

SUPPLEMENTARY TABLES

Supplementary Table S1. Clinical characteristics of patients with Alport syndrome.

Symptoms	Patient 1	Patient 2 (the biological brother of patient)	Patient 3	Patient 4
Gender	female	male	male	female
Age at examination	8 years	13 years	7 years	13 years
Age at disease onset	6 years	1,5 years	7 years	7 years
Family history	The biological siblings have a maternal uncle with the same condition.		A maternal uncle is affected by a similar condition.	The mother has chronic kidney disease stage 1 (CKD1) and is receiving dialysis.
Initial symptoms	Microscopic haematuria (10–15 red blood cells per field of view), Proteinuria (protein 0.33 g/L).	Microscopic haematuria (25–30 red blood cells per field of view) during an episode of obstructive bronchitis. From the age of 3 years, gross haematuria. Proteinuria (protein 2.3 g/L).	Gross haematuria (numerous red blood cells), Proteinuria up to 0.33–0.44 g/L.	Gross haematuria (30–40 red blood cells per field of view), Proteinuria (protein 0.3 g/L). Facial and lower limb edema
Referral diagnosis:	Hereditary nephritis. Chronic kidney disease, stage G1A2.	Hereditary nephropathy, not classified elsewhere.	Nephritic syndrome. Hereditary nephritis.	Chronic nephritic syndrome, CKD stage 1.
Final diagnosis:	Hereditary nephropathy. Alport syndrome, X-linked type.	Hereditary nephropathy. Alport syndrome, X-linked type	Nephritic syndrome.	Chronic nephritic syndrome. Membranoproliferative glomerulonephritis with moderate fibrosis and lymphocytic infiltration.
Renal function impairment				
Haematuria	+	+	+	+
Proteinuria	+	+	+	+
Sensorineural hearing loss	Not observed	Not observed	Not observed	Not observed
Ophthalmologic abnormalities	Not observed	Retinal angiopathy	Not observed	Not observed

Ultrasound of the kidneys	Bilateral thickening of the renal pelvicalyceal system walls.	Not detected	Bilateral nephroptosis, bilateral thickening of the renal pelvicalyceal system walls, and bilateral calyceal ectasia.	Not detected
Kidney biopsy	Focal global and segmental glomerulosclerosis. Interstitial fibrosis and tubular atrophy, grade 1. Electron microscopy: Findings consistent with a congenital/hereditary type IV collagen disorder.	Focal global and segmental glomerulosclerosis. Interstitial fibrosis and tubular atrophy, grade 1. Electron microscopy: Findings consistent with a congenital/hereditary type IV collagen disorder.	Not performed	Morphological features of membranoproliferative glomerulonephritis with moderate fibrosis and lymphocytic infiltration of the interstitial tissue (focal sclerosis – 85.2%, hyalinosis and sclerosis – 14.8%).
Complete blood count (CBC)	Haemoglobin: 131 g/L; Red blood cells: $4.93 \times 10^{12}/L$; Platelets: $365 \times 10^9/L$; White blood cells: $6.71 \times 10^9/L$; ESR: 4 mm/h.	Haemoglobin: 123 g/L; Red blood cells: $4.36 \times 10^{12}/L$; Platelets: $341 \times 10^9/L$; White blood cells: $7.71 \times 10^9/L$; ESR: 10 mm/h.	White blood cells: $6.08 \times 10^9/L$; Red blood cells: $4.8 \times 10^{12}/L$; Haemoglobin: 132 g/L; Platelets: $277 \times 10^9/L$; ESR: 25 mm/h.	No data
Urinalysis	Colour: straw-yellow; clarity: clear; specific gravity: 1.003; pH: 6.5; protein: negative; Leukocytes: $2/\mu L$; Erythrocytes: $80/\mu L$.	Protein: 2.0; Erythrocytes: $40/\mu L$; Protein: 2.31 g/L; Erythrocytes: 15–20 per field of view; Leukocytes: 8–12 per field of view; Colour: “meat-wash” (dark red/brown); Specific gravity: 1.002; pH: 7.0.	Colour: straw-yellow; clarity: clear; specific gravity: 1.030; Protein: 0.264 g/L; Leukocytes: numerous; Erythrocytes (unchanged): 10–15 per field of view; Reaction (pH): acidic.	No data

