https://doi.org/10.48047/AFJBS.6.15.2024.3713-3748

African Journal of Biological Sciences

Journal homepage: http://www.afjbs.com

Research Paper Open Access

Hepatocyte Nuclear Factor-1 β (HNF1β) Gene Mutations in Patients with Congenital Anomalies of the Kidney and Urinary Tract (CAKUT)

Sara Yahia Gaber, Heba Mostafa Ahmed, Manal M. Thomas, Osama Ezzat Botrous, Tamer HAM, Rehab Muhammed Abd Elkareem,

Assistant lecturer, Faculty of Medicine, Beni-Suef University, Egypt
Sara.y.gaber@gmail.com

Professor of pediatrics, Faculty of Medicine, Beni-Suef University, Egypt
Heba_most@yahoo.com

Professor of clinical genetics, Human genetics and genomic research institute, National research center
nula_m@hotmail.com

Ass. Professor of Clinical Pathology, Faculty of Medicine, Beni-Suef University, Egypt
drrehablashin@gmail.com

Professor of pediatrics, faculty of Medicine, Beni-Suef University, Egypt
Osama.botrous@yahoo.com

Researcher of Medical Molecular genetics at Medical Molecular genetics Department, Human Genetics & Genome Research institute National Research center
Mohamedammar6121981@gmail.com
(corresponding author): Sara Yahia Gaber

Article History

Volume 6, Issue 15, 2024

Received : 03 May 2024

Accepted : 28 Aug 2024

Published: 07 Sep 2024

doi:10.48047/AFJBS.6.15.2024.3713-3748

Abstract

Background: Congenital anomalies of the kidney and urinary tract (CAKUT) are the leading cause of chronic kidney disease in children, occurring in 4 to 60 out of 10,000 births. They represent 20-30% of all congenital anomalies, with severity ranging from mild, transient hydronephrosis to kidney failure requiring early dialysis. Approximately 40% of children on kidney replacement therapy in the UK are affected by congenital kidney anomalies, particularly renal hypoplasia or dysplasia. CAKUT encompasses a wide range of phenotypes, including multicystic dysplastic kidney (MCDK) and renal agenesis, with severity often correlating with the timing of developmental disruption. Over 200 syndromes include CAKUT, such as Eagle-Barrett syndrome and PUV. Radiologic studies, especially prenatal ultrasonography, are pivotal for early detection. Genetic testing, particularly for HNF1B and PAX2, is being explored to identify genes implicated in CAKUT, despite limitations in current research strategies. This study aims to further investigate HNF1B mutations in CAKUT to accentuate understanding of its genetic basis and extra-renal clinical manifestations.

Methods: This cross-sectional study followed the STROBE checklist and was conducted at Beni Suef University Hospital between 2020 and 2022. It involved 20 Egyptian pediatric patients from 19 unrelated families, divided into two groups: those with multicystic dysplastic kidney (MCDK) and those with congenital anomalies of the kidney and urinary tract (CAKUT) other than MCDK. After obtaining informed consent, detailed histories, clinical examinations, and laboratory tests, including HNF1-beta gene mutation analysis, were performed. Radiological assessments classified CAKUT patients using the SFU grading system. DNA was extracted from blood samples for PCR and sequencing. Statistical analysis was conducted using SPSS, employing tests such as the unpaired T-test, Chi-Square, Mann-Whitney, and Kruskal-Wallis, with a p-value < 0.05 considered significant. The study's sample size was calculated using Epi Info, targeting 20 patients for adequate power.

Results: This study analyzed socio-demographic data and genetic variations in 20 patients with CAKUT from 19 unrelated families. Patients ranged in age from 11 months to 27 years, with a significant percentage having a family history of diabetes mellitus (75%) and renal problems (40%). Genetic screening of the HNF1B gene revealed no mutations but identified eight SNPs. The most frequent polymorphisms were rs3110641 (70%), rs1800929 (55%), rs2229295 (50%), and rs35913775 (40%). No significant association was found between SNPs and CAKUT when comparing the allele and genotype frequencies with controls. However, rs3110641 was associated with a higher occurrence of CAKUT in the recessive model ($p=0.03$). Furthermore, the AGT haplotype was significantly linked to CAKUT. SNPs rs2229295 and rs1800929 were also associated with severe CAKUT. Additionally, rs2229295, rs1800929, and rs35913775 were significantly associated with CAKUT in patients with T1DM, suggesting a potential genetic predisposition for the co-occurrence of these conditions.

Conclusion: In conclusion, our study explored the genetic basis of HNF1B mutations in CAKUT. While no pathogenic mutations were found, we identified several SNPs associated with CAKUT, including rs3110641 and rs2229295, which showed significant associations with the disease in recessive models. An AGT haplotype was also significantly linked to CAKUT, and the rs35913775 mutant genotype was linked to CAKUT with T1DM. Our findings offer new insights into SNPs related to CAKUT, but further research is needed to fully understand mutations of HNF1B in CAKUT development and improve prenatal detection of renal anomalies.

Keywords: congenital anomalies of the kidney and urinary tract, single-nucleotide polymorphisms, hepatocyte nuclear factor-1 beta, type 1 diabetes mellitus, AGT, multicystic dysplastic kidney.

Introduction

Congenital anomalies of the kidney and the urinary tract (CAKUT) are the most common cause of chronic kidney disease in children [1, 2]. They are observed in 4 to 60 out of 10000 births, representing 20-30% of overall congenital anomalies [3]. The spectrum of severity of CAKUT is vast, ranging from subclinical transient hydronephrosis to kidney failure necessitating dialysis in early neonatal life [4]. As per reports by the UK Renal Registry, 40% of children on kidney replacement therapy (KRT) are impacted by a type of congenital kidney anomaly, namely renal hypoplasia or dysplasia [5].

CAKUT also manifests a wide range of phenotypes that correspond to the severity of the disease, including one of the most common forms multicystic dysplastic kidney (MCDK), and the most severe phenotype, renal agenesis [4]. The severity of congenital renal anomalies typically escalates in proportion to the time of developmental disruption, with earlier disruptions leading to more severe disease [4]. There are nearly 200 syndromes which encompass a type of CAKUT [6]. Clinical phenotypes of CAKUT include renal agenesis, MCDK, renal hypoplasia, renal dysplasia, Eagle-Barrett syndrome, and posterior urethral valves, amongst others [5, 7].

Histopathologic studies of postmortem renal tissue refined our understanding of CAKUT, expounding on the architectural distortion that occurs with congenital renal dysplasia [4]; however, radiologic studies remain the pillars of diagnosis and management of CAKUT as they entail detecting CAKUT at its earliest phase, prenatally, usually by ultrasonography. Prenatal hydronephrosis detected on ultrasonography may be the initial sign of either a transient problem or may indicate a significant anomaly that persists postnatally [5], though the serious end of the spectrum of prenatal hydronephrosis predominates. As relayed by a study of 307 Indian children, out of 80% of CAKUT cases, 70% of those children exhibited prenatal hydronephrosis [8].

This propensity for prenatally detected hydronephrosis to represent a serious renal pathology corroborates the need for detection methods that may mitigate the severity of chronic kidney disease in childhood, as well as offer families the choice of informed genetic counseling and future family planning [4, 9]. Of those methods, genetic testing in patients with kidney malformations has been explored as a potential tool to identify genes which may be implicated in the development of CAKUT, several genes have been identified, most notably PAX2 and HNF1B, which were implicated in 5-15% of dominantly inherited renal anomalies [10]. Concerning HNF1B, it was found to be involved in multitudes of disorders, comprising both renal (renal hypoplasia) and extra-renal disorders (MODY, type 2 DM, and hypomagnesemia), as well as several cancers (renal cell carcinoma, prostatic cancer) [11-16].

Single nucleotide polymorphisms (SNPs) are considered the most common type of genetic variation, characterized by a variation in a single nucleotide in the DNA which affects gene

expression and subsequently impacts physiological processes. Genome-wide association studies (GWAS) provided insights showing that SNPs lead to pathological outcomes such as congenital anomalies and congenital diseases [17]. SNPs have been implicated in the development of CAKUT in many instances in the literature, with some studies attempting to explain the mechanisms through which SNPs influence gene expression to disrupt nephrogenesis. For example, mutations in genes that express fibroblast growth factor receptors (FGFR) have been linked to the development of CAKUT [10, 18]. Regarding HNF1B, SNPs of this gene have been identified as possible catalysts for the development of CAKUT, though evidence is insufficient [19].

Studies that attempted to explore the role of SNPs and CNVs in CAKUT have been met with several limitations such as the use of targeted exome sequencing (TES) instead of whole exome sequencing (WES), which makes the identification of novel mutations impossible. Other limitations included the lack of resources, the lack of sequencing of noncoding regions, as well as insufficient uniform testing for all involved subjects leading to discrepancy in the presented data. Other studies which reported on the role of SNPs in inherited diseases have been limited by the variable miRNA expression, which is much higher in vivo, leading to false positive results in experimental models. Furthermore, discordance between changes at the mRNA levels and protein levels posed an unavoidable limitation since monitoring of protein levels in each cell line was not always conducted [11, 20]. These limitations affected the overall significance of these studies and failed to offer robust evidence of the hypothetical implication of SNPs on CAKUT.

