Mapping the epigenetic basis of kidney disease

Katalin Susztak, MD, PhD
University of Pennsylvania
ksusztak@mail.med.upenn.edu
The Problem
End Stage Kidney Disease

ESRD care is ~ $30 billion/year ~10% of Medicare budget
Why do people develop Kidney Disease?
Chronic kidney disease; typical gene environmental disease

ATTCGAGTCA
GTCGGTTCAG
TTTCGG

Genome

0.33-0.7

Environment

Transcriptome

Kidney disease
GWAS for CKD in EUR population

[Graph showing genetic markers associated with CKD in EUR population]
How do they lead to kidney disease development?

- Causal SNP
- Target cell type
- Target gene
- Mode of dysregulation

88% non-coding regions

- intron 56%
- Intergenic 33%
- 3' UTR 2%
- TTS 2%
- exon 7%
1. The causal variant is localized to a regulatory region in a disease relevant cell type (kidney)

2. The variant alters target expression in disease relevant cell type (the kidney) via altering transcription factor binding

3. The target is expressed in disease relevant tissue (kidney)

4. The expression of target changes in kidney disease

5. Alterations in target expression causes kidney disease. The target is functional in the kidney
Integrated translational approach for target identification for chronic kidney disease

**Histopathology:** Standardized scoring 20 independent parameter

**Clinical data:**
- Kidney function
- CVD
- DM

**Transcriptome:** Microdissected glomeruli, tubuli
- RNAsequencing (n=500)

**Genotyping:**
- Affymetrix Axiom Biobank Chips

**Epigenome analysis:**
- Cytosine methylation (Infinium arrays)
- 500 blood samples
- 150 dissected tubule samples
- Histone modification
- Cell type ChIPs (podocytes, endo, mesangial cells, fibroblasts, tubule cells)
- Control vs. CKD tissue samples

**Human Kidney Tissue**
> 1,200
1. The causal variant is localized to a regulatory region in a disease relevant cell type (kidney)
CKD SNPs are enriched on kidney-specific enhancers in comparison to non-kidney cell lines.

Kidney-specific epigenome: Insulator, Enhancer, Promoter, Transcribed Regions.

Specific epigenomes of 9 different cell types:
- Adult Human Kidney
- H1 ES Cells
- NHLF (lung fibroblast)
CKD SNPs are enriched on tubule cell specific enhancers when comparing kidney cell lines.
Experimental validation
The causal variant is expressed on disease relevant cell type
Our framework to understand the genetics of kidney disease

1. The causal variant is localized to a regulatory region in a disease relevant cell type (kidney)

2. The variant alters target expression in disease relevant cell type (the kidney) via altering transcription factor binding

3. The target is expressed in disease relevant tissue (kidney)

4. The expression of target changes in kidney disease

5. Alterations in target expression causes kidney disease. The target is functional in the kidney
The variant alters target expression in disease relevant tissue (the kidney)

expression quantitative trait locus (eQTL)
Genotype driven gene expression changes (99 CEU kidneys)

- 435 eGenes (p-value < $10^{-8}$)
- 339 eGenes (p-value < $10^{-9}$)

Red spot: *cis*-eqtl
Distance: 1 Mb
P-value: 4.0E-10
Which gene is the target of the polymorphism?
Our framework to understand the genetics of kidney disease

1. The causal variant is localized to a regulatory region in a disease relevant cell type (kidney)

2. The variant alters target expression in disease relevant cell type (the kidney) via altering transcription factor binding

3. The target is expressed in disease relevant tissue (kidney)

4. The expression of target changes in kidney disease

5. Alterations in target expression causes kidney disease. The target is functional in the kidney
Which gene is expressed in the kidney?

<table>
<thead>
<tr>
<th>Gene Name</th>
<th>kidney</th>
</tr>
</thead>
<tbody>
<tr>
<td>GP2</td>
<td>3</td>
</tr>
<tr>
<td>UMOD</td>
<td>1362</td>
</tr>
<tr>
<td>PDILT</td>
<td>4</td>
</tr>
<tr>
<td>ACSM5</td>
<td>4</td>
</tr>
<tr>
<td>ACSM2A</td>
<td>106</td>
</tr>
<tr>
<td>ACSM2B</td>
<td>119</td>
</tr>
<tr>
<td>ACSM1</td>
<td>0.7</td>
</tr>
</tbody>
</table>
Our framework to understand the genetics of kidney disease

1. The causal variant is localized to a regulatory region in a disease relevant cell type (kidney)

2. The variant alters target expression in disease relevant cell type (the kidney) via altering transcription factor binding

3. The target is expressed in disease relevant tissue (kidney)

4. Target expression changes in kidney disease

5. Alterations in target expression causes kidney disease. The target is functional in the kidney
Expression of the target correlates with kidney function.

