3' UTR structural elements in *CYP24A1* are associated with Infantile hypercalcaemia type 1

Nicole Ball[1], Susan Duncan[2], Yueying Zhang[2], Rocky Payet[3], Isabelle Piec[1, 4], Eloise Whittle[1], Jonathan C. Y. Tang[1, 4], Inez Schoenmakers[1], Berenice Lopez[4], Allison Chipchase[4], Arun Kumar[5], Leslie Perry[6], Heather Maxwell[7], Yiliang Ding[2], William D. Fraser[1, 4]*, Darrell Green[1]*

1. Biomedical Research Centre, Norwich Medical School, University of East Anglia, Norwich Research Park, Norwich, UK
2. Cell and Developmental Biology, John Innes Centre, Norwich Research Park, Norwich, UK
3. School of Biological Sciences, University of East Anglia, Norwich Research Park, Norwich, UK
4. Clinical Biochemistry, Diabetes and Endocrinology, Norfolk and Norwich University Hospital, Norwich Research Park, Norwich, UK
5. Paediatrics, Croydon University Hospital, Croydon, UK
6. Clinical Biochemistry, Croydon University Hospital, Croydon, UK
7. Renal Department, Royal Hospital for Children, Glasgow, UK

*Joint corresponding authors: Professor William D. Fraser (w.fraser@uea.ac.uk); Dr Darrell Green (d.green@uea.ac.uk)

ABSTRACT

Loss-of-function mutations in the *CYP24A1* protein-coding region causing reduced 25OHD and 1,25$(OH)_2$D catabolism have been observed in some cases of Infantile hypercalcaemia type 1 (HCINF1), which can manifest as nephrocalcinosis, hypercalcaemia and adult-onset hypercalciuria and renal stone formation. Some cases present with apparent *CYP24A1* phenotypes but do not exhibit pathogenic mutations. Here, we assessed the molecular mechanisms driving apparent HCINF1 where there was a lack of *CYP24A1* mutation. We obtained blood samples from 47 patients with either a single abnormality of no obvious cause or a combination of hypercalcemia, hypercalciuria and nephrolithiasis as part of our metabolic and stone clinics. We used liquid chromatography tandem mass spectrometry (LC-MS/MS) to determine serum vitamin D metabolites and direct sequencing to confirm *CYP24A1* genotype. Six patients presented with profiles characteristic of altered CYP24A1 function but lacked protein-coding mutations in *CYP24A1*. Analysis up- and downstream of the coding sequence showed single nucleotide variants (SNVs) in the *CYP24A1* 3' untranslated region (UTR). Bioinformatics approaches revealed that these 3' UTR abnormalities did not result in microRNA silencing but altered the *CYP24A1* messenger RNA (mRNA) secondary structure, which negatively impacted translation. Our experiments showed that mRNA misfolding driven by these 3' UTR sequence-dependent structural elements was associated with normal 25OHD but abnormal 1,25$(OH)_2$D catabolism. Using CRISPR-Cas9, we developed an *in vitro* mutant model for future *CYP24A1* studies. Our results form a basis for future studies investigating structure-function relationships and novel *CYP24A1* mutations producing a semi-functional protein.

Keywords: CYP24A1, vitamin D, renal, bone, mRNA, 3' UTR

ABBREVIATIONS

1,24,25(OH)$_3$D 1,24,25 trihydroxyvitamin D

1,25(OH)$_2$D 1,25 dihydroxyvitamin D

3' UTR 3 prime untranslated region

4PL four-parameter logistic

24,25(OH)$_2$D 24,25 dihydroxyvitamin D

25OHD 25 hydroxyvitamin D

aCa albumin-adjusted calcium

Ca total calcium

CCD charge-coupled device

CYP2R1 cytochrome P450 family 2 subfamily R member 1

CYP24A1 cytochrome P450 family 24 subfamily A member 1

CYP27B1 cytochrome P450 family 27 subfamily B member 1

dPCR digital PCR

ECLIA electrochemiluminescence immunoassay

FGF23 Fibroblast growth factor 23

gRNA guide RNA

HCINF1 Infantile hypercalcaemia type 1

LC-MS/MS liquid chromatography-tandem mass spectrometry

MFE minimum free energy

miRNA microRNA

MRE microRNA recognition element

mRNA messenger RNA

PBMCs Peripheral blood mononuclear cells

POLR2A RNA polymerase II subunit A

PTH parathyroid hormone

smFISH single molecule fluorescence *in situ* hybridisation

SNV single nucleotide variant

VMRs Vitamin D metabolite relative ratios

INTRODUCTION

Vitamin D plays a key role in classical calciotropic processes including calcium and bone metabolism and is postulated to contribute to non-classical disorders including cancer, diabetes, multiple sclerosis and coronavirus disease (COVID-19)[1–3]. Vitamin D is obtained from either the diet (ergocalciferol) or in the skin (cholecalciferol) after photochemical conversion of 7-dehydrocholesterol. Vitamin D is transported to the liver where it is hydroxylated by cytochrome P450 family 2 subfamily R member 1 (CYP2R1) to form 25 hydroxyvitamin D (25OHD). A further hydroxylation in the kidney by cytochrome P450 family 27 subfamily B member 1 (CYP27B1) generates the active systemic metabolite 1,25 dihydroxyvitamin D ($1,25(OH)_2D$) essential for calcium homeostasis. Vitamin D metabolism is regulated by the activity of CYP27B1 and cytochrome P450 family 24 subfamily A member 1 (CYP24A1). CYP24A1 converts the precursor 25OHD into 24,25 dihydroxyvitamin D ($24,25(OH)_2D$) and $1,25(OH)_2D$ into 1,24,25 trihydroxyvitamin D ($1,24,25(OH)_3D$). Both $24,25(OH)_2D$ and $1,24,25(OH)_3D$ are subject to further hydroxylation followed by bile and urinary excretion to prevent vitamin D toxicity. CYP27B1 and CYP24A1 activity is controlled by $1,25(OH)_2D$, calcium, parathyroid hormone (PTH) and fibroblast growth factor 23 (FGF23). Vitamin D metabolite relative ratios (VMRs) such as $25OHD:24,25(OH)_2D$ and $1,25(OH)_2D:24,25(OH)_2D$ are critical in the differential diagnosis of the vitamin D hydroxylation pathways[4].