In this cross-sectional study, we aimed to address the systematic mutational analyses of HNF1B in CAKUT, primarily focusing on elucidating the prevalence of HNF1B mutations and understanding the clinical manifestations of HNF1B-assoicated diseases, in hopes to augment the available evidence regarding the role of genetics in CKD.

Methods

During the reporting process of this manuscript, we adhered to the checklist of items of the STROBE statement [21]. This cross-sectional study was conducted at the Pediatric Department, Beni Suef University Hospital from the year 2020 to 2022 following the approval of the Scientific and Ethical Committee at Beni Suef University Hospital. It included 20 Egyptian patients from 19 unrelated families. Patients were recruited from the Pediatric Nephrology Clinic, Beni Suef University Hospital and divided into two groups, those who had MCDK (5 patients) and another group with CAKUT other than MCDK (15 patients). Informed consent was obtained from the guardians of all patients included in the study prior to commencing the study.

Our study included patients with fetal bilateral hyperechogenic kidneys, multicystic dysplastic kidney, renal agenesis, hypoplastic or dysplastic kidneys, cysts of unknown origin, ectopic kidney, as well as vesico-ureteral reflux (VUR) and hydronephrosis. We

excluded patients who had autosomal recessive polycystic kidney disease (AR-PCKD) as well as those who had multiple extra-renal congenital anomalies.

All patients were subjected to full history taking, which encompassed personal history (name, age, sex, residence), maternal conception age, and mood of conception, with specific focus on consanguinity, family history of DM or renal disease, sibling death, perinatal history, time of detection of the disease, extra-renal associations, any treatment received, and pedigree analysis—which included a three-generation pedigree analysis construction for all patients, shedding light on consanguinity and similarly affected siblings or other family members.

A thorough clinical examination was conducted, including general examination of all patients in terms of measurements, blood pressure, and any noticeable abnormal facies. Systemic examination was also done focusing on abdominal examination to detect organomegaly or previous operations.

All participants were subjected to laboratory investigations including serum creatinine levels, eGFR, FBG, liver enzymes, serum albumin, serum magnesium, and serum uric acid. Additionally, specific laboratory investigations relevant to our study were done which included a study of HNF1-beta gene mutations with PCR amplification of extracted genomic DNA using capillary electrophoresis.

Radiological investigations were performed (abdominal US, CTUT, and DMSA DPTA), and these were used to classify patients who had CAKUT other than MCDK using the SFU grading system. The SFU grading system of hydronephrosis can be found in Appendix (I) of the supplementary materials.

Sample Collection

Blood samples were collected from 20 patients and their parents and stored in sterile EDTA tubes at 4 C for two days then subjected to DNA extraction. Genomic DNA was extracted from the white blood cells of all patients and their family members using the salting out procedure [22]. Details regarding the reagents used and the procedural technique can be found in Appendix (I) of supplementary materials.

The purity of the extracted DNA was determined using NanoDrop 2000, Thermo Scientific, USA. The nine exons of HNF1B gene were amplified with PCR. Further details on the amplification process can be found in Appendix (I) of supplementary materials. Extracted DNA was subjected to direct sequencing using QIAquick PCR purification kit (QIAGEN, Germany), BigDye Terminator kit (Applied Biosystem, Foster City, CA, USA), and CENTRI-SEP purification spin columns (Applied Biosystem, Foster City, CA, USA). For further details on the principle, steps, and procedures of DNA sequencing, refer to Appendix (I) of supplementary materials.

Statistical Analysis

Statistical calculations were done using computer programs Microsoft Excel and the Statistical Package for Social Sciences (SPSS). Quantitative variables were presented in the form of mean ± standard deviation (SD). Qualitative variables were described as frequencies and percentages.

Inferential analyses were done for quantitative variables using unpaired T-test in cases of two dependent groups. For parametric variables for comparing categorical data, Chi-Square test (X2) was performed. The Mann-Whitney test was used to compare two unrelated samples. The Kruskal-Wallis test was used to compare three or more unrelated samples. The p-value was calculated, with values greater than or equal to 0.05 considered non-significant, and those less than 0.05 deemed statistically significant.

Sample Size

The sample size is determined using Epi info (version 3.5.1, 2008). Based on confidence level of 90%, power 80%, and 10% detection rate of HNF1B mutations in CAKUT cases design effect 1, the minimum required sample size is 20 patients required to be enrolled in the study [23].

Sample size calculation:

$$n = [DEFF*Np(1-p)]/ [(d^2/Z^2_{1-\alpha/2}*(N-1)+p*(1-p)]$$

Extensive details on the statistical tests done, as well as data analysis and calculation of quantitative data, are found in Appendix (I) of supplementary materials.

Results

The distribution of the studied groups is depicted in Table S1. A total of 20 patients were studied and distributed according to their diagnosis. 5 of 20 patients had MCDK, while the other 15 had other types of CAKUT.

The age and sex of the studied groups are shown in Table S2. A total of 20 patients from 19 unrelated families were divided according to their diagnosis into those who had 5 cases with MCDK and another group of 15 with CAKUT other than MCDK. The mean age was 3.76 in the MCDK group, while that of the CAKUT group was 6.34 (p=0.343). Regarding the sex, in the MCDK group there were 4 males and only 1 female and in the CAKUT group there were 11 males and only 4 females (p=0.313). There were no significant differences between both groups regarding age and sex.

Table (I): Association between SNPs and the development of CAKUT:

SNPs			CAKUT (n=20)	Controls (n=50)	p
rs2229295	Allele	C	28 (70%)	78 (78%)	0.49
		A	12 (30%)	22 (22%)	
	Genotype	C/C	10 (50%)	30 (60%)	0.24
		C/A	7 (35%)	18 (36%)	
		A/A	3 (15%)	2 (4%)	
rs1800929	Allele	A	27 (67.5%)	75 (75%)	0.42
		G	13 (32.5%)	25 (25%)	
	Genotype	A/A	9 (45%)	24 (48%)	0.88
		A/G	9 (45%)	23 (46%)	
		G/G	2 (10%)	3 (6 %)	
rs3110641	Allele	C	22 (55%)	64 (64%)	0.61
		T	18 (45%)	36 (36%)	
	Genotype	C/C	6 (30%)	12 (24%)	0.44
		C/T	10 (50%)	36 (72%)	
		T/T	4 (20%)	2 (4%)	

The most frequent polymorphisms detected on DNA sequencing were rs3110641 (70%), rs1800929 (55%), and rs2229295 (50%). No considerable difference was found regarding the allele frequency of the three mentioned SNPs between the cases and controls (p=0.61, p=0.42, p=0.49, Table I). When comparing the CAKUT group to the healthy controls, we noted no significant differences regarding the frequency of genotype in the three reported SNPs (p=0.44, p=0.88, p=0.24, Table I). Although we noted that the mutant homozygous genotypes were more frequent in the CAKUT group across all three SNPs as opposed to the controls, the results were not statistically significant (Table I). More detailed sequencing results can be found in Appendix II (Supplementary materials).

Table (II): Genotype frequency between cases and controls:

SNP	Model	Genotype	CAKUT	controls	OR (95% CI)	P-value
rs222295	General	C/C	10 (50%)	30 (60%)	1.00	0.21
		C/A	7 (35%)	18 (36%)	0.69 (0.21-2.21)	
		A/A	3 (15%)	2 (4%)	0.18 (0.03-1.25)	
	Dominant	C/C	10 (50%)	30 (60%)	1.00	0.26
		C/A-A/A	10 (50%)	20 (40%)	0.53 (0.18-1.58)	
	Recessive	C/C-C/A	17 (85%)	48 (96%)	1.00	0.1
		A/A	3 (15%)	2 (4%)	0.21 (0.03-1.37)	
rs1800929	General	A/A	9 (45%)	24 (48%)	1.00	0.7
		A/G	9 (45%)	23 (46%)	0.75 (0.24-2.33)	
		G/G	2 (10%)	3 (6%)	0.44 (0.06-3.16)	
	Dominant	A/A	9 (45%)	24 (48%)	1.00	0.5
		A/G-G/G	11 (55%)	26 (52%)	0.69 (0.23-2.07)	
	Recessive	A/A-A/G	18 (90%)	47 (94%)	1.00	0.49
		G/G	2 (10%)	3 (6%)	0.51 (0.08-3.34)	
rs3110641	General	C/C	6 (30%)	12 (24%)	1.00	0.091
		C/T	10 (50%)	36 (72%)	1.20 (0.32-4.54)	
		T/T	4 (20%)	2 (4%)	0.17 (0.02-1.28)	
	Dominant	C/C	6 (30%)	12 (24%)	1.00	0.88

		C/T-T/T	**14 (70%)**	**38 (76%)**	**0.90 (0.25-3.28)**	
	Recessive	**C/C-C/T**	**16 (80%)**	**48 (96%)**	**1.00**	**0.03**
		T/T	**4 (20%)**	**2 (4%)**	**2.46 (1.05-5.88)**	

The three SNPs were analyzed after investigating genotype-based associations comparing cases to controls in terms of genotype under the general, dominant, and recessive models; and it was found that across the three SNPs, the only significant results pertained to the rs3110641 SNP in the recessive model of inheritance, being notably more frequent in the T/T recessive mode in the CAKUT group (20%) as opposed to the controls (4%, p=0.03; Table II). The odds ratio (OR) of the CAKUT mutant genotype was 2.6 times higher in reference to the pooled count of heterozygous and wild allele homozygous in the recessive model of the disease (p=0.03, OR 2.46, 95% CI 1.05-5.88; Table II). Haplotype association with CAKUT showed that the AGT haplotype was significantly associated with CAKUT (p=0.045, OR 2.4, p5% CI 1.3 – 4.7). While the other two SNPs failed to show any significant results across all modes (p>0.05).

