BARDET-BIEDL SYNDROME: MOLECULAR GENETICS OF A RARE RECESSIVE DISORDER

By

Ashvinder Kaur Bhogal B.Sc.

THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF
MASTER OF SCIENCE

In the Department of
Molecular Biology and Biochemistry

SIMON FRASER UNIVERSITY

September 2003

APPROVAL

Name: **Ashvinder Kaur Bhogal**

Degree: **Master of Science**

Title of Thesis: **Bardet-Biedl Syndrome: Molecular Genetics of a Rare Recessive Disorder**

Examining Committee:

Chair: **Dr. Peter J. Unrau**
Assistant Professor

Dr. William S. Davidson
Supervisor
Professor

Dr. David L. Baillie
Professor

Dr. Esther M. Verheyen
Assistant Professor

Dr. Nancy Hawkins
Internal Examiner
Assistant Professor
Department of Molecular Biology and Biochemistry
Simon Fraser University

Date Approved: September 9, 2003

PARTIAL COPYRIGHT LICENSE

I hereby grant to Simon Fraser University the right to lend my thesis, project or extended essay (the title of which is shown below) to users of the Simon Fraser University Library, and to make partial or single copies only for such users or in response to a request from the library of any other university, or other educational institution, on its own behalf or for one of its users. I further agree that permission for multiple copying of this work for scholarly purposes may be granted by me or the Dean of Graduate Studies. It is understood that copying or publication of this work for financial gain shall not be allowed without my written permission.

Title of Thesis/Project/Extended Essay:

Bardet-Biedl Syndrome: Molecular Genetics of a Rare Recessive Disorder

Author:

(signature)

Ashvinder Kaur Bhogal

(name)

Sept. 15, 2003

(date)

ABSTRACT

Bardet-Biedl Syndrome, or BBS, is a rare disorder whose cardinal manifestations are retinal dystrophy, dysmorphic extremities, renal structural abnormalities, obesity, and hypogenitalism in males. Eight BBS loci have been identified: BBS1 (11q13), BBS2 (16q21), BBS3 (3p12-13), BBS4 (15q123), BBS5 (2q31), BBS6 (20p12), BBS7 (4q27), and BBS8 (14q32.1), and genes for all of these loci have been identified except for BBS3 and BBS5. BBS has traditionally been modeled as an autosomal recessive disorder; however, there is evidence that BBS can be inherited in a triallelic fashion.

The frequency of BBS in the Newfoundland population is approximately ten fold higher than in Europe or North America. There are at least six genetic loci involved and a minimum of eight mutations in BBS genes in this population. The purpose of this thesis was to screen the Newfoundland BBS families for known mutations in the BBS6 gene to determine if triallelism is common in this population and to identify carriers and non-carriers of BBS6 mutations in specific families as part of a study to investigate the relationships between genotype and phenotype. Candidate genes for BBS3 and BBS5 were also identified and analyzed.

Twenty-one Newfoundland BBS families including four previously uncharacterized families (NF-B20, NF-B21, NF-B23, and NF-B25) were screened for the four known BBS6 mutations in this population: 429ΔCT433ΔAG (fs1), 280ΔT (fs2), L277P, and A242S. Affected individuals in NF-B20 and NF-B25 were

homozygous for fs1 and fs2, respectively. Only one family (NF-B14) showed a potential triallelic inheritance pattern (homozygous for BBS2 Y24X and heterozygous for BBS6 A242S).

Two genes, MYO3B and GORASP2, were selected as candidates for BBS5 based on their expression patterns and the function of their proteins. No mutations were found in the exons or exon-intron boundaries in these genes in an affected individual from NF-B9, a BBS5 family. The gene, GPR15, was predicted to be a candidate for BBS3 based on a bioinformatics approach which involved phylogenetic analysis and information derived from a comparison of BBS homologs in *C. elegans*. Sequencing the single predicted exon for this gene failed to identify any mutations in an affected individual in NF-B2, a BBS3 family.

DEDICATION

To Mom, Dad, and Jag.

ACKNOWLEDGEMENTS

An enormous thanks to Dr. Davidson (Willie) for allowing me to come for a co-op term, and inviting me to return to work on this thesis. Its amazing to think of how many doors have opened since then. I'd also like to thank everyone in the lab for making these two years fly by in the blink of an eye – they say that only happens when you are having fun, right?! So to Amit, Andrea, Carlo, Colin, Darren, Deanna, Evelyn, Laura, Leslie, Jon, Milly, Sarah, Siemon, and Yanli (Team Davidson) – thanks for everything. I know being here has changed me – professionally and personally – for the better.

A special thanks to the families, all the people working on this project at MUN, and to the Leroux lab for providing information that was incorporated into this thesis.

And since I can never thank my family enough for all they have done for me, for supporting me, for always being there – thanks Mom, Dad, and Jag.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF APPENDICES

Chapter 1: Introduction

1.1 Bardet-Biedl Syndrome and Related Disorders

Bardet-Biedl Syndrome, or BBS (OMIM 209900), is characterized by several phenotypic features including dystrophic extremities, obesity, renal structural abnormalities, male hypogenitalism, retinal dystrophy, and neurological deficits. Other features which are more variable include renal failure, learning difficulties, diabetes, and hypertension (Green *et al.*, 1989; Beales *et al.*, 1999; Moore *et al.*, 2003). BBS was first described by Bardet (1920) and Biedl (1922) more than fifty years after a similar disease was described by Laurence and Moon. Laurence and Moon (1866) described a family in which three brothers and one sister displayed obesity and blindness due to retinitis pigmentosa. The brothers walked with slouching gaits, and they also had hypogenitalism. This syndrome was named Laurence-Moon syndrome, or LMS (OMIM 245800).

Other diseases similar to BBS and LMS were described later, including Alstrom syndrome, or ALMS (OMIM 203800), and Biemond syndrome (OMIM 210350). Alstrom *et al.* (1959) described a disease in which patients had diabetes mellitus, nerve deafness, retinal dystrophy, obesity, and hypogonadism. The gene for ALMS was recently found on chromosome 2. This novel gene, ALMS1, encodes a 12.9 kb transcript coding for a protein of 4,169 amino acids (Hearn *et al.*, 2002; Collin *et al.*, 2002). Biemond syndrome is also phenotypically very similar to BBS with patients manifesting obesity, polydactyly,

hypogonadism, hydrocephalus, and iris coloboma. Very few patients have been studied with Biemond syndrome, and thus it has been difficult to locate the gene or genes causing the disease or make any genotypic links with any of the other similar syndromes. McKusick-Kaufman syndrome, or MKKS (OMIM 604896), though not as phenotypically similar to BBS as ALMS and Biemond syndrome, is caused by mutations in the same gene as BBS6 (Katsanis *et al.*, 2000). The major phenotypic features of MKKS that are shared with BBS are genital abnormalities and polydactyly. Though the two diseases have been found to be caused by mutations to the same gene, there has been no genotype-phenotype correlation established. This is largely due to MKKS being very rare in the general population.

Though in the past BBS and LMS were considered similar, they were classified as different syndromes. Moore *et al.* (2003) presented strong evidence that BBS and LMS are indeed the same disorder. Both diseases appear to have widespread systemic involvement and do not display significantly different phenotypes. A person is diagnosed with BBS if they display four of the five cardinal features (retinal dystrophy, obesity, renal abnormalities, hypogenitalism, and dystrophic extremities) or if an individual has three cardinal manifestations and has a sibling who has been diagnosed with BBS. A person affected with LMS would display very similar features, though a key difference would be the absence of polydactyly and the presence of paraplegia. Table 1.1 summarizes the differences noted between BBS, LMS, ALMS, Biemond syndrome, and MKKS.

Table 1.1 Comparison of the phenotypes of Bardet-Biedl Syndrome (BBS) (OMIM 209900), Laurence-Moon Syndrome (LMS) (OMIM 245800), Alstrom Syndrome (ALMS) (OMIM 203800), Biemond Syndrome (OMIM 210350), and McKusick-Kaufman Syndrome (MKKS) (OMIM 604896).

	BBS	LMS	Alstrom	Biemond	MKKS
Obesity	+	+	+	+	-
Retinal dystrophy	+	+	+	-	-
Diabetes mellitus	+	-	+	-	-
Hypogonadism/ Hypogenitalism/Genital Abnormalities	+	+	+	+	+
Polydactyly/Syndactyly/ Brachdactyly	+	-	-	+	+
Renal abnormalities	+	-	+	-	-
Paraplegia	-	+	-	-	-
Mental retardation	+	+	-	+	-
Iris coloboma	-	-	-	+	-

Moore *et al.* (2003) showed that LMS is indeed the same as BBS through clinical examination of patients. Some BBS patients were found to have paraplegia as was expected of LMS patients. In addition, examination of the molecular results of BBS patients in Newfoundland revealed that two patients who met the diagnostic criteria for LMS had the molecular genetic results indicative of BBS – thus implying that on a molecular basis, BBS and LMS are the same syndrome. It has been proposed that BBS and LMS be combined and called LMBBS.

Though eight BBS genes have been cloned and six have been identified, little is known about the molecular basis of LMBBS. This thesis recognizes that LMS and BBS are the same syndrome, but throughout BBS will be used to refer to LMBBS. The purpose of the thesis is to investigate the molecular basis of BBS.

1.2 Bardet-Biedl Syndrome Loci

1.2.1 BBS1

Leppert *et al.* (1994) identified a BBS locus by studying a cohort of 31 BBS families and, through a linkage study, found that in 17 of the families BBS was associated with a 26 cM region on chromosome 11q13 that included markers PYGM and D11S913. This work also revealed that there was more than one gene involved in BBS.

Young *et al.* (1999) used intrafamilial recombinations to decrease the critical region from 26 cM to 15 cM using six families from Newfoundland.

Linkage disequilibrium analysis was then used to define a homozygous region in the area surrounding the PYGM marker. The critical region was defined as 1 cM between markers D11S1883 and D11S4940. This result was consistent with a large scale family study by Katsanis *et al.* (1999) who narrowed the critical region to a 2.6 Mb interval on the basis of recombinations in several families. Using two consanguineous pedigrees, this interval was farther narrowed to a 1.8 Mb region based on loss of identity by descent. However, when the BBS1 gene was isolated, it was found to lie outside the critical region defined independently by these two groups.

Mykytyn *et al.* (2002) used several extended families to redefine a critical region for BBS1, which lay distal to the critical regions defined by Young *et al.* (1999) and Katsanis *et al.* (1999). Positional cloning was performed to predict which genes were located in this new region. The candidate genes were sequenced in a brute force manner and one gene which had slight similarity to the BBS2 gene, in the UniGene cluster Hs.54890, was sequenced in families suspected of carrying a BBS1 mutation. A total of four mutations were found in these families: G1655T, T1179G, G432+1A, and 851delA (Mykytyn *et al.*, 2002). The second mutation listed, T1179G, is the most common BBS1 mutation accounting for 80 per cent of all known cases. It causes a change in amino acid from a methionine to an arginine (M390R) at position 390. The BBS1 gene spans 23 kb and is composed of 17 exons. It appears that BBS1 is ubiquitously expressed in the retina, testes, and fetal tissues.

1.2.2 BBS2

In the search for a BBS locus, Kwitek-Black *et al.* (1993) linked the disease to chromosome 16q using a large inbred Bedouin family. Amplification of eighty short tandem repeats was used in a genome wide scan, and linkage was detected at marker D16S408. Due to the fact that affected individuals were suspected to carry a founder mutation, it was not surprising that all nine affected individuals were found to be homozygous at one specific locus. The critical region, determined by observing intrafamilial recombinations, extended 18 cM near D15S408.

Nishimura *et al.* (2001) narrowed the Kwitek-Black *et al.* (1993) critical region to 2 cM. The gene was eventually found through physical mapping and sequence analysis. Many candidate genes were predicted and were prioritized depending on their expression patterns, and what was known about the predicted function of their protein products. The UniGene EST cluster Hs.24809 was analyzed due to it mapping to the critical region and having an appropriately broad expression pattern. Two contigs were formed from the cluster, each containing a gene. One of the genes ended up being the BBS2 gene after sequencing of the gene in families that mapped to the BBS2 locus. Five pathogenic mutations were found: T224G, 940delA, C823T, C814T, and 1206insA. The BBS2 gene is composed of 17 exons and gives an mRNA transcript of 3.0 kb. To confirm its expression in humans, BBS2 cDNA was amplified from a human fetal cDNA library. Northern blotting showed a band of

lower molecular weight in tracheal tissue, suggesting the possibility of alternative splicing. Homologs of BBS2 have been found in mice sharing 90% identity, in rat sharing 89% identity, and in zebrafish sharing 74% identity.

1.2.2 BBS3

Researchers were finding that BBS was demonstrating non-allelic heterogeneity, illustrated by BBS mapping to different regions of the genome in different families. Sheffield *et al.* (1994) analyzed an inbred Bedouin family using homozygosity mapping (i.e., pooling DNA samples from unaffected individuals and affected individuals). Genotyping was conducted on these pooled samples using more than 200 short tandem repeats. Linkage was finally detected by observing a shift in the number of alleles in the unaffected pooled samples compared to a single allele in the affected pooled sample. The critical region was reported to be 11 cM between D3S1254 and D3S1302 (Sheffield *et al.*, 1994).

Young *et al.* (1998) further narrowed the critical region to 6 cM. Using haplotype analysis on a family with five affected individuals, it was determined that the critical region lies between D3S1595 and D3S1753 based on recombination of the disease chromosome in two of the affected individuals (Young *et al.*, 1998). The gene for BBS3 has yet to be identified.

1.2.3 BBS4

Carmi *et al.* (1995) found the locus for BBS4 in a Bedouin family using homozygosity mapping as described above for BBS3. DNA was pooled from parents of affected individuals as well as from the affected offspring and their unaffected siblings. The three samples were then genotyped using 300 short tandem repeat polymorphism markers. Researchers looked for an allele shift towards a single homozygous allele in the affected pool compared to the control pools and found eight allele shifts. After genotyping the eight markers in all individuals and those markers closely surrounding them, seven were considered to be false positives, whereas one marker on chromosome 15 remained of interest. A region of 18.7 cM between markers D15S125 and D15S99 was proposed as the critical region.

Mytytyn *et al.* (2001) identified BBS4 by searching the NCBI EST database using the critical region. They found an EST cluster that, when compared to a human cDNA library, revealed a gene with 519 codons spanning 16 exons across 52 kb. When this gene was sequenced in affected individuals from different families several mutations were found: G884C, G220+1C, and A406-2C. BBS4 was shown to exhibit a broad expression pattern using northern blot analysis and amplification of cDNA from human tissues. BBS4 is most highly expressed in the kidney and is also expressed in fetal tissue, adipose tissue, and the retina.

BBS4 is believed to belong to the N-acetylglucosamine transferase (OGT) gene family (Katsanis *et al.*, 2002). BBS4 is the smallest contributor to BBS

families, with only one to two per cent of pedigrees showing linkage (Beales *et al.*, 2002).

1.2.4 BBS5

The BBS5 gene was mapped to 2q31 by Young *et al.* (1999) using homozygosity mapping with an extended Newfoundland family with five affected individuals. Recombinations observed through haplotype analysis allowed a critical region of 13 cM between markers D2S156 and D2S1238 to be determined. The gene has yet to be found.

1.2.5 BBS6

The gene for BBS6 is also the one that causes MKKS (Katsanis *et al.*, 2000). Before the BBS6 gene was discovered, five Newfoundland BBS families had been excluded from BBS1 - BBS5. One consanguineous family was used for homozygosity mapping, and a potential critical region based on homozygosity in affected individuals was observed after screening with about 150 markers. A recombination event defined the distal boundary at D20S851. Loss of identity by descent defined the proximal boundary at D20S189. The critical region was then defined as a 1.9 cM segment on chromosome 20. Knowing that MKKS and BBS display overlapping phenotypes and that the gene for MKKS had recently been discovered in this region, the MKKS gene became a candidate gene and was analyzed in the uncharacterized families. Several mutations in the MKKS gene were found in BBS families: 429ΔCT433ΔAG, 280ΔT, L277P, Y37C, and T57A.

Eventually, it was determined that 34 per cent of BBS families in Newfoundland contain a BBS6 mutation (Parfrey *et al.*, 2002). This percentage appears high when one considers that in general population surveys BBS6 only accounts for four per cent of BBS cases (Beales *et al.*, 2002).

The BBS6/MKKS gene is composed of six exons, four of which encode the protein product. The transcript is 2.4 kb encoding 570 amino acids and is present in both adult and fetal tissues (Stone *et al.*, 2000). The BBS6 protein likely functions as a type 2 chaperonin (Katsanis *et al.*, 2000).

1.2.6 BBS7

Badano *et al.* (2003) located BBS7 at 4q27 using a pure bioinformatics approach. In the search of BBS7, Badano *et al.* (2003) used the BBS2 protein sequences from humans and zebrafish and compared these sequences to the NCBI EST database. Five contigs were formed from the database, each with 25 per cent to 45 per cent similarity to BBS2. Two of the contigs had an overlapping region of 124 bp that was 99.8 per cent identical, allowing researchers to form a 1,048 bp sequence. This sequence did not appear to have any promoter elements or stop codons, and was thus suspected to be part of a larger gene. Through exon prediction analysis and searching the dbEST again, a contig of 2,580 bp was assembled. The putative transcript was named BBS2L1. The researchers also identified the mouse ortholog of BBS2L1, which is 2,594 bp and is 94.1 per cent similar and 91.5 per cent identical to the human ortholog.

To determine tissue expression patterns in humans, RT-PCR and northern blotting were carried out on human adult and fetal tissue. BBS2L1 is ubiquitously expressed, though a longer isoform of BBS2L1 was found after comparison of human and mouse BBS2L1 sequences. The 3' end of BBS2L1 differs in the two species, so the unique sequence at the end of the mouse transcript was compared against the human dbEST and human genome sequence to reveal an alternative 3' end for BBS21. This other isoform was confirmed to exist in human tissue by performing RT-PCR. The longer form of BBS2L1 is not as widely expressed as the shorter form (Badano *et al.*, 2003).

BBS2L1 was confirmed to be BBS7 through sequencing of the predicted gene in affected individuals from 84 families of primarily European ancestry. Three pedigrees revealed potential pathogenic mutations. Two families carry a homozygous H323R mutation, and the remaining family carried a homozygous T211I alteration. Neither mutation was found in 192 control chromosomes. Researchers also found a region of homozygosity in the area surrounding BBS2L1 in a Saudi Arabian pedigree. After sequencing the BBS2L1 gene in members of the family, it was observed that four base pairs had been deleted in the affected individuals, causing premature termination of translation and a deletion of 65 per cent of the protein (Badano *et al.*, 2003).

1.3.9 BBS8

Ansley *et al.* (2003) recently discovered the BBS8 gene on chromosome 14 in the region 14q32.1. The gene was discovered using a bioinformatics approach similar to the method used to find BBS7. The gene for BBS4 was split into eight random overlapping fragments that were then compared with the human genome and an EST database. The tetratricopeptide repeats (TPRs) of BBS4 aligned with the C-terminus of the predicted protein TTC8. The gene encoding TTC8 is composed of 14 exons and codes for a 60.4 kDa protein. The predicted gene was sequenced in affected individuals from three families and the following mutations were observed: 187-188delEY and IVS10+2-4delTGC. The first mutation causes a six base pair exonal deletion, the second a three base pair deletion interrupting splicing of exon 10. The researchers supported their molecular results by showing that the BBS8 transcript is widely distributed in both adult and fetal tissues (Ansley *et al.*, 2003).

1.3 Protein Structure and Function of BBS Proteins

BBS6 was the first BBS gene to be identified. It was shown to be the same gene that, when mutated, causes MKKS (Katsanis *et al.*, 2000). The MKKS/BBS6 protein had been described as a type II chaperonin (Stone *et al.*, 2000) and therefore, it was thought that other BBS genes might encode additional chaperonins that worked in harmony with the BBS6/MKKS protein. This was the basis for examining several candidate genes but all to no avail. When the genes for BBS1, BBS2, BBS4, BBS7, and BBS8 were identified, none of these were found to code for proteins resembling a chaperonin. Indeed, the functions of these gene products are still not understood.

Based on sequence similarity, it has been proposed that the BBS4 gene product is part of the N-acetylglucosamine transferase (OGT) gene family (Mytytyn *et al.,* 2001). As illustrated in Figures 1.1 and 1.2, the BBS1, BBS2, and BBS7 proteins are similar over a small section of their respective sequences (Mykytyn *et al.*, 2002; Badano *et al.*, 2003).

Badano *et al.* (2003) noted that BBS2 does not have any known domains except for a putative coiled-coil domain between residues 322 and 365. The BBS2 sequence was analyzed using SCOP (Murzin *et al.,* 1995 and Lo Conte *et al.*, 2002). Between residues 171 and 315 there appears to be a structure best described as a six-bladed ß-propellor structure, which is also found in BBS1 and BBS7. There is some similarity between this region and the transcriptional regulator *zraR* from *E. coli,* as well as similarity to the integrin family.

Figure 1.1 Adapted from Badano *et al.* (2003). Regions of sequence similarity between the BBS1, BBS2, and BBS7 proteins. The numbers refer to amino acid residues and corresponding mutations are as labeled.

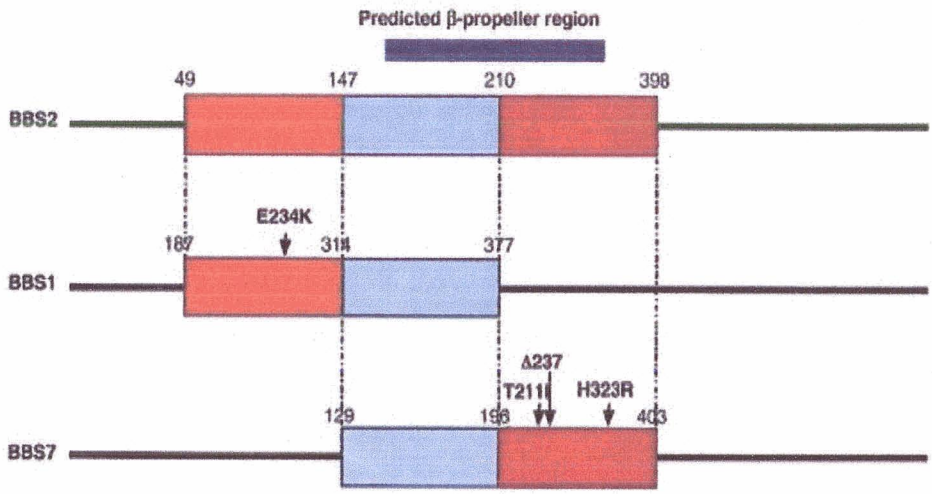
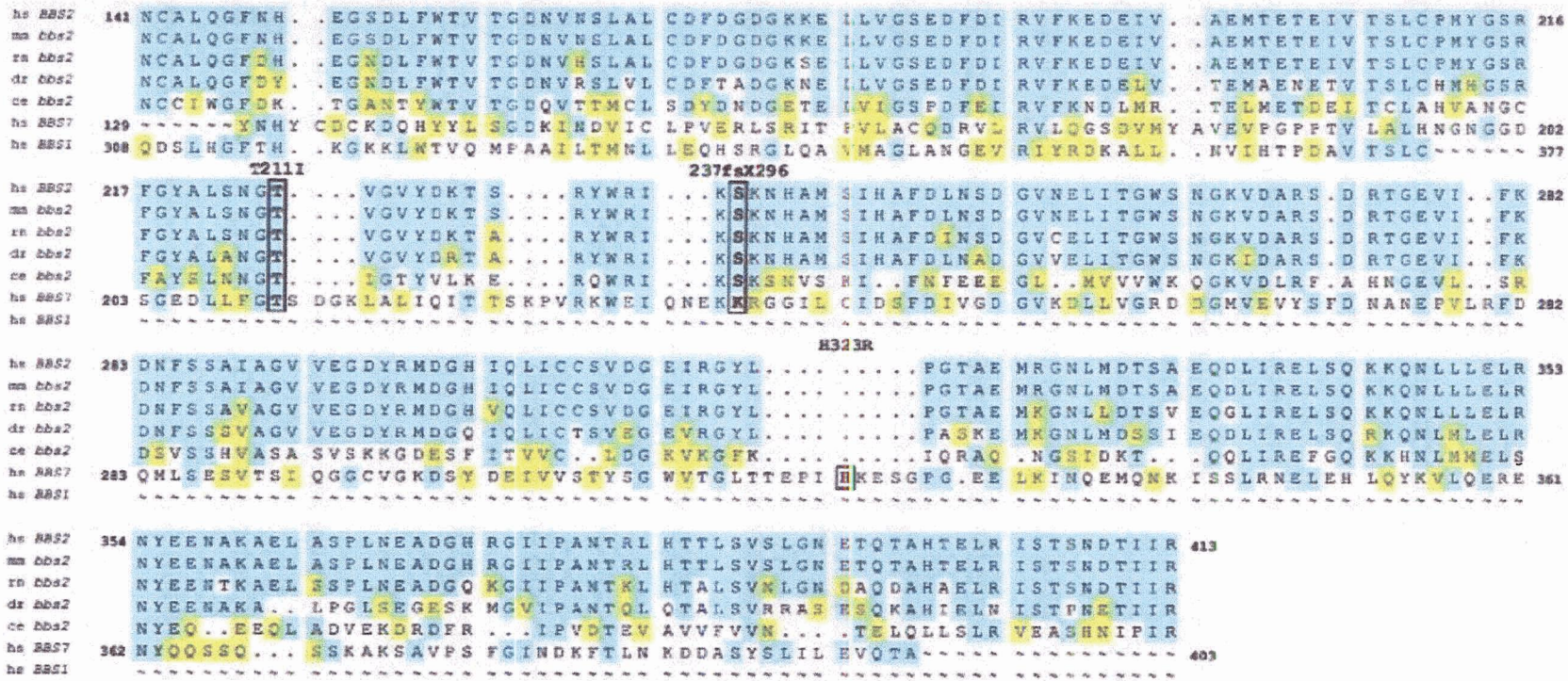

Predicted β-propeller region
49
147
210
398
BBS2
E234K
187
314
377
BBS1
Δ237
T211I
H323R
129
196
403
BBS7

Figure 1.2 Adapted from Badano *et al.* (2003). Homologous regions between BBS1, BBS2, and BBS7 in *Homo sapiens (hs), Mus musculus (mm), Rattus norvegicus (rnn), Danio rerio (dr),* and *Caenorhabditis elegans (ce).* Mutations are as boxed.


T211I
237fsX296
H323R

Ansley *et al.* (2003) showed that in the 5' untranslated regions of the *C. elegans* genes for BBS1, BBS2, BBS7, and BBS8 there appears to be a 14 bp imperfect repeat called the X box about 100 bp from the start codon (Figure 1.3). It is also known as the RFX box. This regulatory region is common to *C. elegans* genes that are expressed in ciliated neurons. This finding helped to build a strong correlation between BBS and ciliary dysfunction. The BBS8 protein also shows similarity to the prokaryotic domain *pilF*, which is present in *E. coli* and other bacteria, and allows for mobility. In addition, it has been shown that BBS4 closely interacts with Pericentriolar Material 1 protein (PCM1). This protein is found in ciliary basal bodies which take part in ciliogenesis.

C. elegans was used as a model organism to demonstrate that the worm ortholog to human BBS8 is exclusively expressed in ciliated neurons. Ansley *et al.* (2003) found that the *C. elegans* ortholog to BBS8 has the same expression pattern as the *C. elegans* ortholog to the mouse *Tg737* polycystic kidney disease (PKD) gene. It is known that defects to the cilia are linked to PKD (Pazour *et al.*, 2000), so it is possible that the wild type BBS8 protein may be involved in maintaining cilia structure and/or function. Another process involving cilia that BBS proteins may play a role in is intraflagellar transport, or IFT. IFT proteins have homologs in *C. elegans*, *D. melanogaster*, mouse, and human. These proteins are noticeably absent in organisms that are not ciliated (Rosenbaum *et al.*, 2002). It is interesting to note that the domain used to identify the BBS8

Figure 1.3 Adapted from Ansley *et al.* (2003). Alignment of X boxes of BBS1, BBS2, BBS7, and BBS8 genes in ciliated neuronal cells of *C. elegans*. The base pair numbers refer to distances between RFX boxes and the start codons of the genes. Conserved nucleotides are indicated in green below the alignment.

bbs-1	GTTCCCATAGCAAC	-99 bp
bbs-2	CTATCCATGGCAAC	-94 bp
bbs-7	GTTGCCATAGTAAC	-107 bp
bbs-8	GTACCCATGGCAAC	-84 bp
	GT(T/A)CCCAT(A/G)GCAAC	consensus sequence
	C T G T	

protein, the tetratricopeptide repeat (TPR), is common in IFT genes such as IFT139 and IFT88, the latter having homologs shared between *C. elegans* and *C. reinhardtii* (Cole *et al.*, 2003).

Though functional studies have not shown co-localization of BBS proteins, it is possible that the proteins may function apart from each other in the cascade. BBS1, BBS2 and BBS8 were found to be expressed in ciliated neuronal cells in *C. elegans*. Additionally, BBS4, BBS6, and BBS8 have all been shown to localize at the centrosome/basal body in mouse primary ciliary cell lines (M. Leroux, personal communication).

1.4 Modes of Inheritance

BBS has traditionally been modeled as a Mendelian autosomal recessive disease. It was on this basis that BBS1-6 were mapped (Parfrey *et al.*, 2002). However, another form of inheritance termed "triallelism" has been proposed by Katsanis *et al.* (2001). This form of inheritance bridges the gap between classical Mendelian genetics and complex diseases. Burghes *et al.* (2001) suggested that triallelism be referred to as "recessive inheritance with a modifier of penetrance."

Triallelism is a form of inheritance where, in order to manifest a disease, one must have two mutations at one locus and one more mutation at a different locus. This is illustrated in Figure 1.4. Katsanis *et al.* (2001) took all 163 members in their BBS cohort of families and screened them for BBS2 and BBS6

mutations. As a result, they found several families that appeared to manifest the disease only if the individual had three mutations, not the typical two mutations as was the case for all BBS affected individuals up until that point in time. There were two cases where the affected person had an unaffected sibling that did not manifest the disease though the unaffected sibling had two mutations for BBS2. Family AR259 had an affected individual who had two BBS2 mutations and was heterozygous for a BBS6 mutation, whereas the unaffected sibling had the same two BBS2 mutations but no BBS6 mutation (Katsanis *et al.*, 2001). The conclusion the authors reached was that in order to manifest the disease in some instances, three allelic mutations are needed instead of two.

During analysis of the data, researchers also found a Newfoundland family, NF-B14, which possibly demonstrated the concept of triallelism. The affected individual in NF-B14 is homozygous for the BBS2 Y24X mutation and is also heterozygous for the BBS6 A242S mutation. Unfortunately, there are no siblings with the homozygous BBS2 mutation and no BBS6 A242S mutation. The only sibling of the affected person is only heterozygous for the BBS6 A242S mutation. As a result, it was impossible for the authors to predict whether the BBS6 A242 mutation was indeed pathogenic in this family. However, it is known that this mutation is found in individuals among the Old Order Amish who have MKKS, and in another BBS patient. After the paper by Katsanis *et al.* (2001) was published, a BBS6 A242 mutant construct was made and it was determined that this mutant BBS6 protein does not function as does the wild type BBS6 protein

(M. Leroux, personal communication). Another family, PB043, was also found to be homozygous for BBS2 and also mapped to BBS4. In a follow up paper by Katsanis *et al.* (2002), they found that BBS4 is also a minor contributor to tetra-allelsim in conjunction with BBS2. The mode of inheritance of tetra-allelism is very similar in idea to that of triallelism, only rather than inheriting three mutant alleles, one inherits four. This can be seen in Figure 1.5. Family PB043 had two mutations for BBS2 and also mapped to BBS4. It was found that the affected individual was homozygous for the BBS4 mutation 1091C->A, and was also homozygous for the BBS2 T5601I mutation. The mother and one unaffected sibling both have a homozygous T5601I mutation and are both heterozygous for the 1091C->A mutation, suggesting that the fourth mutant allele is needed for one to manifest the disease.

BBS7 may also play a part in triallelism. Badano *et al.* (2003) sequenced the other known BBS genes in BBS7 patients and found an affected individual with the T211I mutation had an E234K mutation in the BBS1 gene. However, there were no unaffected siblings with the homozygous T211I mutation and no

Figure 1.4 Illustration of the mode of inheritance that is termed triallelism. One yellow bar represents one locus, and each parent has two. Each green bar represents another locus. Alleles are represented by blue (normal) and red (mutant) bars at each locus. The eight possible inheritance combinations that could be passed on to offspring are shown. If a disease is inherited in a triallelic fashion, the only individual to contract the disease would be the one inheriting the indicated "affected" allelic combination (i.e., three mutant alleles).

