Growth Hormone Receptor and PCSK9 Define a New Paradigm in the Initiation and Development of Chronic Kidney Disease as Revealed by Exome Sequencing on Illumina Platform

Edem Nuglozeh1,2,*, Mohammad Feroze Fazaludeen1, Sumukh Deshpande3, Aymen Ahmed Warille5, and Mustafa Kürşat Şahin6

ABSTRACT

Chronic kidney disease is a combination of many vascular diseases involving mutations of many genes. Hypertension diabetes and atherosclerosis are the most common causes of kidney disease, with hypertension causing just over a quarter of all cases of kidney failure and diabetes causing one-third of them. Other much less common conditions that can cause CKD include inflammation, infections, genetic factors, or longstanding blockage to the urinary system (such as enlarged prostate or kidney stones). In many cases, the causes remained unknown, albeit the manifestation of the diseases with clear phenotypes and biochemical profiles. Heredity and genetic determinants play major roles in the initiation, development, and establishment of CKD. Kidney disease phenotypes can be dissected into many underlying causing candidates' genes and many molecular genetics approaches are striving to lift the veil on this nagging disease. Recent studies using genetic testing have demonstrated that Mendelian etiologies account for approximately 20% of cases of kidney disease of unknown etiology. CKD is known to be plagued with many genes mutations like mutation in Autosomal Dominant Polycystic Kidney Disease (ADPKD) and mutations in MYH9 and APOL1 genes, COL4A3, COL4A4, and COL4A5 genes playing important roles in the CKD picture. Genetic testing has modernized and revolutionized many areas of medical practices and diagnosis of many diseases and the field of nephrology is not an exception. The advance in Next-generation Sequencing, including whole exome sequencing has proven to be a powerful tool in personalized medicine and for potential noninvasive decryption for biomarkers in kidney disease thereby paving the way for better diagnostic purposes. In this regard, we run whole exome sequencing on whole blood genomic DNA from CKD patients. Bioinformatics analysis led us to uncover a total of more than 3000 single nucleotide polymorphisms (SNPs). To sort out these flurries of targeted SNPs, we undertook filtration using an R-algorithm in combination with the diseases association Clinvar database. This approach led us to 12 combined diagnostic missense variants scattered on different chromosomes. Combined missense reduction after FDR filtration with a Cellrate of 0.75 generated two missense variants located on PCSK9 and GHR genes on chromosomes 1 and 5 and lastly, reduction variants after Filtration by spliced region bring us to a single SNP located on the PCSK9 gene.

Keywords: CKD, GHR, PCSK9, Whole exome sequencing.
1. Introduction

The kidney is a complex organ consisting of millions of filter structures and a myriad of vasculatures that fulfill many functions and are thus prone to a wide variety of abnormalities from the tubules involved with sodium and other ions drainage to atypical growth development, all culminating to structural defects. A typical kidney can be divided into functional compartments such as vasculatures, interstitium, and nephron, and each of these elements is affected by genetic variations. At a smaller scale, the nephron can be subdivided into tubules and glomeruli with subsequent subdivision into cell-type structures with highly specialized cell-type populations susceptible to structural and functional anomalies related to genetic variation. The number of genes that have been involved in human kidney diseases or affecting kidney function is diverse. Not all kidney diseases are diseases of the kidney per se, as other diseases affecting human beings can evolve into kidney diseases thereby invading the kidney. For example, Spontaneous hemolytic uremic syndrome can be triggered by mutation in complement factor [1], and mutations in the protease encoded by ADAMTS13 can cause thrombotic thrombocytopenic purpura [2]. Reciprocally, systemic diseases, such as hypertension and perturbation in renal salt handling may be driven by genes that primarily affect kidney function [3]. Common variation in immune regulatory genes, HLA class II molecules seems to affect the risk of various glomerular phenotypes [4]. Diabetes, the most common determinant of CKD, is the culmination in most parts of the coalescence of variation in non-kidney genes and the debate is still out there as to whether susceptibility to nephropathy in the development of diabetes is primarily triggered by genes directly affecting the kidney [5], [6]. For example, some people with mumps get severe diabetes [7], [8] and the mumps virus easily disseminates to the kidney culminating in kidney diseases. So, from one thing leading to another one, it is a clear demonstration that the development of mumps infection can lead to chronic kidney disease and differences in the kidney’s susceptibility to systemic disease originating outside the kidney may reflect genetically determined differences in the kidney’s response to such diseases. CKD has a heritable element and a global threat to public health with high morbidity and mortality [9], [10]. Although genome-wide association studies (GWAS) have delineated myriad common genetic variants influencing all causes of CKD and kidney function, these association studies have only explained a small fraction of the variance in these traits [11]–[13]. Up to date, no studies have systematically examined the contribution of rare genetic variation (<1% minor allele frequency [MAF]) with a presumed larger effect on the risk of CKD. Detection of rare independent variants clustered in a single gene can provide significant clues into the disease biology and improve clinical therapy in many aspects, even if pathogenic variants in a particular gene only explain a small proportion of all cases. First, in a population with both acquired and inherited multifactorial disease, an accurate estimate of the proportion of cases caused by known genes can inform the use of existing therapies and diagnostic tests. Second, rare mutations can lead to the identification of widely applicable therapeutic drug targets, such as the discovery of PCSK9 mutations leading to the development of treatment for general forms of hypercholesterolemia [15], [16]. Third, it is increasingly recognized that the validation of such drug targets in genetic studies of human populations improves the probability of the success of drug development in clinical trials.