Table (III): Association between SNPs and severe CAKUT:

SNP			**Non-severe CAKUT (N=12)**	**Severe CAKUT (N=8)**	**p**
rs2229295	**Allele**	**C**	**19 (79%)**	**8 (50 %)**	**0.07**
		A	**5 (21%)**	**8 (50 %)**	
	Genotype	**C/C**	**7 (58%)**	**3 (38%)**	**0.04**
		C/A	**5 (42%)**	**2 (25%)**	
		A/A	**0 (0%)**	**3 (38%)**	
rs1800929	**Allele**	**A**	**18 (75%)**	**9 (56%)**	**0.43**
		G	**6 (25%)**	**7 (44%)**	
	Genotype	**A/A**	**6 (50 %)**	**3 (38%)**	**0.13**
		A/G	**6 (50 %)**	**3 (38%)**	
		G/G	**0 (0%)**	**2 (25%)**	
rs3110641	**Allele**	**C**	**15 (62%)**	**7 (44%)**	**0.23**

		T	**9 (38%)**	**9 (56%)**	
	Genotype	**C/C**	**4 (33%)**	**2 (25%)**	**0.28**
		C/T	**7 (58%)**	**3 (38%)**	
		T/T	**1 (8%)**	**3 (38%)**	
rs35913775	**Allele**	**G**	**19 (79%)**	**11 (69%)**	**0.31**
		C	**5 (21%)**	**5 (31%)**	
	Genotype	**G/G**	**8 (67%)**	**5 (62%)**	**0.53**
		G/C	**3 (25%)**	**1 (12%)**	
		C/C	**1 (8%)**	**2 (25%)**	

The frequency of different alleles and genotypes were comparable between the severe CAKUT group and the non-severe CAKUT group regarding the four most frequent SNPs; however, rs2229295 genotype A/A was significantly associated with severe CAKUT with 38% of patients in the severe CAKUT group exhibiting that genotype in the recessive mode as opposed to 0% of the non-severe group (p=0.04; Table III)

Table (IV): Genotype frequency between severe CAKUT and non-severe CAKUT:

	Model	Genotype	Non-severe CAKUT (N=12)	Severe CAKUT (N=8)	OR (95% CI)	P-value
rs2229295	General	C/C	7 (58.3%)	3 (37.5%)	1.00	0.042
		C/A	5 (41.7%)	2 (25%)	0.93 (0.11-7.82)	
		A/A	0 (0%)	3 (37.5%)	NA (0.00-NA)	
	Dominant	C/C	7 (58.3%)	3 (37.5%)	1.00	0.36
		C/A-A/A	5 (41.7%)	5 (62.5%)	2.33 (0.37-14.61)	
	Recessive	C/C-C/A	12 (100%)	5 (62.5%)	1.00	0.012
		A/A	0 (0%)	3 (37.5%)	0.02 (0.00-0.05)	
rs1800929	General	A/A	6 (50%)	3 (37.5%)	1.00	0.13
		A/G	6 (50%)	3 (37.5%)	1.00 (0.14-7.10)	
		G/G	0 (0%)	2 (25%)	NA (0.00-NA)	
	Dominant	A/A	6 (50%)	3 (37.5%)	1.00	0.58
		A/G-G/G	6 (50%)	5 (62.5%)	1.67 (0.27-10.33)	
	Recessive	A/A-A/G	12 (100%)	6 (75%)	1.00	0.045
		G/G	0 (0%)	2 (25%)	NA (0.00-NA)	
rs3110641	General	C/C	4 (33.3%)	2 (25%)	1.00	0.28
		C/T	7 (58.3%)	3 (37.5%)	0.86 (0.10-7.51)	

		T/T	**1 (8.3%)**	**3 (37.5%)**	**6.00 (0.35-101.57)**	
	Dominant	**C/C**	**4 (33.3%)**	**2 (25%)**	**1.00**	**0.69**
		C/T-T/T	**8 (66.7%)**	**6 (75%)**	**1.50 (0.20-11.09)**	
	Recessive	**C/C-C/T**	**11 (91.7%)**	**5 (62.5%)**	**1.00**	**0.11**
		T/T	**1 (8.3%)**	**3 (37.5%)**	**6.60 (0.54-80.24)**	
rs35913775	**General**	**G/G**	**8 (67%)**	**5 (62%)**	**1.00**	**0.53**
		G/C	**3 (25%)**	**1 (12%)**	**0.40 (0.03-5.15)**	
		C/C	**1 (8%)**	**2 (25%)**	**2.40 (0.16-34.93)**	
	Dominant	**G/G**	**8 (67%)**	**5 (62.5%)**	**1.00**	**0.91**
		G/C-C/C	**4 (33%)**	**3 (37.5%)**	**0.90 (0.13-6.08)**	
	Recessive	**G/G-G/C**	**11 (92%)**	**6 (75%)**	**1.00**	**0.4**
		C/C	**1 (8%)**	**2 (25%)**	**3.00 (0.22-40.93)**	

An analysis of models of disease penetrance is shown in Table IV, showing significantly higher rate of the rs2229295 SNP in the general (p=0.042, OR 0.93, 95% CI 0.11-7.82; Table IV) and recessive A/A models in the severe CAKUT group (p=0.012, OR 0.02, 95% CI 0.00-0.05; Table IV), while rs1800929 was only significantly higher in severe CAKUT group in the recessive G/G model (p=0.045, OR NA, 95% CI 0.00-NA; Table IV). The other SNPs were not significantly different between both groups.

Table (V): Association between SNPs and CAKUT with and without T1DM:

SNP			**CAKUT without T1DM (N=17)**	**CAKUT WITH T1DM (N=3)**	**p**
rs2229295	**Allele**	**C**	**25 (73.5%)**	**2 (33%)**	**0.14**
		A	**9 (26.5%)**	**4 (67%)**	
	Genotype	**C/C**	**9 (53%)**	**1 (33%)**	**0.03**
		C/A	**7 (41%)**	**0 (0%)**	

		A/A	**1 (6%)**	**2(67%)**	
rs1800929	**Allele**	**A**	**26 (76.5%)**	**1 (17 %)**	**0.02**
		G	**8 (23.5%)**	**5 (83 %)**	
	Genotype	**A/A**	**9 (53%)**	**0 (0 %)**	**0.001**
		A/G	**8 (47%)**	**1 (33 %)**	
		G/G	**0 (0 %)**	**2(67 %)**	
rs3110641	**Allele**	**C**	**21 (61.7%)**	**1 (17 %)**	**0.1**
		T	**13 (38.3%)**	**5 (83 %)**	
	Genotype	**C/C**	**6 (35%)**	**0 (0 %)**	**0.07**
		C/T	**9 (53%)**	**1 (33 %)**	
		T/T	**2 (12%)**	**2 (67 %)**	
rs35913775	**Allele**	**G**	**27 (79.5%)**	**2 (33 %)**	**0.06**
		C	**7 (20.5%)**	**4 (67 %)**	
	Genotype	**G/G**	**11 (64.7%)**	**1 (33.3%)**	**0.02**
		G/C	**5 (29.4%)**	**0 (0.0%)**	
		C/C	**1 (6%)**	**2 (66.7%)**	

The SNP rs2229295 showed a significant association with CAKUT with T1DM being higher in that group (67%) unlike the group without T1DM (6%, p=0.03). The mutant allele of SNP rs1800929 was significantly higher in the group without T1DM, but the mutant genotype was considerably higher in the group with T1DM (p=0.001). Concerning the SNP rs35913775 mutant genotype (C/C), it was revealed to be notably higher in the CAKUT with T1DM group (66.7%) as opposed to those without T1DM (6%, p=0.02; Table V).

