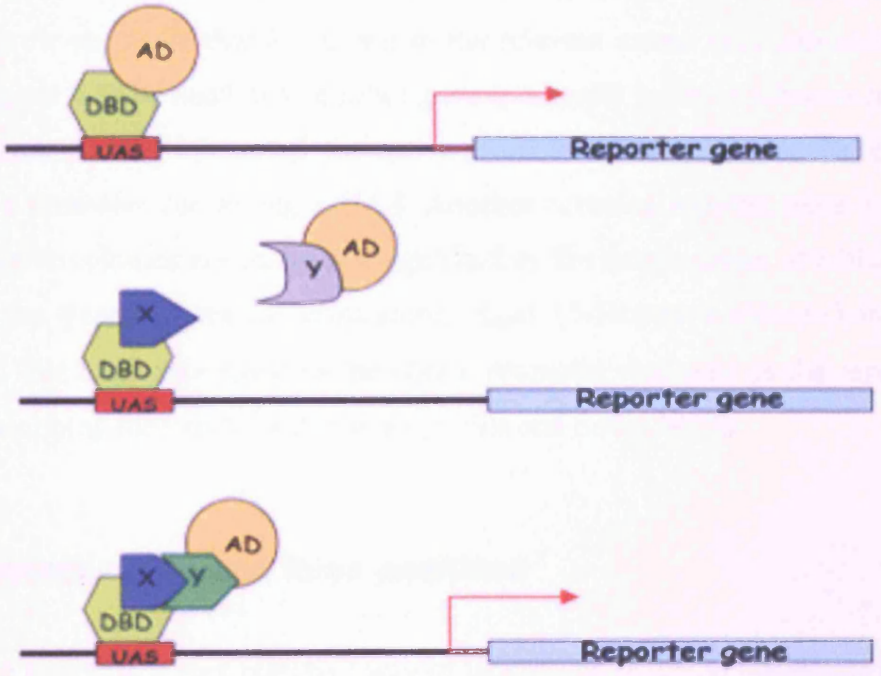


Figure 7. 1: The yeast two-hybrid system. a) The GAL4 protein is a yeast transcriptional activator with separate DNA binding (DBD) and activation (AD) domains. b) and c) Binding of the GAL4 DBD-bait hybrid protein alone is not sufficient to activate transcription of the reporter gene (b), it is only when there is an interaction between the bait and prey proteins (c) that the reporter gene will be transcribed.

The system is technically very simple to carry out, but a specially modified yeast strain is required. Firstly, the yeast chromosomal copies of the *GAL4* and *GAL80* genes must be deleted. Endogenous expression of *GAL4* would result in constitutive expression of the reporter gene regardless of whether there is an interaction between the bait and prey proteins; as GAL80 is a repressor of GAL4, the gene must be deleted to allow transcription of the reporter gene(s). The yeast must also be deleted or mutated for any endogenous copies of genes that are to be used for selection of the bait plasmid, prey plasmid or positive colonies in which a protein-protein interaction has occurred. Genes encoding enzymes required for biosynthesis of amino acids such as *ADE2*, *HIS3* and *URA3* are commonly used for selection; the ability to grow on media that is lacking in the relevant amino acid indicates a positive colony. If the gene is to be used as a reporter gene to identify positive colonies in which the bait and prey proteins have interacted, the native promoter of the gene must be removed and replaced with a promoter containing a UAS. Another common reporter gene is the *E. coli lacZ* gene; positive colonies are easily distinguished by the development of a blue colour on culturing of the yeast in media containing Xgal (5-Bromo-4-chloro-3-indolyl-β-D-galactosidase). The *lacZ* gene fused to the *GAL1* promoter was used as the reporter in the initial study describing the two-hybrid system (Fields and Song, 1989).

7.1.1.1 Autoactivation and false positives

The two-hybrid system is a very effective way of identifying potential interactors of a protein of interest, but there are a number of problems that can occur when using the system, and any interaction detected through a library screen should be confirmed by other means such as coprecipitation or GST pulldown. Before embarking on a library screen, it is essential to test the ability of the bait protein alone to activate the GAL4 system. Approximately 5-10% of proteins, particularly those containing a high proportion of acidic residues, are able to induce transcriptional activation when fused to a DBD (Toby and Golemis, 2001). If autoactivation of the system does occur, it is necessary to map the autoactivation domain of the bait protein and use fragments of the protein that do not contain this region to perform the screen. Another problem is the high frequency of false positives that are often detected. False positives are colonies in which the reporter gene is transcribed even when the bait and prey

do not themselves interact. To overcome the problem of interaction between a prey protein and the promoter of the reporter gene, it is possible to use a number of different selective markers, each under the control of a different GAL4–inducible promoter. Lists of commonly occurring false positives that have been detected in screens using a variety of different proteins are available online (www.fccc.edu/research/labs/golemis/InteractionTrapInWork.html).

7.1.2 BBS4 function

As the polypeptide sequence did not show homology to any known proteins, very little was known about the function of the protein at the time of cloning of the *BBS4* gene (Mykytyn et al., 2001). Analysis of the protein using the SMART protein domain database identified a number of tandem tetratricopeptide repeat (TPR) domains within the protein (Kim et al., 2004; Schultz et al., 2000, Appendix 3). TPR domains typically consist of 34 amino acids, have been found in a number of organisms (bacteria, fungi, plants and humans) and perform a variety of functions including protein-protein interactions (Blatch and Lassle, 1999).

7.1.3 Conclusions

As very little can be deduced about the function of the BBS4 protein by analysis of the polypeptide sequence, it is therefore necessary to perform functional studies to deduce a putative function for the protein. As the protein is known to consist of a number of TPR domains, it is highly likely that BBS4 is involved in protein-protein interactions of some form. A yeast two-hybrid library screen was performed, using BBS4 as the bait, to identify interactors of the protein.

7.2 Methods

7.2.1 pGBDU-BBS4 bait plasmid

Full-length BBS4 cDNA was cloned into the *Eco*RI and *Bam*HI sites of the multiple cloning site (MCS) of pGBDU by Dr. Alison Ross (James et al., 1996). The pGBDU plasmid contains an Ampicillin resistance gene (Amp^r) for selection in bacteria and a nutritional (*URA3*) marker for selection in yeast.

7.2.2 Library screen

The library screen was carried out as described in *2.2.5 Yeast two-hybrid* using the PJ69-4A strain of *S. cerevisiae* (James et al., 1996), the BBS4 protein as the bait and the Human kidney MATCHMAKER cDNA library as the prey (Clontech). The transfection efficiency was tested, as described in *2.2.5.3 Transfection efficiency*, to ensure that complete coverage of the library was achieved.

7.2.3 Identification of positive clones

Prey plasmids were extracted from positive clones as described in *2.2.5.6 Isolation of prey plasmid*. The insert was amplified using pACT2 primers and the PCR product sequenced as described in *2.2.5.7 Amplification and identification of the prey cDNA sequence*. The identity of the clone was determined using BLAST (www.ncbi.nlm.nih.gov/BLAST/Blast.cgi). Following identification of the clone, the correct reading frame was determined to ensure that the cDNA was in frame with the GAL4 AD within the pACT2 vector.

7.3 Results

7.3.1 Transfection efficiency

The transfection efficiency of the library plasmid was calculated to be $4.8x10^7$, indicating that complete coverage of the ~$3.5x10^6$ independent clones from the cDNA library was likely to be achieved.

7.3.2 First round of selection

Following the introduction of both the bait and prey plasmids into the yeast, cells were plated on SD-Ura-Leu-His+3AT agar for the first round of selection. If an interaction between the bait (BBS4) and a clone from the cDNA library (the prey) had occurred, the *GAL1-HIS3* reporter gene was transcribed and the yeast able to survive on the media, in the absence of histidine. After 4-5 days growth at 30°C, 300 colonies were present on the plates.

7.3.3 Second round of selection

After the initial selection for positives in which a protein-protein interaction had occurred, the colonies were transferred to SD-Ura-Leu-Ade plates for the second round of selection. Following a 4-5 day incubation at 30°C, 13 of the 300 colonies, had failed to grow in the absence of adenine. A further 14 colonies were able to grow but developed a red colour, indicating that they were false positives, able to synthesize adenine through an alternative pathway. In order to identify the interacting clones, the prey plasmid was isolated from the 273 positive colonies, the insert amplified by PCR and sequenced.

7.3.4 Identity of positive clones

Although 273 positive clones were found through the library screen, it was not possible to identify all clones; only 62 clones were successfully amplified and sequenced. Of these 62 clones, 40 were in frame with the GAL4 AD polypeptide and therefore presumed to generate a functional polypeptide. See Table 7. 1 for a summary of the clones that were identified.

	Interactor	Length of canonical ORF (Amino acids)	Amino acids present in clone	Number of copies
1	Pericentriolar material 1 (PCM1)	2024	1845 – 2024	3
2	Pericentriolar material 1 (PCM1)	2024	1967 – 2024	2
3	Nuclear receptor interacting protein 1 (NRIP1)	1158	442 – ?	1
4	Nuclear receptor interacting protein 1 (NRIP1)	1158	547 – 716	1
5	Nuclear receptor interacting protein 1 (NRIP1)	1158	647 – 953	3
6	Nuclear receptor interacting protein 1 (NRIP1)	1158	648 – 810	3
7	SMRT/HDAC1 Associated repressor protein (SHARP)	3651	1104 – ?	1
8	SMRT/HDAC1 Associated repressor protein (SHARP)	3651	1148 – ?	1
9	SMRT/HDAC1 Associated repressor protein (SHARP)	3651	1243 – ?	1
10	Betaine-homocysteine methyltransferase (BHMT)	406	230 – 384	1
11	Betaine-homocysteine methyltransferase (BHMT)	406	237 – ?	1
12	Betaine-homocysteine methyltransferase (BHMT)	406	252 – 406	1
13	SKI-interacting protein (SKIP)	536	1 – 159	2
14	SKI-interacting protein (SKIP)	536	241 – 461	1
15	Ecotropic viral integration site 1 protein (EVI1)	1395	539 – 694	2
16	Ecotropic viral integration site 1 protein (EVI1)	1395	614 – ?	1
17	Nuclear receptor coactivator 4 (NCoA4)	615	231 – ?	1
18	Nuclear receptor coactivator 4 (NCoA4)	615	253 – ?	1
19	*Aldolase B fructose-bisphosphate (ALDOB)*	*364*	*108 –?*	*1*
20	*Aldolase B fructose-bisphosphate (ALDOB)*	*364*	*115 –248*	*1*
21	*DNA mismatch repair protein (MLH3)*	*1437*	*982 - ?*	*1*
22	*DNA mismatch repair protein (MLH3)*	*1437*	*1033 – ?*	*1*
23	Estrogen receptor binding protein (ERBP)	736	373 – 504	1
24	Estrogen receptor binding protein (ERBP)	736	510 - 634	1
25	SF21 protein	513	329 – 513	2
26	p150 ^{Glued}	1170	1022 – 1170	1
27	*Phosphoenolpyruvate carboxykinase 2 (mitochondrial)*	*640*	*478 – 640*	*1*
28	*Catalase (CAT)*	*527*	*330 – 527*	*1*
29	Glycine decarboxylase (P-protein)	1020	883 – 1020	1
30	Alpha thalassemia/mental retardation syndrome X-linked	2492	2182 - 2310	1

Table 7. 1: Summary of BBS4 interactors identified through a yeast two-hybrid screen. Interactors in italics are predicted to be false positives. ? denotes clones in which it was not possible to determine the exact size of the clone due to poor sequence quality at the C-terminal end.

7.4 Discussion

Through a yeast two-hybrid screen of a human kidney cDNA library it was possible to identify several potential interactors of the BBS4 protein.

7.4.1 Pericentriolar material 1 (PCM1)

One of the most common interactors identified through the two-hybrid screen was pericentriolar material 1 (PCM1). Two independent clones that were in frame with the GAL4 AD were identified. Both clones were from the C-terminal portion of the protein; the larger clone of 179 amino acids (1,845 – 2,024) was detected three times, and the smaller clone of only 57 amino acids (1,967 – 2,024) was detected twice. Interestingly, a further eight PCM1 clones were also identified; analysis of these fragments revealed that they were not in frame with the GAL4 AD (ORF1), or each other (four were ORF2 and four, ORF3); in total, 21% of all clones that were successfully sequenced were found to be PCM1, but less than half were in frame. Despite appearing to be out-of-frame, it is likely that the levels of PCM1 that are transcribed from these clones by low-level ribosomal frameshifting are sufficient to rescue the yeast phenotype.

PCM1 is a 228kDa protein with a distinct cellular localization; the protein is localised to centriolar satellites throughout the cell cycle, except during the G_2/M phase when the cell is preparing for mitosis, at which point PCM1 disperses into cytoplasmic foci (Balczon et al., 1994). The concentration of the protein also fluctuates during the cell cycle, with high levels in interphase and lower levels during mitosis (Zimmerman and Doxsey, 2000). The centrosome is the largest non-membrane bound organelle in most cells and is involved in several important cellular functions including spindle function, determination of cell shape and the organization and transport of cytoplasmic organelles (Zimmerman et al., 1999). It has a complex structure consisting of a pair of centrioles surrounded by a matrix of centrosome-associated proteins such as PCM1, pericentrin, γ-tubulin and centrin making up the PCM (Andersen et al., 2003; Zimmerman and Doxsey, 2000). PCM1 is transported to the

centrosome by dynein-mediated transport through indirect binding of PCM1 to the p150^{Glued} component of dynactin. Adaptor proteins mediating the link between PCM1 and dynactin include Huntingtin-associated protein 1 (HAP1), which has been shown to interact with both PCM1 and p150^{Glued} by two-hybrid analysis and is thought to be involved in vesicle trafficking within cells (Engelender et al., 1997). Due to their association with pericentriolar satellites and increased concentration during experimental induction of cilia, PCM1 particles are also thought to be involved in the process of ciliogenesis (Kubo et al., 1999).

Following the two-hybrid screen, several lines of evidence have been found to support the interaction between BBS4 and PCM1. In a parallel yeast two-hybrid screen using BBS4 as bait to screen a fetal brain cDNA library using the CytoTrap method (Aronheim et al., 1997), two of five positive clones were found to be PCM1 (Kim et al., 2004). The clones identified in the CytoTrap screen were independent from each other, and the clones identified using the GAL4 method, encoding slightly larger fragments of PCM1 (1,574-2,024 and 1,744-2,024 amino acids). Confirmation of the interaction between BBS4 and PCM1 was provided by the coprecipitation of HA-tagged PCM1 (PCM1-HA) with Myc-tagged BBS4 (BBS4-Myc) in mammalian kidney cells (HEK293) (Kim et al., 2004). Cellular localisation studies in a number of different cultured mammalian cells also confirmed that the BBS4 and PCM1 proteins co-localise to centriolar satellites and the basal body of ciliated cells. As BBS4 was consistently found to be associated with the centrosome throughout the cell cycle, unlike PCM1, these results suggest that BBS4 localises to centriolar satellites in a PCM1-independent manner (Kim et al., 2004).

7.4.2 p150^{Glued}

The p150^{Glued} subunit of the dynactin complex was also found to be a potential interactor of BBS4. A single clone encoding a C-terminal fragment (amino acids 1,022 to 1,170) was identified.

A p150^{Glued} dimer, with nine other proteins including the actin-related protein centractin (Arp-1), forms the macromolecular dynactin complex (Waterman-Storer et al., 1995), a complex that has been found to interact with the dynein complex and be required for dynein-mediated vesicle motility along microtubules in an *in vitro* assay (Gill et al., 1991; Vaughan and Vallee, 1995). p150^{Glued} is one of the best characterised components of the dynactin complex and has been shown to bind directly to microtubules, centractin and the intermediate chains (ICs) of cytoplasmic dynein through distinct domains within the protein. The microtubule-binding domain is located at the N-terminal of the peptide (amino acids 39-150), with the centractin-binding domain defined by a highly conserved cluster of charged amino acids (KKEK), known to be involved in actin binding, towards the C-terminus (Waterman-Storer et al., 1995). The IC-interacting domain is located between amino acids 200 and 811, a region containing a predicted coiled-coiled domain. The N terminal region of ICs that is responsible for the association with p150^{Glued}, also contains a coiled-coiled domain and a serine-rich cluster, indicating that this interaction may be regulated by phosphorylation (Vaughan and Vallee, 1995). The domain of p150^{Glued} mediating the interaction with BBS4 is C-terminal to all of these known interaction domains.

The coprecipitation of GFP-tagged p150^{Glued} with Myc-tagged BBS4 in HEK293 cells and the mislocalisation of BBS4 in cells overexpressing Myc-tagged p50-dynamitin, an antagonist of dynactin, confirmed the interaction between p150^{Glued} and BBS4. These results also indicate that the dynein-dynactin machinery is responsible for transporting BBS4 to the PCM and that the BBS4 protein may function as an adaptor protein to transport PCM1 to its appropriate cellular location (Kim et al., 2004).

7.4.3 Nuclear receptor interacting proteins

In addition to PCM1 and p150^{Glued}, BBS4 was also found to interact with a number of nuclear receptor interacting proteins.

7.4.3.1 Nuclear receptor interacting protein 1 (NRIP1)

A total of eight clones of nuclear receptor interacting protein 1 (NRIP1, also known as RIP140) were identified in the two-hybrid screen. Two clones of the 1,158 amino acid polypeptide, one beginning at codon 442 and the second at codon 547, were present in single copies. The remaining two clones, beginning at amino acids 647 and 648, were each identified three times.

NRIP1 was first identified as an estrogen receptor (ER) interacting protein (Cavailles et al., 1995), and has since been shown to interact with a variety of nuclear receptors, including the retinoic acid receptor (RAR), through different interaction domains (Wei et al., 2001). RAR is a member of the nuclear hormone receptor (NHR) superfamily and, like many other members of this group, forms heterodimers with the retinoid X receptor (RXR). The interaction between NRIP1 and RAR:RXR dimers is mediated through a motif present in the C-terminal region of NRIP1. Rather than activate expression of target genes, the strong ligand–dependent interaction between NRIP1 and RAR:RXR results in suppression of RA-regulated genes (Lee and Wei, 1999). A potential mechanism for this suppressive role of NRIP1 is through interaction of the N-terminal region of NRIP1 (amino acids 78-303) with components of the histone deacetylase complex (HDAC). NRIP1 is able to interact directly with HDAC1 and HDAC3, resulting in the recruitment of HDACs to RA-responsive promoters by the NRIP1:RAR:RXR complex, and subsequent suppression of transcription of target genes (Wei et al., 2000).

Nrip1 is widely expressed in tissues and cells (Lee et al., 1998) and has been shown to be essential for female fertility (White et al., 2000). The study of transgenic mice has confirmed the importance of nuclear receptors in female fertility and ovarian function. Mice lacking the gene for ERα are viable, but female mice are infertile and males have a greatly reduced fertility (Lubahn et al., 1993). ERβ knockout male mice have a normal level of fertility whilst females have fewer and smaller litters in comparison to wt mice, due to a reduced rate of ovulation (Krege et al., 1998). Homozygous Nrip1 knockout mice (RIPKO) are morphologically normal but were found on average to be 15-20% smaller than their

heterozygous and wt littermates. Mature homozygous female RIPKO mice were found to be completely infertile whereas male RIPKO mice were able to sire offspring. Histological analysis of the ovaries from RIPKO mice identified the cause of the infertility to be a failure in the release of oocytes at ovulation. This defect in ovulation in RIPKO mice suggests that Nrip1 plays an important part in the spatial and temporal activation of nuclear receptors required in ovarian cells for ovulation to occur (White et al., 2000).

7.4.3.2 SMRT/HDAC1 associated repressor protein (SHARP)

Three independent clones of SMRT/HDAC1 associated repressor protein (SHARP) were identified, beginning at amino acids 1,104, 1,148 and 1,243 of the 3,651 amino acid protein. Each clone was identified once.

SHARP was initially cloned through a yeast two-hybrid screen using the C-terminal region of the nuclear corepressor SMRT (silencing mediator of retinoic acid and thyroid hormone receptor) (Shi et al., 2001). Like SMRT, and the related nuclear corepressor N-CoR, SHARP is able to bind to the ligand-binding domain of RAR in the absence of its ligand, all trans retinoic acid (ATRA), and recruit components of the HDAC, bringing about repression of RA-inducible genes. Binding of ATRA to RAR results in a change in conformation, disrupting this association and allowing transcription to occur (Ordentlich et al., 1999). In addition to acting as a transcriptional repressor, SHARP can also act as an activator, through interaction with the steroid receptor RNA activator (SRA). SRA is an RNA coactivator that binds to SHARP through three RNA recognition motifs (RRMs) (Shi et al., 2001). Interaction between SHARP and SRA attenuates the hormone response from the ER by a combination of sequestration of SRA and recruitment of SMRT (Lanz et al., 2002).