Vitamin D toxicity and/or sensitivity manifesting as hypercalcemia, hypercalciuria and/or nephrolithiasis caused by *CYP24A1* loss-of-function mutations resulting in elevated serum

105 $1,25(OH)_2D$[5] is a disorder known as Infantile hypercalcaemia type 1 (HCINF1, OMIM #143880). Type
106 2 (HCINF2, OMIM #616963) differs in that there is a loss-of-function in solute carrier family 34 member
107 1 (*SLC34A1*) and the phenotype includes hypophosphatemia. HCINF1 infant presentation includes
108 vomiting, failure to thrive, colic and in rare cases death. HCINF1 adult presentation can include flu-like
109 symptoms, hypercalciuria and renal stone formation. In some female patients these symptoms are
110 triggered in pregnancy, likely uncovered by vitamin D supplementation rather than the pregnancy
111 itself[6]. A future discussion required in the field is the removal of the word "infantile" from HCINF1
112 given the increasing adult presentation frequency.
113
114 In the current study we used the VMR to select patients with suspected HCINF1 for genetic analysis to
115 confirm the diagnosis as part of their care in metabolic and stone former clinics. We also investigated
116 2 paediatric patients referred for investigation of nephrocalcinosis and calcium metabolism
117 abnormalities. Through *CYP24A1* direct sequencing, we identified in the 2 children and 4 out of 47
118 adults (8.5%) single nucleotide variants (SNVs) in the 3' untranslated region (UTR) with unknown
119 clinical significance, which prompted non-clinical studies described here. The 3' UTR is of significant
120 messenger RNA (mRNA) regulatory importance plus single-stranded RNAs fold into complex three-
121 dimensional structures that are critical for their function/regulation including post-transcriptional
122 modification, nuclear export, cellular localisation, translation and degradation[7–12]. We hypothesised
123 that SNVs in the 3' UTR affected the *CYP24A1* mRNA secondary structure with mRNA misfolding
124 leading to the heterogeneous phenotypes observed in some HCINF1 cases. After performing
125 bioinformatics and computational modelling to demonstrate mRNA structural abnormalities in these
126 patients, we generated a CRISPR-Cas9 mutant HEK293T cell line to mimic patients with a *CYP24A1* 3'
127 UTR variant. This mutant model provides a tool for *in vitro* investigation into non-canonical and
128 pathogenic *CYP24A1* phenotypes.
129
130 **MATERIALS AND METHODS**

Clinical samples. Forty-seven patient blood samples were collected as part of routine requests for 25OHD LC-MS/MS analysis from the Department of Laboratory Medicine at the Norfolk and Norwich University Hospital between 2016 and 2017. Patients were referred from the metabolic or stone former clinics. There were 21 patients with a diagnosis of primary hyperparathyroidism including 5 who formed renal stones. There were 8 patients with osteoporosis including 2 who had high urinary calcium excretion. There were 5 patients who had hypercalcaemia of non-parathyroid and non-malignant origin. Eight patients were being investigated for renal stones, 3 were hypoparathyroid and 2 were under investigation for secondary hyperparathyroidism. Blood samples were collected into serum gel separator tubes (BD Vacutainer) and immediately centrifuged. The serum layer was aliquoted and stored at -20 °C. Two samples from children were referred to the regional metabolic biochemistry service for investigation of causes of nephrocalcinosis and abnormal calcium biochemistry.

Biochemical analysis. Serum vitamin D metabolites 25OHD and 24,25$(OH)_2$D were measured simultaneously by liquid chromatography tandem mass spectrometry (LC-MS/MS) using the Micromass Quattro Ultima Pt electrospray ionisation (ESI) tandem mass spectrometer (Waters Corp., Milford, MA, USA) as previously described[4]. MassLynx v4.1 and QuanLynx (Waters Corp., Milford, MA, USA) were used for assay performance, system control, data acquisition, baseline integration and peak quantification. The LIAISON® XL 1,25$(OH)_2$D chemiluminescent immunoassay (DiaSorin, Saluggia, Italy) method was used to measure 1,25$(OH)_2$D in serum samples. This sandwich assay utilises a recombinant fusion protein for 1,25$(OH)_2$D capture and a murine monoclonal antibody detection system. The assay measures total 1,25$(OH)_2$D between 12–480 pmol/L. The inter/intra-assay CV was ≤9.2% and the mean assay recovery was 94 ± 2% across the analytical range. Cell line 1,25$(OH)_2$D catabolism was measured by LC-MS/MS using the Xevo TQ-XS (Waters Corp., Milford, MA, USA) as previously described[13]. Intact PTH and albumin-adjusted calcium (ACa) were analysed on the

157 COBAS® 6000 (Roche Diagnostics, Burgess Hill, UK) platform. EDTA-plasma PTH was measured using
158 electrochemiluminescence immunoassay (ECLIA). The inter-assay CV was ≤3.8% across the analytical
159 range of 1.2–5,000 pg/mL. Total calcium (Ca) and albumin were measured using spectrophotometric
160 methods. The inter-assay CV for Ca was ≤1.6%, albumin was ≤1.1% across the working ranges of the
161 assays.
162
163 **Study participants.** The University of East Anglia Faculty of Medicine and Health Sciences Research
164 Ethics Committee approved the collection and study of DNA samples for non-clinical procedures (ref:
165 ETH2122-1884), i.e. the 4 adult and 2 paediatric patients. We obtained whole blood from the
166 metabolic and stone clinics for genetic analysis from patients presenting with elevated $1,25(OH)_2D$
167 and high/low VMR plus HCINF1 clinical presentation and who were negative for protein-coding
168 mutations (Patients 1-4) (Table 1). We obtained genomic DNA from an infant presenting with Williams
169 syndrome and nephrocalcinosis (Patient 5, Croydon Hospital) and an infant presenting with
170 nephrocalcinosis and polyuria (Patient 6, Royal Hospital for Children) (Table 1). All adults or infant
171 parents/guardians provided written informed consent to donate samples for this study. Anonymised
172 negative control blood samples were collected at the Norfolk and Norwich University Hospital blood
173 typing service (n=10). Exclusion criteria for the negative control samples were those with a vitamin D,
174 calcium or metabolic clinical history.
175
176 **DNA and RNA extraction.** Genomic DNA was isolated from nucleated blood cells using the Purelink
177 genomic DNA kit (Invitrogen, Massachusetts, USA). Total RNA was extracted using the miRNeasy mini
178 kit (Qiagen, Manchester, UK). DNA and RNA concentration and integrity was measured on the
179 NanoDrop 8000 (Thermo Fisher Scientific, Loughborough, UK). DNA was stored at -20 °C. RNA was
180 stored at -80 °C.
181

Sequencing and variant calling. *CYP24A1* direct sequencing was performed using primers as previously described[14]. For direct sequencing analysis we aligned FASTA reads with *CYP24A1* transcript variant 1 (NM_000782.5) using Sequence Alignment Editor (BioEdit). MutationTaster, ClinVar and dbSNP were used to assess SNV disease-causing potential.