In our study dealing with exome sequencing of patients suffering from chronic kidney disease, Chromosomes HeatMap analysis data led us to initially report that some specific genes located on specific chromosomes [17] co-segregate with CKD. Subsequently, we also conducted Codons HeatMap and derived Point of Accepted Mutation (PAM) matrices for all amino acids stemming from exome sequencing data using various bioinformatics tools and identified specific amino acid exchanges associated with this disease herein. These data will be reported separately. Here, in this part of this study, we specifically put the focus on sorting out the myriad SNPs we uncovered from this study and run the disease association algorithm for more than 3000 SNPs resulting from bioinformatics analysis of raw data using the GATK platform. After filtration and hard filter application, we end up with two diagnostic SNPs involved in the symptomatology and pathophysiology of CKD from these patients. We later directed our efforts to establish a pattern or a clinical diagnostic profile for making use of the Illumina platform as a clinical diagnostic and investigative tool. Our results
clearly established a clear pattern of nucleotide substitution specifically associated with chronic kidney diseases (data not shown).

2. Materials and Procedure

   Genomic DNA specimens are purified from samples as published elsewhere [17].

   2.1. Sample Population

   This study is part of a subgroup of an investigation started a while ago on hemodialysis patients from outpatient clinics of King Khaled Hospital and published with ethical approval issued to Dr. Alaraj [17], [18].

   2.2. Libraries Preparation

   We generated human genomic DNA (gDNA) robust libraries as reported elsewhere [19]. Brief, high-quality gDNA was purified from whole blood using Qiaegen kits (QIAamp DNA Blood Mini Kit from QIAGEN, Hilden Germany). Libraries were built using the Illumina Kits system based on Transposase enzymology. We later proceeded to QC control of the libraries using Bioanalyzer (Agilent Technology, Santa Clara, CA, USA). The libraries were subsequently sequenced on Illumina MiSeq Platform using MiSeq Reagent Kit optimized v3 Chemistry, 150-cycles to increase cluster density and read length as well as improve quality (Q) scores. We run the sequencing forward and reverse.

3. Bioinformatics Methodology

   Fig. 1 presented a statistical distribution of mutagenic variants obtained in Bayesian allele caller which produces scores and allele positions in the genome. In Figs. 2 and 3, we described the bioinformatics pipeline used to analyze FASTQ data obtained from CKD patients, and this generated more than 300 SNPs per patient (data not shown).
Growth Hormone Receptor and PCSK9 Define a New Paradigm in the Initiation

Naglozeh et al.

Variant Calling File (VCF) obtained from TVC, Freebayes, and GATK

Remove variants (MAF < 0.05) from 1000 Genomes database

Remove variants (MAF < 0.05) from ESP 65000 database

Remove variants (MAF < 0.05) from dbSNPv138 database

Remove variants (MAF < 0.05) from Complete Genomics 46 database

Variants Annotation

Filter variants associated with disease

Remove synonymous variants

Fig. 3. Variant reduction pipeline performed using ANNOVAR tool.

Fig. 4. Variants reduction filtering flowchart using Clinvar diseases association database illustrating the report of diagnostic clinical variants observed from CKD patients.