7.4.3.3 Ski-interacting protein (SKIP)

Three clones, two beginning at the start methionine and a third beginning at codon 241, of Ski-interacting protein (SKIP) were found to interact with BBS4.

SKIP is a bifunctional nuclear receptor coregulator that interacts with components of a number of key signalling pathways, indicating that it may play an important role in development and oncogenesis. SKIP was identified through its ability to interact with the oncoprotein Ski (Dahl et al., 1998) and is involved in vitamin D-mediated transcription (Barry et al., 2003; Baudino et al., 1998). Like RAR, the vitamin D receptor (VDR) is a member of the NHR superfamily and forms heterodimers with RXR. SKIP is able to bind VDR:RXR heterodimers in a ligand-enhanced manner and, acting synergistically with the coactivator GRIP1, augment VDR-mediated transcription. In the absence of ligand (1,25-dihydroxyvitamin D_3 ($1,25(OH)_2D_3$)), both SKIP and VDR interact with the corepressors N-CoR and SMRT, thereby repressing VDR-mediated transcription (Barry et al., 2003). In addition to its bifunctional role in NHR-dependent transcription, SKIP is also known to be a coregulator of the Notch signalling pathway. CBF1 (a DNA binding protein) binds promoters containing the sequence GTGGGAA and also interacts with SKIP. NotchIC competes with an SMRT repressor complex including Sin3A, HDAC1 and HDAC2 for binding sites on CBF1 and SKIP; when the SMRT repressor complex is bound, transcription from the corresponding promoter will be repressed, whereas when NotchIC is bound, transcription will be active (Zhou et al., 2000).

7.4.4 Problems and further work

Although 273 positives were present after the second round of selection, only 22.7% (62/273) of interacting clones were successfully identified by sequencing. Of the 62 clones that were successfully sequenced, 40 (65%) were found to be in frame with the GAL4 AD. If a similar percentage of the unidentified clones were found to be in frame, 137 additional clones would be expected to generate an active peptide. Although a number of these clones are likely to be replicates of the interactors already identified, it is possible that, due to the low rate of identification, a number of potentially interesting novel interactors of BBS4 have been missed. As a large number of clones failed at the PCR amplification stage, prior to sequencing, it is possible that the size of some of the inserts were too large to be amplified by

PCR. As the Human Kidney MATCHMAKER cDNA library (Clontech) contains clones ranging in size from 0.5 – 4.0kb, a large number of these will be too large to amplify by conventional PCR. Sequencing drectly from the prey plasmid would eliminate this problem and would allow identification of a greater number of potential interactors.

Repeating the screen, possibly in an alternative library, would be useful for both the identification of novel interactors and also to confirm the interaction between BBS4 and some of the potential interactors identified through this screen; if an interactor is identified in two separate screens carried out in different prey libraries it is much more likely that it is a true interactor of the bait. Direct interaction studies between BBS4 and some of the commonly occurring interactors such as NRIP1, SHARP and SKIP are required to confirm whether these interactions are real and whether BBS4 may interact with a subgroup of nuclear receptor interacting proteins in addition to its confirmed interactions with both PCM1 and p150^{Glued}.

7.4.5 Summary

By performing a yeast two hybrid screen it has been possible to establish a putative function for the BBS4 protein through identification of proteins with which it interacts. Interaction with both PCM1 and p150^{Glued} suggest that BBS4 is involved in the transport of PCM1 to centriolar satellites of cells by the dynein-associated machinery. The colocalisation of BBS4 with PCM1 at centrosomes and at the basal bodies of ciliated cells indicates that BBS4 may also be involved in the process of ciliogenesis. The identification of a number of nuclear receptor interacting proteins as potential BBS4 interactors is also interesting, but requires further work to establish if these interactions are real and how BBS4 may contribute to the activating and/or repressive effects of nuclear receptors.

Chapter 8 Expression analysis of *BBS4* and *BBS8* in mouse tissues

8.1 Introduction

Expression analysis of a novel gene at both the cellular and tissue level is an important step in the characterisation of the gene and its corresponding protein product. Expression of a gene in a particular organ can be demonstrated by reverse transcriptase PCR (RT-PCR) on total RNA extracted from the tissue. These data however, only give an indication of the gross expression of a transcript in the organ or tissue and is therefore unable to provide information about individual cell populations in which the gene is expressed within the organ. To obtain detailed data on the spatial expression pattern of the gene, analysis of RNA or protein levels within the tissue of interest is required.

Analysis of RNA in tissues by *in situ* hybridisation (ISH) can be used to determine levels of gene expression within a tissue sample and to establish the specific population of cells in which the gene is transcribed. In practice, RNA probes (riboprobes) ranging in length from a few bases to 1-2kb are synthesized from a vector using RNA polymerase and labelled with radioactivity or fluorescence. The probe is then incubated on the tissue sample, allowing the probe to anneal to its complementary DNA target sequence (Harvey, 2001). Although ISH can be a very useful tool for studying gene expression within tissues, it does also have its disadvantages. RNA suffers from an inherent instability due to the presence of RNases in the environment. Despite good laboratory practice, and the use of certified RNA-free solutions and equipment, RNases can be very difficult to control and inactivate, making the riboprobe susceptible to degradation. An alternative to ISH for the analysis of gene expression in tissue samples is to use antibodies to observe the localisation of the protein (the antigen) within tissues by immunohistochemistry.

8.1.1 Antibody structure and production

Antibodies consist of four polypeptide chains, two light chains and two heavy chains, held together by covalent bonds and, depending on the type of heavy chain, can be divided into five classes; IgA, IgD, IgE, IgG and IgM. Both light and heavy chains consist of a constant C-terminal region (~110 amino acids long in the light chain and 330-440 amino acids in the heavy chain) and a variable N-terminal region (~110 amino acids in both chains). Each antibody contains two antigen-binding sites, made up of the variable regions of both the light and heavy chains (Figure 8. 1), allowing for a high degree of diversity in antigen-binding sites (Alberts et al., 1994). The hinge region provides the antibody with flexibility, allowing the distance between the binding sites to vary.

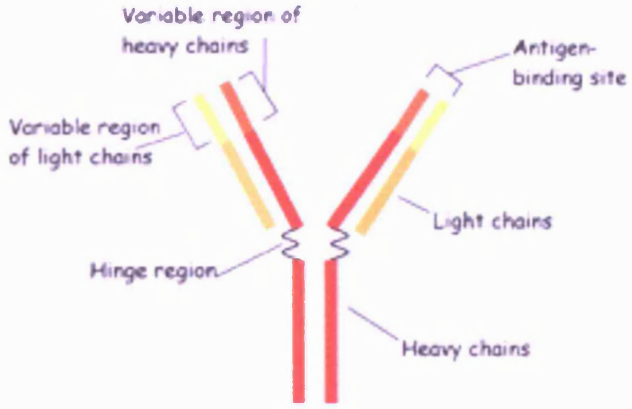


Figure 8. 1: Schematic showing the main characteristics of an antibody.

Antibodies to human proteins are routinely produced by repeated injection of an animal, usually a rabbit, with a suitable immunogen and subsequent recovery of antibodies from the serum. Commonly used immunogens include synthetic peptides of 20-50 amino acids in length or fusion proteins generated from a short N-terminal bacterial sequence and the mRNA for the protein of interest. Both techniques have advantages and disadvantages. The synthetic peptide technique is very simple but the success rate in producing a functional, specific antibody is hard to predict as the short peptide may not fold correctly and the amino

acid sequence to be used in the peptide must be carefully selected to ensure that the antibody will not cross-react with other related proteins. As a fusion protein includes the majority of the polypeptide, the chance of producing a specific antibody is much higher. This method however, involves more work and, although short, the bacterial sequence may interfere with the study protein (Strachan and Read, 1999).

Although IgG is the most abundant class of antibody in the blood during the secondary immune response, IgM is the first detectable antibody in the serum following immunisation of the animal. The period from injection of the immunogen to production of the first IgM in the blood is referred to as the latent period and usually lasts approximately one week. After a second injection, IgG will begin to predominate in the serum and, although IgG has a longer half-life (approximately three weeks compared to five days for IgM), regular injections of immunogen are required to maintain high production levels of the antibody (Boenisch, 2001a). Following bleeding of the animal, polyclonal antibodies, raised to a number of different epitopes on the antigen, can be isolated in the form of stabilised antisera. The immunoglobulins can then be purified as required; salt precipitation and ion exchange chromatography can be used to remove additional proteins from the serum followed by affinity chromatography to isolate antibodies specific to the immunogen, reducing the risk of cross-reactivity with other antigens.

8.1.2 Immunohistochemistry staining methods

Initial immunohistochemistry experiments used a direct method of staining where the antibody to the protein of interest (the primary antibody) is labelled (Figure 8. 2a). Coon *et al.* (1951) developed the immunofluorescent method by the addition of a fluorescent label to the primary antibody allowing visualisation of the antibody using a fluorescent microscope. An alternative technique developed by Nakane and Pierce (1967) uses enzyme-substrate reactions to convert colourless chromogens into coloured end products. Immunologically and enzymatically active conjugates are generated and used to localise the antigen within the tissue. A good enzyme for immunoenzymatic staining must be stable in solution, maintain its

activity when conjugated to an antibody or protein and have an end product that is both stable and easily detectable (Boenisch, 2001c). Two enzymes that fulfill these criteria, and are commonly used for staining, are calf intestinal alkaline phosphatase and horseradish peroxidase (HRP). HRP, isolated from the root of the horseradish plant, forms a complex with its substrate, hydrogen peroxide, causing it to decompose into water and oxygen. In the presence of an electron donor, the enzyme-substrate complex will bring about oxidation of the donor. Electron donors that, when oxidised, are converted from a colourless substance to a coloured one and become insoluble are also referred to as chromogens and are useful in immunohistochemistry. 3,3-diaminobenzidine tetrahydrochloride (DAB) produces a brown end product and is highly insoluble in organic and aqueous solvents, making it a popular chromogen.

Direct methods, although very quick and simple to perform, do not allow a very high level of sensitivity due to little amplification of the signal. The majority of methods in use today are indirect, and incorporate at least one signal amplification step. In the simplest indirect methods, an unlabelled primary antibody is applied to the sample followed by a labelled antibody to the primary antibody (the secondary antibody, Figure 8. 2b). As the secondary antibody is raised against the IgG of the animal in which the primary antibody was produced (for example anti-rabbit IgG), the antibody will bind to several epitopes on the primary antibody, providing amplification of the signal. To further amplify signals, methods such as the avidin-biotin complex (ABC) method are used (Figure 8. 2c). Avidin is a glycoprotein tetramer found in egg whites that contains four binding sites for biotin, a coenzyme of decarboxylase that is present in a number of organs. The high affinity of avidin for biotin, and the ease with which biotin can be conjugated to a range of antibodies or enzymatic and fluorescent labels, makes ABC a highly effective and adaptable method for immunohistochemical staining. Initial incubation of the primary antibody on the sample is usually followed by application of a biotinylated secondary antibody. Avidin-biotin complexes and avidin-HRP conjugates are then added to the sample, followed by the final addition of the chromogen. The ABC technique was initially developed for improved sensitivity in electron microscopy of proteins in the membranes of *Acholeplasma laidlawii* (Heitzmann and Richards, 1974) but is now routinely used in both electron and light microscopy and is available in kit form.

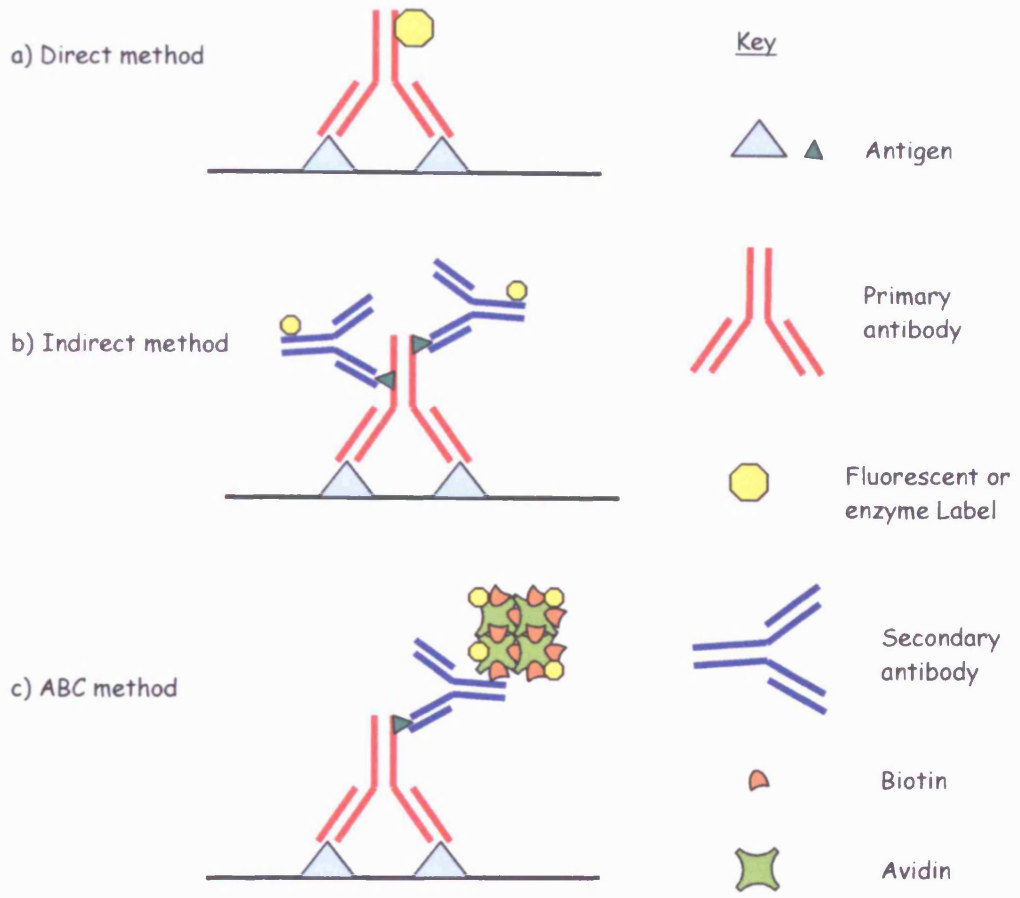


Figure 8. 2: Staining methods. a) The direct method, in which the primary antibody is labelled, b) The indirect method, in which a labelled secondary antibody is used, and c) The ABC method.

8.1.3 Sample preparation

Immunohistochemistry can be carried out on samples prepared in a variety of different ways. In most cases, formalin fixed tissues are embedded in paraffin wax and sectioned on a microtome prior to staining. For small blocks of tissue (<5mm thick) it is also possible to perform whole mount staining to allow a 3D analysis of protein localisation, but this

technique requires extensive optimisation of the staining procedure to ensure complete penetration of the sample with staining reagents. As formalin fixation and embedding can sometimes damage antigens, tissue samples may also be frozen and cryosectioned. Cryosection processing is the best way to preserve antigens but suffers from poor morphological detail and resolution, particularly at high magnifications. As an alternative it is also possible to treat a formalin-fixed sample by either enzyme digestion or by the use of microwaves to increase exposure of the antigen whilst preserving the histology.

8.1.4 Background staining

Non-specific background staining is a common problem in immunohistochemistry and can be caused by a number of factors (Boenisch, 2001b). Some of the major causes of non-specific staining include:

- *Hydrophobic interactions* – Hydrophobic interactions occur between macromolecules when their surface tensions are less than water. Hydrophobicity is a property shared by a number of different types of proteins, mediated predominantly through the neutral aromatic amino acids phenylalanine, tyrosine and tryptophan. The shared hydrophobicity of these residues causes them to link to each other, expelling water and conferring stability to the polypeptide. Tissues fixed in formalin or glutaraldehyde exhibit an increased hydrophobicity resulting in background staining in certain tissues including connective tissues, squamous epithelium and adipocytes. Immunoglobulins, particularly the subclasses IgG_1 and IgG_3, also tend to have a high hydrophobicity that increases if the antibody is not stored correctly, leading to the formation of aggregates. The cross-linking between hydrophobic tissue samples and antibody aggregates can be controlled by optimisation of fixation procedures for individual tissue samples and appropriate storage and use of the antibody. Appropriate blocking, both before and during the primary antibody incubation step, with a protein that will compete sufficiently with IgG for hydrophobic sites within the tissue, is also an effective means of reducing background. The addition of 1% BSA to

a solution containing serum from the animal in which the secondary antibody was raised is a commonly used blocking buffer.

- *Endogenous enzyme activity* – Endogenous enzyme activity can also be a cause of background staining when using enzyme-mediated detection methods. Haemoglobin, myoglobin, cytochrome and catalases all exhibit peroxidase activity with interstitial activity resulting from diffusion of the blood prior to fixation of tissue. An effective means of blocking endogenous peroxidase activity is by incubation of 3% hydrogen peroxide on the sample prior to the blocking of non-specific sites.

- *Endogenous avidin binding activity (EABA)* – The use of the ABC staining method also introduces additional forms of background. Non-specific staining can be caused by binding of avidin-HRP conjugates to endogenous biotin and also to lectin-like and negatively charged tissue components. High levels of biotin are present in the liver, kidney and lymphoid tissue, and, due to its four biotin binding sites, binding of avidin molecules to endogenous biotin increases levels of the protein in tissue by the recruitment of a further three biotin molecules with the binding of each avidin molecule. The high isoelectric point (pI) and carbohydrate composition (10%) of avidin can lead to non-specific binding of the glycoprotein by ionic interactions with tissue proteins. Artificial reduction of the pI of avidin or the substitution of avidin for streptavidin, an analogous protein found in the bacteria *Streptomyces avidinii*, in ABC staining kits has largely eliminated this type of background staining.

8.1.5 Conclusions

Immunohistochemistry can be used to perform a detailed analysis of expression patterns of a gene and its protein product within tissues, providing information about particular cell populations in which the gene is expression. Knowledge of spatial and temporal expression patterns of *BBS* genes may provide evidence for possible functions of the proteins and give

clues as to how mutations in a number of different genes can lead to the complex pleiotropic phenotype of BBS.

8.2 Methods

8.2.1 Sample preparation

Slides of mouse tissues were either made from paraffin embedded blocks of tissue (BBS4 - eyes) or purchased from Novagen (BBS8 - testis, eye, E14 embryo and olfactory epithelium). Eyes were dissected from adult wt female mice and fixed as described in *2.2.6.1 Fixation of mouse tissues*. Four different fixatives (100% methanol, 2% paraformaldehyde (PFA), 4% PFA and Carnoy's fluid) were tested to determine what effect, if any, the fixative had on the quality of the staining of BBS4. The microwave antigen retrieval step was also tested; slides from each different fixative were microwaved for 0, 4, 7 or 10 minutes in 0.01M citric acid as described in *2.2.6.4 Antigen retrieval*. Slides that were not microwaved were placed in 0.01M citric acid for 10 minutes

8.2.2 Staining protocol

BBS4 and BBS8 affinity purified antisera were used as the primary antibody in the staining experiments. The ABC method was used for signal amplification (ABComplex, DAKO), as described in *2.2.6.5 Immunohistochemistry staining*. To allow smaller quantities of antibody to be used and different antibody dilutions on sections on the same slide, a PAP pen (Ted Pella, Inc.) was used to draw a water repellent barrier around each section. As a control, one section on each slide was designated 'no primary'; in place of the primary antibody, block solution was applied to the section.

8.3 Results

8.3.1 Fixatives - BBS4

The type of fixative used (100% methanol, 2% PFA, 4% PFA and Carnoy's fluid) did not appear to have a large effect on BBS4 staining, but did affect the quality of sections produced from the tissue. 100% methanol is a mild fixative and was therefore expected to have the least damaging effect on the BBS4 antigen within tissue. During the sectioning process, tissue fixed in 100% methanol appeared to fragment, resulting in torn and damaged sections. The sections were used in staining experiments and did show *BBS4* expression, but due to the poor quality of the sections, the results were hard to interpret. Tissues fixed in PFA (2% and 4%) produced good sections and a good staining pattern. The best histology however was seen with tissues fixed in Carnoy's fluid. This may be due to the fixative alone or may also be influenced by other steps in the embedding and sectioning process. As only two eyes were tested for each fixative, a larger number of samples of different tissue types are required to establish if Carnoy's fluid is the best fixative for routine use.