Bioinformatics. RNAfold[15] was used to determine the predicted *CYP24A1* mRNA structures. For mRNA secondary structure prediction, RNAfold provides three dynamic programming algorithms including (i) minimum free energy (MFE), which generates a single optimal structure using thermodynamic predictions based on the MFE generated by the nucleotide composition of the input sequence. (ii) Partition function that calculates base pair probabilities in the thermodynamic ensemble. (iii) Suboptimal folding that generates all suboptimal structures within a given energy range of the optimal energy. For quantifying mRNA secondary structure comparison (to the control/wild type), the package contains several measures of distance (i.e. dissimilarities) using either string alignment or tree-editing. RNAfold performance was extensively tested and validated by comparing MFE predictions between RNAfold 1.8.5, RNAfold 2.1.8, UNAfold 3.8 and RNAstructure 5.7 including accuracy; sensitivity, positive predictive value, Phi coefficient and F-measure. The test set was based on 1,919 non-multimer sequence/structure pairs obtained from the RNAstrand database (all without pseudoknots in the reference structure). Both versions of RNAfold were run with -d2 option whereas UNAfold and RNAstructure were run with default options[16]. RBPmap[17] was used to determine whether the 5' and 3' UTR mutations impaired protein-RNA interaction. RBPmap employs an algorithm for mapping protein binding motifs on RNA transcripts whilst considering the motif clustering propensity and the overall tendency of the regulatory region to be conserved. miRDB[18,19] was used to elucidate whether the 3' UTR variants altered or introduced microRNA (miRNA) recognition elements (i.e. target sites). miRDB is a database for miRNA target prediction and functional annotation. All mRNA targets in miRDB were predicted by miRTarget, which was developed by

analysing thousands of miRNA-mRNA functional interactions from next generation sequencing studies.

Digital PCR (dPCR). Total RNA was quantified by density measurement after separation by agarose gel electrophoresis with ethidium bromide staining. Equal RNA amounts were reverse transcribed using the high-capacity RNA to cDNA kit (Thermo Fisher Scientific, Loughborough, UK). *CYP24A1* transcript expression was quantified in triplicate using a TaqMan gene expression assay (Thermo Fisher Scientific, Loughborough, UK). dPCR was performed on the QuantStudio 3D Digital PCR System using the GeneAmp PCR System 9700 (Thermo Fisher Scientific, Loughborough, UK). After PCR, chips were imaged on the QuantStudio 3D instrument to convert raw data into the concentration of the cDNA sequence targeted by FAM and VIC labelled probes according to Poisson distribution[20]. The QuantStudio 3D AnalysisSuite was used to convert the data into copies per µL. Experiments were performed in triplicate in three independent experiments.

CYP24A1 protein expression. Peripheral blood mononuclear cells (PBMCs) were separated from whole EDTA blood samples by Ficoll-Paque PLUS (GE Life Sciences). PBMCs were lysed using M-PER lysis buffer (Thermo Fisher Scientific, Loughborough, UK) with complete protease inhibitor cocktail (Roche) and were clarified by centrifugation. Protein concentrations were determined from the supernatants using the BCA protein assay system (Thermo Fisher Scientific, Loughborough, UK). Proteins were separated on a precast 4-12% gradient SDS-PAGE gel (Thermo Fisher Scientific, Loughborough, UK) and transferred onto immobilon PVDF (Millipore, Watford, UK) to blot. CYP24A1 monoclonal antibody #WH0001591M7 (Sigma Aldrich, Gillingham, UK) was used to probe membranes for 48 h at 4 °C. Actin and GAPDH were included as loading controls. IRDye labelled secondary antibodies were used to detect primary antibodies. Proteins were visualised using the Odyssey infrared system (LI-COR, Cambridge, UK). CYP24A1 was also quantified by ELISA (Cusabio, Houston, USA). The detection range of this sandwich assay was 7.8-500 pg/mL and the sensitivity was 1.95

pg/mL. The inter/intra-assay CV was ≤10% and the mean assay recovery was 86 ± 6% across the analytical range. Sample absorbance was detected at 450 nm using a plate reader. Concentrations were calculated using a four-parameter logistic (4PL) curve ranging 0-500 pg/mL. ELISAs were performed in duplicate according to manufacturer's instructions.

CRISPR-Cas9. Plasmid guide RNA (gRNA) Cas9 constructs based on the vector GE100002 and donor 100 bp ssDNA oligos (Origene, Herford, Germany) were designed to introduce *CYP24A1* 3' UTR alterations (Suppl. File 1). To anneal forward/reverse ssDNA oligos, 2 µg each oligo were combined with 50 µL annealing buffer (10 nM Tris, 1 mM EDTA, 50 mM NaCl), heated at 95 °C for 2 min and then allowed to cool to 25 °C over 1 h. Human embryonic kidney (HEK293T) cells were cultured as a monolayer in DMEM GlutaMax medium (Thermo Fisher Scientific, Loughborough, UK) supplemented with 1% (v/v) penicillin-streptomycin (Thermo Fisher Scientific, Loughborough, UK) and 10% (v/v) fetal bovine serum (Sigma Aldrich, Gillingham, UK). HEK293T cells were seeded at 3.5 x 10^5 cells/well into 6 well plates and were transfected with 7.5 µL Lipofectamine 3000 (Thermo Fisher Scientific, Loughborough, UK), 1 µg gRNA/Cas9 and 1 µg annealed DNA oligo in 250 µL Opti-MEM. Cells were incubated for 48 h at 37 °C in 5% CO_2. Control cells were treated with Lipofectamine 3000 minus the gRNA/Cas9 vector or oligos. After 48 h, cells were split 1:10 with fresh DMEM GlutaMax medium. Array dilution was performed to isolate single cell colonies. To confirm CRISPR-Cas9 transfection in each single cell colony, *CYP24A1* direct sequencing was performed (c.2026_2032del, c.2035_2037del and c.2040del).

Mutant 25OHD and 1,25$(OH)_2$D catabolism. Both wild type HEK293T (i.e. controls) and CRISPR-Cas9 mutant cell lines were treated with 200 nM 25OHD or 10 nM 1,25$(OH)_2$D over a 48 h period. Culture medium was collected at intervals (1, 4, 12, 24, and 48 h) to assess vitamin D metabolite metabolism as described under biochemical analysis.