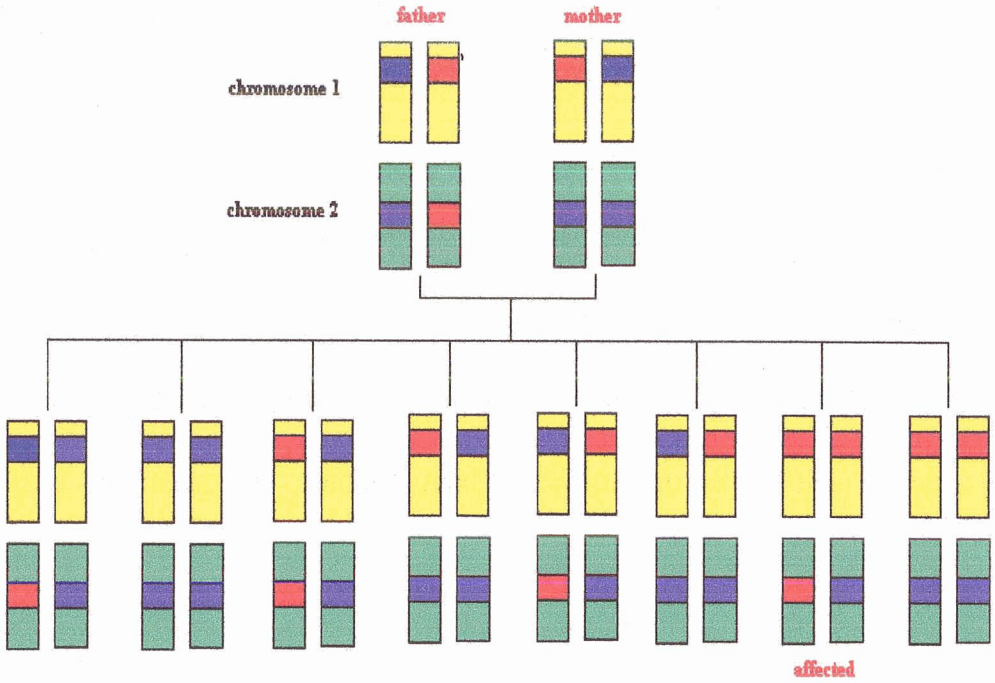
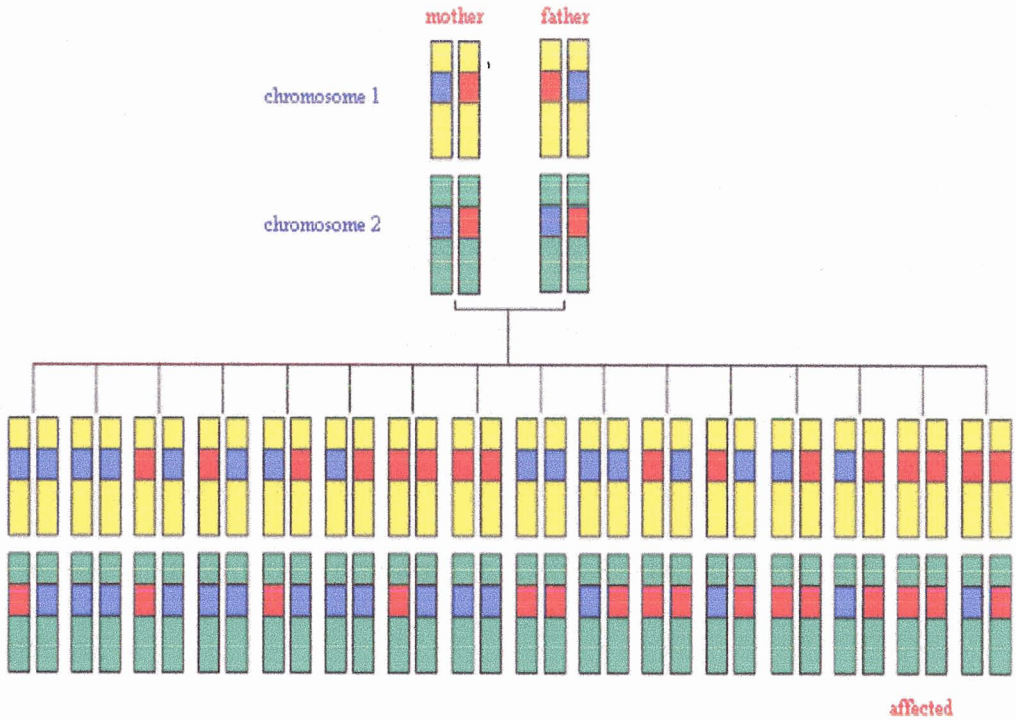

father
mother
chromosome 1
chromosome 2
affected

Figure 1.5 Illustration of the mode of inheritance that is termed tetra-allelism. One yellow bar represents one allele, and each parent has two. Each green bar represents another locus. Alleles are represented by blue (normal) and red (mutant) bars at each locus. The sixteen possible inheritance combinations that could be passed on to offspring are shown. If a disease is inherited in a tetra-allelic fashion, the only individual to contract the disease would be the one inheriting the indicated "affected" allelic combination. (i.e., four mutant alleles).


mother
father
chromosome 1
chromosome 2
affected

E234K mutation to confirm the importance of the latter.

Badano *et al.* (2003) have recently described intrafamilial variation in three families. There was recognizably faster progression of symptoms in the affected persons with three mutations as opposed to two mutations. All three families – AR768, PB009, and PB061 each have at least two affected individuals, and one of them had a third mutation. Clinical findings suggested that the affected individuals with a third mutation had more severe phenotypes, such as severe retinal dystrophy, compared to their affected siblings with only two mutations.

In summary, Katsanis *et al.* (2002) determined that of the 19 families with BBS2 mutations, 47.3 per cent of them had some involvement of another BBS locus. For the cohort of BBS6 families, 37.5 per cent have involvement of the BBS2 locus. It is difficult to tell how significant a tertiary or quaternary mutant allele is in manifestation of BBS in triallelic families. One possibility is that there could be other mutations in other genes that are necessary in manifesting the disease, and that they simply have not been found yet. There is also the possibility that people with mutations do not manifest the disease due to modifier genes. Multifactorial diseases manifest under a variety of genetic and environmental conditions, and certain conditions must be met at or above their respective thresholds in order for the disease to present itself in some individuals.

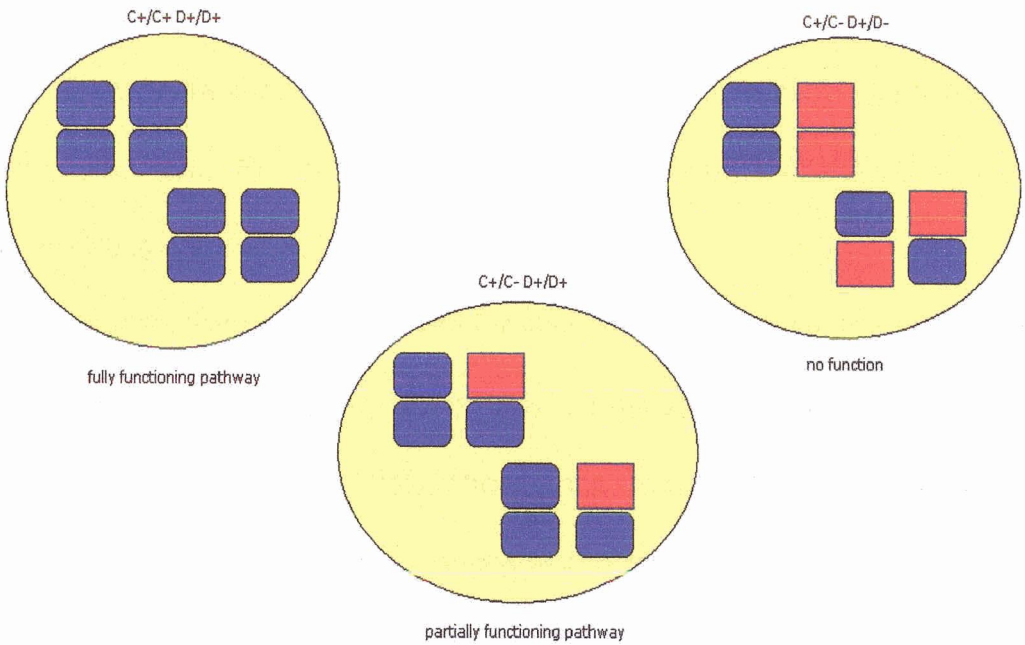
There are other cases where multiple mutations are required for a disease to be manifest. The first case of complex inheritance involved Charcot-Marie-Tooth neuropathy type 1A, or CMT1A. Lupski *et al.* (1991) found that the disease is caused by a duplication on chromosome 17, inherited as a homozygous mutation. Digenic inheritance is a related type of inheritance that waivers from traditional Mendelian inheritance patterns, where one manifests a disease through inheritance of mutated alleles of seemingly unrelated genes (Helwig *et al.*, 1995). A prime example of digenic inheritance is in Retinitis Pigmentosa (RP), which will manifest when there are heterozygous mutations in the genes ROM1 and peripherin/RDS (Kajiwara *et al.*, 1994). RP can be inherited in a monogenic fashion as well. Digenic inheritance is also occasionally seen in the case of inheritance of Waardenburg syndrome type 2 and autosomal recessive ocular albinism (Morell *et al.*, 1997). If BBS were inherited in a digenic fashion, then the father depicted in Figure 1.4 would be affected. The review by Katsanis *et al.* (2002) summarizes the research done in this field and touches on the bodily effects of various models of inheritance such as the poison model. The poison model theorizes how higher concentrations of mutated proteins may reach a certain threshold that would allow a disease to manifest. This model helps to explain why inheritance patterns such as digenic inheritance would be necessary to cause a disease in an individual, as seen in Figure 1.6. There are not enough functioning protein complexes in the third illustration to allow the

cascade to function normally in the cell, allowing for manifestation of the disease.

1.5 Purpose of Thesis

The purpose of this thesis was to provide information that could be used to better understand the molecular basis for BBS. The approach taken was: (1) to characterize newly diagnosed Newfoundland BBS families and to investigate the prevalence of triallelism involving BBS6 in the Newfoundland BBS population, and (2) to screen candidate genes for BBS3 and BBS5.

Figure 1.6 Adapted from Katsanis *et al.* (2002). Illustration of the poison model. The ovals represent separate cells with different genotypes, as is indicated above them. The blue octahedrons represent wild type proteins and the red squares represent mutated proteins.


C+/C+ D+/D+
fully functioning pathway
C+/C- D+/D-
no function
C+/C- D+/D+
partially functioning pathway

Chapter 2: Materials and Methods

2.1 Subjects

Twenty-six Newfoundland families with BBS have been identified, named NF-B1 to NF-B26. However, DNA was not available for 7 of the 26 families so 19 BBS kindreds were examined as part of this thesis.

Clinical information on these families was collected by Jane Green and colleagues at Memorial University of Newfoundland (Harnett *et al.*, 1988; Green *et al.*, 1989; O'Dea *et al.*, 1996). The Human Investigations Committee of the Faculty of Medicine at Memorial as well as the Medical Advisory Council of the St. John's General Hospital approved the protocols for clinical investigation of these BBS patients and their relatives. The research has also been approved by the Research Ethics Board at Simon Fraser University. The patients gave informed consent, and formal diagnosis of affected individuals was performed. A person was diagnosed with BBS if they demonstrated four cardinal features, or if the person had three cardinal features and was also the sibling of an affected individual.

Of the Newfoundland families, there were 174 individuals out of the 19 BBS families from whom DNA was available. Of these, 39 were affected individuals.

2.2 Extraction of DNA

2.2.1 Extraction of DNA from Whole Blood

Extraction of DNA from whole blood was conducted via the Puregene DNA Isolation Kit by Gentra Systems.

2.2.2 Extraction of DNA from Cell Lines

DNA was extracted from Epstein-Barr Virus transformed B-cell lines. The cells were on dry ice on arrival from Newfoundland, with a minimum of three million cells per tube. The cells were contained in freezing media composed of 90 per cent fetal bovine serum and 10 per cent DMSO. The cells were stored at SFU by submergence in liquid nitrogen before DNA extraction.

All preparation of cells was performed in a Laminar flow hood. Cells were thawed rapidly in warm water with mild agitation. After only a pellet of frozen freezing media was left in the tube, the tube was again plunged into ice. This was done so that the DMSO in the media would not become warmer than 4 °C, preventing unnecessary lysing of cells by DMSO. The tube was then decontaminated by drenching the cap and top part of the tube with 70 per cent ethanol. The contents were then removed with a sterile transfer pipet. The cells were transferred to a 15 ml conical tube containing 10 ml culture media. The culture media contained 500 ml of 90 per cent RPMI 1640 (Gibco), 50 ml of inactivated 10 per cent fetal bovine serum, 5.5 ml penicillin streptomycin, 5.5 ml

L-glutamine, 5.5 ml sodium pyruvate, and 5.5 ml HEPES. The last four ingredients were one per cent of the final media preparation.

The conical tube was then centrifuged at 1500 RPM for approximately 5 minutes. The supernatant was then decanted, and the cell pellet was again gently suspended in 10 ml of culture media. Resuspension with culture media was performed three times. Five ml of the cell culture was then transferred to a culture flask containing 5 ml of fresh culture media. The cells were then incubated for twenty-four hours in a carbon dioxide controlled incubator, set to 90-95 per cent humidity at 37^{o}C. Cells were grown to a concentration of about three million per culture flask.

After cells had reached a density of three million per flask, the cells were diluted out. 5 ml of the culture was added to 5 ml of fresh culture media in a culture flask. The cells were then placed in the incubator until they too grew to a density of about three million cells per flask. This procedure was repeated several tmes, allowing for the growth of millions of cells.

Cells from a flask with a density of three million were frozen by pelleting the cells several times with culture media at 1500 RPM. The cells were spun for 5 minutes and supernatant decanted before resuspension in fresh culture media. After repeating resuspension in culture media, cells were resuspended in 1 ml of freezing media. Cells were later transferred to a freezing tube that had been pre-cooled. The tube was capped and transferred to a –80^{o}C freezer. After 24 hours, the tube was transferred to liquid nitrogen. Extraction of DNA from the

cell culture was conducted via the Puregene DNA Isolation Kit by Gentra Systems.

2.3 Mutation Analysis

All mutational analysis in this project was carried out using direct sequencing. In order to sequence a segment of the genome, primers were made that annealed to the sides of the region in the genome to be amplified. These primers were between 18 and 35 base pairs in length, and designed using the program C-primer (http://iubio.bio.indiana.edu:7780/archive/00000017/). The primers were tested for their optimum annealing temperatures by using PCR in combination with a temperature gradient. The PCR reactions were composed of 6.5 ul dH_2O, 1.0 ul 10x buffer, 0.05 uM dNTP, 0.1 uM forward primer, 0.1 uM reverse primer, 0.5 units of Taq, and 40 ng of DNA. The PCR reaction was then subjected to the following PCR cycle: 95°C for 4 minutes, 95°C for 45 seconds (denaturation,) specified annealing temperature for 45 seconds, 72°C for 45 seconds (extension,) 72 °C for 10 minutes, and ending at 4 °C. There were 29 cycles between steps 4 and 2 before being held at 4 °C. The reactions were then run out on a two per cent agarose gel to determine which annealing temperature gave the best product.

After the best annealing temperatures were established, the primers were used to amplify regions of interest from DNA from individuals of BBS families. Depending on the primers being used and the quality of DNA, the PCR reactions

and PCR cycles were modified to allow for optimal amplification of the region of interest. About 10 ul of the PCR reaction was mixed with 5 ul of loading dye and run out on a two per cent agarose gel. If the product was good, the remaining PCR reaction was purified using the QIAgen PCR purification kit. The purified PCR was then subjected to a sequencing reaction, using the dideoxy terminator method carried out with the DYEnamic ET terminator Cycle Sequencing Kit from Amersham Biosciences. The resulting sequences were precipitated using the following procedure: adding 2 ul of EDTA/sodium acetate along with 80 ul 95 per cent ethanol and spinning down at 13,000 RPM for 20 minutes. This was followed by pipeting off the supernatant and adding 200 ul of 70 per cent ethanol, spinning down at 13,000 RPM for 5 minutes, then pipeting the supernatant and letting the pellets dry. After drying, 4 ul of loading dye was added before loading the sample into the ABI 373 or 377 DNA sequencer (Applied Biosystems). Sequences were analyzed using the program Sequencer 3.1, which provided corresponding sequence chromatograms. These chromatograms were analyzed for quality, and good sequences were compared to the human genome available at BLAST on the NCBI homepage (Altschul *et al.*, 1990). Comparisons were carried out by pasting the sequences into the BLAST program and comparing them against the non-redundant database. The best matches would then be listed and visual comparisons could then be made.

2.4 Computational Methods

A couple of computational programs were used in the bioinformatics portion of this project. The web sites for these programs are as follows:

BLAST programs: http://www.ncbi.nlm.nih.gov/BLAST/ (Altschul *et al.*, 1990)

SMART program: http://smart.embl-heidelberg.de/ (Schultz *et al.*, 1998)

Chapter 3: Mutational Analysis of Newfoundland BBS Families

3.1 Rationale

Considerable research has been carried out on BBS using several Newfoundland families. Initially this concentrated on the clinical features of BBS but progressed to molecular studies and now this is culminating in analyzing the relationships between phenotype and genotype. The first paper published was by Harnett *et al.* (1988), who described 20 patients in 17 families, showing that renal abnormalities were a part of the phenotype of BBS. A follow up paper by Green *et al.* (1989) defined the cardinal manifestations of BBS as: renal abnormalities, obesity, dysmorphic extremities, severe retinal dystrophy, and hypogenitalism in males. Other features included mental retardation, reproductive abnormalities in women, and altered pituitary function measured by blood glucose levels.

Further studies investigated the molecular basis of BBS. Basing the disease on a Mendelian autosomal recessive model, one of the first papers on the molecular basis of BBS in Newfoundland was by Woods *et al.* (1999) who surveyed 17 Newfoundland BBS families. Besides confirming the existence of more than one locus and thus the likelihood of there being more than one founder, it was determined that there was a BBS5 locus because of the exclusion of several families from BBS1-4.

Woods *et al.* (1999) also showed that family NF-B2 was the first BBS3 family of northern European descent to be identified, and only the second BBS3 family described worldwide. The BBS3 locus was confirmed and the critical region reduced by Young *et al.* (1998) using haplotype analysis on family NF-B2. The critical region was determined to be between D3S1595 and D3S1753. The phenotype differed from the first BBS3 family as described by Carmi *et al.* (1995), which alluded to there being no locus-specific phenotype.

To follow up on the work done on BBS5 by Woods *et al.* (1999), Young *et al.* (1999) analyzed all Newfoundland families, and it was apparent that exactly six families were excluded from BBS1 to BBS4 as Woods *et al.* (1999) described. One family, NF-B9, was known to be consanguineous. Using a homozygosity mapping approach, a homozygous region at marker D2S1353 was found. Linkage of BBS5 to 2q31 was then confirmed by showing affected individuals in the family were homozygous by descent for an ancestral haplotype (Young *et al.*, 1999). The research was extended by surveying all uncharacterized families in the Newfoundland cohort for BBS5. It was found that five families were excluded from BBS1-5, thus presenting strong evidence for the existence of BBS6.

The BBS6 gene was found using the Newfoundland cohort as well as other BBS families from different regions of the world. After finding

that one Newfoundland family mapped to BBS5, there were still five Newfoundland families that did not map to BBS1-5, thus leaving room for the

existence of yet another BBS locus. In collaboration with researchers at the Baylor College of Medicine, Katsanis *et al.* (2000) used uncharacterized BBS families in a linkage analysis that helped to identify linkage to the region between D20S851 and D20S189. The MKKS gene had been recently discovered, and Katsanis *et al.* (2000) were aware of the phenotypic similarities between MKKS and BBS. As a result, the MKKS gene was screened for mutations in the non-BBS1-4 Newfoundland affected individuals. Refer to Figure 3.1 for a depiction of the gene and to Table 3.1 for a list of primers used to amplify the gene. Three different BBS6 mutations were found in several Newfoundland families: fs1, fs2, and L277P. These families are NF-B1, NF-B3, NF-B4, NF-B5, NF-B13, and NF-B16, as shown in Figures 3.2 to 3.8.

Young *et al.* (1999) narrowed the critical region for BBS1 between markers D11S1883 and D11S4940 by using haplotype analysis. It was determined that six of these families segregated in the region of 11q13, in the area of the BBS1 critical region.

All the research done on Newfoundland BBS families has given rise to the question "if BBS exists in Newfoundland due to the founder effect, then why are there so many BBS types on the island?" This question is aptly termed the Newfoundland paradox. It was first thought that only one of the founders from southeast Ireland and southwest England who settled on the island had introduced a BBS mutation to the population. If this were the case, one would

Figure 3.1 MKKS/BBS6 gene structure. BBS6 is located on chromosome 20p12, and is comprised of six exons and five introns. Sizes are indicated in base pairs. Forward and reverse primers are shown with rightward and leftward arrows and names displayed above. Arrows show start and stop codons. Locations of known mutations fs1, fs2, L277P, and A242S are as indicated. The exons and portion of exon 3 in yellow are untranslated, whereas the blue portions are translated.

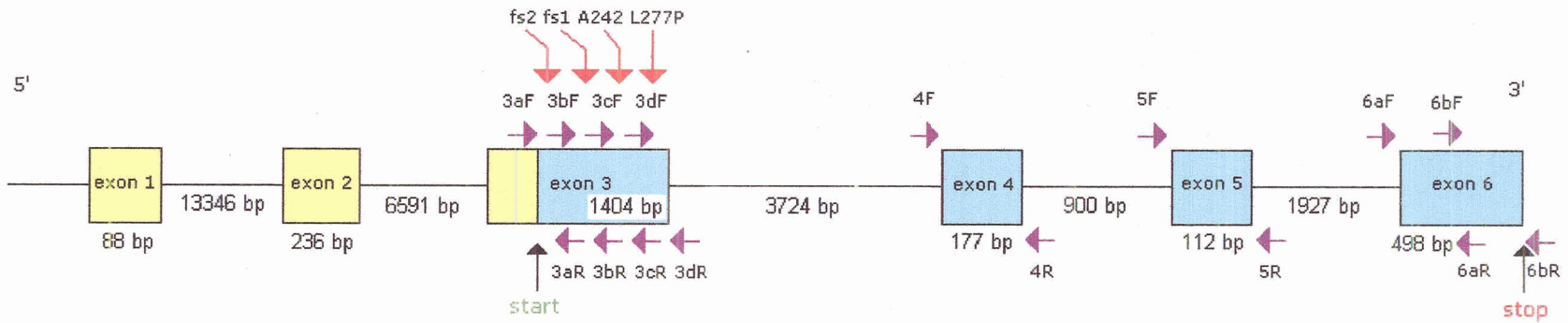

fs2 fs1 A242 L277P
5'
3aF 3bF 3cF 3dF
4F
5F
6aF 6bF
3'
exon 1
13346 bp
exon 2
6591 bp
exon 3
1404 bp
3724 bp
exon 4
900 bp
exon 5
1927 bp
exon 6
88 bp
236 bp
177 bp
112 bp
498 bp
3aR 3bR 3cR 3dR
4R
5R
6aR
6bR
start
stop

Table 3.1 Table of oligonucleotide primers used in amplifying the BBS6 gene (Katsanis *et al.*, 2001).

Exon	Primer Name	Sequence	Size (bp)	Annealing Temp (°C)
3a	BBS6X3a-F BBS6X3a-R	GATTTTATAGCCACAATGCT ATGACAGTGGTGGGTGTCAA	451	63.1
3b	BBS6X3b-F BBS6X3b-R	TCTGGTGAGCATACAGGCAG CGTTTGGAAGCTAAGAAGCC	457	53.8
3c	BBS6X3c-F BBS6X3c-R	GATCCTCCTTTGTTTGGTGC GGTTAAGCAGCTGGTCCAAG	351	65.0
3d	BBS6X3d-F BBS6X3d-R	AATCAACTGCCCTCAAGGTG CCTTTGCTGCCAGAAATGAT	381	53.8
4	BBS6X4-F BBS6X4-R	ATGCTTGTGGGGCTTTTATG AATGGCAACACATGCCAAAT	435	63.1
5	BBS6X5-F BBS6X5-R	GCACCACACAAGTTTTGTTC CCTATACATGCACCCCTGAA	339	53.8
6a	BBS6X6a-F BBS6X6a-R	GTGCCAGACCCCAAATTAAA CCAGTTGAGTTCTTCCTGGC	352	63.1
6b	BBS6X6b-F BBS6X6b-R	GGCAGATTCTCCCTGTGTTG GCATTTCCATTCACGAATCA	407	58.5

Figure 3.2 Pedigree of family NF-B1 with corresponding BBS6 haplotypes. Individual PID's are indicated under each individual.

Key:

unaffected individual

affected individual

carrier

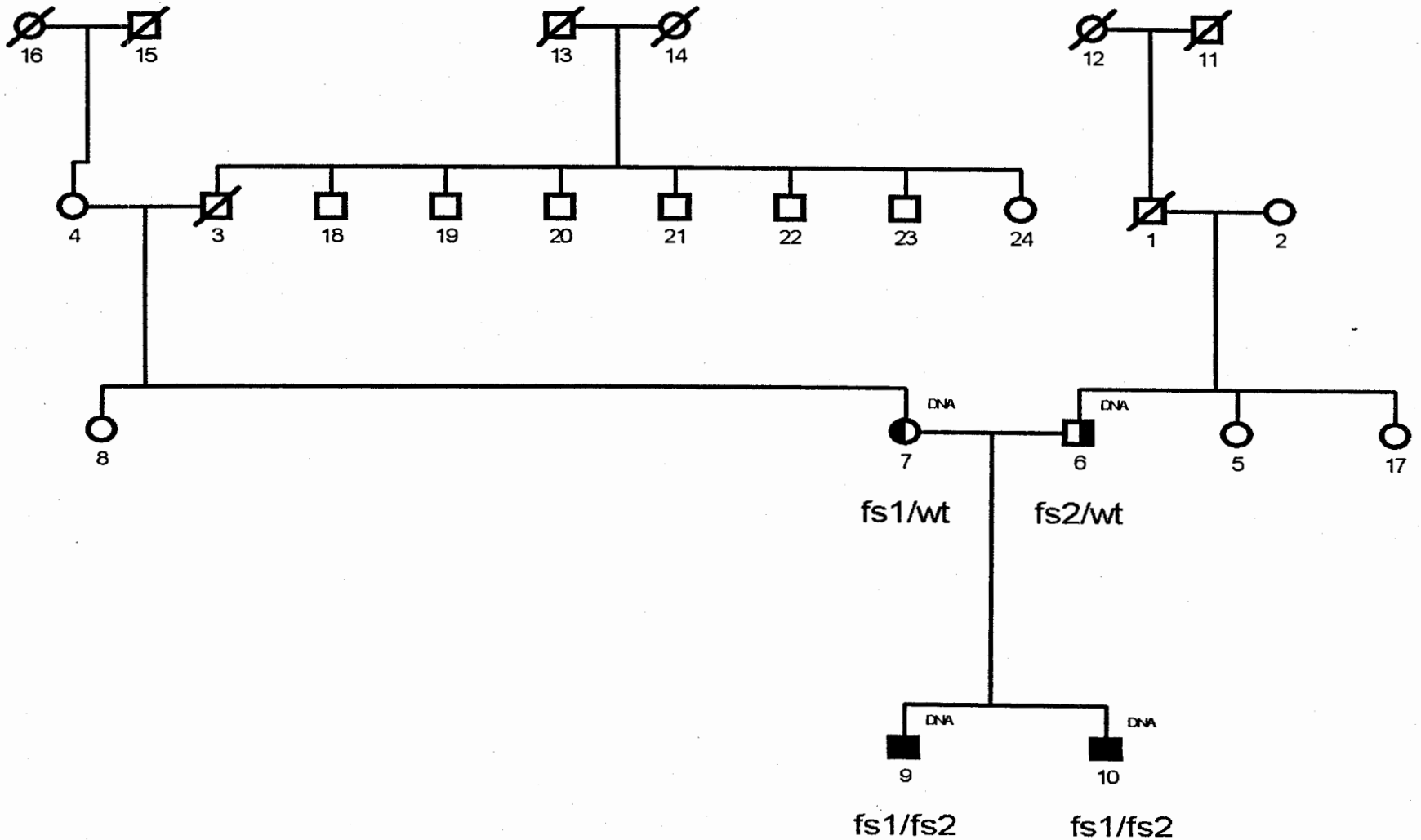

16
15
13
14
12
11
4
3
18
19
20
21
22
23
24
1
2
8
DNA
7
DNA
6
5
17
fs1/wt
fs2/wt
DNA
DNA
9
10
fs1/fs2
fs1/fs2

Figure 3.3 Pedigree of family NF-B3 with corresponding BBS6 haplotypes. Individual PID's are indicated under each individual.

Key:

unaffected individual

affected individual

carrier

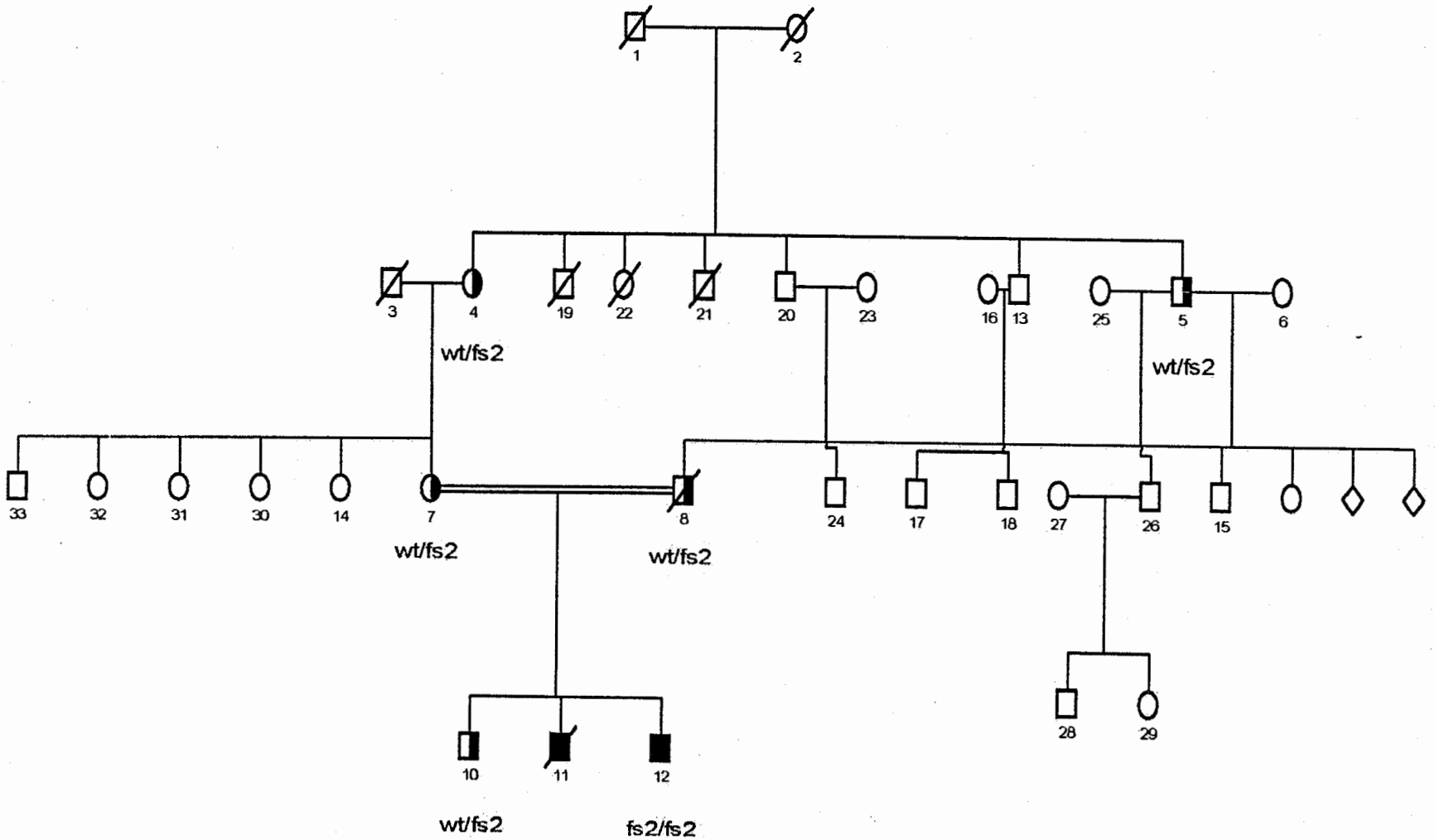

1
2
3
4
19
22
21
20
23
16
13
25
5
6
wt/fs2
wt/fs2
33
32
31
30
14
7
8
24
17
18
27
26
15
wt/fs2
wt/fs2
28
29
10
11
12
wt/fs2
fs2/fs2

Figure 3.4 Pedigree of family NF-B4 with corresponding BBS6 haplotypes. Individual PID's are indicated under each individual.

Key:

unaffected individual

affected individual

carrier

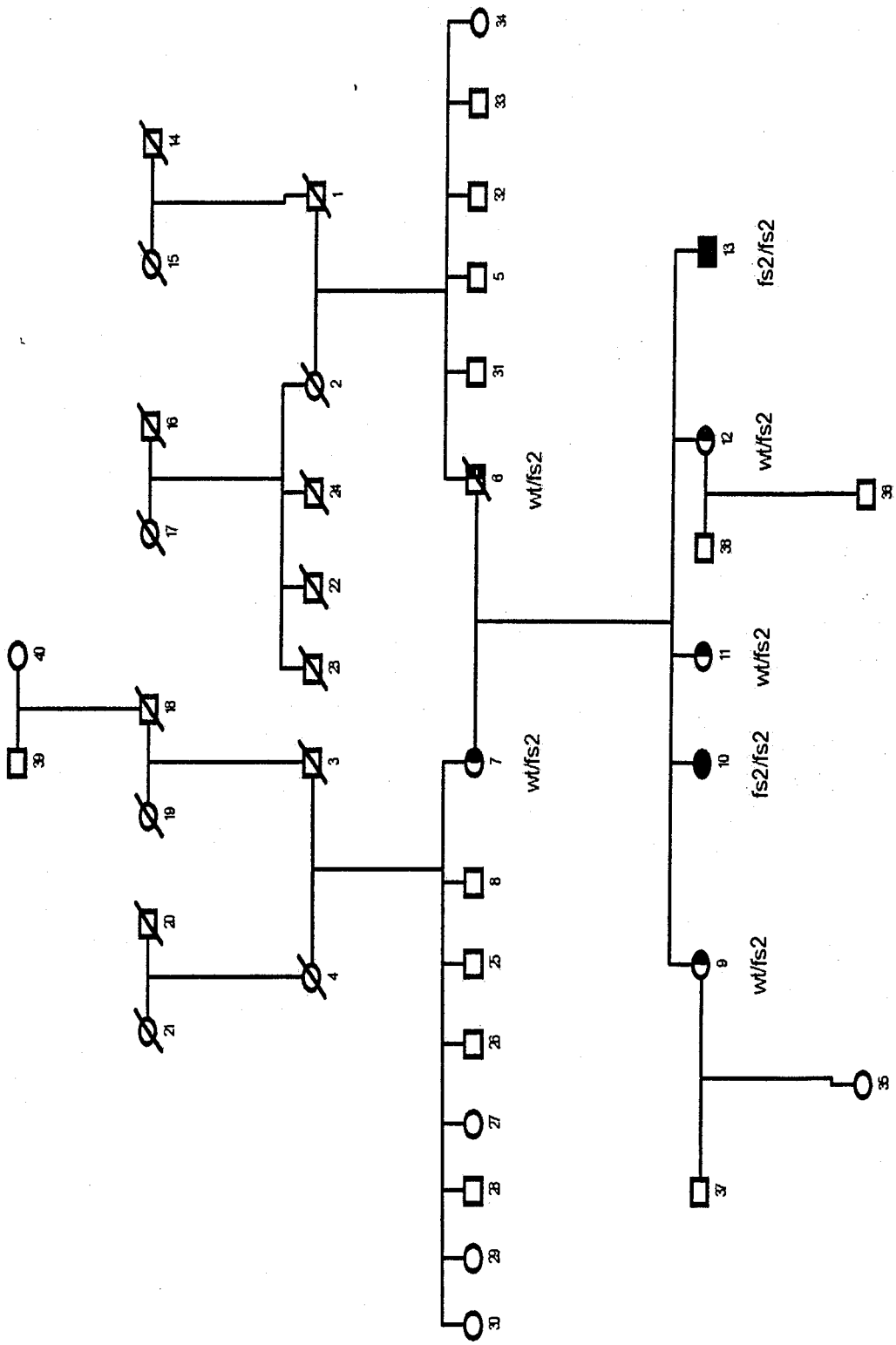

fs2/fs2
wt/fs2
wt/fs2
wt/fs2
wt/fs2
fs2/fs2
wt/fs2

Figure 3.5 Pedigree of family NF-B5 with corresponding BBS6 haplotypes. Individual PID's are indicated under each individual.

