NGS Podocytopathy Panel Screening in Adults with CKD: Findings from the GCKD Study

CME Course “Diagnosis and management of inherited kidney diseases: What's New?”

ERA-EDTA Congress 2017
June 3rd, 2017

Anna Köttgen, MD MPH; anna.koettgen@uniklinik-freiburg.de
Disclosures

None
Background

Introduction

- Kidney function and disease have a genetic component

- Two “extremes” of human genetics:
  - Monogenic: often severe manifestation, childhood (recessive) or adulthood (dominant genes)
  - Complex (polygenic): chronic kidney disease (CKD) in adults, etiology often unknown

- But: genetic predisposition of a person really combination of common susceptibility alleles (GWAS) and rare mutations (NGS)
Knowledge Gaps

Introduction

1. Prevalence of rare pathogenic mutations in “monogenic” kidney disease genes in adult patients with CKD

2. Prevalence and combinations of rare hypomorphic alleles in “monogenic” kidney disease genes in adult CKD patients

3. Range of presentation of adult CKD patients with mutations in “monogenic” kidney disease genes

4. Effect of combinations of rare mutations and common susceptibility variants on phenotype
Project Aims

Introduction

• Address knowledge gaps for glomerular diseases through NGS panel exome sequencing of 37 known or suspected podocytopathy genes among adult patients with CKD attributed to primary glomerular disease
Patient Selection
Subjects and Methods

GCKD Study: 5,217 CKD patients aged 18-74 years with eGFR 30-60 ml/min/1.73m² or UACR >300 mg/g, nephrological care

Of these, selected those

• With biopsy

• Without secondary causes of FSGS
diabetes, hepatitis B and C, HIV, SLE, Wegener’s, sclerodermia, Tbc, GPA, TTP/ aHUS, amyloidosis, gout, sarcoidosis, other systemic diseases

• With leading cause of CKD: primary glomerular disease
membranoprolif. GN, RPGN (pauci-immune), RPGN (anti-GBM), post-infect. GN, FSGS, membranous GN, minimal change disease, or other primary GN

→ 345 individuals, 341 with GWAS data
Panel Design (EURenOmics)

Subjects and Methods

PREPARATION

MASTR assay

SCREENING

DESIGN

- Version 1-2 training sets
- Version 3 31 genes coverage 99.7% >30x
- Version 4 31 genes (+2, -2) coverage 99.3% >30x
- Version 5 37 genes (+8, -2) coverage 98.1% >30x

ANALYSIS

VALIDATION

- 68 patients with known mutations
- sensitivity 98.5%

Slide courtesy of B. Lipska, EURenOmics Consortium
Panel Content & Sequencing

Subjects and Methods

**Known genes** (n=29)

**AD genes:**
ACTN4, ANLN, INF2, TRPC6, ARHGAP24

**AR genes:**
NPHS1, NPHS2, PLCE1, MYO1E, DGKE, CD2AP, PTPRO, CRB2, EMP2

**Syndromic genes:**
LAMB2, LMX1B, GLA, SMARCAL1, WT1, PAX2, WDR73

**Alport-associated genes:**
COL4A3, COL4A4, COL4A5

**Mitochondropathies:**
COQ2, COQ6, PDSS2, MT-TL1, ADCK4

**Candidate genes** (n=8)

APOL1, C14orf142, MAGI2, MYH9, CD151, TTC21B, SCARB2, ARHGDIA

- Panel in use in 4C, PodoNet, EURenOmics
- Covers coding exons (7-8 problematic), 720 amplicons, mean coverage 98%
- Illumina MiSeq reagent kit v3, sample prep F. Schaefer lab
- Illumina MiSeq, Dept. of Human Genetics, DKFZ Heidelberg
- Raw data (.fastq files)
Data Analysis

Subjects and Methods

PRE-PROCESSING

Raw Reads

Map to Reference

BWA mem

Mark Duplicates

Picard

Base Recalibration

Analysis-Ready Reads

Non-GATK

VARIANT DISCOVERY

Analysis-Ready Reads

Var. Calling

HC-ERC GVCF

Genotype Likelihoods

Joint Genotyping

Raw Variants

SNPs

Indels

Filter Variants

Analysis-Ready Variants

SNPs

Indels

CALLSET REFINEMENT

Analysis-Ready Variants

SNPs & Indels

Refine Genotypes

Annotate Variants

Evaluate Callset

look good?