Brief,
- FASTQC files were concatenated from multiple runs of the libraries.
- Adapter trimming and base quality scores (less than Q30) were removed using Cut adapt (V1.7.1).
- FASTQC (V0.11.2) was used to check primary and post-trimmed sequences.
- Alignments to the reference human genome (hg19) were conducted using BWA (version 0.7.15).
- The Genome Analysis Tool Kit, GATK (version 3.0.0), was used for base quality score recalibration, and variant calling followed by hard filtering to identify high-quality variants for downstream analyses.
- SnpEffv4.1 was exploited to determine in silico impacts upon the protein function of candidate genes. The Fast QC files analyses generated more than 300 SNPs per patient and these SNPs were subsequently later committed to reduction variants analysis that we named disease association filtration.

4. Data Filtration and Diseases Association

The combined SNPs obtained from the analysed FASTQ data were filtrated following the logic on the diagram depicted in Fig. 4. Brief, data from samples of the CKD patients were used to filter the variants based on the disease association mutation database. The annotated variants data were filtered based on five keywords: “hypertension”, “kidney diseases”, “heart”, “diabetes”, “hypercholesterolemia”, and “diabetes”. The filtered variants were used to annotate the variants and to further analyze the association with disease.
and “hypercholesterolemia”. The automated variant filtration was applied using R base language to narrow down the number of candidate diagnostics SNVs using the following rules as illustrated in Fig. 4. The combined SNPs were later submitted to hard filtration by merging them with Clinvar database leading us to generate the diseases association SNPs database. The disease keywords mentioned above were used to filter the variants for each sample on the Clinvar annotation column. The variants matching the filter criteria showed a total number of 12 SNPs, all associated with cardiovascular diseases, renal diseases, diabetes, and thrombosis. Those variants matching with filter criteria for each sample were then combined to produce a final list of missense variants selected for all samples in Table 1.

5. Results Presentation

In Fig. 1, we presented the distribution of mutagenic signature effects on genomic features as the total number of variants per effect type. This figure projects an atlas of somatic mutagenesis in CKD patients and identifies the disease-prone genomic regions, with most variants concentrated in intervening sequences and downstream regions. The mechanisms underlying those mutagens that shape this distribution related to CKD remain unknown but merit further in-depth investigations to establish a clear picture between genomic fragment distribution and disease. In Figs. 2 and 3, we presented different algorithms and pipelines used in the analysis of FASTQ data analysis with the complete description of different algorithms. These pipelines call the variants based on the aligned reads that the sequencer generated and variant calling in this ongoing process reviews the sequence alignment, typically in the form of a BAM file, to identify that, the target loci in question differ from the reference genome. The most common type of variation will come in the forms of SNPs or SNVs followed by insertions and deletions and these variant calls are stored in the VCF file and the total combined SNPs obtained and stored in the VCF file is roughly 3600. In Fig. 4, we presented the results of our data filtration using the diseases association database.

Table I presents the resultant of combined missense reduction variants following hard Filtration following. The 12 variants we retained in this sorting process are those that fit with the criteria of filtration using Clinvar database. All the remaining variants here are symptomatic cardiovascular diseases, renal diseases, diabetes, and thrombosis. Table II harbors two variants resulting from missense reduction variants after FDR Filtration and Table III is the results of combined missense reduction variants after Filtration by spliced region. All these tables, vis, Tables I–III represent a hierarchy of filtration, starting from FASTQC sequencing data representing more than