Table (VI): Genotype frequency between CAKUT and T1DM:

	Model	Genotype	group=CAKUT	group=T1DM	OR (95% CI)	P-value
rs2229295	General	C/C	7 (46.7%)	1 (33.3%)	1.00	0.02
		C/A	7 (46.7%)	0 (0%)	0.34 (0.03-4.67)	
		A/A	1 (6.7%)	2 (66.7%)	14.00 (1.58-338.80)	
	Dominant	C/C	7 (46.7%)	1 (33.3%)	1.00	0.67
		C/A-A/A	8 (53.3%)	2 (66.7%)	1.75 (0.13-23.70)	
	Recessive	C/C-C/A	14 (93.3%)	1 (33.3%)	1.00	0.025
		A/A	1 (6.7%)	2 (66.7%)	28.00 (1.21-648.84)	
rs1800929	General	A/A	9 (52.9%)	0 (0%)	1.00	0.0049
		A/G	8 (47.1%)	1 (33.3%)	0.7 (0.05-2.7)	
		G/G	0 (0%)	2 (66.7%)	0.02 (0.01-0.56)	
	Dominant	A/A	9 (52.9%)	0 (0%)	1.00	0.045

		A/G-G/G	**8 (47.1%)**	**3 (100%)**	**1.03 (1.32-17.56)**	
	Recessive	**A/A-A/G**	**17 (100%)**	**1 (33.3%)**	**1.00**	**0.0024**
		G/G	**0 (0%)**	**2 (66.7%)**	**0.2 (0.02-0.51)**	
rs3110641	**General**	**C/C**	**6 (35.3%)**	**0 (0%)**	**1.00**	**0.088**
		C/T	**9 (52.9%)**	**1 (33.3%)**	**0.17 (0.001-3.87)**	
		T/T	**2 (11.8%)**	**2 (66.7%)**	**0.02 (0.003-0.77)**	
	Dominant	**C/C**	**6 (35.3%)**	**0 (0%)**	**1.00**	**0.12**
		C/T-T/T	**11 (64.7%)**	**3 (100%)**	**0.34 (0.12-1.67)**	
	Recessive	**C/C-C/T**	**15 (88.2%)**	**1 (33.3%)**	**1.00**	**0.049**
		T/T	**2 (11.8%)**	**2 (66.7%)**	**15.00 (1.90-251.07)**	
rs35913775	**General**	**G/G**	**11 (64.7%)**	**1 (33.3%)**	**1.00**	**0.045**
		G/C	**5 (29.4%)**	**0 (0%)**	**0.003 (0.006-1.7)**	
		C/C	**1 (5.9%)**	**2 (66.7%)**	**22.00 (6.94-515.90)**	
	Dominant	**G/G**	**11 (64.7%)**	**1 (33.3%)**	**1.00**	**0.31**

		G/C-C/C	**6 (35.3%)**	**2 (66.7%)**	**3.67 (0.27-49.29)**	
	Recessive	**G/G-G/C**	**16 (94.1%)**	**1 (33.3%)**	**1.00**	**0.019**
		C/C	**1 (5.9%)**	**2 (66.7%)**	**32.00 (1.39-737.49)**	

Analysis of the models of inheritance showed that all SNPs (rs2229295, rs1800929, rs3110641, and rs35913775) were significantly associated with T1DM in the recessive model (p=0.025, p=0.0024, p=0.049, p=0.019), while only rs2229295 and rs1800929 were associated with T1DM in the general model (p=0.02, p=0.0049), and only rs1800929 was associated with T1DM in the dominant model (p=0.045; Table VI).

Discussion

Congenital anomalies of the kidney and the urinary tract represent a large spectrum of disorders that result in structural anomalies of the kidneys and urinary tract [24]. They may be associated in some instances with other extrarenal manifestations such as liver disease, diabetes, and hypomagnesemia [11, 25-27]. GWAS have not yet yielded a robust body of evidence regarding the genetic basis of CAKUT due to lack of sufficient sample size [10]; however, the genetics of CAKUT have been repeatedly explored in the available literature, with many SNPs and CNVs being identified and implicated in the development of CAKUT [11, 28]. CAKUT can be syndromic or non-syndromic, of the first genes implicated in CAKUT were HNF1B and PAX2 [29]. HNF1B is a transcription factor involved in the embryogenesis of various tissues including the pancreas, liver, and the Wolffian duct, from which the urinary tract and the kidneys develop [30]. HNF1B along with PAX2 seem to be collectively associated with at least 15% of cases of CAKUT, with HNF1B gene mutations being significantly associated with cystic kidneys [31-33].

In this cross-sectional study, we aimed to address the systematic mutational analyses of HNF1B in CAKUT, primarily focusing on elucidating the prevalence of HNF1B mutations and understanding the clinical manifestations of HNF1B-assoicated diseases.

In our current study, we distributed the participants over two groups according to the type of CAKUT and found that the percentage of those who were diagnosed with MCDK was 25% and 75% for those who were diagnosed with other types of CAKUT. Stonebrook et al. reported that MCDK is one of the most common congenital kidney anomalies,

representing a rather heavy burden in the context of bilaterality, which causes either intrauterine fetal death or end-stage renal disease (ESRD) in children who survive postnatally [34]. Feldenberg et al. expounded on the incidence and clinical course of MCDK, revealing that it was the most common CAKUT occurring in 1 in 3,640 births; with bilaterality or association with a contralateral different anomaly resulting in a worse prognosis and leading to 50% of patients reaching ESRD or renal failure in childhood [35]. We can explain the discrepancy in the number of patients who had MCDK as opposed to those who had other types of CAKUT, with those who had MCDK being lower in numbers, by taking into consideration the sociodemographic data of the studied sample, as those who were diagnosed with MCDK were mostly males with only 1 female in that group, which aligns with the findings of Rudnik-Schoneborn et al. who noted a male preponderance for MCDK [36]. Also, the age of patients in the MCDK group was at a mean of 3.76 years, while that of the CAKUT group was, though not significantly (p=0.343), higher at 6.34 years, bearing in mind that cases who are affected with bilateral MCDK either perish intrauterine or die shortly after birth or early in childhood [34, 35]. Furthermore, there's a tendency for unilateral MCDK to regress with time. Hains et al. reported that in 74% of patients who were diagnosed with unilateral MCDK, complete involution was observed in the first two years [37].

We screened 20 patients from 19 unrelated families for the HNF1B gene who were impacted by CAKUT, yet we found no mutations. However, screening revealed 8 SNPs. We found an intronic variant (intron 5) c.1207-51C>A present in heterozygous form in patient (1) of family (1). Another SNP was a homozygous intronic variant (intron 6) c.1339+27T>C, (rs2107133) in two patients, patients (2, 3) of (family 2). Moreover, a heterozygous synonymous Variant c.1413C>T, P471 (Pro)> P471 (Pro) in (exon 7), (rs140781855) in patient (4) of family (3). There is also an intronic variant (intron 8) c.1653+21G>A, (rs375700762) present in heterozygous form in the patient (5) of family (4). an intronic variant c.1653+47_1653+48insC (intron 8) rs35913775. It was detected in homozygous form in three patients, patient (2, 3) of family (2), and patient (6) family (5), and detected in heterozygous form in five patients; patient (7), patient (8), patient (9), patient (10), patient (11) of family (6-10). Moreover, c.1654-22C>T intronic variant (intron 8) (rs3110641) was detected in homozygous form in four patients; patient (2), patient (3), patient (7), patient (9) families (2,6, 8) and was detected in heterozygous form in 10 patients; patients (1, 5, 6, 8, 12-17) families (1, 4, 5, 7, 11-16). c.*99C>A, (rs2229295) detected in homozygous state in three patients while in heterozygous state in six patients. c.*100A>G, (rs1800929) detected in homozygous state in two patients while in heterozygous state in 9 patients. The most frequent polymorphisms were rs3110641 (70%), rs1800929 (55%), rs2229295 (50%), and rs35913775 (40%). In line with our findings, Goda et al. reported on SNP C>A in rs2229295 in HNF1B, ascertaining that this SNP specifically could have a role in the binding of two miRNAs, hsa-miR-214-5p and has-miR-550a-5p, which affects luciferase activity on luciferase reporter assay. This is highly

relevant to our study as it suggests that the studied SNP in rs2229295 affects the expression of the target gene (HNF1B) by suppressing its expression, leading to different outcomes, some of which may be anomalous. They also studied another SNP in the gene HNF1B, which was of interest to us, SNP rs1800929; however, they did not find that it significantly affected luciferase activity, i.e. it did not bind to the HNF1B 3'UTR [12]. Edghill et al. and Cereghini et al. both inferred that HNF1B expression indeed affects multiple tissues, such as the pancreas, liver, and kidneys [38, 39]. Edghill et al. further revealed that HNF1B gene expression has multitudes of impact on various tissues including the kidney, often leading to renal cysts or chronic renal impairment [38]. There is currently no solid report on the association between our studied SNPs and CAKUT. Goda et al., in their report on the impact of HNF1B SNPs on type 2 diabetes mellitus (T2DM), in which they concluded that there is a possibility that dysregulation of the HNF1B gene expression due to single-nucleotide variations leads to a variance in susceptibility to T2DM, also had no predecessor reports, so relevance is drawn from predictions based on the current knowledge on gene expression and the use of luciferase reporter assay in that context [12, 40].

Regarding the association between the identified SNPs and the development of CAKUT, we found no significant differences in terms of the allele frequency in different SNPs between both the CAKUT group and the controls (p=0.49, p=0.42, p=61). Moreover, when comparing the most frequent mutant homozygous genotypes between CAKUT group and controls, it was revealed to us that despite the higher frequency in CAKUT, the results were not statistically significant (p=0.24, p=0.88, p=0.44). The current literature lacks any reports on the association between the SNPs we identified and CAKUT. However, Ahn et al. reported in the context of CAKUT on single-nucleotide variants that were deemed pathogenic to the gene HNF1B, which they noted was the most commonly identified genetic cause of CAKUT at a rate of 46.2% of patients with pathogenic SNPs, namely c.541C>T, which lead to bilateral renal hypodysplasia in one patient causing ESRD at 1 month of age, as well as c1103_1116del nucleotide which also negatively impacted the expression of HNF1B leading to complex MCDK, with left MCDK and right renal hypodysplasia with renal cysts, causing CKD [11]. Gad et al. who studied the SNPs we identified as significant, found that rs2229295, rs1800929, and rs3110641, all were associated with homozygous genotypes that were strongly associated with chromophobe renal cell carcinoma ($p<0.02$), especially c.1654-22C>T (rs3110641) and c*99C>A (rs2229295), which accounted for 12% and 7.6% of cases, respectively, out of 46 patients with chromophobe RCC [41]. Adding on, Calderon-Margalit et al. reported in a large cohort that involved 2959 recruits, that CAKUT was associated with a significantly higher risk of urinary tract cancer, especially before the age of 45 in women, and after the age of 45 in men [42]. These findings can be explained by the fact that the HNF1B gene, which is involved in renal and urinary tract development, is situated in the 17q12 region, a region which is highly prone to genomic rearrangement, making it susceptible to mutations that change the course of gene expression, often towards pathogenicity [43-46].