Key:

unaffected individual

affected individual

carrier

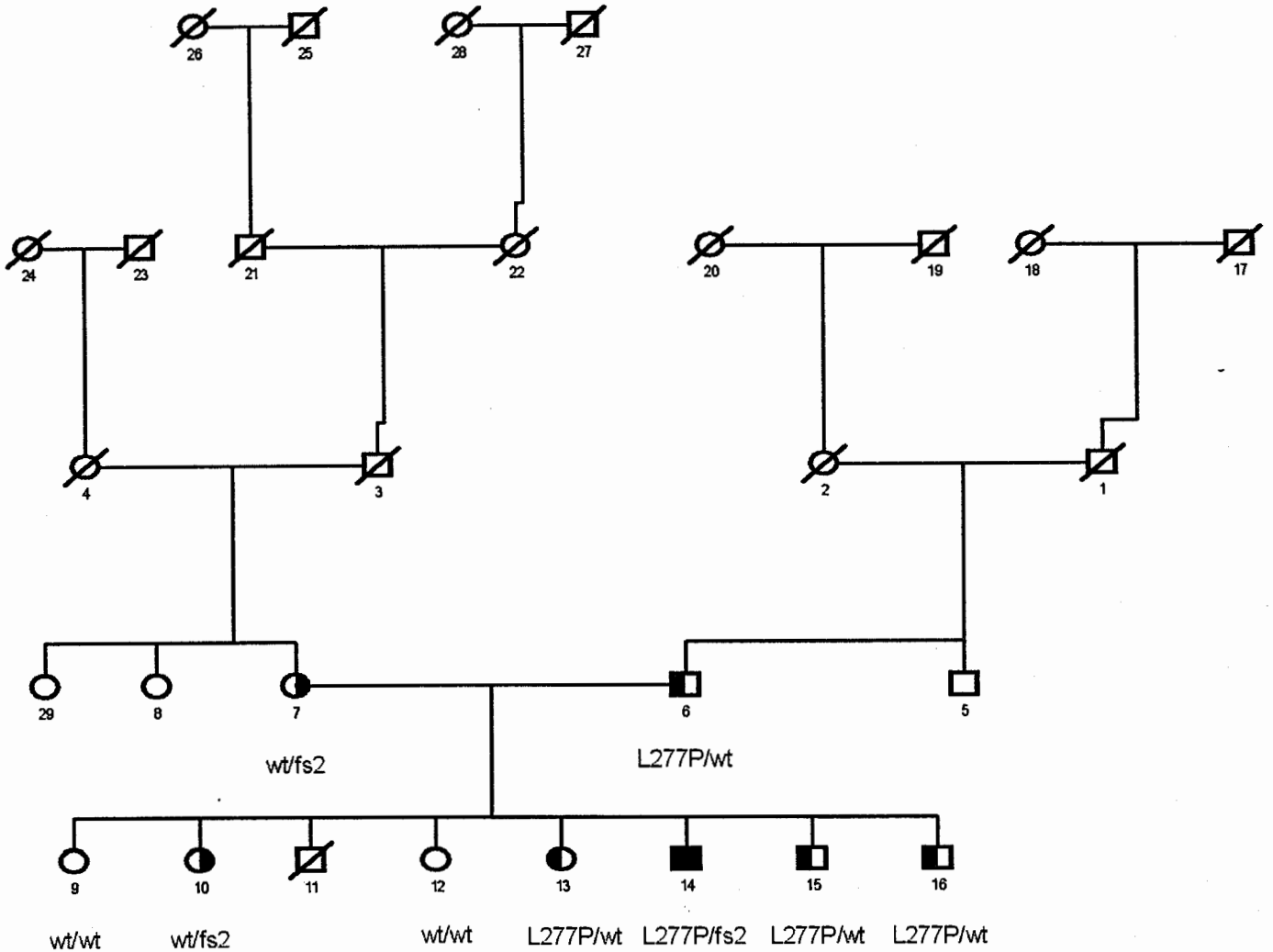
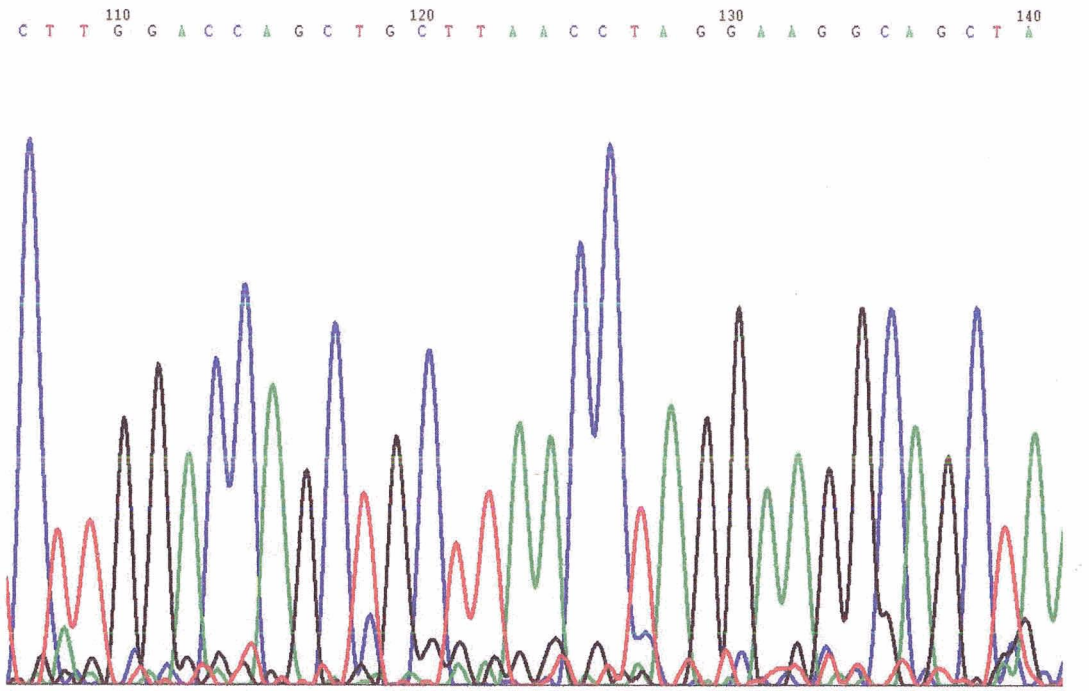

26
25
28
27
24
23
21
22
20
19
18
17
4
3
2
1
29
8
7
6
5
wt/fs2
L277P/wt
9
10
11
12
13
14
15
16
wt/wt
wt/fs2
wt/wt
L277P/wt
L277P/fs2
L277P/wt
L277P/wt

Figure 3.6 Chromatograms showing the BBS6 mutation L277P in family NF-B5. A wild type sequence is shown in the first chromatogram on the coding strand. The sequence from affected individual PID 14 (heterozygote; L277P/fs2) is shown in the second chromatogram. N = C and T.


110
120
130
140
C T T G G A C C A G C T G C T T A A C C T A G G A A G G C A G C T A

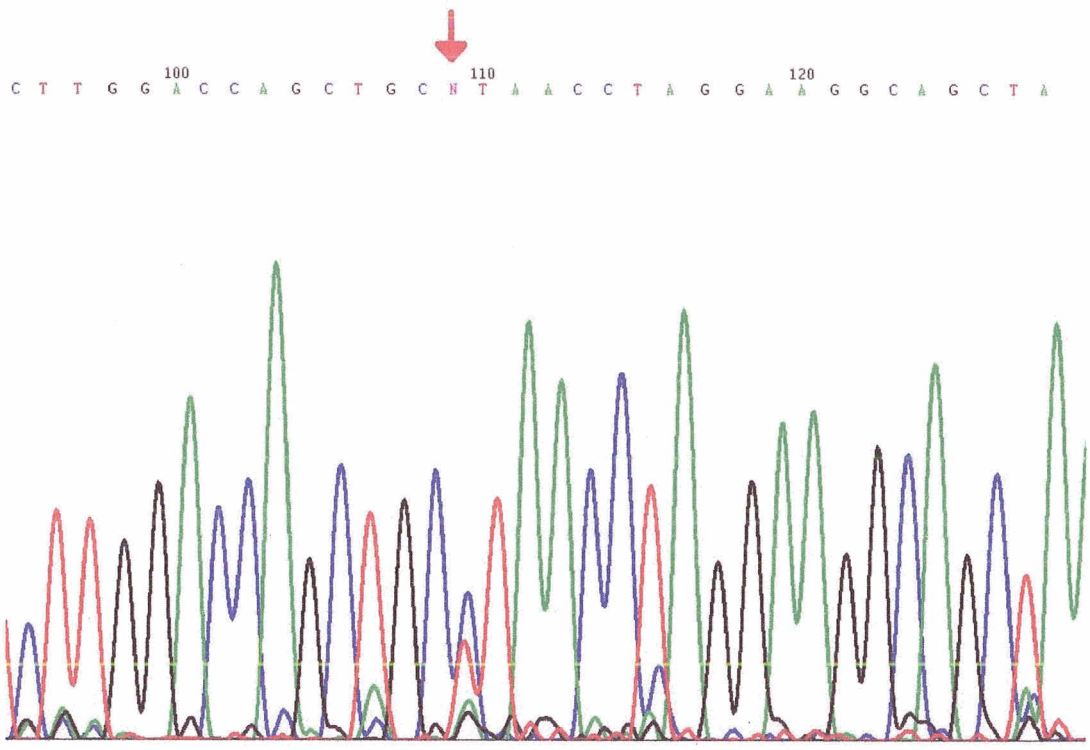

100
110
120
C T T G G A C C A G C T G C N T A A C C T A G G A A G G C A G C T A

Figure 3.7 Pedigree of family NF-B13 with corresponding BBS6 haplotypes. Individual PID's are indicated under each individual.

Key:

unaffected individual

affected individual

carrier

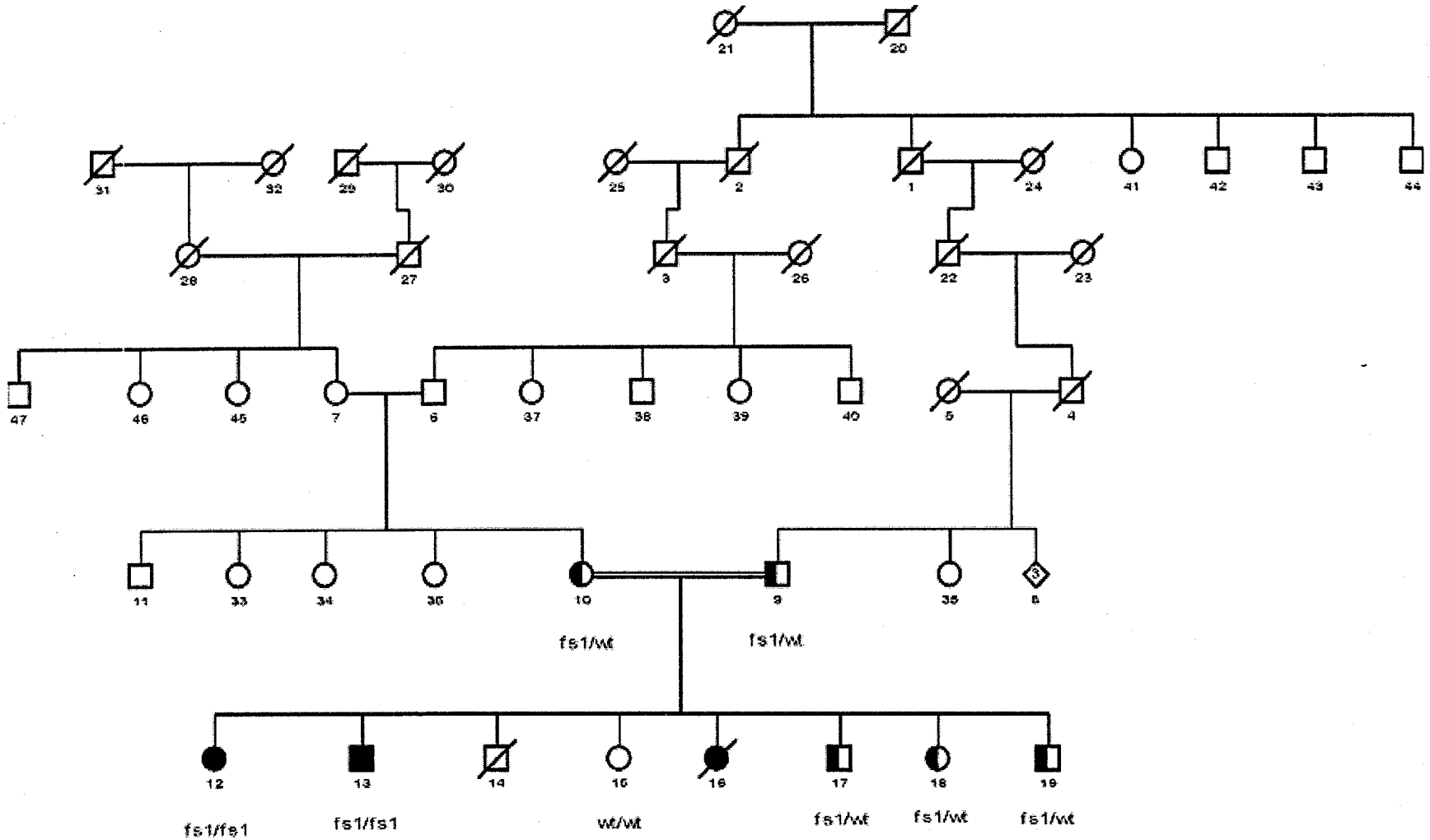

21
20
31
32
29
30
25
2
1
24
41
42
43
44
28
27
3
26
22
23
47
46
45
7
6
37
38
39
40
5
4
11
33
34
36
10
9
35
8
fs1/wt
fs1/wt
12
13
14
15
16
17
18
19
fs1/fs1
fs1/fs1
wt/wt
fs1/wt
fs1/wt
fs1/wt

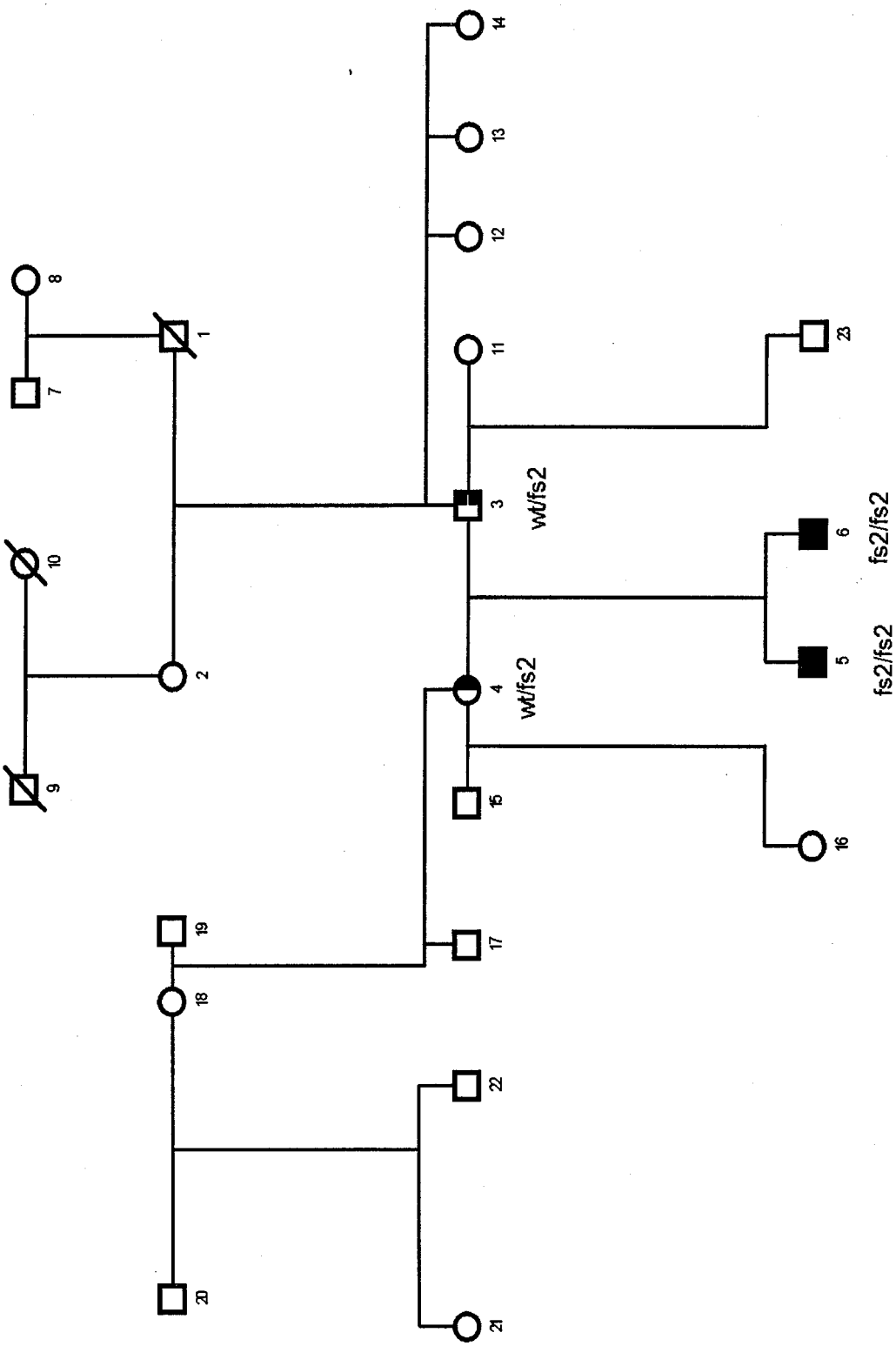
Figure 3.8 Pedigree of family NF-B16 with corresponding BBS6 haplotypes. Individual PID's are indicated under each individual.

Key:

unaffected individual

affected individual

carrier


14
13
12
8
1
11
23
7
wt/fs2
3
6
fs2/fs2
10
fs2/fs2
5
2
wt/fs2
4
9
15
16
19
17
18
22
20
21

expect most BBS patients to have the same mutation at the same locus. This is clearly not the case, as five BBS loci have been found within the affected population. It is difficult to explain why BBS is ten times more prevalent on the island than in the general population if the theory of a common founder or founders is indeed unfounded.

Shortly after BBS6 was found, Katsanis *et al.* (2001) screened all known BBS families for BBS6 mutations, irrespective of previous characterizations. Triallelism was thus discovered, and family NF-B14 helped to demonstrate this concept. The affected individual in this family is homozygous for the Y24X BBS2 mutation, and is also heterozygous for the BBS6 A242S mutation. Section 3.2 will describe this topic in greater detail.

As part of the triallelism study, members of families whose affected individuals were characterized with a specific BBS mutation were screened to determine if they carried one of the four mutations in the BBS6 gene known to occur in this population. The DNA of four new families (NF-B20, NF-B21, NF-B23, and NF-B25) became available during the course of this thesis, and they were thus screened for known BBS mutations. The distribution of the locations of all known BBS families on the island of Newfoundland is shown in Figure 3.9. The frequency of the different BBS types is depicted in Figure 3.10. The known mutations in the Newfoundland population are listed in Table 3.2. All mutational analyses of the previously identified families and the new families are summarized in this section.

It is important to fully screen BBS families for several reasons. Screening of known mutations helps to determine the prevalence of certain mutations in different regions around the world. It also aids in determining if modified modes of inheritance such as triallelism and tetra-allelism are indeed real forms of inheritance. Since there has been no genotype-phenotype correlation found in the case of BBS, finding more families with known mutations and new mutations could help to statistically support the conclusion that there is no correlation, or may provide evidence to support some correlations.

3.2 Triallelism in the Newfoundland Population

As part of the continuing study of Newfoundland BBS families, analysis was carried out by sequencing all known BBS mutations in the Newfoundland cohort. Analysis of the four known BBS6 mutations is summarized in Table 3.3. This was carried out in conjunction with screens of BBS1 and BBS2 mutations (Yanli Fan, unpublished data). This analysis was done for a number of reasons, as listed in the previous section. One area of importance was the issue of triallelism and its prevalence in the Newfoundland population. Katsanis *et al.* (2001) discuss family NF-B14, which is homozygous for the BBS2 Y24X mutation, and is also heterozygous for the BBS6 A242S mutation. No other Newfoundland families to date have been found to demonstrate such an inheritance pattern.

Figure 3.9 Map of the island portion of the province of Newfoundland and Labrador. Locations of BBS families are indicated. Families are referred to as "NF-" before their numbers throughout the thesis: i.e., NF-B1 for the B1 family.

Key: red – BBS1
orange – BB
green – BBS3
purple – BBS5
blue – BBS6
brown - unknown

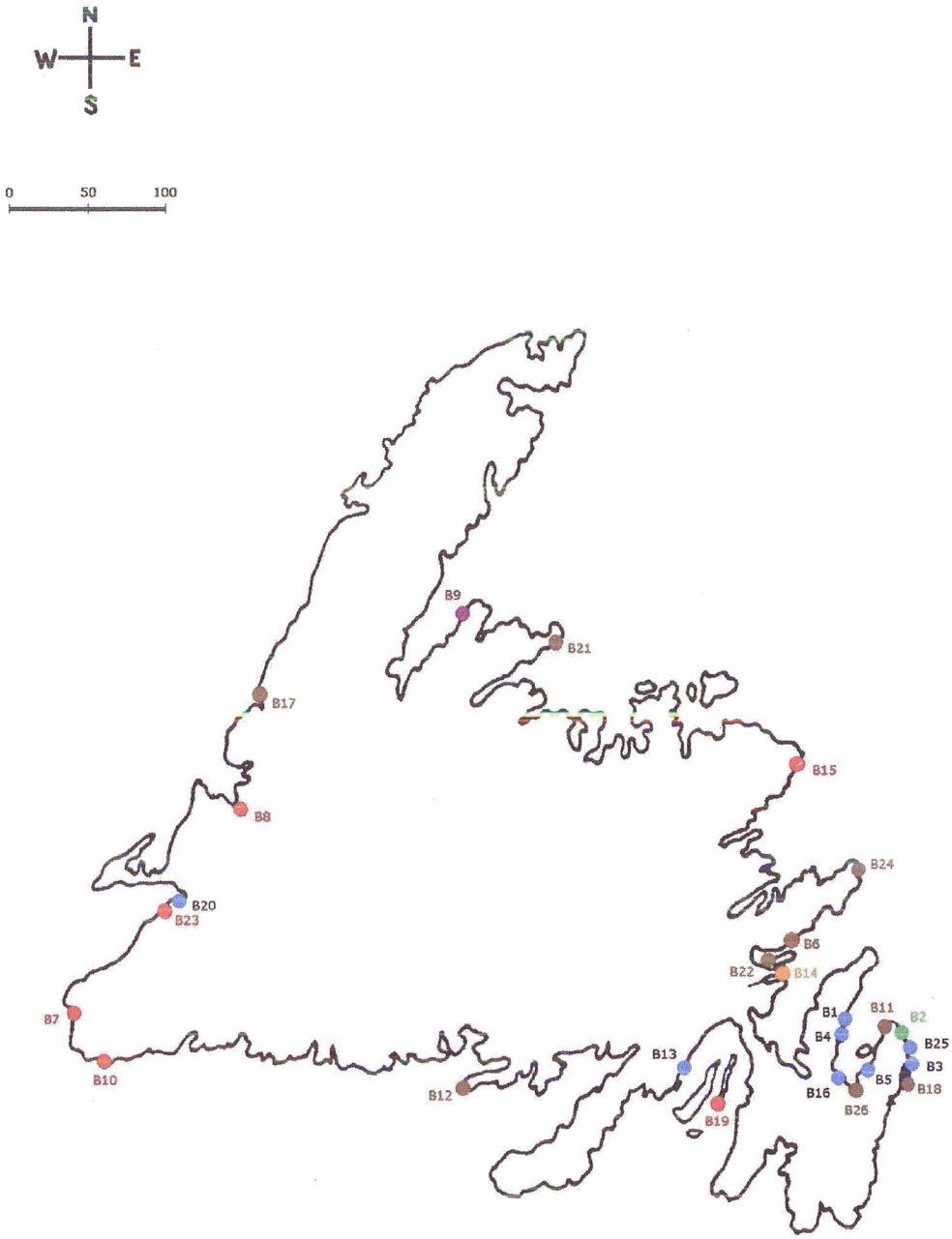

N
W
E
S
0
50
100
B9
B21
B17
B15
B8
B24
B20
B23
B6
B22
B14
B1
B11
B2
B7
B4
B25
B13
B3
B10
B5
B18
B12
B16
B26
B19

Figure 3.10 Pie chart summarizing the number of BBS types on the island of Newfoundland. The chart summarizes the findings from 26 families that have been diagnosed with BBS.

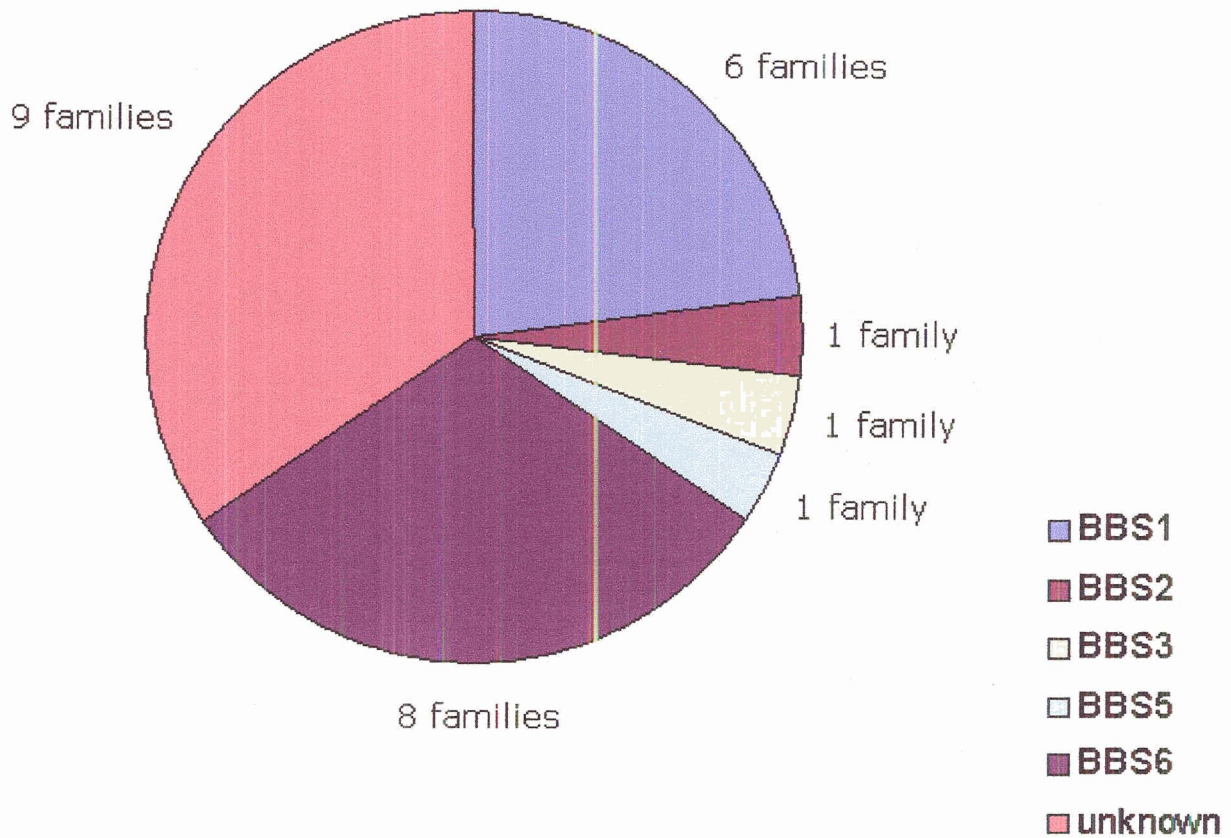

6 families
1 family
1 family
1 family
8 families
9 families
BBS1
BBS2
BBS3
BBS5
BBS6
unknown

Table 3.2 Known BBS mutations in the Newfoundland population.

BBS Gene	Mutations in the Newfoundland Population
BBS1	M390R
BBS2	Y24X
BBS6	429ΔCT433ΔAG (fs1), 280ΔT (fs2), L277P, A242S

Table 3.3 Summary of results after screening affected individuals for the following BBS6 mutations: fs1 and fs2 using BBS6X3d primers, L277P using BBS6X3d primers, and A242S using BBS6X3c primers. Known BBS6 mutations in affected individuals and carriers are as indicated. Any results indicated in blue italics are from Katsanis *et al.* (2000).

DNA Number	Family	known mutations	L277P	fs1	fs2	A242S
681	NF-B1	fs1/fs2	*wt*	hetero	*hetero*	*wt*
115/738B	NF-B2		wt	wt	wt	wt
112	NF-B3	fs2/fs2	wt	*wt*	homo	wt
126	NF-B4	fs2/fs2	wt	*wt*	homo	wt
176	NF-B5	L277P/fs2	*hetero*	*wt*	hetero	wt
645	NF-B6		wt	wt	wt	wt
669/739A	NF-B7		wt	wt	wt	wt
689	NF-B8		wt	wt	wt	wt
671	NF-B9		wt	wt	wt	wt
683	NF-B10		wt	wt	wt	wt
668	NF-B11		wt	wt	wt	wt
685	NF-B13	fs1/fs1	wt	homo	wt	wt
208	NF-B15		wt	wt	wt	wt
385	NF-B16	fs2/fs2	*wt*	wt	homo	wt
475	NF-B19		wt	wt	wt	wt
648	NF-B20	fs1/fs1	wt	homo	wt	wt
649	NF-B21		wt	wt	wt	wt
725	NF-B23		wt	wt	wt	wt
730	NF-B25	fs2/fs2	wt	wt	homo	wt

In the process of examining the possibility of the existence of other Newfoundland triallelic families, families NF-B20 to NF-B26 were diagnosed and recorded as new Newfoundland BBS families. NF-B20, NF-B21, NF-B23, and NF-B25 had DNA available and were thus screened for BBS6 mutations. As a result, affected individuals in NF-B21 and NF-B25 were found to be homozygous for the BBS6 mutations fs1 and fs2, respectively. NF-B21 was excluded from BBS1, BBS2, or BBS6. Affected individuals in NF-B23 were found to be homozygous for the BBS1 M390R mutation (Yanli Fan, unpublished results). This analysis helped to determine that 31 percent of Newfoundland families are considered to be BBS6 families, excluding the B14 family which is the only family considered to be a BBS2 family. This can be seen in Figure 3.9. This statistic is significantly different from the worldwide prevalence of BBS6, which is closer to four per cent (Beales *et al.*, 2001).

The frequency of the different mutations of the BBS6 gene varies considerably. There are eight known BBS6 families, each carrying two BBS mutations. The triallelic family, NF-B14, also carries one mutation. Out of the seventeen mutated alleles, five are fs1, ten are fs2, one is L277P, and one is A242S. Though the number of alleles is small, it is easily recognized that 59 per cent of BBS6 alleles in the Newfoundland population are fs2.

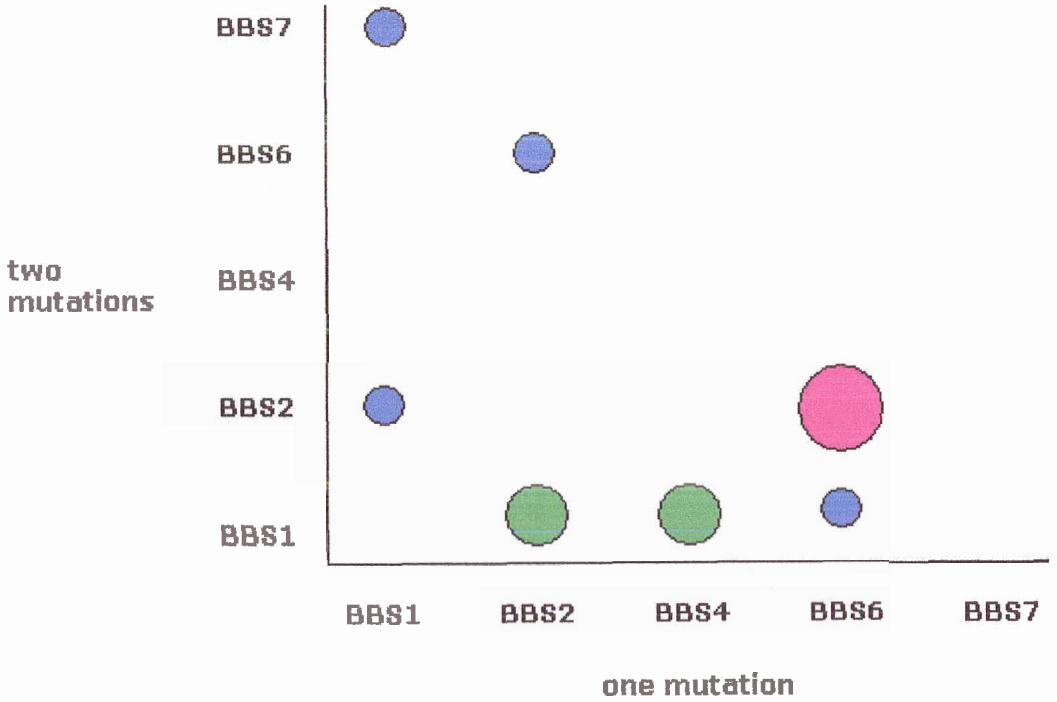
A summary of the different combinations of mutations in identified triallelic BBS families is shown in Table 3.4. A simplified depiction is illustrated in Figure 3.11.

Beales *et al.* (2003) recently reported on three families named PB056, AR396, and AR241. All families had a minimum of two affected individuals. They found that affected siblings did not have the same genotype. In these families, all siblings had two mutations at one locus, but one of the affected siblings had yet another mutation at a second locus. The researchers noted that the phenotype of the sibling with triallelism was more severe than that of their affected sibling(s) with two mutations. Information on these families is summarized in Figure 3.5.

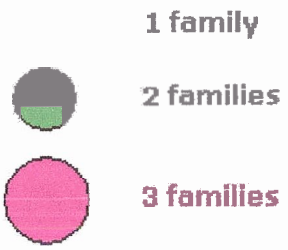
Table 3.4 Summary of examples of BBS triallelism found to date. Eleven families have been described with various combinations of BBS mutations, as more clearly seen in Figure 3.11.

