



HAL
open science

Targeted high-throughput sequencing for diagnosis of genetically heterogeneous diseases: efficient mutation detection in Bardet-Biedl and Alström Syndromes

Claire Redin, Stéphanie Le Gras, Oussema Mhamdi, Véronique Geoffroy, Corinne Stoetzel, Marie-Claire Vincent, Pietro Chiurazzi, Didier Lacombe, Ines Ouertani, Florence Petit, et al.

► To cite this version:

Claire Redin, Stéphanie Le Gras, Oussema Mhamdi, Véronique Geoffroy, Corinne Stoetzel, et al.. Targeted high-throughput sequencing for diagnosis of genetically heterogeneous diseases: efficient mutation detection in Bardet-Biedl and Alström Syndromes. *Journal of Medical Genetics*, 2012, 49 (8), pp.502-512. 10.1136/jmedgenet-2012-100875 . hal-04542601

HAL Id: hal-04542601

<https://hal.science/hal-04542601>

Submitted on 11 Apr 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Targeted high-throughput sequencing for diagnosis of genetically heterogeneous diseases: efficient mutation detection in Bardet-Biedl and Alström Syndromes

Claire Redin,¹ Stéphanie Le Gras,² Oussema Mhamdi,³ Véronique Geoffroy,⁴ Corinne Stoetzel,⁵ Marie-Claire Vincent,⁶ Pietro Chiurazzi,⁷ Didier Lacombe,⁸ Ines Ouertani,³ Florence Petit,⁹ Marianne Till,¹⁰ Alain Verloes,¹¹ Bernard Jost,² Habiba Bouhamed Chaabouni,³ Helene Dollfus,^{5,12} Jean-Louis Mandel,^{1,6,13} Jean Muller^{1,6}

► Additional materials are published online only. To view these files please visit the journal online (<http://dx.doi.org/10.1136/jmedgenet-2012-100875/content/early/recent>).

For numbered affiliations see end of article.

Correspondence to

Professor Jean-Louis Mandel, Department of Neurogenetics & Translational medicine, IGBMC, 1 rue Laurent Fries, Illkirch cedex 67404, France; jlmandel@igbmc.fr

Received 7 March 2012

Revised 15 May 2012

Accepted 28 May 2012

ABSTRACT

Background Bardet-Biedl syndrome (BBS) is a pleiotropic recessive disorder that belongs to the rapidly growing family of ciliopathies. It shares phenotypic traits with other ciliopathies, such as Alström syndrome (ALMS), nephronophthisis (NPHP) or Joubert syndrome. BBS mutations have been detected in 16 different genes (*BBS1-BBS16*) without clear genotype-to-phenotype correlation. This extensive genetic heterogeneity is a major concern for molecular diagnosis and genetic counselling. While various strategies have been recently proposed to optimise mutation detection, they either fail to detect mutations in a majority of patients or are time consuming and costly.

Method We tested a targeted exon-capture strategy coupled with multiplexing and high-throughput sequencing on 52 patients: 14 with known mutations as *proof-of-principle* and 38 with no previously detected mutation. Thirty genes were targeted in total including the 16 BBS genes, the 12 known NPHP genes, the single ALMS gene *ALMS1* and the proposed modifier *CCDC28B*.

Results This strategy allowed the reliable detection of causative mutations (including homozygous/heterozygous exon deletions) in 68% of BBS patients without previous molecular diagnosis and in all *proof-of-principle* samples. Three probands carried homozygous truncating mutations in *ALMS1* confirming the major phenotypic overlap between both disorders. The efficiency of detecting mutations in patients was positively correlated with their compliance with the classical BBS phenotype (mutations were identified in 81% of 'classical' BBS patients) suggesting that only a few true BBS genes remain to be identified. We illustrate some interpretation problems encountered due to the multiplicity of identified variants.

Conclusion This strategy is highly efficient and cost effective for diseases with high genetic heterogeneity, and guarantees a quality of coverage in coding sequences of target genes suited for diagnosis purposes.

INTRODUCTION

Bardet-Biedl syndrome (BBS; OMIM# 209900) is a pleiotropic recessive disorder with high non-allelic genetic heterogeneity. Its incidence varies from an estimated 1:160 000 in northern Europe to 1:13 500–17 000 in Bedouins and Newfoundlanders, respectively.¹ BBS belongs to the large and growing family of ciliopathies and, therefore, shares phenotypic traits with Joubert (JBTS), Alström (ALMS) and Meckel (MKS) syndromes.^{1–2} Differential clinical diagnosis may thus be difficult, especially in young probands who do not yet show some later onset-specific manifestations.^{3–4} In particular, recent reports highlight a significant clinical overlap between BBS and ALMS.^{3–5}

The main phenotypic features of BBS comprise retinal dystrophy, polydactyly, obesity, mild developmental delay, polycystic kidneys and hypogonadism. Other minor features can also be observed in patients, such as cardiac abnormalities, other digit or eye anomalies, diabetes, hypertension, hearing defects, anosmia.^{6–7} Up to now, mutations have been detected in 16 different genes (*BBS1-BBS16*), but no clear genotype-to-phenotype correlation could be observed, besides the suggested exception of *BBS16*.⁸

Alström syndrome (OMIM #203800) was reported to be much less prevalent than BBS, with an estimated incidence of 1:1 000 000. Its phenotypic features overlap with those of BBS in early infancy and include: cone-rod dystrophy, obesity, type 2 diabetes mellitus, hearing loss but also hypertriglyceridemia, dilated cardiomyopathy, and progressive pulmonary, hepatic, or renal dysfunction.⁹ To date, only one gene (*ALMS1*) has been identified, but recent reports showed some families with suggestive ALMS-carrying mutations in BBS genes.^{3–5} The large size of *ALMS1* coding sequence appears to have impaired widespread diagnostic testing of ALMS.

Exhaustive conventional Sanger sequencing for BBS diagnosis is prohibitively expensive because of the large number of genes involved, and so also for ALMS due to the large size of *ALMS1* coding sequence (12 kb, 24 exons; table 1). Alternative



This paper is freely available online under the BMJ Journals unlocked scheme, see <http://jmg.bmj.com/site/about/unlocked.xhtml>

Table 1 Genes included in the targeted enrichment strategy and their associated disorders

Official gene symbol	BBS#	NPHP#	Other disease-related symbols	# Exons in consensus	# Total exons in all isoforms	Size of coding exons (bp)	Gene size (bp)	Targeted region size
BBS1*	BBS1	—	—	17	17	1782	22 966	23 285*
BBS2	BBS2	—	—	17	17	2166	35 750	3801
ARL6	BBS3	—	RP55	8	11	561	33 779	2123
BBS4*	BBS4	—	LCA	16	18	1560	52 292	52 611*
BBS5	BBS5	—	—	12	17	1026	27 160	3997
MKKS	BBS6	—	—	6	7	1713	29 034	3326
BBS7	BBS7	—	—	19	20	2148	46 008	5308
TTC8	BBS8	—	RP51	15	16	1518	53 358	3137
BBS9	BBS9	—	—	23	24	2559	476 529	5196
BBS10	BBS10	—	—	2	2	2172	3957	3941
TRIM32	BBS11	—	LGMD2H; STM	2	2	1962	13 999	4077
BBS12	BBS12	—	—	2	3	2133	12 242	3829
MKS1	BBS13	—	MKS1	18	20	1680	14 170	3745
CEP290	BBS14	NPHP6	MKS4; JBTS5; LCA10; SLSN6	54	55	7440	93 204	10 510
WDPCP	BBS15	—	—	18	19	2469	467 317	4405
SDCCAG8	BBS16	NPHP10	SLSN7	18	21	2141	244 087	3599
ALMS1	—	—	ALMS; LCA	23	24	12 510	224 161	13 682
NPHP1	NPHP1	NPHP1	JBTS4; SLSN1	20	22	2202	81 726	3264
INVS	NPHP2	NPHP2	—	17	19	3198	201 916	4103
NPHP3	NPHP3	NPHP3	MKS7; RHPD	27	27	3993	41 823	5328
NPHP4	NPHP4	NPHP4	SLSN4	30	30	4281	129 662	5693
IQCB1	NPHP5	NPHP5	SLSN5	15	15	1797	65 317	2585
GLIS2	NPHP7	NPHP7	—	8	8	1575	7374	2175
RPGRIPL1	NPHP8	NPHP8	MKS5; JBTS7; CS	27	27	3708	103 954	5243
NEK8	NPHP9	NPHP9	—	15	15	2079	13 953	3096
TMEM67	NPHP11	NPHP11	MKS3; JBTS6; CS	28	30	2745	64 389	4797
TTC21B	NPHP12	NPHP12	JBTS11; ATD4	29	29	3951	79 894	5414
TMEM216	—	—	MKS2; JBTS2	5	5	447	6504	1795
AHI1	—	—	JBTS3	27	29	3591	213 794	5175
CCDC28B	—	—	—	6	6	603	4790	1079
			Total	483	545	79 781	2 865 109	200 319

*Sequence of the entire gene (coding/non-coding exons, introns) was targeted.