<table>
<thead>
<tr>
<th>Chr</th>
<th>Start</th>
<th>End</th>
<th>Ref</th>
<th>Alt</th>
<th>Type</th>
<th>ID</th>
<th>Genes names</th>
<th>Functions</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chr1</td>
<td>55529186</td>
<td>55529187</td>
<td>G</td>
<td>A</td>
<td>rs505151</td>
<td>PCSK9</td>
<td>Chol metabolism</td>
<td>[22]</td>
<td></td>
</tr>
<tr>
<td>Chr2</td>
<td>21233971</td>
<td>21233972</td>
<td>T</td>
<td>C</td>
<td>rs533617</td>
<td>APOB</td>
<td>Chol, thrombosis vascular biol</td>
<td>[29]</td>
<td></td>
</tr>
<tr>
<td>Chr2</td>
<td>21224852</td>
<td>21224853</td>
<td>C</td>
<td>T</td>
<td>rs1801695</td>
<td>APOB</td>
<td>Chol, thrombosis vascular biol</td>
<td>[30]</td>
<td></td>
</tr>
<tr>
<td>Chr3</td>
<td>38645419</td>
<td>38645420</td>
<td>T</td>
<td>C</td>
<td>rs1805124</td>
<td>SCN5A</td>
<td>Heart diseases</td>
<td>[31]</td>
<td></td>
</tr>
<tr>
<td>Chr4</td>
<td>42719238</td>
<td>42719239</td>
<td>A</td>
<td>C</td>
<td>rs6180</td>
<td>GHR</td>
<td>Reg lipoproteins metabolism</td>
<td>[37]</td>
<td></td>
</tr>
<tr>
<td>Chr6</td>
<td>26091178</td>
<td>26091179</td>
<td>C</td>
<td>G</td>
<td>rs1799945</td>
<td>HFE</td>
<td>Homochromatosis and kidney disease</td>
<td>[41], [42]</td>
<td></td>
</tr>
<tr>
<td>Chr7</td>
<td>94946083</td>
<td>94946084</td>
<td>A</td>
<td>T</td>
<td>rs854560</td>
<td>PONI</td>
<td>Renal functions and diabetes</td>
<td>[43], [44]</td>
<td></td>
</tr>
<tr>
<td>Chr7</td>
<td>1.51E+08</td>
<td>1.51E+08</td>
<td>T</td>
<td>G</td>
<td>rs1799983</td>
<td>NOS3</td>
<td>Coronary disease atherosclerosis</td>
<td>[47]</td>
<td></td>
</tr>
<tr>
<td>Chr8</td>
<td>11405575</td>
<td>11405576</td>
<td>G</td>
<td>A</td>
<td>rs5575876</td>
<td>BLK</td>
<td>β-cell dysfunction and diabetes</td>
<td>[49]</td>
<td></td>
</tr>
<tr>
<td>Chr9</td>
<td>1.08E+08</td>
<td>1.08E+08</td>
<td>T</td>
<td>C</td>
<td>rs2230806</td>
<td>ABCA1</td>
<td>Coronary artery diseases</td>
<td>[51], [52]</td>
<td></td>
</tr>
<tr>
<td>Chr12</td>
<td>1.21E+08</td>
<td>1.21E+08</td>
<td>A</td>
<td>G</td>
<td>rs1169305</td>
<td>HNF1A</td>
<td>Type 2 diabetes</td>
<td>[55]</td>
<td></td>
</tr>
<tr>
<td>Chr20</td>
<td>43058266</td>
<td>43058267</td>
<td>A</td>
<td>G</td>
<td>rs14763845</td>
<td>HNF4A</td>
<td>Diabetes mellitus</td>
<td>[56]</td>
<td></td>
</tr>
</tbody>
</table>
3600 variants and culminating in Table III with a single variant on PCSK9 located on chromosome 1.

6. DISCUSSION

Chronic kidney disease (CKD) is a worldwide global threat with more than 850 million individuals affected [20], [21]. Because of the magnitude of the population affected by the disease worldwide, a global search for an array of genes involved with the disease will be of great help toward the design of sound treatments and therapies. Understanding the molecular mechanisms and the genetic variations associated with CKD becomes imperative and represents an important step toward new drug development. To make a step toward this goal, we run exome sequencing from CKD patients’ genomic DNA (gDNA) on the Illumina platform. After bioinformatics analysis of FASTQ data, the INDELs and SNPs stemming from this study are scattered on many chromosomes (data not shown). Most of them are linked to cardiovascular diseases, hypertension lipids metabolism, and diabetes. Using the Heatmapper algorithm (http://www.heatmapper.ca), we characterized the chromosomal positions for some critical genes that we presented under the format of Chromosomes HeatMap in CKD patients that we already published [17]. We also derive amino acids exchange matrices from CKD patients against those obtained from normal human databases 1000 kg and gnomAD and we will be reporting this separately. From the results of the bioinformatics pipeline we used, we characterized more than 3600 SNP variants from CKD patients as well as the genomic positions of these variants. We sorted out these variants in regard to various diseases associated with kidney diseases using Clinvar database. This filtering generated 12 SNPs scattered on chromosomes: 1, 2, 3, 5, 6, 7, 8, 9, 12, and 20.