Figure 3.

UMOD and ACSM2A/2B expressions correlate with renal function. The expressions of (A) UMOD and (F) ACSM2A correlate with eGFR in tubule samples. The x axes represent eGFR (ml/min per 1.73 m²), whereas the y axes represent the normalized gene expression values of the transcript. Each dot represents transcript levels and eGFR values from a single kidney sample. The lines are the fitted correlation value (Pcorr, P value after Benjamini-Hochberg multiple testing correction). Immunohistochemistry of the samples with low and high mRNA expression showed differences of (B–E) the UMOD and (G–J) the ACSM2A expression on protein level. Scale bars, 50 mm.

Pearson R = 0.526
R² = 0.2768
Pcorr = 2.45 x 10⁻⁶
Our framework to understand the genetics of kidney disease

1. The causal variant is localized to a regulatory region in a disease relevant cell type (kidney)

2. The variant alters target expression in disease relevant cell type (the kidney) via altering transcription factor binding

3. The target is expressed in disease relevant tissue (kidney)

4. The expression of target changes in kidney disease

5. Alterations in target expression causes kidney disease. The target is functional in the kidney
Functional studies in model organism

Morpholino knock down in zebrafish embryo

Acsn3 200 uM

Pericardial edema (sign of kidney damage)
Functional studies in model organism
CONCLUSION 1.

Our Roadmap to understand GWAS associated hits
- Human tissue samples are needed
- Epigenome maps to identify regulatory DNA
- Model organisms to validate causal variant
- eQTL maps to identify target genes
- Examine kidney expression, correlation with kidney function
- Model organisms to validate gene function (zfish to mouse)

Our analysis indicate that ACSM gene family are likely targets of a common CKD GWAS hit on Chr16

Fatty acid metabolism might be the target of common CKD associated GWAS variant
These variants explain small fraction of heritability

Can be explained by sequence variants (SNP)

Inherited but we can not identify DNA sequence variation

Missing heritability

Larger sample size
Different ethnic groups
Deeper sequencing
Epigenetics
Different ethnic groups...  
...admixture study in YRB for kidney disease...

![Graph showing admixture analysis and chromosome 22 gene localization.](image)

Kopp et al. Nature Genetics
Association of Trypanolytic ApoL1 Variants with Kidney Disease in African Americans

Giulio Genovese,1,2* David J. Friedman,1,3* Michael D. Ross,4 Laurence Lecordier,5 Pierrick Uzureau,5 Barry I. Freedman,6 Donald W. Bowden,7,8 Carl D. Langefeld,8,9 Taras K. Oleksyk,10 Andrea L. Uscinski Knob,4 Andrea J. Bernhardy,1 Pamela J. Hicks,7,8 George W. Nelson,11 Benoit Vanhollebeke,5 Cheryl A. Winkler,12 Jeffrey B. Kopp,11 Etienne Pays,5† Martin R. Pollak1,13†
APOL1 variants cause kidney disease in mice

Beckerman, Park and Susztak unpublished
These variants explain small fraction of heritability

Can be explained by sequence variants (SNP)

Inherited but we can not identify DNA sequence variation

Missing heritability

EPIGENETICS
# Demographics of the research participants

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Control</th>
<th>Hypertension</th>
<th>DM</th>
<th>DKD</th>
</tr>
</thead>
<tbody>
<tr>
<td>n</td>
<td>23</td>
<td>23</td>
<td>20</td>
<td>21</td>
</tr>
<tr>
<td>Age (years) Mean ± SD</td>
<td>60.1±10.4</td>
<td>61.9±10.7</td>
<td>65.3±12.1</td>
<td>65.8±12.3</td>
</tr>
<tr>
<td>Ethnicity</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asian, Pacific Islander</td>
<td>1</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>White, non-Hispanic</td>
<td>2</td>
<td>5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Black, non-Hispanic</td>
<td>6</td>
<td>7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hispanic</td>
<td>2</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other&amp;Unknown</td>
<td>9</td>
<td>6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BMI (kg/m2) Mean ± SD</td>
<td>28.9±5.53</td>
<td>32.5±7.76</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Diabetes</td>
<td></td>
<td></td>
<td>20</td>
<td>21</td>
</tr>
<tr>
<td>Hypertension</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Proteinuria (dipstick)</td>
<td>0.09±0.29</td>
<td>0.5±0.8</td>
<td>0.74±1.19</td>
<td>3.0±1.7</td>
</tr>
<tr>
<td>Serum BUN (mg/dL) Mean ± SD</td>
<td>15.30±5.38</td>
<td>15.56±7.23</td>
<td>17.29±1.06</td>
<td>35.20±15.82</td>
</tr>
<tr>
<td>Serum creatinine (mg/dL) Mean ± SD</td>
<td>0.9±1.19</td>
<td>1.18±3.22</td>
<td></td>
<td></td>
</tr>
<tr>
<td>eGFR (ml/min/1.73 m²) Mean ± SD</td>
<td>74.47±10.20</td>
<td>30.32±17.21</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Histology</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tubular atrophy (%)</td>
<td>4.5±4.02</td>
<td>31.8±22.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Interstitial fibrosis (%)</td>
<td>5.5±4.18</td>
<td>29.6±19.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Glomerulosclerosis (%)</td>
<td>7.13±7.52</td>
<td>31.11±32.25</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mesangial matrix Expansion</td>
<td>0.32±0.40</td>
<td>1.38±1.18</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Arteriosclerosis Intima</td>
<td>1.0±0.7</td>
<td>2.0±0.8</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

