npg

ARTICLE

Pitfalls of homozygosity mapping: an extended consanguineous Bardet–Biedl syndrome family with two mutant genes (*BBS2*, *BBS*10), three mutations, but no triallelism

Virginie Laurier[1,2,8], Corinne Stoetzel[1,8], Jean Muller[3,4], Christelle Thibault[3], Sandra Corbani[5], Nadine Jalkh[5], Nabiha Salem[5], Eliane Chouery[5], Olivier Poch[3], Serge Licaire[3], Jean-Marc Danse[1], Patricia Amati-Bonneau[6], Dominique Bonneau[6], André Mégarbané[5], Jean-Louis Mandel[3,7] and Hélène Dollfus*[1,2]

[1]Laboratoire de Génétique Médicale EA 3949, Faculté de Médecine de Strasbourg, Université Louis Pasteur, Strasbourg, France; [2]Service de Génétique Médicale et Centre de Référence pour les Affections Génétiques Ophtalmologiques, Hôpitaux Universitaires de Strasbourg, Strasbourg, France; [3]Institut de Génétique et de Biologie Moléculaire et Cellulaire, CNRS/INSERM/ULP Illkirch Cedex, CU de Strasbourg, France; [4]Laboratoire de Biologie Moléculaire, d'Analyse Génique et de Modélisation, Centre de Recherche Public-Santé, Luxembourg, Luxembourg; [5]Unité de Génétique Médicale, Université Saint Joseph Faculté de Médecine, Beirut, Lebanon; [6]Service de Génétique, CHU, Angers, France; [7]Collège de France, Paris, France

The extensive genetic heterogeneity of Bardet–Biedl syndrome (*BBS*) is documented by the identification, by classical linkage analysis complemented recently by comparative genomic approaches, of nine genes (*BBS1–9*) that account cumulatively for about 50% of patients. The *BBS* genes appear implicated in cilia and basal body assembly or function. In order to find new *BBS* genes, we performed SNP homozygosity mapping analysis in an extended consanguineous family living in a small Lebanese village. This uncovered an unexpectedly complex pattern of mutations, and led us to identify a novel *BBS* gene (*BBS10*). In one sibship of the pedigree, a *BBS2* homozygous mutation was identified, while in three other sibships, a homozygous missense mutation was identified in a gene encoding a vertebrate-specific chaperonine-like protein (*BBS10*). The single patient in the last sibship was a compound heterozygote for the above *BBS10* mutation and another one in the same gene. Although triallelism (three deleterious alleles in the same patient) has been described in some *BBS* families, we have to date no evidence that this is the case in the present family. The analysis of this family challenged linkage analysis based on the expectation of a single locus and mutation. The very high informativeness of SNP arrays was instrumental in elucidating this case, which illustrates possible pitfalls of homozygosity mapping in extended families, and that can be explained by the rather high prevalence of heterozygous carriers of *BBS* mutations (estimated at one in 50 in Europeans).
European Journal of Human Genetics (2006) 14, 1195–1203. doi:10.1038/sj.ejhg.5201688; published online 5 July 2006

*Correspondence: Professor H Dollfus, Laboratoire de Génétique Médicale, EA 3949, Faculté de Médecine, Université Louis Pasteur, Hôpitaux Universitaires de Strasbourg, 11 rue Humann, 67085 Strasbourg, France.
Tel: +33 3 90 24 35 62 or + 33 3 88 12 81 20; Fax: + 33 3 88 12 81 25;
E-mail: helene.dollfus@medecine.u-strasbg.fr
[8]These authors contributed equally to this work.

Introduction

Bardet–Biedl syndrome (*BBS* (MIM 209900)) is characterized by early-onset retinitis pigmentosa, childhood-onset obesity, polydactyly, hypogonadism, cognitive impairment and kidney dysplasia.[1,2] Positional cloning studies

complemented recently by comparative genomic approaches have proven the extensive genetic heterogeneity of this autosomal-recessive condition. Nine genes have been identified before 2006 BBS1 (11q13);[3] BBS2 (16q21);[4] BBS3/ARL6 (3p12–13);[5,6] BBS4 (15q22.3–q23);[7] BBS5 (2q31);[8] BBS6/MKKS (20 p12);[9,10] BBS7 (4q27);[11] BBS8/TTC8(14q32.11)[12] and BBS9 /PTHB1 (7p14).[13] Mutations in these genes account only for about 50% of BBS patients, suggesting that there are several more genes to be found for this condition. To add to the genetic complexity, three mutated alleles in two genes have been identified in patients[14–18] in whom the third allele appears to modify penetrance or severity of the clinical phenotype.

Studies of other sets of families have however failed to document unambiguous triallelic inheritance, suggesting that incomplete penetrance is rare in individuals carrying two bona fide mutations in a single BBS gene.[19–22]

The identification of BBS8 pointed to a defect in the assembly or function of cilia or basal bodies.[12] This was confirmed for the other BBS proteins that appear to be also implicated in the important developmental process of planar cell polarity.[23]

For diagnostic and genetic counselling purposes, it is of major importance to identify the remaining BBS genes. This would also allow the testing in an unbiased way of the incidence and clinical consequences of triallelism, by searching systematically in patients for mutations in all the BBS genes, even if two deleterious mutations in a single gene have been already detected.[15,17,18,21] Finally, identification of novel genes will help in the understanding of the biology of cilia and basal body.

We are performing linkage analysis in multiplex and/or consanguineous BBS families that do not present mutations in the known genes. The family that appeared a priori the most informative was an extended consanguineous Lebanese family that however failed to show linkage on a whole-genome microsatellite scan. Further study using SNP microarray (Affymetrix Genechip 10K) analysis revealed

the involvement of two genes in the family: BBS2 (chromosome 16q) for one sibship, and for the rest of the family, a novel gene localized on chromosome 12q21.1 encoding a vertebrate-specific chaperonin-like protein.[24] Instead of the expected homozygous mutation in a single gene, we thus found in this family three mutations in two BBS genes, but no evidence of triallelism.