ALMS, Alström syndrome; ATD, asphyxiating thoracic dystrophy; BBS, Bardet-Biedl syndrome; CS, COACH syndrome; JBTS, Joubert syndrome; LCA, Leber congenital amaurosis; LGMD, limb-girdle muscular dystrophy; MKKS, McKusick-Kaufman syndrome; MKS, Meckel-Gruber syndrome; NPHP, nephronophthisis; RHPD, renal-hepatic-pancreatic dysplasia; RP, retinitis pigmentosa; SLSN, Senior-Loken syndrome; STM, sarcotubular myopathy.

cost-conscious strategies have been proposed for BBS diagnosis, such as: initial screening of recurrent mutations and frequently mutated genes (*BBS1*, *BBS10*, *BBS12*) combined with homozygosity mapping for consanguineous families^{10–11}; or primer extension arrays to test a series of known BBS mutations.⁵ Another approach recently proposed is the pooling of patients' DNAs with subsequent PCR-amplification and massive parallel resequencing of *BBS1-12* coding exons, followed by heteroduplex screening to identify the mutation carrier.¹² Such a method presents some limitations as it will miss exon deletions and may not be suited for diagnostic purposes. Considering the clinical overlap with other ciliopathies, another approach would be to test, systematically and simultaneously, all corresponding genes for such overlapping syndromes, which would be particularly relevant for patients with atypical or incomplete clinical phenotypes. We describe here the results of such an approach, based on a targeted exon capture of 30 genes coupled to next-generation sequencing (NGS).

SUBJECTS AND METHODS

Detailed protocols are available in Supplementary Methods.

Subjects

DNA samples from 52 unrelated patients were collected. Most patients had been addressed to the diagnostic laboratory, or to

the National Reference Center for rare ophthalmogenetic diseases in Strasbourg. Eleven DNA samples stemmed from Tunisian patients included in an independent BBS epidemiology study.

The *proof-of-principle* cohort included 14 non-Tunisian patients with a confirmed BBS molecular diagnosis (identified prior to this study by Sanger sequencing). Twenty-six out of the 38 patients without known mutations, and recruited in Strasbourg, had been initially screened for *BBS1* and *BBS10* recurrent mutations, plus the entire coding sequence of *BBS12*.

For all patients, a written consent for genetic testing was obtained, either from adult probands or from the legal representative in case of minors.

Library preparation, targeted capture and sequencing

DNA samples were prepared and controlled following standard procedures.

The capture design was performed with eArray following the manufacturer's instructions (Agilent).

DNAs (3 µg) were sheared mechanically using Covaris E220 (duty cycle: 10%; intensity: 5; cycles per burst: 200; time: 300 s).

For the *proof-of-principle* experiment, sequencing adaptors were added on 500 ng of sheared DNA using the SPRIworks Fragment Library System I (Beckman Coulter). After amplification and quality assessments, targeted capture was performed on individual samples using the *in-solution* SureSelect Target Enrichment System (Agilent) on 500 ng of DNA-prepped library. Additional

steps of washing, purification and elution were performed, and multiplexing adaptors (TruSeq Illumina DNA indexes) were added by PCR during the post-capture amplification step.

For all following experiments, multiplexing adapters were added simultaneously to sequencing adapters using the SPRI-works system. Equimolar amounts of two tagged libraries were then pooled prior to the capture reaction. All other following steps prior to sequencing remained identical. A 72-bp single-read sequencing was performed on a Genome Analyser IIx (GAIIx, Illumina).

Bioinformatic pipeline

Read mapping and variant calling were performed following standard procedures. Variant filtering was performed using VaRank, an in-house software which collects variant-specific information to rank them according to their predicted pathogenicity (figure 1, Supplementary Methods).

Copy-number variation (CNV) detection method

CNVs were identified using a depth-of-coverage method.^{13 14} For each patient, read counts in non-overlapping windows of 20 nucleotides were computed, normalised and then compared randomly with eight other samples from the same experiment (considered as replicates) using the Bioconductor package DEseq (initially designed for RNA-seq data).¹⁵ Candidate regions for

CNVs were retrieved when log₂ ratios (controls/sample) were either ≥ 0.84 (fold change >1.8 , potential deletion) or ≤ -0.51 (fold change <0.7 , potential duplication), and if p values adjusted for multiple testing (Benjamini and Hochberg procedure)¹⁶ were smaller than 0.1.

Statistical methods

Confidence intervals were computed for proportion estimates and indicated in brackets. Fisher's exact test was computed to compare distributions of small populations. Subsequent p value is given at $\alpha=0.05$.

RESULTS AND DISCUSSION

Targeted regions: design strategy

Our primary goal was to develop an efficient mutation-screening strategy for the diagnosis of patients with phenotypes evocative of BBS, or of clinically overlapping ciliopathies. We chose a target enrichment approach coupled with NGS in order to focus the sequencing on genomic regions of interest. We targeted all exons (including 5' and 3' UTRs) of the 16 known BBS genes (table 1). Because of the known clinical overlap, we also included coding exons of *ALMS1*, and of all 12 known nephronophthisis genes (*NPHP1-12*), since retinal degeneration can often be observed in this kidney-specific disease.^{9 17} Coding sequences of *AHI1/JBTS3*, *TMEM216/MKS2/JBTS2*, and of the proposed BBS-modifier *CCDC28B/MGC1203*, were also targeted.¹⁸ Because some of these genes are associated with multiple phenotypes, our design includes 6 MKS, 7 JBTS and 4 Senior-Loken syndrome (SLSN) genes (see table 1).

With this first design, we wanted to investigate whether including intronic sequences could favour both, the detection and sizing of exon deletions. We therefore included baits-targeting intronic sequences of *BBS1* and *BBS4*. This choice was dictated by two observations: an apparent excess of patients heterozygous for the *BBS1*-recurrent mutation M390R with no second mutation detected,¹¹ and multiple reports of *BBS4* exon deletions in patients.^{4 11 19} A maximal threshold of 200 kb for cumulated targeted regions was set because of the manufacturer's pricing limits.

Presence of repeated sequences precluded bait tiling in 19.7% of initially targeted regions. This concerned, almost exclusively, introns of *BBS1* and *BBS4*, besides a small number of 3' UTRs, and only 128 bp of protein coding regions (within first exons of *ALMS1* and *NPHP3*; table S1).

Proof-of-principle and technical results

In our *proof-of-principle* experiment, we selected 16 DNA samples, of which 14 were with known BBS mutations. In this first trial, after barcoding the target-enriched libraries, we sequenced pools of four or eight libraries per lane of a GAIIx (see Supplementary Methods). This *proof-of-principle* analysis was carried blind, that is, without knowledge of implicated BBS genes and their associated mutations. A constellation of all mutation types (missenses, nonsense, splice mutations, large deletions and complex rearrangements) at different allelic dosage was tested (figure S1). All 14 previously identified mutations, including two heterozygous *BBS1*-deletions (figure 2A), were detected in their correct heterozygous/homozygous state (table S2). In particular, in patient AKE12, we could detect an abnormal local drop of coverage in *BBS12* due to a rare mutation type (insertion of an Alu sequence, figure S1A) although the exact nature of the mutation could only be determined by Sanger sequencing. A

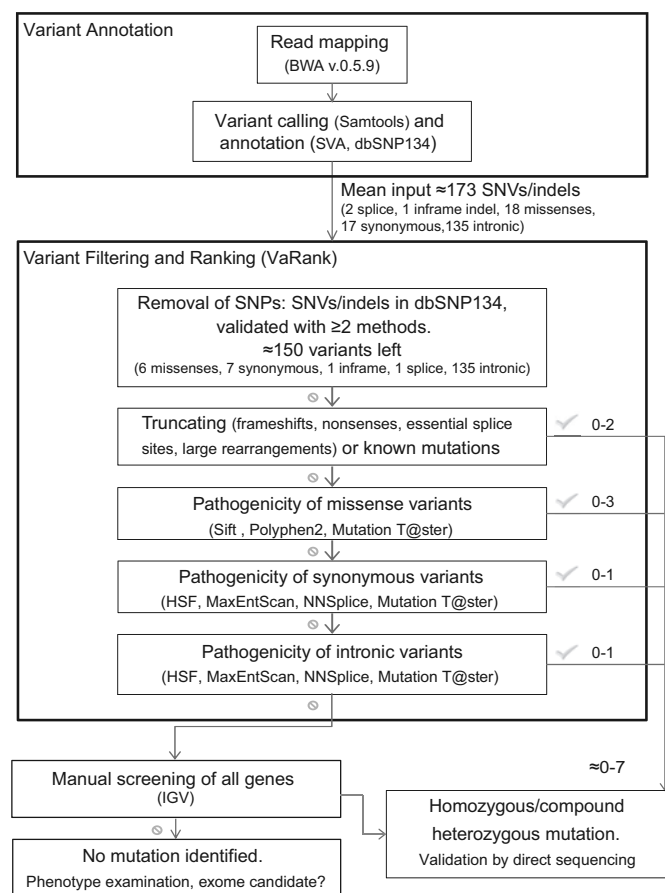
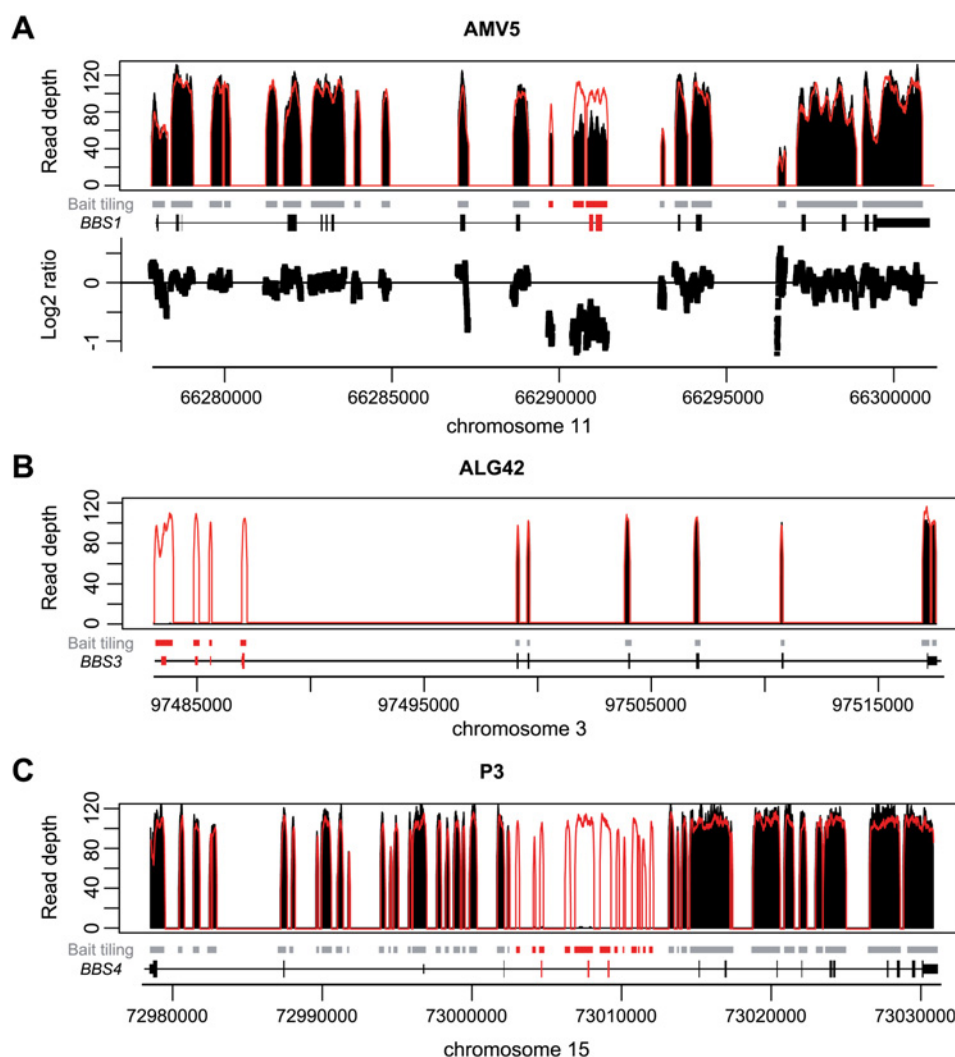