According to the results of our genotype-based investigations, our analysis revealed that there was statistical significance regarding the proportion of the mutant homozygous genotype of rs3110641 in the recessive model of disease penetrance in the CAKUT group (p=0.03). Gad et al. reported on the same SNP in the context of RCC and ascertained that a mutant homozygous genotype of rs3110641 was significantly associated with a pathogenic expression of HNF1B and subsequently to chromophobe RCC in 12% of patients, as well as 2.2% of cases with oncocytoma, and 1.1% of cases with papillary RCC. Although these findings do not address CAKUT explicitly, they are aligned with our results regarding the pathogenic effect of a mutant homozygous genotype of rs3110641 in renal tissues [41].

Haplotype association with CAKUT showed that AGT haplotype was notably related to CAKUT (OR 2.4, 95% CI 1.3-4.7; p=0.045). Capone et al. relayed data aligning with our findings, stating that mutations in the AGT and AGTR1 genes, which are genes encoding parts of the renin-angiotensin system, are involved in renal tubular dysgenesis (RTD) as well as other types of CAKUT [29]. Gribouval et al., who investigated the RAS genes in the context of kidney disease, ascertained in agreement with our data that they detected a homozygous G→A transition causing a missense mutation in the AGT gene which they hypothesized to have caused a severe change in splicing, causing a truncating mutation of the serpin domain of the protein, which is a crucial domain for the cleavage of the corresponding peptide. Based on these findings, they determined that in the gene AGT, a nucleotide alteration (1124G→A) at the last nucleotide of the exon, was responsible for RTD in the studied family, comparing this with 120 healthy controls with a similar ethnic background (Turkish), who did not manifest this mutation [28]. These findings are suggested to stem from the fact that angiotensin is an essential part of kidney development, which was reported in both human and mouse models, hence, the impaired production of angiotensin or mutations in the corresponding genes may severely impact tissue development by defects in the production of angiotensin [28, 47, 48]. However, according to the work of Mahieu-Caputo et al., autosomal recessive renal tubular dysgenesis was observed in those without an impairment in the RAS system, suggesting that there may be another explanation. Hypothetically, RAS dysfunction leads to chronic low perfusion pressure in the fetal kidney leading to renal dysgenesis [28, 49].

Furthermore, we analyzed our data attempting to draw correlations between the different alleles and genotypes, of the four most common SNPs, and the severity of CAKUT and we found that the frequency of alleles and genotypes was in fact comparable between the two groups and no statistical significance was noted (p>0.05); however, we reported that the SNP rs2229295 showed a strong association with severe CAKUT in both the general model of inheritance, as well as the recessive one (p=0.04, p=0.012, respectively). Moreover, we noted that the SNP rs1800929, although not in the general model, demonstrated a significant association with severe CAKUT in the recessive model of inheritance

(p=0.045). Ahn et al. expounded on patients with single-nucleotide variants of several genes including HNF1B and the correlation between pathogenic variants of these genes and the phenotypic severity of CAKUT, and they found that patients did not differ significantly in terms of CAKUT phenotypes in terms of the genetic abnormalities. However, they observed a more frequent rate of bilateral anomalies of the kidneys in patients who pathogenic SNPs (76.9%) as opposed to those without pathogenic SNPs (51.9%), but still, the data was not statistically significant (p=0.091). Notably, they found that patients who had pathogenic SNPs exhibited syndromic CAKUT (p<0.001) [11]. Nishimura et al. studied the genetic basis of CAKUT, extrapolating data from both mouse and human models; although not in the context of HNF1B or the SNPs we shed light on, they found that the inheritance of CAKUT did not follow a typical Mendelian pattern and instead showed an incomplete penetrance [50].

Given the rate of diabetes mellitus reported in the family history of the study subjects (75% of the cases), as well as the documentation of type 1 diabetes mellitus (T1DM) in 15% of the cases, we found it appropriate to study the association between the four most frequent SNPs and T1DM. We noted that the frequency of the mutant allele in the SNP rs2229295 was, indeed, significantly higher in CAKUT with T1DM (p=0.03). Similarly, in the general model of inheritance, we found that the frequency of the mutant allele and genotype of SNP rs1800929 was also significantly higher in CAKUT with T1DM (p=0.02, p=0.001, respectively). Adding to that, the SNP rs35913775 mutant genotype was considerably more frequent in the CAKUT with T1DM group (p=0.02). Upon analyzing the disease penetrance model, we showed that all four SNPs were associated with CAKUT with TIDM in the recessive mode of inheritance (p=0.025, p=0.0024, p=0.049, p=0.019, respectively). Simiilarly, Moszynska et al. inferred that the SNP rs2229295 (C>A) in HNF1B appears to be associated with the development of diabetes; however, they reported on its association with type 2 diabetes mellitus, not type 1 [20]. Ovsyannikova et al. put out a case report on a 27-year-old female patient who, upon finding a mutation in the HNF1B gene (rs138986885), was diagnosed with maturity-onset diabetes of the young type 5 (MODY-5); however, the patient was also a carrier of heterozygous SNP rs2476601 in the tyrosine phosphatase gene, which is a risk factor for type 1 diabetes mellitus. Follow up revealed rapid deterioration in pancreatic cell function with reduced fasting C-peptide, which necessitated insulin therapy in this patient, leading to a diagnosis of type 1 diabetes mellitus as well [51]. HNF1B mutations could be associated with diabetes leading to a picture of MODY as well as renal cysts, which is referred to as MODY-5. Francis et al. described a case who was previously diagnosed with T1DM, without obesity or any features of metabolic syndrome, as well as liver disease and hypomagnesemia [52].

LIMITATIONS

Conclusion

In conclusion, our study aimed to explore the genetic basis of HNF1B mutations in the context of CAKUT. Although we found no pathogenic mutations regarding HNF1B, we managed to identify several SNPs, four of which were revealed to be associated with CAKUT to various degrees. Notably, we observed one of the reported SNPs to be associated with the development of CAKUT with significance in the recessive model of inheritance, namely rs3110641. Moreover, we found that another SNP, rs2229295, was significantly associated with a severe form of CAKUT in the general and the recessive models of inheritance. Additionally, we found an association between an AGT haplotype and CAKUT, which was statistically significant. Finally, we concluded that the rs35913775 mutant genotype was significantly associated with CAKUT-T1DM. While other studies investigated some of the SNPs we chose as the center of our study in relation to diabetes and renal cancer; to our knowledge, no previous studies explored these SNPs in the context of CAKUT. Thus, more comprehensive research is crucial to illuminate the role of HNF1B mutations in the development of CAKUT to validate our findings and augment the current body of literature on the genetics of CAKUT, in hopes to establish a definitive understanding of the genes involved, the mechanisms of anomalous development, and prenatal identification of complex severe renal anomalies that may pose unfavorable outcomes in the affected fetus.