Single molecule fluorescence *in situ* hybridisation (smFISH). smFISH for *CYP24A1* mRNA transcripts was performed in HEK293T cells, *CYP24A1* mutant cells and Human PBMCs using a pool of 48 Stellaris RNA FISH probes (LGC Biosearch Technologies, Petaluma, CA). Probes recognising the housekeeping gene RNA polymerase II subunit A *(POLR2A)* labelled with Quasar™ 570 dye (#SMF-2003-1) were used as a positive control following the manufacturer's instructions available online at https://www.biosearchtech.com/support/resources/stellaris-protocols (access date 2021). Parameter design for *POLR2A* probes were against the coding sequence of NM_000937.4 (NCBI gene ID: 5430, nucleotides 387-6,299). Custom probes were designed against *CYP24A1* (NM_000782.5; NCBI gene ID: 1591; nucleotides 1-643 and 1,842-1,962) (Suppl. File 2) using the Stellaris RNA FISH probe designer available online at www.biosearchtech.com/stellarisdesigner and labelled with Quasar™ 670 dye. Cells were grown to ~80% confluency on 18 mm round glass coverslips before fixing with 4% paraformaldehyde for 30 min. Fixed cells were permeabilised with 90% (v/v) ethanol for 1 h. Hybridisation with 125 nM each probe was carried out for 18–24 h in a blackout humidified chamber at 37 °C. Coverslips were placed into fresh 12-well plates containing wash buffer A and incubated at 37 °C for 30 min. Samples were counterstained with DAPI for nuclear detection. Coverslips were then washed with wash buffer B and incubated for 5 min. Coverslips were mounted with 15 µL vectashield mounting medium (Vector Laboratories, Burlingame, CA) onto a microscope slide cell side down. Images were acquired with a Zeiss Elyra PS1 inverted microscope using 100x oil-immersion objective (1.46 NA) and cooled electron multiplying-CCD (charge-coupled device) Andor iXon 897 camera (512 × 512 QE>90%). Images were acquired using standard widefield rather than super-resolution mode. For fluorescence detection, Quasar™ 570 probes were detected using an excitation line of 561 nm with the signal detected at 570–640 nm. Quasar™ 670 probes were detected using an excitation line of 642 nm and signal detection at 655–710 nm. A 405 nm excitation line was used to detect the nuclear stain DAPI with emission detected between 420–480 nm. For all experiments, z-steps of 0.2 µm series were collected. mRNA counting was performed using ImageJ[21].

Statistics. Unless otherwise stated in the methods, statistical analyses and graphical representations were performed using SPSS Statistics v25.0.0.1 (IBM, Armonk, NY, USA) and Prism v9.0 (GraphPad Software, Inc., USA). Statistical significance was considered as a two-tailed p value <0.05.

RESULTS

High/low VMRs were associated with *CYP24A1* genetic abnormalities. Forty-seven patient biochemical profiles were analysed using LC-MS/MS as part of their standard of care from our metabolic and stone former clinics. Patients with elevated $1,25(OH)_2D$ and/or high/low VMR concentrations were indicative for HCINF1. Cases were sent for genetic screening to support the *CYP24A1* diagnosis. Four out of 47 (8.5%) displayed a lack of *CYP24A1* protein-coding region mutations despite the clinical presentation and biochemistry, which prompted non-clinical studies. The other 43 patients were positive for coding sequence mutations in *CYP24A1* consistent with previous reports and/or other genes including *PHEX* associated with their condition. The four adult cases (Patients 1-4) were examined alongside two infants with suspected *CYP24A1* pathology (Patients 5 and 6) (Table 1); Patient 5 also had Williams Syndrome. Except for one adult (Patient 1) who was within the reference range, $1,25(OH)_2D$ was elevated in all patients providing evidence for reduced $1,25(OH)_2D$ catabolism (Table 1). Patient 1, who was within the $1,25(OH)_2D$ reference range, had markedly elevated adjusted calcium (Table 1). The two infants (Patients 5 and 6) also displayed abnormal calcium handling (Table 1). Except for Patient 1, adults presented a 25OHD:$24,25(OH)_2D$ VMR in the lower 25^{th} percentile, but within the reference range demonstrating some CYP24A1 protein function (Table 1). In both infants (Patients 5 and 6) and Patient 1, the 25OHD:$24,25(OH)_2D$ VMR was in the upper 75^{th} percentile or above the upper limit, but within the reference range (Table 1). In summary, all six patients presented with either an abnormality of $1,25(OH)_2D$ or calcium handling with elevated $1,25(OH)_2D$ plus low normal 25OHD:$24,25(OH)_2D$ VMR associated with hypercalciuria and nephrolithiasis. Markedly elevated $1,25(OH)_2D$ plus high normal and elevated 25OHD:$24,25(OH)_2D$ VMR was associated with

hypercalcemia (in three patients) and nephrocalcinosis/renal stones (all patients). Five patients displayed normal phosphate levels (Table 1). Patient 6 was not tested for phosphate. All six patients harboured SNVs in the *CYP24A1* 3' UTR (Patient 1 c.2083T>C; Patient 2 c.1993C>T and c.2658C>G; Patient 3 c.2083T>C and c.2512T>A; Patient 4 c.2658C>G; Patient 5 c.2691G>A; Patient 6 c.2083T>C) with additional mutations in the protein-coding region in Patient 6 (c.368insC and c.1144insT) (Table 1).

***CYP24A1* 3' UTR mutations do not introduce *de novo* miRNA target sites.** We first speculated that the 3' UTR SNVs might have introduced *de novo* and/or mutated endogenous miRNA recognition elements (MREs) causing hypersilencing by miRNAs known to target *CYP24A1*, such as miR-30b and miR-125b, however our results did not support the hypothesis that there was MRE interruption with the 3' UTR mutations.

***CYP24A1* 3' UTR mutations are associated with mRNA misfolding.** Given the regulatory importance of the 3' ends of mRNA transcripts, we assessed all six patients using RNAfold, which produced graphical (Fig. 1A-G) and quantitative (Fig. 2A-G) outputs visualising the predicted mRNA structures. Each patient's predicted mRNA structure was visibly misfolded when compared to the reference human genome (Fig. 1A-G) and measurably misfolded after producing mountain plots when compared to the reference human genome (Fig. 2A-G). We investigated both the thermodynamic mRNA structures and the centroid structures for each genotype. Both parameters consistently showed that the patients' genotypes altered the mRNA structure.

***CYP24A1* mRNA misfolding does not alter transcript expression but is associated with higher CYP24A1 protein abundance.** Information movement from genotype to phenotype includes abundant regulatory stages so we tested whether mRNA misfolding caused a reduction of *CYP24A1* mRNA through mRNA degradation and/or decay leading to loss-of-function and reduced $1,25(OH)_2D$

catabolism. We extracted total RNA from whole blood from controls (n=5) plus the six patients and performed *CYP24A1* dPCR analysis. This experiment showed a non-significant difference in mRNA levels between the controls and patients (p=0.07) indicating that the mRNA structural changes were not causing reduced mRNA expression or stability (Fig. 3A). We next sought to determine whether mRNA misfolding was associated with abnormal translation. Western blot analysis performed on controls and three of the adult patients available for re-sampling showed a significant increase and/or accumulation of CYP24A1 (p=0.023) (Fig. 3B, C). To support this result, we also performed ELISAs for CYP24A1, which confirmed a significant increase in CYP24A1 protein associated with *CYP24A1* mRNA misfolding (p = 0.008) (Fig. 3D).