Materials and methods
Patients
Members of family III.8 are Lebanese Arabs from the Muslim Sunni community and originate from a village in the North of Lebanon in which all members live to date. The family displays extensive consanguinity (Figure 1). Seven affected individual were examined by two of us (HD, AM) (Table 1). DNA was available for nine patients as well as for some parents and non-affected siblings. Signed informed consents specific for the genetic study were obtained for all patients. One hundred and seven DNAs from Lebanese individuals and 96 DNA from French individuals were used as controls.

Methods
Whole-genome scan microsatellite analysis
A whole-genome scan was performed at Centre National de Genotypage (Evry, France) using a microsatellite marker set comprising 400 fluorescent-labelled microsatellite markers with an average spacing of 10 centi-morgans (cM) and an average heterozygosity of 75%. At the time of analysis, DNA was available for eight patients and three unaffected sibs (sibhips A–D), but not from parents.

SNP homozygosity mapping
Family III.8 was then studied with the Affymetrix Gene-Chip® Mapping 10K Array (Affymetrix, Santa Clara, CA, USA) on eight patients, one parent and four unaffected sibs
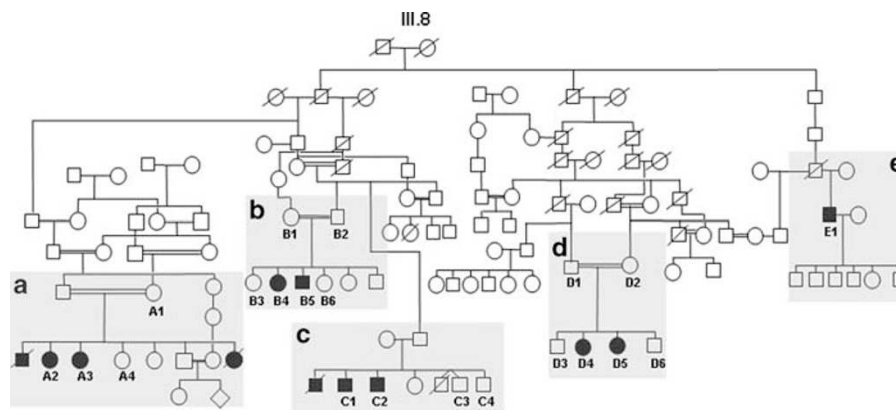
**Figure 1** Pedigree of family III.8.

(Figure 2). Sample processing and labelling were performed according to the manufacturer's instructions (Affymetrix Mapping 10K 2.0 Assay Manual, Version 1.0, 2004). The arrays were hybridized on a GeneChip Hybridization Oven 640, washed with the GeneChip Fluidics Station 450 and scanned with a GeneChip Scanner 3000. Data were processed by the GeneChip DNA Analysis Software version 3.0. 2 (GDAS) to generate SNP allele calls. An average call rate greater than 99% was obtained. Homozygosity regions were identified as regions of homozygosity longer than 25 adjacent SNPs. Affymetrix genechip SNP arrays allow analysis of 10 000 SNPS with a mean genetic gap distance of 0.32 cM and an average heterozygosity of 0.37.

Table 1 Clinical manifestations in members of family III.8

Patient	Sex	Age	Retinitis pigmentosa	BMI	Polydactyly	Cognitive impairment	Renal involvement
A2	F	50	+	30.8	Appendage right hand	+/−	Unknown
A3	F	47	+	28	Hands feet	+/−	+
B4	F	22	+	32.6	Left foot	+/−	Unknown
B5	M	33	+	46.18	Hands feet	+/−	Unknown
D4	F	23	+	55	Hands feet	+/−	−
D5	F	18	+	50.21	Hands feet	+/−	−
E1	M	70	+	32.9	Hands feet	−	−

BMI = body mass index.
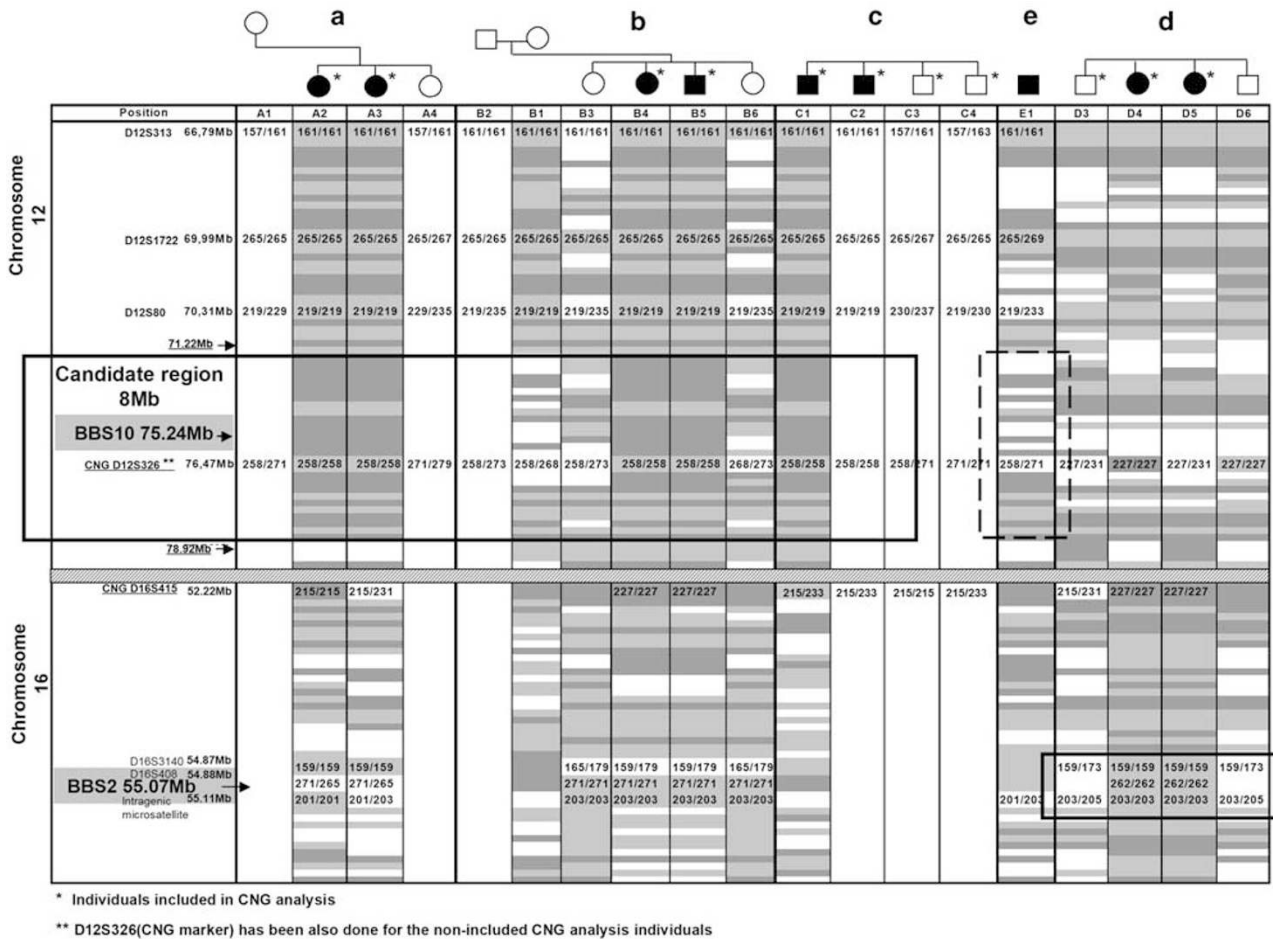
Figure 2 SNP, microsatellite segregation and homozygosity mapping defines two chromosomal regions of interest. Light and dark grey shading represent homozygous SNPS (respectively AA and BB) while white regions indicate heterozygous alleles (AB).