Reference	Family	Allele 1	Allele 2	Allele 3
Katsanis *et al.* (2001)	AR124	Homozygous haplotype (BBS1)	Homozygous haplotype (BBS1)	fsX200 (BBS2)
Katsanis *et al.* (2001)	AR238	Homozygous haplotype (BBS1)	Homozygous haplotype (BBS1)	D104A (BBS2)
Katsanis *et al.* (2002)	AR153	Homozygous haplotype (BBS1)	Homozygous haplotype (BBS1)	L327P (BBS4)
Beales *et al.* (2003).	PB056	M390R (BBS1)	M390R (BBS1)	M472V (BBS4)
Beales *et al.* (2003)	AR396	M390R (BBS1)	Q291X (BBS1)	S236P (BBS6)
Katsanis *et al.* (2001)	NFB14	Y24X (BBS2)	Y24X (BBS2)	A242S (BBS6)
Katsanis *et al.* (2001)	AR579	fsX200 (BBS2)	R316X (BBS2)	C499S (BBS6)
Katsanis *et al.* (2001)	AR259	Q59X (BBS2)	Y24X (BBS2)	Q147X (BBS6)
Beales *et al.* (2003)	AR241	IVSx2 with R315Q (BBS2)	R315Q (BBS2)	M390R (BBS1)
Katsanis *et al.* (2001)	AR237	Y37C (BBS6)	Y37C (BBS6)	N70S (BBS2)
Badano *et al.* (2003)	AR69	T211I (BBS7)	T211I (BBS7)	E234K (BBS1)

Figure 3.11 Graph depicting the different combinations of mutations in triallelic BBS families.


BBS7
BBS6
BBS4
BBS2
BBS1
two mutations
BBS1
BBS2
BBS4
BBS6
BBS7
one mutation

Key:


1 family
2 families
3 families

3.3 NF-B20 Family

Individual with PID 36 in the NF-B20 family was diagnosed with BBS by Dr. Jane Green at the Memorial University of Newfoundland. The family had not been analyzed previously by molecular methods, so bloods and pedigree information were provided to the Davidson lab. The pedigree for NF-B20 is shown in Figure 3.12. As part of the search for other triallelic families in the Newfoundland population, BBS6 mutations were screened in the family through direct sequencing as described in Chapter 2 (Methods; section 2.4). This family was found to have 429ΔCT433ΔAG, named fs1, as seen in Figure 3.13. Sequencing results of known BBS6 mutations are summarized in Table 3.3.

PID 36 was also screened for the BBS1 M390R mutation, as well as the BBS2 Y24X mutation, but no mutations were found (Yanli Fan, unpublished results).

3.4 NF-B21 Family

Individual with PID 15 in the NF-B25 family was diagnosed with BBS by Dr. Jane Green at Memorial University of Newfoundland. The family had not been analyzed previously by molecular methods, so blood samples and pedigree information were provided to the Davidson lab. The pedigree for NF-B21 is shown in Figure 3.14. As part of the search for other triallelic families in the Newfoundland population, BBS6 mutations were screened in the family through direct sequencing as described in Chapter 2 (Methods; section 2.4). None of the

four known BBS6 mutations were found in this family. Sequencing results of BBS6 known mutations are summarized in Table 3.3.

PID 15 was also screened for the BBS1 M390R mutation, as well as the BBS2 Y24X mutation, but no mutations were found (Yanli Fan, unpublished results).

3.5 NF-B23 Family

Individual with PID 18 in the NF-B23 family was diagnosed with BBS by Dr. Jane Green at Memorial University of Newfoundland. The family had not been analyzed previously by molecular methods, so blood samples and pedigree information were provided to the Davidson lab. The pedigree for NF-B25 is shown in Figure 3.15. As part of the search for other triallelic families in the Newfoundland population, BBS6 mutations were screened in the family through direct sequencing as described in Chapter 2 (Methods; section 2.4). Sequencing results of BBS6 known mutations are summarized in Table 3.3.

PID 18 was also screened for the BBS1 M390R mutation, as well as the BBS2 Y24X mutation. The affected individual is homozygous for the BBS1 mutation (Yanli Fan, unpublished results).

Table 3.5 Three BBS families have been described by Badano *et al.* (2003), whose affected individuals share two mutations at one locus but may differ by the presence or absence of a third mutation at a second locus.

Reference	Family	Allele 1	Allele 2	Allele 3
Badano *et al.* (2003)	AR768	M390R (BBS1)	L548fsX579 (BBS1)	T325F (BBS6)
Badano *et al.* (2003)	PB009	M390R (BBS1)	M390R (BBS1)	L349W (BBS2)
Badano *et al.* (2003)	PB061	R275X (BBS2)	R275X (BBS2)	IVS115+2T→C(BBS1)

Figure 3.12 Pedigree of family NF-B20 with corresponding BBS6 haplotypes. Individual PID's are indicated under each individual.

Key:

unaffected individual

affected individual

carrier

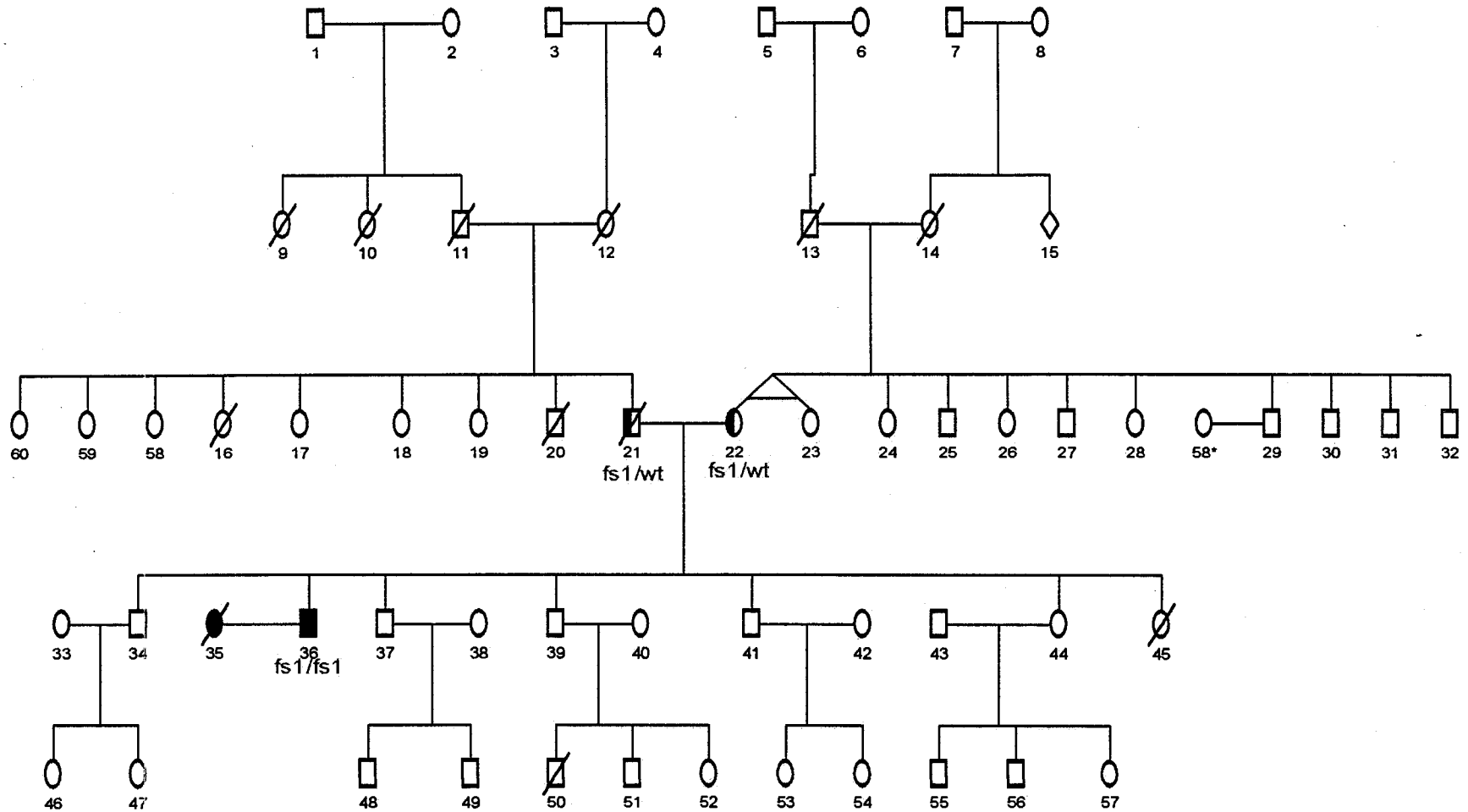
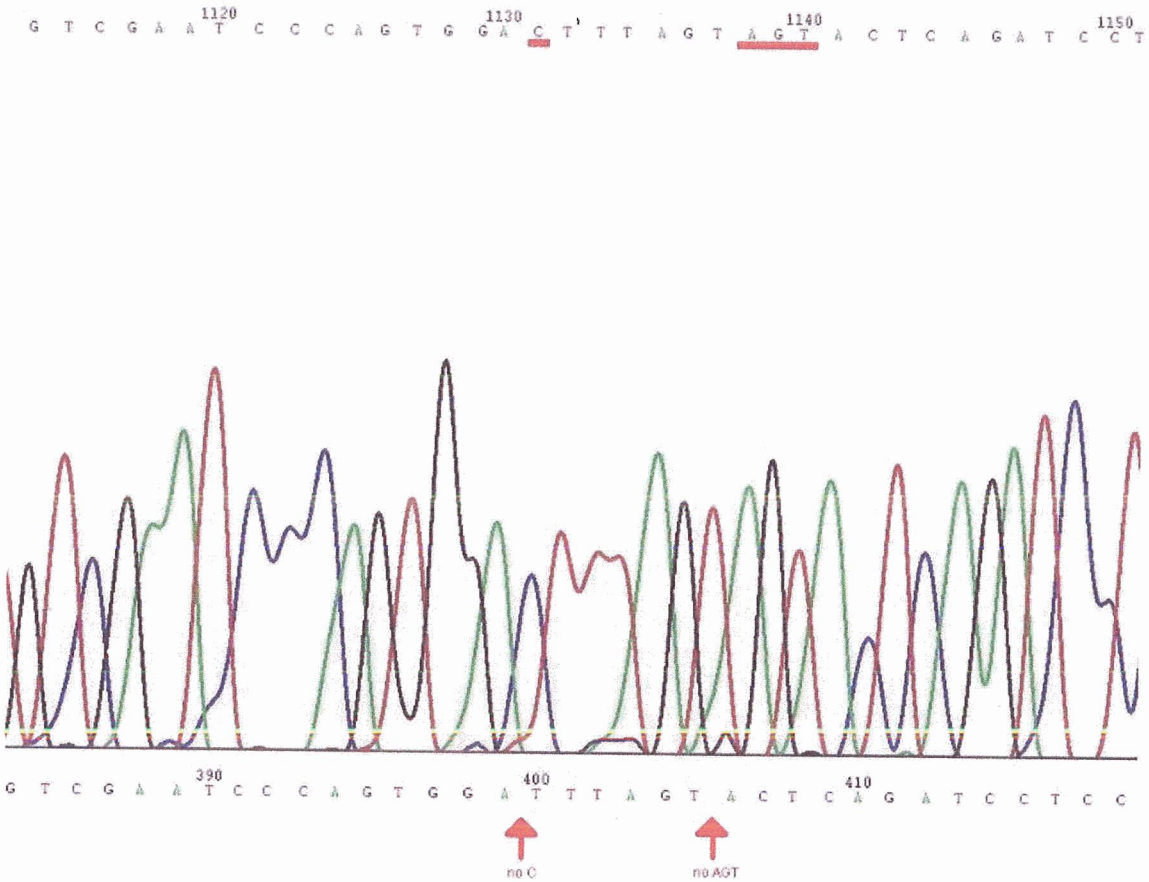

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
60
59
58
16
17
18
19
20
21
fs1/wt
22
fs1/wt
23
24
25
26
27
28
58*
29
30
31
32
33
34
35
36
fs1/fs1
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57

Figure 3.13 Chromatograms showing the BBS6 mutation 429ΔCT433ΔAG (fs1) in family NF-B20. A wild type sequence is shown in the first chromatogram. The sequence of affected individual PID 36 (homozygote) is shown in the second chromatogram, and an example of a heterozygote is shown in the third chromatogram. Base pairs missing in the affected individual are underlined in the wild type sequence.


1120
1130
1140
1150
390
400
410
no C
no AGT

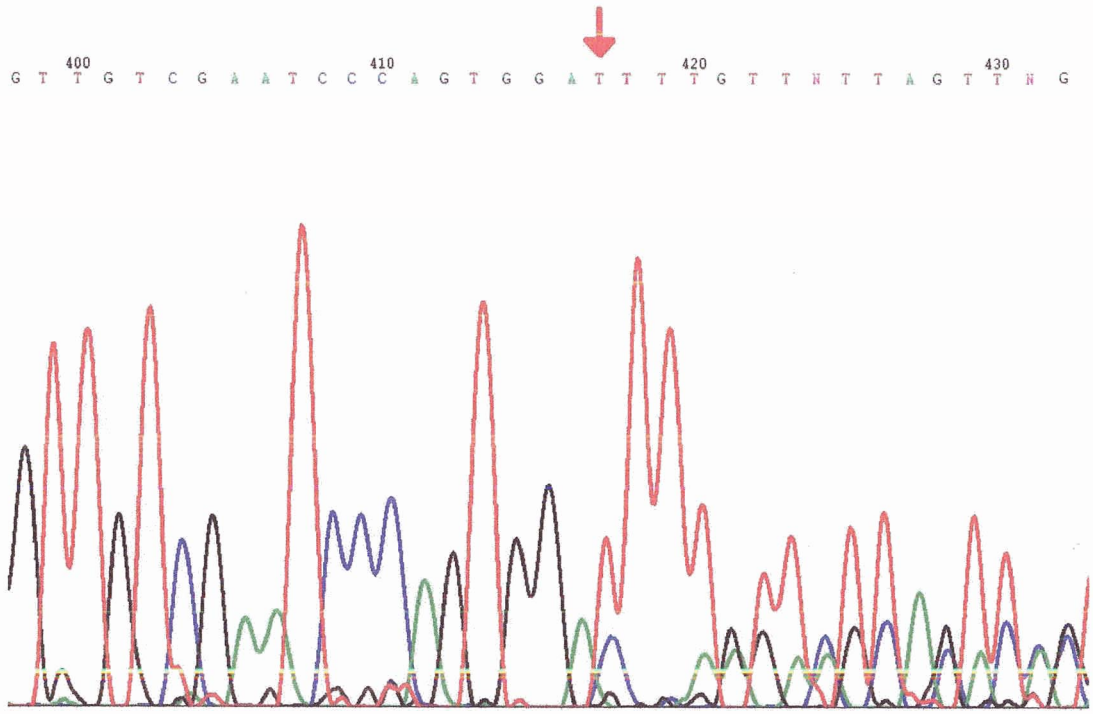

400
410
420
430
G T T G T C G A A T C C C A G T G G A T T T T G T T N T T A G T T N G

Figure 3.14 Pedigree of family NF-B21. Individual PID's are indicated under each individual.

Key:

unaffected individual

affected individual

carrier

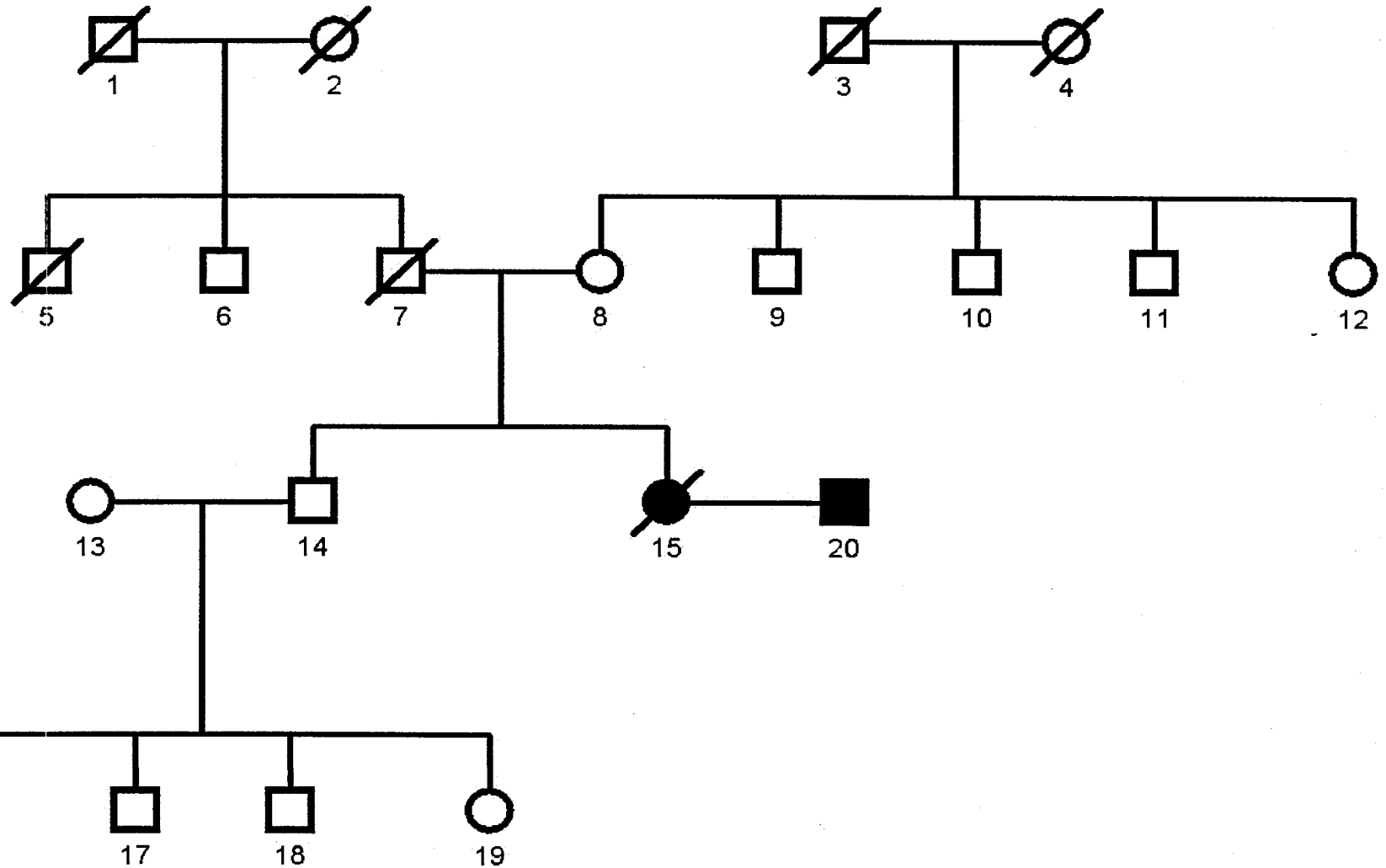
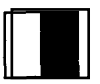

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
20
16
17
18
19

Figure 3.15 Pedigree of family NF-B23. Individual PID's are indicated under each individual.

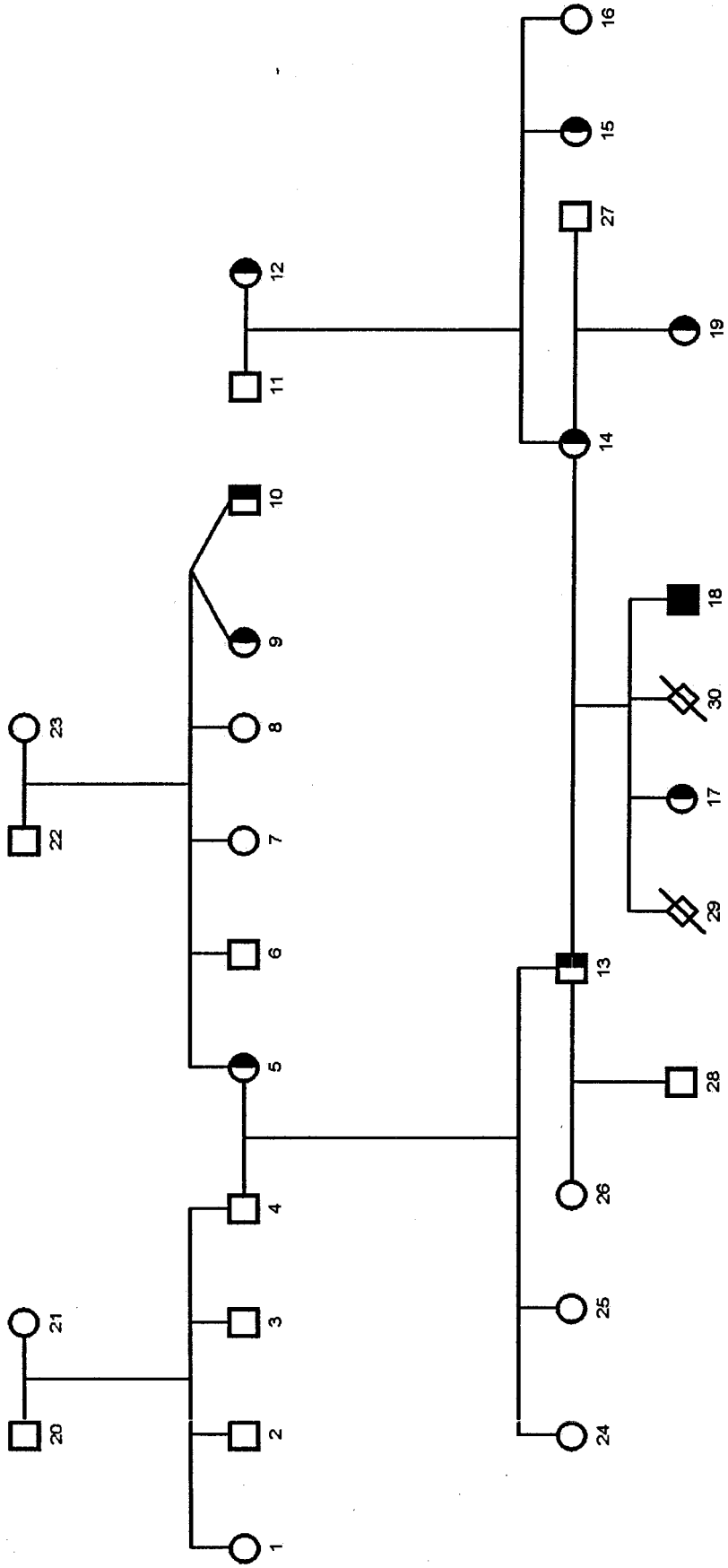
Key:

unaffected individual

affected individual



carrier


16
15
27
12
19
11
14
10
18
9
30
23
8
17
22
7
29
6
13
5
28
4
26
21
3
25
20
2
24
1

3.6 NF-B25 Family

Individual with PID 19 in the NF-B25 family was diagnosed with BBS by Dr. Jane Green at Memorial University of Newfoundland. The family had not been analyzed previously by molecular methods, so blood samples and pedigree information were provided to the Davidson lab. The pedigree for NF-B25 is shown in Figure 3.16. As part of the search for other triallelic families in the Newfoundland population, BBS6 mutations were screened in the family through direct sequencing as described in Chapter 2 (Methods; section 2.4). This family was found to have 280ΔT, named fs2, as seen in Figure 3.17. Sequencing results of BBS6 known mutations are summarized in Table 3.3.

PID 19 was also screened for the BBS1 M390R mutation, as well as the BBS2 Y24X mutation, but no mutations were found (Yanli Fan, unpublished results).

Figure 3.16 Pedigree of family NF-B25 with corresponding BBS6 haplotypes. Individual PID's are indicated under each individual.

Key:

unaffected individual

affected individual

carrier

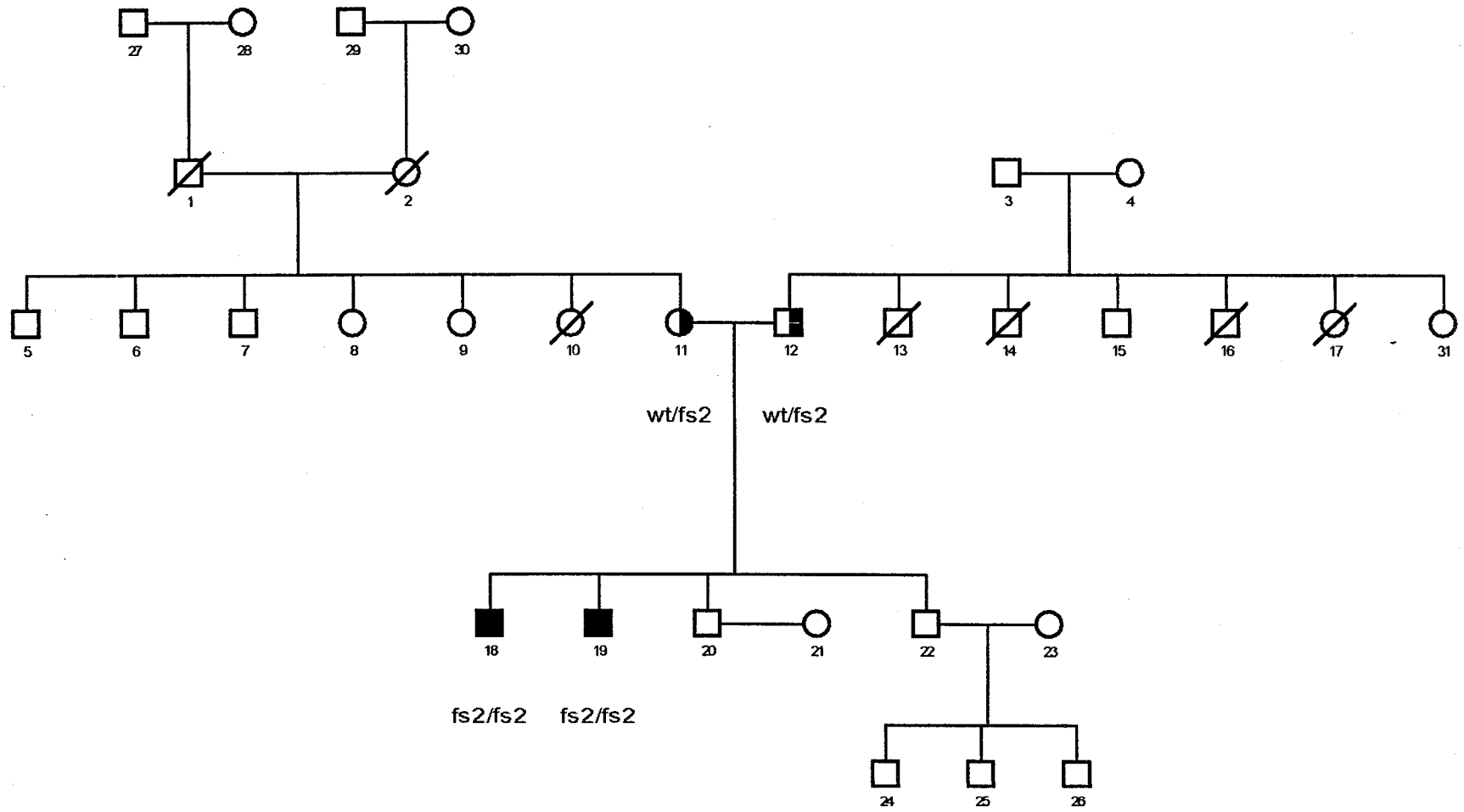
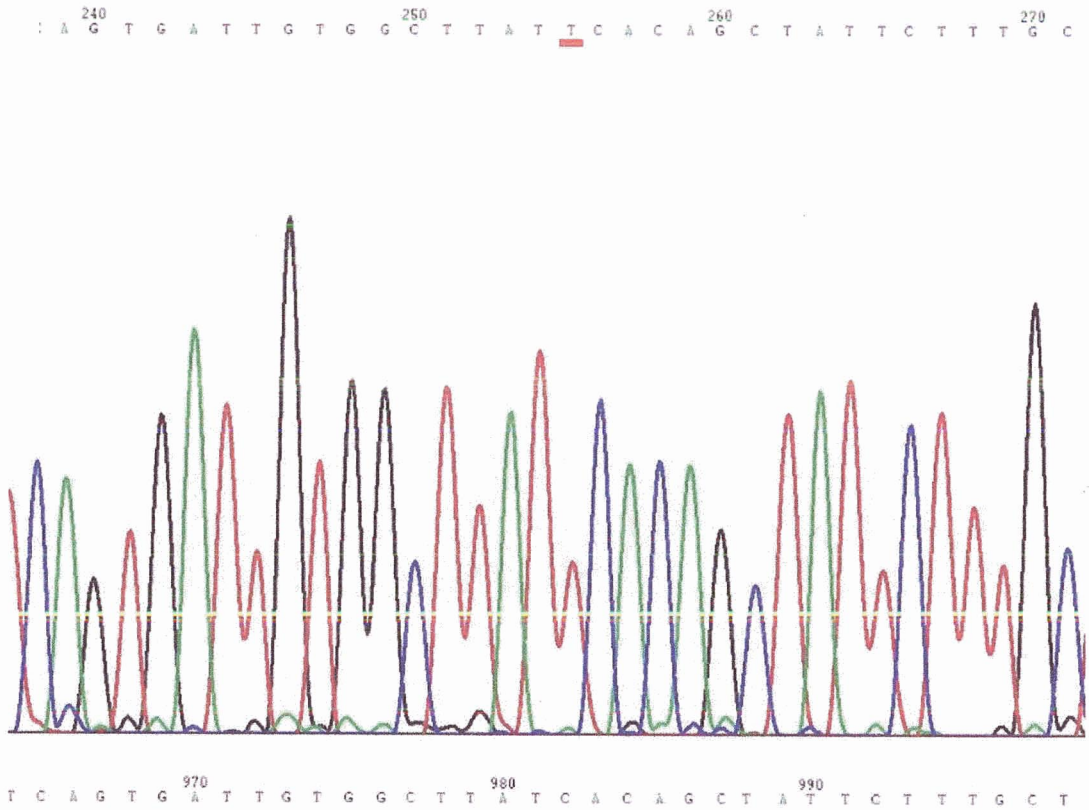
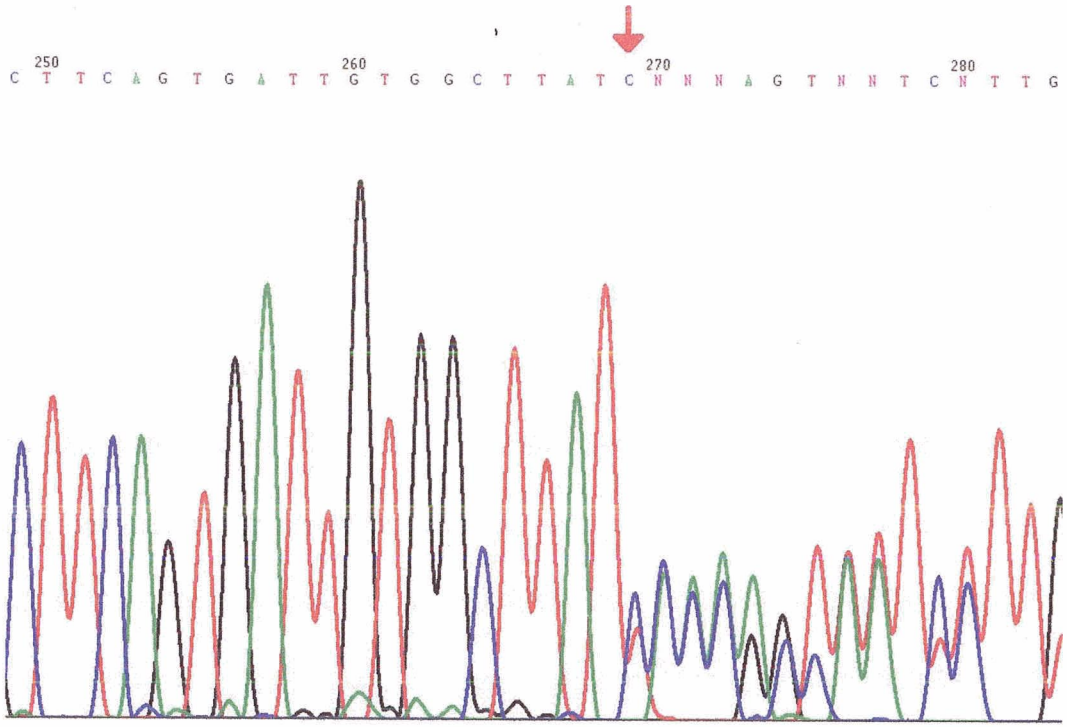

27
28
29
30
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
31
wt/fs2
wt/fs2
18
19
20
21
22
23
fs2/fs2
fs2/fs2
24
25
26

Figure 3.17 Chromatograms showing the BBS6 mutation 280ΔT (fs2) in family NF-B25. A wild type sequence is shown in the first chromatogram. The sequence from affected individual, PID 19, is shown in the second, and an example of a heterozygote is shown in the third. The base pair that is missing in the affected individual is underlined in the wild type sequence.


240
250
260
270
: A G T G A T T G T G G C T T A T T C A C A G C T A T T C T T T G C
970
980
990
T C A G T G A T T G T G G C T T A T C A C A G C T A T T C T T T G C T


250
260
270
280
C T T C A G T G A T T G T G G C T T A T C N N N A G T N N T C N T T G

3.7 Phenotype of the Newfoundland Cohort

The Newfoundland BBS cohort has been followed for twenty-two years and as part of the study, three assessments have been carried out on forty-six patients from twenty-three families in 1986, 1993, and 2001. In order to consider an individual affected with BBS, the person would have to manifest four cardinal features or a sibling of an affected person would have to have three of these manifestations to be considered affected. The study did not find any genotype-phenotype correlation. However, the study has provided valuable information about the phenotype of the disease in the Newfoundland population.

Ninety-one percent of patients were blind by the end of their teen years, with the median age of onset being eighteen years. It is known that the most common phenotypic display is retinal dystrophy in BBS patients (Green *et al.,* 1989). All patients demonstrated brachydactyly in the feet, and eighty-six per cent in the hands. Brachdactyly is where one has abnormally shortened digits. Also, ninety-five per cent showed syndactyly, or webbing between digits, and sixty-three per cent showed polydactyly, or extra digits. Fourty-eight per cent of patients had diabetes mellitus, with the median age of onset being fourty-three years. Obesity was measured by body mass index, or BMI. The mean BMI over time for forty-four of the forty-six patients was 35.5 kg/m^2. All patients were obese at some point in their lives. Morbid obesity was present in twenty-five per cent of patients. Twenty per cent of patients had moderate chronic renal failure at the median age of 57.6 years. Seven of the fourty-six patients had mild

chronic renal failure, and four had end stage renal disease. The BBS2 patient did not display any signs of renal failure. Sixty-seven per cent of patients had hypertension with a median age of onset at thirty-four years.