Development of an *in vitro* model system for HCINF1 and *CYP24A1* studies. For studies to explore how *CYP24A1* mRNA misfolding was associated with CYP24A1 abundance, which given the patients' phenotype appeared to be semi-functional (i.e. normal 25OHD and 24,25$(OH)_2$D but abnormal 1,25$(OH)_2$D catabolism), we used CRISPR-Cas9 to modify the *CYP24A1* 3' UTR in a HEK293T cell line. The Human Protein Atlas (www.proteinatlas.org) confirmed that CYP24A1 is enriched in kidney cells, which is why we selected an immortalised human kidney cell line. We designed the mutant cell model to contain three deletions (c.2026_2032del, c.2035_2037del and c.2040del), which was a trade-off between ensuring that the 3' UTR structure was altered and designing CRISPR-Cas9 oligos that would be successful, which was confirmed and validated by direct sequencing and RNAfold (Fig. 4A, B). The mountain plot and entropy for each nucleotide of the mutant sequence with comparison to the wild type was measured (Fig. 4C, D). Western blotting showed that CYP24A1 was still expressed in the mutant when compared to wild type HEK293T cells (Fig. 4E). We then compared mutant cell responses to 25OHD and 1,25$(OH)_2$D stimulation to validate the model. We observed a non-significant (p=0.18) production of 24,25$(OH)_2$D in 25OHD-treated cells over a 48 h period (Fig. 4F) (Table 1). 1,25$(OH)_2$D catabolism was significantly impacted after 24 h (p=0.04) in 1,25$(OH)_2$D-treated mutant cells (Fig. 4G).

***CYP24A1* mRNA transcript subcellular location *in vitro*.** We developed a single molecule fluorescent *in situ* hybridisation (smFISH) protocol to detect single *CYP24A1* mRNA transcripts. This protocol was developed to visualise whether mRNA misfolding interfered with mRNA nuclear export, trafficking and/or localisation leading to abnormal protein translation. We stimulated HEK293T cells with 10 nM $1,25(OH)_2D$ for 12 h before fixing and probing, which showed single *CYP24A1* mRNA transcripts (Fig. 5A). These mRNAs were significantly upregulated in response to $1,25(OH)_2D3$ ($p=0.008$) (Fig. 5B). As well as observing a significant increase in total *CYP24A1* mRNA transcripts when compared to the housekeeping gene *POLR2A* ($p=0.005$), we observed markedly high cell-to-cell variation and that most *CYP24A1* mRNA transcripts resided in the nucleus (Fig. 5C). *POLR2A* detection was consistent across all experiments, supporting this molecule as an appropriate control for smFISH studies (Fig. 5C). *CYP24A1* localisation variation included cytoplasmic accumulation in cells with minimal nuclear expression (Fig. 5D), low abundance *CYP24A1* absent from both the nucleus and the cytoplasm (Fig. 5D) plus several *CYP24A1* transcription sites in the nucleus with minimal detection in the cytoplasm (Fig. 5D). We also performed smFISH in the mutant and observed *CYP24A1* localisation in both the nucleus and cytoplasm (Fig. 5D). Strong *CYP24A1* probe signal was identified in mutant cells similar to transcription sites observed in the wild type (Fig. 5D). We observed a higher *CYP24A1* mRNA cytoplasmic retention in the mutant, but this was non-significant (cytoplasm $p=0.54$; nuclear $p=0.43$) when compared to wild type HEK293T cells (Fig. 5E).

DISCUSSION

LC-MS/MS and immunoassay-based vitamin D analysis and VMR determination of 47 patients who presented at our metabolic and renal stone clinics and 2 children being investigated for nephrocalcinosis and calcium abnormalities suggested possible *CYP24A1* loss-of-function, consistent with a HCINF1 diagnosis. A small group of patients with suspected *CYP24A1* pathology due to abnormal biochemical profiles were absent of protein-coding mutations, which prompted further

investigation. A lack of protein-coding mutations in *CYP24A1* despite phenotypes was consistent with some other previous reports[22,23]. We extended our mutational analysis to include the 5' and 3' UTR regions in the six patients as these loci are highly important in gene regulation independent of the coding sequence. Direct sequencing identified SNVs in the 3' UTR in all six patients. The differing 3' UTR SNVs in our cohort might explain some of the heterogeneous phenotypes reported in the literature in patients with HCINF1[5].

We observed mRNA structural alterations in each patient with 3'UTR variants. mRNA structural changes triggered by the presence of 3' UTR sequence variants (or structural elements) have been shown to induce translational heterogeneity or impair translation completely by mRNA destabilisation[24], abnormal mRNA trafficking[24,25] and/or reduced ribosome scanning efficacy[26]. Our *in silico* mRNA structural alterations driven by the described 3' UTR variants may impair proper production of functional CYP24A1 protein resulting in inappropriate $1,25(OH)_2D$ concentration with a low/high VMR.

We did not observe MRE abnormalities in the 3' UTR mutations in our cohort, which was the most obvious starting point for non-clinical studies. We observed no significant effect on *CYP24A1* mRNA transcription in patients with 3' UTR structural elements. The lack of evidence supporting transcriptional or post-transcriptional alteration observed in our cohort directed our new hypotheses towards a translational impairment, compromising CYP24A1 enzyme function giving rise to the elevated $1,25(OH)_2D$ and high/low VMR[27,28]. We observed a significant increase and/or retention of CYP24A1 in patients with misfolded mRNAs. Our data suggested an accumulation of a partially functional CYP24A1 protein, demonstrated by a somewhat normal catabolism of 25OHD into $24,25(OH)_2D$ but abnormal $1,25(OH)_2D$ catabolism.

Molecular investigation into the consequences of 3' UTR mRNA structural alterations led to two new hypotheses that needed to be addressed. (i) CYP24A1 upregulation is the expected homeostatic response to increased serum calcium therefore increased protein on a western blot, while mRNA structural elements signal for an unknown but pathogenic post-translational modification hindering protein function. (ii) mRNAs are trafficked from the nucleus to subcellular regions where the subsequent protein will be required more rapidly; mRNA localisation to specific regions within a cell provide regulation of protein expression but mRNA misfolding interferes with this trafficking process[29]. Given that CYP24A1 is functional in the inner mitochondrial membrane in $1,25(OH)_2D$ target cells and that some RNA species, particularly long non-coding RNAs, are known to act as structural components to the mitochondrial membrane[30], it is possible that mRNA structural abnormalities physically anchor translational machinery and prevent the protein from proper localisation. Improper localisation affecting translational machinery could go some way to describe increased CYP24A1 expression with little effect on mRNA transcription as was observed. We investigated some but not all components of these hypotheses here using our newly developed *CYP24A1* mutant model.