From chromosome 1, we identified rs505151 SNP characterized by the mutation of G→A PCSK9. This mutation is a long-range haplotype signature of nonsynonymous allele SNP rs505151 (E670G) characterized by a gain of function on PCSK9 and this gain of function provokes an elevation of LDL cholesterol in familial hypercholesterolemia [22]. We have to recall that, PCSK9, (Prohormone convertase subtilisin kexin 9) is a plasma membrane glycoprotein whose major function is the metabolism the LDL-cholesterol. The enzyme is characterized by multiple alleles divided into gain-of-functions and others with loss-of-functions alleles. The gain-of-function mutations of PCSK9 reduce LDL receptor levels in the liver, resulting in high levels of LDL cholesterol in the plasma whereas, Loss-of-function mutations lead to higher levels of the LDL receptor, thereby lowering LDL cholesterol levels and protection from coronary heart disease [23]. Indeed, human kidneys synthesized PCSK9 [24] and the number of LDLRs available on the surface of renal cells depends on the amounts of PCSK9 [25]. High levels of PCSK9 can provoke excessive accumulation of lipids in kidneys leading to renal fibrosis and antibodies against PCSK9 used in these pathological abolish and reverse the disease state [26]. The corroborating fact is the discovery that, hypercholesterolemia has been discovered in patients with nephrotic syndrome in association with high levels of PCSK9 [27], [28] and this corroborates with our current observations, that after filtering the 3600 SNPs obtained from exome sequencing in CKD patients genomic DNA, missense variant SNP located on PCSK9 (G > A, p. Gly670Glu) characterized by gain of function emerged as one of causative agent leading to the establishment and development of chronic kidney disease.

Continuing with our chromosomes walking, on chromosome 2, with our filtering methodology, we retained two variants’ SNPs (rs533617 and rs1801695) located on the APOB gene and those variants also associated with cholesterol metabolism reinforcing the assertion we have drawn from PCSK9 biology. The SNP (rs533617 T > C) is a missense mutation and its gene product sequence is characterized by mutation of Hist → Arg at the position 1923. The SNP rs533617 (p. His1923Arg) was revealed as a potential causal variant in lipids metabolism and is present in low frequency [29]. The second variant (rs1801695 C > T), a nonsynonymous variant still on the APOB gene was found to be most significantly SNP associated With High-Density Lipoprotein Cholesterol metabolism [30]. Together with the precedent information about cholesterol metabolism from PCSK9, this led us to evidence that impairment in lipids metabolism constitutes one of the major factors leading to chronic kidney diseases.

On chromosome 3, we characterized the variant SNP (rs1805124 T > C) which is a missense variant located on the gene SCN5A and its protein sequence is characterized by a mutation of Hist → Arg at position 558. This is a disease-causing mutation in familial dilated cardiomyopathy [31]. Indeed, this variant was earlier reported as one of the haplotypes affecting the members of Brugada family [32]. The His558Arg missense variant alters a conserved residue in the sodium channel inter-domain and has been reported as a modulator of arrhythmia-causing SCN5A variants. Moreover, transient Brugada-type electrocardiographic abnormalities in renal failure were reversed by dialysis liking the functionality of sodium/potassium channels in this disease [33]. Hyperkalemia is one of the renal impairments in chronic kidney disease (CKD) and hyperkalemia is one of the dangerous complications of renal impairment (CKD) and is directly linked to the missense Hist558Arg missense variant [34].