In terms of genital and reproductive abnormalities, ninety-two per cent of males demonstrated small penile length, eleven per cent had undescended testes, eight per cent had hypospadias, eight per cent had phimosis, eight per cent had recurrent urethral strictures, and four per cent had posterior urethral valves. In females, ten per cent had vaginal atresia, twenty-five per cent had hypoplastic Labia Minora, and five per cent did not have a urethral opening. Two women with BBS1 each gave birth to one offspring who did not display any features of BBS.

Other medical conditions the researchers found were speech and neurological abnormalities such as impaired limb co-ordination, abnormal facial movements, and spasticity. Twenty-eight percent of patients also had asthma and thirty per cent had some form of psychiatric disease. The median age of death for affected individuals is 62.7 years.

These results are described in more detail in a manuscript (Moore *et al.*, 2003) that has been submitted for publication.

Chapter 4: Candidate Gene Screening for Bardet-Biedl Syndrome 5

4.1 Rationale

The locus for BBS5 was identified by Young *et al.* (1999) using homozygosity mapping with family NF-B9 (Figure 4.1). Recombinations observed through haplotype analysis allowed a critical region of 13 cM between markers D2S156 and D2S1238 to be determined (Figure 4.2). Beales *et al.* (2001) confirmed the BBS5 locus of Young *et al.* (1998) by showing that the same homozygous region existed in affected individuals in a series of BBS families (AR-199, AR-274, and AR-082) which had been excluded at the known BBS loci. However, the critical region was not narrowed any further as no recombinations were observed in these families in this segment of the genome. As the BBS5 gene has not been identified to date, it was decided to start screening candidate genes. Two genes were screened: MYO3B and GORASP2.

4.1 MYO3B – Myosin 3B

Class III myosins are highly expressed in retinal photoreceptors. It is known that a mutation in the class III myosin NINAC causes retinal degeneration in *Drosophila melanogaster* (Dose *et al.*, 2002). The human homolog of NINAC is MYO3B, and thus Dose *et al.* (2002) proposed that this gene is a prime candidate gene for diseases whose phenotypes

Figure 4.1 Pedigree of family NF-B9. Individual PID's are indicated under each individual.

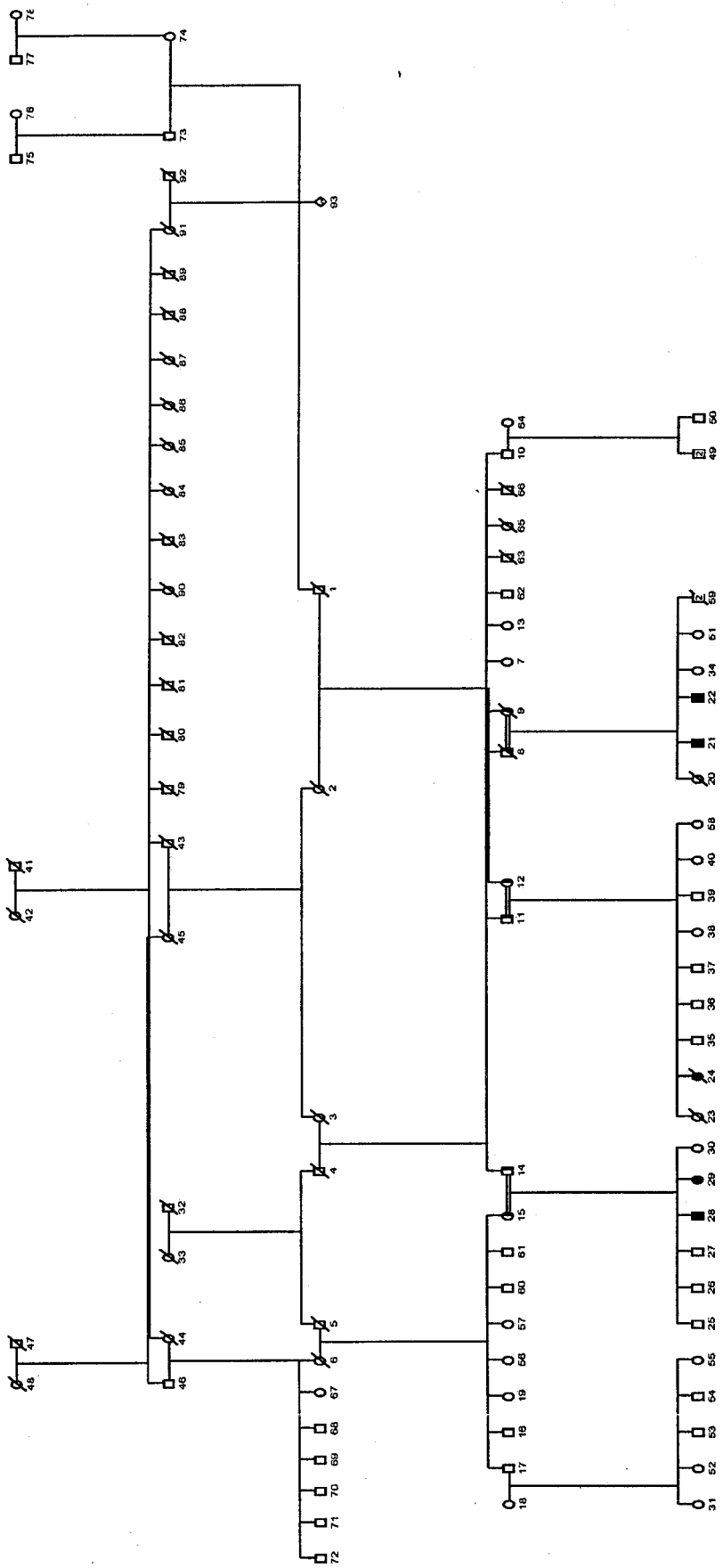
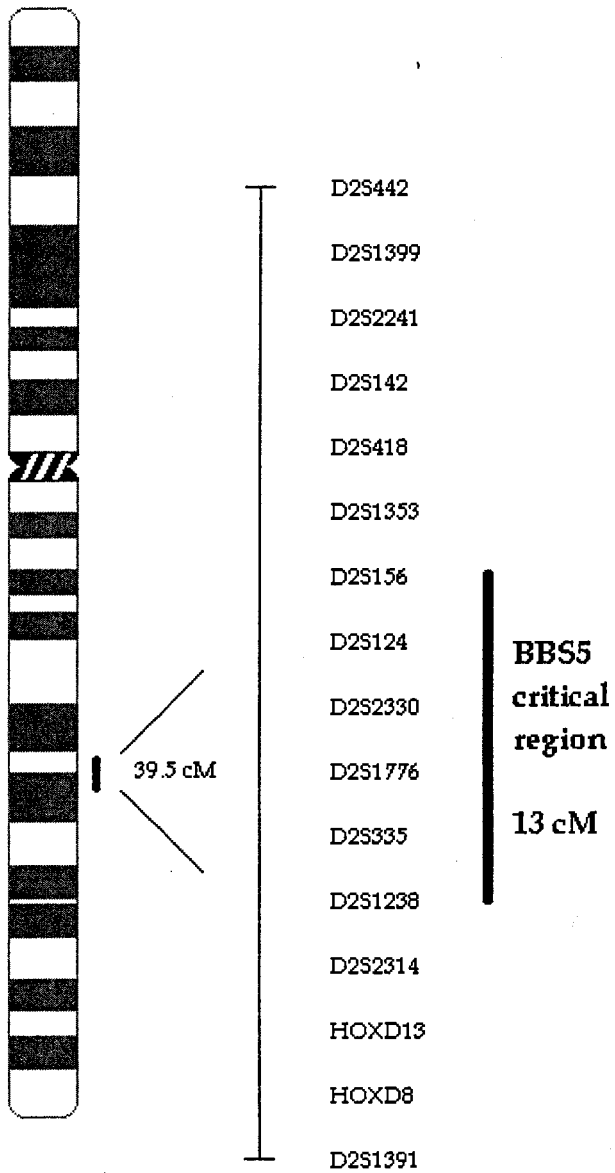


Figure 4.2 Adapted from Young *et al.* (1998). The BBS5 critical region on chromosome 2 as determined by Young *et al.* (1998) using family NF-B9. The distance 13 cM translates to 14 Mb.

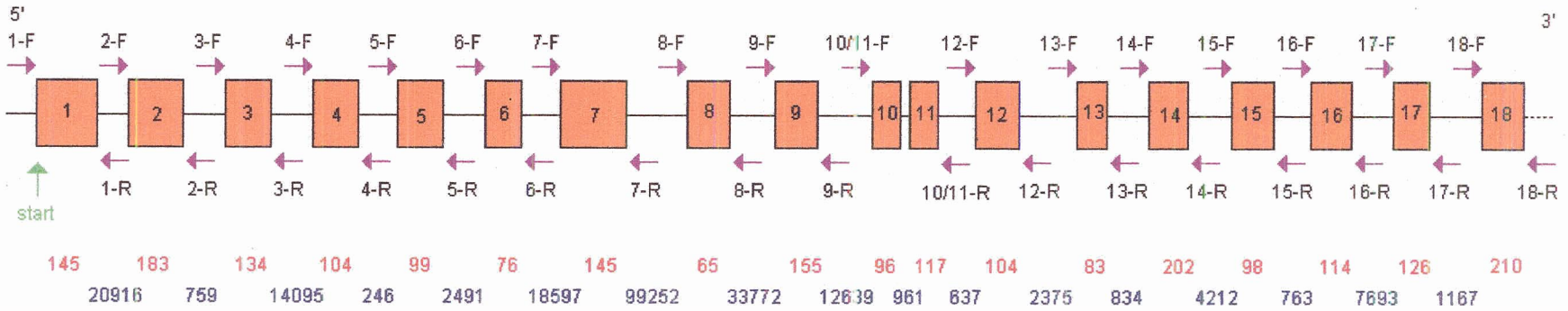

D2S442
D2S1399
D2S2241
D2S142
D2S418
D2S1353
D2S156
D2S124
D2S2330
39.5 cM
D2S1776
D2S335
D2S1238
D2S2314
HOXD13
HOXD8
D2S1391
BBS5
critical
region
13 cM

include retinal degeneration. Dose *et al.* (2002) found the MYO3B transcript present in the human retina, kidney, and testis through northern blotting. The researchers noted these tissues are affected in BBS, and they also noted that the MYO3B gene maps to 2q31.1-q31.2, which overlaps with the critical region for BBS5 (Figure 4.3). Thus, the authors proposed that MYO3B is a prime candidate for BBS5 but they did not carry out any mutation analyses of BBS5 patients. A review of the literature did not reveal any other proposed candidate genes for BBS5.

4.3 GORASP2 – Golgi reassembly stacking protein 2

The gene GORASP2 was considered a candidate gene for BBS5 because it maps to the critical region of BBS5 and its function, when mutated, could cause a phenotype similar to that of BBS (Figure 4.4). GORASP stands for Golgi reassembly stacking protein, and encodes the protein commonly known as GRASP55. Factors such as GRASP55 have been shown *in vitro* to recreate the Golgi apparatus after mitosis through reassembly of the cisternae (Short *et al.*, 2001). The Golgi apparatus is known to play a key role in the modification and distribution of proteins throughout the cell (Voet *et al.*, 1998). Therefore, it is plausible that disruption of proper Golgi reassembly could lead to the improper distribution of proteins within the cell, thus creating a pleiotropic phenotype phenotype similar to what might be anticipated with mutations in factors associated with protein folding (e.g., chaperonins, BBS6).

Figure 4.3 MYO3B gene structure. This open reading frame is located on chromosome 2q31, and is comprised of thirty-six exons. The red numbers are the sizes of the exons in base pairs. The blue numbers are the sizes of the introns in base pairs. Forward and reverse primers used to amplify segments and then used for sequencing are shown with rightward and leftward arrows and names displayed above. Arrows show start and stop codons.


5'
3'
1-F
2-F
3-F
4-F
5-F
6-F
7-F
8-F
9-F
10/11-F
12-F
13-F
14-F
15-F
16-F
17-F
18-F
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
start
1-R
2-R
3-R
4-R
5-R
6-R
7-R
8-R
9-R
10/11-R
12-R
13-R
14-R
15-R
16-R
17-R
18-R
145
183
134
104
99
76
145
65
155
96
117
104
83
202
98
114
126
210
20916
759
14095
246
2491
18597
99252
33772
12639
961
637
2375
834
4212
763
7693
1167

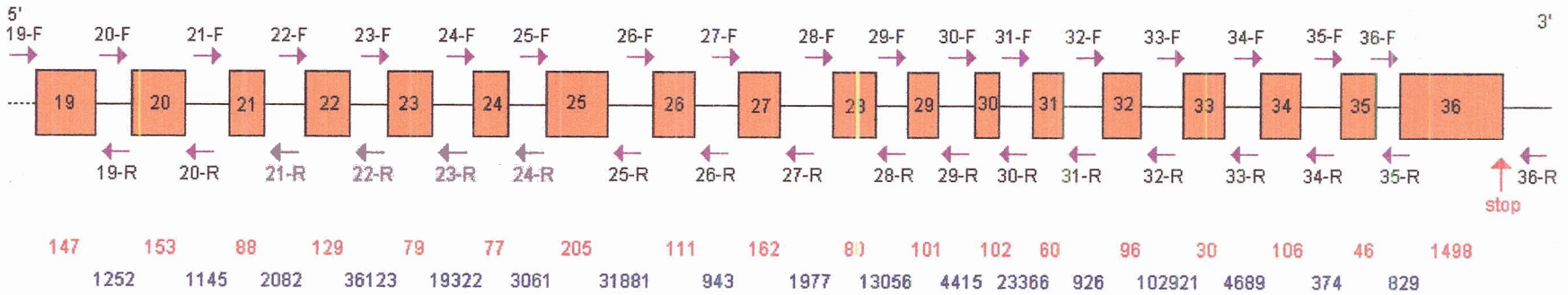
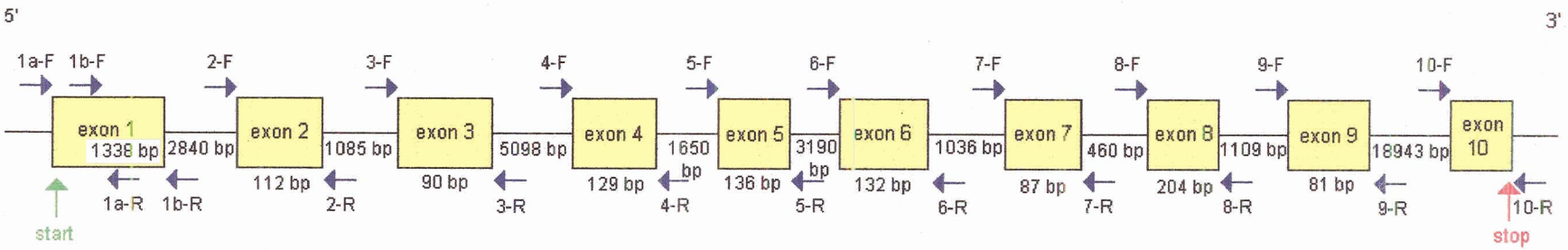

5'
3'
19-F
20-F
21-F
22-F
23-F
24-F
25-F
26-F
27-F
28-F
29-F
30-F
31-F
32-F
33-F
34-F
35-F
36-F
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
19-R
20-R
21-R
22-R
23-R
24-R
25-R
26-R
27-R
28-R
29-R
30-R
31-R
32-R
33-R
34-R
35-R
36-R
stop
147
153
88
129
79
77
205
111
162
80
101
102
60
96
30
106
46
1498
1252
1145
2082
36123
19322
3061
31881
943
1977
13056
4415
23366
926
102921
4689
374
829

Figure 4.4 GORASP2 gene structure. This open reading frame is located on chromosome 2q31, and comprises ten exons. Sizes of exons and introns are indicated in base pairs. Forward and reverse primers are shown with rightward and leftward arrows and names displayed above. Arrows show start and stop codons.


5'
3'
1a-F
1b-F
2-F
3-F
4-F
5-F
6-F
7-F
8-F
9-F
10-F
exon 1
1338 bp
2840 bp
exon 2
1085 bp
exon 3
5098 bp
exon 4
1650 bp
exon 5
3190 bp
exon 6
1036 bp
exon 7
460 bp
exon 8
1109 bp
exon 9
18943 bp
exon 10
112 bp
90 bp
129 bp
136 bp
132 bp
87 bp
204 bp
81 bp
1a-R
1b-R
2-R
3-R
4-R
5-R
6-R
7-R
8-R
9-R
10-R
start
stop

4.4 Results and Discussion

The exons and exon-intron boundaries of MYO3B and GORASP2 were sequenced using the primers described in Tables 4.1 and 4.2. No variation was found in the coding regions or intron-exon boundaries of MYO3B or GORASP2 in affected individuals from BBS5 family, NF-B9. It was assumed that if there was a mutation in MYO3B or GORASP2, the affected individual would be homozygous for the mutation as this family is known to be highly consanguineous. If any variation had been found, it would have been screened in other family members and controls. Association of the variation with BBS would have helped to confirm the variation as the causative mutation, and unaffected individuals would be screened to determine if they were carriers of the mutation. Mutational analyses of the MYO3B and GORASP2 genes indicate that neither of these genes is probably involved in BBS5.

Though screening exons as well as exon-intron boundaries is a valid way of screening a gene, it is also plausible that there may be mutations in non-coding regions. Examples of these non-coding regions would be large introns with cryptic splice sites, enhancer sequences, and promoter regions.

Table 4.1 Oligonucleotide primers used in amplifying segments of the MYO3B gene in affected member 671 from the BBS5 family B9.

Exon	Primer Name	Sequence	Size (bp)	Annealing Temp (°C)
1	MYO3BX30-F MYO3BX30-R	CAC CAA GGA TTT GGT GGA ATC TGC CCT TTG CAG AAA A	249	45.4
2	MYO3BX31-F MYO3BX31-R	CTG CCC CCA GTT CTG TG AAG TGG AGT CCC TTC ATT CAG	359	60.9
3	MYO3BX32-F MYO3BX32-R	TTT GAA GGT ATG GTA ATG AGA TTG ACC TGT AAG AAC TGT GGA GGA G	391	60.9
4	MYO3BX33-F MYO3BX33-R	TCA TCA CCT CAA ATA AAA AAT ACA TAC AAC AAC AGG AAA GAG AAA GGG	487	51.4
5	MYO3BX34-F MYO3BX34-R	CAA AGG TCT ACT CAG ATG TGG C AAG TCC CTC AGC CTA TGG TG	491	60.9
6	MYO3BX35-F MYO3BX35-R	AGT CTG CCG ATT GCT GG GAA GGT GCC CTG TGG G	328	60.9
7	MYO3BX36-F MYO3BX36-R	AAA ACT GAG AAC GGG GCT AAG AAA AGC TGA GTG CTA AAA ATA GGC	444	53.8
8	MYO3BX37-F MYO3BX37-R	TTG CCT CAC CCC TCA TTA G ATG GAA CGC CTT ACT TTT CAC	502	53.8
9	MYO3BX29-F MYO3BX29-R	GGT TGC TTT CAT ACG GTT ACC TAT TCA TAG GGG AAC AGG GTG	535	58.5
10	MYO3BX28-F MYO3BX28-R	CAA CTG AAT GCT CCC TGA AC TGT GCC TGA CAC TGA GCA G	539	64.5
11	MYO3BX27-F MYO3BX27-R	TAT TAC ATT ACT ATC TGC AGT TTG GAG TGG GAT CAT TAG GGT CAT TG	361	51.4
12	MYO3BX26-F MYO3BX26-R	TCT CCC GAT GCT ATG GC CTA ATA AGT GAC AGA TTC TTG GTT TC	338	56.2
13	MYO3BX25-F MYO3BX25-R	TCT CCT CCC TAC GTG AAG G TCT CCT GCA GAT GAA CCT G	561	56.2
14	MYO3BX24-F MYO3BX24-R	ATC CCT AAA GCA CGC CC CTA ACG TCC CCC AAA TGG	515	64.5
15	MYO3BX23-F MYO3BX23-R	TCC TGC CAC CTC ACC C GTA GTA GGC GTT TGG GCT G	502	64.5
16	MYO3BX22-F MYO3BX22-R	CAT TTG AAA GCA AGT GTC TTA TCC TGG GCA ACA GAG TGA GGT C	290	56.2
17	MYO3BX21-F MYO3BX21-R	AGC GGA GAT TGC ACC AC TCT AAG GCA TCC TCA TCA GTG	579	64.5
18	MYO3BX20-F MYO3BX20-R	CTA AGA GAC GGC GGG G CAT CAA GCA TCT ATG AAG TAA GAG C	446	60.9
19	MYO3BX19-F MYO3BX19-R	ATT TAG CCA GGA CTG CCT TAC GTG GGT GAA TTG GGT ACT TG	356	53.8
20	MYO3BX18-F MYO3BX18-R	AAG TGG GAA GGT GTT AGG AAT AG TTA CTG GTT TAG CTG CTG CC	542	65.0

21	MYO3BX17-F MYO3BX17-R	GCA ACT TTC TGT GAG TGG C CTA AAA CTA ACT GGC AGG GC	521	63.1
22	MYO3BX16-F MYO3BX16-R	AGT TGC CAG TGT CAC CAA TG AAG TAG AAA TAA TCT TCC TGA TAA CTT TTG	523	46.8
23	MYO3BX15-F MYO3BX15-R	AGG ATT GTA AGG ATT CCA TGA TG AAA ATA CAA GGA CCC AAA GGG	315	60.9
24	MYO3BX14-F MYO3BX14-R	TGG GCA GGG TTT TGT G GGG TTC ATT AGA AGA TGT AGA AGT ATA G	489	60.9
25	MYO3BX13-F MYO3BX13-R	AGA GCA CGC AGG AGG AAC TTG GCA CTA GGT GAT GTT TAA AG	467	60.9
26	MYO3BX12-F MYO3BX12-R	TAC ATA CAG AAC AGG AGG CAA AC CAA ATG GTG ACT GAA AAG GC	317	60.9
27	MYO3BX10/11-F MYO3BX10/11-R	GGT TAG CAC CTT TAT CAG AGG AAG AGA AAC AGT TGT GTA TTC CC	582	53.8
28	MYO3BX9-F MYO3BX9-R	GTG TGG CAA ACA GCA AAT G TTC ATC AGT GGA ACC CAG AG	509	60.9
29	MYO3BX8-F MYO3BX8-R	CCC AAG ACA CTA AAC TTT CCT C AAG TGT GGA TTT TCC TGG C	542	60.9
30	MYO3BX7-F MYO3BX7-R	CCC CCA CAA AGT GCT G CTG CCG GGA CCC AAA TC	569	60.9
31	MYO3BX6-F MYO3BX6-R	GCA GCC GCA GAC AAG C AGG CAG ACT GTG GAT GTG C	332	64.5
32	MYO3BX5-F MYO3BX5-R	CCT GCC CAA ACA ACA CAC TTG CCC CAA GGA GAG G	330	56.2
33	MYO3BX4-F MYO3BX4-R	GTA CTT TGC TGG ATT TTG TGG GCA CTT CAG GAG GCA GG	599	64.5
34	MYO3BX3-F MYO3BX3-R	GGC TGG AAC AGG TAG TAA TAA CAG TTT TTA GGA TGG CAT TTC TCC	302	58.5
35	MYO3BX2-F MYO3BX2-R	TCC AAT GTC AAT AAG AGT GGC ATG TCA GGA TGT CTG CCT TC	366	53.8
36.1	MYO3BX1e-F MYO3BX1e-R	TAG AGG ATA GGA TAA GAT GTT TTG G GCC CCA GTG ATG ATT GTA AC	333	53.8
36.2	MYO3BX1d-F MYO3BX1d-R	ATA GAA AAG CCC AGC AAA GG GAA CAA CTC AGC CCA CCC	498	63.1
36.3	MYO3BX1c-F MYO3BX1c-R	CAC TGT AGC CAC TGA TAG CCT C TGG ACC CTG ATA ATA TTG CTT G	448	53.8
36.4	MYO3BX1b-F MYO3BX1b-R	GAT TAC ATC ACA CTG CAA TTT GTC GAT GCT GAG AAA GCC TCT GTC	464	58.5
36.5	MYO3BX1a-F MYO3BX1a-R	GAG TCT TGC CCT GTC GC GTC ACT TTT CAT ACA CTA TAA CCT CTG	485	65.0

Table 4.2 Oligonucleotide primers used in amplifying segments of the GORASP2 gene in affected member 671 from the BBS5 family B9.

Exon	Primer Name	Sequence	Size (bp)	Annealing Temp (°C)
1A	GORASP2X1A-F GORASP2X1A-R	GAC AGG AAA CCT AGT GGA CG GTC ACC TTA ACT TTG AAC CAT TC	296	63.1
1B	GORASP2X1B-F GORASP2X1B-R	TTT CTG ACA TCT CAC CCT AGC ATA AAA ATA GCC CTT TGA CAA TG	579	63.1
2	GORASP2X2-F GORASP2X2-R	TTC AGT AGA GAT GGG GTT TTG ATT GTT ATG ATA CAC AGC TGC C	479	65.0
3	GORASP2X3-F GORASP2X3-R	GGA GAC AGT GGA GAA TGG AC AAA GGG TGA TGG ACA GTC AG	529	65.0
4	GORASP2X4-F GORASP2X4-R	TGC TGG TTT AGA ACT CCT GAG TAT CTG CAG GGC AAG TAA GAG	327	65.0
5	GORASP2X5-F GORASP2X5-R	TCT AGC CCT GTC CTT TTT CC CTT GAA TCT ATG CTC TTG GGT TAC	400	63.1
6	GORASP2X6-F GORASP2X6-R	ATG CTG TTA GCT CTT AGG TTA GTA AAG TGA TAG CCA GGT TTG A	397	58.5
7	GORASP2X7-F GORASP2X7-R	CAC TGA ACA AAA AGG GGA AG GAC ACA TCT CTA GCA AGA CAG AAG	493	63.1
8	GORASP2X8-F GORASP2X8-R	ACT GCT TCA TGA ACC TCC AG ATG TTG GTG AAA CCT TTG TTG	338	58.5
9	GORASP2X9-F GORASP2X9-R	ATA CCC ATC TCA GGT CAT CTG GAT TCG GCT TAC AAT TTT CAC	187	65.0
10	GORASP2X10-F GORASP2X10-R	AGG GGG GGT GGA GGA G CGA CGG CGG CCG GGG AGC	285	49.0

The critical region for BBS5 is 13 cM (14 Mb), a very large region to search for candidate genes. See Figure 4.2. Only five BBS5 families have been described worldwide, and all have been examined in an attempt to narrow the critical region. The critical region of 13 cM found by Young *et al.* (1999) has not been narrowed any farther using three new families as described by Beales *et al.* (2001).

Using bioinformatics, the gene for BBS1 was discovered by positional cloning and then comparing the sequences of candidate genes to the human BBS2 gene. BBS7 was also found using the sequence of the human BBS2 gene, comparing the human and zebrafish BBS2 sequences to the NCBI EST database. BBS8 was found in a similar manner by comparing pieces of the human BBS4 gene sequence with the human genome and EST databases. In this thesis, a bioinformatics approach was used to find BBS3 candidate genes, as described in the next section. No gene for BBS5 was identified via literature searches, and no gene for BBS3 was identified using bioinformatics. Though literature searches still have value, it appears that bioinformatics is an even more useful tool for uncovering BBS candidate genes. Narrowing the 14 Mb region of BBS5 would make bioinformatics easier to apply - but new BBS5 families would have to be found and haplotype analysis done. Any new BBS5 families would have to have fortuitous recombinations to narrow the critical region further. Due to the rareness of BBS5 and the greater rareness of finding recombinations in the critical region, it is not likely that the critical region will be narrowed anytime

soon. Innovative bioinformatic approaches must be developed to facilitate discovery of the unknown BBS genes, as discussed in Chapter 5.

Chapter 5: Candidate Gene Screening for Bardet-Biedl Syndrome 3

5.1 Rationale

The locus for BBS3 was originally identified by Sheffield *et al.*(1994) in an inbred Bedouin family. Linkage was detected by observing a shift in the number of alleles in the unaffected pooled sample compared to a single allele in the affected pooled sample. The critical region was reported to be 11 cM between D3S1254 and D3S1302 (Sheffield *et al.*, 1994). Young *et al.* (1998) narrowed the region even further to an interval of 6 cM using family NF-B2, as seen in Figure 5.1. The pedigree of NF-B2 can be seen in Figure 5.2. Haplotype analysis determined that the new critical region lay between D3S1595 and D3S1753 based on recombination of the disease chromosome in two of the affected individuals. NF-B2 was the second BBS3 family found worldwide after the discovery of the Bedouin BBS3 family by Sheffield *et al.* (1994).

The critical region was refined even further by Beales *et al.* (2001) using haplotype analysis on family AR-201. The new critical region was said to be between D3S1603 and D3S1251, defining a region of 1.1 Mb. However, it has now been shown that the critical region is actually between D3S1566 and D3S1271, defining a region of 9 Mb (Figure 5.3). An explanation of how this was determined is as follows.

Figure 5.1 Adapted from Young *et al.* (1998). The BBS3 critical region on chromosome 3 as determined by Young *et al.* (1998) using family NF-B2.

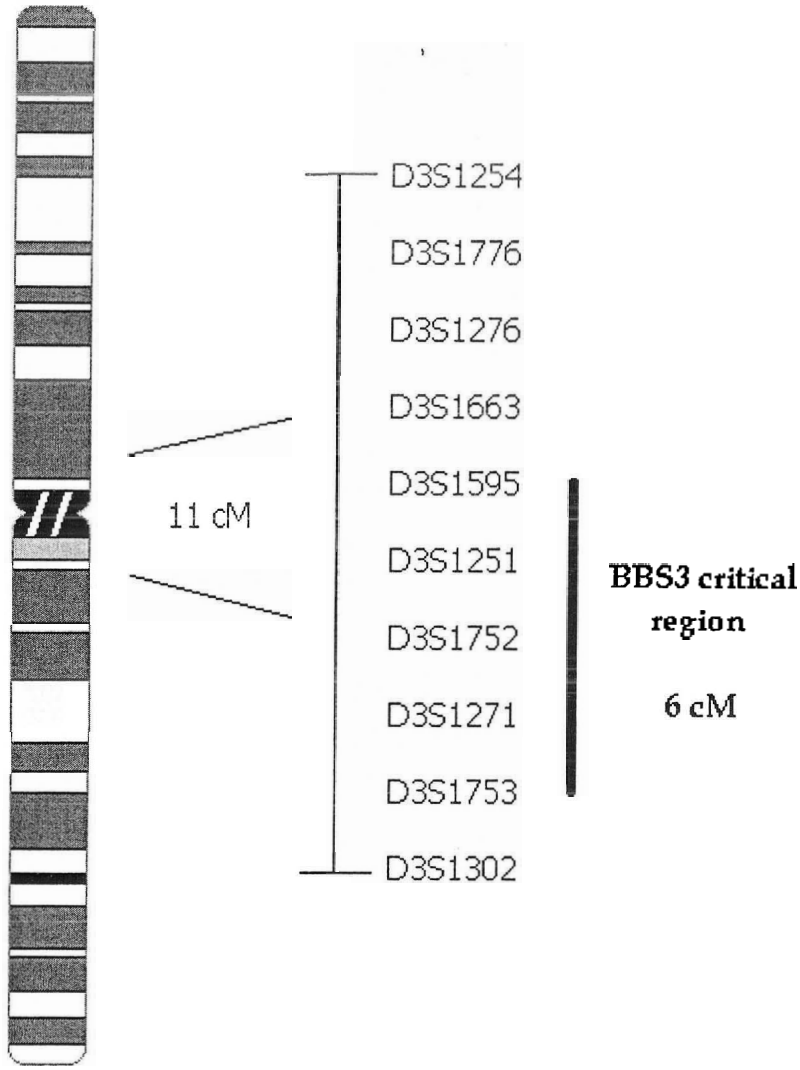

D3S1254
D3S1776
D3S1276
D3S1663
D3S1595
D3S1251
D3S1752
D3S1271
D3S1753
D3S1302
11 cM
BBS3 critical region
6 cM

Figure 5.2 Pedigree of family NF-B2. Individual PID's are indicated under each individual.