While the *CYP24A1* knockout mouse models can be useful in further understanding patients with complete loss-of-function *CYP24A1* mutations[31], our patient cohort presented with hypomorphic mutations associated with compromised CYP24A1 function. Currently no models are commercially available that comprise *CYP24A1* 3' UTR variants leading to mRNA structural alterations and partially functional CYP24A1. RNAfold demonstrated that the 3' UTR deletions transfected into our newly generated CRISPR-Cas9 mutant cell line altered the *CYP24A1* mRNA secondary structure. There was a non-significant decrease in catabolism of 25OHD to $24,25(OH)_2D$ but significantly decreased $1,25(OH)_2D$ clearance over 48 h. Decreased catabolism rates were consistent with the patient phenotype of detectable $24,25(OH)_2D$ concentration and elevated $1,25(OH)_2D$ concentration due to

partially functional CYP24A1. Our CRISPR-Cas9 mutant allows for future *in vitro* investigation into non-canonical CYP24A1 disease pathogenesis.

Using the mutant model we investigated the effect that 3' UTR structural elements had on CYP24A1 cellular localisation by novel single cell visualisation of individual *CYP24A1* mRNA transcripts *in vitro* using smFISH. We observed significant *CYP24A1* cell-to-cell expression and localisation variability in HEK293T and mutants. The variability in *CYP24A1* localisation would have been undetected using routine mRNA quantification methods such as qPCR that analyse total mRNA abundance. The cell-to-cell variability could be explained by the progression of the cell cycle, which has been shown to affect mRNA transcription rates[32] or fixing at different stages of the cell cycle. smFISH analysis indicated that alterations to the 3' UTR likely had little to no effect on *CYP24A1* localisation *in vitro* as no significant difference was observed in comparison to wild type HEK293T cells. This finding supports the lack of *CYP24A1* mRNA transcription variability observed in our patient samples analysed by dPCR. Although no significant localisation or abundance differentiation was observed between the mutant and HEK293T cell lines, this work provides insight into the localisation of *CYP24A1* in cells, which has not been previously reported. *CYP24A1* visualisation can be applied on a patient-by-patient basis to assess the effect that different *CYP24A1* mutations (both protein-coding and non-coding) have on mRNA stability, expression and cellular localisation. Future Human *ex vivo* plus *in vitro* studies using smFISH will allow for the accurate visualisation of mRNA abundance and localisation relating to structure-function relationships, which is a major advancement on current qPCR techniques in RNA biology. smFISH is therefore a compelling tool in investigating the mRNA lifecycle of *CYP24A1*[33].

This study investigated the role of mRNA misfolding in Human disease. We identified a cohort of patients with *CYP24A1*-mediated abnormal calcium handling and/or HCINF1 phenotypes harbouring non-coding genetic abnormalities. We expanded *CYP24A1* analysis to include the UTRs, which revealed a possible mechanism for six patients who lacked coding sequence pathology. Further studies

on this group revealed that whilst the amino acid sequence was unaffected (except for Patient 6), the mRNA molecule itself was affected, importantly the key regulatory region 3' UTR, which resulted in predicted three-dimensional structural changes that interfered with proper protein translation. Our findings provide insight into the potential effect mRNA structural abnormalities caused by sequence-dependent structural elements have on biological function. Though we could not fully describe the underlying mechanism linking mRNA structure with partially functional CYP24A1 protein activity, our work presents a pathway for employing ribosome frameshift profiling and/or protein sequencing to determine the core C-term region required for catabolising $1,25(OH)_2D$, which only differs from 25OHD by the additional OH group. Some mRNAs carry specific structural elements in their 3' ends that cause ribosomes to slip and then readjust the reading frame[34]. The frameshift results from a change in the reading frame by one or more bases in either the 5' (−1) or 3' (+1) directions during translation[34]. The newly developed CRISPR-Cas9 mutated HEK293T cell line could provide the foundation for large scale *in vitro* studies and will continue to support our understanding of vitamin D metabolism in patients with novel 3' UTR mutations. The findings of this research provide a framework that can be used to better understand the molecular basis of pathogenesis in patients lacking protein-coding region abnormalities.

AUTHOR CONTRIBUTIONS

Study conception and design: WDF and DG. Experiments: NB, SD, YZ, RP, IP, EW and JCYT. Data analysis and interpretation: NB, RP, IS, YD, WDF and DG. Samples, medical classification and pathology: BL, AC, AK, LP, HM and WDF. Wrote the manuscript: NB and DG. Revised and approved the final manuscript: all authors.

DATA AVAILABILITY

All data supporting the findings of this study are available within the article and supplementary files or from the corresponding authors on request.

ACKNOWLEDGEMENTS

The Michael Davie Research Foundation and Norwich Medical School PhD Programme funded this study. EW is funded by the Bone Cancer Research Trust. WDF and DG are supported by The Difference Campaign. We thank the John Innes Centre bioimaging facility and staff for their contribution to this work and we thank Matthew Jefferson, Yingxue (Sophia) Wang and Gabriella Oliver-Wilkins for technical support. We are indebted to the patients for their participation in this study.

COMPETING INTERESTS

The authors declare no competing interests.

REFERENCES

1. Fraser WD, Milan AM. Vitamin D Assays: Past and Present Debates, Difficulties, and Developments. Calcif Tissue Int. 2013 Feb;92(2):118–27.
2. Fraser WD, Tang JCY, Dutton JJ, Schoenmakers I. Vitamin D Measurement, the Debates Continue, New Analytes Have Emerged, Developments Have Variable Outcomes. Calcif Tissue Int. 2020 Jan;106(1):3–13.
3. Teshome A, Adane A, Girma B, Mekonnen ZA. The Impact of Vitamin D Level on COVID-19 Infection: Systematic Review and Meta-Analysis. Front Public Health. 2021 Mar 5;9:624559.

4. Tang JCY, Nicholls H, Piec I, Washbourne CJ, Dutton JJ, Jackson S, et al. Reference intervals for serum 24,25-dihydroxyvitamin D and the ratio with 25-hydroxyvitamin D established using a newly developed LC–MS/MS method. J Nutr Biochem. 2017 Aug 1;46:21–9.
5. Schlingmann KP, Kaufmann M, Weber S, Irwin A, Goos C, John U, et al. Mutations in CYP24A1 and Idiopathic Infantile Hypercalcemia. N Engl J Med. 2011 Aug 4;365(5):410–21.
6. Schoenmakers I, Piec I, Baban S, Bärebring L, Green D, Washbourne CJ, et al. Gestational hypercalcemia: Prevalence and biochemical profile. J Steroid Biochem Mol Biol. 2020 May 1;199:105611.
7. Sharp PA. The Centrality of RNA. Cell. 2009 Feb 20;136(4):577–80.
8. Pan T, Sosnick T. Rna Folding During Transcription. Annu Rev Biophys Biomol Struct. 2006;35(1):161–75.
9. Warf MB, Berglund JA. The role of RNA structure in regulating pre-mRNA splicing. Trends Biochem Sci. 2010 Mar;35(3):169–78.
10. Martin KC, Ephrussi A. mRNA Localization: Gene Expression in the Spatial Dimension. Cell. 2009 Feb 20;136(4):719–30.
11. Kozak M. Regulation of translation via mRNA structure in prokaryotes and eukaryotes. Gene. 2005 Nov 21;361:13–37.
12. Garneau NL, Wilusz J, Wilusz CJ. The highways and byways of mRNA decay. Nat Rev Mol Cell Biol. 2007 Feb;8(2):113–26.
13. Tang J. Methods for the Measurement of Vitamin D Metabolites and Studies on Their Relationships in Health and Disease. University of East Anglia; 2019.
14. Sayers J, Hynes A, Rice S, Hogg P, Sayer J. Searching for CYP24A1 mutations in cohorts of patients with calcium nephrolithiasis. OA Nephrol. 2013;1(1).
15. Gruber AR, Lorenz R, Bernhart SH, Neuböck R, Hofacker IL. The Vienna RNA websuite. Nucleic Acids Res. 2008 Jul 1;36(Web Server issue):W70-74.