DNA analysis with microsatellite markers

Genotyping of additional fluorescent microsatellite markers (Figure 2) was performed on a CEQ8800 genetic analysis system (Beckman Coulter). Experimental conditions are available on request. Microsatellite sequences were obtained from the UCSC Genome Browser Bioinformatics site (http://genome.ucsc.edu/cgi-bin/hgGateway).

DNA sequencing and mutation screening

PCR amplifications were performed with 50 ng of genomic DNA template. Bidirectional sequencing of the purified PCR products was carried out using the ABI Big Dye Terminator 3.1 Sequencing kit on an ABI 3130xl automated capillary sequencer (Applied Biosystems). Detailed protocols are available on request.

Splice sites scoring programs (http://l25.itba.mi.cnr.it/~webgene/wwwspliceview.html, http://www.fruitfly.org/seq_tools/splice.html) and exonic splicing enhancer predictor program rescue ESE web server (http://genes.mit.edu/burgelab/rescue-ese) were used to evaluate the effect of mutations on splicing.

Results

SNP homozygosity mapping in family III.8

Affected members of family III.8 belong to five sibships related by complex consanguinity loops (Figure 1). Three sibships (A, B, D) correspond to first cousin unions with added more distant consanguinity (inbreeding coefficient is thus >1/16). Consanguinity was not documented for sibships C and E (the latter with a single affected 70-year-old patient). The standard microsatellite linkage scan performed at CNG (Evry) (400 markers with an average spacing of 10 cM) on the four multiplex sibships (A–D) did not allow us to find a common region of linkage and/or significant homozygosity, even considering only the three sibships with documented consanguinity (A, B, D) (not shown). Analysis of the same family by a different genotyping center gave the same negative result (P Beales and N Katsanis, personal communication). This could have been owing to trivial reasons (sample mix-up or diagnostic errors, for instance), and thus the family was resampled and diagnosis was rechecked on site by two clinical geneticists (HD, AM). We then analysed DNAs from available affected members with the Affymetrix gene-chip SNP arrays that allow analysis of 10 000 SNPs with a mean genetic gap distance of 0.32 cM and an average heterozygosity of 0.37. This is equivalent to an about eightfold increase in informativeness and marker density compared to a 400 microsatellites scan. Again, no common candidate region of homozygosity could be documented.

We then analysed separately the most informative sibships (A–D). In sibship D, a region, defined by 81 successive SNPs (from 26.35 to 60.45 Mb on chromosome 16q), was homozygous by state in the two affected sibs, but not in the unaffected ones (Figure 2). The region included *BBS2* (16q21) (lod score $\leqslant 2.05$, the latter value is calculated assuming an inbreeding coefficient of 1/16; however, inbreeding is in reality higher owing to the multiple consanguinity loops, and the true lod score is thus lower).[25]

A second region of interest was identified for sibships A, B and C. The overlapping region of homozygosity spanned from 65.64 to 78.92 Mb on chromosome 12. Further analysis, either using the SNP array for some parents or unaffected sibs, or by microsatellite analysis, including the genome scan data and additional markers allowed us to reduce the region of interest to an interval of 7.7 Mb ranging from 71.22 to 78.92 Mb. (Figure 2).

No significant overlapping region of homozygosity with either sibship D or sibships A–C was observed for patient E1, although the data were consistent with heterozygosity for the haplotype found in the homozygous state in sibships A–C.

Sequencing of *BBS2*

The *BBS2* gene was sequenced in all the patients of the III.8 family. A homozygous missense variant G139V was identified in the two affected patients of sibship D. This change affects the first G of a four-amino-acid sequence (GGNC) completely conserved in vertebrates and the nematode *Caenorhabditis elegans* (Figure 3), and is also predicted to abolish a putative exonic splicing enhancer (an SRP40 motif with a score of 3.38, according to the ESE web server). The effect on splicing remains, however, to be tested. The variant was not found in any other patients of the remaining sibships of the family even in the heterozygous state, and has not been reported previously (http://www.hgmd.cf.ac.uk/ac/search.html). This change was not found in 48 sequenced Lebanese controls. We conclude that the G139V change is a pathogenic mutation.