Key:

unaffected individual

affected individual

carrier

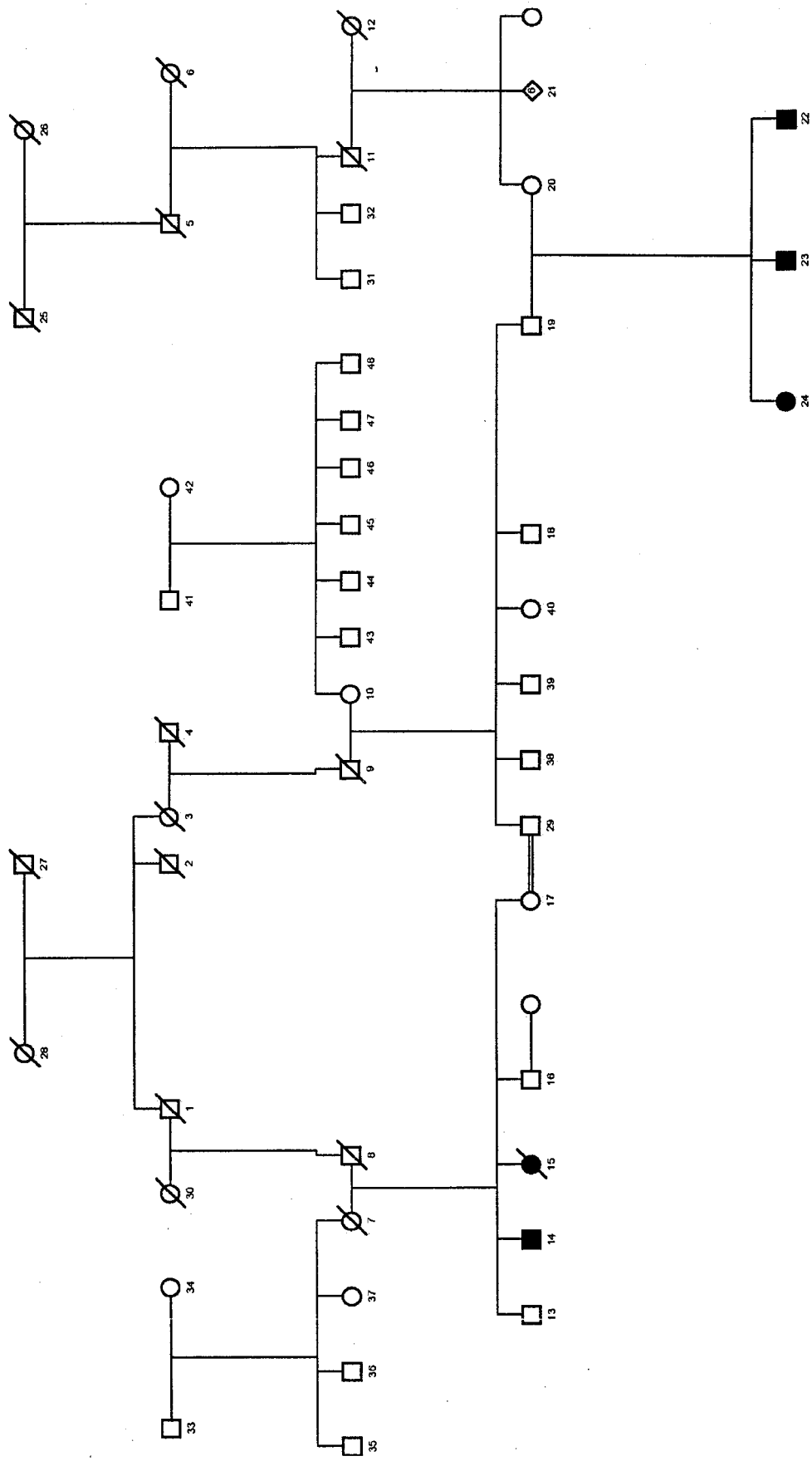
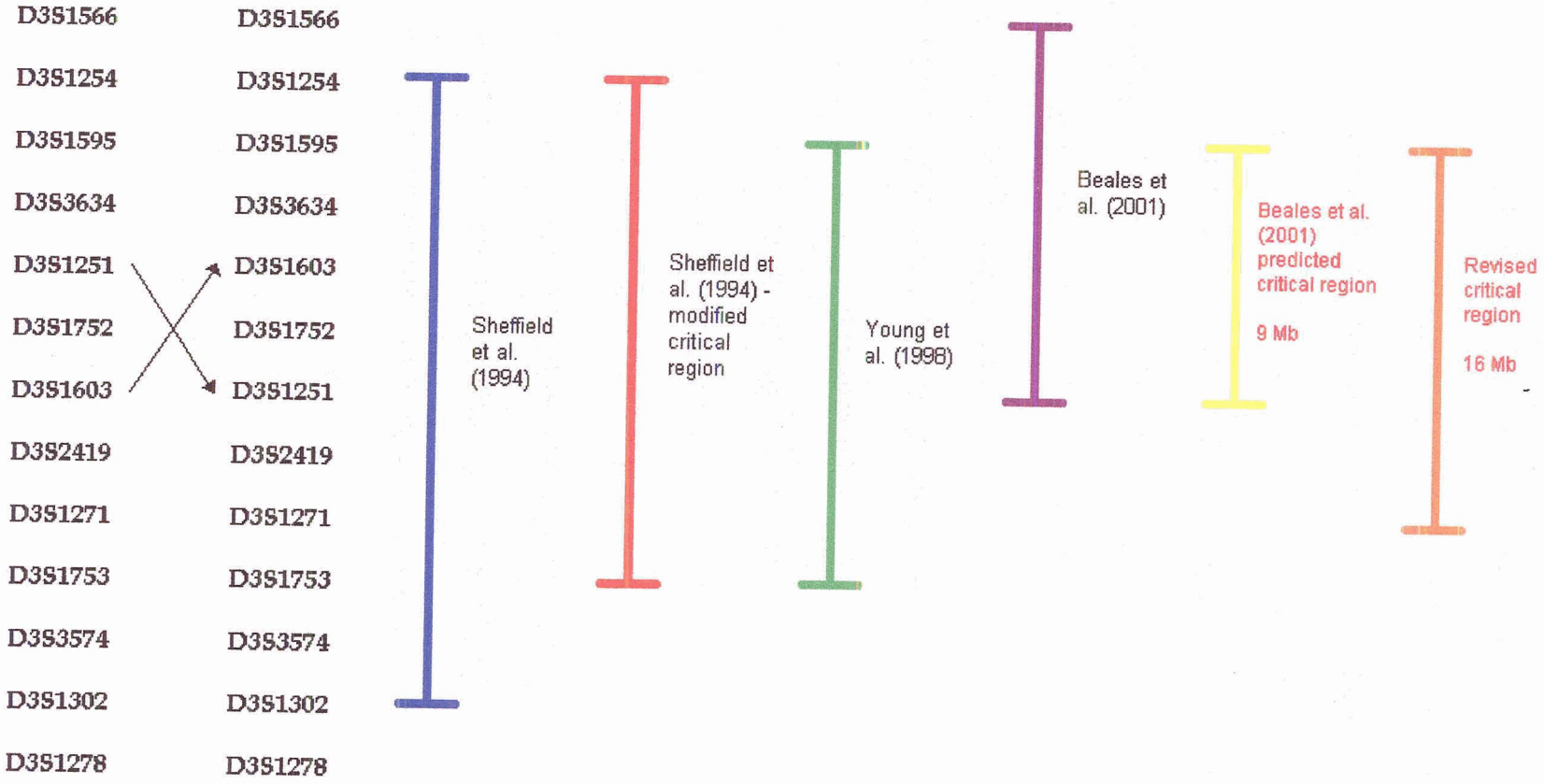


Figure 5.3 Illustration of determination of the revised 16 Mb BBS3 critical region.


D3S1566
D3S1254
D3S1595
D3S3634
D3S1251
D3S1752
D3S1603
D3S2419
D3S1271
D3S1753
D3S3574
D3S1302
D3S1278
D3S1566
D3S1254
D3S1595
D3S3634
D3S1603
D3S1752
D3S1251
D3S2419
D3S1271
D3S1753
D3S3574
D3S1302
D3S1278
Sheffield et al. (1994)
Sheffield et al. (1994) - modified critical region
Young et al. (1998)
Beales et al. (2001)
Beales et al. (2001) predicted critical region
9 Mb
Revised critical region
16 Mb

Sheffield *et al.* (1994) concluded that the critical region was between D3S1254 and D3S1302; however, when looking at the haplotypes of the Bedouin family used to find the critical region, it is apparent that there is a recombination in affected individual PID 27 between markers D3S1271 and D3S1753. Therefore, the distal marker is D3S1753, not D3S1302. Young *et al.* (1998) refined the proximal boundary of the critical region from D3S1254 to D3S1595 by observing a recombination in an affected individual from family NF-B2. Combining these two sets of observations would put the critical region between D3S1595 and D3S1753. Later, Beales *et al.* (2001) observed intra-familial recombinations in BBS3 family AR-201, narrowing the critical region to between D3S1566 and D3S1251. In combination with former critical regions, a 2 cM critical region was determined between markers D3S1603 and D3S1251. However, the order of the markers has been revised since the Beales *et al.* (2001) paper was published, as illustrated in Figure 5.3. Due to this change, the critical region has become larger. It is now between markers D3S1595 and D3S1271, which is a region of approximately 16 Mb. Changing the order of the markers has changed the pattern of intra-familial recombinations in family AR-201, thus widening the region.

5.2 Screening of Candidate Genes for BBS

It was predicted that BBS proteins may share similar domains, and this hypothesis enabled the identification of BBS7 and BBS8. After their identification, it was found that human BBS1, BBS2 and BBS7 appear to share similarity across a β-propeller region (Ansley *et al.*, 2003). It is also known that the BBS1, BBS2, BBS7, and BBS8 homologous genes in *C. elegans* share a common regulatory element called the RFX box. The RFX box, also known as the X box, is an imperfect 14 base pair repeat that is part of a promoter region (Figure 5.4). It is approximately 100 base pairs upstream from the start codon of genes involved in ciliation (Ansley *et al.*, 2003). It was found that when the RFX boxes for BBS2 and BBS7 were altered, there was considerably reduced expression of these genes. Loss of function experiments testing the importance of the RFX box in BBS homologs were carried out in ciliated neurons in *C. elegans.* GFP promoter fusions were constructed with genes exclusively expressed in ciliated neurons. The RFX box was modified in three ways: the region at the 5′ end of the RFX box was deleted; the RFX box and corresponding 5′ end were deleted; and the RFX box was replaced with a nonspecific base pair sequence. The first and third constructs, when injected into *C. elegans*, showed some expression. However, when the second construct was injected, there was no expression. Though the RFX box plays a critical role in regulation of expression, it is likely that there are other elements upstream of the RFX box

Figure 5.4 Adapted from Ansley *et al.* (2003). Illustration of the 14 bp RFX-box consensus sequence in *C. elegans.*

key: G - guanine
T - thymine
C - cytosine
A - adenosine

that contribute to such regulation (Oliver Blacque, personal communication). This information, fused with the speculation that BBS is caused by mutations to ciliary genes (Blacque *et al.,* 2003), gives validation to screening *C. elegans* BBS candidate genes by taking into account whether or not they have an RFX box.

Other *C. elegans* genes that have the RFX box are *che-3, daf-19, osm-1,* and *osm-6,* and all are involved in ciliation. It was noted that genes that are specific to subsets of ciliated cells do not have an upstream RFX box. The purpose of the RFX box is to permit the binding of transcription factors such as *daf-19,* which regulates expression of this set of genes (Swoboda *et al.*, 2000).

5.3 GPR15 – G-protein coupled receptor 15

Using a bioinformatic approach, a search for the BBS3 gene was conducted. An outline of the procedure can be seen in Figure 5.5. All known BBS genes were compared against the nonredundant database as part of a phylogenetic analysis. Results are as compiled in Table 5.1. Those organisms that have homologs to known BBS genes are listed. Note that *C. reinhardtii* and *C. elegans* have homologs of all known BBS genes excluding BBS6. It is also suspected that *C. elegans* does not have a convincing BBS4 homolog. *S. cerevisiae* has no known equivalents to the human BBS genes. Using this information, a search for BBS3 was conducted.

Figure 5.5 Flow chart summarizing the approach used to find candidate BBS3 genes.

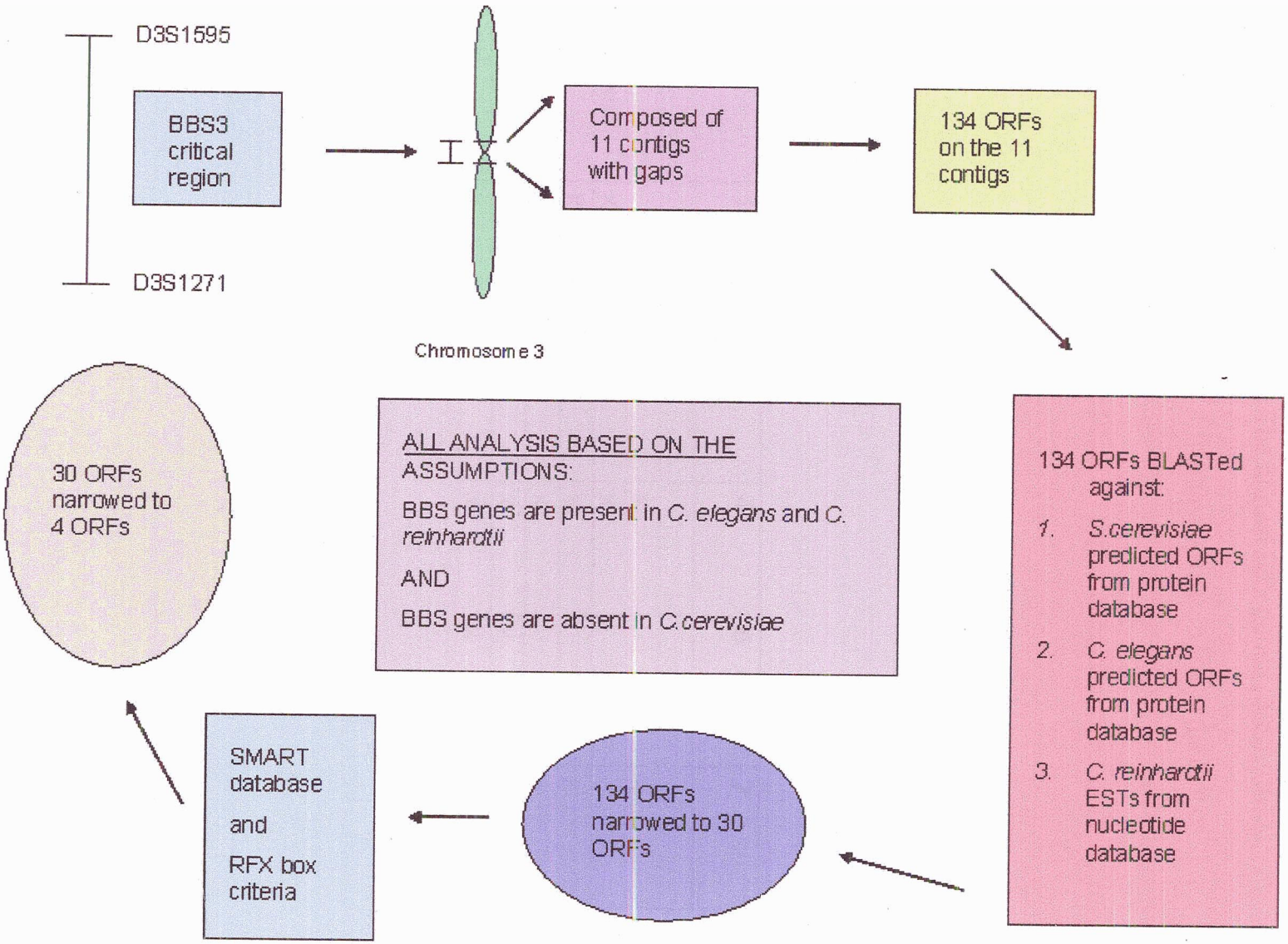

D3S1595
BBS3 critical region
D3S1271
Chromosome 3
Composed of 11 contigs with gaps
134 ORFs on the 11 contigs
134 ORFs BLASTed against:
1. S.cerevisiae predicted ORFs from protein database
2. C. elegans predicted ORFs from protein database
3. C. reinhardtii ESTs from nucleotide database
134 ORFs narrowed to 30 ORFs
SMART database
and
RFX box criteria
30 ORFs narrowed to 4 ORFs
ALL ANALYSIS BASED ON THE ASSUMPTIONS:
BBS genes are present in C. elegans and C. reinhardtii
AND
BBS genes are absent in C.cerevisiae

Table 5.1 Taxonomy of BBS proteins. The known BBS genes (BBS1, BBS2, BBS4, BBS6, and BBS7) are crossed off if homologs exist in the organism or organisms listed to the left.

Organism	**BBS Genes**				
	1	**2**	**4**	**6**	**7**
S. cerevisiae					
G. max			x		x
D. melanogaster	x		x		
C. reinhardtii	x	x	x		x
C. elegans	x	x	x		x
Z. rerio	x	x	x	x	x
H. sapiens	x	x	x	x	x

Eleven contigs that encompassed most of the critical region for human BBS3 on chromosome 3 were used to narrow in on candidate BBS3 genes. Several BLAST analyses were conducted using predicted open reading frames in BLAST searches against the non-redundant database as well as the genomes of *C. reinhardtii, C. elegans,* and *S. cerevisiae*. After the searches were conducted using the NCBI database, it was found that the comparison with the *C. reinhardtii* genome yielded 134 possible candidate open reading frames. Human open reading frames were considered to be candidate genes only if a homolog was present in *C. reinhardtii* and was not present in *S. cerevisiae*. This is action is justified from the finding that there are no known *S. cerevisiae* BBS homologs, whereas *C. reinhardtii* has homologs of all known BBS genes except BBS6. Therefore, any human homologs of *S. cerevisiae* open reading frames were excluded from the search. After the searches were conducted with the genomes of *C. reinhardtii* as well as that of *C. elegans*, it was found that there was only one open reading frame that was similar in *C. reinhardtii, C. elegans* and humans that did not exist in *S. cerevisiae*. Three more open reading frames, exclusively found in *C.* elegans, were also considered as their translated products had functions that were suggestive of BBS proteins; i.e., involved in cilia formation or function.

The revised critical region of the BBS3 locus was almost completely covered using eleven contigs. The region around the centromere has not been completely sequenced and therefore could not be included in the analysis. The

BLAST analysis was carried out twice, between January and February in 2003. After the first analysis, candidate open reading frames were again compared with sequences in the NCBI database.

The four open reading frames chosen to be analyzed were screened using the internet program SMART, which searches for known domains (Schultz *et al.*, 1998). The presence of the RFX box was considered, helping to prioritize the field of candidate genes. Only one of the open reading frames had such a structure in the promoter region, at -119 bp from the start codon. This gene was candidate gene GPR15.

As a result of the analysis, four open reading frames were chosen for sequencing. All were found through the comparison of the human BBS3 contigs with the *C. elegans* genome. All had informative domains as found by the SMART program. These four candidates were G-protein coupled receptor 15 (GPR15), G-protein coupled receptor 128 (GRP128), vitamin k dependent protein S, and an unknown protein called XTALbg domain protein, which may be axoneme associated. GPR15 was the gene with an RFX-like element in the promoter region in *C. elegans*, and thus it was chosen as the best candidate (Figure 5.6 and Table 5.2). The second choice was GPR128, the third the XTALbg domain protein, and lastly the axoneme associated protein.

A GFP-promoter fusion of F41E6.3 *C. elegans* homolog to GPR15 was constructed, and it appears that this gene is involved in ciliated neuron expression, though this has yet to be confirmed (Muneer Esmail, personal

communication). Using a bioinformatics approach to select GPR15, in combination with results from localization studies, illustrates that GPR15 was an excellent BBS3 candidate gene.

GPR15 was sequenced in one affected individual, PID 24, from the NF-B2 family. The exon and exon-intron boundaries were sequenced, but no variation was found. Therefore, GPR15 does not appear to be the cause of BBS3 in this family.

Figure 5.6 GPR15 gene structure. This open reading frame is located on chromosome 2q31, and consists of one exon (1083 bp). Forward and reverse primers are shown with rightward and leftward arrows and names displayed above. Arrows show start and stop codons.

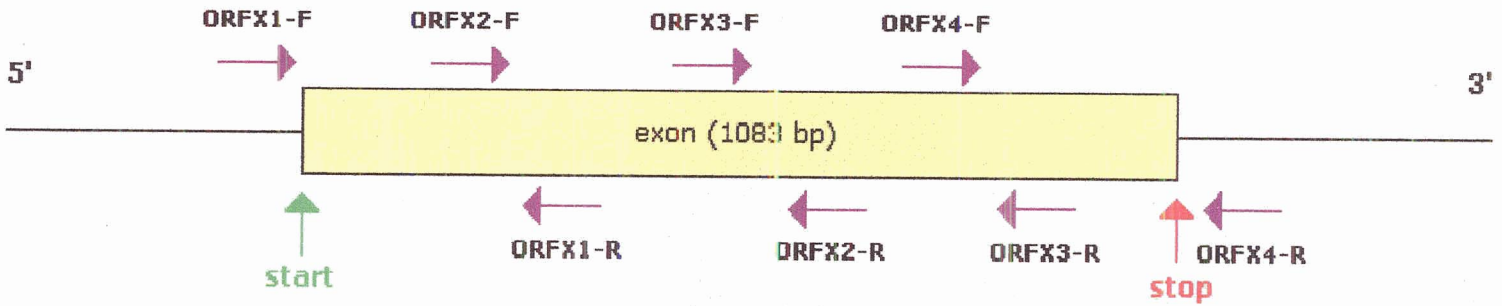

ORFX1-F
ORFX2-F
ORFX3-F
ORFX4-F
5'
3'
exon (1083 bp)
start
ORFX1-R
ORFX2-R
ORFX3-R
stop
ORFX4-R

Table 5.2 Table of Oligonucleotide primers used in amplifying GPR15 in affected member 125 from the BBS3 family B2.

Exon portion	Primer Name	Sequence	Size (bp)	Annealing Temp (°C)
1	BBS3ORFX1-F BBS3ORFX1-R	ATG AAG CAA TGT GAA TCC TAT C GAA AAA CTA TGA CTT TGG GAG TAG	176	65.0
2	BBS3ORFX2-F BBS3ORFX2-R	ATG AAT GAA GGT GGA GAG AGC ATT GAT GAT AAG CCA TAC TGT GC	467	58.5
3	BBS3ORFX3-F BBS3ORFX3-R	ATA TGA GTT TAA TTG GAG TTG CC CTT GTC ACA TTG CCT CTC TG	295	65.0
4	BBS3ORFX4-F BBS3ORFX4-R	CTT CTG AAT TTC CTG GAT ACG CCC TTT CTG AGG ATT GGT AG	591	58.5

Chapter 6: Discussion

BBS is a relatively rare disorder. Eight genetic loci have been identified, each on a different chromosome, based on an autosomal recessive disease model. Using this model, researchers used genome scanning, homozygosity mapping, linkage disequilibrium analysis, and loss of identity by descent to investigate the molecular basis of BBS.

A genome wide scan involves looking for an unknown locus by using markers such as microsatellites to scan the human genome to detect haplotypes that show association with the disease. Disease associated haplotypes would be shared among affected individuals. However, if a haplotype is shared between affected and unaffected individuals or if affected siblings have different haplotypes in a particular region, it would be excluded. If the family is known to be consanguineous, homozygosity mapping can be used as a form of genome wide scanning. This technique involves pooling DNA from affected individuals and unaffected individuals and then performing a genome wide scan on these two samples. If there was a homozygous region shared between affected individuals, there would be a visible allele shift between the two samples – the affected sample would show one allele and the unaffected sample would show more than one allele. This technique was used to find the BBS5 critical region as well as to identify a BBS locus on chromosome 20 around the MKKS gene, making it a candidate gene for BBS6. If there is a suspicion that a disease in a population is caused by a single disease causing mutation descending from a

common ancestor, then a linkage disequilibrium (LD) study could be carried out on affected individuals from apparently unrelated families. Affected individuals are expected to share the mutation and be homozygous at the surrounding markers. Loss of identity by descent (IBD) would determine the ends of the critical region, where the ancestral haplotype is interrupted by recombination events that may not be observed within the families. In the case of BBS6, linkage disequilibrium studies would not have been useful since there are three BBS6 mutations present in the eight Newfoundland BBS6 families – there are simply too many mutations for all of them to be descended from a common ancestor and so no relationship would have been found.

When BBS was first studied, recombination frequencies allowed linkage maps to be constructed and thus measurements were made in cM. The definition of a cM is when recombination occurs 1 per cent of the time between two loci. At present, due to the human genome being sequenced, measurements can be directly taken in Mb. The exact positions of markers relative to one another are now known. It was generally accepted that 1 cM was equivalent to 1 Mb, but this does not always hold true as recombination rates vary across the human genome.

A more complex mode of inheritance termed triallelism was proposed by Katsanis *et al.* (2001) after a mutational screen of many BBS families. Triallelism was also described as "autosomal recessive inheritance with a modifier of penetrance" (Burghes *et al.*, 2001), and this is more in line with recent

observations by Badano *et al.* (2003) who described three families in which two mutations at one locus were sufficient to cause the disease but siblings with an additional mutation at another locus were more severely affected. There is still debate about the frequency and importance of triallelism in BBS (Mykytyn *et al.*, 2003).

BBS is more common in Newfoundland than elsewhere in North America or Europe and it is known that the population carries at least six BBS loci and a minimum of nine different mutations. As members of the BBS Newfoundland families have been followed clinically for more than twenty years, this is an excellent cohort to investigate the frequency of triallelism.

All Newfoundland BBS families were screened for the known Newfoundland BBS mutations in BBS1, BBS2, and BBS6. BBS4 has not been observed in the Newfoundland population, and was thus excluded from this particular study. When this study was conducted neither BBS7 nor BBS8 had been discovered so they were not included. The M390R BBS1 and Y24X BBS2 mutations were screened in all Newfoundland affected individuals by Yanli Fan, and the four BBS6 mutations: fs1, fs2, A242S, and L277P were screened as part of this thesis. The mutations in six BBS6 families NF-B1, NF-B3, NF-B4, NF-B5, NF-B13, and NF-B16 had been detailed previously (Katsanis *et al.*, 2001). All other Newfoundland families with DNA available from an affected individual were screened for the four BBS6 mutations. Besides the NF-B14 family (Katsanis *et al.*, 2001), no other families were found to exhibit triallelism. Except for this

known case, the results obtained during this thesis did not reveal any additional instances where a BBS6 mutation was associated with an affected individual who was linked to BBS1, BBS2, BBS3, or BBS5. This suggests that triallelism is a rare event. A review of the literature revealed that only eleven out of two hundred and fifty-nine families (including NF-B14) exhibit true examples of triallelism. This translates to 4.2 per cent, which is similar to the Newfoundland average of 5.3 per cent (1 out of 19 families).

Triallelism has yet to be fully accepted as a *bona fide* complex mode of inheritance. Researchers have doubted its existence, looking for other ways to explain why individuals in certain families only manifest BBS if they have a third mutation. There is the possibility that the third mutation in 4.2 per cent of the population is present by chance and plays no part in manifesting the disease. It is also possible that in some cases the third mutation will only have an effect if the translated protein directly interacts with the other mutated protein translated from the gene with two mutated alleles. If BBS proteins do not work together in the same cascade, then triallelism would be much harder to explain.

Triallelism is already being challenged less than two years after it was first recognized by Katsanis *et al.* (2001). Badano *et al.* (2003) have shown the existence of three families which have affected siblings with differing genotypes (Table 3.5). It appears that the third mutation is a modifier of the other two mutations, and it is only with this third mutation in certain individuals that the disease will be allowed to manifest, or will manifest the disease to a greater

extent in one affected individual compared to the other. This exhibits what can be called "autosomal recessive inheritance with a modifier of expressitivity".

The study looking at triallelism in the Newfoundland population permitted the characterization of new families. Two new BBS6 families were found: NF-B20 and NF-B25. The affected individual in NF-B20 was homozygous for mutation 429ΔCT433ΔAG, otherwise named fs1. In NF-B25 the affected individual was homozygous for mutation 280ΔT, named fs2. The six BBS6 families that were formerly identified: NF-B1, NF-B3, NF-B4, NF-B5, NF-B13, and NF-B16 - were all from the east coast of the island of Newfoundland (Figure 3.8). NF-B25 is also from the east coast; however, NF-B20 is the first BBS6 family to be seen on the west coast. Pedigree analysis and a more extensive search of the ancestry of the BBS6 families may reveal single founder events for the different mutations and allow extended family trees to be constructed. This would help identify potential carriers, who could then be recruited for genotype-phenotype studies through genetic counseling.

Eight BBS loci have been identified: BBS1 (11q13), BBS2 (16q21), BBS3 (3p12-13), BBS4 (15q123), BBS5 (2q31), BBS6 (20p12), BBS7 (4q27), and BBS8 (14q32.1), and genes for all of these loci have been identified except for BBS3 and BBS5. Due to the large number of genes implicated in the disease, extensive studies searching for genotype-phenotype correlations have been carried out, but none have been found. BBS6 is likely a chaperonin, but the actual functions of the others remain something of a mystery. BBS proteins

BBS1, BBS2, and BBS8 are found in ciliated neuronal cells in *C.* elegans. As well, BBS4, BBS6, and BBS8 have all been shown to localize at the centrosome/basal body in mice (M. Leroux, personal communication).

The two genes that have not yet been found - BBS3 and BBS5 – have had critical regions reduced by using haplotype analysis on Newfoundland families NF-B2 and NF-B9. In this study candidate genes for BBS3 and BBS5 were identified using two different methods: bioinformatics and a literature search.

Myosin 3B was suggested as a candidate gene for BBS5 by Dose *et al.* (2002) for three reasons: it is highly expressed in retinal photoreceptors, it is expressed in tissues that are adversely affected in BBS, and the gene maps to the critical region of 2q31.1-q31.2. GORASP2 was also suggested as a candidate gene based on two factors: it mapped to the BBS5 critical region as well as having a function that when disrupted, could potentially cause a BBS-like phenotype. No mutations were found in the NF-B9 family for either candidate gene.

A more promising method to find BBS genes seems to lie within the field of bioinformatics. BBS1, BBS7, and BBS8 were all found by using databases readily available on the internet. BBS1 and BBS7 were found by using the BBS2 gene sequence, and the BBS4 sequence was used to identify candidate BBS8 genes. The premise for using BBS2 and BBS4 gene sequences was that other unknown BBS genes will share portions of their sequences with known genes,

corresponding to domains that the BBS proteins have in common. This assumes that some BBS proteins have similar functions, and this assumption substantiates the triallelic hypothesis – whereas more components of a complementary pathway are interrupted, the more likely disease will be manifested. This can be seen in the poison model as shown in Figure 1.6.

This idea of comparing BBS gene sequences was taken and modified to search for BBS3 candidate genes. Instead of using human BBS sequences, a phylogenetic analysis was performed using known human BBS gene sequences and screening each against the non-redundant, *C. reinhardtii*, *C. elegans,* and *S. cerevisiae* databases. Under the premise that BBS genes are seen in *C. reinhardtii* and *C. elegans* but not *S. cerevisiae*, four candidate BBS3 genes were proposed. GPR15 was sequenced as part of this thesis, and three more are in the process of being sequenced. No mutations have been found so far, it has yet to be seen if this phylogenetic bioinformatic will prove informative for finding BBS genes.

Besides looking for mutations in translated portions of a gene, it is also important to realize that mutations may be found in introns (e.g., cryptic splice sites,) enhancer sequences, and promoter regions. Though splice sites and 5′ regions are usually recognizable, regulatory elements may be much more difficult to identify. As an example, Lettice *et al.* (2003) have described the *Shh* enhancer, which is found in intron 5 of the *Lmbr1* gene which is 1 Mb from the *Shh* gene, much farther away than many researchers would have anticipated.

Comparing genomes is helping to identify potential regulatory regions and this will be applied to BBS genes in the future.

Critical regions must be examined closely and reviewed as sequence data becomes available. A change in marker order can drastically affect a critical region, as shown with the BBS3 critical region, which expanded from 6 cM to 9 cM based on a complete review of current literature. The order of markers on linkage maps is being superceded by actual sequence information from the genome.

BBS6 was found by recognizing that the BBS6 critical region mapped to 20p12 – the same region of the genome where the MKKS gene resides. Sequencing the MKKS gene in individuals diagnosed with BBS lead to the discovery that MKKS and BBS are strongly linked due to both having mutations in the same gene. Referring to Table 1.1 in the beginning of the thesis, it can be seen that Alstrom and Biemond syndromes are more similar to BBS and LMS than is MKKS. It is possible that unknown BBS families which have not been linked to BBS1-8 may harbor mutations in genes causing similar syndromes. It would be wise to screen new BBS families for BBS1-8, and if they did not map to any of the known loci, examine if any of them map to the *ALMS1* gene in region 2p13, which causes Alstrom syndrome (Collin *et al.*, 2002; Hearn *et al.*, 2002). The gene for Biemond syndrome has yet to be found.

Currently, it is strongly believed that BBS is caused by mutations in genes involved in ciliation. Mutations made to the BBS8 gene in *C. elegans* have

created a phenotype were the cilia are not functioning normally (O. Blaque, personal communcation). Observing that BBS proteins appear to localize at the base of cilia helps to support this claim.

Usher syndrome is similar to BBS in that affected individuals have retinitis pigmentosa. The product of the Usher 1B gene is myosin VIIa, and this product is found in cilia of photoreceptor cells (Liu *et al.*, 1997), possibly affecting cilia formation and/or function. A winged helix transcription factor named HFH-4 in mice regulates dynein expression. Chen *et al.* (1998) found that when HFH-4 is mutated this leads to the absence of cilia. It was hypothesized that the absence of cilia leads to the absence of movement of cells within the embryo, possibly leading to Kartagener's syndrome, which exhibits *situs inversus*. Extending this observation, it is possible that mutations deterring the proper movement of cells during gastrulation would lead to dystrophy of the extremities, as seen in BBS. This theory is supported by Pennarum *et al.* (2002) who found the human gene *hPF20* is responsible for the disease Primary ciliary dyskinesia (PCD), which also exhibits *situs inversus*, which has also been found in some BBS patients (Ansley *et al.*, 2003). The PCD gene has a worm orthologue named *pf20*, which codes for a protein containing a WD domain, as is commonly seen in genes involved in intraflagellar transport (IFT) (Cole *et al.*, 2003). IFT proteins with WD domains are seen just as commonly as those with tetratricopeptide repeats (TPRs), which were described in BBS4. The BBS4 TPR was matched to that of TTC8, which was later found to be BBS8. Knowing that TPRs were common to IFT genes,

researchers examined the effect of mutations in the *C. elegans* BBS8 gene on ciliary function.

Screening for BBS genes in the future should involve comparing domains common to other *C. elegans* ciliary genes (e.g., WD domains) to the human genome.

Appendix A: GenomeScan output file for contigs NT_006931, NT_022419, NT_022642, NT_022435, NT_022497, NT_037573, NT_022475, NT_033050, NT_033041, NT_037574, NT_005863.

MapViewer

Data As Table View Download All !