8.3.2 Antigen retrieval – BBS4

In addition to different fixatives, the microwave oven antigen retrieval step was also tested. BBS4 staining was comparable in sections that had been microwaved for 0, 4 or 7 minutes. Some sections from slides that had been microwaved for 10 minutes, in particular 100% methanol sections, began to become detached from slides, suggesting that times of seven minutes or less is the optimum time for microwaving, depending on the antigen.

8.3.3 *BBS4* expression

BBS4 expression was studied in the mouse eye and was present throughout the layers of the retina (Figure 8. 3a and b). High levels of expression were also seen in the cornea (Figure 8. 3c) and the ciliary muscles (Figure 8. 3d). DAB staining was not seen in these layers in the 'no primary' control sections. Western blot analysis was not performed which would have given an indication of the antibody specificity but subsequent experiments using the same BBS4 antibody have shown a more specific pattern of expression in the outer nuclear layer and the inner segments, where BBS4 colocalses with PCM1 (Kim et al., 2004). This does not however demonstrate or confirm the specificity of the antibody.

8.3.4 *BBS8* expression

BBS8 expression was observed in ciliated structures in mouse tissues. In adults, BBS8 was present in maturing flagellated spermatids (Figure 8. 4a) and also in the connecting cilium (CC) of the retina between the inner and outer segments of photoreceptors (Figure 8. 4b). In the embryo, BBS8 was seen in the developing telencephalon (TE, Figure 8. 4c), olfactory epithelium (OE) and olfactory sensory neurons (OSN, Figure 8. 4c and d) (Ansley et al., 2003).

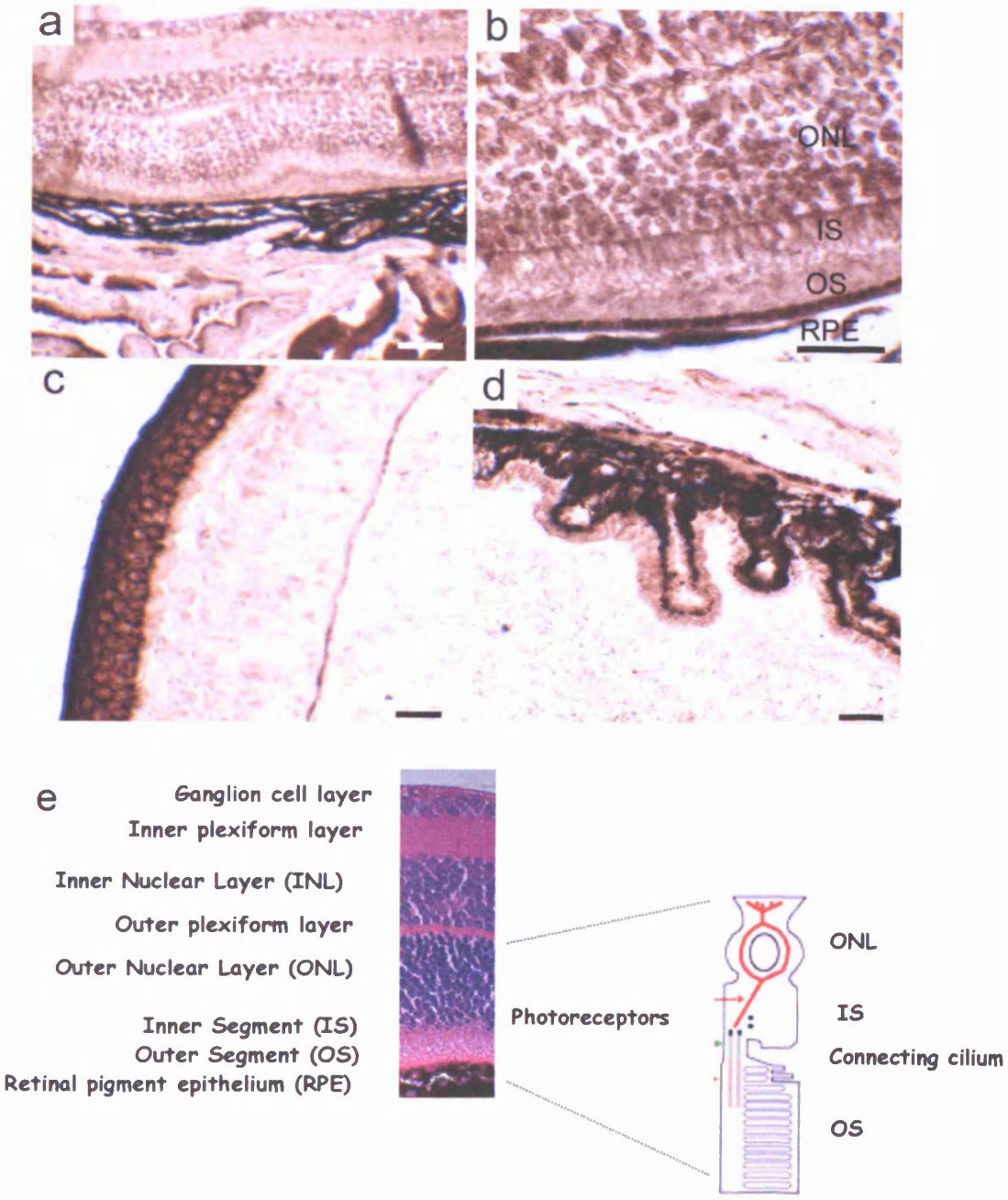


Figure 8. 3: Expression of BBS4 in the mouse eye. a) Retina, showing highest levels of expression in the inner (IS) and outer segments (OS) of the photoreceptors and the outer nuclear layer (ONL) (RPE – retinal pigment epithelium). b) Higher magnification of a BBS4 stained retina. c) High levels of BBS4 expression was also seen in the cornea and d) the ciliary muscles. All sections are from Carnoy's fluid fixed tissue with four minutes microwaving and five minutes DAB reaction. Scale bars represent 100µm in a), c) and d), and 50µm in b). e) Schematic of retinal layers.

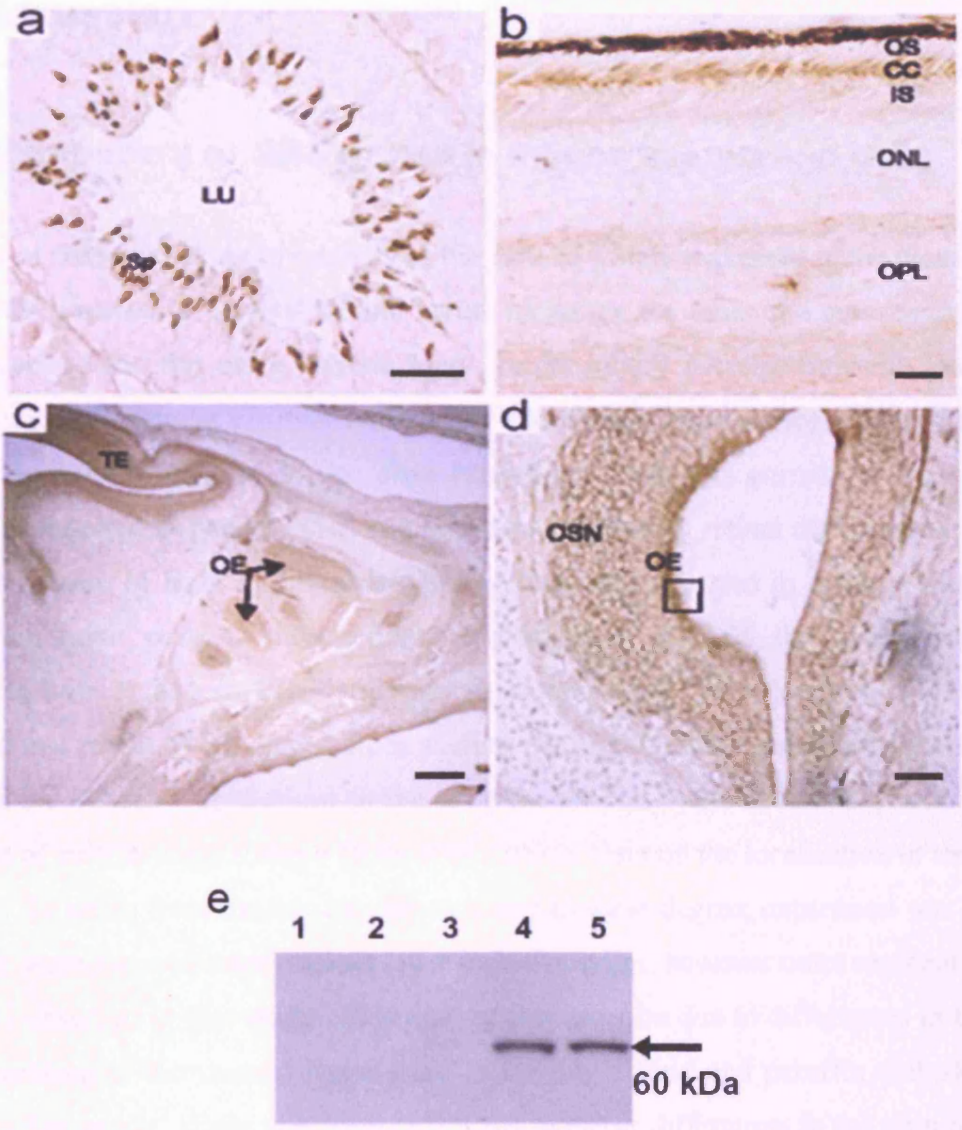


Figure 8. 4: *BBS8* expression in adult and embryonic mouse tissues. a) *BBS8* expression was seen in maturing spermatids (Sp) surrounding the lumen (LU) of the seminiferous tubule from a 12 day old mouse. b) In the retina, BBS8 was found to be exclusively expressed in the connecting cilium (CC). c) In embryonic tissue expression was seen in the olfactory epithelium (OE) and developing telencephalon (TE) in an E14 embryo, and also, d) in the OE and olfactory sensory neurons (OSN) in an E16 embryo (Ansley et al., 2003). All images are from Novagen slides with a four minute microwave step and five minutes DAB reaction. Scale bars represent 200μm in a) and b), 500μm in c) and 100μm in d). e) Western to show specificity of the BBS8 antibody. The BBS8 antibody, but not the preimmune serum, was able to specifically immunoprecipitate (IP) BBS8-Mycfrom BBS8-Myc transfected HEK293 cells. Lane 1 – Untransfected cells, 2 - Empty vector, 3 – Preimmune IP, 4 – BBS8 antibody IP, 5 – cell lysate (Ansley et al., 2003).

8.4 Discussion

8.4.1 Expression of *BBS* genes in mouse tissues and cells

Analysis of *BBS4* in tissues revealed that the gene is widely expressed in the mouse eye, with high levels observed in several retinal layers including the inner and outer segments of the photoreceptors and the outer nuclear layer, in the ciliary muscles (muscles that relax the fibres holding the lens in position and enable it to change shape during accommodation) and also in the cornea (Figure 8. 3). This expression pattern is consistent with the ocular phenotype reported in patients (See *1.4.1 Retinal dystrophy*); retinal degeneration is the most common feature of BBS and strabismus has also been reported in patients (Beales et al., 1999). In a larger, parallel immunohistochemical study of BBS4, the protein was shown to colocalise with PCM1 (a BBS4-interacting protein, see *7.4.1 Pericentriolar material 1 (PCM1)*) in a subset of cell populations such as the hippocampus, columnar epithelium of the lung and the olfactory epithelium in the adult mouse and in the developing pericardium and limb bud of an E16 stage embryo (Kim et al., 2004). Data on the localisation of the protein in layers of the retina from the two studies overlaps to some degree; expression was detected in the inner segments and outer nuclear layer in both studies, however outer segment expression was only observed in this study. This discrepancy may be due to differences in the fixation and processing of the tissue (tissue fixed in Carony's fluid and paraffin embedded versus 'hybridisation-ready' slides purchased from Novagen) or differences in the staining protocol. Further work is required to determine whether the expression in the outer segments represents true expression of the protein or is background staining. The absence of DAB staining in the outer segment in control sections suggests that this signal is more than merely background.

BBS8 was found to be almost exclusively expressed in ciliated tissues (Figure 8. 4), consistent with the cellular localisation of the protein (Ansley et al., 2003). In adult mouse testes, *BBS8* expression was observed in flagellated spermatids surrounding the lumen of seminiferous tubules. Defects in the length or function of sperm flagella caused by mutations in *BBS* genes may explain the lower levels of fertility observed in some male patients

(Mykytyn et al., 2004). In the retina *BBS8* was found to have a more restricted expression pattern than *BBS4*, present only in the connecting cilium of the photoreceptors. The connecting cilium, located between the inner and outer segments, is the region of the photoreceptor through which newly synthesised components of the rods and cones are transported from their site of synthesis (the inner segment) to the outer segment where they are required to maintain the cell. As photoreceptor components are continually turned over at a high rate, ~2,000 opsin molecules are required every minute to maintain the rods, defects in this form of intraflagellar transport (IFT) leads to a degeneration of the rods and cones (Rosenbaum and Witman, 2002). In the embryo, expression was seen in the developing telencephalon, particularly at the ciliated ependymal cell layer, and the olfactory epithelium and olfactory sensory neurons.

Immunocytochemical analysis of the cellular localisation of BBS4 and BBS8 revealed that both proteins colocalise to the centrosome in a number of cultured mammalian cells, including HEK 293 and NIH 3T3 cells (Ansley et al., 2003; Kim et al., 2004). In ciliated cells, one of the pair of centrioles (the mother centriole) is recruited to the cell surface to form the basal body of the cilium. Analysis of BBS4 and BBS8 in the ciliated murine kidney cell line IMCD3 revealed expression of both proteins at the basal body of the cilium and also at the centrosome, associated with the daughter centriole.

8.4.2 Improvements to the staining protocol

Although promising results were achieved with BBS4 and BBS8 antibodies, further experiments are required to confirm that the staining pattern observed is a true reflection of the expression pattern of the genes and are not affected by non-specific background staining. Blocking of endogenous peroxidase activity and non-specific sites was carried out prior to the primary antibody incubation step, eliminating certain types of background. Blocking of EABA was not carried out in these experiments but, as EABA-induced DAB staining would be uniform between sections incubated with primary antibody and those without, and no DAB staining was present in the control sections, it is likely that EABA in tissues such as the

eye, testis and olfactory epithelium is not sufficient to cause background staining. 'No primary' is a commonly used negative control however, a better control is to test whether staining is inhibited by adsorption of the antibody with the peptide to which the antibody was raised, but not with an unrelated peptide. If staining is successfully blocked by adsorption, it indicates that staining is specific to the protein in question, rather than being caused by a contaminating antibody also present in the antisera. Additional experiments using the EABA block and the adsorption control would be helpful to establish the reliability of the results already achieved and the specificity of the BBS4 and BBS8 antibodies.

8.4.3 Summary

Immunohistochemical staining methods are an effective way of analysing expression of genes within tissues and determining specific cell populations in which expression levels are highest. Optimisation of fixation and staining protocols are essential to ensure a high level of specific expression with a low level of non-specific background staining. *BBS4* expression was observed throughout the eye, whereas *BBS8* expression was restricted to the connecting cilium. *BBS8* expression was also studied in the testis and embryonic tissues and was present in maturing spermatids and the developing telencephalon and olfactory epithelium respectively.

Chapter 9 General discussion

9.1 Cloning of *BBS* genes

BBS is a highly heterogeneous disease with eight cloned loci (*BBS1-8*). In recent years considerable progress has been made in the identification of the disease-causing genes by conventional and more novel gene mapping methods (Table 9. 1). The first *BBS* genes to be cloned (*BBS2*, *4* and *6*) were identified using conventional positional cloning methods involving mutation screening of candidate genes or transcripts in the pedigrees used to map the loci (Katsanis et al., 2000; Mykytyn et al., 2001; Nishimura et al., 2001; Slavotinek et al., 2000). Following the successful cloning of the most common *BBS* gene, *BBS1*, by a combination of positional cloning and sequence homology to BBS2 (Mykytyn et al., 2002), the *BBS7* and *BBS8* genes were recently cloned based on their homology to BBS2 and BBS4, respectively (Ansley et al., 2003; Badano et al., 2003a).

At the time of cloning, the majority of *BBS* genes (*BBS1*, *2*, *4* and *7*) were novel transcripts of unknown function (Badano et al., 2003a; Mykytyn et al., 2001; Mykytyn et al., 2002; Nishimura et al., 2001). The *BBS6* gene, mutations in which also cause the related disorder MKKS, is predicted to be a type II chaperonin based on its similarity to archeobacterial chaperonins (Stone et al., 2000). Little knowledge of potential targets of the BBS6/MKKS protein however makes verification of this putative function very difficult. Recent functional studies on other BBS proteins have been more successful. BBS4, BBS6 and BBS8 have all been shown to localise at the centrosome and basal body in ciliated cells (Ansley et al., 2003; Kim et al., 2004). *BBS4*, *BBS7* and *BBS8* are expressed in a specific set of ciliated head and tail neurons in the nematode *C. elegans* and, like many other genes expressed in ciliated structures in the organism, contain an upstream 14bp sequence known as the X-box which is a binding site for the RFX transcription factor DAF-19, suggesting a role for BBS proteins in cilia assembly, maintenance or function (See *9.5 The role of cilia in BBS*) (Ansley et al., 2003; Blacque et al., 2004; Kim et al., 2004).

A greater understanding of the putative function of known BBS proteins has also aided the cloning of the *BBS3* and *BBS5* genes that, until very recently, remained unidentified. Due to the small number of families linked to both loci and, in the case of *BBS3*, the position of the gene at the centromere, an area of little recombination, it was only with the aid of functional data on other BBS proteins that it has been possible to clone both genes. Li *et al.* (2004) used a comparative genomics approach to clone the *BBS5* gene. The proteome of the non-flagellated plant *Arabidopsis* was subtracted from the proteome shared by *Chlamydomonas* (a single-celled organism with flagella) and humans, giving a flagellar apparatus-basal body (FABB) proteome. Using this method, 688 genes specifically expressed in organisms with basal bodies or flagella were identified. In addition to five of the previously cloned *BBS* genes (*BBS1*, *2*, *4*, *7* and *8*), two genes, one known to be involved in IFT (*IFT139*) and the other unknown (NM_152384), located within the *BBS5* critical interval were present in the gene list. Mutational analysis of both transcripts resulted in the identification of pathogenic alterations including splice site (IVS6+3A>G), insertion/deletion (c.263-272indelGCTCTTA), nonsense (W59X and L142X) and missense (N184S and R207H) mutations, in the novel gene (NM_152384), now referred to as *BBS5*. Analysis of the *C. elegans* ortholog of *BBS5* (*bbs-5*) revealed the presence of an X-box in the promoter of the gene and an expression pattern similar to that of other *BBS* orthologs, with specific expression of the gene in the ciliated head and tail neurons (Li et al., 2004).

Shortly after the cloning of *BBS5*, the *BBS3* gene was identified by two groups (Chiang et al., 2004; Fan et al., 2004a). Chiang *et al.* (2004) used a similar approach to that of Li *et al.* (2004), identifying *ARL6,* a member of the ARL (ADP-ribosylation factor (ARF) - like) family of small GTP-binding proteins, as a candidate for *BBS3* by comparative genomics of ciliated and non-ciliated organisms. Mutation screening of the gene in a large Bedouin pedigree linked to the locus revealed a homozygous nonsense mutation (R122X) segregating with disease in 13 affected individuals. Using an alternative approach, Fan *et al.* (2004a) analysed the *C. elegans* genome for genes containing X-box sequences, based on the assumption that all *C. elegans* orthologs of *BBS* genes will be under the control of DAF-19. Of 368 *C. elegans* genes, 168 were found to have a human ortholog. Three of these genes, *ESRRBL1* (probably the human ortholog of the *C. elegans* gene *che-13*), *ARL6* and *ARL2L1*,

were located in the *BBS3* critical interval. Sequencing of a large cohort of patients identified several *ARL6* missense mutations (T31M, T31R, G169A and L170W) in pedigrees of various ethnicities (Saudi Arabian, Irish, Newfoundland and North American), providing independent confirmation that *ARL6* is *BBS3* (Fan et al., 2004a).