Identifying FLJ23560 gene as *BBS10*

The 8 Mb region of chromosome 12 was investigated for potential candidate genes. Twenty-three genes were listed in the ENSEMBL database for this interval, of which 18 are well characterized and five are not characterized. Eight genes were selected to be screened as a priority either because they belong to a cilia–proteome database[6,8] or because of the putative functional relevance. Four genes in the interval of interest were previously reported as potentially cilia-related by comparative genomic analysis of ciliated/non-ciliated organisms: *RAB21* (Ras-related protein Rab-21, 70.43 Mb), *TPH2* (tryptophan 5-monooxygenase *2*, 70.26 Mb); *TBC1D15* (TBC1 domain family member 15, 70.51 Mb) and *HRB2* (HIV-1 Rev binding protein 2, 74.18 Mb).[6,8]

Four additional genes of the interval were considered as priority candidate genes because of their function: *SYT1* (synaptotagmin-1, 77.76 Mb); *ZDHHC17* (Huntingtin-
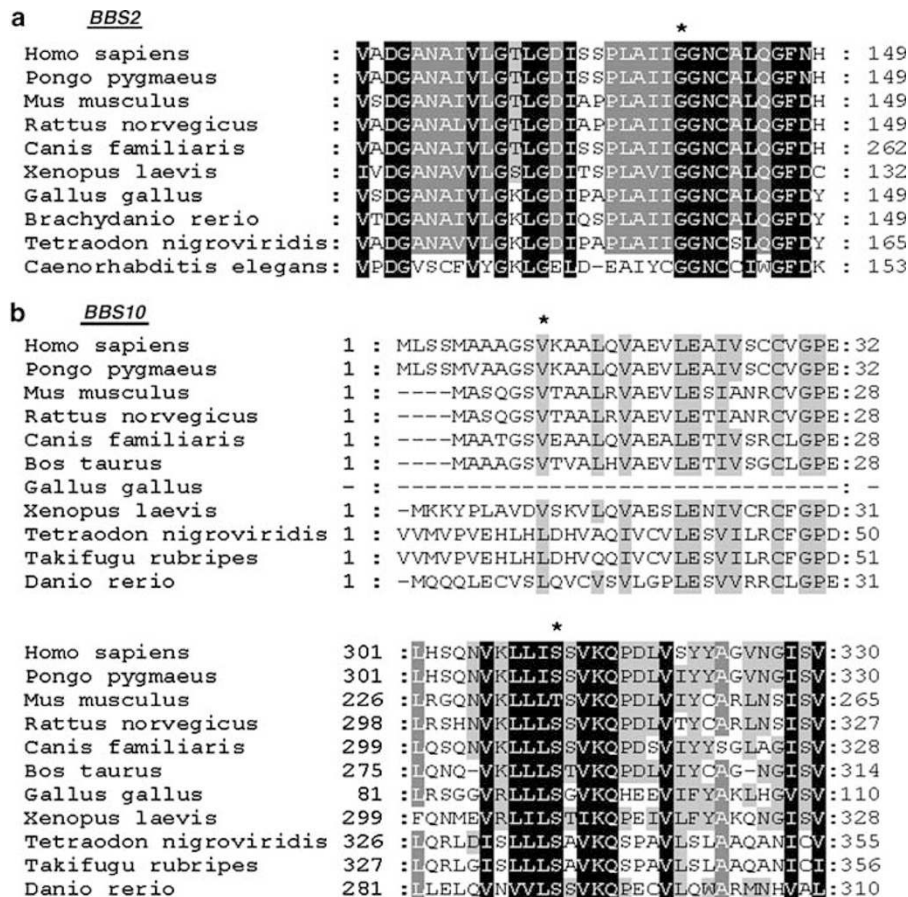
Figure 3 Amino-acid sequence conservation around residues affected by missense changes at *BBS*2 and *BBS*10 loci, identified in family III.8. The sequences of BBS proteins/or predicted translation products from several species have been compared and aligned. The *BBS*2 (G139V) and *BBS*10 (V11G and S311A) mutations are noticed with a black star. (**a**) *BBS*2 sequences are from Uniprot (*Homo sapiens* bbs2_human; *Pleioblastus pygmaeus* q5r9u3_ponpy; *Mus musculus* bbs2_mouse; *Rattus norvegicus* bbs2_rat; *Gallus gallus* q5zi17_chick; *Danio rerio* bbs2_brare; *Tetraodon nigroviridis* q4rgw3_tetng) and NCBI (*C. familiaris* xp_535296 ; *C. elegans* np_501325). Genome prediction has been performed only for *Xenopus laevis* using ensembl database. (**b**) *BBS*10 sequences are from Uniprot (*H. sapiens* q8tam1_human; *P. pygmaeus* q5r8p3_ponpy; *M. musculus* q9dbi2_mouse; *X. laevis* q5fwq1_xenla; *T. nigroviridis* q4rew1_tetng). Genome predictions have been performed using ensembl database for *R. norvegicus, C. familiaris, G. gallus, D. rerio* and *T. rubripes* and N-terminal correction for *H. sapiens* and *M. musculus*.

interacting protein 14; 75.66); *Loc* 387869 (similar to microtubule, 73.34 Mb) and NP_078961.2 (FLJ 23560, described as 'ATP binding; chaperone activity' in NCBI-Contig NT_086796 database).

No mutations were detected for SYT1, ZDHHC17 and Loc 387869. The fourth tested gene, FLJ23560, consists of two exons and spans 3.7 kb. Although only exon 2 was annotated as protein coding in genome databases, a bioinformatic analysis indicated that the coding sequence starts in exon 1.[24] A homozygous missense change (S311A) was found in exon 2, in all the affected members of sibships A, B and C. Affected individual E1 was heterozygous for this mutation. Further sequencing of FLJ23560 revealed another missense variant V11G in exon 1 of patient E1, indicating that he is a compound heterozygote. Affected individuals of sibship D did not show any mutation in the FLJ 23560 gene.