Blood biochemistry panel	Total protein: 61 g/L; ALT: 10.6 U/L; AST: 21 U/L; Urea: 3.5 mmol/L; Creatinine: 44.7 µmol/L; Glucose: 4.5 mmol/L; Total bilirubin: 4.8 µmol/L.	Total protein: 49 g/L; ALT: 7.6 U/L; AST: 19.6 U/L; Urea: 8.2 mmol/L; Creatinine: 69 µmol/L; Glucose: 4.6 mmol/L; Total bilirubin: 3 µmol/L.	Creatinine: 46.8 µmol/L; AST: 44.0 U/L; Total protein: 65.0 g/L; Total bilirubin: 11.2 µmol/L; Albumin: 39 g/L; Total cholesterol: 3.3 mmol/L; Glucose: 3.9 mmol/L; Creatinine: 34.3 µmol/L; Urea: 3.53 mmol/L; ALT: 16.0 U/L.	Total protein: 48–60 g/L; Albumin: 28–40 g/L; Total cholesterol: 2.4–9.79 mmol/L. ANA and ANCA: within normal limits; ASO: 56 IU/L; Complement C3: 0.1 g/L (reference range 0.9–1.7); Complement C4: 0.23 g/L (reference range 0.5–0.9).
24-hour urinary protein excretion	Protein- 0,2475 g/day; 1,32 g/day.	Protein-0,528 g/L. 0,436 g/L; 1,815g/day	Protein- 2,31 g/day	Protein-0,231 g/day
Estimated glomerular filtration rate (eGFR) by the Schwartz formula	GFR- 129ml/min/1,73m ²	GFR-82 mL/min/1,73 m ² ; 47 mL/min/1,73m ²	Not determined	GFR 96 mL/min/1,73m ²
Chronic kidney failure (CKF)	Not observed	CKD stage 1 for 12 years	Not observed	Not observed

Supplementary Table S2. Optional parameters for design.

Parameter	Description	Example
Basic Parameters		
minPCR	Minimum amplicon size (bp)	minPCR=250
maxPCR	Maximum amplicon size (bp)	maxPCR=500
minLen	Minimum primer length (nt)	minLen=18
maxLen	Maximum primer length (nt)	maxLen=24
minTm	Minimum melting temperature (°C)	minTm=60
maxTm	Maximum melting temperature (°C)	maxTm=62
minLC	Minimum linguistic complexity (%)	minLC=75
maskedBases	Repeat-masked positions tolerated inside a primer (0-40)	maskedBases=0
repeatKmer	k-mer size for de novo repeat detection	repeatKmer=32
repeatMinCount	Minimum occurrence count for de novo repeat detection	repeatMinCount=2
3end	3' end constraint	3end=w
5end	5' end constraint	5end=n
Forwardtail=	5' adapter for forward primers	Illumina P5 adapter forwardtail=TCGTCGGCAGCGTCAGATGTGTATAAGAGACAG
Reversetail=	5' adapter for reverse primers	Illumina P7 adapter reversetail=GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAG
Input/Output Paths		
target_path	Path to individual target file (can be repeated)	target_path=D:\Panels\MODY\NG_008690.gb
target_primers	Path to existing primer/probe list	target_primers=C:\PCRpanel\test\primers.txt
folder_path	Path to folder containing multiple target files	folder_path=E:\genes\
folder_out	Output directory for results	folder_out=E:\report\

Using GenBank/RefSeq and FASTA inputs, and automated sequence retrieval for large-scale PCR design

PCRpanel supports two complementary classes of input sequence sources: annotated GenBank/RefSeq records and FASTA sequences. This dual-mode input strategy enables both

annotation-aware primer design (e.g., exon- and promoter-centric panels) and annotation-free design for arbitrary targets, whole genes, or whole genomes.