Figure 1 Global flowchart of the bioinformatic pipeline implemented for mutation detection. Software acronyms: BWA, Burrows-Wheeler Aligner; SVA, Sequence Variant Analyser; SIFT, Polyphen2; HSF, Human Splicing Finder; MaxEntScan, Maximum Entropy Scanning; NNSplice; Mutation Taster; IGV, Integrative Genome Viewer.

Figure 2 Detection of large deletions in three patients using a depth-of-coverage method. Black peaks: normalized depth of coverage from patients' DNA samples. Empty peaks: normalized mean depth of coverage across samples from the same sequencing lane. Grey squares: bait-covered regions. Black peaks: normalized depth of coverage from patients' DNA samples. Empty peaks: normalized mean depth of coverage across samples from the same sequencing lane. Grey squares: bait-covered regions. Highlighted squares: deleted regions. Gene representation: black squares: exons, dashed lines: introns. Genomic positions are given according to the human reference genome hg19/ GRCh37. (A) Heterozygous deletion of *BBS1* (exon #10, 11) in AMV5 patient. Corresponding Log₂ ratios between both depths of coverage (normalised mean and AMV5 patient) further highlight the presence of the deletion. (B) Homozygous deletion of *BBS3* (exon #1, 2a, 2b, 3) in ALG42 patient. (C) Homozygous deletion of *BBS4* (exon #4, 5, 6) in P3 patient. (A and C): targeting intronic sequences allows restricting the deletion breakpoints. (B) and (C): Log₂ ratios between both depths of coverage (normalised mean and corresponding patients) could also allow detecting both deletions but are not shown (supplementary figure S2).



similar drop in coverage was observed for a second patient, AHX91, with another complex mutation detected previously by Sanger sequencing (insertion/inversion in *BBS5*).

In this first experiment, we almost systematically reached the maximal theoretical coverage of 144x illustrated by a mean coverage of $127 \pm 4x$ after removal of duplicate reads (table 2). Due to this global saturating coverage when considering unique reads, we used all reads, including duplicates, when applying our depth-based method for the detection of CNVs.

These promising depth-of-coverage results (table 2, table S3) encouraged us to further increase the number of pooled samples. In the next experiments we used a single capture reaction for two barcoded libraries, allowing both cost and bench-time savings, and pooled 12 libraries per sequencing lane (maximum number of barcodes proposed at that time by Illumina).

This new protocol was performed on a second cohort of 36 patients with unknown mutations. Sequencing resulted in

a mean coverage of $78 \pm 17x$ ($283 \pm 153x$ before discarding duplicate reads) with $91.4 \pm 6.4\%$ of targeted regions being covered more than 40x (table 2). This relative drop of coverage appears to be a consequence of a lower capture efficiency that might be due to: (1) an input amount of individual library reduced by half, due to the pre-capture pooling and (2) the addition of barcodes before capture, leading to less efficient blocking and unspecific hybridisation. The resulting coverage still guarantees a reliable detection of variants and of their homozygous/heterozygous state.

A small proportion of targeted regions was weakly covered in some patients (ie, depth $< 10x$ after duplicates filtering), with very few of them in a systematic way in other patients (table S4). This only concerned $0.63 \pm 0.68\%$ of protein coding regions, and mostly included intronic GC-rich sequences (GC content: $68.3 \pm 5\%$ vs $40.2 \pm 10\%$ across all targeted regions), or some first exons (tables S3 and S4).

Table 2 Sequencing statistics of both coverage (in captured regions) and capture efficiency

	Mean coverage (x)		Targeted regions with 5x coverage	Targeted regions with 40x coverage	Targeted regions with 80x coverage	Reads in targeted regions	
	Before filters	After filters				Before filters	After filters
Pool of 4*	2208 ± 416	130 ± 3	100 ± 0%	99.9 ± 0%	98.3 ± 0.6%	76 ± 2%	35 ± 5%
Pool of 8*	1024 ± 151	125 ± 3	99.9 ± 0.3%	99.7 ± 0.3%	96.9 ± 0.9%	74 ± 5%	40 ± 6%
Pool of 12†	283 ± 153	78 ± 17	99.6 ± 0.4%	91.4 ± 6.4%	48.7 ± 28%	25 ± 9%	14 ± 4%

*First set of experiment. The capture-enrichment step was done individually, on untagged DNA libraries.

†Second set of experiment. The capture-enrichment step was realised on equimolar pools of 2 barcoded DNA libraries.

Variant filtering: importance of databases and frequency data

In targeted regions, we detected, on average, 170 variants (Single Nucleotide Variants (SNVs) and indels) per patient. All were systematically analysed for putative effect on protein structure and splice sites using VaRank (figure 1, Supplementary Methods). About 130 of these variants were recorded in dbSNP134 (table S5), but only 20 were validated with at least two independent methods and, therefore, filtered out. Indeed, in the context of a rare recessive disorder, some true mutations can be present at very low frequency in a heterozygous state in controls.

Potential pathogenicity of the remaining 150 variants was assessed using bioinformatic tools and considering their allele frequency in a European-American population, as reported in the Exome Variant Server database (EVSdb). This yielded from zero to six interesting variants per sample, among which were obvious truncating or known mutations in some patients.

The new 'clinical significance' field introduced in dbSNP134 has to be considered with caution since established mutations can now be reported in the database but are not systematically flagged as pathogenic (example: rs179363897, p.R138H) mutation in *BBS5*). Conversely, we detected some false-positive annotations: rs4784677 (p.N70S) in *BBS2*—initially reported as a third allele according to the triallelic hypothesis²⁰—is flagged as pathogenic, but is too frequent to be a fully penetrant mutation (0.77% in EVSdb). Filters have to be carefully adapted to the disorder of interest, and to the constantly evolving updates of databases.

Detection of exon deletions

One advantage of NGS-based strategies, as opposed to Sanger sequencing, is the opportunity to detect—in addition to SNVs and small indels—CNVs affecting one or more exons (figure 2 and S2). In the *proof-of-principle* experiment, two heterozygous deletions could be detected in *BBS1*. Among the unknown samples, two homozygous deletions in *BBS3/ARL6* and *BBS4* were identified. To our knowledge, we provide here the first report of large deletions in *BBS1* and *BBS3/ARL6*, while several deletions affecting *BBS4* have been previously observed.^{4 11 19} Since we also targeted intronic sequences of *BBS1* and *BBS4*, we were able to narrow the boundaries of subsequent detected deletions (figure 2). For patient AMV5, by using coordinates of affected exons, the estimated size of *BBS1* deletion would be between 466 and 4707 bp, while with our design, we could restrict it to 1862–3841 bp (figure 2A, table S2). In patient P3, we could similarly reduce the assessed size of the *BBS4* deletion from 4626–12 975 bp based on exon positions down to 9376–10 469 bp (figure 2C, table 3A). Lastly, since the *BBS3/ARL6* deletion in patient ALG42 encompasses the first three exons of the gene (figure 2B), we tested whether it may extend and affect *EPHA6* located upstream, encoding an ephrin receptor. Direct PCR testing excluded such extended deletion (data not shown).

Thanks to this method, in the six patients in whom we detected a single heterozygous potentially pathogenic mutation, we can ascertain that no heterozygous deletion is present in *trans*, or at least none encompassing exonic sequences.

Distribution of detected BBS mutations in the 38 unknown patients

Of the 38 samples with unknown BBS mutations (36+2 from the *proof-of-principle* experiment) we detected clearly pathogenic biallelic mutations in 26 cases (68.4%; table 3A). To our

knowledge, seventeen of these mutations have not been reported previously, indicating that the BBS mutation spectrum is far from being saturated in spite of numerous BBS mutation reports. Homozygous mutations were found in 88% (23/26) of mutated patients, coherent with the large number of consanguineous probands included in the cohort (75%; 25/33). In two patients of consanguineous origin, the BBS mutation was located outside the homozygous regions detected by prior SNP array analysis and would, therefore, have been missed using a homozygosity mapping strategy (patients AGL23, AKX44; table 3A).