The S311A and V11G variants were not found in 107 Lebanese controls or in 96 French controls. The S311 residue belongs to a four-amino-acid motif (LLIS) highly conserved in vertebrates (the serine residue is replaced in mouse by a threonine, also an hydroxylated amino acid) (Figure 3). The region around V11 is less conserved, but the V itself is conserved down to *Xenopus*, and replaced by another hydrophobic amino acid (leucine) in fishes (Figure 3). Furthermore, the V11G change might also affect splicing as it abolishes a potential SC35 exonic splice enhancer (score 2.66). The pathogenic nature of these variants could however be disputed, as they correspond to missense mutations affecting a gene that is not well conserved in evolution, as opposed to eight of the nine other *BBS* genes[24] and induce significant, but not drastic chemical changes (polar/nonpolar for S311A, hydrophobic/ nonhydrophobic for V11G). Further study of the FLJ 23560

gene identified clearly deleterious mutations in many *BBS* families (although none with the S311A or V11G mutations) confirming its implication in *BBS*.[24] The limited segregation data in the family suggested typical recessive inheritance.

Detailed clinical information was only available for seven patients (five with *BBS*10, two with *BBS*2 mutations), and we detected no major *BBS* gene-related phenotypic differences in this very small set of patients, although one can note that the two patients with the highest body mass index carry the *BBS2* mutation (but this might be also influenced by social factors).

Discussion

Homozygosity mapping was proposed in 1987 as an efficient strategy for locating the genes implicated in rare recessive diseases, when families available for linkage analysis are limited in number and size.[26] Implementation of this strategy awaited the construction of the first whole-genome microsatellite map[27] and the first successes were reported in 1993.[28–30] This approach was used extensively as, by scanning the genome with microsatellite panels. Very recently, it was shown that SNP microarrays provide a faster and much more informative technique.[31] Homozygosity mapping is particularly essential in diseases with extensive nonallelic heterogeneity, as a single large consanguineous family may be sufficient to locate the gene, with an lod score >3.0, whereas smaller ones can document heterogeneity if they do not show homozygosity for the same region. *BBS* was in fact the first case where linkage and heterogeneity could be demonstrated at the same time by this strategy.[30] Indeed, it would have been impossible to disentangle the extreme genetic complexity of *BBS* without such an approach, applied to families from the Middle East, Newfoundland, Turkey or Puerto-Rico.[3,30,32–35] However, when a single or very few families are used, the candidate region identified by homozygosity mapping is in general very large and contains many genes. This led, in the case of *BBS*, to use in addition comparative genomic approaches. When it became clear that *BBS* genes code for proteins implicated in cilia assembly or function,[12] comparison of sequenced genomes allowed selection of candidate genes that have orthologs in ciliated organisms (vertebrates, *drosophila*, *C. elegans*, chlamydia, trypanozoma) but not in nonciliated ones (*Arabidopsis*, yeast) or in *Giardia lamblia*, an organism that does not contain well-conserved orthologs to known *BBS* genes.[13] This combination of homozygosity mapping and comparative genomics allowed the identification of *BBS3*, *BBS5* and *BBS9*.[5,6,8,13]

Single sibship consanguineous families that have enough informativeness to reach an lod score of 3 are quite rare, even in countries with high consanguinity and large sibship size (one requires, in a first cousin marriage at least 3 affected sibs, or 2 affected and $\geqslant 5$ nonaffected). More distant consanguinity is potentially more informative, but the risk is then to miss the relevant homozygous region, if the genome scan is performed at the usual density of one marker every 10 cM. More distant consanguinity also increases the risk, especially if the disease is not very rare, and that a second mutation is introduced by a married in ancestor, on a different haplotype. To overcome limited informativeness, extended families related by multiple consanguinity loops originating from inbred populations (like family III.8 studied herein) are thus often used. The success of homozygosity mapping relies on the assumption that, for a rare recessive disease, patients in a consanguineous family will be homozygous for a mutation derived from a common ancestor of the parents. Our results for family III.8 present an outstanding departure from this expectation. Initial attempts to map the affected gene by microsatellite scan (performed independently by two laboratories) failed. Retrospectively, analysing sibships separately, indications for linkage to chromosome 12 were rather good for sibship A (four contiguous homozygous microsatellites) and even better for sibship B (nine contiguous homozygous microsatellites) (Supplementary Figure 1), corresponding to lod scores of <1.8 for each sibship (1.8 would correspond to an inbreeding coefficient of 1/16), but as these pedigrees contain additional consanguinity loops, the increased inbreeding results in decreased maximum lod score.[25] Parents of sibship C were not known to be consanguineous, and the two patients shared homozygosity for only one microsatellite from the initial genome scan on chromosome 12. The expected lod score (elod) in this sibship would be only 0.85, in the absence of documented consanguinity. For sibship D, the maximum elod for the initial genome scan was <1.92. In this family, only two adjacent microsatellites were homozygous around the *BBS*2 locus in the genome scan (Supplementary Figure 1), insufficient to prove homozygosity by descent, so that the observed lod score was 1.57 (also an overestimate given the additional consanguinity loop). One can notice that sibships A, B and C are more closely related to each other (four generations separating them from the ancestral couple in generation II), than to sibships D or E, with whom they share a common ancestral couple in generation I. The much greater informativeness of the 10K SNP microarray (increased by a factor of about 8) allowed us to more clearly define the significant regions of homozygosity, and revealed that sibship C shows consanguinity (but more distant than first cousin), as patients present five homozygous tracts of 10–22 Mb (compared to tracts of up to 60–70 Mb in the documented first cousin sibships A, B and D). In an outbred population, true homozygosity for a region of more than 5 Mb occurs only once in 35 individuals,[36] and the maximum length of homozygosity by state expected by chance in a

non-consanguineous family, using a 10K SNP array is about 27 SNPs ≈ 8 Mb. Individual E1 presents a similar level of consanguinity, as detected by the SNP microarray, with three tracts of homozygosity of 8–18 Mb, suggesting an inbreeding coefficient ≤ 1/75. Based on the SNP array analysis, we could then identify *BBS2* as mutated in sibship D, and identify another gene (FLJ23560 or *BBS10*) as mutated at the homozygous state in sibships A, B and C and at the heterozygous state in affected individual E1 (despite the indication that he also had distantly related parents, second cousins or more).