References

1. Harambat, J., et al., *Epidemiology of Chronic Kidney Disease in Children.* 2011. **27**(3): p. 363-373.
2. Kang, H.G., et al., *KNOW-Ped CKD (KoreaN Cohort Study for Outcomes in Patients With Pediatric CKD): Design and Methods.* 2016. **17**(1).
3. Murugapoopathy, V. and I.R. Gupta, *A Primer on Congenital Anomalies of the Kidneys and Urinary Tracts (CAKUT).* 2020. **15**(5): p. 723-731.
4. Walawender, L., B. Becknell, and D.G. Matsell, *Congenital Anomalies of the Kidney and Urinary Tract: Defining Risk Factors of Disease Progression and Determinants of Outcomes.* 2023. **38**(12): p. 3963-3973.
5. Kerecuk, L., M.F. Schreuder, and A.S. Woolf, *Renal Tract Malformations: Perspectives for Nephrologists.* 2008. **4**(6): p. 312-325.
6. Stein, D. and E. McNamara, *Congenital Anomalies of the Kidneys and Urinary Tract.* 2022. **49**(3): p. 791-798.
7. Chiodini, B., et al., *Clinical Outcome of Children With Antenatally Diagnosed Hydronephrosis.* 2019. **7**.
8. Kumar, B.H., et al., *Clinical Spectrum of Congenital Anomalies of Kidney and Urinary Tract in Children.* 2019. **56**(7): p. 566-570.
9. Walker, E.Y.X., P. Winyard, and M. Marlais, *Congenital Anomalies of the Kidney and Urinary Tract: Antenatal Diagnosis, Management and Counselling of Families.* 2023. **39**(4): p. 1065-1075.
10. Nicolaou, N., et al., *Genetic, Environmental, and Epigenetic Factors Involved in CAKUT.* 2015. **11**(12): p. 720-731.
11. Ahn, Y.H., et al., *Targeted Exome Sequencing Provided Comprehensive Genetic Diagnosis of Congenital Anomalies of the Kidney and Urinary Tract.* 2020. **9**(3): p. 751.
12. Goda, N., et al., *Polymorphism in microRNA-binding Site in HNF1B Influences the Susceptibility of Type 2 Diabetes Mellitus: A Population Based Case–control Study.* 2015. **16**(1).
13. Madariaga, L., et al., *Variable Phenotype in HNF1B Mutations: Extrarenal Manifestations Distinguish Affected Individuals From the Population With Congenital Anomalies of the Kidney and Urinary Tract.* 2018. **12**(3): p. 373-379.
14. Stiles, C.E., et al., *De Novo HNF1 Homeobox B Mutation as a Cause for Chronic, Treatment-Resistant Hypomagnesaemia.* 2018. **2018**.
15. Chandra, S., S. Srinivasan, and J. Batra, *Hepatocyte Nuclear Factor 1 Beta: A Perspective in Cancer.* 2021. **10**(5): p. 1791-1804.
16. Larson, N.B., et al., *Network-Directed Cis-Mediator Analysis of Normal Prostate Tissue Expression Profiles Reveals Downstream Regulatory Associations of Prostate Cancer Susceptibility Loci.* 2017. **8**(49): p. 85896-85908.
17. Vashisht, S., C. Parisi, and C.L. Winata, *Computational Analysis of Congenital Heart Disease Associated SNPs: Unveiling Their Impact on the Gene Regulatory System.* 2024.
18. Bates, C.M., *Role of Fibroblast Growth Factor Receptor Signaling in Kidney Development.* 2011. **301**(2): p. F245-F251.
19. Yosypiv, I.V., *Congenital Anomalies of the Kidney and Urinary Tract: A Genetic Disorder?* 2012. **2012**: p. 1-10.

20. Moszyńska, A., et al., *SNPs in microRNA Target Sites and Their Potential Role in Human Disease.* 2017. **7**(4): p. 170019.
21. Elm, E.v., et al., *Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) Statement: Guidelines for Reporting Observational Studies.* 2007. **335**(7624): p. 806-808.
22. Miller, S.A., D.D. Dykes, and H.F. Polesky, *A Simple Salting Out Procedure for Extracting DNA From Human Nucleated Cells.* 1988. **16**(3): p. 1215-1215.
23. Raaijmakers, A., et al., *Criteria for HNF1B Analysis in Patients With Congenital Abnormalities of Kidney and Urinary Tract.* 2014. **30**(5): p. 835-842.
24. Ichikawa, I., et al., *Paradigm Shift From Classic Anatomic Theories to Contemporary Cell Biological Views of CAKUT.* 2002. **61**(3): p. 889-898.
25. Vivante, A., et al., *Single-Gene Causes of Congenital Anomalies of the Kidney and Urinary Tract (CAKUT) in Humans.* 2014. **29**(4): p. 695-704.
26. Dressler, G.R., *The Cellular Basis of Kidney Development.* 2006. **22**(1): p. 509-529.
27. Sanna-Cherchi, S., et al., *Genetic Approaches to Human Renal Agenesis/Hypoplasia and Dysplasia.* 2007. **22**(10): p. 1675-1684.
28. Gribouval, O., et al., *Mutations in Genes in the Renin-Angiotensin System Are Associated With Autosomal Recessive Renal Tubular Dysgenesis.* 2005. **37**(9): p. 964-968.
29. Capone, V., et al., *Genetics of Congenital Anomalies of the Kidney and Urinary Tract: The Current State of Play.* 2017. **18**(4): p. 796.
30. Coffinier, C., et al., *Essential Role for the Homeoprotein vHNF1/HNF1β in Visceral Endoderm Differentiation.* 1999. **126**(21): p. 4785-4794.
31. Madariaga, L., et al., *Severe Prenatal Renal Anomalies Associated With Mutations in HNF1B or PAX2 Genes.* 2013. **8**(7): p. 1179-1187.
32. Weber, S., et al., *Prevalence of Mutations in Renal Developmental Genes in Children With Renal Hypodysplasia.* 2006. **17**(10): p. 2864-2870.
33. Thomas, R., et al., *HNF1B and PAX2 Mutations Are a Common Cause of Renal Hypodysplasia in the CKiD Cohort.* 2011. **26**(6): p. 897-903.
34. Stonebrook, E., M. Hoff, and J.D. Spencer, *Congenital Anomalies of the Kidney and Urinary Tract: A Clinical Review.* 2019. **5**(3): p. 223-235.
35. Feldenberg, L.R. and N.J. Siegel, *Clinical Course and Outcome for Children With Multicystic Dysplastic Kidneys.* 2000. **14**(12): p. 1098-1101.
36. Rudnik-Schöneborn, S., et al., *Clinical Features of Unilateral Multicystic Renal Dysplasia in Children.* 1998. **157**(8): p. 666-672.
37. Hains, D.S., et al., *Management and Etiology of the Unilateral Multicystic Dysplastic Kidney: A Review.* 2009. **24**(2): p. 233-241.
38. Edghill, E.L., *Mutations in Hepatocyte Nuclear Factor-1 and Their Related Phenotypes.* 2005. **43**(1): p. 84-90.
39. Cereghini, S., *Liver-enriched Transcription Factors and Hepatocyte Differentiation.* 1996. **10**(2): p. 267-282.
40. Liu, Y., J. Liao, and Q. Lu, *Laboratory Methods in Epigenetics.* 2015: p. 7-35.
41. Gad, S., et al., *Mutations in BHD and TP53 Genes, but Not in HNF1β Gene, in a Large Series of Sporadic Chromophobe Renal Cell Carcinoma.* British Journal of Cancer, 2006. **96**(2): p. 336-340.

42. Calderon-Margalit, R., et al., *Congenital Anomalies of the Kidney and Urinary Tract and Adulthood Risk of Urinary Tract Cancer.* Kidney International Reports, 2021. **6**(4): p. 946-952.
43. Sharp, A.J., et al., *Discovery of previously unidentified genomic disorders from the duplication architecture of the human genome.* Nature Genetics, 2006. **38**(9): p. 1038-1042.
44. Mefford, H.C., et al., *Recurrent Reciprocal Genomic Rearrangements of 17q12 Are Associated with Renal Disease, Diabetes, and Epilepsy.* The American Journal of Human Genetics, 2007. **81**(5): p. 1057-1069.
45. Lokmane, L., et al., *vHNF1 functions in distinct regulatory circuits to control ureteric bud branching and early nephrogenesis.* Development, 2010. **137**(2): p. 347-357.
46. Massa, F., et al., *Hepatocyte nuclear factor 1β controls nephron tubular development.* Development, 2013. **140**(4): p. 886-896.
47. Tufro-McReddie, A., et al., *Ontogeny of Type 1 Angiotensin II Receptor Gene Expression in the Rat.* 1993. **91**(2): p. 530-537.
48. Okubo, S., et al., *Angiotensinogen Gene Null-Mutant Mice Lack Homeostatic Regulation of Glomerular Filtration and Tubular Reabsorption.* 1998. **53**(3): p. 617-625.
49. Mahieu-Caputo, D., et al., *Twin-to-Twin Transfusion Syndrome.* 2000. **156**(2): p. 629-636.
50. Nishimura, H., et al., *Role of the Angiotensin Type 2 Receptor Gene in Congenital Anomalies of the Kidney and Urinary Tract, CAKUT, of Mice and Men.* 1999. **3**(1): p. 1-10.
51. Ovsyannikova, A., et al., *Two Types Diabetes Mellitus: Clinical Case of HNF1B–mody and Type 1 Diabetes in One Patient.* 2020.
52. Francis, Y., et al., *HNF1B-MODY Masquerading as Type 1 Diabetes: A Pitfall in the Etiological Diagnosis of Diabetes.* 2022. **6**(8).