Locus	Mapping	Position	Cloning	Contribution*	Function
BBS1	Linkage in outbred pedigrees	11q13	Positional cloning/ homology to BBS2	23.2%	Unknown
BBS2	IBD in Bedouin	16q21	Positional cloning	8.1%	Unknown
BBS3	IBD in Bedouin	3p13	Positional cloning/ comparative genomics	0.4%	GTP-binding protein
BBS4	IBD in Bedouin	15q23	Positional cloning	2.3%	Centrosomal/ basal body
BBS5	IBD in Newfoundland	2q31	Positional cloning/ comparative genomics	0.4%	Basal body
BBS6	IBD in Newfoundland	20p12	Positional cloning	5.8%	Putative chaperonin
BBS7	Not mapped	4q32	Homology to BBS1/2	1.5%	Basal body
BBS8	Not mapped	14q31	Homology to BBS4	1.7%	Centrosomal/ basal body

Table 9. 1: Summary of known BBS genes. * contribution data from Katsanis (2004). See Appendices 1 - 6 for schematics of the genomic orginisation, position of mutations and protein structure of BBS1, 2, 4, 6, 7 and 8.

9.2 Mapping of novel *BBS* loci

Despite the recent cloning of eight *BBS* genes (Table 9. 1), over half the cases of BBS are unaccounted for by mutations in the known genes (Katsanis, 2004). A genome-wide homozygosity screen carried out in this project attempting to map a novel *BBS* locus in pedigrees of Middle Eastern origin was unsuccessful. With the exception of *BBS1*, the main *BBS* locus, each of the remaining loci are responsible for a small fraction of the disease, with four of the loci (*BBS3*, *5*, *7* and *8*) estimated to each account for fewer than 2% of cases (Katsanis, 2004). This high degree of genetic heterogeneity makes combining data from a number of different families in genome-wide screens very difficult, as it is unlikely that a

region of homozygosity shared by a number of pedigrees will be identified. Large consanguineous pedigrees in which locus homogeneity can be reliably assumed are therefore required, as were used to successfully map a number of the known loci (Carmi et al., 1995b; Katsanis et al., 2000; Kwitek-Black et al., 1993; Sheffield et al., 1994; Young et al., 1999a). The lack of such a large pedigree in our cohort meant that it was necessary to combine data from a number of small consanguineous pedigrees, each with only one or two affecteds. As all the pedigrees used in the screen were of Middle Eastern origin (Indian, Kurdish, Pakistani and Turkish), it was hoped that the screen would define a novel locus, common to Middle Eastern pedigrees. No significant regions of homozygosity shared between all 17 pedigrees were identified, suggesting that a degree of genetic heterogeneity exists even within this population group. Surprisingly high levels of heterogeneity have also been reported in the two populations in which BBS is most common, on the island of Newfoundland and in the Bedouin of Israel. At the time of the initial mapping studies, when all cases of BBS were assumed to be caused by mutations at the same genetic locus, three large Bedouin pedigrees were studied, expecting to allow the mapping and cloning of the causative gene. Linkage analysis in the different kindreds however revealed that each was mapped to a different locus, resulting in the mapping of *BBS2*, *3* and *4* (Carmi et al., 1995b; Kwitek-Black et al., 1993; Sheffield, 2004; Sheffield et al., 1994). The high incidence of BBS on Newfoundland was initially thought to be a result of a founder effect. Results from haplotype and mutational analysis of the BBS pedigrees on the island have in fact shown this to not be entirely the case. In addition to the *BBS1*, *2*, *3*, *5* and *6* genes being represented on the island (Beales et al., 2001; Katsanis et al., 2000; Young et al., 1999a; Young et al., 1999b; Young et al., 1998), multiple different mutations in the genes in Newfoundland patients have demonstrated the presence of both genetic and allelic heterogeneity in this BBS population (Beales et al., 2001; Fan et al., 2004a; Fan et al., 2004b; Katsanis et al., 2000; Li et al., 2004).

9.3 Complex inheritance in BBS

Genetic diseases are commonly classified as monogenic, caused by mutations at a single locus, or polygenic, resulting from the combination of mutations at multiple loci. For a

growing number of disorders that were initially classed as monogenic, recent research has shown that this may not be true, and that the disease phenotype is influenced by mutations at more than one locus (Badano and Katsanis, 2002; Burghes et al., 2001). Suspicions about whether BBS might be inherited in a non-Mendelian manner were first raised due to the relatively high frequency of single mutant *BBS6* alleles identified on mutation screening of the gene in a large patient cohort (Beales et al., 2001). In seven of eight pedigrees in which a *BBS6* mutation was identified, only a single mutant allele was detected. It was not until the cloning of the second *BBS* gene, *BBS2*, that the hypothesis of multiallelic inheritance in BBS could be tested. Mutation screening of the ORF of *BBS2* in the same cohort of 163 patients provided evidence for digenic triallelic inheritance, two mutant alleles at one locus with a third mutation at a second locus (Rivolta et al., 2002), in a subset of families (Katsanis et al., 2001a). Mutation screening of all newly cloned *BBS* genes in the patient cohort has shown that triallelic inheritance is not restricted to only *BBS2* and *6*; incidences of triallelic inheritance have been reported in all known *BBS* genes, with the exception of *BBS8* (Beales et al., 2003; Fan et al., 2004a; Katsanis et al., 2001a; Katsanis et al., 2002; Li et al., 2004). Participation in triallelic inheritance varies for the different *BBS* loci; *BBS2* and *6* are frequently involved in complex inheritance, whereas over 70% of *BBS1* mutations are inherited in a recessive manner (Katsanis, 2004).

As an extension to the triallelic hypothesis, in a consanguineous Kurdish pedigree two homozygous mutations, T558I in *BBS2* and A364E in *BBS4* were found to segregate with disease (Chapter 5, Katsanis et al., 2002). As both mutations in this case are missense, functional studies are required to confirm the pathogenicity of the mutant alleles. The nature of mutations involved in triallelic inheritance is variable, with nonsense, frameshift, splice site and missense mutations all reported in association with complex inheritance (Table 9. 2). In the majority of cases, at least one or two missense mutations are involved, raising the question of whether the alterations are pathogenic or merely polymorphisms. In one North American pedigree, AR259 (Fig 5.5), the segregation of three *bona fide* loss of function mutations (Y24X and Q59X in *BBS2*, and Q147X in *BBS6*) with disease, leaves little doubt about the pathogenicity of each of the mutations present in this family (Katsanis et al., 2001a).

In the majority of cases of complex inheritance in BBS, three mutant alleles are both necessary and sufficient to cause disease, with siblings or parents in some cases segregating two mutant alleles at one locus despite being asymptomatic (Beales et al., 2003; Katsanis et al., 2001a). In some cases however, the presence of the third allele has a modifying rather than causative effect (Table 9. 2). Badano *et al.* (2003b) described three pedigrees, each with multiple affected sibs and a high degree of intrafamilial variation in the phenotype; a third allele was found to segregate with the more affected sib in each case. A modifying allele in *ARL6* (*BBS3*) was also found to segregate with the more severely affected of two sisters with BBS, each of whom also carry a homozygous M390R (*BBS1*) mutation (Fan et al., 2004a).

The occurrence of complex inheritance in a proportion of BBS families makes genetic counselling challenging. In a recessive monogenic condition, following the identification of homozygous or compound heterozygous mutations at one locus, mutational analysis is usually considered to be complete and the pedigree assigned to the respective locus. With the lack of any genotype-phenotype relationship in BBS, added to the lack of any phenotypic characteristics allowing recessive pedigrees to be distinguished from triallelic ones, it is necessary to screen a new cases of BBS for mutations in all known *BBS* genes to ensure that no potentially pathogenic mutations will be missed. Sensitive, quick and cheap mutation screening methods that will detect both heterozygous and homozygous, and known and novel mutations is therefore required. The MCHA technique has been shown to have a high throughput and sensitivity at a considerably lower cost than direct sequencing, making it highly applicable for this function (Chapter 6, Hoskins et al., 2003). However, the development of a mutation or SNP chip will aid future diagnosis.

Although questioned by some (Mykytyn et al., 2003), supporting evidence for a digenic inheritance pattern in BBS has been reported by Fauser *et al.* (2003). In their cohort of 21 patients no cases of triallelic inheritance were found, but two cases of suspected digenic diallelic inheritance between *BBS2* and *4* were identified (Table 9. 2). In addition, a single heterozygous *BBS6* mutation was also identified in one patient. As the *BBS3* and *5* genes had not yet been cloned at the time of this study, it is possible that this patient may have additional mutations in either of these genes or in a novel, as yet unidentified *BBS* gene.

Allele 1	Allele 2	Allele 3	Phenotype	Inheritance	Reference
BBS1 (R)	*BBS1* (R)	*BBS2* (M)	BBS	Modifier	(Badano et al., 2003b)
BBS1 (R)	*BBS1* (R)	*BBS3* (M)	BBS	Modifier	(Fan et al., 2004a)
BBS1 (R)	*BBS1* (R)	*BBS4* (M)	BBS	Triallelic	(Beales et al., 2004)
BBS1 (R)	*BBS1* (R)	*BBS5* (M)	BBS	Triallelic	(Li et al., 2004)
BBS1 (N)	*BBS1* (R)	*BBS6* (M)	BBS	Triallelic	(Beales et al., 2004)
BBS1 (fs)	*BBS1* (R)	*BBS6* (M)	BBS	Triallelic	(Beales et al., 2004)
BBS1 (fs)	*BBS1* (R)	*BBS6* (M)	BBS	Modifier	(Badano et al., 2003b)
BBS2 (Sp)	*BBS2* (M)	*BBS1* (R)	BBS	Triallelic	(Beales et al., 2003)
BBS2 (N)	*BBS2* (N)	*BBS1* (Sp)	BBS	Modifier	(Badano et al., 2003b)
BBS2 (M)	*BBS2* (M)	*BBS4* (M)*	BBS	Tetra-allelic	(Katsanis et al., 2002)
BBS2 (N)	*BBS2* (N)	*BBS6* (N)	BBS	Triallelic	(Katsanis et al., 2001a)
BBS2 (fs)	*BBS2* (N)	*BBS6* (M)	BBS	Triallelic	(Katsanis et al., 2001a)
BBS2 (N)	*BBS2* (N)	*BBS6* (M)	BBS	Triallelic	(Katsanis et al., 2001a)
BBS7 (M)	*BBS7* (M)	*BBS1* (M)	BBS	Triallelic	(Beales et al., 2003)
NPHP1 (del)	*NPHP1* (del)	*BBS4* (M)	SLS	Triallelic	Chapter 5
NPHP1 (N)	*BBS4* (M)	-	SLS	Digenic diallelic ?	Chapter 5
NEK8 (M)	*BBS6* (M)	-	SLS	Digenic diallelic ?	Chapter 5
BBS2 (N)	*BBS4* (M)	-	BBS	Digenic diallelic	(Fauser et al., 2003)
BBS2 (M)	*BBS4* (M)	-	BBS	Digenic diallelic	(Fauser et al., 2003)

Table 9. 2: Summary of triallelic allele combinations. R – M390R, M – missense mutation, N – nonsense mutation, fs – frameshift mutation, Sp – splice site mutation, del – whole gene deletion, SLS – Senior Löken syndrome.

In addition to contributing to non-Mendelian inheritance in BBS patients, *BBS4* and *6* have also been found to be associated with complex inheritance in a related disorder, Senior Löken syndrome (SLS). SLS is composed of nephronophthisis (NPHP) and RP and, like BBS, is a heterogenic condition caused by mutations in several known genes (*NPHP1*, *INVS*, *NPHP3* and *NPHP4*) and also additional, as yet unmapped, genes. Evidence for non-Mendelian inheritance has been reported recently in SLS and isolated NPHP; single mutant alleles were

detected in patients from mutation screens for *INVS* (NPHP2), *NPHP3* and *NPHP4* (Mollet et al., 2002; Olbrich et al., 2003; Otto et al., 2003). The similarity in phenotype, and also protein function (See *9.6.3 Nephronophthisis (NPHP))*, prompted the screening of a cohort of NPHP and SLS patients for mutations in *BBS* genes. One case of triallelic inheritance between *NPHP1* and *BBS4* was detected and two cases of potential digenic diallelic inheritance, one between *NPHP1* and *BBS6* and the other between *NEK8*, a recently cloned *NPHP* gene, and *BBS4*, were also found (Table 9. 2, Chapter 5).

9.4 BBS protein function

9.4.1 BBS4

The BBS4 protein is composed of several TPR domains and in recent functional studies, has been shown to be a centriolar protein that colocalises and interacts with pericentriolar material 1 (PCM1) (Chapter 7, Kim et al., 2004). In ciliated IMCD3 cells, BBS4 is associated with both the mother centriole at the centrosome and also with the daughter centriole at the basal body of the cilium. Depletion of BBS4 by RNA interference (RNAi) results in dispersal of PCM1 into the cytoplasm of the cell, a failure of microtubule anchoring at the centrosome, and increased levels of apoptosis. In addition to PCM1, BBS4 also interacts with the $p150^{glued}$ subunit of dynactin (Chapter 7, Kim et al., 2004). As the dynactin complex interacts with dynein and is required for dynein-associated movement along the microtubules, these results suggest that BBS4 may function as an adaptor molecule to transport PCM1 to centrosomal satellites. Yeast-two-hybrid analysis also identified a number of nuclear receptors as potential interactors of BBS4 (Chapter 7). As these results have not yet been confirmed, their significance is hard to interpret at present.

9.4.2 BBS8

As *BBS8* was cloned based on its homology to BBS4, the two proteins share several functional features. BBS8 colocalises with BBS4 and PCM1 at the centrosome and basal body and interacts directly with the same region of PCM1 that mediates the BBS4-PCM1 interaction (the C-terminal) (Ansley et al., 2003). BBS8 expression in mouse tissue samples showed that the protein is expressed in ciliated structures and tissues in the developing and adult mouse, including the connecting cilium of the retina, flagellated spermatids and the olfactory epithelium (Chapter 8, Ansley et al., 2003). *BBS8* is highly conserved and, like several other BBS genes (*BBS1*, *2*, *3*, *5* and *7*), has an ortholog in *C. elegans* (*bbs-8*). The expression pattern of *bbs-8* is very specific; the protein is present exclusively in the ciliated structures of the head and tail neurons and the midbody PDE neuronal cell in hermaphrodites and also in the tail-ray neurons of the male. Analysis of the 5' untranslated region (UTR) of *bbs-8* identified an X-box sequence, indicating that *bbs-8* is under the control of DAF-19, or RFX in mammals (Ansley et al., 2003; Blacque et al., 2004).

9.4.3 ARL6 (BBS3), BBS5 and BBS7

Like *bbs-8*, *C. elegans* orthologs of *ARL6* (*BBS3*), *BBS5* and *BBS7* (*arl-6*, *bbs-5* and *bbs-7*) are specifically expressed in head and tail neurons and are under the control of DAF-19 (Ansley et al., 2003; Blacque et al., 2004; Fan et al., 2004a; Li et al., 2004).

9.5 The role of cilia in BBS

Several lines of evidence from functional studies on BBS proteins support a role for these proteins in cilia assembly or function, implicating ciliary dysfunction as a possible cause of BBS (Ansley et al., 2003; Blacque et al., 2004; Fan et al., 2004a; Kim et al., 2004; Li et al., 2004). Cilia are divided into two classes: motile, which have an internal structure composed of two central microtubules surrounded by nine outer microtubule doublets (9+2 structure),

or non-motile, which contain the nine outer doublets but lack the central microtubule pair (9+0 structure) (Rosenbaum and Witman, 2002). Non-motile, or primary cilia, are present on almost all mammalian cells and are thought to play important roles in development, homeostasis and sensation (Snell et al., 2004). In *C. elegans* cilia perform a chemosensory role, allowing the animal to sense its surroundings and locate potential mates (Igarashi and Somlo, 2002). They also perform a sensory role in humans with the outer segments of the rod cells, a modified cilium, responsible for photoreception and the odorant receptors located on the cilia of the olfactory neurons responsible for odorant signalling. Renal primary cilia act as mechanosensors in the kidney epithelium, causing an increase in intracellular calcium on bending (Snell et al., 2004).

A cilium is made up of a basal body, derived from the mother centriole at the centrosome, and the ciliary axoneme along which protein components are transported by IFT for assembly and maintenance of the cilium. Movement towards the tip of the cilium (anterograde transport) is mediated by kinesin II, a plus end-directed microtubule motor, whereas dynein is responsible for movement towards the cell body (retrograde transport) (Rosenbaum and Witman, 2002). IFT proteins in *C. elegans* localise to the axoneme and transition zone of the cilium, a structure at the base of the cilium that is analogous to the basal body. Using time-lapse microscopy, the proteins have also been shown to move bidirectionally along the axoneme (Blacque et al., 2004). Mutations in IFT genes result in chemosensory defects and abnormal cilia in *C. elegans* (Snell et al., 2004).

Analysis of *BBS* orthologs in *C. elegans* has identified several features that support the hypothesis for the involvement of BBS proteins in cilia biogenesis and/or maintenance (Ansley et al., 2003; Blacque et al., 2004; Fan et al., 2004a; Li et al., 2004):

- The expression patterns of *bbs-1*, *2*, *3*, *5*, *7* and *8* overlap with those of known *C. elegans* IFT genes including *osm-5*, *osm-6*, *xbx-1*, and *che-13*.
- All *C. elegans bbs* genes contain an X-box sequence, a binding site for DAF-19, which is present in cilia-specific genes.

- Loss of *bbs-7* and/or *bbs-8* result in structural and functional ciliary defects, including shortened, abnormal cilia, chemosensory defects and decreased IFT.
- Knockdown of *Bbs5* in *Chlamydomonas* results in an inability to form flagella.

In addition to an overlapping expression pattern, BBS-7, BBS-8 and ARL-6 (BBS3) have also been observed moving in an anterograde and retrograde direction along the axoneme of the cilium in a manner similar to that of other IFT proteins (Blacque et al., 2004; Fan et al., 2004a). ARL6 is known to be involved in membrane-associated intracellular trafficking processes in its role as a small GTP-binding protein. The observation of ARL-6 moving along the ciliary axoneme therefore extends this trafficking property to include IFT (Fan et al., 2004a). To establish if BBS proteins are involved in the movement of other IFT proteins in *C. elegans*, Blacque *et al.* (2004) analysed the movement of three known IFT proteins, OSM-5, CHE-2 and CHE-11, in *bbs-7* and *bbs-8* mutant nematodes. In the *bbs-8* mutant, OSM-5 and CHE-11 levels in the axoneme were greatly reduced, with a number of animals showing accumulation of OSM-5 in the transition zone. Similarly, in the *bbs-7* mutant, axonemal OSM-5 appeared disorganised and levels of CHE-11 were significantly reduced. Expression of CHE-2 was not affected by loss of function of *bbs-7* or *bbs-8*. The ability of BBS-7 and BBS-8 to transport proteins within the cilia, and for BBS4 to transport PCM1 to the centrosome and basal body (Kim et al., 2004), suggests that BBS proteins may share a common function as adaptor proteins mediating trafficking of a number of cellular components. A common pathway is one explanation as to how mutations in a number of different genes can give the indistinguishable phenotype seen in BBS.

The expression of *BBS8* in ciliated structures in mammalian tissues and the colocalisation of BBS4 and BBS8 with PCM1 at centriolar satellites and basal bodies in cultured cells also support the cilia/basal body hypothesis (Ansley et al., 2003; Kim et al., 2004). Certain features of the BBS phenotype can be easily explained by the ciliary dysfunction hypothesis. Structural and functional renal abnormalities are known to be caused by defects in cilia (See *9.6 Ciliary dysfunction in human disease*). IFT is essential to maintain the outer segments of the photoreceptor; mice with mutations in IFT genes *Kif3a* and *Polaris (Tg737)* develop retinal degeneration (Pazour et al., 2002). *Situs inversus*, a feature seen in some patients

(Ansley et al., 2003; Lorda-Sanchez et al., 2000), results from defects in the nodal cilia or nodal flow, resulting in abnormalities in left/right patterning (Capdevila et al., 2000). Anosmia, a recently recognised secondary feature of BBS, results from a reduced ciliated border of the olfactory epithelium and disorganisation of the dendritic microtubule network in *Bbs1* and *Bbs4* knockout mice (Kulaga et al., 2004).

9.6 Ciliary dysfunction in human disease

BBS is not the only human disease thought to be caused by a defect in cilia. Recently ciliary dysfunction has been implicated in several diseases, primarily conditions with a strong renal component (Snell et al., 2004; Watnick and Germino, 2003).