The findings of two mutant genes and three mutations in family III.8 were unexpected and indicate some pitfalls of homozygosity mapping. Two previous publications had pointed out some of these pitfalls. Investigating for S-cone syndrome gene, Miano *et al*,[25] found a false-positive linkage of 3.69 and they ascribed it to a combination of type 1 error (expected rate of false positive of one in 50), and more importantly to the inflated lod score resulting from underestimating the true inbreeding coefficient, by not taking into account additional more ancient consanguinity loops in a highly inbred population. They also reported in an extended family with three affected sibships, the presence in one of them of compound heterozygosity. There too, as in family III8, while the two homozygous sibships shared common ancestors distant by only three generations, the third compound heterozygous patient was more distantly related to the common ancestor of all three couples (five generations), increasing the probability of introducing another mutation in the family. A similar finding of homozygous and compound heterozygous sibships was reported in an extended inbred Amish family with congenital hypothyroidism owing to mutations in the thyroid peroxidase gene.[37] There the common ancestral couple in this case was even further removed (7–8 generations).

The finding of an extensive heterogeneity of calpain 3 mutations in the isolated population of the Reunion island (the so-called Reunion paradox) is only superficially similar to these observations.[38,39] LGMD2A that results from calpain 3 mutations is a much more frequent disease than each *BBS* subtype, increasing the probability of multiple mutations in a founder population, and the admixture of different ethnic groups in this island favours allelic heterogeneity. The putative founder couple was also much more distant (13 generations) than for family III8, and such genealogic reconstructions are often biased.

The finding of two genes segregating in different branches of the family, but corresponding to the same clinical phenotype, has not been reported previously to our knowledge, and thus appears even more unexpected, especially for a disease as rare as *BBS*: but is this indeed so unexpected?

Prevalence of *BBS* has been estimated at 1/125 000 to 1/160 000 in European outbred populations.[40,41] It appears,

however, much more frequent in inbred populations (1/13 500 among Bedouin of Koweit, and 1/18 000 in Newfoundland).[42,43] This could be owing to the combination of inbreeding, which raises the frequency of very rare recessive diseases, and possible founder effects.

Given that *BBS* is characterized by extreme nonallelic heterogeneity, each individual *BBS* gene will be responsible for a much smaller prevalence of corresponding patients, and intuitively one would not expect to find two mutated genes in a given extended family from a very homogeneous background. However, a simple calculation indicates that this is not true. In an outbred population, assuming a 1/150 000 prevalence accounted for by two genes representing each 20% of patients (*BBS*1 and *BBS*10)[3,17,24] and 15 genes each accounting for 4% (the other eight known genes fit well with this figure,[18] and the data of Nishimura *et al*[13] suggest that at least five other genes remain to be identified), one can deduce that the cumulative carrier frequency of *BBS* mutations is one in 50, more than half of the frequency of CFTR mutation carriers! Although in inbred populations the relative contribution of different *BBS* genes may vary, given the number of genes involved, the final outcome should be similar (if there is no heterozygote selection). Thus, it is not so surprising that in an extended family, mutations in two different *BBS* genes might be segregating, and that the consanguinity may result in bringing them both to homozygosity. This also suggests that unlike recessive diseases that are owing to mutations in a single gene, where the presence of a founder mutation in one, but not in another population will result in wide disparity in disease prevalence, in the case of the numerous *BBS* genes, we predict that cumulative *BBS* prevalence should be rather high in all populations with a high rate of consanguinity. This estimation of the frequency of *BBS* carriers in the general population raises the question of the possible impact of *BBS* mutations on risk of some common diseases corresponding to clinical manifestations of *BBS* (notably obesity) or possible interaction with mutations associated with retinitis pigmentosa or kidney dysplasia (for instance, nephronophthisis or polycystic kidney disease). Only two association studies of the *BBS1* common mutation (M390R) and of missense SNPs in *BBS6* have been performed in relation to obesity, which suggested that heterozygosity for the tested variants does not contribute significantly to the risk of obesity in the population (with the possible exception of the rare *BBS6* variant A242S).[44,45]

The cumulative prevalence of *BBS* mutation carriers indicates that the finding of a third mutated allele can occur by chance in 2% of *BBS* patients (or 1.5% in the present state of incomplete identification of *BBS* genes).

Given that a majority of the third alleles suggesting triallelism reported previously correspond to missense changes of uncertain or even in some cases dubious pathogenic significance (for instance, the A242S *BBS6*

variant),[14,45] one should compare the cumulative frequency of truly pathogenic third allele to this chance expectation: a higher frequency is expected if indeed a significant number of individuals with two pathogenic mutations at a single locus do not express the disease, or in such attenuated or incomplete form that they will not be diagnosed as BBS.

Although a few cases of nonpenetrance in patients with two *bona fide* mutations in a single *BBS* gene have been described,[14,17] in particular for the *BBS*1 M390R missense (a missense that may in fact be a milder mutation that truncating ones), the recent study by Stoetzel *et al*[24] found only three cases with three convincing mutations out of 65 families with one or two *BBS10* mutations (two of these cases implicated the *BBS*1 M390R mutation). This low rate (close to the chance expectation) explains, whereas some studies failed to document unambiguous triallelism.[19–22] Third alleles may, however, act as modifiers of the severity, as reported by Badano *et al*.[15,46]

Another unexpected finding from the analysis of family III.8 was the nonconservation of the *BBS10* genes in ciliated organisms, apart from vertebrates.[24] This gene, like *BBS6*, was thus not included in the lists of candidates obtained by comparative genomic approaches.[5,6,8,13,47] One can wonder why *BBS10*, which is with *BBS1* the most frequently mutated *BBS* gene, at least in patients of European origin, escaped positional cloning until now, whereas genes that are mutated in a much smaller proportion of patients have been identified. The absence of *BBS10* on lists of candidate genes might explain why it was not detected by Nishimura *et al*.[13] who had two small consanguineous families consistent with the chromosome 12 localization of *BBS10*. Also as the number of large enough families suitable for initial mapping (ie that would singly yield an lod score >3) is small. *BBS*10 may not have been represented in the pool of very informative families reported previously. Furthermore, the high frequency of *BBS10* mutations is due in part to a major mutation (C91fsX95),[24] which may be more prevalent in Europeans than in other populations (notably Middle East) from which most of the informative families were derived. As consanguinity and sibship size are in general low in Europeans, this decreases the probability of an informative family in such population (apart from Amish and Newfoundland families). Finally, the use of extended families might have generated in some instances similar problems to those encountered in family III.8, as we now see that these were not so unexpected, preventing successful mapping. Nishimura *et al*[13] have convincingly shown that with a 10K SNP array, smaller families can be analysed to suggest candidate regions (albeit large ones) that will include false positives (as each family may be consistent with a few candidate regions), but that should help in finding new genes (especially those included in the comparative or functional genomics lists of candidates).