GenBank/RefSeq records (annotation-aware mode). When a GenBank flat file is provided (e.g., RefSeqGene NG_, genomic NC_, or transcript records NM_* / NR_*), PCRpanel can parse the FEATURES table and the underlying nucleotide sequence (from the ORIGIN section). This allows the pipeline to derive biologically meaningful target intervals directly from standard annotations—most importantly exon blocks (often encoded via join(...) within mRNA, ncRNA, or CDS features) and upstream regulatory segments such as promoters. In practice, GenBank/RefSeq input is preferable whenever exon/promoter boundaries and transcript structure are important for the design objective, because primer placement can be constrained to specific annotated intervals rather than the full locus.

FASTA sequences (sequence-only mode). PCRpanel also accepts FASTA as a lightweight input format when annotation is not required or not available. Both single-record and multi-record FASTA files are supported, and each FASTA record can be treated as an independent design target. This mode is useful for curated target regions (e.g., consensus sequences, assembled contigs, pathogen genomes, synthetic constructs), for rapid prototyping, and for workflows where annotations are external or intentionally ignored. FASTA input therefore provides a universal fallback that keeps the primer-design workflow usable even in the absence of reliable feature annotation.

Automated retrieval from NCBI by gene symbols or RefSeq accessions. To scale from pilot panels to large gene sets, PCRpanel can be coupled with an automated sequence retrieval step from NCBI GenBank/RefSeq. Two common entry points are supported in practice:

Gene symbol lists (e.g., TP53, BRCA1, EGFR) together with an organism identifier (recommended: TaxID). Gene symbols can be resolved to NCBI GeneIDs, which are then used to discover linked RefSeq accessions for genomic and transcript records (e.g., NG_* for RefSeqGene loci, optionally NM_*/NR_** for transcript sequences).

Direct accession lists, most commonly RefSeqGene NG_* identifiers, which can be retrieved without symbol resolution.

A typical high-throughput pipeline is: (i) resolve gene symbols to GeneIDs (if needed), (ii) collect the relevant RefSeq accessions, (iii) download GenBank flat files (preferably with full annotations and sequence), (iv) run PCRpanel in batch mode over the downloaded folder to automatically extract target intervals (exons/promoters) and design primers. This “download → parse annotations → extract targets → design” workflow reduces manual sequence collection, improves reproducibility, and enables fully automated re-analysis when gene lists, organisms, or design constraints change.

Target Region Detection

PCRpanel designs primers **only within target coordinates**. These can be derived automatically from GenBank annotations or specified as arbitrary regions.

Exon Recognition Priority

Many RefSeq/GenBank records lack explicit exon features. PCRpanel uses this recognition hierarchy:

Priority	Feature Type	Notes
1	exon	Used directly if present
2	mRNA, ncRNA, rRNA, tRNA	join(...) blocks extracted as exons
3	CDS	join(...) blocks used when transcript features absent
4	Fallback	Full sequence treated as one target region

Example Configuration File:

```
# Target sequences (multiple files supported)
target_path=C:\PCRpanel\test\NG_008690.txt
target_path=C:\PCRpanel\test\NG_011731.txt
target_path=C:\PCRpanel\test\NG_013019.txt
target_path=C:\PCRpanel\test\NG_008847.txt
target_path=C:\PCRpanel\test\NC_000002.txt

# Optional: existing primers to incorporate
target_primers=C:\PCRpanel\test\primers.txt

# Amplicon size constraints
minPCR=250
maxPCR=500

# Primer parameters
minLen=18
maxLen=24
minTm=60
maxTm=62
minLC=75
maskedBases=0
repeatKmer=32
repeatMinCount=2

# Primer end constraints
3end=w
5end=

# Adapter tails (Illumina example)
forwardtail=ACACTCTTTCCCTACACGACGCTCTTCCGATCT
reversetail=GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT
```