Among the remaining 12 patients with no biallelic mutations identified (table 3B), one patient (AHR2) had a heterozygous clearly pathogenic splicing mutation in *BBS3*. Five patients had heterozygous missenses predicted to be damaging by SIFT, Polyphen2 and/or Mutation T@ster (AIY87, AIX45, AMO77, AMA28, AHL86) with the latter two carrying such variants in two different genes. One consanguineous Melanesian patient, AKE98, presented with classical BBS-inclusion features, including polydactyly. He carried two homozygous variants which initially appeared as potentially pathogenic: a distal frameshift in *INVS/NPHP2* predicted to add 14 amino-acids to the C-terminus of the longer protein isoform, and a non-reported missense P2679L affecting a conserved residue in *ALMS1* (figure S3). Subsequent segregation analysis ruled out their implication in the disease since both variants were heterozygous in a similarly affected brother.

In five patients, no potentially pathogenic variant could be identified in any of the 30 targeted genes. These patients are thus candidate for exome sequencing which might either help in identifying novel genes or in reconsidering the clinical diagnosis.

Mutation load in BBS and other targeted genes: importance of *ALMS1*

The mutation load among BBS genes in our cohort appears consistent with previous reports.^{7 11} Observed occurrences for *BBS1* (7/38, 18.4%) and *BBS2* mutations (5/38, 13.2%), the most frequently mutated genes in our study, are similar to the respective reported figures of 16.9% and 12%.⁷ Considering frequently mutated genes, our study was strongly biased against *BBS1*, *BBS10* and *BBS12*, since two-thirds of the patients' DNAs were previously tested negative for *BBS1* and *BBS10* recurrent mutations, plus all *BBS12* protein coding sequence. This explains the total absence of *BBS12* mutations in our cohort, and the relatively low contribution of *BBS10* (5/38, 13.2%) as compared with the literature ($\geq 20\%$).^{7 11} The contribution of other BBS genes was low, with frequently only one proband involved.

We did not find any mutation in the 'new' BBS genes (*BBS13–16*) suggesting that, cumulatively, they have a small contribution to the total mutation load. *BBS13/MKS1* was indeed shown to be mostly implicated in MKS since only one BBS patient was reported with two heterozygous mutations p.[C492W]; [F371del] others carried only heterozygous missenses, sometimes in addition to homozygous truncating *BBS1* mutations.²¹ Likewise, for *BBS14/CEP290*, a homozygous truncating mutation (p.E1903*) was found in a single BBS patient,²¹ while other mutations are much more often implicated in Joubert, Senior Loken, Leber Congenital Amaurosis or Meckel syndromes. Similar observations can be made for *BBS15/WDPCP* and *BBS16/SDCCAG8*.^{22 23} Like in all other studies of BBS cohorts, no mutation was identified in *BBS11/TRIM32* raising the question of its real implication in BBS: only one homozygous missense mutation was described in a single

Table 3 Identified mutations and other potentially pathogenic variants in the 38 patients with previously unknown genotype. A) Patients with two clearly pathogenic variants in one gene; B) Patients with a single or no clear pathogenic variant in one gene. Mutations are described according to the latest nomenclature conventions described in HGVS

Patient#	Principal mutations (p.)	Principal mutations (c.)	Sanger validation/ segregation	Mutation, as predicted		Additional potentially pathogenic variants	Frequency of additional variants in EVSub	Geographic origin	Inbred	BBS inclusion criteria
				by (among SIFT, PPhen2 & Mutation T@ster)	PPhen2 & Mutation T@ster					
A)										
BBS1										
AL047	p.[R160Q]; [R160Q]	c.[479G→A]; [479G→A]	SV	PP, MT	—	—	—	Turkey	Yes (BBS1, BBS4)	No
P1	p.?: ?	c.[1473+4A→G]; [1473+4A→G]	SV + S	MT (Ex14 splice site altered)	—	—	—	Tunisia	Yes*	Yes
P9	p.?: ?	c.[1110G→A]; [1110G→A]	SV + S	MT (Ex11 splice site altered)	AH11/JBTS3: p.[R830W]; [—]	2.83%	—	Tunisia	Yes*	No
AMK19	p.?: [M390R]	c.[951+1G→A]; [1169T→G]	SV + S	MT (Ex10 splice site altered); S	TTC21B/NPHP12: p.[R713I]; [—], BBS7: p.[D412G]; [—]	NF, 0.23%	—	NA	NA	No
P12	p.[A14Lfs*28]; [A14Lfs*28]	c.[39del]; [39del]	SV + S	MT	—	—	—	Tunisia	Yes*	Yes
P11	p.[R146*]; [R146*]	c.[436C→T]; [436C→T]	SV + S	MT	CCDC28B: Ex2 splice site altered (F110F); [—]	2.07%	—	Tunisia	Yes*	Yes
AI057	p.[E224K]; [E224K]	c.[670G→A]; [670G→A]	SV	S, PP	NPHP3: p.[R1074H]; [—], ALMS1: p.[S2102L]; [—]	0.01%, 3.07%	—	Morocco	Yes (BBS1, BBS3, NPHP5, JBTS2)	No
BBS2										
P2	p.?: ?	c.[345+5G→A]; [345+5G→A]	SV + S	(Ex2 splice site altered)	—	—	—	Tunisia	Yes*	Yes
AGL23	p.[H665Tfs*11]; [H665Tfs*11]	c.[1992del]; [1992del]	SV + S	MT	TTC21B/NPHP12: p.[R616C]; [—]	0.31%	—	India	Yes (ALMS1, NPHP1, NPHP3, NPHP6, NPHP8)	No
P7	p.[R189*]; [R189*]	c.[565C→T]; [565C→T]	SV + S	MT	NPHP4: p.[R674H]; [R674H]	NF	—	Tunisia	Yes*	Yes
AKX44	p.[R272*]; [R272*]	c.[814C→T]; [814C→T]	SV	MT	SDCCAG8/BBS16: p.[Q505E]; [—]	NF	—	Algeria	Yes (BBS7, BBS12, NPHP8)	Yes
ALG76	p.[L209P]; [L209P]	c.[626T→C]; [626T→C]	SV + S	S, PP, MT	—	—	—	Turkey	Yes (BBS2, NPHP8)	Yes
BBS3/ARL6										
ALG42	p.[del Ex1-3]; [del Ex1-3]	c.[(?-30)_(123+?)del]; [(?-30)_(123+?)del]	SV + S	—	—	—	—	Turkey	Yes (ALMS1, BBS3, BBS4, BBS9, BBS14, NPHP2, NPHP3)	Yes
BBS4										
P3	p.[del Ex4-5-6]; [del Ex4-5-6]	c.[(157-?)_(405+?)del]; [(157-?)_(405+?)del] (g.[24029..24285]_ (33661..34498))del]; g.[24029..24285]_ (33661..34498))del]	SV + S	—	BBS5: p.[N184S]; [—]	0.57%	—	Tunisia	Yes*	Yes
BBS5										
P13	p.[L50R]; [L50R]	c.[149T→G]; [149T→G]	SV + S	S, PP, MT	—	—	—	Tunisia	Yes*	Yes
ALG5	p.[R138H]; [R138H]	c.[413G→A]; [413G→A]	SV + S	S, PP, MT	—	—	—	France	No	No
BBS6/MKKS										
P10	p.?: ?	c.[1272+1G→A]; [1272+1G→A]	SV + S	MT (Ex5 splice site altered)	NPHP4: p.[R959Q]; [—]	NF	—	Tunisia	Yes*	Yes
BBS8/TTCS8										
P14	p.?: ?	c.[329+1G→A]; [329+1G→A]	SV + S	MT (Ex4 splice site altered)	AH11/JBTS3: p.[R830W]; [—]	2.83%	—	Tunisia	Yes*	Yes