Supplementary Materials

Appendix (I): Methods

SFU system:

It is a grading system of hydronephrosis developed by the Society of Fetal Urology (SFU) that was devised to assess the degree of hydronephrosis. This particular system is thought to be the most common in use and was originally designed for grading neonatal and infant pelvicalyectasis:

- **grade 0**
  - no dilatation, calyceal walls are apposed to each other
- **grade 1** (mild)
  - dilatation of the renal pelvis without dilatation of the calyces (can also occur in the extrarenal pelvis)
  - no parenchymal atrophy
- **grade 2** (mild)
  - dilatation of the renal pelvis (mild) and calyces (pelvicalyceal pattern is retained)
  - no parenchymal atrophy
- **grade 3** (moderate)
  - moderate dilatation of the renal pelvis and calyces
  - blunting of fornices and flattening of papillae
  - mild cortical thinning may be seen
- **grade 4** (severe)
  - gross dilatation of the renal pelvis and calyces, which appear ballooned
  - loss of borders between the renal pelvis and calyces
  - renal atrophy seen as cortical thinning

Reagents of salting out technique:

- **Sucrose Triton (2X): pH 7.6**

0.64 M Sucrose (Q-BIO-gene, France)*

0.02 M Tris base (s.d.fine-Chem LTD, India)**

0.01 M $MgCl_2$ (Q-BIO-gene, France)

2% (v/v) Triton X-100 (Eastern Kodak Company, USA)

- **Nuclei Lysis Buffer: pH 8.2**

10 mM Tris base

400 mM NaCl

2 mM EDTA (Ethylene diamine tetraacetic acid)

- **20% Sodium Dodecyl Sulfate (SDS) (Q-BIO-gene, France)**
- **Proteinase K (5 µg/µl) (Finzyme, Finland)***
- **Saturated NaCl (35%)**
- **Absolute Ethanol, Molecular biology grade (95-99%)**
- **Sterile water**

Procedure of salting out:

800µl of blood was taken and put in an EDTA containing polypropylene tube (pH 8.0), so no clotting takes place. Then add 800 µl of cold 2X sucrose triton. Make vigorous vortex shaking. Centrifuge the samples at 13000 rpm for 3 minutes at 4ºC, use a cooling centrifuge (SIGMA-1K15-am Hanz, Germany) **then get rid of the supernatants keeping only the nuclear pellet. Add another 800 µl 1X sucrose triton for further washing of the sample. Revortex, and recentrifuge with the same steps. Remove the supernatant. And keep the pellet. Add 515 µl nuclei lysis buffer, 30 µl of 20% Sodium dodecyl sulphate (SDS) and 20 µl of proteinase K (5µg/µl). Then good mixing. Incubate the mixture 37 ºC overnight to complete the digestion ***(Sambrook, et al., 1989).***

---

* *Q.BIO-Gene:Biomedicals Rue Geiler de Kaysersberg 67400 ILLKIRCH GRAFFENSTADEN, France.*

** *S.D. Fine Chem. LTD: T V Industrial Estate 248 Worli Road, Mumbai, Maharashtra, India.*

**Finzyme: FINNZYMES OY, Keilaranta16 A, 02150 Espoo, Finland*.

***Sigma Laborzentrifugen GmbH An der UnterenSoese 50, 37520 Osterode, Germany.*

The second day after digestion, Add 160µl saturated NaCl solution, do vigorous shaking then centrifuge at 13000 rpm for 10 minutes at 4 ºC, collect the supernatants into sterile tubes. Add double volume of 100% ethanol. Precipitate the DNA by quiet inverting of the tubes.

Use a glass Pasteur pipette to collect DNA. Wash DNA in 70% Ethanol to remove excess salts. Leave the pipette to dry. Dissolve DNA IN 150 µL sterile water at room temperature.

c) Determination of DNA Quality and Quantity and integrity:

With the help of (NanoDrop 2000, Thermo Scientific, USA) *DNA purity and concentration can be determined. The apparatus measures the absorbance by a spectrophotometer system at wavelengths 280 nm and 260 nm ***(Sambrook et al., 1989).*** DNA concentration is estimated by absorbance at 260 nm because DNA absorbs light most strongly at this wave length.

Protein contamination can be estimated. The absorbance of tyrosine and tryptophan is the best at 280 nm. So taking the ratio of DNA absorbance at 260 nm and 280 nm is an estimation of DNA purity. Pure DNA should have a ratio of 1.7-2

The integrity of DNA is determined by examination of DNA on 1% agarose gel stained with (3µ/ml) ethidium bromide. Mix1µl of DNA with suitable volume of loading dye (0.1% bromophenol blue, 0.1M EDTA, 1X electrophoresis buffer, 50% glycerol). Then allow to be electrophoresed and use a UV light for checking DNA integrity. DNA will make clear white bands as it binds to ethedium bromide.

d) Polymerase Chain Reaction (PCR):

The nine exons of HNF1B gene were PCR amplified using the primers listed in the table (S1) and Taq DNA polymerase. The primers were designed by https://bioinfo.ut.ee/primer3-0.4.0/ & checked for overlapping amplification by https://www.ncbi.nlm.nih.gov/tools/primer-blast/

PCR amplification of HNF1B genomic DNA was performed in nine fragments. The suitable annealing temperatures for each primer were listed in ***table (S1).***

Reagents used in PCR

- 10X buffer (Finzyme, Finland)

100 mM Tris-HCl, pH 8.8

*3411Silverside Rd # 100, Wilmington, DE 19810, United States

500 mM KCL

150 mM MgCl2

1% Triton X-100

- Taq DNA Polymerase (2.0 U/µl) (Finzyme, Finland)
- dNTPs (100 mM each) (Finzyme, Finland)
- Oligonucleotide primers (MWG-Biotech, Germany)[*]

Method of PCR amplification:

The volume of PCR reaction mixture is 25 µl: It consists of the following components:

- 1 µg genomic DNA
- 10x buffer
- 0.25 mM dNTPs
- 1 pmol of each primer (MWG-Biotech, Germany)
- 2.5 U of Thermus aquaticus polymerase.

Perkin Elmer thermal cycler (Applied Biosystem 2720, Singapore) [*]was used for PCR procedure.

The cycling parameters were: 94°C for 5 minutes (for initial denaturation), followed by 30 cycles of [denaturation at 94ºC for 30 seconds, annealing at suitable annealing temperature listed in **table(20)** for each primer pair for 30 seconds, and elongation at 72 ºC for 1 minute] then a final elongation step at 72 ºC for 7 min.

e) Using Gel Electrophoresis to show PCR Products

Reagents of gel electrophoresis:

- **Agarose Biotechnology grade** (BioShop, Canada)
- **10X TBE (Tris-borate EDTA buffer) pH 8.3** (Q.BIO-gene, France)

0.9 M Tris base
0.89 M Boric acid

[*]*Applied Biosystem Lincoln Centre Drive, Foster City, CA 94404 USA*

100 mM EDTA (pH 8.0)

[*] *MWG-Biotech: EurofinsMWG Operon, Anzingerstr. 7a 85560Ebersberg, Germany*

- **6X Loading dye solution** (Fermentas, Germany)*

60 mM EDTA (pH 8.0)
0.09% Bromophenol blue
0.09% Xylene cyanol FF
60% Glycerol

- **10 mg/ml Ethedium bromide** (Q.BIO-gene, France)
- **Phi-X174RF Hae III Digest (**Fermentas, Germany)
- **Gene Ruler Low Range DNA Ladder** (Fermentas, Germany)

Procedure of Gel Electrophoresis:

2% agarose gel electrophoresis was appropriate to separate the PCR products of HNF1B amplified coding sequences. To prepare the gel add 1 gram agarose to 50 ml of half X TBE buffer. Mix well then heat in microwave till the agarose dissolve. Cool to 37 ºC. Add Ethedium bromide. Mix well then pour into the gel chamber. Allow the gel to solidify at room temperature. Add enough amount of buffer (0.5X TBE) to cover the gel.

Mix 2 µl of PCR products with 2 µl of the loading dye. Pipette the mix into gel slots. Pipette 1µl of [100 bp marker which is formed of 11 fragments (1500, 1000, 900, 800, 700, 600, 500, 400, 300, 200, 100-bp)]. PCR products were allowed to electrophoresed for 0.5 hour at 100 volt. Visualization was done using UV transilluminator (TFX-20-M-Vilber Lourmat, France) *and apparatus of (Biometra Photo documentation System, Germany)** used for photographing the gel. Then we know the length of PCR products by comparing with sizes of DNA 100 bp marker.

f) DNA Sequencing:

Reagents used for direct sequencing:

- QIAquick PCR purification kit (QIAGEN, Germany)***
- BigDye® Terminator kit (Applied Biosystem, Foster City, CA, USA)****

Fermentas: FERMENTAS GMBH, Opelstrasse 9, Leon-Rot 68789 St, Germany

* Surplus2D Ltée., 813 St Jacques, St-Jean-sur-Richelieu, Qc, Canada J3B 2N2.
**Rudolf-Wissell-Straße 30, 37079 Göttingen, Germany.
***QiagenStraße 1, 28104Hilden, Germany.
****Applied Biosystems, 850 Lincoln Centre Drive, Foster City, CA 94404, USA.

***** Applied Biosystems, 850 Lincoln Centre Drive, Foster City, CA, 901100, USA.

- CENTRI-SEP purification spin columns (Applied Biosystem, Foster City, CA, USA) *****

Principle:

To know the accurate sequence of PCR product called DNA sequencing. Our study used the chain termination or dideoxy method reviewed by Sanger et al. (1977). It is used commonly uptill now. 2', 3'-dideoxynucleoside-5-triphosphates (ddNTP) work as specific termination of DNA chain elongation. It can be incorporated into the growing PCR chain as it contains 5'-triphosphate but after binding cannot form phosphodiester link with the next base as it lacks 3'- hydroxyl group. So termination of chain elongation occurs whenever an analogue tries to incorporate ***(Sanger et al., 1977).***

Before sequencing, separate DNA fragments to be in the form of single strand. Normal four (deoxy) nucleotides are added to DNA chain in the sequence mixture. Put limited amount of dideoxynucleotides of the 4 types each labeled with a different fluorescent tag. Add DNA polymerase.