As *BBS10* accounts for about 30–40% of previously unassigned families, its identification will restrict the number of such candidate regions suggested by analysis of small consanguineous families, and should facilitate the identification of additional *BBS* genes.

Note added in proof

A report of a 11th BBS,[48] describing a homozygous missense mutation in one family, appeared while this article was going to press.

References

1 Green JS, Parfrey PS, Harnett JD *et al*: The cardinal manifestations of Bardet–Biedl syndrome, a form of Laurence–Moon–Biedl. *N Engl J Med* 1989; **321**: 1002–1009.

2 Beales PL, Elcioglu N, Woolf AS, Parker D, Flinter FA: New criteria for improved diagnosis of Bardet–Biedl syndrome : results of a population survey. *J Med Genet* 1999; **36**: 437–446.

3 Mykytyn K, Nishimura DY, Searby CC *et al*: Identification of the gene (*BBS*1) most commonly involved in Bardet–Bield syndrome, a complex human obesity syndrome. *Nat Genet* 2002; **31**: 435–438.

4 Nishimura DY, Searby CC, Carmi R *et al*: Positional cloning of a novel gene on chromosome 16q causing Bardet–Biedl syndrome. *Hum Mol Genet* 2001; **10**: 865–874.

5 Fan Y, Esmail MA, Ansley SJ *et al*: Mutations in a member of the Ras superfamily of small GTP-binding proteins causes Bardet–Biedl syndrome. *Nat Genet* 2004; **36**: 989–993.

6 Chiang AP, Nishimura D, Searby C *et al*: Comparative genomic analysis identifies an ADP-ribosylation factor-like gene as the cause of Bardet–Biedl syndrome (*BBS*3). *Am J Hum Genet* 2004; **75**: 475–484.

7 Mykytyn K, Braun T, Carmi R *et al*: Identification of the gene that, when mutated, causes the human obesity syndrome BBS4. *Nat Genet* 2001; **28**: 188–191.

8 Li JB, Gerdes JM, Haycraft CJ *et al*: Comparative genomics identifies a flagellar and basal body proteome that includes the *BBS*5 human disease gene. *Cell* 2004; **117**: 541–552.

9 Slavotinek AM, Stone EM, Mykytyn K *et al*: Mutations in MKKS cause Bardet–Biedl syndrome. *Nat Genet* 2000; **26**: 15–16.

10 Katsanis N, Beales PL, Woods MO *et al*: Mutations in MKKS cause obesity, retinal dystrophy and renal malformations associated with Bardet–Biedl syndrome. *Nat Genet* 2000; **26**: 67–70.

11 Badano JL, Ansley SJ, Leitch CC, Lewis RA, Lupski JR, Katsanis N: Identification of a novel Bardet–Biedl syndrome protein, *BBS7*, that shares structural features with *BBS*1 and *BBS*2. *Am J Hum Genet* 2003; **72**: 650–658.

12 Ansley SJ, Badano JL, Blacque OE *et al*: Basal body dysfunction is a likely cause of pleiotropic Bardet–Biedl syndrome. *Nature* 2003; **425**: 628–633.

13 Nishimura DY, Swiderski RE, Searby CC *et al*: Comparative genomics and gene expression analysis identifies *BBS9*, a new Bardet–Biedl syndrome gene. *Am J Hum genet* 2005; **77**: 1021–1033.

14 Katsanis N, Ansley SJ, Badano JL *et al*: Triallelic inheritance in Bardet–Biedl syndrome, a Mendelian recessive disorder. *Science* 2001; **293**: 2256–2259.

15 Badano JL, Kim JC, Hoskins BE *et al*: Heterozygous mutations in *BBS*1, *BBS*2 and *BBS*6 have a potential epistatic effect on Bardet–Biedl patients with two mutations at a second *BBS* locus. *Hum Mol Genet* 2003; **12**: 1651–1659.

16 Fauser S, Munz M, Besch D: Further support for digenic inheritance in Bardet–Biedl syndrome. *J Med Genet* 2003; **40**: e104.

17 Beales PL, Badano JL, Ross AJ *et al*: Genetic interaction of *BBS*1 mutations with alleles at other *BBS* loci can result in non-mendelian Bardet–Biedl syndrome. *Am J Hum Genet* 2003; **72**: 1187–1199.

18 Katsanis N: The oligogenic properties of Bardet–Biedl syndrome. *Hum Mol genet* 2004; Spec. No. **1**: R65–R71.

19 Mykytyn K, Nishimura DY, Searby CC *et al*: Evaluation of complex inheritance involving the most common Bardet–Biedl syndrome locus (*BBS*1). *Am J Hum Genet* 2003; **72**: 429–437, (E-pub ahead of print Jan 10 2003).

20 Moore SJ, Green JS, Fan Y *et al*: Clinical and genetic epidemiology of Bardet–Biedl syndrome in Newfoundland: a 22-year prospective, population-based, cohort study. *Am J Med Genet A* 2005; **132**: 352–360.

21 Hichri H, Stoetzel C, Laurier V *et al*: Testing for triallelism: analysis of six *BBS* genes in a Bardet–Biedl syndrome family cohort. *Eur J Hum Genet* 2005; **13**: 607–616.