16. Lorenz R, Bernhart SH, Höner zu Siederdissen C, Tafer H, Flamm C, Stadler PF, et al. ViennaRNA Package 2.0. Algorithms Mol Biol. 2011 Nov 24;6(1):26.
17. Paz I, Kosti I, Ares M, Cline M, Mandel-Gutfreund Y. RBPmap: a web server for mapping binding sites of RNA-binding proteins. Nucleic Acids Res. 2014 Jul 1;42(W1):W361–7.
18. Chen Y, Wang X. miRDB: an online database for prediction of functional microRNA targets. Nucleic Acids Res. 2020 Jan 8;48(D1):D127–31.
19. Liu W, Wang X. Prediction of functional microRNA targets by integrative modeling of microRNA binding and target expression data. Genome Biol. 2019 Dec;20(1):18.
20. De St. Groth SF. The evaluation of limiting dilution assays. J Immunol Methods. 1982 Mar 12;49(2):R11–23.
21. Bahry E, Breimann L, Zouinkhi M, Epstein L, Kolyvanov K, Long X, et al. RS-FISH: Precise, interactive, fast, and scalable FISH spot detection. bioRxiv; 2021. p. 2021.03.09.434205.
22. Dauber A, Nguyen TT, Sochett E, Cole DEC, Horst R, Abrams SA, et al. Genetic Defect in CYP24A1, the Vitamin D 24-Hydroxylase Gene, in a Patient with Severe Infantile Hypercalcemia. J Clin Endocrinol Metab. 2012 Feb 1;97(2):268–74.
23. Molin A, Baudoin R, Kaufmann M, Souberbielle JC, Ryckewaert A, Vantyghem MC, et al. *CYP24A1* Mutations in a Cohort of Hypercalcemic Patients: Evidence for a Recessive Trait. J Clin Endocrinol Metab. 2015;100(10):1343–52.
24. Ainger K, Avossa D, Diana AS, Barry C, Barbarese E, Carson JH. Transport and Localization Elements in Myelin Basic Protein mRNA. J Cell Biol. 1997 Sep 8;138(5):1077–87.
25. An JJ, Gharami K, Liao GY, Woo NH, Lau AG, Vanevski F, et al. Distinct Role of Long 3'UTR BDNF mRNA in Spine Morphology and Synaptic Plasticity in Hippocampal Neurons. Cell. 2008 Jul 11;134(1):175–87.
26. Steri M, Idda ML, Whalen MB, Orrù V. Genetic Variants in mRNA Untranslated Regions. Wiley Interdiscip Rev RNA. 2018;9(4):1474.

27. Moor CH de, Meijer H, Lissenden S. Mechanisms of translational control by the 3' UTR in development and differentiation. Semin Cell Dev Biol. 2005 Feb 1;16(1):49–58.
28. Scheper GC, van der Knaap MS, Proud CG. Translation matters: protein synthesis defects in inherited disease. Nat Rev Genet. 2007 Sep;8(9):711–23.
29. Biever A, Glock C, Tushev G, Ciirdaeva E, Dalmay T, Langer JD, et al. Monosomes actively translate synaptic mRNAs in neuronal processes. Science. 2020 Jan 31;367(6477):4991.
30. Vendramin R, Verheyden Y, Ishikawa H, Goedert L, Nicolas E, Saraf K, et al. SAMMSON fosters cancer cell fitness by concertedly enhancing mitochondrial and cytosolic translation. Nat Struct Mol Biol. 2018 Nov;25(11):1035–46.
31. St-Arnaud R. Targeted inactivation of vitamin D hydroxylases in mice. Bone. 1999 Jul;25(1):127–9.
32. Perez-Carrasco R, Beentjes C, Grima R. Effects of cell cycle variability on lineage and population measurements of messenger RNA abundance. J R Soc Interface. 2020 Jul;17(168):20200360.
33. Janes KA. Cell-to-cell transcript variability—seeing signal in the noise. Cell. 2015 Dec 17;163(7):1566–8.
34. Rampersad S, Tennant P. Chapter 3 - Replication and Expression Strategies of Viruses. In: Viruses. Academic Press; 2018. p. 55–82.

FIGURE LEGENDS

Fig. 1. 3' UTR mutations and mRNA structures. A. Reference human genome *CYP24A1* transcript. **B.** Patient 1 (c.2083T>C). **C.** Patient 2 (c.1993C>T; c.2658C>G). **D.** Patient 3 (c.2083T>C; c.2512T>A). **E.** Patient 4 (c.2658C>G). **F.** Patient 5 (c.2691G>A). **G.** Patient 6 (c.368insC; c.1144insT; c.2083T>C). Red circles highlight the misfolded region when compared to the reference.