GenomeScan On Sequence next
Total Models On Chromosome: 3052 [12 not localized]
Region Displayed: 85M-93M bp
Models in Region: 49

Download/View Sequence/Evidence All Sequence Maps Download Data

start	stop	name	orient	protein gi
85771306	85778779	Hs3_6188_31_2_2	+	19923268
85938638	86012976	Hs3_6188_31_3_1	+	-
86242086	86511026	Hs3_22575_31_1_3	-	-
89784809	89867903	Hs3_22798_31_3_1	-	12840558
89881557	89883097	Hs3_22798_31_4_1	-	548638
89919395	89919868	Hs3_22798_31_5_1	-	-
89956522	90004142	Hs3_22798_31_5_2	+	-
90020448	90031617	Hs3_22798_31_5_3	-	-
90057888	90084189	Hs3_22798_31_6_1	+	7661634
90090261	90106829	Hs3_22798_31_7_1	-	488129
90153635	90154345	Hs3_22798_31_8_1	+	105815
90154471	90155090	Hs3_22798_31_8_2	+	307083
90164613	90319329	Hs3_22798_31_8_3	-	-
90371996	90425486	Hs3_22798_31_8_4	+	13129148
90460720	90461142	Hs3_22798_31_8_5	-	16741033
90461708	90461986	Hs3_22798_31_8_6	-	16741033
90623251	90659882	Hs3_22798_31_9_1	-	-
90720397	90764698	Hs3_22798_31_9_2	-	-
90821071	90822114	Hs3_22798_31_10_1	+	10835197
90885733	90886240	Hs3_22798_31_12_1	-	21361099
90889584	90969572	Hs3_22798_31_13_1	+	6912632
90969692	90971389	Hs3_22798_31_14_1	+	23955470
~~90980323~~	~~90986899~~	~~Hs3_22798_31_15_1~~	~~+~~	~~18848553~~
~~91018707~~	~~91019021~~	~~Hs3_22798_31_15_2~~	~~+~~	~~-~~
~~91058841~~	~~91087686~~	~~Hs3_22798_31_15_3~~	~~-~~	~~-~~
91147420	91230056	Hs3_22798_31_15_4	+	10179834
91295622	91339773	Hs3_22798_31_15_5	+	-
91429351	91495111	Hs3_22591_31_1_2	-	-
91597544	91614359	Hs3_22591_31_1_3	+	-
91666182	91698452	Hs3_22591_31_1_4	-	-
91746426	91747160	Hs3_22591_31_1_5	+	-
91845110	91846033	Hs3_22591_31_1_7	-	7669492
92010199	92148747	Hs3_22653_31_2_1	+	21361241
92162806	92216028	Hs3_22653_31_3_1	-	-
92256207	92293456	Hs3_22653_31_3_2	-	5835658
92532080	92566316	Hs3_37577_31_1_2	-	-
92744573	92776910	Hs3_37577_31_1_4	+	-
92806181	92895090	Hs3_22631_31_1_1	-	131086
92917343	92925365	Hs3_22631_31_3_3	+	12855722
92935838	92936722	Hs3_22631_31_3_4	-	21961211
92948306	92974716	Hs3_22631_31_3_5	+	21450707
92982401	92982964	Hs3_22631_31_4_1	-	13097774
92990579	93003939	Hs3_22631_31_5_2	-	15929379
93015113	93047943	Hs3_22631_31_5_3	+	11545785
93066138	93168899	Hs3_22631_31_6_1	+	-
93188088	93215177	Hs3_22631_31_6_2	+	-
93292903	93378134	Hs3_22631_31_6_3	+	-
93388957	93409908	Hs3_22631_31_6_4	+	-
93454469	93543530	Hs3_22631_31_8_1	-	-

Genes On Sequence prev next
Total Genes On Chromosome: 1371 [7 not localized]
Region Displayed: 85M-93M bp
Genes in Region: 22
Download/View Sequence/Evidence All Sequence Maps Download Data

start	stop	symbol	orient.	links	evidence	cyto.	full name
85725465	85832789	LOC253559	+	sv ev - seq mm	C	3p11.1	nectin-like protein 3
89784809	89867903	LOC286672	-	sv ev - seq mm	PE	3q11.1	LOC286672
89879571	89883153	LOC285232	-	sv ev - seq mm	I	3q11.1	similar to Amidophosphoribosyltrans
90057628	90085865	DKFZP564O123	+	sv ev hm seq mm	C	3q11.1	DKFZP564O123 protein
90090000	90106954	POU1F1	-	sv ev hm seq mm	C	3p11	POU domain, class 1, transcription fa
90153635	90154345	LOC285233	+	sv ev - seq mm	P	3q11.1	similar to keratin 8, type II cytoskele
90154471	90155090	LOC285234	+	sv ev - seq mm	P	3q11.1	similar to keratin 8, type II cytoskele
90371996	90425486	LOC285235	+	sv ev - seq mm	P	3q11.1	similar to hypothetical protein MGC4
90460720	90461142	LOC285236	-	sv ev - seq mm	P	3q11.1	similar to proteasomal ATPase (S4) [
90461708	90461986	LOC200899	-	sv ev - seq mm	P	3q11.1	similar to proteasomal ATPase (S4) [
90821014	90822114	HTR1F	+	sv ev hm seq mm	C	3p12	5-hydroxytryptamine (serotonin) rec
90882388	90980149	CGGBP1	-	sv ev - seq mm	C	3p12-p11.1	CGG triplet repeat binding protein 1
90969376	90972567	FLJ10997	+	sv ev - seq mm	C	3q11.1	hypothetical protein FLJ10997
90980007	90988227	LOC285237	+	sv ev - seq mm	I	3q11.1	similar to CG13876-PA [Drosophila
91845110	91846033	LOC285208	-	sv ev - seq mm	P	3q11.1	similar to Glyceraldehyde 3-phospha
92010159	92151377	EPHA3	+	sv ev hm seq mm	C	3p11.2	EphA3
92256207	92293456	LOC253953	-	sv ev - seq mm	P	3q11.1	similar to ATP synthase F0 subunit 6
92532080	92566316	LOC286675	-	sv ev - seq mm	PE	3q11.2	LOC286675
92806177	92895319	PROS1	-	sv ev hm seq mm	C	3p11-q11.2	protein S (alpha)
92901608	92976730	DKFZp761H079	+	sv ev - seq mm	C	3q11.2	hypothetical protein DKFZp761H07
92979558	92984352	LOC200895	-	sv ev - seq mm	I	3q11.2	similar to Dihydrofolate reductase
92984366	93049998	FLJ22609	+	sv ev - seq mm	C	3q11.2	hypothetical protein FLJ22609

Transcripts On Sequence prev next
Total Transcripts On Chromosome: 1484 [7 not localized]
Region Displayed: 85M-93M bp
Transcripts in Region: 22
Download/View Sequence/Evidence All Sequence Maps Download Data

start	stop	accession	locus	orient	links	evidence
85725465	85832113	NM_153184.1	LOC253559	+	sv ev	?
89784809	89867903	XM_210531	LOC286672	-	sv ev	PE
89879571	89883153	XM_208302	LOC285232	-	sv ev	I
90057659	90085281	NM_014043.1	DKFZP564O123	+	sv ev	?
90090000	90106954	NM_000306.1	POU1F1	-	sv ev	?
90153635	90154345	XM_210532	LOC285233	+	sv ev	P
90154471	90155090	XM_210533	LOC285234	+	sv ev	P
90371996	90425486	XM_210534	LOC285235	+	sv ev	P
90460720	90461142	XM_210535	LOC285236	-	sv ev	P
90461708	90461986	XM_115995	LOC200899	-	sv ev	P
90821014	90822114	NM_000866.1	HTR1F	+	sv ev	C
90882388	90889205	NM_003663.2	CGGBP1	-	sv ev	?
90969589	90972336	NM_018293.1	FLJ10997	+	sv ev	?
90980007	90988227	XM_209524	LOC285237	+	sv ev	I
91845110	91846033	XM_210526	LOC285208	-	sv ev	P
92010159	92151377	NM_005233.2	EPHA3	+	sv ev	?
92256207	92293456	XM_172255	LOC253953	-	sv ev	P
92532080	92566316	XM_212509	LOC286675	-	sv ev	PE
92806177	92895314	NM_000313.1	PROS1	-	sv ev	?
92901650	92976730	NM_144996.1	DKFZp761H079	+	sv ev	?

MapViewer

92979558	92984352	XM_114312	LOC200895	-	sv ev	?
92984472	93048200	NM_022072.1	FLJ22609	+	sv ev	?

Contig prev next

Total Contigs On Chromosome: **107** [3 not localized]

Region Displayed: **85M-93M bp**

Contigs in Region: 7

Download/View Sequence/Evidence All Sequence Maps Download Data

start	stop	accession	orient
85706398	86101556	NT_006031.8	+
86111557	86596811	NT_022419.8	+
89696812	91355572	NT_022642.5	+
91365573	91892176	NT_022435.6	+
91902177	92309191	NT_022497.11	+
92319192	92793455	NT_037573.1	+
92803456	98065601	NT_022475.10	+

STS prev

Total STSs On Chromosome: **6849** [48 not localized]

Region Displayed: **85M-93M bp**

STSs in Region: 118

Download/View Sequence/Evidence All Sequence Maps Download Data

start	stop	marker	Kbp	STS	thon	Marsh	G-G3	G-GB4	NCBI	S-G3	TNG	W-RH	W-YAC	deCODE	polymorphism	in_gene
85763442	85763734	D3S2315	85763	O		O								O	Y	-
85832552	85832709	SHGC-77127	85833	O				O		O	O				Y	-
85835970	85836149	SHGC-77126	85836	O				O		O	O				Y	-
85836432	85836585	RH12384	85836	O				O		O					Y	-
85837368	85837505	G32302;A005W32	85837	O				O		O					Y	-
85837368	85837505	G32302;A005W32	85837	O											-	-
~~85837443~~	~~85837569~~	~~D3S3302~~	~~85837~~	O												
~~85838131~~	~~85838327~~	~~A007K07~~	~~85838~~	O				O		O					~~Y~~	
~~85885794~~	~~85885934~~	~~D3S1595~~	~~85886~~	O			O	O	O	O	O				~~Y~~	
85930603	85930878	SHGC-105131	85931	O							O				-	-
85971851	85972153	SHGC-146678	85972	O							O				-	-
86184445	86184718	RH123037	86184	O							O				-	-
86191366	86191582	SHGC-85103	86191	O							O				-	-
86194317	86194651	SHGC-147325	86194	O							O				-	-
86225328	86225491	D3S3679	86225	O	O	O	O		O	O	O			O	Y	-
86322968	86323304	SHGC-147609	86323	O							O				-	-
86323166	86323456	SHGC-148213	86323	O							O				-	-
86355396	86355525	SHGC-37062	86355	O			O				O				-	-
86381096	86381387	SHGC-85727	86381	O							O				-	-
89743271	89743584	SHGC-89213	89743	O							O				-	-
89748667	89748884	D3S3671	89749	O	O	O	O		O		O		O	O	Y	-
89768372	89768513	D15S1269	89768	O											-	-
89768467	89768567	A004P25	89768	O				O		O					Y	-
89771067	89771250	G20437;A005R17	89771	O											-	-
89771067	89771250	G20437;A005R17	89771	O				O		O					Y	-
89771658	89771747	SHGC-68425	89772	O							O				-	-
89771702	89771807	G20685;A006B39	89772	O											-	-
89771702	89771807	G20685;A006B39	89772	O				O		O					Y	-
89771857	89771977	RH69879	89772	O				O		O					Y	-
89772786	89773063	SHGC-77133	89773	O				O		O	O				Y	-
89773928	89774051	RH45063	89774	O				O		O					Y	-
89777275	89777403	RH102693	89777	O				O							-	-
89818818	89818983	SHGC-77128	89819	O				O		O	O				Y	-
89879599	89879804	WI-20631	89880	O				O				O			Y	-
89881384	89881600	RH70521	89881	O											-	-
89968685	89968998	RH119233	89969	O							O				-	-
90058584	90058865	SHGC-155470	90059	O							O				-	-

start	stop	marker	Kbp	STS	Chen	Marsh	G-G3	G-GB4	NCBI	S-G3	TNG	W-RH	W-YAC	deCODE	polymorphism	in_gene
90084978	90085221	SHGC-77130	90085	O	·	·	·	O	·	O	O	·	·	·	Y	-
90085681	90085801	RH66762	90086	O	·	·	·	O	·	O	·	·	·	·	Y	-
90089986	90090527	G67534	90090	O	·	·	·	·	·	·	·	·	·	·	-	-
90090132	90090238	D3S3933	90090	O	·	·	·	·	·	·	·	·	O	·	Y	-
90091370	90091870	G67533	90091	O	·	·	·	·	·	·	·	·	·	·	-	-
90092344	90092672	G67532	90092	O	·	·	·	·	·	·	·	·	·	·	-	-
90094536	90094931	G67531	90095	O	·	·	·	·	·	·	·	·	·	·	-	-
90103661	90103971	G67530	90104	O	·	·	·	·	·	·	·	·	·	·	-	-
90106580	90107027	G67529	90107	O	·	·	·	·	·	·	·	·	·	·	-	-
90106717	90106878	RH17368	90107	O	·	·	·	O	·	O	·	·	·	·	Y	-
90130433	90130734	G59369	90130	O	·	·	·	·	·	·	O	·	·	·	-	-
90292173	90292358	SHGC-13656	90292	O	·	·	·	·	O	O	·	·	·	·	Y	-
90308287	90308582	SHGC-142691	90308	O	·	·	·	·	·	·	O	·	·	·	-	-
90593149	90593563	G54723	90593	O	·	·	·	·	·	·	·	·	·	·	-	-
90593372	90593561	G59441	90593	O	·	·	·	·	·	·	O	·	·	·	-	-
90622137	90622255	D3S2318	90622	O	·	O	·	·	·	·	·	·	·	O	Y	-
90699569	90699859	D3S2386	90700	O	·	O	·	·	·	·	·	·	·	·	Y	-
90703265	90703453	D3S1338	90703	O	·	·	·	·	·	·	·	·	·	·	-	-
90779948	90780286	G63651	90780	O	·	·	·	·	·	·	O	·	·	·	-	-
90790122	90790242	RH68753	90790	O	·	·	·	O	·	O	·	·	·	·	Y	-
90790206	90790338	G32985;A009Y24	90790	O	·	·	·	O	·	O	·	·	·	·	Y	-
90790206	90790338	G32985;A009Y24	90790	O	·	·	·	·	·	·	·	·	·	·	-	-
90792461	90792745	SHGC-84017	90792	O	·	·	·	·	·	·	O	·	·	·	-	-
90822010	90822241	RH68987	90822	O	·	·	·	O	·	O	·	·	·	·	Y	-
90822084	90822225	RH46950	90822	O	·	·	·	O	·	O	·	·	·	·	Y	-
90843716	90843839	SHGC-77135	90844	O	·	·	·	O	·	O	O	·	·	·	Y	-
90887751	90887916	RH102190	90888	O	·	·	·	O	·	·	·	·	·	·	-	-
~~90938295~~	~~90938419~~	~~SHGC-35057~~	~~90938~~	O	·	·	O	·	O	O	·	·	·	·	~~Y~~	
~~90951167~~	~~90951316~~	~~SHGC-31360~~	~~90951~~	O	·	·	O	·	·	·	·	·	·	·		
~~[illegible]~~	~~[illegible]~~	~~SHGC-[illegible]~~	~~90959~~	O	·	·	·	O	·	·	O	·	·	·		
90964802	90965090	SHGC-86193	90965	O	·	·	·	·	·	·	·	·	·	·	-	-
90972288	90972422	RH44266	90972	O	·	·	·	O	·	O	·	·	·	·	Y	-
90977983	90978327	SHGC-144965	90978	O	·	·	·	·	·	·	O	·	·	·	-	-
90987123	90987270	RH92935	90987	O	·	·	·	O	·	·	·	·	·	·	-	-
91097596	91097900	SHGC-144708	91098	O	·	·	·	·	·	·	O	·	·	·	-	-
91144318	91144634	SHGC-111666	91144	O	·	·	·	·	·	·	O	·	·	·	-	-
91292636	91292927	G63351	91293	O	·	·	·	·	·	·	O	·	·	·	-	-
91431697	91431836	D3S1552	91432	O	·	·	O	O	O	O	O	·	·	·	Y	-
91431704	91431839	D3S1552	91432	O	O	O	·	·	·	·	·	·	O	O	Y	-
91560697	91560992	SHGC-142032	91561	O	·	·	·	·	·	·	O	·	·	·	-	-
91588252	91588465	D3S3445	91588	O	·	·	·	·	·	·	·	·	·	·	-	-
91647432	91647645	D3S3445	91647	O	·	·	·	·	·	·	·	·	·	·	-	-
91674905	91675200	SHGC-142032	91675	O	·	·	·	·	·	·	O	·	·	·	-	-
91682560	91682737	WI-11235	91683	O	·	·	·	·	·	·	·	·	O	·	Y	-
91791278	91791442	AFMA049XG9	91791	O	·	·	·	·	·	·	·	·	O	·	Y	-
91870426	91870726	SHGC-84484	91870	O	·	·	·	·	·	·	O	·	·	·	-	-
91915468	91915812	SHGC-79129	91915	O	·	·	·	·	·	·	O	·	·	·	-	-
91936445	91936675	D3S2366	91936	O	·	·	·	·	·	·	·	·	O	·	Y	-
91976125	91976448	RH123819	91976	O	·	·	·	·	·	·	O	·	·	·	-	-
92004797	92005144	RH121633	92005	O	·	·	·	·	·	·	O	·	·	·	-	-
92154838	92155015	D3S3690	92155	O	O	O	O	O	O	O	·	·	·	O	Y	-
92160806	92161111	SHGC-112262	92161	O	·	·	·	·	·	·	O	·	·	·	-	-
92163169	92163449	RH122116	92163	O	·	·	·	·	·	·	O	·	·	·	-	-
92264841	92265155	SHGC-81552	92265	O	·	·	·	·	·	·	O	·	·	·	-	-
92369101	92369511	D3S1536	92369	O	·	O	·	·	·	·	·	·	·	·	Y	-
92409868	92410117	SHGC-13127	92410	O	·	·	·	·	·	·	·	·	·	·	-	-
92411590	92411843	SHGC-13127	92412	O	·	·	·	·	·	·	·	·	·	·	-	-
92588324	92588619	SHGC-147715	92588	O	·	·	·	·	·	·	O	·	·	·	-	-

92791854	92792195	SHGC-106507	,92792	O							O				-	-
92974675	92974817	RH48040	92975	O											-	-
92979596	92979854	RH69370	92980	O											-	-
92980208	92980335	SHGC-30466	92980	O			O		O	O					Y	-
92985741	92985841	SHGC-36500	92986	O			O								-	-
93014508	93014635	RH69243	93015	O											-	-

start	stop	marker	Kbp	STS	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	polymorphism	in_gene
93019006	93019347	SHGC-142686	93019	O							O				-	-
93020950	93021089	RH35958	93021	O											-	-
93021035	93021184	SHGC-33570	93021	O			O		O	O	O				Y	-
93021094	93021177	RH80683	93021	O											-	-
93047891	93048056	RH41657	93048	O											-	-
93049989	93050109	114YA9	93050	O									O		Y	-
93142014	93142298	RH119955	93142	O							O				-	-
93154510	93154740	D3S3398	93155	O								O	O		Y	-
93185391	93185730	SHGC-149730	93185	O							O				-	-
93245576	93245751	D3S3556	93246	O	O	O					O			O	Y	-
93245663	93245913	D3S3634	93246	O	O		O		O	O	O		O		Y	-
93284399	93284731	RH121031	93284	O							O				-	-
93295651	93295900	SHGC-12195	93296	O											-	-
93412723	93412955	SHGC-14457	93413	O					O	O					Y	-
93479742	93479871	D3S1251	93480	O		O								O	Y	-
93486562	93486867	SHGC-84550	93487	O							O				-	-
93550104	93550411	SHGC-83675	93550	O							O				-	-

All Sequence Maps Download [1]

Region Displayed: 85M-93M bp

~~[1] Max first 5000 available for download. For downloading all records, see our ftp site~~

MapViewer

Data As Table View Download All

GenomeScan On Sequence next
Total Models On Chromosome: 3052 [12 not localized]
Region Displayed: 92M-98M bp
Models in Region: 61

Download/View Sequence/Evidence All Sequence Maps Download Data

start	stop	name	orient	protein gi
92806181	92895090	Hs3_22631_31_1_1	-	131086
92917343	92925365	Hs3_22631_31_3_3	+	12855722
92935838	92936722	Hs3_22631_31_3_4	-	21961211
92948306	92974716	Hs3_22631_31_3_5	+	21450707
92982401	92982964	Hs3_22631_31_4_1	-	13097774
92990579	93003939	Hs3_22631_31_5_2	-	15929379
93015113	93047943	Hs3_22631_31_5_3	+	11545785
93066138	93168899	Hs3_22631_31_6_1	+	-
93188088	93215177	Hs3_22631_31_6_2	+	-
93292903	93378134	Hs3_22631_31_6_3	+	-
93388957	93409908	Hs3_22631_31_6_4	+	-
93454469	93543530	Hs3_22631_31_8_1	-	-
93678150	93700560	Hs3_22631_31_8_2	+	-
93959013	93959960	Hs3_22631_31_9_1	-	21618713
94323626	94363030	Hs3_22631_31_9_4	+	-
94424258	94657245	Hs3_22631_31_10_1	+	-
94677467	94829276	Hs3_22631_31_10_2	-	6678952
94905602	94988488	Hs3_22631_31_11_1	+	-
95212555	95215357	Hs3_22631_31_11_3	+	-
95371937	95373209	Hs3_22631_31_12_1	+	631471
95588617	95589407	Hs3_22631_31_13_2	+	9454362
95606693	95642088	Hs3_22631_31_13_3	+	11415026
95699334	95756537	Hs3_22631_31_13_4	+	-
95775517	95777119	Hs3_22631_31_13_5	+	-
95799192	95800016	Hs3_22631_31_13_6	+	8923710
95835577	95836106	Hs3_22631_31_13_7	-	-
95889082	96266544	Hs3_22631_31_14_1	+	6679661
96399039	96758379	Hs3_22631_31_15_1	+	23574684
96774473	96786892	Hs3_22631_31_15_3	+	-
96787080	96814083	Hs3_22631_31_16_1	+	14149815
96843596	96908127	Hs3_22631_31_17_1	+	-
96911612	96965571	Hs3_22631_31_18_1	+	23957694
96967441	96989850	Hs3_22631_31_19_1	-	23346418
96994559	97006440	Hs3_22631_31_20_1	+	-
97008940	97009230	Hs3_22631_31_20_2	-	6434195
97011861	97047920	Hs3_22631_31_20_3	-	20302089
97075711	97123078	Hs3_22631_31_20_4	+	7211485
97127010	97141414	Hs3_22631_31_20_5	+	18479486
97154954	97155895	Hs3_22631_31_20_6	+	21928607
97171642	97172574	Hs3_22631_31_20_7	+	21928607
97190956	97191897	Hs3_22631_31_20_8	+	21928607
97229679	97230597	Hs3_22631_31_20_10	+	21928607
97244300	97252258	Hs3_22631_31_20_11	+	21928607
97286589	97287518	Hs3_22631_31_20_12	+	21928607
97305159	97306088	Hs3_22631_31_20_13	+	21928607
97334169	97407068	Hs3_22631_31_20_14	+	22128945
97412922	97456489	Hs3_22631_31_20_15	+	22128945
97491833	97494631	Hs3_22631_31_21_1	+	21928228
97519937	97528497	Hs3_22631_31_21_2	+	21928226
97538915	97543680	Hs3_22631_31_22_1	-	7512789

start	stop	name	orient	protein gi
97545054	97553305	Hs3_22631_31_23_1	+	-

http://www.ncbi.nlm.nih.gov/mapview/maps.cgi?ORG=hum&CHR=3&MA

97554402	97555356	Hs3_22631_31_24_1	+	-	
97602948	97615760	Hs3_22631_31_26_1	-	7431644	
97647438	97650690	Hs3_22631_31_27_2	+	1706902	
97680387	97681941	Hs3_22631_31_28_1	+	13654239	
97682053	97682826	Hs3_22631_31_28_2	+	18204432	
97755460	97757189	Hs3_22631_31_30_1	+	-	
97790697	97796243	Hs3_22631_31_30_2	+	5174697	
97804580	97814249	Hs3_22631_31_30_3	+	5174697	
97821628	97923582	Hs3_22631_31_31_1	-	21314772	
97989181	98012733	Hs3_22631_31_32_2	-	-	

Genes On Sequence prev next

Total Genes On Chromosome: 1371 [7 not localized]

Region Displayed: 92M-98M bp

Genes in Region: 31

Download/View Sequence/Evidence All Sequence Maps Download Data

start	stop	symbol	orient.	links	evidence	cyto.	full name
92806177	92895519	PROS1	-	sv ev hm seq mm	C	3p11-q11.2	protein S (alpha)
92901608	92976730	DKFZp761H079	+	sv ev - seq mm	C	3q11.2	hypothetical protein DKFZp761H
92979558	92984352	LOC200895	-	sv ev - seq mm	I	3q11.2	similar to Dihydrofolate reductase
92984366	93049998	FLJ22609	+	sv ev - seq mm	C	3q11.2	hypothetical protein FLJ22609
93984847	94013594	LOC255025	+	sv ev - seq mm	I	3q11.2	LOC255025
94676689	94705459	LOC285216	-	sv ev - seq mm	I	3q11.2	similar to methylenetetrahydrofol
95371628	95374426	LOC285217	+	sv ev - seq mm	I	3q11.2	similar to Heterogeneous nuclear
95588617	95589407	LOC131167	+	sv ev - seq mm	P	3q12.1	similar to chromatin assembly fac
95606693	95642088	LOC285218	+	sv ev - seq mm	P	3q12.1	similar to ribosomal protein L18a,
95699334	95756537	LOC253194	-	sv ev - seq mm	PE	3q12.1	LOC253194
95799192	95800016	LOC200842	+	sv ev - seq mm	PE	3q12.1	similar to hypothetical protein H4
95836843	96266957	LOC203806	+	sv ev - seq mm	I	3q12.1	similar to Eph receptor A6 [Mus r
96443981	96447063	LOC285219	+	sv ev - seq mm	I	3q12.1	similar to putative p150 [Homo sa
96461877	96758547	LOC285220	+	sv ev - seq mm	I	3q12.1	similar to Eph receptor A6 [Mus r
96786863	96823364	ARL6	+	sv ev hm seq mm	C	3q12.1	ADP-ribosylation factor-like 6
96910671	96967222	DKFZp667G2110	+	sv ev - seq mm	C	3q12.1	hypothetical protein DKFZp667G
96966382	96994685	MINA53	-	sv ev hm seq mm	C	3q12.1	myc-induced nuclear antigen, 53 k
97008940	97009230	LOC151973	-	sv ev - seq mm	P	3q12.1	similar to GABA-C receptor rho3
97011861	97047920	LOC200959	-	sv ev - seq mm	P	3q12.1	similar to GABA receptor rho-3 si
97171642	97172574	LOC257089	+	sv ev - seq mm	P	3q12.1	similar to olfactory receptor MOR
97190956	97191897	OR5H1	+	sv ev - seq mm	C	3q12.1	olfactory receptor, family 5, subfa
97229679	97230597	LOC257091	+	sv ev - seq mm	P	3q12.1	similar to seven transmembrane h
97244300	97252258	LOC201543	+	sv ev - seq mm	P	3q12.1	similar to olfactory receptor MOR
97286589	97287518	LOC131284	+	sv ev - seq mm	P	3q12.1	similar to olfactory receptor MOR
97305159	97306088	LOC201537	+	sv ev - seq mm	P	3q12.1	similar to seven transmembrane h
97528731	97545220	C3orf4	-	sv ev - seq mm	C	3p11-q11	chromosome 3 open reading fram
97554290	97555372	GPR15	+	sv ev - seq mm	C	3q11.2-q13.1	G protein-coupled receptor 15
97601707	97615961	CPO	-	sv ev hm seq mm	C	3q12	coproporphyrinogen oxidase (copr
97647438	97650690	LOC201540	+	sv ev - seq mm	P	3q12.1	similar to Ferritin heavy chain 2 (
97754542	97818101	ST3GALVI	+	sv ev hm seq mm	C	3q12.1	alpha2,3-sialyltransferase
97818226	97923945	ESDN	-	sv ev - seq mm	C	3q12.1	endothelial and smooth muscle ce

Contig prev next

Total Contigs On Chromosome: 107 [3 not localized]

Region Displayed: 92M-98M bp

MapViewer

Contigs in Region: 1

Download/View Sequence/Evidence | All Sequence Maps | Download Data

start	stop	accession	orient
92803456	98065601	NT_022475.10	+

Transcripts On Sequence prev

Total Transcripts On Chromosome: 1484 [7 not localized]

Region Displayed: 92M-98M bp

Transcripts in Region: 31

Download/View Sequence/Evidence | All Sequence Maps | Download Data

start	stop	accession	locus	orient	links	evidence
92806177	92895314	NM_000313.1	PROS1	-	sv ev	?
92901650	92976730	NM_144996.1	DKFZp761H079	+	sv ev	?
92979558	92984352	XM_114312	LOC200895	-	sv ev	?
92984472	93048200	NM_022072.1	FLJ22609	+	sv ev	?
93984847	94013594	XM_173083	LOC255025	+	sv ev	I
94676689	94705459	XM_209520	LOC285216	-	sv ev	I
95371628	95374426	XM_209517	LOC285217	+	sv ev	I
95588617	95589407	XM_067237	LOC131167	+	sv ev	P
95606693	95642088	XM_210529	LOC285218	+	sv ev	P
95699334	95756537	XM_174349	LOC253194	-	sv ev	PE
95799192	95800016	XM_115974	LOC200842	+	sv ev	PE
95836843	96266957	XM_114973	LOC203806	+	sv ev	I
96443981	96447063	XM_209518	LOC285219	+	sv ev	I
96461877	96758547	XM_209519	LOC285220	+	sv ev	I
96787008	96823364	NM_032146.2	ARL6	+	sv ev	?
96910671	96967222	NM_153605.1	DKFZp667G2110	+	sv ev	?
96966817	96994344	NM_032778.3	MINA53	-	sv ev	?
97008940	97009230	XM_093538	LOC151973	-	sv ev	P
97011861	97047920	XM_116036	LOC200959	-	sv ev	P
97171642	97172574	XM_172287	LOC257089	+	sv ev	P
97190956	97191897	XM_172288	OR5H1	+	sv ev	C
97229679	97230597	XM_172289	LOC257091	+	sv ev	P
97244300	97252258	XM_116076	LOC201543	+	sv ev	P
97286589	97287518	XM_067296	LOC131284	+	sv ev	P
97305159	97306088	XM_116070	LOC201537	+	sv ev	P
97537739	97544770	NM_019895.1	C3orf4	-	sv ev	?
97554290	97555372	NM_005290.1	GPR15	+	sv ev	C
97601707	97615820	NM_000097.2	CPQ	-	sv ev	?
97647438	97650690	XM_116082	LOC201540	+	sv ev	P
97754977	97816248	NM_006100.1	ST3GALVI	+	sv ev	?
97818226	97923945	NM_080927.2	ESDN	-	sv ev	?

All Sequence Maps Download [1]

Region Displayed: 92M-98M bp

[1] Max first 5000 available for download. For downloading all records, see our ftp site

MapViewer

Data As Table View Download All !

GenomeScan On Sequence next
Total Models On Chromosome: 3052 [12 not localized]
Region Displayed: 97M-100M bp
Models in Region: 40

Download/View Sequence/Evidence All Sequence Maps Download Data

start	stop	name	orient	protein gi
97790697	97796243	Hs3_22631_31_30_2	+	5174697
97804580	97814249	Hs3_22631_31_30_3	+	5174697
97821628	97923582	Hs3_22631_31_31_1	-	21314772
97989181	98012733	Hs3_22631_31_32_2	-	-
98173396	98213603	Hs3_33226_31_1_2	-	-
98324365	98477060	Hs3_33226_31_1_3	-	21753825
98479106	98535761	Hs3_33226_31_1_4	-	-
98574984	98598166	Hs3_33226_31_2_1	+	-
98675776	98746164	Hs3_33226_31_4_1	+	-
98788708	98794161	Hs3_33226_31_4_2	+	17738302
98815957	98827081	Hs3_33226_31_6_1	+	-
98828439	98929045	Hs3_33226_31_6_2	-	7657037
98933593	99001120	Hs3_33226_31_6_3	-	-
99071945	99098924	Hs3_33226_31_6_5	-	-
99109198	99112470	Hs3_33226_31_6_6	-	-
99122713	99174030	Hs3_33226_31_6_7	+	14150167
99183712	99209791	Hs3_33226_31_7_1	+	15208195
99223399	99224098	Hs3_33226_31_7_2	+	5454166
99244076	99258965	Hs3_33226_31_7_3	-	-
99259044	99318996	Hs3_33226_31_8_1	+	8922842
99332817	99353293	Hs3_33226_31_10_1	+	9910460
99363589	99398974	Hs3_33226_31_12_1	-	7662673
99399142	99399504	Hs3_33226_31_13_1	+	-
99421754	99429535	Hs3_33226_31_13_2	+	-
99455995	99460200	Hs3_33217_31_1_1	+	-
99504823	99509397	Hs3_33217_31_2_1	+	-
99525872	99530612	Hs3_33217_31_2_2	-	-
99561036	99581292	Hs3_33217_31_2_3	+	8922243
99607735	99609683	Hs3_33217_31_3_1	+	-
99614131	99665452	Hs3_33217_31_4_1	+	14249454
99669324	99686848	Hs3_33217_31_4_2	-	12851526
99717960	99752805	Hs3_33217_31_6_1	+	21361370
99754769	99835182	Hs3_33217_31_7_1	-	22539628
99843746	99846745	Hs3_33217_31_7_2	+	-
99868875	99996718	Hs3_33217_31_9_1	-	13376177
100083452	100131286	Hs3_37578_31_1_1	+	-
100195705	100200592	Hs3_37578_31_1_3	+	-
100218609	100222220	Hs3_37578_31_1_5	-	-
100317396	100362047	Hs3_6020_31_1_2	+	-
100375028	100466046	Hs3_6020_31_2_1	-	7706717

97821447	97823905	BU171889.1	-				
97821447	97822097	exon		862	207	98.63	1195.0
97822835	97822972	exon		206	69	100.0	266.0
97823856	97823905	exon		68	19	100.0	96.8
97821573	97833573	BM450083.1	-				
97821573	97822097	exon		943	407	96.28	892.0
97822835	97822972	exon		406	269	100.0	266.0
97823856	97823905	exon		268	219	100.0	96.8
97830337	97830430	exon		218	125	100.0	181.0
97833450	97833573	exon		124	1	100.0	243.0
97821586	97822972	BQ014739.1	-				
97821586	97822097	exon		24	535	99.61	983.0
97822835	97822972	exon		536	673	100.0	262.0
97821888	97821991	AW583777.1	-				
97821888	97821991	exon		104	1	100.0	204.0
97822030	97830401	BE925427.1	-				
97822030	97822097	exon		317	250	100.0	131.0
97822835	97822972	exon		249	112	100.0	266.0
97823856	97823905	exon		111	64	92.0	68.0
97830337	97830401	exon		63	1	96.92	111.0
97822214	97822460	AA457544.1	+				
97822214	97822460	exon		248	2	100.0	484.0

start	stop	Hits:UniGene	...				
97823117	97823576	AW102917.1	-				
97823117	97823576	exon		1	460	98.91	870.0
97823126	97823532	AI246663.1	-				
97823126	97823532	exon		1	407	99.75	792.0
97823561	97843243	BG536257.1	+				
97823561	97823711	exon		2	152	100.0	291.0
97823805	97823873	exon		153	221	100.0	133.0
97843044	97843243	exon		222	421	100.0	385.0
[illegible]							
97823646	97823905	exon		18	277	100.0	500.0
97830337	97830416	exon		278	357	98.75	150.0
97823646	97833500	AI302412.1	-				
97823646	97823905	exon		10	269	99.62	502.0
97830337	97830430	exon		270	363	98.94	175.0
97833450	97833500	exon		364	414	100.0	100.0
97823646	97840058	BU634176.1	-				
97823646	97823905	exon		19	278	100.0	500.0
97830337	97830430	exon		279	372	100.0	181.0
97833450	97833575	exon		373	498	100.0	242.0
97833724	97833810	exon		499	585	100.0	167.0
97834588	97834738	exon		586	737	94.87	260.0
97840031	97840058	exon		738	765	96.43	45.9
97827190	97827819	AV649685.1	+				
97827190	97827819	exon		630	1	97.95	1154.0
97827412	97827896	N42165.1	-				
97827412	97827896	exon		503	1	95.43	815.0
97827421	97827819	AV651165.1	+				
97827421	97827819	exon		399	1	100.0	783.0
97827459	97827725	AA373072.1	-				
97827459	97827725	exon		267	1	99.25	514.0
97827962	97828331	H78999.1	+				
97827962	97828331	exon		375	1	98.13	683.0

!!! Sorry, too many records, only first 1000 returned !!!