PCR process will happen normally, till by chance a dideoxynucleotide is inserted instead of the normal deoxynucleotide. Percentage of normal deoxynucleotides is very high in comparison of dideoxynucleotides. So some DNA single strands can be completed before the dideoxynucleotides interfere with the process. Finally, after the incubation period has ended, all DNA fragments will be separated from longest to shortest. One nucleotide difference in length can be easily separated. Different colors will be fluoresced from different dideoxynucleotides. A laser beam is used in their illumination. There is an automatic scanner for perfect results of sequencing ***(Sanger et al., 1977).***

Steps of DNA sequencing:

1. Enzymatic PCR Cleanup using Exonuclease I and Shrimp Alkaline Phosphatase

Enzymatic PCR cleanup method offers an easy way to remove the remaining primers and dNTP left from a PCR reaction. Two enzymes are needed to complete the process: Exonuclease I (Exo I, NEB #M0293) which degrades the residual PCR primers, and Shrimp Alkaline Phosphatase (rSAP, NEB #M0371) which dephosphorylates the remaining dNTP. This method enables direct downstream applications, such as Sanger sequencing, the two enzymes are added directly to the PCR reaction after thermal cycling, without changing buffer condition or additional additives. Further, these enzymes are 100% compatible with all commonly used PCR reaction buffers.

Steps:

1- Add 0.5 µl of Exo I and 1 µl of rSAP to 5 µl of PCR product

2- Incubate the mix at 37°C for 15 minutes.

3. Inactivate both enzymes at 80°C for 15 minutes.

4. PCR products are ready for downstream application.

2. Cycle sequencing of the purified PCR products:

BigDye Terminator kit is used in cycle sequence PCR. Each ddNTPS is labeled by a different fluorescent dye. Kit contains premixed required components for doing the reaction of sequencing. The composition of buffer is suitable to sequence single-stranded PCR products in one tube. Different lengths of DNA chain are produced, each terminated by ddNTPs which is labbled by different fluorescing dye. For each position of template, the termination occurs once at least. This is done by adjusting the ratio of dNTPs/ddNTPs. Last step is running reaction products in an automated sequencer, so final sequence results can be obtained.

Component for preparation of reaction mixtures:

- Big dye (4.0 µl)
- purified PCR products (0.5-2µl) (according to concentration)
- 3pmol of primer of 10 µl /100ml conc. (forward or reverse primer)
- Final reaction volume complete by adding Sterile H2O.

Perkin Elmer thermal cycler (Applied Biosystem 2720, Singapore) is used for Cycle sequencing PCR. The procedure: 1 min at 96 ºC (initial denaturation), 25 cycle [10s at 96 ºC (denaturation), 5s at 61 ºC (annealing), 10s at 60 ºC (elongation)]

3-Removal of dye terminators (big dye) before sequencing by 3. BigDye XTerminator™ Purification Kit (Catalog number: **4376484):**

The BigDye® XTerminator™ Purification Kit is a fast, simple purification method for DNA sequencing reactions that improves the sequencing workflow and removes unincorporated BigDye terminators.

The BigDye XTerminator Purification Kit requires the addition of only two reagents, which can be added sequentially or premixed:

- XTerminator™ Solution—Scavenges unincorporated dye terminators and free salts from the post-sequencing reaction
- SAM™ Solution—enhances the performance of the XTerminator™ Solution and stabilizes the post-purification reactions

Procedure:

1. After cycle sequencing, pipette the SAM™ Solution into each tube.

2. Pipette the correct volume of XTerminator Solution into each tube.

3. Centrifuge tubes for 30 minutes at 3000 rpm.

4. Transfer 10µl of the product to the sequencing plate.

5. Place the reaction plate in ABI PRISM 310 Genetic Analyzer (Applied Biosystems). Select run module and run plate.

6. Place the tubes of sequencing reaction samples in a tray in the instrument's autosampler.

The apparatus has an electric field of anode and cathode. Sample is brought by the autosampler into cathode. Polymer must fill one end of a glass capillary (Performance Optimized Polymer, 6%"POP-6"). The other end of capillary is attached to the anode which is immersed in buffer (1X Genetic Analyzer Buffer). Sample entered the capillary from cathode to anode as current flew (electrokinetic injection). Place the capillary end at the cathode in buffer. Apply current again to continue electrophoresis.

The fluorescent dye labels are excited by the laser beam. When tagged sequences become in contact with a detector window of the capillary coating, CCD camera will collect the emitted dye fluorescence. The result interpreted by suitable software.

There were 2 software used:

1- The raw data are controlled, then monitored; finally collected by (ABI PRISM® 310 Data Collection Software) it contains standard data as reference for process to succeed.

2- ABI PRISM® DNA Sequencing Software which was used in analyzing raw sequencing data and calling bases.

g. Functional analysis through free online programs

Finch TV Version 1.4.0 is the programs which shows the data of sequencing

Free online programs:

Blast ® Basic local alignment search tool: a program from NCBI for comparison of new sequence of the patients and that present in the human genome project. (http:/blast.ncbi.nlm.nih.gov) is the home page. https://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE_TYPE=BlastSearch&BLAST_SPEC=OGP__9606__9558&LINK_LOC=blasthome

Mutations are tested by program of http://www.mutationtaster.org/

Statistical Analysis

Tests used:

- **Descriptive analysis** of the results in the form of percentage distribution for qualitative data (minimum, maximum, mean and standard deviation) calculation for quantitative data.

$$\text{-Mean} = \frac{\sum x}{n}$$

$$\text{-S.D} = \sqrt{\frac{\sum (x - \bar{x})^2}{n-1}}$$

Where: x = observation.

$\bar{x}$ = mean.

n = number.

n2 = number of group 2.

- **P:** The probability/significance value.

For all the above-mentioned statistical tests done, the threshold of significance is fixed at 5% level (p-value). P-value of more than 0.05 was considered non-significant while P-value less than 0.05 indicated a significant result. The smaller the P-value obtained the more significant were the results.

- **P value > 0.05 ⇒** (NS) Not significant.
- ***P value <0.05*** *⇒Significant at 0.05 level.
- ***P value < 0.01*** **⇒ Significant at 0.01 Level.

Data analysis:

The data will be coded to fit the program of statistical analysis (SPSS) Statistical Package for Special Sciences version 22 under windows 7.

A random sample of 10% of cases will be selected and reviewed to ensure an adequate quality of data.

Statistical tests:

- ✓ Description of qualitative variables will be by frequency and percentage.
- ✓ Description of quantitative variables will be in the form of mean and standard deviation (mean ± SD).
- ✓ Chi-square (χ^2) test will be used for comparison of qualitative variables with each other.
- ✓ Comparison between quantitative variables will be carried by using :
  - Student t-test of two independent samples.
  - One way ANOVA test (analysis of variance) was used instead of t-test
- ✓ Correlation and multivariable logistic regression analysis.

Appendix (II): Results

Table (S1): The distribution of the studied participants according to their diagnosis; (N= 20):

		Frequency	Percent
Diagnosis	**MCDK**	**5**	**25%**
	CAKUT	**15**	**75%**
	Total	**20**	**100.0**

Table (S2): Socio-Demographic data of the studied participants; (N= 20):

		Diagnosis		Total Population N= 20	p-value
		MCDK N= 5	CAKUT N= 15		
Age; (years)	**Mean ±SD**	**3.76 ±2.15**	**6.34 ±8.15**	**5.22 ±6.3**	**0.343**
	Min – Max	**1 – 7**	**0.70 – 27.0**	**0.70 – 27**	
Sex	**Male**	**4**	**11**	**15**	**0.313**
		80.0%	**73.3%**	**75%**	
	Female	**1**	**4**	**5**	
		20.0%	**26.7%**	**25%**	

Sequencing results:

Screening of the *HNF1B* gene in 20 patients from 19 unrelated families affected with CAKUT did not reveal any mutations but it revealed 8 SNPs as follow: an intronic variant (intron 5) c.1207-51C>A present in heterozygous form in the patient (1) of (family 1) (5% of cases). Another SNP was a homozygous intronic variant (intron 6) c.1339+27T>C, **rs2107133** in two patients (10%), patients (2, 3) of (family 2)**.** Moreover, a heterozygous synonymous Variant c.1413C>T, P471 (Pro)> P471 (Pro) in exon 7, rs140781855 in patient 4 of (family 3), 5% of cases. There is also an intronic variant (intron 8) c.1653+21G>A **rs375700762** present in heterozygous form in the patient (5) of (family 4), 5% of cases**.**

Other common SNPs were detected in our study. The first is an intronic variant c.1653+47_1653+48insC (intron8) **rs35913775.** It was detected in homozygous form in three patients (15%), patient (2, 3) of (family 2), and patient (6) family (5), and detected in heterozygous form in five patients (25 %); patient7, patient 8, patient 9, patient10, patient 11, of (family 6-10).

Moreover, c.1654-22C>T intronic variant (intron 8) **rs3110641** was detected in homozygous form in four patients (22.2%); patient2, patient3, patient7, patient9 (families 2,6. 8) and was detected in heterozygous form in 10 patients (50%); (patients 1,5,6, 8, 12-17) (families 1,4,5, 7, 11-16)**.**

c.*99C>A, **rs2229295** detected in homozygous state in three patients while in heterozygous state in six patients**.**

c.*100A>G, **rs1800929** detected in homozygous state in two patients while in heterozygous state in 9 patients**.** The most frequent polymorphisms were rs3110641 (70%), rs1800929 (55%), rs2229295 (50%), and rs35913775 (40%).