Continued

Table 3 Continued

Patient#	Principal mutations (p.)	Principal mutations (c.)	Sanger validation/ segregation	Mutation, as predicted by PPhen2 & Mutation T@ster	Additional potentially pathogenic variants	Frequency of additional variants in EVSdb	Geographic origin	Inbred	BBS inclusion criteria
BBS10									
P8	p.[S396Lfs*6]; [S396Lfs*6]	c.[1171_1181dupGCATTATACC]; [1171_1181dupGCATTATACC]	SV + S	MT			Tunisia	Yes*	Yes
JSL	p.[V707*]; [R95S]	c.[2119-2120delGT]; [285A→T]	SV + S	MT; S, PP, MT	CCDC28B: Ex2 splice site altered (F110F); [=]	2.07%	NA	NA	NA
BBH64	p.[T483Nfs*8]; [?]	c.[1448-1452delCTCAA]; [?]	S	MT			NA	No	No
AMR64	p.[L414S]; [L414S]	c.[1241T→C]; [1241T→C]	SV	S, PP	CCDC28B: Ex2 splice site altered (F110F); [=]	2.07%	NA	NA	Yes
AKR68	p.[L414S]; [L414S]	c.[1241T→C]; [1241T→C]	SV	S, PP	CEP290/BBS14: p.[K1870Nfs*4]; [=]	NF	NA	Yes (BBS10, JBTS3)	No
ALMS1									
AIA84	p.[E1114Rfs*9]; [E1114Rfs*9]	c.[3340del]; [3340del]	SV	MT			Turkey	Yes (ALMS1, BBS14)	Yes
AK026	p.[R3629*]; [R3629*]	c.[10885C→T]; [10885C→T]	SV + S	MT	WDPCP/BBS15: p.[V329M]; [=], AH1/JBTS3: p.[R548H]; [=]	0.13%, 1.65%	France	Yes (ALMS1, BBS12)	Yes
ALB64	p.[S577*]; [S577*]	c.[1730C→G]; [1730C→G]	SV + S	MT	IQG81/NPHP5: p.[E481K]; [=]	0.17%	Portugal	Yes (BBS7, BBS12, BBS14, ALMS1)	No
Patient#	Gene	Potential mutations (p.)	Potential mutations (c.)	Frequency in EVSdb	Mutation, as predicted by (among SIFT, PPhen2 & mutation T@ster)	Geographic origin	Consanguinity	BBS inclusion criteria	
B)									
AHR2	BBS3/ARL6	p.[D179N]; [=]	c.[535G→A]; [=]	NF	MT (Ex8 splice site altered)	Reunion	Yes: BBS3, JBTS3	No	
AIY87	BBS9	p.[P641S]; [=]	c.[1921C→T]; [=]	NF	S, PP, MT	NA	Yes: BBS5, ALMS1, NPHP12, NPHP5, BBS14	No	
AKE98	ALMS1	p.[P2679L]; [P2679L]	c.[8036C→T]; [8036C→T]	NF	S, P	Melanesia	Yes: BBS12, BBS9, CEP290, NPHP1, MKS1	Yes	
AIX45	NPHP4	p.[V228G]; [=]	c.[3683T→G]; [=]	0.10%	S, P			Yes	
AMA28	NPHP3	p.[A519G]; [=]	c.[1556C→G]; [=]	NF	P, MT	France	No	Yes	
	INVS/NPHP2	p.[N1061Kfs*20]; [N1061Kfs*20]	c.[3176_3177insA]; [3176_3177insA]	NF	—				
	TIME67/ NPHP11	p.[V486C]; [=]	c.[1457A→G]; [=]	0.01%	P, MT	NA	Yes	Yes	
	MKKS/BBS6	p.[A242S]; [=]	c.[724G→T]; [=]	0.58%	S, P	France	No	Yes	
	NPHP1	p.[E569G]; [=]	c.[1706A→G]; [=]	NF	S, P, MT				
AM077	TTC21B/ NPHP12	p.[R863W]; [=]	c.[2587C→T]; [=]	0.23%	S, P, MT	NA	NA	Yes	
AHL86	AH1/JBTS3	p.[R830W]; [=]	c.[2488C→T]; [=]	2.83%	S, P, MT	NA	NA	No	
	INVS/NPHP2	p.[T122P]; [=]	c.[364A→C]; [=]	NF	S, P, MT				
ALA60						Caucasian	No	Yes	
AKT96						France	No	No	
AMT10						France	No	Yes	
AIF95						NA	No	No	
AKK32						France	No	No	

*Patients from consanguineous families, but not genotyped on Affymetrix SNP arrays. In bold: previously reported mutations in other studies. Mutation prediction software equivalents: possibly/probably damaging (polyphen2), damaging (SIFT), disease causing (mutation taster). Ex, exon; S, SIFT; PP, PolyPhen2; MT, mutation T@ster.
 (?): The exact nature of the second heterogeneous mutation could not be identified by high-throughput sequencing. An abnormal loss of coverage is observed at the very end of BBS10. The exact nature of this apparently complex mutation is still under investigation by direct sequencing, but appears to involve an alu insertion coupled with a duplication/inversion. Amino acid conservation of non-reported missense mutations is shown in supplementary figure S4. SV, Sanger validation; S, segregation validated. Consanguinity was documented by clinicians, and in most cases, BBS patients from consanguineous families were genotyped on Affymetrix 250K SNP arrays. When so, BBS and other targeted genes located within homozygous regions are thus indicated in ().

consanguineous family, while several other mutations were identified in recessive forms of limb girdle muscular dystrophy.²⁴

One noteworthy result is the finding of homozygous truncating *ALMS1* mutations in 3/38 patients (AIA84, AKO26, ALB64; 7.9%). In particular, the nonsense found in AKO26 patient p.R3629* seems to be a recurrent *ALMS1* mutation, since already reported in five other *ALMS* patients.^{25–27} The phenotypic overlap between BBS and *ALMS* seems to be larger than previously thought, as recently suggested with examples of Alström patients with mutations in BBS genes,⁵ and the reverse situation, such as in our study, of *ALMS1* mutations in patient with suspected BBS.³

Lastly, no clearly pathogenic mutation was found in any NPHP or JBTS genes in the cohort.

Correlation between mutation detection efficiency and clinical phenotype

Comparison of clinical phenotype between patients with two clearly pathogenic mutated alleles ($n=26$) and those with either a single possible pathogenic variant or no suspicious variant detected ($n=12$) showed a clear correlation between the number of major BBS clinical features and the probability of detecting two BBS mutated alleles in patients (figure 3). Biallelic mutations were detected in 81% (CI (60% to 92%)) of patients meeting BBS inclusion criteria. In particular, in Tunisian patients recruited upon strict clinical criteria, mutations were found in 11/11 cases and in seven different BBS genes, ruling out a potential founder effect. On the contrary, for some of the 12 patients without clear mutations, BBS was only one suspected

diagnosis among others. Furthermore, our initial selection of patients without recurrent mutations in *BBS1* or *BBS10*, and without any mutation in *BBS12* may have enriched our cohort in patients with non-typical BBS phenotypes. The current widely quoted estimation that known BBS genes account for only 70–75% of the total mutation load in BBS patients may thus be underestimated if considering only patients with strictly defined BBS phenotype.

The distribution of BBS inclusion features appears different between patients with two BBS mutations, two *ALMS1* mutations or no biallelic mutation identified (table 4). Patients with no detected mutation presented with significantly less polydactyly, a major BBS clinical sign: only 25% versus 70% in patients with detected BBS mutations ($p=0.029^*$). The other clinical features seem to follow the trend of classical BBS patients.

Regarding *ALMS1*-mutated probands, 2/3 had been sent for suspected BBS (Prader-Willi or *ALMS* were also considered for AIA84 and AKO26, respectively) and satisfied BBS diagnostic criteria; the last one (ALB64) was addressed for syndromic retinal dystrophy (table 4).⁶ AIA84 presents a classical BBS with retinal dystrophy, obesity, cognitive defects, hypogonadism and brachydactyly. AKO26 presents an atypical BBS with the same features along with abnormal severe deafness, specific for *ALMS*. Lastly, ALB64 presents a typical *ALMS* with severe deafness and retinal dystrophy. None of them presented with polydactyly. As previously suggested,³ both, the absence of polydactyly and the prevalence of deafness in *ALMS1*-mutated patients, are keys for genotype-phenotype discrimination between *ALMS* and BBS mutated patients.

Assessment of oligogenism in BBS

The presence and potential effect of triallelism or oligogenism in BBS has been widely discussed and appears controversial (^{18 20 28 29} vs^{30–32}). In our approach, the simultaneous sequencing of all 16 BBS genes, and of 14 other genes involved in overlapping ciliopathies, allows the systematic detection of most additional potentially pathogenic variants in those genes and, consequently, an unbiased assessment of oligogenism.

Out of the 52 patients analysed, we found only one heterozygous truncating mutation (p.K1870Nfs*4) as a third allele in *BBS14/CEP290* in patient AKR68 who carries a pathogenic missense mutation in *BBS10*. Such a frequency is in fact in the range of what to expect by chance. We previously calculated the probability of carrying a true BBS-pathogenic mutation to be about 1:50.³¹ Since we also included in our design other ciliopathy genes, the probability to carry a pathogenic mutation in one of the 30 genes is rather between 1:20 and 1:30 (calculation based on each disease incidence and reported contributions of targeted genes in the mutation load). Potentially, pathogenic heterozygous missenses (not previously reported in patients or in the EVSdb) were also found in eight patients (three of the *proof-of-principle*, table S2; five of the ‘unknown’ cohort, table 3). Such variants might act as modifiers, but it is unlikely that they are required for full expression of a classical BBS phenotype. Conversely, in some patients where a single clearly pathogenic mutation was found, variants in other genes of the same pathway (especially those encoding proteins of the same complex, such as the BBSome or the BBS-chaperonin complex)^{33 34} might contribute to the disease state in a digenic mode of inheritance proposed in few BBS families.^{20 35 36}

Potential case of triallelism is illustrated by patient AIZ62, who is compound heterozygous or a nonsense p.E191* and a missense p.A242S in *BBS6/MKKS*. The pathogenicity of A242S variant has

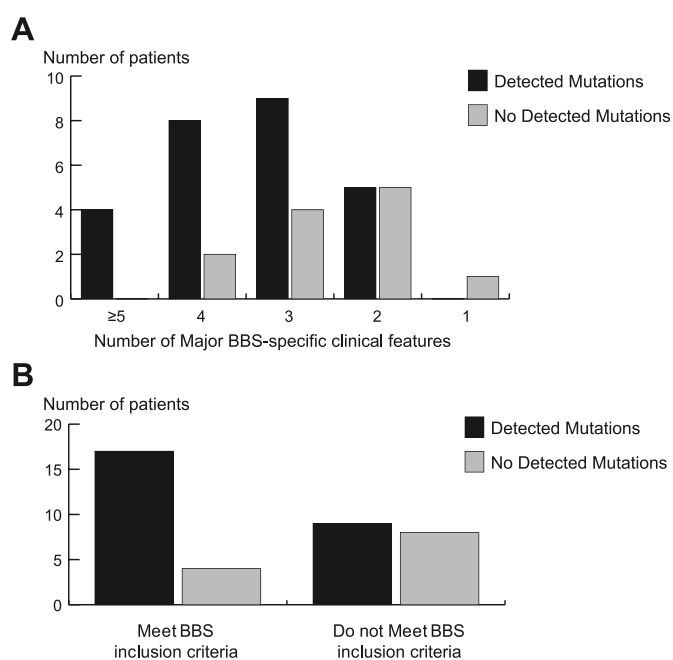


Figure 3 Compliance with classical BBS phenotype is positively correlated to the efficiency to detect principal mutations in BBS genes. (A) The number of BBS diagnostic major inclusion criteria⁶ in patients is correlated to an efficient detection of BBS mutations. (B) Efficiency of detecting mutation in patients fulfilling BBS the phenotypic inclusion criteria or not. BBS inclusion criteria presenting with three major features plus at least two minors, or presenting with four major features and more.⁶ Primary criteria include: rod-cone dystrophy, polydactyly, obesity, learning disabilities, hypogonadism and renal anomalies. Secondary features comprise speech delay, other eye anomalies, brachydactyly or syndactyly, ataxia, diabetes, developmental delay, dental anomalies, cardiac anomalies and hepatic fibrosis.⁶