9.6.1 Autosomal dominant polycystic kidney disease (ADPKD)

Autosomal dominant polycystic kidney disease (ADPKD) is a common cause of renal failure, affected between 1 in 500 and 1 in 1,000 children and adults. The condition is characterised by the growth of fluid-filled cysts in the kidneys and also other organs including the liver, ovaries and occasionally the pancreas. ADPKD can be caused by mutations in either *PKD1*, which encodes the protein polycystin 1 (PC1), or *PKD2*, encoding polycystin 2 (PC2). The polycystins, which are known to heterodimerise, are involved in a number of important processes including fertilisation, ion translocation and mechanosensation and their *C. elegans* orthologs (LOV-1 and PKD-2) are essential for mating (Igarashi and Somlo, 2002). PC1 and PC2 have both been shown to colocalise with the ciliary markers acetylated α-tubulin, in the axoneme of the cilium, and γ-tubulin at the basal body (Nauli et al., 2003). PC2 functions as a calcium channel and PC1 may function as a mechanosensor, sensing the bending of the cilium by fluid flow, resulting in a conformational change in the protein and subsequent activation of PC2 channels. Loss of function of PC1 or PC2 may bring about cyst development by the inability to sense mechanical cues and hence regulate tissue morphogenesis (Nauli et al., 2003).

9.6.2 Autosomal recessive polycystic kidney disease (ARPKD)

Autosomal recessive polycystic kidney disease (ARPKD) combines renal cysts with congenital hepatic fibrosis and is rarer than ADPKD, with an incidence of between 1 in 6,000 and 1 in 40,000. The phenotype is variable and can present as perinatal, neonatal, infantile or juvenile (Igarashi and Somlo, 2002). The homozygous *orpk* mouse, which has a mutation in *Polaris (Tg737)*, is a model for ARPKD and provided the first indication of a link between polycystic kidneys and ciliary defects (Moyer et al., 1994). The *Chlamydomonas* ortholog of *Polaris*, *IFT88*, is essential for flagellar assembly; cells deficient in *IFT88* grow at the same rate as wt cells but lack flagella (Pazour et al., 2000). Analysis of the cilia in *orpk* mice revealed that the renal cilia are shorter than in wt mice. Rod outer segments in these mice are also affected and a progressive retinal degeneration develops caused by apoptosis of the photoreceptor cells (Pazour et al., 2002). Mice lacking the gene completely die *in utero* due to absence of the nodal cilia (Pazour et al., 2000).

9.6.3 Nephronophthisis (NPHP)

Mutations in *INVS*, a gene known to cause cystic kidney disease and *situs inversus* in mice (Mochizuki et al., 1998), were recently found to be responsible for NPHP2, the infantile form of NPHP (Otto et al., 2003). Inversin is localised to varicosities along the length of the ciliary axoneme and shows partial colocalisation with nephrocystin, the protein product of the *NPHP1* gene. In addition to interacting with nephrocystin (Otto et al., 2003), inversin has also been shown to interact with Apc2, a subunit of the anaphase promoting complex responsible for the ubiquitintation of cell regulators at metaphase-anaphase and mitosis-G1 transitions (Morgan et al., 2002). The identification of *NPHP3*, the gene responsible for adolescent NPHP, as being orthologous to the gene mutated in the *pcy* polycystic kidney mouse model, further supports the hypothesis of NPHP being caused by a defect in cilia (Olbrich et al., 2003). In addition, the *NPHP4* gene was found to be present in the FABB proteome generated by Li *et al.* (2004).

9.7 BBS mouse models

In order to further understand the pathogenesis of BBS and determine how defects in cilia contribute to the obesity, polydactyly and genital abnormalities present in patients, it is necessary to study mouse models of the disease. Preliminary studies on a BBS4 knockout mouse (*Bbs4*$^{-/-}$) demonstrated that the animal develops some characteristics of BBS; mice develop obesity, retinal degeneration and fail to reproduce (Mykytyn et al., 2004). At birth *Bbs4*$^{-/-}$ mice appear smaller than their wt littermates but by 12 weeks of age weigh significantly more than wt mice. Further examination of the retina of the mice confirmed that retinal development is normal, but photoreceptors begin to degenerate at six weeks as a result of increased apoptosis. Analysis of renal epithelium from both knockout and wt mice showed that normal cilia form in the absence of Bbs4, suggesting that the protein is not required for cilia assembly. Bbs4 may however be required for the formation of flagella; histological examination of the seminiferous tubules in *Bbs4*$^{-/-}$ mice revealed a complete absence of flagella, providing an explanation for the inability of male *Bbs4*$^{-/-}$ mice to sire offspring (Mykytyn et al., 2004). A more recent study on anosmia in mice deficient for Bbs4 or Bbs1 (*Bbs1*$^{-/-}$) reported these mice to have a highly variable phenotype, with obesity observed in only ~10% of animals by 10 weeks of age and retinal degeneration in 30% (Kulaga et al., 2004). Examination of the olfactory epithelium in the mice revealed a reduced ciliated border and trapping of cilia-enriched olfactory proteins such as type III adenylyl cyclase (ACIII) and the signal transduction protein $G_{\gamma 13}$. In both studies, polydactyly and renal abnormalities were absent (Kulaga et al., 2004; Mykytyn et al., 2004). It will also be interesting to use mouse models to test the hypothesis of triallelic inheritance. Crossing of different mutant mice may result in the development of additional features of the disease, producing a phenotype that more fully models the human disease.

9.8 Candidate genes for BBS

Owing to the recent success in mapping *BBS* genes by homology based approaches, it is likely that these mapping methods could be used to clone the gene(s) responsible for the cases of BBS which are unaccounted for by the eight known genes (*BBS1-8*). Genes encoding flagellar or basal body proteins identified by Li *et al.* (2004) are good candidates for *BBS* genes. As there was a strong enrichment for genes containing X-box motifs and proteins containing TPR domains in this dataset, combined with the fact that both features are present in a number of the known BBS proteins, transcripts predicted to include such sequences would also be good candidates for BBS. The identification of incidences of potential multiallelic inheritance involving mutations in *BBS4* and *6* in cases of SLS, combined with the shared function of the NPHP and BBS proteins, raises the possibility of mutations in NPHP genes accounting for some cases of BBS. Interestingly, a Libyan BBS patient has been reported with a pericentric inversion of chromosome 1 (p36.3q23) in the region in which *NPHP4* is located (Tayel et al., 1999). It is possible therefore that the disease in this patient is a result of a heterozygous mutation in combination with a disruption of *NPHP4* caused by the translocation.

References

Al-Awadi, S. A., Moussa, M. A., Naguib, K. K., Farag, T. I., Teebi, A. S., M, e. K., and el Dossary, L. (1985). Consanguinity among the Kuwaiti population. Clin Genet *27*, 483-486.

Alberts, B., Bray, D., Lewis, J., Raff, M., Roberts, K., and Watson, J. D. (1994). Molecular biology of the cell, 3rd edn (New York, Garland Publishing Inc.).

Ammann, F. (1970). Investigations clinique et genetique sur le syndrome de Bardet-Biedl en Suisse. J Genet Hum *18 (Suppl)*, 1-310.

Andersen, J. S., Wilkinson, C. J., Mayor, T., Mortensen, P., Nigg, E. A., and Mann, M. (2003). Proteomic characterization of the human centrosome by protein correlation profiling. Nature *426*, 570-574.

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.

Antonarakis, S. E., Krawczak, M., and Cooper, D. N. (2001). The nature and mechanism of human gene mutation. In The metabolic and molecular basis of inherited disease, Sly., ed., pp. 343-377.

Aronheim, A., Zandi, E., Hennemann, H., Elledge, S. J., and Karin, M. (1997). Isolation of an AP-1 repressor by a novel method for detecting protein-protein interactions. Mol Cell Biol *17*, 3094-3102.

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., and Katsanis, N. (2002). Beyond Mendel: An evolving view of human genetic disease transmission. Nat Rev Genet *3*, 779-789.

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.

Balczon, R., Bao, L., and Zimmer, W. E. (1994). PCM1, a 228kD centroosme autoantigen with a distinct cell cycle distibution. J Cell Biol *124*, 783-789.

Bardet, G. (1920) Sur un syndrome d'obesite infantile avec polydactylie et retinite pigmentaire (contribution a l'etude des formes cliniques de l'obesite hypophysaire), PhD Thesis, Sorbonne University of Paris, Paris.

Barnett, S., Reilly, S., Carr, L., Ojo, I., Beales, P. L., and Charman, T. (2002). Behavioural phenotype of Bardet-Biedl syndrome. J Med Genet *39*, e76.

Barry, J. B., Leong, G. M., Church, W. B., Issa, L. L., Eisman, J. A., and Gardiner, E. M. (2003). Interactions of SKIP/NCoA-62, TFIIB, and retinoid X receptor with vitamin D receptor helix H10 residues. J Biol Chem *278*, 8224-8228.

Baudino, T. A., Kraichely, D. M., Jefcoat, S. C., Jr., Winchester, S. K., Partridge, N. C., and MacDonald, P. N. (1998). Isolation and characterization of a novel coactivator protein, NCoA-62, involved in vitamin D-mediated transcription. J Biol Chem *273*, 16434-16441.

Bauman, M. L., and Hogan, G. R. (1973). Laurence-Moon-Biedl syndrome. Report of two unrelated children less than 3 years of age. American Journal of Disease in Children *126*, 119-126.

Beales, P., Parfrey, P. S., and Katsanis, N. (2004). The Bardet-Biedl and Alstrom syndromes. In Genetics of Renal Disease, F. A. Flinter, E. R. Maher, and A. Saggar-Malik, eds. (Oxford, Oxford University Press).

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.

Beales, P. L., Elcioglu, N., Woolf, A. S., Parker, D., and Flinter, F. A. (1999). New criteria for improved diagnosis of Bardet-Biedl syndrome: results of a population survey. J Med Genet *36*, 437-446.

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., Reid, H. A., Griffiths, M. H., Maher, E. R., Flinter, F. A., and Woolf, A. S. (2000). Renal cancer and malformations in relatives of patients with Bardet-Biedl syndrome. Nephrol Dial Transplant *15*, 1977-1985.

Beales, P. L., Warner, A. M., Hitman, G. A., Thakker, R., and Flinter, F. A. (1997). Bardet-Biedl syndrome: a molecular and phenotypic study of 18 families. J Med Genet *34*, 92-98.

Bear, J. C., Nemec, T. F., Kennedy, J. C., Marshell, W. H., Power, A. A., Kolonel, V. M., and Burke, G. B. (1988). Inbreeding in outport Newfoundland. Am J Med Genet *29*, 649-660.

Bell, J. (1958). The Laurence-Moon syndrome. In The treasury of human inheritance, L. S. Penrose, ed. (Cambridge, Cambridge Univesity press), pp. 51-96.

Benson, D. M., Shepherd, S. R., and Dash, B. (1988). Idiopathic thrombocytopenic purpura in a patient with Laurence-Moon-Bardet-Biedl syndrome. J Am Osteopath Assoc *88*, 779-783.

Bhattacharyya, A., and Lilley, D. M. (1989). The contrasting structures of mismatched DNA sequences containing looped-out bases (bulges) and multiple mismatches (bubbles). Nucleic Acids Res *17*, 6821-6840.

Biedl, A. (1922). Ein Geschwisterpaar mit adiposo-genitaler dystrophie. Deutscher Medizinischen Wochenschrift *48*, 1630.

Blacque, O. E., Reardon, M. J., Li, C., McCarthy, J., Mahjoub, M. R., Ansley, S. J., Badano, J. L., Mah, A. K., Beales, P. L., Davidson, W. S., *et al.* (2004). Loss of C. elegans BBS-7 and BBS-8 protein function results in cilia defects and compromised intraflagellar transport. Genes Dev *18*, 1630-1642.

Blatch, G. L., and Lassle, M. (1999). The tetratricopeptide repeat: a structural motif mediating protein-proteininteractions. Bioessays *21*, 932-939.

Boenisch, T. (2001a). Antibodies. In Immunochemical staining methods handbook, T. Boenisch, ed. (Carpinteria, Califormia, DakoCytomation).

Boenisch, T. (2001b). Background. In Immunochemical staining methods handbook, T. Boenisch, ed. (Carpinteria, California, DakoCytomatin).

Boenisch, T. (2001c). Basic enzymology. In Immunochemical staining methods handbook, T. Boenisch, ed. (Carpinteria, California, DakoCytomation).

Boichis, H., Passwell, J., David, R., and Miller, H. (1973). Congenital hepatic fibrosis and nephronophthisis. A family study. Q J Med *42*, 221-233.

Borgstrom, M. K., Riise, R., Tornqvist, K., and Granath, L. (1996). Anomalies in the permanent dentition and other oral findings in 29 individuals with Laurence-Moon-Bardet-Biedl syndrome. J Oral Pathol Med *25*, 86-89.

Bruford, E. A., Riise, R., Teague, P. W., Porter, K., Thomson, K. L., Moore, A. T., Jay, M., Warburg, M., Schinzel, A., Tommerup, N., *et al.* (1997). Linkage mapping in 29 Bardet-Biedl syndrome families confirms loci in chromosomal regions 11q13, 15q22.3-q23, and 16q21. Genomics *41*, 93-99.

Burghes, A. H., Vaessin, H. E., and de La Chapelle, A. (2001). Genetics. The land between Mendelian and multifactorial inheritance. Science *293*, 2213-2214.

Capdevila, J., Vogan, K. J., Tabin, C. J., and Belmonte, J. C. I. (2000). Mechanisms of left-right determination in vertebrates. Cell *101*, 9-21.

Carmi, R., Elbedour, K., Stone, E. M., and Sheffield, V. C. (1995a). Phenotypic differences among patients with Bardet-Biedl syndrome linked to three different chromosome loci. Am J Med Genet *59*, 199-203.

Carmi, R., Rokhlina, T., Kwitek-Black, A. E., Elbedour, K., Nishimura, D., Stone, E. M., and Sheffield, V. C. (1995b). Use of a DNA pooling strategy to identify a human obesity syndrome locus on chromosome 15. Hum Mol Genet *4*, 9-13.

Cassart, M., Eurin, D., Didier, F., Guibaud, L., and Avni, E. F. (2004). Antenatal renal sonographic anomalies and postnatal follow-up of renal involvement in Bardet-Biedl syndrome. Ultrasound Obstet Gynecol *24*, 51-54.

Cavailles, V., Dauvois, S., L'Horset, F., Lopez, G., Hoare, S., Kushner, P. J., and Parker, M. G. (1995). Nuclear factor RIP140 modulates transcriptional activation by the estrogen receptor. EMBO J *14*, 3741-3751.

Chiang, A. P., Nishimura, D., Searby, C., Elbedour, K., Carmi, R., Ferguson, A. L., Secrist, J., Braun, T., Casavant, T., Stone, E. M., and Sheffield, V. C. (2004). Comparative Genomic Analysis Identifies an ADP-Ribosylation Factor-like Gene as the Cause of Bardet-Biedl Syndrome (BBS3). Am J Hum Genet *75*, 475-484.

Chien, C., Bartel, P. L., Sternglanz, R., and Fields, S. (1991). The two-hybrid system: A method to identify and clone genes for proteins that interact with a protein of interest. Proc Natl Acad Sci *88*, 9578-9582.

Churchill, D. N., McManamon, P., and Hurley, R. M. (1981). Renal disease - a sixth cardinal feature of the Laurence-Moon-Biedl syndrome. Clinical Nephrology *16*, 151-154.

Collins, C. M., Mendoza, S. A., Griswold, W. R., Tanney, D., Lieberman, E., and Reznik, V. M. (1994). Pediatric renal transplantation in Laurence-Moon-Biedl syndrome. Pediatr Nephrol *8*, 221-222.

Collins, F. S. (1992). Positional cloning: Let's not call it reverse anymore. Nat Genet *1*, 3-6.

Coons, A. H. (1951). Fluorescent antibodies as histochemical tools. Fed Proc *10*, 558-559.

Cotton, R. G. (1997). Slowly but surely towards better scanning for mutations. Trends Genet *13*, 43-46.

Cramer, B., Green, J., Harnett, J., Johnson, G. J., McManamon, P., Farid, N., Pryse-Phillips, W., and Parfrey, P. S. (1988). Sonographic and urographic correlation in Bardet-Biedl syndrome (formerly Laurence-Moon-Biedl syndrome). Urol Radiol *10*, 176-180.

Croft, J. B., Morrell, D., Chase, C. L., and Swift, M. (1995). Obesity in heterozygous carriers of the gene for the Bardet-Biedl syndrome. Am J Med Genet *55*, 12-15.

Croft, J. B., and Swift, M. (1990). Obesity, hypertension, and renal disease in relatives of Bardet-Biedl syndrome sibs. Am J Med Genet *36*, 37-42.

Dahl, R., Wani, B., and Hayman, M. J. (1998). The Ski oncoprotein interacts with Skip, the human homolog of Drosophila Bx42. Oncogene *16*, 1579-1586.

Dar, P., Sachs, G. S., Carter, S. M., Ferreira, J. C., Nitowsky, H. M., and Gross, S. J. (2001). Prenatal diagnosis of Bardet-Biedl syndrome by targeted second-trimester sonography. Ultrasound Obstet Gynecol *17*, 354-356.

David, A., Bitoun, P., Lacombe, D., Lambert, J. C., Nivelon, A., Vigneron, J., and Verloes, A. (1999). Hydrometrocolpos and polydactyly: a common neonatal presentation of Bardet-Biedl and McKusick-Kaufman syndromes. J Med Genet *36*, 599-603.

Deininger, P. L., and Batzer, M. A. (1999). Alu repeats and human disease. Mol Gen Metab *67*, 183-193.

Devarajan, P. (1995). Obesity and genitourinary anomalies in Bardet-Biedl syndrome after renal transplantation. Pediatr Nephrol *9*, 397-398.

Dippell, J., and Varlam, D. E. (1998). Early sonographic aspects of kidney morphology in Bardet-Biedl syndrome. Pediatr Nephrol *12*, 559-563.

Dyer, D. S., Wilson, M. E., Small, K. W., and Pai, G. S. (1994). Alstrom syndrome: a case misdiagnosed as Bardet-Biedl syndrome. J Pediatr Ophthalmol Strabismus *31*, 272-274.

Engelender, S., Sharp, A. H., Colomer, V., Tokito, M. K., Lanahan, A., Worley, P., Holzbaur, E. L. F., and Ross, C. A. (1997). Huntingtin-associated protein (HAP1) interacts with the p150^{Glued} subunit of dynactin. Hum Mol Genet *6*, 2205-2212.

Fan, Y., Esmail, M. A., Ansley, S. J., Blacque, O. E., Boroevich, K., Ross, A. J., Moore, S. J., Badano, J. L., May-Simera, H., Compton, D. S., *et al.* (2004a). Mutations in a member of the Ras superfamily of small GTP-binding proteins causes Bardet-Biedl syndrome. Nat Genet *36*, 989-993.

Fan, Y., Rahman, P., Peddle, L., Hefferton, D., Gladney, N., Moore, S. J., Green, J. S., Parfrey, P. S., and Davidson, W. S. (2004b). Bardet-Biedl syndrome 1 genotype and obesity in the Newfoundland population. Int J Obes Relat Metab Disord.

Fanconi, G., Hanhart, E., von, A. A., Uhlinger, E., Dolivo, G., and Prader, A. (1951). [Familial, juvenile nephronophthisis (idiopathic parenchymal contracted kidney)]. Helv Paediatr Acta *6*, 1-49.

Farag, T. I., and Teebi, A. S. (1989). High incidence of Bardet Biedl syndrome among the Bedouin. Clin Genet *36*, 463-464.

Fauser, S., Munz, M., and Besch, D. (2003). Further support for digenic inheritance in Bardet-Biedl syndrome. J Med Genet *40*, e104.

Fields, S., and Song, O. (1989). A novel genetic system to detect protein-protein interactions. Nature *340*, 245-246.

Ganguly, A. (2002). An update on conformation sensitive gel electrophoresis. Hum Mutat *19*, 334-342.

Ganguly, A., Rock, M. J., and Prockop, D. J. (1993). Conformation-sensitive gel electrophoresis for rapid detection of single-base differences in double-stranded PCR products and DNA fragments: evidence for solvent-induced bends in DNA heteroduplexes. Proc Natl Acad Sci U S A *90*, 10325-10329.

Gershoni-Baruch, R., Nachlieli, T., Leibo, R., Degani, S., and Weissman, I. (1992). Cystic kidney dysplasia and polydactyly in 3 sibs with Bardet-Biedl syndrome. Am J Med Genet *44*, 269-273.