On chromosome 5, we reported the missense variant (rs61810 c.1630 A > C) on Growth Hormone Receptor (GHR). This variant is a nonsynonymous single-nucleotide polymorphism (SNP) rs6180 (p. Ile544Leu) in exon 3 of the GHR gene. Initially, some nonsynonymous variations in the GHR gene were identified in the cytoplasmic domain of the GHR protein [35]. This consisted of three coding SNPs reported as: (rs6182: p. Cys440Phe, rs6180: p. Ile544Leu, rs6184: p. Pro579Thr) on GHR. Among them, only rs6180 (p. Ile544Leu; c.1630 A>C) in exon 3 remains highly polymorphic [36]. The rs6180 variant has been reported in Clinvar database and is associated with familial hypercholesterolemia and Laron-type isolated somatotropin defect. The growth hormone receptor (GHR) belongs to the cytokine receptor superfamily and mediates the majority of growth hormone effects. Growth hormone (GH) and its mediator insulin-like growth factor-1 (IGF-1) have important effects on the
kidneys, including glomerular and tubular function, as well as the synthesis of 1,25 (OH)₂ vitamin D₃. Observation in patients with acromegaly demonstrated that GH excess can impact kidney health, including glomerular hyperfiltration, hypertrophy, and glomerulosclerosis [37], [38]. In poorly controlled type 1 diabetes patients, elevation in GH was shown to produce injury in podocytes thereby inducing diabetic nephropathy. GH accumulation is due to GHR dysfunction. GHR mediates the autoregulation of GH and the disruption of GHR interferes with post-translational processing, maturation, ligand binding, and signal transduction. All these processes lead to the accumulation of GH that impacts the kidneys’ concomitant development of nephropathic diabetes which culminates in CKD. Indeed, the mutation in GHR codon 49 via histidine-to-leucine substitution significantly impairs glycosylation-mediated receptor processing, maturation, ligand binding, and signaling causing GH-excess secretion [39]. Although not yet demonstrated with our current variant of (p. Ile544Leu; c.1630 A > C), it is possible that this mutation may induce other forms of nephropathic diabetes leading to CKD.

On chromosome 6, we reported, the missense variant (rs1799945 c.187 C > G) with its protein features, (p. His63Asp) on the HFE gene. This variant was reported on Clinvar with more than 100 citations and the most common genotypes are associated with iron metabolism in hemochromatosis and type 2 diabetes. It has been known for a long time that, between 50% and 80% of patients with hemochromatosis have type 2 diabetes [40], [41]. Indeed, more recent studies have confirmed the former observation by reporting that missense mutation H63D in the HFE gene confers risk for the development of type 2 diabetes mellitus [42].

On chromosome 7, we reported two missense mutations: (rs854560 A > T) with protein features: (p. Leu55M) located PON1 gene and (rs1799983) on chromosome 7 bearing the characteristic of c.894T > G or (G894T) with protein features (p. Asp298Glu) or (E298D) [Glu298Asp] on NOS3 gene. The variant characterized by the mutation of A > G; c.1813A > T with protein features (p. Ser605Gly) on the HNF1A gene [38]. In poorly controlled type 1 diabetes patients, elevation in GH was shown to produce injury in podocytes thereby inducing diabetic nephropathy. GH accumulation is due to GHR dysfunction. GHR mediates the autoregulation of GH and the disruption of GHR interferes with post-translational processing, maturation, ligand binding, and signaling. All these processes lead to the accumulation of GH that impacts the kidneys’ concomitant development of nephropathic diabetes which culminates in CKD. Indeed, the mutation in GHR codon 49 via histidine-to-leucine substitution significantly impairs glycosylation-mediated receptor processing, maturation, ligand binding, and signaling causing GH-excess secretion [39]. Although not yet demonstrated with our current variant of (p. Ile544Leu; c.1630 A > C), it is possible that this mutation may induce other forms of nephropathic diabetes leading to CKD.

On chromosome 8, we reported the missense mutation rs55758736 characterized by c.211G > A with protein features: (p. Ala71Thr) on the gene BLK reported on Clinvar registry. The gene product is a nonreceptor tyrosine-kinase of the src family of proto-oncogenes involved in cell differentiation and proliferation. Maturity-onset in diabetes of the young (MODY) is linked to β-lymphocyte kinase (BLK) gene locus mutation and few mutations were uncovered from this gene locus, among which, Ala71Thr substitution. BLK gene product expression enhances insulin synthesis and secretion in response to glucose and these actions are substantially abrogated by the Ala71Thr variant [49] and all these metabolic dysfunctions are symptomatic of chronic kidney disease. Moreover, reduction in Blk expression Enhances proinflammatory cytokine production and induces nephrosis in association with β-cell dysfunction [50].