troubleshoot

use in project

Best Practices for Germline SNPs and Indels in Whole Genomes and Exomes - June 2016

Data Analysis

Subjects and Methods

Consensus Calling: GATK, SamTools, FreeBayes, Platypus

Best Practices for Germline SNPs and Indels in Whole Genomes and Exomes - June 2016
Variant Filtering and Annotation

Subjects and Methods

Annotate every variant in VCF with information from (subset):
- dbSNP
- UCSC
- ClinVar
- OMIM
- KEGG
- Pfam
- Ensembl
- ESP
- 1000G
- CADD
- Polyphen
- SIFT
- ENCODE
- HPRD
- COSMIC
- GERP
- FitCons
- VISTA

Variants, annotations, phenotypes & genotypes together in a database

Prioritize genetic variants in various disease contexts based on genome annotations, sample genotypes, and sample relationships.

Variant Filtering and Annotation

Subjects and Methods

Annotate every variant in VCF with information from (subset):

- dbSNP
- UCSC
- ClinVar
- OMIM
- KEGG
- Pfam
- Ensembl
- ESP
- 1000G
- CADD
- Polyphen
- SIFT
- ENCODE
- HPRD
- COSMIC
- GERP
- FitCons
- VISTA

Variants, annotations, phenotypes & genotypes together in a database

Prioritize genetic variants in various disease contexts based on genome annotations, sample genotypes, and sample relationships.

Variant already known?
Variant frequency?
Clinical significance?
Variant Filtering and Annotation

Subjects and Methods

Bioinformatic ("in silico") impact prediction

Annotate every variant in VCF with information from (subset):
- dbSNP
- UCSC
- ClinVar
- OMIM
- KEGG
- Pfam
- Ensembl
- ESP
- 1000G

gemini

Variants, annotations, phenotypes & genotypes together in a database

Prioritize genetic variants in various disease contexts based on genome annotations, sample genotypes, and sample relationships.
Variant Filtering and Annotation

Subjects and Methods

Annotate every variant in VCF with information from (subset):

- dbSNP
- UCSC
- ClinVar
- OMIM
- KEGG
- Pfam
- Ensembl
- ESP
- 1000G
- CADD
- Polyphen
- SIFT

Variants, annotations, phenotypes & genotypes together in a database

Prioritize genetic variants in various disease contexts based on genome annotations, sample genotypes, and sample relationships.

Known regulatory variant? Evolutionary conservation?
Results: Project Status

- As of May 30\textsuperscript{th}, 2017: data for 288/341 patients (84%)

- Found 1584 (unfiltered) variants
  - 1377 SNPs, 207 InDels
  - 946 ts, 431 tv, 80 ins, 127 del
  - impact severity: 26 high, 354 medium, 1214 low
  - in dbSNP: 1310/1584
  - in ExAC: 930/1584
  - allele frequency (any source) < 1%: 858/1584, < 0.1%: 533/1584
### Variant Filtering

#### Results

<table>
<thead>
<tr>
<th>allele frequency</th>
<th>impact</th>
<th>clinvar</th>
<th>patient genotype</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- **Allele frequency**
- **Impact**
- **Clinvar**
- **Patient Genotype**

<table>
<thead>
<tr>
<th>Gene</th>
<th>Chromosome</th>
<th>Position</th>
<th>Allele Frequency</th>
<th>Impact</th>
<th>ClinVar</th>
<th>Patient Genotype</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Gene</th>
<th>Chromosome</th>
<th>Position</th>
<th>Allele Frequency</th>
<th>Impact</th>
<th>ClinVar</th>
<th>Patient Genotype</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

---

**Table Columns**
- **Gene**: Gene symbol.
- **Chromosome**: Chromosome number.
- **Position**: Position on the chromosome.
- **Allele Frequency**: Frequency of the allele.
- **Impact**: Type of impact.
- **ClinVar**: ClinVar status.
- **Patient Genotype**: Genotype of the patient.