Gill, S. R., Schroer, T. A., Szilak, I., Steuer, E. R., Sheetz, M. P., and Cleveland, D. W. (1991). Dynactin, a conserved ubiquitously expressed component of an activator of vesicle motility mediated by cytoplasmic dynein. J Cell Biol *115*, 1639-1650.

Grace, C., Beales, P., Summerbell, C., Jebb, S. A., Wright, A., Parker, D., and Kopelman, P. (2003). Energy metabolism in Bardet-Biedl syndrome. Int J Obes Relat Metab Disord *27*, 1319-1324.

Green, J. S., Parfrey, P. S., Harnett, J. D., Farid, N. R., Cramer, B. C., Johnson, G., Heath, O., McManamon, P. J., O'Leary, E., and Pryse-Phillips, W. (1989). The cardinal manifestations of Bardet-Biedl syndrome, a form of Laurence-Moon-Biedl syndrome. N Engl J Med *321*, 1002-1009.

Harnett, J. D., Green, J. S., Cramer, B. C., Johnson, G., Chafe, L., McManamon, P., Farid, N., Pryse-Phillips, W., and Parfrey, P. S. (1988). The spectrum of renal disease in Laurence-Moon-Biedl syndrome. N Engl J Med *319*, 615-618.

Harvey, R. (2001). In situ hybridization. In Immunochemical staining methods handbook, T. Boenisch, ed. (Carpinteria, California, DakoCytomation).

Heitzmann, H., and Richards, F. M. (1974). Use of the avidin-biotin complex for specific staining of biological membranes in electron microscopy. Proc Natl Acad Sci *71*, 3537-3541.

Highsmith, W. E., Jr., Jin, Q., Nataraj, A. J., O'Connor, J. M., Burland, V. D., Baubonis, W. R., Curtis, F. P., Kusukawa, N., and Garner, M. M. (1999). Use of a DNA toolbox for the characterization of mutation scanning methods. I: construction of the toolbox and evaluation of heteroduplex analysis. Electrophoresis *20*, 1186-1194.

Hildebrandt, F. (1999). Juvenile nephronophthisis. In Pediatric nephrology, W. E. Harmon, ed. (Baltimore).

Hildebrandt, F., Nothwang, H. G., Vossmerbaumer, U., Springer, C., Strahm, B., Hoppe, B., Keuth, B., Fuchshuber, A., Querfeld, U., Neuhaus, T. J., and Brandis, M. (1998). Lack of large, homozygous deletions of the nephronophthisis 1 region in Joubert syndrome type B. APN Study Group. Arbeitsgemeinschaft fur Padiatrische Nephrologie. Pediatr Nephrol *12*, 16-19.

Hildebrandt, F., Otto, E., Rensing, C., Nothwang, H. G., Vollmer, M., Adolphs, J., Hanusch, H., and Brandis, M. (1997a). A novel gene encoding an SH3 domain protein is mutated in nephronophthisis type 1. Nat Genet *17*, 149-153.

Hildebrandt, F., Strahm, B., Nothwang, H. G., Gretz, N., Schnieders, B., Singh-Sawhney, I., Kutt, R., Vollmer, M., and Brandis, M. (1997b). Molecular genetic identification of families with juvenile nephronophthisis type 1: rate of progression to renal failure. APN Study Group. Arbeitsgemeinschaft fur Padiatrische Nephrologie. Kidney Int *51*, 261-269.

Hildebrandt, F., Waldherr, R., Kutt, R., and Brandis, M. (1992). The nephronophthisis complex: clinical and genetic aspects. Clin Investig *70*, 802-808.

Hoskins, B. E., Thorn, A., Scambler, P. J., and Beales, P. L. (2003). Evaluation of multiplex capillary heteroduplex analysis: a rapid and sensitive mutation screening technique. Hum Mutat *22*, 151-157.

Iannaccone, A., Vingolo, E. M., Rispoli, E., De Propris, G., Tanzilli, P., and Pannarale, M. R. (1996). Electroretinographic alterations in the Laurence-Moon-Bardet-Biedl phenotype. Acta Ophthalmol Scand *74*, 8-13.

Igarashi, P., and Somlo, S. (2002). Genetics and pathogenesis of polycystic kidney disease. J Am Soc Nephrol *13*, 2384-2398.

James, P., Halladay, J., and Craig, E. A. (1996). Genomic libraries and a host strain designed for highly efficient two-hybrid selection in yeast. Genetics *144*, 1425-1436.

Katsanis, N. (2004). The oligogenic properties of Bardet-Biedl syndrome. Hum Mol Genet *13 Suppl 1*, R65-71.

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. (2001a). Triallelic inheritance in Bardet-Biedl syndrome, a Mendelian recessive disorder. Science *293*, 2256-2259.

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., 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.

Katsanis, N., Lewis, R. A., Stockton, D. W., Mai, P. M., Baird, L., Beales, P. L., Leppert, M., and Lupski, J. R. (1999). Delineation of the critical interval of Bardet-Biedl syndrome 1 (BBS1) to a small region of 11q13, through linkage and haplotype analysis of 91 pedigrees. Am J Hum Genet *65*, 1672-1679.

Katsanis, N., Lupski, J. R., and Beales, P. L. (2001b). Exploring the molecular basis of Bardet-Biedl syndrome. Hum Mol Genet *10*, 2293-2299.

Kiess, W., Galler, A., Reich, A., Muller, G., Kapellen, T., Deutscher, J., Raile, K., and Kratzsch, J. (2001). Clinical aspects of obesity in childhood and adolescence. Obes Rev *2*, 29-36.

Kim, J. C., Badano, J. L., Sibold, S., Esmail, M. A., Hill, J., Hoskins, B. E., Leitch, C. C., Venner, K., Ansley, S. J., Ross, A. J., *et al.* (2004). The Bardet-Biedl protein BBS4 targets cargo to the pericentriolar region and is required for microtubule anchoring and cell cycle progression. Nat Genet *36*, 462-470.

Klein, D., and Ammann, F. (1969). The syndrome of Laurence-Moon-Bardet-Biedl and allied diseases in Switzerland. Clinical, genetic and epidemiological studies. J Neurol Sci *9*, 479-513.

Kobrin, J. L., Ternand, C. L., Knobloch, W. H., and Johnson, D. D. (1990). Dental abnormalities as a component of the Laurence-Moon-Bardet-Biedl syndrome. Ophthalmic Paediatr Genet *11*, 299-303.

Kopelman, P. (1994). Investigation of obesity. Clinical Endocrinology *41*, 703-708.

Krawczak, M., Ball, E. V., and Cooper, D. N. (1998). Neighbouring-nucleotide effects on the rates of germ-line single-base-pair substitution in human genes. Am J Hum Genet *63*, 474-488.

Krege, J. H., Hodgin, J. B., Couse, J. F., Enmark, E., Warner, M., Mahler, J. F., Sar, M., Korach, K. S., Gustafsson, J., and Smithies, O. (1998). Generation and reproductive phenotypes of mice lacking estrogen receptor ß. Proc Natl Acad Sci *95*, 15677-15682.

Kubo, A., Sasaki, H., Yuba-Kubo, A., Tsukita, S., and Shiina, N. (1999). Centriolar satellites: molecular characterization, ATP-dependent movement toward centrioles and possible involvement in ciliogenesis. J Cell Biol *147*, 969-979.

Kulaga, H. M., Leitch, C. C., Eichers, E. R., Badano, J. L., Lesemann, A., Hoskins, B. E., Lupski, J. R., Beales, P. L., Reed, R. R., and Katsanis, N. (2004). Loss of BBS proteins cause anosmia in humans and defects in olfactory cilia structure and function in the mouse. Nat Genet *36*, 994-998.

Kwitek-Black, A. E., Carmi, R., Duyk, G. M., Buetow, K. H., Elbedour, K., Parvari, R., Yandava, C. N., Stone, E. M., and Sheffield, V. C. (1993). Linkage of Bardet-Biedl syndrome to chromosome 16q and evidence for non-allelic genetic heterogeneity. Nat Genet *5*, 392-396.

Lanz, R. B., Razani, B., Goldberg, A. D., and O'Malley, B. W. (2002). Distinct RNA motifs are important for coactivation of steroid receptors by steroid receptor RNA activator (SRA). Proc Natl Acad Sci *99*, 16081-16086.

Laurence, J. Z., and Moon, R. C. (1866). Four cases of 'retinitis pigmentosa' occurring in the same family, and accompanied by general imperfections of development. Ophthalmical Review *2*, 32-41.

Lee, C. H., Chinpaisal, C., and Wei, L. J. (1998). Cloning and characterisation of mouse RIP140, a corepressor for nuclear orphan receptor TR2. Mol Cell Biol *18*, 6745-6755.

Lee, C. H., and Wei, L. N. (1999). Characterization of receptor-interacting protein 140 in retinoid receptor activities. J Biol Chem *274*, 31320-31326.

Leppert, M., Baird, L., Anderson, K. L., Otterud, B., Lupski, J. R., and Lewis, R. A. (1994). Bardet-Biedl syndrome is linked to DNA markers on chromosome 11q and is genetically heterogeneous. Nat Genet *7*, 108-112.

Li, J. B., Gerdes, J. M., Haycraft, C. J., Fan, Y., Teslovich, T. M., May-Simera, H., Li, H., Blacque, O. E., Li, L., Leitch, C. C., *et al.* (2004). Comparative genomics identifies a flagellar and basal body proteome that includes the BBS5 human disease gene. Cell *117*, 541-552.

Løken, A. C., Hanssen, O., Halvorsen, S., and Jolster, N. J. (1961). Hereditary renal dysplasia and blindness. Acta Paediatr *50*, 177-184.

Lorda-Sanchez, I., Ayuso, C., and Ibanez, A. (2000). Situs inversus and hirschsprung disease: two uncommon manifestations in Bardet-Biedl syndrome. Am J Med Genet *90*, 80-81.

Lorda-Sanchez, I., Ayuso, C., Sanz, R., and Ibanez, A. (2001). Does Bardet-Biedl syndrome have a characteristic face? J Med Genet *38*, E14.

Lubahn, D. B., Moyer, J. S., Golding, T. S., Couse, J. F., Korach, K. S., and Smithies, O. (1993). Alteration of reproductive function but not prenatal sexual development after insertional disruption of the mouse estrogen receptor gene. Proc Natl Acad Sci *90*, 11162-11166.

Manouvrier-Hanu, S., Moerman, A., and Lefevre, J. (1999). Bardet-Biedl syndrome with preaxial polydactyly. Am J Med Genet *84*, 75.

McKusick, V. A. (2002). History of medical genetics. In Principles and practice of medical genetics, D. L. Rimoin, J. M. Connor, R. E. Pyeritz, and B. R. Korf, eds. (Churchill Livingstone), pp. 3-36.

Mishriki, Y. Y. (2000). Can you name this congenital syndrome? Bardet-Biedl syndrome. Postgrad Med *108*, 207-209.

Mochizuki, T., Saijoh, Y., Tsuchiya, K., Shirayoshi, Y., Takai, S., Taya, C., Yonekawa, H., Yamada, K., Nihei, H., Nakatsuji, N., *et al.* (1998). Cloning of inv, a gene that controls left/right asymmetry and kidney development. Nature *395*, 177-181.

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.

Morgan, D., Eley, L., Sayer, J., Strachan, T., Yates, L. M., Craighead, A. S., and Goodship, J. A. (2002). Expression analysis and interaction with the anaphase promoting complex protein Apc2 suggest a role for inversin in primary cilia and involvement in the cell cycle. Hum Mol Genet *11*, 3345-3350.

Morton, N. E. (1955). Sequential tests for the detection of linkage. Am J Hum Genet *7*, 277-318.

Moyer, J. H., Lee-Tischler, M. J., Kwon, H. Y., Schrick, J. J., Avner, E. D., Sweeney, W. E., Godfrey, V. L., Cacheiro, J. E., Wilkinson, J. E., and Woychik, R. P. (1994). Candidate gene associated with a mutation causiing recessive polycystic kidney disease in mice. Science *264*, 1329-1333.

Musarella, M. A. (2001). Patient with Bardet-Biedl syndrome presenting with nystagmus at fifteen months of age. J Aapos *5*, 262-264.

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., Mullins, R. F., Andrews, M., Chiang, A. P., Swiderski, R. E., Yang, B., Braun, T., Casavant, T., Stone, E. M., and Sheffield, V. C. (2004). Bardet-Biedl syndrome type 4 (BBS4)-null mice implicate Bbs4 in flagella formation but not global cilia assembly. Proc Natl Acad Sci U S A *101*, 8664-8669.

Mykytyn, K., Nishimura, D. Y., Searby, C. C., Beck, G., Bugge, K., Haines, H. L., Cornier, A. S., Cox, G. F., Fulton, A. B., Carmi, R., *et al.* (2003). Evaluation of complex inheritance

involving the most common Bardet-Biedl syndrome locus (BBS1). Am J Hum Genet *72*, 429-437.

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.

Nagamine, C. M., Chan, K., and Lau, Y. F. (1989). A PCR artifact: generation of heteroduplexes. Am J Hum Genet *45*, 337-339.

Nakane, P. K., and Pierce, G. B. (1967). Enzyme-labeled antibodies for the light and electron microscopic localization of tissue antigens. J Cell Biol *33*, 307-318.

Nauli, S. M., Alenghat, F. J., Y., L., Williams, E., Vassilev, P., Li, X., Elia, A. E. H., Lu, W., Brown, E. M., Quinn, S. J., *et al.* (2003). Polycystins 1 and 2 mediate mechanosensation in the primary cilium of kidney cells. Nat Genet *33*, 129-137.

Ng, P. C., and Henikoff, S. (2002). Accounting for human polymorphisms predicted to affect protein function. Genome Research *12*, 436-446.

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.

Norden, G., Friman, S., Frisenette-Fich, C., Persson, H., and Karlberg, I. (1991). Renal transplantation in the Bardet-Biedl syndrome, a form of Laurence-Moon-Biedl syndrome. Nephrol Dial Transplant *6*, 982-983.

O'Dea, D., Parfrey, P. S., Harnett, J. D., Hefferton, D., Cramer, B. C., and Green, J. (1996). The importance of renal impairment in the natural history of Bardet-Biedl syndrome. Am J Kidney Dis *27*, 776-783.

Olbrich, H., Fliegauf, M., Hoefele, J., Kispert, A., Otto, E., Volz, A., Wolf, M. T., Sasmaz, G., Trauer, U., Reinhardt, R., *et al.* (2003). Mutations in a novel gene, NPHP3, cause adolescent nephronophthisis, tapeto-retinal degeneration and hepatic fibrosis. Nat Genet *34*, 455-459.

Ordentlich, P., Downes, M., Xie, W., Genin, A., Spinner, N. B., and Evans, R. M. (1999). Unique forms of human and mouse nuclear receptor corepressor SMRT. Proc Natl Acad Sci *96*, 2639-2644.

Orkin, S. H. (1986). Reverse genetics and human disease. Cell *47*, 845-850.

Otto, E., Betz, R., Rensing, C., Schatzle, S., Kuntzen, T., Vetsi, T., Imm, A., and Hildebrandt, F. (2000). A deletion distinct from the classical homologous recombination of juvenile nephronophthisis type 1 (NPH1) allows exact molecular definition of deletion breakpoints. Hum Mutat *16*, 211-223.

Otto, E., Hoefele, J., Ruf, R., Mueller, A. M., Hiller, K. S., Wolf, M. T., Schuermann, M. J., Becker, A., Birkenhager, R., Sudbrak, R., *et al.* (2002). A gene mutated in nephronophthisis and retinitis pigmentosa encodes a novel protein, nephroretinin, conserved in evolution. Am J Hum Genet *71*, 1161-1167.

Otto, E. A., Schermer, B., Obara, T., O'Toole, J. F., Hiller, K. S., Mueller, A. M., Ruf, R. G., Hoefele, J., Beekmann, F., Landau, D., *et al.* (2003). Mutations in INVS encoding inversin cause nephronophthisis type 2, linking renal cystic disease to the function of primary cilia and left-right axis determination. Nat Genet *34*, 413-420.

Parfrey, P. S., Davidson, W. S., and Green, J. S. (2002). Clinical and genetic epidemiology of inherited renal disease in Newfoundland. Kidney Int *61*, 1925-1934.

Pazour, G. J., Baker, S. A., Deane, J. A., Cole, D. G., Dickert, B. L., Rosenbaum, J. L., Witman, and Besharse, J. G. (2002). The intraflagellar transport protein IFT88 is essential for vertebrate photoreceptor assembly and maintenance. J Cell Biol *157*, 103-113.

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.

Riise, R., Andreasson, S., Borgastrom, M. K., Wright, A. F., Tommerup, N., Rosenberg, T., and Tornqvist, K. (1997). Intrafamilial variation of the phenotype in Bardet-Biedl syndrome. Br J Ophthalmol *81*, 378-385.

Rivolta, C., Sharon, D., DeAngelis, M. M., and Dryja, T. P. (2002). Retinitis pigmentosa and allied diseases: numerous diseases, genes, and inheritance patterns. Hum Mol Genet *11*, 1219-1227.

Rosenbaum, J. L., and Witman, G. B. (2002). Intraflagellar transport. Nature Mol Cell Biol *3*, 813-825.

Rozycka, M., Collins, N., Stratton, M. R., and Wooster, R. (2000). Rapid detection of DNA sequence variants by conformation-sensitive capillary electrophoresis. Genomics *70*, 34-40.

Ruddle, F. H. (1984). Reverse genetics and beyond. Am J Hum Genet *36*, 944-953.

Schaap, C., ten Tusscher, M. P., Schrander, J. J., Kuijten, R. H., and Schrander-Stumpel, C. T. (1998). Phenotypic overlap between McKusick-Kaufman and Bardet-Biedl syndromes: are they related? Eur J Pediatr *157*, 170-171.

Schachat, A. P., and Maumenee, I. H. (1982). Bardet-Biedl syndrome and related disorders. Arch Ophthalmol *100*, 285-288.

Schultz, J., Copley, R. R., Doerks, T., Ponting, C. P., and Bork, P. (2000). SMART: a web-based tool for the study of genetically mobile domains. Nucleic Acids Res *28*, 231-234.

Senior, B., Friedmann, A. I., and Braudo, J. L. (1961). Juvenile familial nephropathy with tapetoretinal degeneration: a new oculo-renal dystrophy. Am J Ophthalmol *52*, 625-633.

Sheffield, V. C. (2004). Use of isolated populations in the study of a human obesity syndrome, the Bardet-Biedl syndrome. Pediatr Res *55*, 908-911.

Sheffield, V. C., Carmi, R., Kwitek-Black, A., Rokhlina, T., Nishimura, D., Duyk, G. M., Elbedour, K., Sunden, S. L., and Stone, E. M. (1994). Identification of a Bardet-Biedl syndrome locus on chromosome 3 and evaluation of an efficient approach to homozygosity mapping. Hum Mol Genet *3*, 1331-1335.

Sheffield, V. C., Nishimura, D., and Stone, E. M. (2001). The molecular genetics of Bardet-Biedl syndrome. Curr Opin Genet Dev *11*, 317-321.

Shi, Y., Downes, M., Xie, W., Kao, H. Y., Ordentlich, P., Tsai, C. C., Hon, M., and Evans, R. M. (2001). Sharp, an inducible cofactor that integrates nuclear receptor repression and activation. Genes Dev *15*, 1140-1151.

Slavotinek, A. M., and Biesecker, L. G. (2000). Phenotypic overlap of McKusick-Kaufman syndrome with bardet-biedl syndrome: a literature review. Am J Med Genet *95*, 208-215.

Slavotinek, A. M., Searby, C., Al-Gazali, L., Hennekam, R. C., Schrander-Stumpel, C., Orcana-Losa, M., Pardo-Reoyo, S., Cantani, A., Kumar, D., Capellini, Q., *et al.* (2002). Mutation analysis of the MKKS gene in McKusick-Kaufman syndrome and selected Bardet-Biedl syndrome patients. Hum Genet *110*, 561-567.

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.

Snell, W. J., Pan, J., and Wang, Q. (2004). Cilia and flagella revealed: from flagellar assembly in Chlamydomonas to human obesity disorders. Cell *117*, 693-697.

Solis-Cohen, S., and Weiss, E. (1925). Dystrophia adiposogenitalis, with atypical retinitis pigmentosa and mental deficiency. The Laurence-Biedl syndrome: a report of four cases in one family. Am J Med Sci *169*, 489-505.

Stone, D. L., Slavotinek, A. M., Bouffard, G. G., Banerjee-Basu, S., Baxevanis, A. D., Barr, M., and Biesecker, L. G. (2000). Mutation of a gene encoding a putative chaperonin causes McKusick-Kaufman syndrome. Nat Genet *25*, 79-82.