Table 4 Report of major BBS clinical features in the 38 patients without previously known molecular diagnosis, with or without detected mutations

	<i>Retinitis pigmentosa</i>	<i>Obesity</i>	<i>Polydactyly</i>	<i>Brachydactyly</i>	<i>Hypogonadism</i>	<i>Cystic kidney</i>	<i>Mild intellectual disabilities</i>
2 BBS mutations	95.65% (22/23)	86.96% (20/23)	69.57% (16/23)	43.48% (10/23)	39.13% (9/23)	21.74% (5/23)	73.91% (17/23)
2 ALMS1 mutations	3/3	1/3 (AIA84)	0/3	2/3 (AIA84, AKO26)	2/3 (AIA84, AKO26)	1/3 (ALB64)	2/3 (AIA84, AKO26)
0 or 1 mutation	81.82% (9/11)	81.82% (9/11)	25.00% (3/12)	33.33% (4/12)	41.67% (5/12)	50% (5/10)	81.82% (9/11)

ALMS1 patients: AIA84 was addressed to Strasbourg Diagnostic Laboratory for Bardet–Biedl or Prader–Willi syndromes, AKO26 for suggestive BBS or ALMS with abnormal cognitive defects and ALB64 for a syndromic retinal dystrophy or suggested ALMS. Patients with a proposed Alström syndrome presented with early deafness: at 5 (ALB64) or 6 (AKO26) years of age.

been a subject of discussion.^{37–39} Analysis in zebrafish indicated that it affects BBS6/MKKS function and suggested a dominant negative effect.⁴⁰ EVSdb allows to infer its frequency at 0.59% (CI (0.43 to 0.80%)), higher than the most frequent BBS mutation, M390R in BBS1 (0.24%, CI (0.15 to 0.39%)). A242S cannot thus be a highly penetrant mutation since it should then be found more frequently in patients than the M390R mutation, which is not the case. In patient AIZ62, a third heterozygous variant was identified in BBS12 (p.Q620R, residue conserved in mammals, but not in more distant vertebrates) thus affecting another subunit of the BBS-chaperonin complex.³⁴ We suggest that A242S is a hypomorphic allele that may lead to a phenotype when in *trans*, with a complete null mutation, and could be further potentiated by a hypomorphic allele affecting another subunit of the same complex. Segregation analysis in AIZ62 family could not be performed to test this hypothesis.

Lastly, we looked in our cohort for the allelic frequency of the previously proposed BBS-modifier variant c.330C→T in *CCDC28B/MGC1203*,¹⁸ and found a frequency of 3.85% (CI (1.56 to 9.47%)), which is not significantly different ($p=0.17$) from the 2.07% (CI (1.76 to 2.43%)) observed in EVSdb.

CONCLUSIONS

The extensive non-allelic genetic heterogeneity of Bardet–Biedl syndrome has been a major problem for molecular diagnostic and genetic counselling applications. Various strategies have been proposed in recent years to optimise mutation detection,^{5 10–12} but have either low sensitivity or are too time consuming and expensive in diagnostic settings. This problem is shared by other disease entities, such as cardiomyopathies, hearing loss, Usher syndrome or Charcot–Marie tooth neuropathies. Those NGS-based alternative strategies can be divided in whole genome/exome sequencing,^{41–43} or targeted sequencing, either by using multiplex PCR,¹² multiple singleplex PCR^{44–46} or capture enrichment approaches.

We implemented an *in-solution* targeted capture strategy for the 16 known BBS genes and 14 other genes implicated in ciliopathies that share overlapping clinical features with BBS. We show here that this is very efficient since in a single sequencing lane of an Illumina GAIIX one can simultaneously analyse more than 99.4% of targeted protein coding sequences of these genes in 12 patients, with sufficient coverage to guarantee reliable detection of heterozygous variants, small indels and exon deletions. In a week-long run, one could thus potentially analyse 96 patients. Investigation of 36 patients could be completed in about 3 weeks when including sample preparation and initial bioinformatic analysis. Data analysis and further mutation validation takes a longer time for cases where only variants of uncertain pathogenicity are present. We estimated the overall consumable costs in our settings at about \$600. This can be even further decreased by using the latest more powerful sequencers, allowing analysis of larger pools of barcoded samples while keeping a high depth of coverage. Indeed, we recently sequenced a new cohort of 24 patients (using 12 capture reactions) in

a single lane of an Illumina HiSeq2000 with even higher coverage than the one obtained in previous experiments with the GAIIX (data not shown). While exome sequencing is clearly more exhaustive, in terms of gene coverage in current implementations, 20–30% of targeted exons are not sufficiently covered for diagnostic accuracy, that is, to ensure low rates of false-positive/false-negative findings and reliable detection of heterozygous mutations or exon deletions. Moreover, the informatics resources needed for exome/genome sequencing data analysis and storage are considerably more important than for targeted sequencing, and can often be a limitation.

This strategy, however, presents some limited pitfalls. Few protein coding regions were not well covered, either because of failure in bait design (presence of repeat motifs) or poor capture efficiency (mostly GC-rich sequences and first exons). Then, our protocol with initial barcoding of libraries followed by capture on pooled samples may be cost effective, but at present, limits the capture efficiency and needs further optimisation, especially if applied to larger gene sets. Finally, like for all targeted exon strategies, deep intronic mutations will be missed. The alternative to targeting entire genes would still miss a high proportion of intronic regions containing repetitive sequences, and would also disproportionately increase the number of rare variants to analyse with splice-prediction bioinformatic tools that are currently not highly reliable. For genes expressed in leucocytes or fibroblasts, another alternative would be selected RNA sequencing (enriched in cognate gene transcripts).

While similar capture strategies have been recently developed for other diseases, most of them included a much smaller cohort, and reported only *proof-of-principle* analysis,^{47–49} with the exception of Walsh *et al.*⁵⁰ Multiplex PCR approaches may have the potential of covering exons more exhaustively,⁴⁶ given that primer design is more flexible than hybridisation bait tiling, but is limited to smaller gene panels than for exon capture. PCR pooling without barcoding has been used for BBS and NPHP,^{12 51} a strategy which may be cost effective for analysing large numbers of samples in epidemiological studies, but appears unsuited for diagnosis where a key preoccupation is to limit false-positive/negative rates.

Regarding the specific case of BBS, our study suggests that when strict clinical criteria are complied with, the frequency of detected mutations is higher than the generally quoted 70% figure.^{7 11} There may thus be only few strict BBS genes remaining to be identified, especially considering that with most strategies, with the exception of whole-gene sequencing, one will miss deep intronic pathogenic mutations. Additional genes to be discovered may correspond rather to variant BBS-like phenotypes than to strictly defined BBS. Indeed, for patients with *BBS16/SDCCAG8* mutations, genotype-phenotype analysis showed for the first time a clear departure from the typical BBS phenotype with absence of polydactyly and systematic and severe renal manifestations (usually present in only 30% of BBS patients).^{7 8} Our finding of *ALMS1* mutations in three patients

Web resources

The URLs for online tools and data presented herein are:
 OMIM: <http://www.omim.org/>
 e-Array (Agilent Technologies): <https://earray.chem.agilent.com/earray/>
 UCSC: <http://genome.ucsc.edu/>
 SVA: www.svaproject.org/
 dbSNP: <http://www.ncbi.nlm.nih.gov/projects/SNP/>
 Mutation Nomenclature: <http://www.hgvs.org/mutnomen/recs.html>
 Exome Variant Server, NHLBI Exome Sequencing Project (ESP): <http://evs.gs.washington.edu/EVS/>
 Integrative Genomics Viewer (IGV): <http://www.broadinstitute.org/igv/>
 Polyphen 2: <http://genetics.bwh.harvard.edu/pph2/>
 Sift Human Protein: http://sift.jcvi.org/www/SIFT_enst_submit.html
 Mutation taster: <http://www.mutationtaster.org/>

confirms the major clinical overlap with BBS. Finally, we found no evidence for triallelism in our cohort of BBS patients.

Author affiliations

- ¹Institut de Génétique et de Biologie Moléculaire et Cellulaire (IGBMC), CNRS UMR7104, INSERM U964, Université de Strasbourg, Illkirch, France
- ²Microarrays and Sequencing Platform, IGBMC, Illkirch, France
- ³Laboratory of Human genetics, University of Medicine of Tunis, Tunis, Tunisia
- ⁴Bioinformatics Platform, IGBMC, Illkirch, France
- ⁵Laboratoire de Génétique Médicale EA3949 Inserm Avenir, Université de Strasbourg, Strasbourg, France
- ⁶Laboratoire de Diagnostic Génétique, Hôpitaux Universitaires de Strasbourg, Strasbourg, France
- ⁷Istituto di Genetica Medica, Università Cattolica, Roma, Italy
- ⁸CHU Bordeaux, University of Bordeaux, Department of Medical Genetics, Bordeaux, France
- ⁹Service de Génétique Clinique, Hôpital Jeanne de Flandre, CHRU de Lille, Lille, France
- ¹⁰Service de Cytogénétique Constitutionnelle, Hospices Civils de Lyon, CBPE, Bron Cedex, France
- ¹¹Department of Genetics, INSERM U676, Assistance Publique Hôpitaux de Paris (AP-HP), Robert Debré University Hospital, Paris, France
- ¹²Service de Génétique Médicale, Centre de Référence pour les Affections Rares en Génétique Ophtalmologique (CARGO), Hôpitaux Universitaires de Strasbourg, Strasbourg, France
- ¹³Chaire de Génétique Humaine, Collège de France, Illkirch, France

Acknowledgements We thank Ngoc-Hoan Nguyen for his help in the bioinformatics setup, and Cécile Pizot for the development of VaRank. We warmly thank Géraldine Greff, Anne-Sophie Jaeger, Manuela Antin, Elisabeth Scherrer, Serge Vicaire and Muriel Philipps for their technical assistance. Lastly, we wish to thank all patients and families included in this study, and our clinician colleagues who addressed patients for diagnostic analysis and provided clinical information, in particular: Drs P Abou-Jaoudé, C Baumann, H Flodrops, T Frébourg, D Genevieve, A Goldenberg, B Leheup, P Parent, P Petitjean, C Poitou-Bernert and S Taque.