590 **Fig. 2. mRNA misfolding quantification using mountain plots.** mRNA structure thermodynamic
591 ensemble and the centroid structures. Mountain plots represent structure in a plot of height versus
592 position where the height m(K) is given by the base pair number enclosing the base at position k, i.e.
593 loops correspond to plateaus, hairpin loops are peaks, helices to slopes. We also show the entropy for
594 each position. Patient genotypes are visibly different when compared to the reference structure. **A.**
595 Reference human genome *CYP24A1*. **B.** Patient 1. **C.** Patient 2. **D.** Patient 3. **E.** Patient 4. **F.** Patient 5.
596 **G.** Patient 6.
597
598 **Fig. 3. mRNA misfolding mechanisms.** Three patient samples were available for *ex vivo* studies. **A.**
599 Total RNA extracted from whole blood samples used for dPCR analysis for *CYP24A1* transcript
600 expression shows no significant difference (p=0.07) between 5 controls and 3 patients. Data is
601 reported as copies/µL as calculated by Poisson distribution. Data presentation includes the 75th and
602 25th percentiles with the median values (50th percentile) indicated by the central line. The x represents
603 the mean value. Due to the accuracy and sensitivity of dPCR, *CYP24A1* expression across the samples
604 is invariable. Statistical significance was calculated using an unpaired t test. **B.** PBMC western blot
605 using a mouse monoclonal anti-CYP24A1 antibody shows an increase, retention and/or accumulation
606 of CYP24A1 protein in patients when compared to controls. Actin was used as a loading control. **C.**
607 Graphical representation of the western blot data to demonstrate CYP24A1 increase when compared
608 to controls. Error bars represent the standard deviation. Statistical significance was calculated using
609 an unpaired t test (p=0.023). **D.** Plasma samples' ELISA confirmed CYP24A1 increase and/or
610 accumulation in patient samples when compared to controls. Error bars represent the standard
611 deviation. Statistical significance was calculated using an unpaired t test (p=0.008). For **C** and **D**, the
612 top and bottom whisker lines denote the maximum and minimum value in the dataset, respectively,
613 the top and bottom of the box denote the 75th and 25th percentiles of the dataset, the median value
614 (50th percentile) is indicated by the central line of the box plot, the x represents the mean value and
615 the white circles indicate data points that are within the interquartile range.

Fig. 4. CRISPR-Cas9 mutant cell line. MFE-based mRNA structures of **A.** Reference human genome *CYP24A1* transcript and **B.** CRISPR-Cas9 mutant *CYP24A1* (c.2026_2032del, c.2035_2037del and 2040-2041del) with the red circle indicating structure alteration. Mountain plot measurements of the MFE-based mRNA structures of **C.** Wild type *CYP24A1* and **D.** CRISPR-Cas9 modified *CYP24A1* (c.2026_2032del, c.2035_2037del, 2040-2041del). **E.** Western blot shows that the CYP24A1 protein is still expressed in the mutant consistent with the patient phenotype. GAPDH was used as a loading control. **F.** LC-MS/MS analysis of HEK293T and mutant 24,25$(OH)_2$D in the culture medium over 48 h when stimulated with 200 nM 25OHD. Results are displayed as average 24,25$(OH)_2$D concentration at each timepoint measured (n=3). **G.** LC-MS/MS analysis of HEK293T and mutant 1,25$(OH)_2$D concentration in culture medium over 48 h after stimulating with 10 nM 1,25$(OH)_2$D (n=3).

Fig. 5. *CYP24A1* mRNA quantification and subcellular localisation. A. Maximum Z-projected smFISH image showing a HEK293T cell stimulated with 10 nM 1,25$(OH)_2$D with markedly more *CYP24A1* mRNAs (green) than control *POLR2A* transcripts (magenta). **A**. High intensity *CYP24A1* spots indicate ongoing *CYP24A1* transcription (white arrow). Nuclear stain DAPI is shown in blue. **B.** Comparison of the average *CYP24A1* mRNA transcripts per cell observed before and after stimulation of wild type HEK293T cells with 10 nM 1,25$(OH)_2$D (p=0.00842). **C.** A significant difference was observed between the frequency of *CYP24A1* and *POLR2A* in the cytoplasm (p=0.03) and the nucleus (p=0.04) in 1,25$(OH)_2$D-stimulated HEK293T cells. There was a significant difference in total *CYP24A1* and *POLR2A* probe frequency between cells (p=0.005). **D.** smFISH images of a group of three HEK293T cells and two mutant cells that demonstrate the observed variation in *CYP24A1* mRNA localisation. Individual control *POLR2A* (magenta) and *CYP24A1* mRNA transcripts (green) are shown along with the nuclear stain DAPI (blue) in the merged image. In HEK293T cell 1, *CYP24A1* mRNA is absent from both the nucleus and the cytoplasm. HEK293T cell 2 has a high abundance of *CYP24A1* mRNA in both the nucleus and cytoplasm with a high intensity spot indicating high levels of ongoing

642 transcription (white arrows). HEK293T cell 3 also has a discrete *CYP24A1* mRNA spot in the nucleus
643 indicating ongoing transcription (white arrow), but relatively fewer *CYP24A1* mRNAs in the cytoplasm.
644 In the mutant cells 1 and 2 *CYP24A1* is also expressed in both the nucleus and the cytoplasm with high
645 intensity spots indicating high levels of ongoing transcription (white arrows) **E.** Data for *CYP24A1* in
646 the wild type HEK293T and mutant cell lines show no significant difference in mRNA localisation. For
647 **B, C** and **E**, the top and bottom whisker lines denote the maximum and minimum value in the dataset,
648 respectively, the top and bottom of the box denote the 75^{th} and 25^{th} percentiles of the dataset, the
649 median value (50^{th} percentile) is indicated by the central line of the box plot, the x represents the
650 mean value and the white circles indicate data points that are within the interquartile range. Scale
651 bars are 10 µm.

Table 1. Serum biochemistry and mutational analysis. A low 25OHD:24,25(OH)$_2$D VMR was associated with hypercalciuria and nephrolithiasis. A high 25OHD:24,25(OH)$_2$D VMR was associated with hypercalcemia and nephrocalcinosis. All but one patient had significantly reduced 1,25(OH)$_2$D catabolism and some patients presented with abnormal calcium handling. Reference ranges are given in brackets. Red font shows abnormal measurement.

Patient	Sex	Age (Y)	Total 25OHD (50-120 nmol/L)	1,25(OH)2D (55-139 pmol/L)	Total 24,25(OH)2D (1.1-13.5 nmol/L)	Total 25OHD:24,25(OH)2D Relative Ratio (7-23)	1,25(OH)2D:24,25(OH)2D Relative Ratio (11-62)	Adjusted Calcium (2.1-2.6 mmol/L)	Phosphate (0.8-1.5 mmol/L) (Adult); 1.45-2.1 (Child)	SNV	Clinical Notes
1	F	33	104	83	3.3	32	25	3.27	1.04	c.2083T>C	Hypercalciuria; recurrent nephrolithiasis
2	F	28	91	262	10.6	9	25	2.33	1.29	c.1993C>T; c.2658C>G	Hypercalciuria; recurrent nephrolithiasis
3	F	33	97	177	7	14	25	2.32	0.93	c.2083T>C; c.2512T>A	Hypercalciuria; recurrent

											nephrolithiasis; osteopenia
4	M	55	73	171	9.4	8	18	2.44	0.93	c.2658C>G	Recurrent nephrolithiasis; osteopenia
5	M	<1	122	616	3.5	35	176	3.41	1.98	c.2691G>A	Williams syndrome; IIH; nephrocalcinosis
6	M	<1	108	175	5.8	19	30	3.2	Not tested	c.368insC; c.1144insT; c.2083T>C	Nephrocalcinosis; high creatinine