Strachan, T., and Read, A. P. (1999). Human molecular genetics 2 (Oxford, BIOS scientific publishers Ltd).

Tayel, S. M., Al-Naggar, R. L., Krishna Murthy, D. S., Naguib, K. K., and Al-Awadi, S. A. (1999). Familial pericentric inversion of chromosome 1 (p36.3q23) and Bardet-Biedl syndrome. J Med Genet *36*, 418-419.

Toby, G. G., and Golemis, E. A. (2001). Using the yeast interaction trap and other two-hybrid-based approaches to study protein-protein interactions. Methods *24*, 201-217.

Vaughan, K. T., and Vallee, R. B. (1995). Cytoplasmic dynein binds dynactin through a direct interaction between the intermediate chains and $p150^{Glued}$. J Cell Biol *131*, 1507-1516.

Waldherr, R., Lennert, T., Weber, H. P., Fodisch, H. J., and Scharer, K. (1982). The nephronophthisis complex. A clinicopathologic study in children. Virchows Arch A Pathol Anat Histol *394*, 235-254.

Waterman-Storer, C. M., Karki, S., and Holzbaur, E. L. F. (1995). The $p150^{Glued}$ component of the dynactin complex binds to both microtubules and the actin-related protein centractin (Arp-1). Proc Natl Acad Sci *92*, 1634-1638.

Watnick, T., and Germino, G. (2003). From cilia to cyst. Nat Genet *34*, 355-356.

Wei, L. N., Farooqui, M., and Hu, X. (2001). Ligand-dependent formation of retinoid receptors, receptor-interacting protein 140 (RIP140), and histone deacetylase complex is mediated by a novel receptor-interacting motif of RIP140. J Biol Chem *276*, 16107-16112.

Wei, L. N., Hu, X., Chandra, D., Seto, E., and Farooqui, M. (2000). Receptor-interacting protein 140 directly recruits histone deacetylases for gene silencing. J Biol Chem *275*, 40782-40787.

White, R., Leonardsson, G., Rosewell, I., Jacobs, M. A., Milligan, S., and Parker, M. (2000). The nuclear receptor co-repressor Nrip1 (RIP140) is essential for female fertility. Nature Medicine *6*, 1368-1373.

Woods, M. O., Young, T. L., Parfrey, P. S., Hefferton, D., Green, J. S., and Davidson, W. S. (1999). Genetic heterogeneity of Bardet-Biedl syndrome in a distinct Canadian population: evidence for a fifth locus. Genomics *55*, 2-9.

Young, T. L., Penney, L., Woods, M. O., Parfrey, P. S., Green, J. S., Hefferton, D., and Davidson, W. S. (1999a). A fifth locus for Bardet-Biedl syndrome maps to chromosome 2q31. Am J Hum Genet *64*, 900-904.

Young, T. L., Woods, M. O., Parfrey, P. S., Green, J. S., Hefferton, D., and Davidson, W. S. (1999b). A founder effect in the newfoundland population reduces the Bardet-Biedl syndrome I (BBS1) interval to 1 cM. Am J Hum Genet *65*, 1680-1687.

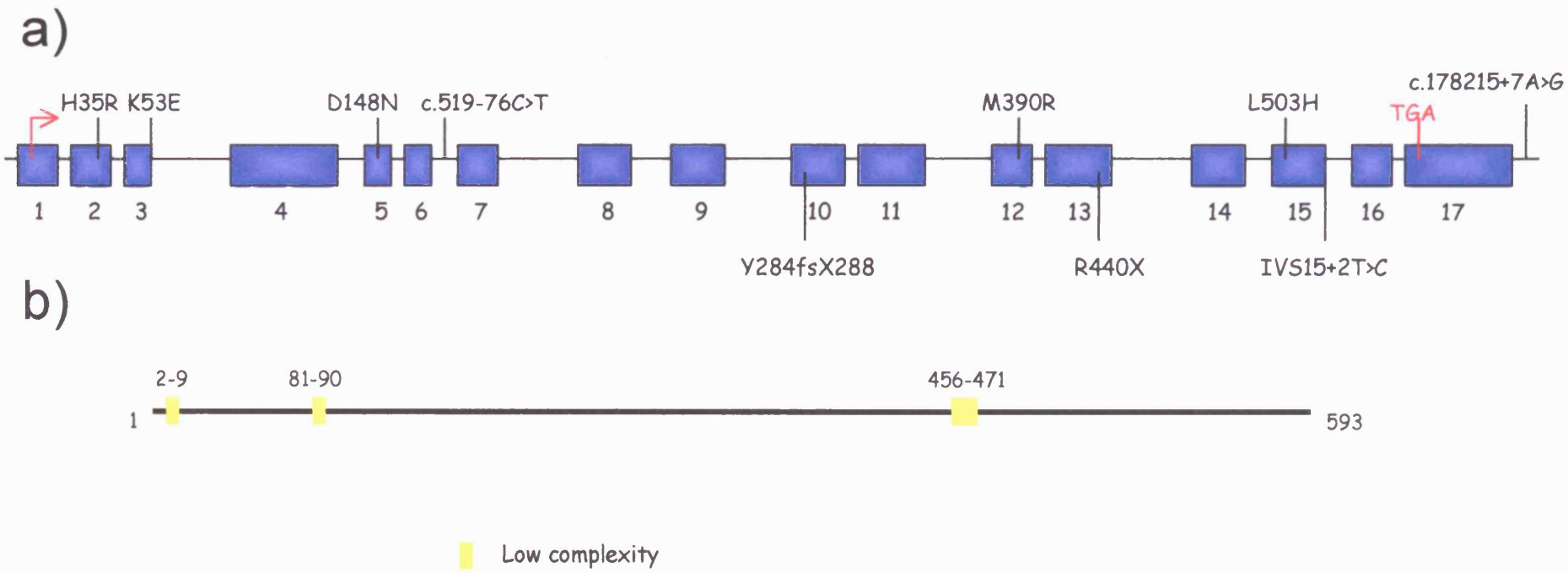
Young, T. L., Woods, M. O., Parfrey, P. S., Green, J. S., O'Leary, E., Hefferton, D., and Davidson, W. S. (1998). Canadian Bardet-Biedl syndrome family reduces the critical region of BBS3 (3p) and presents with a variable phenotype. Am J Med Genet *78*, 461-467.

Zhou, S., Fujimuro, M., Hsieh, J. D., Chen, L., Miyamoto, A., Weinmaster, G., and Hayward, S. D. (2000). SKIP, a CBF1-associated protein, interacts with the ankyrinrepeat domain of NotchIC to facilitate NotchIC function. Mol Cell Biol *20*, 2400-2410.

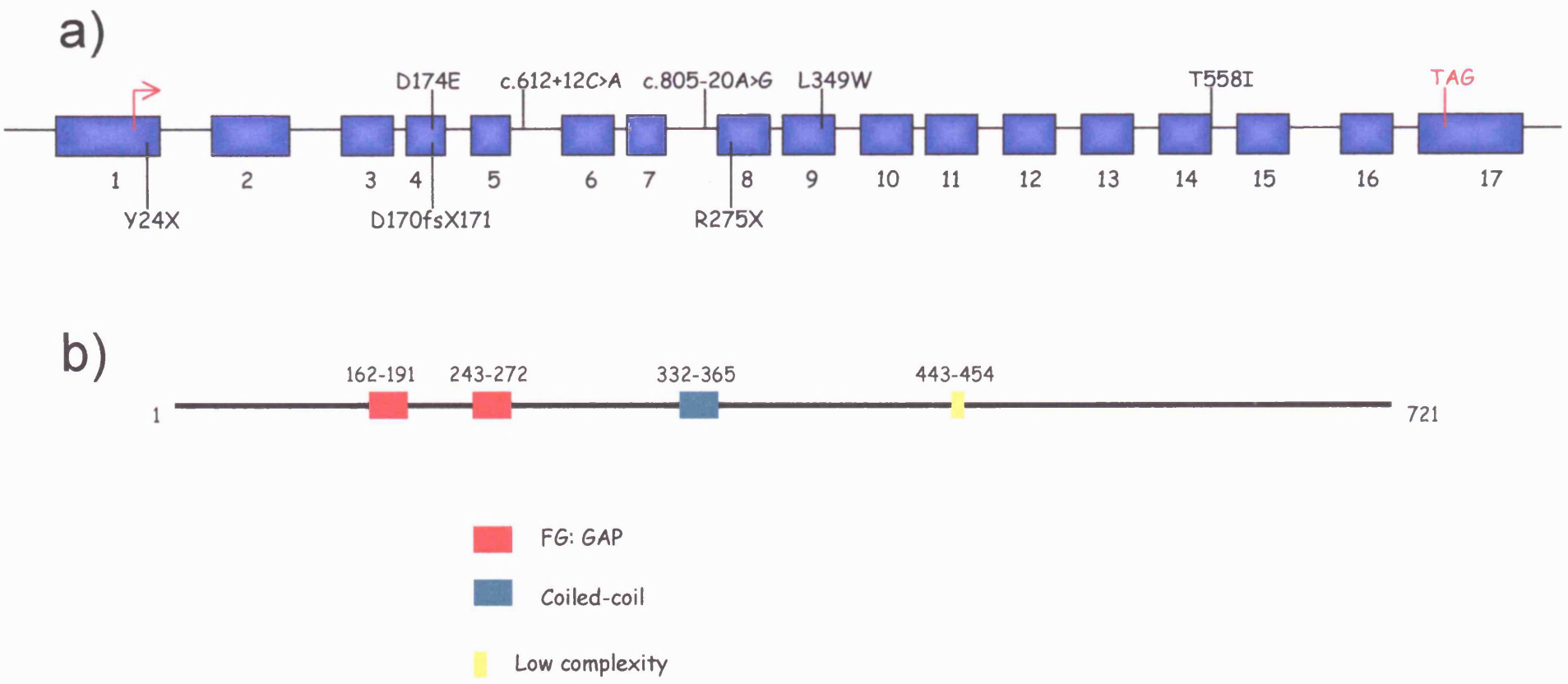
Zimmerman, W., and Doxsey, S. J. (2000). Construction of centrosomes and spindle poles by molecular motor-driven assembly of protein particles. Traffic *1*, 927-934.

Zimmerman, W., Sparks, C. A., and Doxsey, S. J. (1999). Amorphous no longer: the centrosome comes into focus. Curr Opin Cell Biol *11*, 122-128.

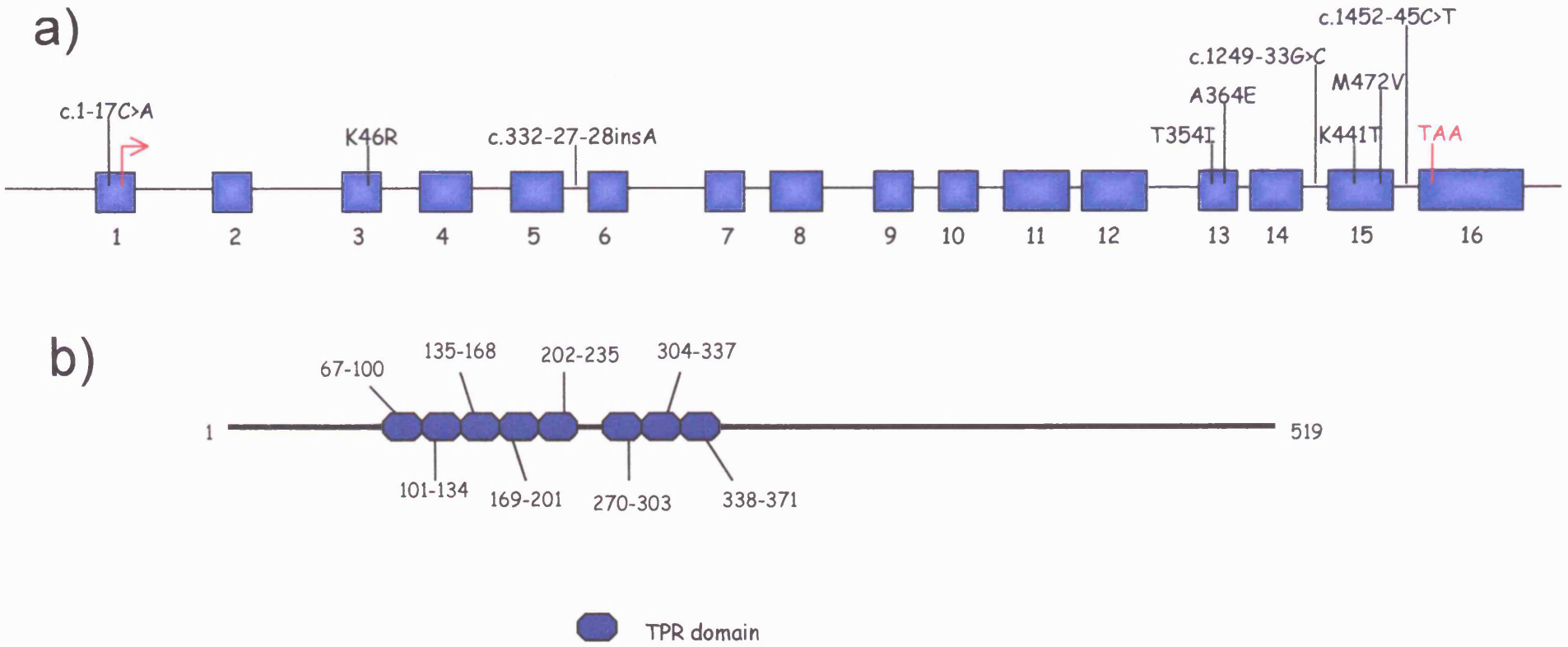
Zlotogora, J. (1998). Selection for carriers of recessive disease: a common phenomenon? Am J Med Genet *80*, 266-268.



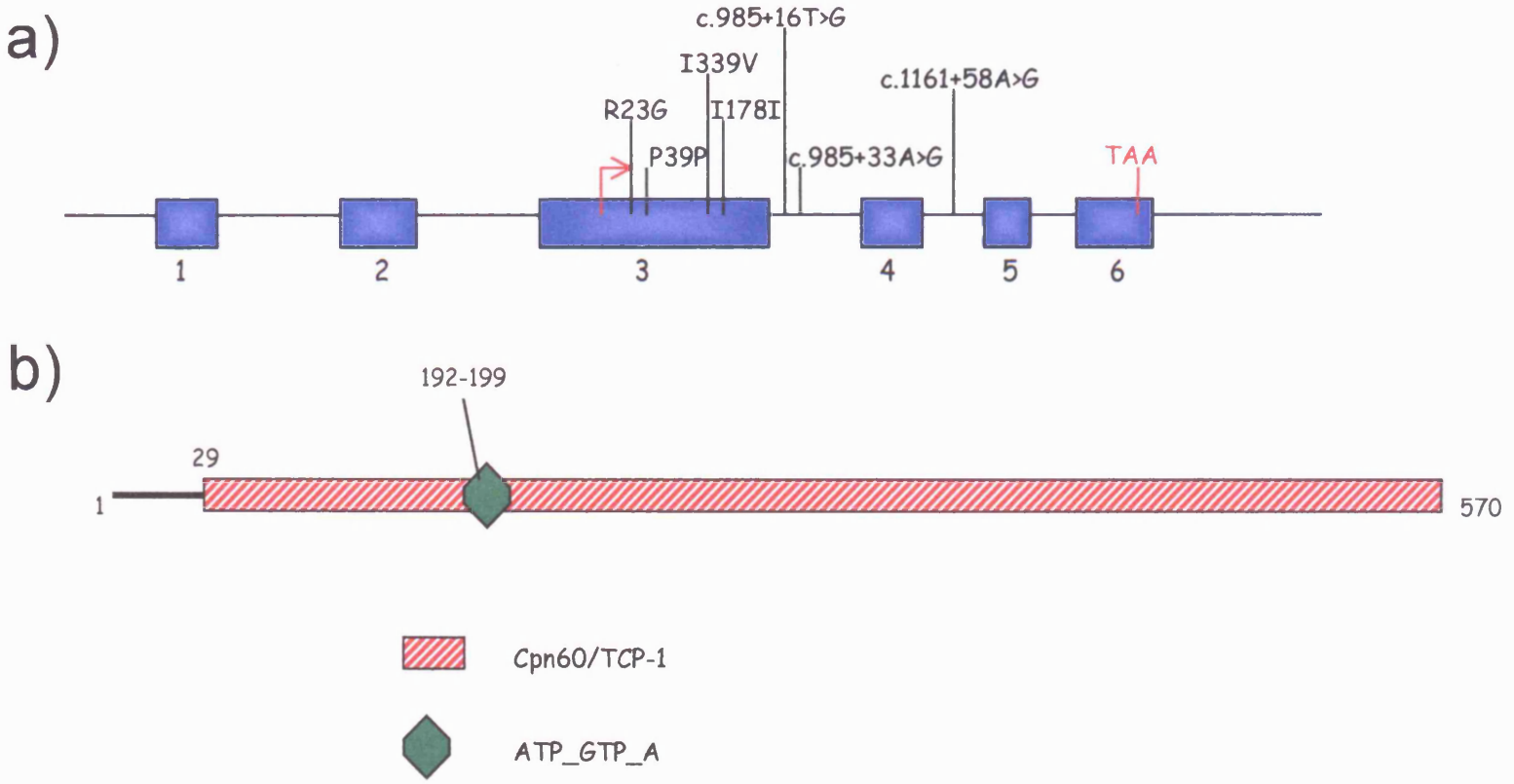
Appendix 1: a) Genomic organisation and position of mutations in *BBS1*. Missense and neutral variants are shown above the gene and deleterious mutations below. b) Primary protein structure of BBS1.



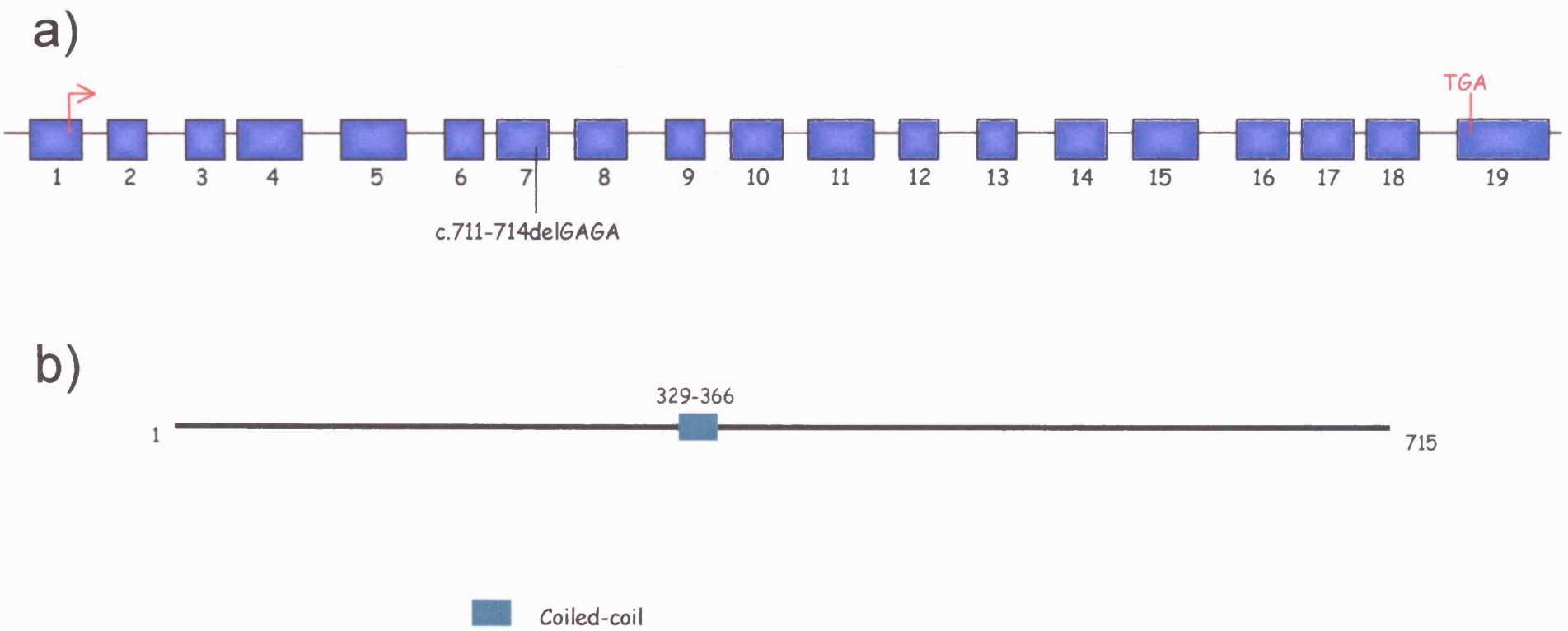
Appendix 2: a) Genomic organisation and position of mutations in *BBS2*. Missense and neutral variants are shown above the gene and deleterious mutations below. b) Primary protein structure of BBS2. The protein contains two FG:GAP repeats (Pfam Accession number: PF01839) which are predicted to fold into beta propeller structures and are commonly found at the N-terminus of integrin alpha chains, and a coiled-coil domain (Pfam Accession number: PF05710).