Contributors CR, JLM and JM designed the study; CR, OM, CS, BJ performed experiments; CR, SLG, VG, JM performed the bioinformatics studies; CR, SLG, OM, VG, CS, JLM and JM collected and analysed data; OM, CS, MCV, PC, DL, IO, FP, MT, AV, HBC and HD provided DNA samples and clinical information; CR, JLM and JM wrote the manuscript; MCV, BJ, HBC, HD, JLM and JM provided technical support, conceptual advice and project coordination.

Funding This work was partially supported by a grant from Agence de Biomédecine to JLM and JM, by funds from APLM and by the Association Française contre les Myopathies (AFM) thanks to its support to the IGBMC sequencing platform.

Competing interests None.

Patient consent For children, consent was signed by the parents.

Provenance and peer review Not commissioned; externally peer reviewed.

REFERENCES

1. Zaghoul NA, Katsanis N. Mechanistic insights into Bardet-Biedl syndrome, a model ciliopathy. *J Clin Invest* 2009;**119**:428–37.

2. Hildebrandt F, Benzing T, Katsanis N. Ciliopathies. *N Engl J Med* 2011;**364**:1533–43.
3. Aliferis K, Helle S, Gyapay G, Duchatelet S, Stoetzel C, Mandel JL, Dollfus H. Differentiating alstrom from Bardet-Biedl syndrome (BBS) using systematic ciliopathy genes sequencing. *Ophthalmic Genet* 2011;**33**:18–22.
4. Karmous-Benailly H, Martinovic J, Gubler MC, Sirot Y, Clech L, Ozilou C, Auge J, Brahim N, Etchevers H, Detrait E, Esculapvit C, Audoulet S, Goudfroye G, Gonzales M, Tantau J, Loget P, Joubert M, Gaillard D, Jeanne-Pasquier C, Delezoide AL, Peter MO, Plessis G, Simon-Bouy B, Dollfus H, Le Merrer M, Munnich A, Encha-Razavi F, Vekemans M, Attie-Bitach T. Antenatal presentation of Bardet-Biedl syndrome may mimic Meckel syndrome. *Am J Hum Genet* 2005;**76**:493–504.
5. Pereira I, Hoskins BE, Marshall JD, Collin GB, Naggert JK, Pineiro-Gallego T, Oitmaa E, Katsanis N, Valverde D, Beales PL. Arrayed primer extension technology simplifies mutation detection in Bardet-Biedl and Alstrom syndrome. *Eur J Hum Genet* 2011;**19**:485–8.
6. 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**:437–46.
7. Deveault C, Billingsley G, Duncan JL, Bin J, Theal R, Vincent A, Fieggen KJ, Gerth C, Noordeh N, Traboulsi EI, Fishman GA, Chitayat D, Knueppel T, Millan JM, Munier FL, Kennedy D, Jacobson SG, Innes AM, Mitchell GA, Boycott K, Heon E. BBS genotype-phenotype assessment of a multiethnic patient cohort calls for a revision of the disease definition. *Hum Mutat* 2011;**32**:610–19.
8. Schaefer E, Zaloszcyc A, Lauer J, Durand M, Stutzmann F, Perdomo-Trujillo Y, Redin C, Bennouna Greene V, Vicaire S, Megarbane A, Kaplan J, Drouin-Garraud V, Hamdani M, Sigaudy S, Francannet C, Roume J, Bitoun P, Goldenberg A, Philip N, Odent S, Green J, Cossee M, Davis EE, Katsanis N, Bonneau D, Verloes A, Poch O, Mandel JL, Dollfus H. Identification of 28 novel mutations in the Bardet-Biedl syndrome genes: the burden of private mutations in an extensively heterogeneous disease. *Hum Genet* 2010;**127**:583–93.
9. Janssen S, Ramaswami G, Davis EE, Hurd T, Airik R, Kasanuki JM, Van Der Kraak L, Allen SJ, Beales PL, Katsanis N, Otto EA, Hildebrandt F. Mutation analysis in Bardet-Biedl syndrome by DNA pooling and massively parallel resequencing in 105 individuals. *Hum Genet* 2011;**129**:79–90.
10. Medvedev P, Stanciu M, Brudno M. Computational methods for discovering structural variation with next-generation sequencing. *Nat Methods* 2009;**6**(11 Suppl):S13–20.
11. Nord AS, Lee M, King MC, Walsh T. Accurate and exact CNV identification from targeted high-throughput sequence data. *BMC Genomics* 2011;**12**:184.
12. Anders S, Huber W. Differential expression analysis for sequence count data. *Genome Biol* 2010;**11**:R106.
13. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc Ser B (Methodological)* 1995;**57**:289–300.
14. Simms RJ, Hynes AM, Eley L, Sayer JA. Nephronophthisis: a genetically diverse ciliopathy. *Int J Nephrol* 2011;**2011**:527137.
15. Badano JL, Leitch CC, Ansley SJ, May-Simera H, Lawson S, Lewis RA, Beales PL, Dietz HC, Fisher S, Katsanis N. Dissection of epistasis in oligogenic Bardet-Biedl syndrome. *Nature* 2006;**439**:326–30.
16. Mykytyn K, Braun T, Carmi R, Haider NB, Searby CC, Shastri M, Beck G, Wright AF, Iannaccone A, Elbedour K, Riise R, Baldi A, Raas-Rothschild A, Gorman SW, Duhl DM, Jacobson SG, Casavant T, Stone EM, Sheffield VC. Identification of the gene that, when mutated, causes the human obesity syndrome BBS4. *Nat Genet* 2001;**28**:188–91.
17. Katsanis N, Ansley SJ, Badano JL, Eichers ER, Lewis RA, Hoskins BE, Scambler PJ, Davidson WS, Beales PL, Lupski JR. Triallelic inheritance in Bardet-Biedl syndrome, a Mendelian recessive disorder. *Science* 2001;**293**:2256–9.
18. Leitch CC, Zaghoul NA, Davis EE, Stoetzel C, Diaz-Font A, Rix S, Alfaridh M, Lewis RA, Eyaad W, Banin E, Dollfus H, Beales PL, Badano JL, Katsanis N. Hypomorphic mutations in syndromic encephalocoele genes are associated with Bardet-Biedl syndrome. *Nat Genet* 2008;**40**:443–8.
19. Kim SK, Shindo A, Park TJ, Oh EC, Ghosh S, Gray RS, Lewis RA, Johnson CA, Attie-Bittach T, Katsanis N, Wallingford JB. Planar cell polarity acts through septins to control collective cell movement and ciliogenesis. *Science* 2010;**329**:1337–40.
20. Otto EA, Hurd TW, Airik R, Chaki M, Zhou W, Stoetzel C, Patil SB, Levy S, Ghosh AK, Murga-Zamalloa CA, van Reeuwijk J, Letteboer SJ, Sang L, Giles RH, Liu Q, Coene KL, Estrada-Cuzcano A, Collin RW, McLaughlin HM, Held S, Kasanuki JM, Ramaswami G, Conte J, Lopez I, Washburn J, Macdonald J, Hu J, Yamashita Y, Maher ER, Guay-Woodford LM, Neumann HP, Obermuller N, Koeneke RK, Bergmann C, Bei X, Lewis RA, Katsanis N, Lopes V, Williams DS, Lyons RH, Dang CV, Brito DA, Dias MB, Zhang X, Cavalcoli JD, Nurnberg G, Nurnberg P, Pierce EA, Jackson PK, Antignac C, Saunier S, Roepman R, Dollfus H, Khanna H, Hildebrandt F.