SUPPLEMENTARY FIGURES

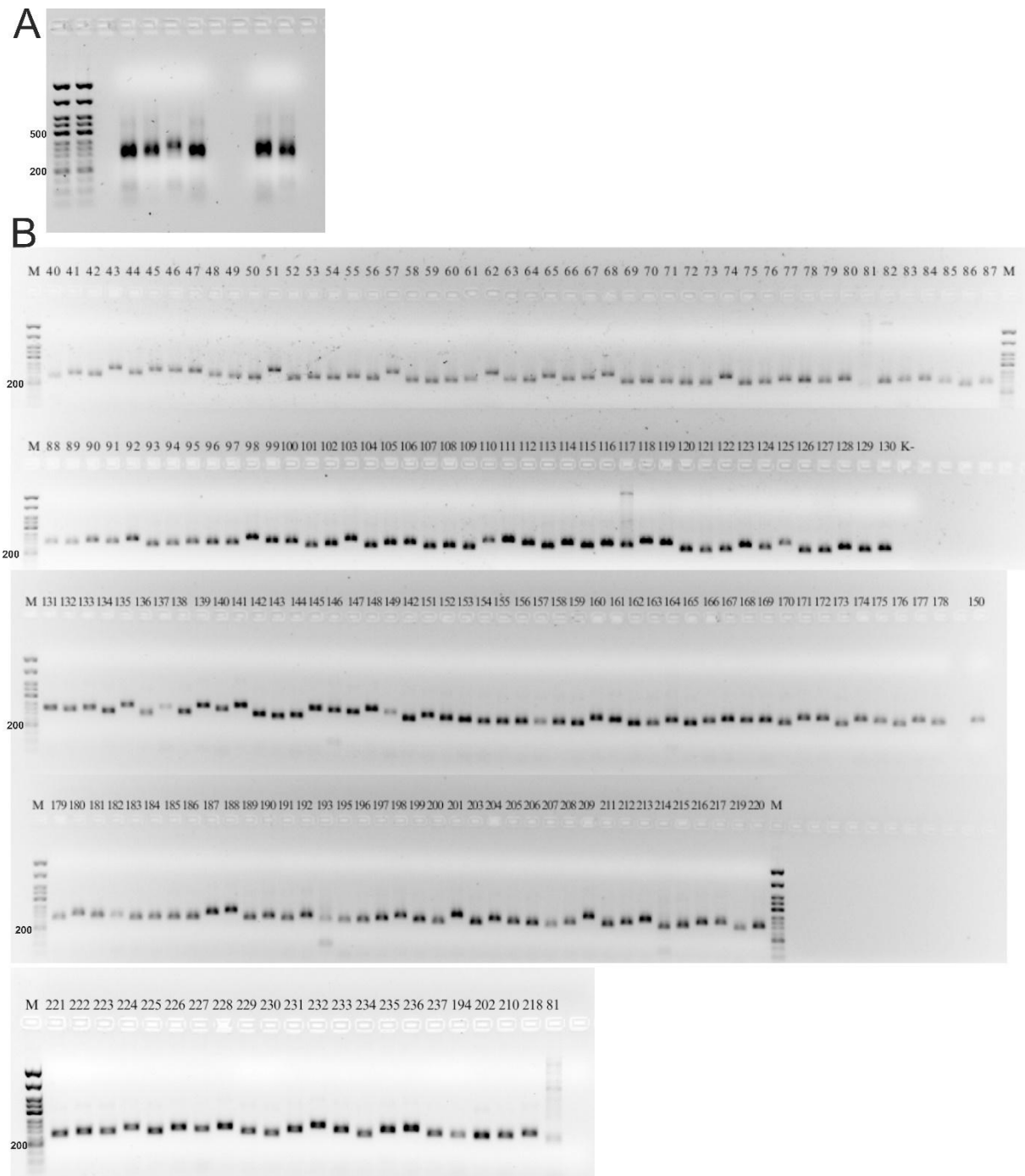


Figure S1. Amplicon validation by PCR amplification. (A) Multiplex PCR for individual pools analyzed by agarose gel electrophoresis. Each lane represents a separate pool, with bands indicating successful amplification of target regions ranging from 240-450 bp. Lane M: DNA size marker (50 bp ladder). (B) Analysis of individual PCR pairs after multiplex PCR for each of the pools. Each lane corresponds to a single primer pair tested separately, confirming specific amplicon generation with expected fragment sizes (indicated above each lane) and optimal band intensity for downstream next-generation sequencing applications. Lane M: DNA size marker.