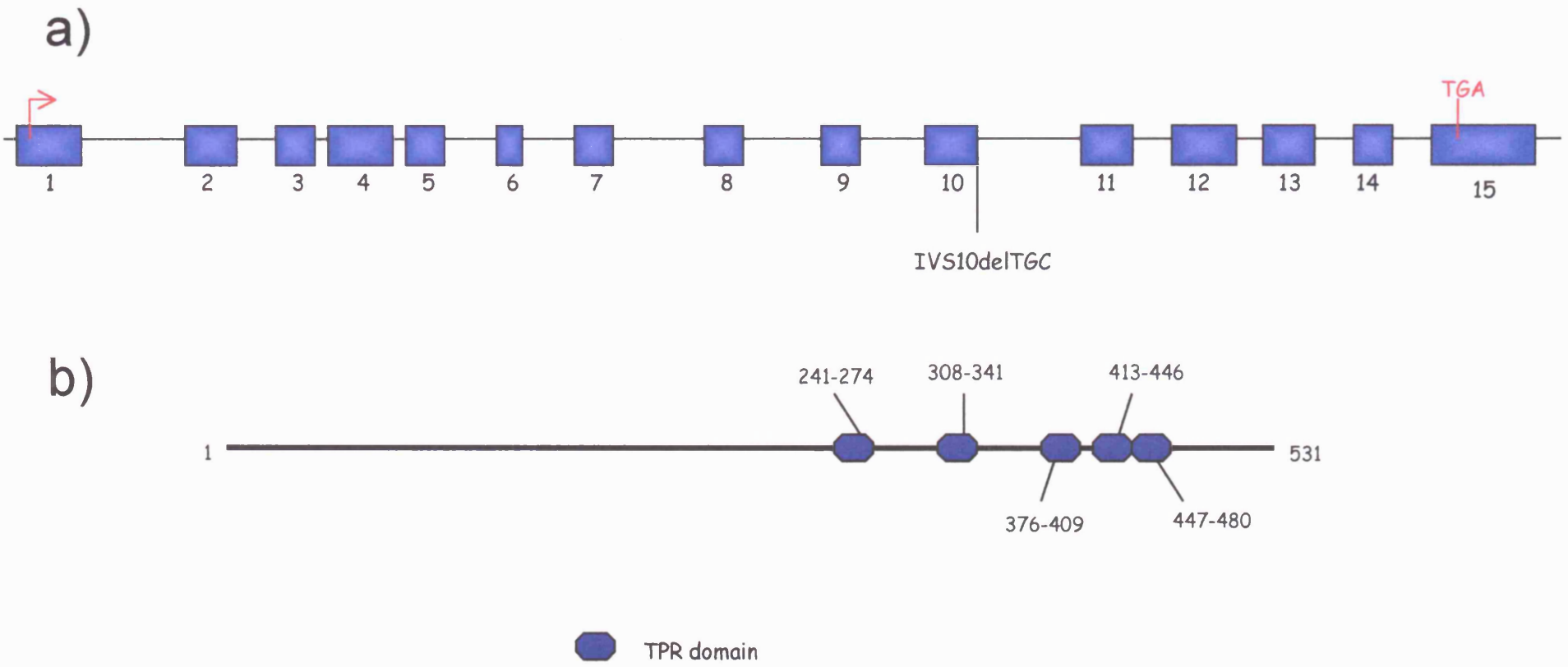
Appendix 3: a) Genomic organisation of *BBS4*. Missense and neutral variants are shown above the gene, no deleterious mutations were identified in *BBS4* in our patient cohort. b) Protein structure of BBS4. The protein contains a number of tandem TPR domains (Pfam Accession number: CL0020) which consist of a degenerate sequence of ~34 amino acids.



Appendix 4: a) Genomic organisation of BBS6. Missense and neutral variants are shown above the gene, no deleterious mutations in BBS6 were identified in our patient cohort. b) Protein structure of BBS6. The protein shows similarity to the Cpn60/TCP-1 chaperonin family (Pfam Accession number: PF00118). The 60kDa chaperonins (Cpn60) form a structure composed of two stacked rings, each made up of seven identical subunits. Mg^{2+}-ATP is required for assembly of the ring structures. The TCP-1 family of proteins act as molecular chaperones for tubulin, actin and probably some other proteins.



Appendix 5: a) Genomic organisation of BBS7. A single frameshift mutation was identified in *BBS7* in the patient cohort. b) The protein contains a coiled-coil domain in a region of shared homology with BBS2.



Appendix 6: a) Genomic organisation of *BBS8*. A single splice site mutation was identified in the *BBS8* gene. b) Protein structure of BBS8. Like BBS4, the BBS8 protein contains a number of TPR domains.

Primer	Sequence	Size (bp)	HA label	PCR
BBS1 x1F *BBS1* x1R	ATTCCTCAACCCAGGAAGGT GCGACCTCCAGACACACATA	401	TET	TD
BBS1 x2+3F *BBS1* x2+3R	CCTGGACTTGTACCCAGACG TTTCAGCCGTCAGGAAATCT	460	FAM	TD
BBS1 x4F *BBS1* x4R	TTGGTGCAGGAATGAATGAA ATAGCTGACCTCCTGCCACC	523	FAM	TD
BBS1 x5+6F *BBS1* x5+6R	GGAGGCAGAGACCAAGAGGT TCCATCATTCTGGCACATTC	417	FAM	TD
BBS1 x7F *BBS1* x7R	CCATCCAGTCACTCAGGTAAG TGGCTGGAAGGGATATAGCA	381	TET	TD
BBS1 x8F *BBS1* x8R	TCTTCTGTCACATCTCTGATATTTCC ACAAGGAACATATATATTACCCAGAAC	396	FAM	TD
BBS1 x9F *BBS1* x9R	GCTTTTGCTAAATGTTGCCC AAATTCCAGCCTTAAAGCCC	417	FAM	TD
BBS1 x10F *BBS1* x10R	GTTTTCCAAGGCCACACATT GAAAGAACGGTTTCTGGGGT	432	FAM	TD
BBS1 x11F *BBS1* x11R	GAGAGAGTCCTCTGGCTTCCC AAGGAGGAGTGAGTGGCAGA	323	TET	PCR71
BBS1 x12F *BBS1* x12R	TATTAGGAGGTTGACCCCA AGGCCTTACTTTCCACACCC	458	FAM	TD
BBS1 x13F *BBS1* x13R	CTCGTCTGGAAGACGGATGT TGGATTTGCAGAGGTGAGTG	460	TET	TD
BBS1 x14F *BBS1* x14R	GCAGACTCCTCCTGCAGCAC CTGAGTTGGCTTGGAATTGG	408	TET	TD
BBS1 x15F *BBS1* x15R	TTGAGTAGGAGGAGGGGACA CTGGTCTGTGGTGGGAGAGT	431	FAM	TD
BBS1 x16F *BBS1* x16R	TGGGAGTCCAGAGGTCTAGG CTCCACTGTGCAGATACCC	457	FAM	TD
BBS1 x17F *BBS1* x17R	GACATCATCAAGGTAGGCCC GGGTGCTTAAGAGAGGGGAG	472	TET	TD

Appendix 7: *BBS1* primer sequences including amplicon size, colour of MCHA label and PCR conditions. M13 tags (Forward – TGTAAAACGACGGCCAGT, Reverse – CAGGAAACAGCTATGACC) were added to all primer pairs to improve stability.

Primer	Sequence	Size (bp)	HA label	PCR
BBS2 x1F *BBS2* x1R	GCGTGAGGCCAGCTCCGCTGC GCGCGGCCGGCGGAGATCCTG	254	FAM	FAILSAFE - H
BBS2 x2F *BBS2* x2R	TTTTAAGGGAATGTAATTAGT TGGACATTAATGAGTAATGAC	325	TET	TD
BBS2 x3F *BBS2* x3R	CTGTTTTACTCAAAATCTGCTCAG TTAGCTACATGAAGGAGAGGATTAC	409	FAM	PCR61
BBS2 x4F *BBS2* x4R	AATCCTCTCCTTCATGTAGCT GGAGAAGCTTACACTTCTGTC	239	FAM	TD
BBS2 x5F *BBS2* x5R	AGAAGCAGCATGCAAAGTACT TCATCTGACAGTACTGATCTA	240	FAM	TD
BBS2 x6F *BBS2* x6R	TATAAAGCCGTACTTGACAGT CAATAACTATCAAGCGCCTGA	222	FAM	TD
BBS2 x7F *BBS2* x7R	TATTGTGAGACTTCTGTGCTA TGTTACTGTTCTAAGTCCTAC	213	TET	TD
BBS2 x8F *BBS2* x8R	AGAATACTCTTGAAAACTGCT ATCTCGGTACAAATACTTCAG	349	FAM	TD
BBS2 x9F *BBS2* x9R	TAAGAGCAGGTAATTGATGAC CCCTGGCAATGACACTCTCAT	308	FAM	TD
BBS2 x10F *BBS2* x10R	GGCTCTGTCTTTTGAAGCTGA CCAAGACAGAGGAAGACTCTG	338	FAM	TD
BBS2 x11F *BBS2* x11R	ACCTCCTGACCTCGTGATCTG CCCCAAGAATCCACTGGGCAT	344	FAM	TD
BBS2 x12F *BBS2* x12R	CCTTAAATATCAATTGATGAC ACTGCTACCAATATAACACAT	344	TET	TD
BBS2 x13F *BBS2* x13R	GAATGTTACTTAAGAGCATAG CTGAATGGTAAACACCACATG	481	FAM	TD
BBS2 x14F *BBS2* x14R	GCTAAGTTTGTCTAACATCTG ACATAAGTACATTTGTAGTAC	281	FAM	PCR55
BBS2 x15F *BBS2* x15R	TTAATTGGTATAAGCGAACAG TTATACTTCTATTGGTAACAT	330	TET	TD
BBS2 x16F *BBS2* x16R	TAAGCTTGCCATATCAACATG ATATGAATTATTGGATGCTAC	291	TET	TD
BBS2 x17F *BBS2* x17R	TTGTTTTAAAACTGACGTCTA ATTCAGCAACAGTACTACTAC	389	FAM	TD

Appendix 8: *BBS2* primer sequences. FAILSAFE – H corresponds to the FAILSAFE PCR program and enzyme mix with buffer H.

Primer	Sequence	Size (bp)	HA label	PCR
BBS4 x1F *BBS4* x1R	GAGCAGACCCTAATCCTCCC CAGTTCCCGGGACAGTAAAA	343	TET	TD
BBS4 x2F *BBS4* x2R	TTGCATAATTGGTGAGCTACTGA AGGTGGCAGTGAGCCAAGAT	300	FAM	TD
BBS4 x3F *BBS4* x3R	TGTGATATTGCAGTATGTTTATGGT TCACTACCATAGCAAAAGATAACCAG	343	FAM	TD
BBS4 x4F *BBS4* x4R	CATAATCTGCCTGCCTTGGT TCACACAATGACAAAATTGCC	371	FAM	TD
BBS4 x5F *BBS4* x5R	TGTCAGCAGTTTTGGTTGTTG ACCATTCCCCCACTTGTGTA	417	FAM	PCR63
BBS4 x6F *BBS4* x6R	GCAGCTTCACTGACCAAACC CAGCAATGCCTGCAAAGTTA	413	TET	PCR63
BBS4 x7F *BBS4* x7R	AAAAAGCTGACTGTAATGCATAGTTT CAAGCAGAAAACAACAGATGAA	376	HEX	PCR63
BBS4 x8F *BBS4* x8R	GGCTGTTTGCTGAAATGTGA TCTACCTTATTGAAGCTACTGGGA	417	FAM	PCR63
BBS4 x9F *BBS4* x9R	CACCAGGGTTGAGATGACCT GCTTCCAATTTAAATGCAGAAAA	418	HEX	PCR63
BBS4 x10F *BBS4* x10R	ATGTTGGTCAGGCTGGTCTC TTTCAAATAAGAAGAGCCACCTTT	389	FAM	TD
BBS4 x11F *BBS4* x11R	CTGATGGGCCTGCTGAGTAT TGCATGGTCAAATGGACAAG	434	FAM	PCR63
BBS4 x12F *BBS4* x12R	TGGAATGTGTTTCTTTGGCA CAGGATCTCTAAGGGAATGGC	420	FAM	PCR63
BBS4 x13F *BBS4* x13R	GGATGCATAGAACCTGGCAACTG CACCTCATGGCTTTGGAGAGTCC	318	FAM	PCR71
BBS4 x14F *BBS4* x14R	TTAACCAGTTTTGTTTTGTTTTTGTG TGGATTTGGATGATCTGGGCTTG	321	HEX	TD
BBS4 x15F *BBS4* x15R	TTTGATAAGTACTTCCTGCCTCAA TCCCCTTGTGGCCAATACTA	473	FAM	TD
BBS4 x16F *BBS4* x16R	GGCAAACTTGACTGTTGCTTT CTTAGGCTCAACTGCTGGCT	429	TET	TD

Appendix 9: *BBS4* primer sequences, amplicon sizes, MCHA label and conditions.

Primer	Sequence	Size (bp)	HA label	PCR
BBS6 x3aF	GATTTTATAGCCACAATGCT	492	FAM	TD
BBS6 x3aR	ATGACAGTGGTGGGTGTCAA			
BBS6 x3bF	TCTGGTGAGCATACAGGCAG	498	FAM	TD
BBS6 x3bR	CGTTTGGAAGCTAAGAAGCC			
BBS6 x3cF	GATCCTCCTTTGTTTGGTGC	392	FAM	TD
BBS6 x3cR	GGTTAAGCAGCTGGTCCAAG			
BBS6 x3dF	AATCAACTGCCCTCAAGGTG	421	HEX	TD
BBS6 x3dR	CCTTTGCTGCCAGAAATGAT			
BBS6 x4F	ATGCTTGTGGGGCTTTTATG	475	HEX	TD
BBS6 x4R	AATGGCAACACATGCCAAAT			
BBS6 x5F	GCACCACACAAGTTTTGTTC	378	TET	TD
BBS6 x5R	CCTATACATGCACCCCTGAA			
BBS6 x6aF	GTGCCAGACCCCAAATTAAA	391	TET	TD
BBS6 x6aR	CCAGTTGAGTTCTTCCTGGC			
BBS6 x6bF	GGCAGATTCTCCCTGTGTTG	447	TET	TD
BBS6 x6bR	GCATTTCCATTCACGAATCA			

Appendix 10: *BBS6* primer sequences, amplicon sizes, MCHA label and conditions.

Primer	Sequence	PCR
BBS7 x1F	GCCCTATCCCTTGGGTTT	TD
BBS7 x1R	GTCTGGGGTCTCTGTGGA	
BBS7 x2F	TGTAAAACGACGGCCAGTGCCAGTCATCTTACACAAC	TD
BBS7 x2R	AGGAAACAGCTATGACCAAATGTCCCTTGGTATTCCAG	
BBS7 x3F	TTACTTTTGTGCTACCCG	TD
BBS7 x3R	GGAACAAAATCTATGGCC	
BBS7 x4F	TTCCAGAAAGCCTATTAA	TD
BBS7 x4R	AAAAACCTGAAGACCTGC	
BBS7 x5F	TTCAGCTTTCAAAATCAA	TD
BBS7 x5R	TTTCCACATGTTTATAAA	
BBS7 x6F	CGTGCTGTTAGTTACTGG	TD
BBS7 x6R	TGCAAAATTGCTAACAAA	
BBS7 x7F *	AGATAAATCAAGGTGTGA	TD
BBS7 x7R	AATGGGGAAATGTCTTAT	
BBS7 x8F	TCCTGCTTAAAGGCAAGA	TD
BBS7 x8R	TTGCTCACCATTCTGAGT	
BBS7 x9F	TGAGGTTTGAGGCTTCCA	TD
BBS7 x9R	TTCCTGCCATTTGTTCAA	
BBS7 x10F	CAGCACTTACGCTAATTT	TD
BBS7 x10R	GATATTGGTGGACAAAGG	
BBS7 x11F	AAAACGACGGCCAGTACTCCAGTCTAGGTGACAGAGT	TD
BBS7 x11R	AAACAGCTATGACCCATTTTGTAAAAGCAAATG	
BBS7 x12F	CCGACACAGATTTTGAAG	TD
BBS7 x12R	TAAGGGGGTGGTGAGAGA	
BBS7 x13F	GAGATTGTGGTGTGGGCT	TD
BBS7 x13R	CATGTTTGAAAAGCGCTG	
BBS7 x14F	TCCAACTCAAACCAGCCT	TD
BBS7 x14R	GGAAACAGCTATGACCTGAATTGGAACAGAGTGGGAC	
BBS7 x15F	TCAATCAGATTACTCACA	TD
BBS7 x15R	AGCACAGGTGCAGGTATA	
BBS7 x16F	GTCCACGACGACACATGT	TD
BBS7 x16R	CAATTTCCTTCTACAGCT	
BBS7 x17F	ACGACGGCCAGTGTCTTTCAAGATGTGCAGGTT	TD
BBS7 x17R	GGAAACAGCTATGAAAAGGAAAATTAACCTC	
BBS7 x18F	TAAAACGACGGCCAACTGATTCATGACTGGTTCA	TD
BBS7 x18R	GAAACAGCTATGACCACATGTGTCGTCGTGGACTGG	

Appendix 11: Primer sequences and conditions for *BBS7*. For MCHA, *BBS7*x7F was labelled with FAM, resulting in a 468bp labelled fragment.

Primer	Sequence	PCR
BBS8 x1F *BBS8* x1R	AGTTCTGCTTGCGGTTGTTT CGCTAGTCACAGCTCCACAA	TD
BBS8 x2F *BBS8* x2R	AGTTTTTGACATGGCCCTTTA TTTTTCATCATTTTTGTTTCATGTC	TD
BBS8 x3F *BBS8* x3R	CATCTCCCTAAAATACATTTCTTGC AATGGGTCTTCCAGCTTGTG	TD
BBS8 x4F *BBS8* x4R	ACGTCTTTGAAACTCCCTGG AATCACCTTCAAAAATCTTCCA	TD
BBS8 x5F *BBS8* x5R	TCCCCTGTTGTTATAGTGGAGAA TTCTGACTGACCTTCCTTAACTCA	TD
BBS8 x6F *BBS8* x6R	TGAGCATTCCAGTTTGTATTCTG CAGAAACTGAGGGTGGATTTT	TD
BBS8 x7F *BBS8* x7R	GTAAGAAGGCCAGTGTGGCT TGAAGAAAAGGATTTTGCTGCT	TD
BBS8 x8F *BBS8* x8R	CCATCTGGAAACATGAGCAA TTCTTATTGCTACCCTGTGATTG	TD
BBS8 x9F *BBS8* x9R	AACATGTTAATTTATGTGTATGTGCAA TCCCTTGCTGTAAAGAAAAACA	TD
BBS8 x10F *BBS8* x10R	CTCCCAAAATGCTGGGAATA TGTTCCCAAAACAAAGAAAAGA	TD
BBS8 x11F *BBS8* x11R	TTCATATTGTATCCCCAGGGTC TCAGTGTTGACATTTGAGTTTCAT	TD
BBS8 x12F *BBS8* x12R	GGAGTATCAAAAATCACAAGATGAA AAGCAATCCTCCTGCTTCAG	TD
BBS8 x13F *BBS8* x13R	TCCACTTACGTAGAATTCACATTG CAACAGAGCTGGGATTCAAA	TD
BBS8 x14F *BBS8* x14R	TCATGAGGAATGTTGTCGGT TGCTGTCCCTTGAAGTAAAGA	TD
BBS8 x15F *BBS8* x15R	TTTGAGTTATGATGTTAGTTGTGGG TGCATTACACATTCTGTGTCACC	TD

Appendix 12: *BBS8* primer sequences and conditions. M13 tags were added to all primers for increased stability.