- Candidate exome capture identifies mutation of SDCCAG8 as the cause of a retinal-renal ciliopathy. *Nat Genet* 2010;**42**:840–50.
24. **Schooser BG**, Frosk P, Engel AG, Klutzny U, Lochmuller H, Wrogemann K. Commonality of TRIM32 mutation in causing sarcotubular myopathy and LGMD2H. *Ann Neurol* 2005;**57**:591–5.
 25. **Bond J**, Flintoff K, Higgins J, Scott S, Bennet C, Parsons J, Mannon J, Jafri H, Rashid Y, Barrow M, Trembath R, Woodruff G, Rossa E, Lynch S, Sheilds J, Newbury-Ecob R, Falconer A, Holland P, Cockburn D, Karbani G, Malik S, Ahmed M, Roberts E, Taylor G, Woods CG. The importance of seeking ALMS1 mutations in infants with dilated cardiomyopathy. *J Med Genet* 2005;**42**:e10.
 26. **Marshall JD**, Hinman EG, Collin GB, Beck S, Cerqueira R, Maffei P, Milan G, Zhang W, Wilson DI, Hearn T, Tavares P, Vettor R, Veronese C, Martin M, So WV, Nishina PM, Naggert JK. Spectrum of ALMS1 variants and evaluation of genotype-phenotype correlations in Alstrom syndrome. *Hum Mutat* 2007;**28**:1114–23.
 27. **Minton JA**, Owen KR, Ricketts CJ, Crabtree N, Shaikh G, Ehtisham S, Porter JR, Carey C, Hodge D, Paisley R, Walker M, Barrett TG. Syndromic obesity and diabetes: changes in body composition with age and mutation analysis of ALMS1 in 12 United Kingdom kindreds with Alstrom syndrome. *J Clin Endocrinol Metab* 2006;**91**:3110–16.
 28. **Katsanis N**. The oligogenic properties of Bardet-Biedl syndrome. *Hum Mol Genet* 2004;**13**:R65–71.
 29. **Katsanis N**, Eichers ER, Ansley SJ, Lewis RA, Kayserili H, Hoskins BE, Scambler PJ, Beales PL, Lupski JR. BBS4 is a minor contributor to Bardet-Biedl syndrome and may also participate in triallelic inheritance. *Am J Hum Genet* 2002;**71**:22–9.
 30. **Hichri H**, Stoetzel C, Laurier V, Caron S, Sigaudy S, Sarda P, Hamel C, Martin-Coignard D, Gilles M, Leheup B, Holder M, Kaplan J, Bitoun P, Lacombe D, Verloes A, Bonneau D, Perrin-Schmitt F, Brandt C, Besancon AF, Mandel JL, Cossee M, Dollfus H. Testing for triallelism: analysis of six BBS genes in a Bardet-Biedl syndrome family cohort. *Eur J Hum Genet* 2005;**13**:607–16.
 31. **Laurier V**, Stoetzel C, Muller J, Thibault C, Corbani S, Jalkh N, Salem N, Chouery E, Poch O, Licaire S, Danse JM, Amati-Bonneau P, Bonneau D, Megarbane A, Mandel JL, Dollfus H. Pitfalls of homozygosity mapping: an extended consanguineous Bardet-Biedl syndrome family with two mutant genes (BBS2, BBS10), three mutations, but no triallelism. *Eur J Hum Genet* 2006;**14**:1195–203.
 32. **Smaoui N**, Chaabouni M, Sergeev YV, Kallel H, Li S, Mahfoudh N, Maazoul F, Kammoun H, Gandoura N, Bouaziz A, Nouiri E, M'Rad R, Chaabouni H, Hejtmancik JF. Screening of the eight BBS genes in Tunisian families: no evidence of triallelism. *Invest Ophthalmol Vis Sci* 2006;**47**:3487–95.
 33. **Nachury MV**, Loktev AV, Zhang Q, Westlake CJ, Peranen J, Merdes A, Slusarski DC, Scheller RH, Bazan JF, Sheffield VC, Jackson PK. A core complex of BBS proteins cooperates with the GTPase Rab8 to promote ciliary membrane biogenesis. *Cell* 2007;**129**:1201–13.
 34. **Seo S**, Baye LM, Schulz NP, Beck JS, Zhang Q, Slusarski DC, Sheffield VC. BBS6, BBS10, and BBS12 form a complex with CCT/TRiC family chaperonins and mediate BBSome assembly. *Proc Natl Acad Sci U S A* 2010;**107**:1488–93.
 35. **Badano JL**, Kim JC, Hoskins BE, Lewis RA, Ansley SJ, Cutler DJ, Castellani C, Beales PL, Leroux MR, Katsanis N. Heterozygous mutations in BBS1, BBS2 and BBS6 have a potential epistatic effect on Bardet-Biedl patients with two mutations at a second BBS locus. *Hum Mol Genet* 2003;**12**:1651–9.
 36. **Beales PL**, Badano JL, Ross AJ, Ansley SJ, Hoskins BE, Kirsten B, Mein CA, Froguel P, Scambler PJ, Lewis RA, Lupski JR, Katsanis N. Genetic interaction of BBS1 mutations with alleles at other BBS loci can result in non-Mendelian Bardet-Biedl syndrome. *Am J Hum Genet* 2003;**72**:1187–99.
 37. **Andersen KL**, Echwald SM, Larsen LH, Hamid YH, Glumer C, Jorgensen T, Borch-Johnsen K, Andersen T, Sorensen TI, Hansen T, Pedersen O. Variation of the McKusick-Kaufman gene and studies of relationships with common forms of obesity. *J Clin Endocrinol Metab* 2005;**90**:225–30.
 38. **Slavotinek AM**, Searby C, Al-Gazali L, Hennekam RC, Schrandt-Stumpel C, Orcana-Losa M, Pardo-Reoyo S, Cantani A, Kumar D, Capellini Q, Neri G, Zackai E, Biesecker LG. Mutation analysis of the MKKS gene in McKusick-Kaufman syndrome and selected Bardet-Biedl syndrome patients. *Hum Genet* 2002;**110**:561–7.
 39. **Stone DL**, Slavotinek A, Bouffard GG, Banerjee-Basu S, Baxevarian AD, Barr M, Biesecker LG. Mutation of a gene encoding a putative chaperonin causes McKusick-Kaufman syndrome. *Nat Genet* 2000;**25**:79–82.
 40. **Zaghoul NA**, Liu Y, Gerdes JM, Gascue C, Oh EC, Leitch CC, Bromberg Y, Binkley J, Leibel RL, Sidow A, Badano JL, Katsanis N. Functional analyses of variants reveal a significant role for dominant negative and common alleles in oligogenic Bardet-Biedl syndrome. *Proc Natl Acad Sci U S A* 2010;**107**:10602–7.
 41. **Choi M**, Scholl UI, Ji W, Liu T, Tikhonova IR, Zumbo P, Nayir A, Bakaloglu A, Ozen S, Sanjad S, Nelson-Williams C, Farhi A, Mane S, Lifton RP. Genetic diagnosis by whole exome capture and massively parallel DNA sequencing. *Proc Natl Acad Sci U S A* 2009;**106**:19096–101.
 42. **Lupski JR**, Reid JG, Gonzaga-Jauregui C, Rio Deiros D, Chen DC, Nazareth L, Bainbridge M, Dinh H, Jing C, Wheeler DA, McGuire AL, Zhang F, Stankiewicz P, Halperin JJ, Yang C, Gehman C, Guo D, Irikat RK, Tom W, Fantin NJ, Muzny DM, Gibbs RA. Whole-genome sequencing in a patient with Charcot-Marie-Tooth neuropathy. *N Engl J Med* 2010;**362**:1181–91.
 43. **Walker RH**, Schulz VP, Tikhonova IR, Mahajan MC, Mane S, Arroyo Muniz M, Gallagher PG. Genetic diagnosis of neuroacanthocytosis disorders using exome sequencing. *Mov Disord* 2011;**27**:539–43.
 44. **Hopp K**, Heyer CM, Hommerding CJ, Henke SA, Sundsbak JL, Patel S, Patel P, Consugar MB, Czarnecki PG, Gliem TJ, Torres VE, Rossetti S, Harris PC. B9D1 is revealed as a novel Meckel syndrome (MKS) gene by targeted exon-enriched next-generation sequencing and deletion analysis. *Hum Mol Genet* 2011;**20**:2524–34.
 45. **Jones MA**, Bhide S, Chin E, Ng BG, Rhodenizer D, Zhang VW, Sun JJ, Tanner A, Freeze HH, Hegde MR. Targeted polymerase chain reaction-based enrichment and next generation sequencing for diagnostic testing of congenital disorders of glycosylation. *Genet Med* 2011;**13**:921–32.
 46. **Mondal K**, Shetty AC, Patel V, Cutler DJ, Zwick ME. Targeted sequencing of the human X chromosome exome. *Genomics* 2011;**98**:260–5.
 47. **Brownstein Z**, Friedman LM, Shahin H, Oron-Karni V, Kol N, Rayyan AA, Parzefall T, Lev D, Shalev S, Frydman M, Davidov B, Shohat M, Rahile M, Lieberman S, Levy-Lahad E, Lee MK, Shomron N, King MC, Walsh T, Kanaan M, Avraham KB. Targeted genomic capture and massively parallel sequencing to identify genes for hereditary hearing loss in middle eastern families. *Genome Biol* 2011;**12**:R89.
 48. **Meder B**, Haas J, Keller A, Heid C, Just S, Borries A, Boisguerin V, Scharfenberger-Schmeer M, Stahler P, Beier M, Weichenhan D, Strom TM, Pfeufer A, Korn B, Katus HA, Rottbauer W. Targeted next-generation sequencing for the molecular genetic diagnostics of cardiomyopathies. *Circ Cardiovasc Genet* 2011;**4**:110–22.
 49. **Shearer AE**, DeLuca AP, Hildebrand MS, Taylor KR, Gurrilo J 2nd, Scherer S, Scheetz TE, Smith RJ. Comprehensive genetic testing for hereditary hearing loss using massively parallel sequencing. *Proc Natl Acad Sci U S A* 2010;**107**:21104–9.
 50. **Walsh T**, Lee MK, Casadei S, Thornton AM, Stray SM, Pennil C, Nord AS, Mandell JB, Swisher EM, King MC. Detection of inherited mutations for breast and ovarian cancer using genomic capture and massively parallel sequencing. *Proc Natl Acad Sci U S A* 2010;**107**:12629–33.
 51. **Otto EA**, Ramaswami G, Janssen S, Chaki M, Allen SJ, Zhou W, Airik R, Hurd TW, Ghosh AK, Wolf MT, Hoppe B, Neuhaus TJ, Bockenauer D, Milford DV, Soliman NA, Antignac C, Saunier S, Johnson CA, Hildebrandt F. Mutation analysis of 18 nephronophthisis associated ciliopathy disease genes using a DNA pooling and next generation sequencing strategy. *J Med Genet* 2011;**48**:105–16.