22 Nakane T, Biesecker LG: No evidence for triallelic inheritance of MKKS/BBS loci in Amish Mckusick–Kaufman syndrome. *Am J Med Genet A* 2005; **138**: 32–34.

23 Ross AJ, May-Simera H, Eichers ER *et al*: Disruption of Bardet–Biedl syndrome ciliary proteins perturbs planar cell polarity in vertebrates. *Nat Genet* 2005; **37**: 1135–1140.

24 Stoetzel C, Laurier V, Davis E *et al*: A novel gene encoding a vertebrate-specific chaperonin-like is a major *BBS* locus. *Nat Genet* 2006; **38**: 521–524.

25 Miano MG, Jacobson SG, Carothers A *et al*: Pitfalls in homozygosity mapping. *Am J Hum Genet* 2000; **67**: 1348–1351.

26 Lander ES, Botstein D: Homozygosity mapping: a way to map human recessive traits with the DNA of inbred children. *Science* 1987; **236**: 1567–1570.

27 Weissenbach J, Gyapay G, Dib C *et al*: A second-generation linkage map of the human genome. *Nature* 1992; **359**: 794–801.

28 Ben Hamida C, Doerflinger N, Belal S *et al*: Localization of Friedreich ataxia phenotype with selective vitamin E deficiency to chromosome 8q by homozygosity mapping. *Nat Genet* 1993; **5**: 195–200.

29 Pollak MR, Chou YH, Cerda JJ, Steinmann B *et al*: Homozygosity mapping of the gene for alkaptonuria to chromosome 3q2. *Nat Genet* 1993; **5**: 201–204.

30 Kwitek-Black AE, Carmi R, Duyk GM *et al*: linkage of Bardet–Biedl syndrome to chromosome 16q and evidence for non-allelic genetic heterogeneity. *Nat Genet* 1993; **5**: 392–396.

31 Janecke AR, Thompson DA, Utermann G *et al*: Mutations in RDH12 encoding a photoreceptor cell retinol dehydrogenase cause childhood-onset severe retinal dystrophy. *Nat Genet* 2004; **36**: 850–854.

32 Leppert M, Baird L, Anderson KL, Otterud B, Lupski JR, Lewis RA: Bardet–Biedl syndrome is linked to DNA markers on chromosome 11q and is genetically heterogeneous. *Nat Genet* 1994; **7**: 108–112.

33 Sheffield VC, Carmi R, Kwitek-Black A *et al*: Identification of a Bardet–Biedl syndrome locus on chromosome 3 and evaluation of an efficient approach to homozygosity mapping. *Hum Mol Genet* 1994; **3**: 1331–1335.

34 Bruford EA, Riise R, Teague PW *et al*: Linkage mapping in 29 Bardet–Biedl syndrome families confirms loci in chromosomal regions 11q13, 15q22–3, and 16q21. *Genomics* 1997; **41**: 93–99.

35 Young T-L, Penney L, Woods MO *et al*: A fifth locus for Bardet–Biedl syndrome maps to chromosome 2q31. *Am J Hum Genet* 1999; **64**: 900–904.

36 Gibson J, Morton NE, Collins A: Extended tracts of homozygosity in outbred human populations. *Hum Mol Genet* 2006; **15**: 789–795.

37 Pannain S, Weiss RE, Jackson CE *et al*: Two different mutations in the thyroid peroxidase gene of a large inbred Amish kindred: power and limits of homozygosity mapping. *J Clin Endocrinol Metab* 1999; **84**: 1061–1071.

38 Allamand V, Broux O, Richard I *et al*: Preferential localization of the limb–girdle muscular dystrophy type 2A gene in the proximal part of a 1-cM 15q15.1–q15.3 interval. *Am J Hum Genet* 1995; **56**: 1417–1430.

39 Richard I, Broux O, Allamand V *et al*: Mutations in the proteolytic enzyme calpain 3 cause limb–girdle muscular dystrophy type 2A. *Cell* 1995; **81**: 27–40.

40 Beales PL, Warner AM, Hitman GA, Thakker R, Flinter FA: Bardet–Biedl syndrome: a molecular and phenotypic study of 18 families. *J Med Genet* 1997; **34**: 92–98.

41 Klein D, Ammann F: The syndrome of Laurence—Moon–Bardet–Biedl and allied diseases in Switzerland. Clinical, genetic and epidemiological studies. *J Neurol Sci* 1969; **9**: 479–513.

42 Farag TI, Teebi AS: High incidence of Bardet–Biedl syndrome among the Bedouin. *Clin Genet* 1989; **36**: 463–464.

43 Green JS, Harnett JD, Farid NR *et al*: The cardinal manifestations of Bardet–Biedl syndrome, a form of Laurence–Moon–Biedl syndrome. *N Engl J Med* 1989; **321**: 1002–1009.

44 Fan Y, Rahman P, Peddle L *et al*: Bardet–Biedl syndrome 1 genotype and obesity in the Newfoundland population. *Int J Obes Relat Metab Disord* 2004; **28**: 680–684.

45 Andersen KL, Echwald SM, Larsen LH *et al*: Variation of the McKusick–Kaufman gene and studies of relationships with common forms of obesity. *J Clin Endocrinol Metab* 2005; **90**: 225–230.

46 Badano JL, Leitch CC, Ansley SJ *et al*: Dissection of epistasis in oligogenic Bardet–Biedl syndrome. *Nature* 2006; **439**: 326–330.

47 Avidor-Reiss T, Maer AM, Koundakjian E *et al*: Decoding cilia function: defining specialized genes required for compartmentalized cilia biogenesis. *Cell* 2004; **117**: 527–539.

48 Chiang AP, Beck JS, Yen HJ *et al*: Homozygosity mapping with SNP arrays identifies TRIM32, an E3 ubiquitin ligase, as a Bardet-Biedl syndrome gene (BBS11). *Proc Natl Acad Sci USA* 2006; **103**: 6287–6292.

Supplementary Information accompanies the paper on European Journal of Human Genetics website (http://www.nature.com/ejhg)