---

**Table Rows**
- **Gene**: various gene symbols.
- **Chromosome**: various chromosome numbers.
- **Position**: various positions on the chromosome.
- **Allele Frequency**: various allele frequencies.
- **Impact**: various types of impact.
- **ClinVar**: various ClinVar statuses.
- **Patient Genotype**: various genotypes of the patient.
### Pathogenic Mutations

<table>
<thead>
<tr>
<th>Patient</th>
<th>Leading Cause</th>
<th>Sex</th>
<th>Age [yr]</th>
<th>eGFR [ml/min/1.73m²]</th>
<th>UACR [mg/g]</th>
<th>Gene</th>
<th>Mutation</th>
<th>Other Variants</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>FSGS</td>
<td>M</td>
<td>54</td>
<td>51</td>
<td>357</td>
<td>COL4A5</td>
<td>c.1871G&gt;A G624D</td>
<td>Y</td>
</tr>
<tr>
<td>2</td>
<td>FSGS</td>
<td>F</td>
<td>56</td>
<td>49</td>
<td>303</td>
<td>COL4A5</td>
<td>c.4877C&gt;T (S1626L) htz</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>FSGS</td>
<td>M</td>
<td>57</td>
<td>30</td>
<td>4</td>
<td>COL4A5</td>
<td>c.4549C&gt;A (P1517T) htz</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>FSGS</td>
<td>F</td>
<td>42</td>
<td>37</td>
<td>1123</td>
<td>COL4A5</td>
<td>GLA</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Other GN</td>
<td>F</td>
<td>44</td>
<td>54</td>
<td>4881</td>
<td>GLA</td>
<td>c.427G&gt;A (A143T) htz</td>
<td>Y</td>
</tr>
<tr>
<td>6</td>
<td>FSGS</td>
<td>M</td>
<td>24</td>
<td>72</td>
<td>2228</td>
<td>INF2</td>
<td>c.641G&gt;A (R214H) htz</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>FSGS</td>
<td>2</td>
<td>45</td>
<td>80</td>
<td>1662</td>
<td>NPHS2</td>
<td>c.413G&gt;A (R138Q) htz</td>
<td>Y</td>
</tr>
<tr>
<td>8</td>
<td>Other GN</td>
<td>F</td>
<td>27</td>
<td>45</td>
<td>1975</td>
<td>NPHS2</td>
<td>p.R229Qhtz p.A284V</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>FSGS</td>
<td>M</td>
<td>39</td>
<td>33</td>
<td>4146</td>
<td>NPHS2</td>
<td>c.413G&gt;A (R138Q) htz</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>RPGN pauci</td>
<td>F</td>
<td>73</td>
<td>63</td>
<td>47</td>
<td>PAX2</td>
<td>c.76G&gt;C (V26L) htz</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>FSGS</td>
<td>F</td>
<td>24</td>
<td>61</td>
<td>292</td>
<td>PAX2</td>
<td>c.69_70insG</td>
<td>Y</td>
</tr>
<tr>
<td>12</td>
<td>Other GN</td>
<td>M</td>
<td>52</td>
<td>38</td>
<td>153</td>
<td>WT1</td>
<td>c.962G&gt;C (G321A) htz</td>
<td></td>
</tr>
</tbody>
</table>
Future Directions

Discussion

• Complete sequencing, joint calling and annotation

• Check extra-renal manifestations supporting pathogenicity, unusual manifestations

• Identify functional collaborators (VUS)

• Assess variant burden in same genes in healthy individuals

• Evaluate combined SNP risk score and mutations
Future Directions

Discussion

• Evaluate combined SNP risk score and mutations

• GWAS data for all patients available

• Check if GWAS alleles can help explain unusual manifestations or predisposition in mutation-negative patients
Summary

Discussion

• Among adult CKD patients seen by nephrologists, with eGFR of >30 ml/min/1.73m$^2$, and with CKD of presumed primary glomerular etiology

• 5% carry definite exonic pathogenic mutations in known podocytopathy genes

• an additional 10% of patients carry VUS

• As expected, most pathogenic mutations map into AD genes, and all $NPHS2$ were compound htz

• All 4 patients with $COL4A5$ mutations presented as FSGS

• Challenges: benign rare variants, VUS, structural variation
Thank You

Project Team Freiburg
• Matthias Wuttke
• Anselm Hoppmann
• Ulla Schultheiß

Project team Heidelberg and Gdansk
• Franz Schaefer
• Beata Lipska-Ziętkiewicz
• All members of the Schaefer lab
• GCKD study
• Kai-Uwe Eckardt
• All Colleagues and Patients

EURenOmics  DFG  KfH  Bundesministerium für Bildung und Forschung  KIDGEM