Genes On Sequence prev next
Total Genes On Chromosome: 1371 [7 not localized]
Region Displayed: 97M-100M bp

MapViewer

Genes in Region: 16

Download/View Sequence/Evidence | All Sequence Maps | Download Data

start	stop	symbol	orient	evidence	cyto.	full name
97754542	97818101	ST3GALVI	+	sv ev hm seq mm	C 3q12.1	alpha2,3-sialyltransferase
97818226	97923945	ESDN	-	sv ev - seq mm	C 3q12.1	endothelial and smooth muscle cell-de
98324365	98477060	LOC131563	-	sv ev - seq mm	P 3q12.2	similar to KIAA1998 protein [Homo s
98636636	99176627	LOC57086	+	sv ev - seq mm	C 3q12.2	smooth muscle cell-expressed and mac
98845954	98874227	DOC1	-	sv ev hm seq mm	C 3q12.2	downregulated in ovarian cancer 1
98931146	98933303	PRO0641	+	sv ev - seq mm	C 3q12.2	PRO0641 protein
99259025	99323273	FLJ11046	+	sv ev - seq mm	C 3q12.2	hypothetical protein FLJ11046
99332749	99353662	NIT2	+	sv ev hm seq mm	C 3q12.2	Nit protein 2
99361490	99399601	TOMM70A	-	sv ev - seq mm	C 3q12.2	translocase of outer mitochondrial mem
99496893	99581715	FLJ10134	+	sv ev - seq mm	C 3q12.2	hypothetical protein FLJ10134
99613875	99699695	FLJ14454	+	sv ev - seq mm	C 3q12.3	hypothetical protein FLJ14454
99713590	99753237	TFG	+	sv ev hm seq mm	C 3q11-q12	TRK-fused gene
99753611	99850823	NESHBP	-	sv ev - seq mm	C 3q12	DKFZP586L2024 protein
99866706	99997744	FLJ21551	-	sv ev - seq mm	C 3q12.3	hypothetical protein FLJ21551
100195705	100200592	LOC285274	+	sv ev - seq mm	PE 3q12.3	LOC285274
100372460	100466795	IMPG2	-	sv ev - seq mm	C 3q12.2-q12.3	interphotoreceptor matrix proteoglycan

Contig prev next

Total Contigs On Chromosome: 107 [3 not localized]

Region Displayed: 97M-100M bp

Contigs in Region: 5

Download/View Sequence/Evidence | All Sequence Maps | Download Data

start	stop	accession	orient
92803456	98065601	NT_022475.10	+
98075602	99441326	NT_033050.3	+
99451327	100061927	NT_033041.3	+
100071928	100225279	NT_037574.1	+
100235280	101803320	NT_005863.13	+

Transcripts On Sequence prev

Total Transcripts On Chromosome: 1484 [7 not localized]

Region Displayed: 97M-100M bp

Transcripts in Region: 18

Download/View Sequence/Evidence | All Sequence Maps | Download Data

start	stop	accession	locus	orient	evidence	
97754977	97816248	NM_006100.1	ST3GALVI	+	sv ev	?
97818226	97923945	NM_080927.2	ESDN	-	sv ev	?
98324365	98477060	XM_067445	LOC131563	-	sv ev	P
98636636	98678487	NM_020351.1	LOC57086	+	sv ev	?
98636643	98794192	NM_001850.2	LOC57086	+	sv ev	?
98815862	99176627	NM_032359.1	MGC4308	+	sv ev	?
98846042	98849100	NM_014890.1	DOC1	-	sv ev	?
98931146	98933303	NM_014135.1	PRO0641	+	sv ev	?
99259037	99321696	NM_018309.1	FLJ11046	+	sv ev	?
99332796	99353376	NM_020202.1	NIT2	+	sv ev	?
99361490	99399065	NM_014820.1	TOMM70A	-	sv ev	C
99496893	99581715	NM_018004.1	FLJ10134	+	sv ev	?
99613875	99699695	NM_032787.1	FLJ14454	+	sv ev	?
99713725	99753237	NM_006070.2	TFG	+	sv ev	?
99753612	99850823	NM_015429.1	NESHBP	-	sv ev	?
99866706	99997724	NM_024801.1	FLJ21551	-	sv ev	?
100195705	100200592	XM_212510	LOC285274	+	sv ev	PE
100372964	100466795	NM_016247.1	IMPG2	-	sv ev	?

All Sequence Maps Download

Region Displayed: 97M-100M bp

Appendix B: List of candidate open reading frames in critical region.

START		NAME	ORIENT	#	PROTEIN GI	YEAST	CHLAMY	WORM		RFX-box	SMART / PFAM DOMAINS IN h_ORF	BEST BLAST HIT
				0								
85771306	85778779	Hs3_6188_31_2_2	+	1	19923268							58kDa interferon induced kinase (p58IPK)
85938638	86012976	Hs3_6188_31_3_1	+	2	-							?
86242086	86511026	Hs3_22575_31_1_3	-	3	-							?
89784809	89867903	Hs3_22798_31_3_1	-	4	12840558	?		W	B0205.10 / Y49E10.10	-187		? / rat amphoterin
89881557	89883097	Hs3_22798_31_4_1	-	5	548638	Y	C	W				phosporibosyl pyrophosphate amidotransferase
89919395	89919868	Hs3_22798_31_5_1	-	6	-		?					cytosolic glutamine synthetase
89956522	90004142	Hs3_22798_31_5_2	+	7	-							?
90020448	90031617	Hs3_22798_31_5_3	-	8	-							?
90057888	90084189	Hs3_22798_31_6_1	+	9	7661634	Y	C	W				#
90090261	90106829	Hs3_22798_31_7_1	-	10	488129	Y		W				pituitar specific POU-domain DNA binding factor
90153635	90154345	Hs3_22798_31_8_1	+	11	105815	Y		W				similar to keratin 8, type II
90154471	90155090	Hs3_22798_31_8_2	+	12	307083			W	F10C1.2a		Intermediate filament	similar to keratin 8, type II
90164613	90319329	Hs3_22798_31_8_3	-	13	-							-
90371996	90425486	Hs3_22798_31_8_4	+	14	13129148							#
90460720	90461142	Hs3_22798_31_8_5	-	15	16741033	Y	C	W				proteasomal ATPase (S4)
90461708	90461986	Hs3_22798_31_8_6	-	16	16741033	Y	C	W				proteasomal ATPase (S4)
90623251	90659882	Hs3_22798_31_9_1	-	17	-							?
90720397	90764698	Hs3_22798_31_9_2	-	18	-							?
90821071	90822114	Hs3_22798_31_10_1	+	19	10835197			W	C02D4.2a / Y22D7AR.13		transmembrane (rhodopsin)	5-hydroxytryptamine (serotonin) receptor 1F
90885733	90886240	Hs3_22798_31_12_1	-	20	21361099							trinucleotide repeat DNA binding protein p20-CGG BP
90889584	90969572	Hs3_22798_31_13_1	+	21	6912832							rearranged L-myc fusion
90969692	90971389	Hs3_22798_31_14_1	+	22	23955470	Y		?				melanoma associated antigen
90980323	90986899	Hs3_22798_31_15_1	+	23	18848333							#
91018707	91019021	Hs3_22798_31_15_2	+	24	-							?
91058841	91087686	Hs3_22798_31_15_3	-	25	-							?
91147420	91230056	Hs3_22798_31_15_4	+	26	10179834	Y	C	W				ATP binding cassette, subfamily F, member 2
91295622	91339773	Hs3_22798_31_15_5	+	27	-							#
91429351	91495111	Hs3_22591_31_1_2	-	28	-							?
91597544	91614359	Hs3_22591_31_1_3	+	29	-							?
91666182	91698452	Hs3_22591_31_1_4	-	30	-							?
91746426	91747160	Hs3_22591_31_1_5	+	31	-							similar to NMDA receptor regulated gene 2
91845110	91846033	Hs3_22591_31_1_7	-	32	7669492	Y	C	W				similar to glyceraldehyde 3-phosphate dehydrogenase
92010199	92148747	Hs3_22653_31_2_1	+	33	21361241	Y	C	W				ephrin receptor Eph A3
92162806	92216028	Hs3_22653_31_3_1	-	34	-							?
92256207	92293456	Hs3_22653_31_3_2	-	35	5835658	Y	C					ATP synthase F0 subunit 6
92532080	92566316	Hs3_37577_31_1_2	-	36	-							?
92744573	92776910	Hs3_37577_31_1_4	+	37	-							?
92806181	92895090	Hs3_22631_31_1_1	-	38	131086		C	W	F56H11.1b / c	-928	GLA / EGF_CA / LamG	vitamin K dependent protein S (alpha)
92917343	92925365	Hs3_22631_31_3_3	+	39	12855722	Y	C	W				ADP ribosylation factor-like 2
92935838	92936722	Hs3_22631_31_3_4	-	40	21961211	Y		W				#
92948306	92974716	Hs3_22631_31_3_5	+	41	21450707	Y		W				ADP ribosylation factor-like 2
92982401	92982964	Hs3_22631_31_4_1	-	42	13097774	Y		W				dihydrofolate reductase
92990579	93003939	Hs3_22631_31_5_2	-	43	15929379	Y	C	W				similar to retinoblastoma binding protein 4
93015113	93047943	Hs3_22631_31_5_3	+	44	11545785	Y		W				#
93066138	93168899	Hs3_22631_31_6_1	+	45	-							?
93188088	93215177	Hs3_22631_31_6_2	+	46	-							?
93292903	93378134	Hs3_22631_31_6_3	+	47	-							?
93388957	93409908	Hs3_22631_31_6_4	+	48	-							-
93454469	93543530	Hs3_22631_31_8_1	-	49	-							?

KEY: ? = No convincing match; # = excellent match but protein not named / hypothetical; - = no blast hit found

START	Stop	NAME	ORIENT	#	PROTEIN GI	YEAST	CHLAMY	WORM	C.elegans	RFX-box	DOMAINS IN HUMAN ORF	BEST BLAST HIT
92806181	92895090	Hs3_22631_31_1_1	-	1	131086		C	W				see BBS3 cr l
92917343	92925365	Hs3_22631_31_3_3	+	2	12855722	Y	C	W				see BBS3 cr l
92935838	92936722	Hs3_22631_31_3_4	-	3	21961211	Y		W				see BBS3 cr l
92948306	92974716	Hs3_22631_31_3_5	+	4	21450707	Y		W				see BBS3 cr l
92982401	92982964	Hs3_22631_31_4_1	-	5	13097774	Y		W				see BBS3 cr l
92990579	93003939	Hs3_22631_31_5_2	-	6	15929379	Y	C	W				see BBS3 cr l
93015113	93047943	Hs3_22631_31_5_3	+	7	11545785	Y		W				see BBS3 cr l
93066138	93166899	Hs3_22631_31_6_1	+	8	-							see BBS3 cr l
93166868	93215177	Hs3_22631_31_6_2	+	9	-							see BBS3 cr l
93292903	93378134	Hs3_22631_31_6_3	+	10	-							see BBS3 cr l
93388957	93409908	Hs3_22631_31_6_4	+	11	-							see BBS3 cr l
[redacted]		Hs3_22631_31_8_1	-	12	-							see BBS3 cr l
93678150	93700560	Hs3_22631_31_8_2	+	13	-							?
93959013	93959960	Hs3_22631_31_9_1	-	14	21618713	Y	C	W			WD40	#
94323626	94363030	Hs3_22631_31_9_4	+	15	-							?
94424258	94657245	Hs3_22631_31_10_1	+	16	-							?
94677467	94829276	Hs3_22631_31_10_2	-	17	6678952	Y	C	W			THF / DHG / CYH	similar to methylenetetrahydrofolate dehydrogenase
94905602	94988488	Hs3_22631_31_11_1	+	18	-							?
95212555	95215357	Hs3_22631_31_11_3	+	19	-							?
95371937	95373209	Hs3_22631_31_12_1	+	20	631471			W	F26B1.2A		2 KH domains (RNA binding)	nuclear ribonucleoprotein K isoform A
95588617	95589407	Hs3_22631_31_13_2	+	21	9454362	Y	C	W				chromatin assembly factor-1 p48 subunit
95606693	95642088	Hs3_22631_31_13_3	+	22	11415028	Y	C	W				ribosomal protein L18a
95699334	95756537	Hs3_22631_31_13_4	-	23	-							?
95775517	95777119	Hs3_22631_31_13_5	+	24	-							?
95799192	95800016	Hs3_22631_31_13_6	+	25	8923710							
95835577	95836106	Hs3_22631_31_13_7	-	26	-							?
95869082	96266544	Hs3_22631_31_14_1	+	27	6679661			W	M03A1.1A		EPH and FN3 domains	Eph receptor
96399039	96758379	Hs3_22631_31_15_1	+	28	23574684	Y	C	W			FN3 / transmem. / STYkc	Eph receptor A6
96774473	96786892	Hs3_22631_31_15_3	+	29	-							?
96787080	96814083	Hs3_22631_31_16_1	+	30	14149815	Y	C	W			ARF / small_GTPase	ADP ribosylation-like 6
96843596	96908127	Hs3_22631_31_17_1	+	31	-			W	ZK783.1 / Y36B5A.1		XTALbg (beta-gamma crystallin)	7 / axomene associated
96911612	96965571	Hs3_22631_31_18_1	+	32	23957694							
96967441	96989850	Hs3_22631_31_19_1	-	33	23346418			W	T28F2.4		JmjC domain	myc-induced nuclear antigen
96994559	97006440	Hs3_22631_31_20_1	+	34	-							?
97008940	97009230	Hs3_22631_31_20_2	-	36	6434195			W			transmembrane	GABA receptor rho-3 subunit
97011861	97047920	Hs3_22631_31_20_3	-	36	20302069			W	ZC462.1 / T21C12.1b	-664	transmembrane	GABA receptor rho-3 subunit
97075711	97123078	Hs3_22631_31_20_4	+	37	7211485			W			transmembrane	olfactory receptor MOR 183-9
97127010	97141414	Hs3_22631_31_20_5	+	38	18479486			W	C30F12.6		transmembrane	olfactory receptor MOR183-9
97154954	97155895	Hs3_22631_31_20_6	+	39	21928607			W				
97171642	97172574	Hs3_22631_31_20_7	+	40	21928607			W	Y39A3B.5			olfactory receptor MOR183-1
97190956	97191897	Hs3_22631_31_20_8	+	41	21928607			W				
97229679	97230597	Hs3_22631_31_20_10		42	21928607			W				
97244300	97252258	Hs3_22631_31_20_11		43	21928607							
97286589	97287518	Hs3_22631_31_20_12		44	21928607			W				
97305159	97306088	Hs3_22631_31_20_13		45	21928607			W				
97334189	97407068	Hs3_22631_31_20_14		46	22128945			W	Y39A3B.5		transmembrane	olfactory receptor MOR182-4
97412922	97456489	Hs3_22631_31_20_15		47	22128945			W				
97491833	97494631	Hs3_22631_31_21_1	+	48	21928228			W	Y39A3B.5		transmembrane	olfactory receptor MOR184-3
97519937	97528497	Hs3_22631_31_21_2	+	49	21928226			W				
97538915	97543680	Hs3_22631_31_22_1	-	50	7512789							
97545054	97553305	Hs3_22631_31_23_1	+	51	-							?
97554402	97555356	Hs3_22631_31_24_1	+	52	-			W	F41E7.3	-119	Rhodopsin GPCR and transmem	G-protein coupled receptor (GPR15)
97602948	97615760	Hs3_22631_31_26_1	-	53	7431644	Y	C					coproporphyrinogen oxidase
97647438	97650690	Hs3_22631_31_27_2	+	54	1706902			W	D1037.3		ferritin	similar to ferritin heavy chain 2
97680387	97681941	Hs3_22631_31_28_1	+	55	13654239	Y	C	W			WW domain	*suppressor of deltex / Nedd-4 ubiquitin ligase*
97682053	97682826	Hs3_22631_31_28_2	+	56	18204432	Y	C	W			HECTc	suppressor of deltex / Nedd-4 ubiquitin ligase
97755460	[redacted]	Hs3_22631_31_30_1	+	57	-							?
[redacted]	97796243	Hs3_22631_31_30_2	+	58	5174697							
97804580	97814249	Hs3_22631_31_30_3	+	59	5174697							
97821628	97923582	Hs3_22631_31_31_1	-	60	21314772			W	ZC116.3		CUB and FA58C domains	smooth muscle neuropilin like protein
97989181	98012733	Hs3_22631_31_32_2	-	61	-							

KEY: ? = No convincing match; # = excellent match but protein not named / hypothetical; - = no blast hit found

III BBS3 cr

START	STOP	NAME	ORIENT	#	PROTEIN GI	YEAST	CHLAMY	WORM		RFX-box	SMART / PFAM DOMAINS IN h_ORF	BEST BLAST HIT
97790697	97796243	Hs3_22631_31_30_2	+		5174697							see BBS3 cr II
97804580	97814249	Hs3_22631_31_30_3	+		5174697							see BBS3 cr II
97821628	97923582	Hs3_22631_31_31_1	-		21314772							see BBS3 cr II
97989181	98012733	Hs3_22631_31_32_2	-		-							see BBS3 cr II
98173396	98213603	Hs3_33226_31_1_2	-	1	-							?
98324365	98477060	Hs3_33226_31_1_3	-	2	21753825							# / (Rho interacting protein2 ?)
98479106	98535761	Hs3_33226_31_1_4	-	3	-							?
98574984	98598166	Hs3_33226_31_2_1	+	4	-							?
98675776	98746164	Hs3_33226_31_4_1	+	5	-							?
98788708	98794161	Hs3_33226_31_4_2	+	6	17738302			W	F01G12.5a / b	-1378	C1Q	similar to alpha-1 type viii collagen
98815957	98827081	Hs3_33226_31_6_1	+	7	-							similar to alpha-1 type viii collagen
98828439	98929045	Hs3_33226_31_6_2	-	8	7657037	Y		W				down reg. in ovarian cancer / GPBP interacting protein 90
98933593	99001120	Hs3_33226_31_6_3	-	9	-							?
99071945	99098924	Hs3_33226_31_6_5	-	10	-							?
99109198	99112470	Hs3_33226_31_6_6	-	11	-							-
99122713	99174030	Hs3_33226_31_6_7	+	12	14150167							#
99183712	99209791	Hs3_33226_31_7_1	+	13	15208195	Y	C	W				#
99223399	99224098	Hs3_33226_31_7_2	+	14	5454166	Y	C	W				vessicle transport v-snare Vit-1 / NSF attach. prot. receptor
99244078	99258965	Hs3_33226_31_7_3	-	15	-							?
99259044	99318996	Hs3_33226_31_8_1	+	16	8922842			W	F20D1.2		TCB / RHOD	?
99332817	99353293	Hs3_33226_31_10_1	+	17	9910460	Y	C	W				Nit protein 2
99363589	99398974	Hs3_33226_31_12_1	-	18	7662673	Y	C	W			TPR	mitochondreal import receptor
99399142	99399504	Hs3_33226_31_13_1	+	19	-							-
99421754	99429535	Hs3_33226_31_13_2	+	20	-							?
99455993	99460200	Hs3_33217_31_1_1	+	21	-							?
99504823	99509397	Hs3_33217_31_2_1	+	22	-							?
99525872	99530612	Hs3_33217_31_2_2	-	23	-							?
99561036	99561292	Hs3_33217_31_2_3	+	24	8922243							# / dermal papilla derived protein 7
99607735	99609683	Hs3_33217_31_3_1	+	25	-							?
99614131	99665452	Hs3_33217_31_4_1	+	26	14249454			W	B0286.2	-573	transmemb. / coiled coil / GPS	G-protein coupled receptor 128
99669324	99686848	Hs3_33217_31_4_2	-	27	12851526							glia maturation factor B
99717960	99752805	Hs3_33217_31_6_1	+	28	21361320							TFG
99754769	99835182	Hs3_33217_31_7_1	-	29	22539628	?		W	Y51B11A.1			Nesh binding protein
99843746	99846745	Hs3_33217_31_7_2	+	30	-							?
99868875	99996718	Hs3_33217_31_9_1	-	31	13376177						FN3	C protein Immunoglobulin / Nesh binding protein
100083452	100131286	Hs3_37578_31_1_1	+	32	-							?
100195705	100200592	Hs3_37578_31_1_3	+	33	-							#
100218609	100222220	Hs3_37578_31_1_5	-	34	-							?
100317396	100362047	Hs3_6020_31_1_2	+	35	-							?
100375028	100466046	Hs3_6020_31_2_1	-	36	7706717	?		W	Zk783.1 / K06A9.1b			interphotoreceptor matrix proteoglycan 2

KEY: ? = No convincing match; # = excellent match but protein not named / hypothetical; - = no blast hit found

REFERENCES

Ansley S.J., Badano J.L., Blacque O.E., Hill J., Hoskins B.E., Kim J.C., Ross A.J., Lewis R.A., Mah A.K., Johnsen R.C., Cavender J.C., Leroux M.R., Beales P.L., Katsanis N. (2003). Ciliary dysfunction is a candidate mechanism for the pleiotropic phenotype of Bardet-Biedl syndrome. Submitted to *Nature*

Altschul S.F., Gish W., Miller W., Myers E.W., Lipman D.J. (1990) Basic local alignment search tool. *Journal of Molecular Biology* 215, 403-410.

Badano J.L., Ansley S.J., Leitch C.C., Lewis R.A., Lupski J.R., Katsanis N. (2003) Identification of a Novel Bardet-Biedl Syndrome Protein, BBS7, That Shares Structural Features with BBS1 and BBS2. *American Journal of Human Genetics* 72, 650-658.

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

Baker ME, French FS, Joseph DR. (1987). Vitamin K-dependent protein S is similar to rat androgen-binding protein. *Biochemistry Journal* 243, 293-6.

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., Lupski J.R., Katsanis N. (2003) Genetic Interaction of BBS1 Mutations with Alleles at Other BBS Loci Can Result in Non-Mendelian Bardet-Biedl Syndrome. *American Journal of Human Genetics* 72, 1187-1199.

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., Lupski J.R. (2001) Genetic and Mutational Analysis of a Large Multiethnic Bardet-Biedl Cohort Reveal a Minor Involvement of BBS6 and Delineate the Critical Intervals of Other Loci. *American Journal of Human Genetics* 68, 606-616.

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

Burford E.A., Riise R., Teague P.W., Porter K., Thomson K.L., Moore A.T., Jay M., Warburg M., Schinzel A., Tommerup N., Tornqvist K., Rosenberg T., Patton M., Mansfield D.C., Wright A.F. (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.M., Vaessin H.E.F., de la Chapelle, A. (2001) The Land Between Mendelian and Multifactorial Inheritance. *Science* 293, 2213-2214.

Burley S.K. (1994) DNA-binding motifs from eukaryotic transcription factors. *Current Opinion in Structural Biology* 4, 3-11.

Carey J.C., Hall B.D. (1978) Confirmation of the Cohen syndrome. *The Journal of Pediatrics* 93, 239-244.

Carmi R., Elbedour K., Stone E.M., Sheffield V.C. (1995) Phenotypic Differences Among Patients With Bardet-Biedl Syndrome Linked to Three Different Chromosome Loci. *American Journal of Medical Genetics* 59, 199-203.

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

Chen J., Knowles H.J., Heberrt J.L., Hackett B.P. (1998) Mutation of the Mouse Hepatocyte Nuclear Factor/Forkhead Homologue 4 Gene Results in an Absence of Cilia and Random Left-Right Asymmetry. *Journal of Clinical Investigation* 102, 1077-1082.

Cohen M.M., Hall B.D., Smith D.W., Graham C.B., Lampert K.J. (1973) A new syndrome with hypotonia, obesity, mental deficiency, and facial, oral, ocular, and limb abnomalies. *The Journal of Pediatrics* 83, 280-284.

Cole D.G. (2003) The Intraflagellar Transport Machinery of *Chlamydomonas reinhardtii* 4, 435-442.

Collin G.B., Marshall J.D., Ikeda A., So W.V., Russell-Eggitt I., Maffei P., Beck S., Boerkoel C.F., Sicolo N., Martin M., Nishina P.M., Naggert J.K. (2002) Mutations in ALMS1 cause obesity, type 2 diabetes and neurosensory degeneration in Alstrome syndrome. *Nature* Genetics 31, 74-78.

Dose A.C., Burnside B. (2002) A Class III Myosin Expressed in the Retina Is a Potential Candidate for Bardet-Biedl Syndrome. *Genomics* 79, 621-624.

Eickers E.R., Green J.S., Stockton D.W., Jackman C.S., Whelan J., McNamara J.A., Johnson G.J., Lupski J.R., Katsanis N. (2002) Newfoundland Rod-Cone Dystrophy, an Early-Onset Retinal Dystrophy, Is Caused by Splice-Junction Mutations in RLBP1. *American Journal of Human Genetics* 70, 955-964.

Emery P., Durand B., Mach B., Reith W. (1996). RFX proteins, a novel family of DNA binding proteins in the eukaryotic kingdom. *Nucleic Acids Research* 24, 803-807.

Emery P.l, Strubin M., Hofmann K., Bucher P., Mach B., Reith W. A Consensus Motif in the RFX DNA Binding Domain and Binding Domain Mutants with Altered Specificity. (1996). *Molecular and Cellular Biology* 16, 4486-4494.

Feitosa M.F., Borecki I.B., Rich S.S., Arnett D.K., Sholinsky P., Myers R.H., Leppert M., Province M.A. (2002) Quantitative-Trait Loci Influencing Body-Mass Index Reside on Chromosomes 7 and 13: The National Heart, Lung, and Blood Institute Family Heart Study. *American Journal of Human Genetics* 70, 72-82.

Gabriel S.B., Salomon R., Pelet A., Angrist M., Amiel J., Fornage M., Attie-Bitach T. Olson J.M., Hofstra R., Buys C., Steffann J., Munnich A., Lyonnet S., Chakravarti A. (2002) Segregation at three loci explains familial and population risk in Hirschsprung disease. *Nature Genetics* 31, 89-93.

Ghadami M., Tomita H.A., Najafi M.T., Damavandi E., Farahvash M.S., Yamada K., Majidzadeh A.K. Niikawa N. (2000) Bardet-Biedl syndrome type 3 in an Iranian family: clinical study and confirmation of disease localization. *American Journal of Medical Genetics* 94, 433-437.

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., Pryse-Phillips W. (1989) The Cardinal Manifestations of Bardet-Biedl Syndrome, a Form of Laurence-Moon-Biedl Syndrome. *The New England Journal of Medicine* 321, 1002-1009.

Harnett, J.D., Green J.S., Cramer B.C., Johnson G., Chafe L., McManamon P., Farid, N.R., Pryse-Phillips W., Parfrey P.S. (1998) The Spectrum of Renal Disease in Laurence-Moon-Biedl Syndrome. *The New England Journal of Medicine* 319, 615-618.

Hearn T., Renforth G.L., Spalluto C., Hanley N.A., Piper K., Brickwood S., White C., Connolly V., Taylor J.F., Russell-Eggitt I., Bonneau D., Walder M., Wilson D.I. (2002) Mutation of ALMS1, a large gene with a tandem repeat encoding 47 amino acids, causes Alstrom syndrome. *Nature Genetics* 31, 79-83.

Helwig U., Imai K., Schmahl W., Thomas B.E., Varnum D.S., Nadeau J.H., Balling R. (1995) Interaction between *undulated* and *Patch* leads to an extreme form of spina bifida in double-mutant mice. *Nature Genetics* 11, 60-63.

Kajiwara K., Berson E.L., Dryja T.P. (1994) Digenic Retinitis Pigmentosa Due to Mutations at the Unlinked Peripherin/RDS and ROM1 Loci. *Science* 264, 1604-1608.

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., Lupski J.R. (2001) 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., Lupski J.R. (2000) Mutations in MKKS cause obesity, retinal dystrophy and renal malformations associated with Bardet-Biedl syndrome. *Nature Genetics* 26, 67-70.

Katsanis N., Eichers E.R., Ansley S.J., Lewis R.A., Kayerili H., Hoskins B.E., Scambler P.J., Beales P.L., Lupski J.R. (2002) BBS4 Is a Minor Contributor to Bardet-Biedl Syndrome and May Also Participate in Triallelic Inheritance. *American Journal of Human Genetics* 71, 22-29.

Katsanis N., Lewis R.A., Stockton D.W., Mai P.M.T., Baird L., Beales P.L., Leppert M., 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. *American Journal of Human Genetics* 65, 1672-1679.

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

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

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

Lettice L.A., Heaney S.J.H., Purdie L.A., Li L., de Beer P., Oostra B.A., Goode D., Elgar G., Hill R.E., de Graaff E. (2003). A long-range *Shh* enhancer regulates expression in the developing limb and fin and is associated with preaxial polydactyly. *Human Molecular Genetics* 12, 1725-1735.

Liu, X., Vansant G., Udovichenko I.P., Wolfrum U., Williams D.S. (1997). Myosin VIIa, the Product of Usher 1B Syndrome Gene, Is Concentrated in the Connecting Cilia of Photoreceptor Cells. *Cell Motility and the Cytoskeleton* 37, 240-252.

Lupski J.R., Montes de Oca-Luna R., Slaugenhaupt S., Pentao L., Guzzetta V., Trask B.J., Saucedo-Cardenas O., Barker D.F., Killian J.M., Garcia C.A., Chakravarti A., Patel P.I. (1991) DNA Duplication Associated with Charcot-Marie-Tooth Disease Type 1A. *Cell* 66, 219-232.

Moore S.J., Green J.S., Fan Y., Bhogal A.K., Dicks E., Stefanelli M., Fernandez B.A., Murphy C., Beales P.L., Katsanis N., Davidson W.S., Parfrey P.S. (2003) The Clinical and Genetic Epidemiology of Laurence-Moon-Bardet-Biedl Syndrome in Newfoundland. Submitted to *The New England Journal of Medicine*.

Morell R., Spritz R.A., Ho L., Pierpont J., Guo W., Friedman T.B., Asher Jr J.H. (1997) Apparent digenic inheritance of Waardenburg syndrome type 2 (WS2) and autosomal recessive ocular albinism (AROA). *Human Molecular Genetics* 6, 659-664.

Murzin, A.G. (1992) Structural Principles for the Propellor Assembly of ß-Sheets: The Preference for Seven-Fold Symmetry. *Protiens: Structure, Function and Genetics* 14, 191-201.

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., Iannaccone A., Jacobson S.G., Weleber R.G., Wright A.F., Riise R., Hennekam R.C.M., Luleci G., Berker-Karauzum S., Biesecker L.G., Stone E.M., Sheffield V.C. (2003) Evaluation of Complex Inheritance Involving the Most Common Bardet-Biedl Syndrome Locus (BBS1). *American Society of Human Genetics* 72, 429-437.

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

Pazour G.J., Dickert B.L., Vucica Y., Seeley E.S., Rosenbaum J.L., Witman G.B., Cole D.G. (2000). *Chlamydomonas IFT*88 and Its Mouse Homologue, Polycystic Kidney Disease Gene *Tg*737, Are Required for Assembly of Cilia and Flagella. *The Journal of Cell Biology* 151, 709-718.

Pennarun, G., Bridoux, A. Escudier E., Dastot-Le Moal, F., Cacheux V., Amselem S., Duriez B. Isolation and Expression of the Human *hPF20* Gene Orthologous to *Chlamydomonas pf*20. *American Journal of Respiratory Molecular Biology* 26, 362-370.

Riise R., Andreasson S., Borgstrom M.K., Wright A.F., Tommerup N., Rosenberg T., Tornqvist K. (1997) Intrafamilial variation of the phenotype in Bardet-Biedl syndrome. *British Journal of Ophthalmology* 82: 378-385.

Riise R., Tornqvist K, Wright A.F., Mykytyn K., Sheffield V.C. (2002) The Phenotype in Norwegian Patients With Bardet-Biedl Syndrome With Mutations in the BBS4 Gene. *Ophthalmic Molecular Genetics* 120, 1364-1367.

Rosenbaum J.L., Witman G.B. (2002) Intraflagellar Transport. *Molecular Cell Biology* 3, 813-825.

Schultz J., Milpetz F., Bork P., Ponting C.P. (1998) SMART, a simple modular architecture research tool: Identification of signaling domains. *PNAS* 95, 5857-5864.

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

Short B., Preisinger C., Korner R., Kopajtich R., Byron O., Barr F.A. (2001) The GRASP55-rab2 effector complex linking Golgi structure to membrane traffic. *Journal of Cell Biology* 155, 877-883.

Swoboda P., Adler H.T., Thomas J.H. (2000) The RFX-Type Transcription Factor DAF-19 Regulates Sensory Neuron Cilium Formation in *C. elegans*. *Molecular Cell* 5, 411-421.

Voet D., Voet J.G., Pratt C.W. Fundamentals of Biochemistry. New York: John Wiley & Sons Inc., 1998.

Young I.D., Moore J.R. (1987) Intrafamilial variation in Cohen syndrome. *Journal of Medical Genetics* 24, 488-492.

Young T.L., Penney L., Woods M.O., Parfrey P.S., Green J.S., Hefferton D., Davidson W.S. (1999) A Fifth Locus for Bardet-Biedl Syndrome Maps to Chromosome 2q31. *American Journal of Human Genetics* 64, 900-904.

Young, T., Woods M.O., Parfrey P.S., Green J.S., Hefferton D., Davidson W.S. (1999) A Founder Effect in the Newfoundland Population Reduces the Bardet-Biedl Syndrome 1 (BBS1) Interval to 1 cM. *American Journal of Human Genetics* 65, 1680-1687.

Young T., Woods M.O., Parfrey P.S., Green J.S., O'Leary E., Hefferton D., Davidson W.S. (1998). Canadian Bardet-Biedl Syndrome Family Reduces the Critical Region of BBS3(3p) and Presents With a Variable Phenotype. *American Journal of Medical Genetics* 78, 461-467.

Woods M.O., Young T., Parfrey P.S., Hefferton D., Green J.S., Davidson W.S. (1999) Genetic Heterogeneity of Bardet-Biedl Syndrome in a Distinct Canadian Population: Evidence for a Fifth Locus. *Genomics* 55, 2-9.