On chromosome 9, we reported a rs2230806 missense variant characterized by C>T substitution. The rs2230806 variant, also known as Arg219Lys or R219K, is a SNP in the ABCA1 ATP-binding cassette 1, sub-family A (ABC1), member 1 gene. The rs2230806(G) allele encodes the arginine (R), and the (A) allele encodes the lysine (K). A meta-analysis published in 2011 comprising 22 studies with 6597 cases and 15,369 controls studied the association between the rs2230806 variant and risk for coronary artery disease. ABCA1, ATP-binding cassette 1 is associated with coronary artery disease (CAD) and atherosclerosis (AS) which is basically considered as the pathological basis of CAD. Studies have shown that the accumulation of cholesterol in the arterial wall triggers AS, which leads to an imbalance between the lipoprotein influx and the cholesterol efflux. The cholesterol efflux pathways are found mainly in the ABCA1 pathway. Macrophages and ABCA1 mediate reverse cholesterol transport (RCT) accounting for approximately 90% of cholesterol excretions [51], [52]. ABCA1, a conserved transmembrane-spanning protein, plays a crucial role in the efflux of cellular cholesterol. In humans, ABCA1 mutations can cause a severe HDL-deficiency syndrome characterized by cholesterol deposition in tissue macrophages. Disruption of Abca1 in mice promotes accumulation of excessive cholesterol in macrophages, and physiological manipulation of ABCA1 expression with concomitant AS and transplantation of bone marrow from Abca1-/- mice into Ldlr-/- or apoE-/- recipients caused an increase in AS [53], [54]. From the above clinical and scientific studies, it becomes clear that ABCA1 is an agent that contributes to AS development and will provoke CKD development. Besides, from our initial analysis form the variant p. Asp298Glu on the NOS3 gene, we successfully demonstrated atherosclerosis remains one of the major causal agents of chronic kidney disease through fibrosis and vascular stasis. Then, the rs2230806 SNP variant, also known as Arg219Lys or R219K in the ABCA1 ATP-binding cassette would contribute to the development of CKD.

On chromosome 12, we reported the rs1169305 missense variant characterized by the mutation of A > G; c.1813A > G with protein features: (p. Ser605Gly) on the HNF1A gene [53]. In this locus, three alleles of this variant have been reported on dbSNP site and none of them have been
reported as pathogenic in any diseases contributing to the development of CKD.

On chromosome 20 we reported (rs147638455, A > G) with protein features: (p.Ile463Val) on the HNF4A gene [56]. This missense variant was reported on Clinvar database. It is a coding sequence variant and its significance in the maturity Type 1, Type 2 diabetes mellitus and familial hyperinsulinism remained uncertain.

7. Conclusion

In summary, we have run Exome sequencing from CKD patient samples genomic g(DNA) on the Illumina Platform. After bioinformatics analysis followed by filtration and hard filtration against Clinvar of the presumptive diagnostic SNPs, we identified causal damageable genes involved in the physiopathology of chronic kidney disease. Of notice is the identification of 12 SNPs on chromosomes: 1, 2, 3, 5, 6, 7, 8, 9, 12, and 20 spanning all the cocktail of diseases contributing to the development of CKD and finally, combined missense reduction after FDR filtration with Cellrate of 0.75 generated two missense variants: PCSK9 and GHR on chromosomes 1 and 5 and reduction variants after Filtration by spliced region brings us to a single SNP, located on PCSK9. More studies need to be conducted using genetically modified animals to ascertain the authenticity of these variants in the initiation and development of CKD. Nonetheless, after having channeled our 3600 SNPs we extracted from the FASTQ file through different bioinformatics algorithms followed by filtrations and hard filtrations, PCSK9 and GHR emerged, and major candidates were involved in the initiation and development of CKD.

Acknowledgment

Authors are grateful to the Center of Molecular Diagnostic and Personalized Medicine, University of Hail where we performed Exome sequencing, and to the Institute of Bioinformatics, McGill University, Montreal Canada for their expertise in sequencing data analysis on the GATK platform.

Author Contributions

Eden Nuglozeh fulfilled the role of Conceptualization; Data curation; Exome sequencing; Formal analysis; Funding acquisition; Investigation; Methodology; Project administration; Resources; Supervision; Roles/Writing—original draft; Writing—review & editing. Feroz M. Fazaludeen fulfilled data acquisition, Exome sequencing, graphic plotting, and manuscript edition. Sumukh Deshpande performed some part of bioinformatics analysis dealing with SNPs filtration and annotation as well as diseases association database with the variants spreadsheet. Aymen A. Warille fulfilled the role of data mining web database search. Mustafa Kursat Sahin provided help gene search database and data analysis.

Funding

Authors are grateful to The University of Hail as well as the Hail Kidney Foundation for the support and realization of this project.

Conflicts of Interest

Authors declare that they do not have any conflict of interest.

References

Nuglozeh et al.

Growth Hormone Receptor and PCSK9 Define a New Paradigm in the Initiation of