## Genome wide Cytosine methylation analysis
- MRE-Chip
- Illumina 450K
Distinct Cytosine Methylation Profiles tubule cells obtained from patients with Diabetic kidney Disease

Cytosine methylation (p<10^{-16} ∆>13%)
Differential methylation occurs on kidney specific enhancers

Kidney specific epigenome

Promoter
Enhancer
Insulator
Transcribed

ENHANCERS

Promoter
Enhancer
Insulator
Tx Regions
Repressed/Heterochromatin

YA Ko.... Susztak Genome Biology 2013
Differentially methylated regions affect kidney specific transcription factor binding

**SIX2 binding site**

```
TTTAAATTTTGGTAGAGACGGGTTTCGCCATGTTCCG
```

---

**Kidney specific epigenome**

- Insulator
- Enhancer
- Promoter
- Transcribed

---

**Motif**

<table>
<thead>
<tr>
<th>Name</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>PITX2</td>
<td>4.40E-06</td>
</tr>
<tr>
<td>FOXP1</td>
<td>1.66E-05</td>
</tr>
<tr>
<td>IKAROS/IKZF</td>
<td>5.20E-05</td>
</tr>
<tr>
<td>ESR1/ER-alpha</td>
<td>7.13E-05</td>
</tr>
<tr>
<td>SIX2/SIX3</td>
<td>8.60E-05</td>
</tr>
<tr>
<td>SIX2</td>
<td>5.66E-04</td>
</tr>
<tr>
<td>EGR22</td>
<td>6.37E-04</td>
</tr>
<tr>
<td>ZFX</td>
<td>8.96E-04</td>
</tr>
<tr>
<td>GFI1</td>
<td>9.13E-04</td>
</tr>
<tr>
<td>GLI3</td>
<td>1.13E-03</td>
</tr>
<tr>
<td>CREB</td>
<td>1.29E-03</td>
</tr>
<tr>
<td>FPM315/ZNF263</td>
<td>1.57E-03</td>
</tr>
<tr>
<td>GKLF2isoform1/KLF4</td>
<td>1.57E-03</td>
</tr>
<tr>
<td>LUN</td>
<td>1.72E-03</td>
</tr>
<tr>
<td>SREBP21b</td>
<td>2.10E-03</td>
</tr>
<tr>
<td>TCFAP2E</td>
<td>2.66E-03</td>
</tr>
<tr>
<td>ZBTB3</td>
<td>2.97E-03</td>
</tr>
<tr>
<td>NKX6A</td>
<td>3.02E-03</td>
</tr>
<tr>
<td>RNF96/TRIM28/KAP1</td>
<td>3.48E-03</td>
</tr>
<tr>
<td>NF2E2Np45</td>
<td>3.63E-03</td>
</tr>
<tr>
<td>MEF22C</td>
<td>3.63E-03</td>
</tr>
<tr>
<td>TCF1/HNF1A</td>
<td>4.98E-03</td>
</tr>
<tr>
<td>HOXD11</td>
<td>7.60E-03</td>
</tr>
</tbody>
</table>

**Motif**

- **TCFAP2E**
- **HOXD11**

---

YA Ko.... Susztak Genome Biology 2013
Cytosine methylation differences correlate with transcript level changes

- Differentially methylated regions (DMR) (HELP): 4,751
- Unique nearby genes: 1,092
- Differentially expressed transcripts: 414

**Gene Ontology Terms:**
- Developmental process
- Cellular process
- Multicellular organismal
- Cellular component
- Biological adhesion
- Localization
- Biological regulation
- Establishment of localization
- Immune system process

**-log p-value:**
- 0
- 0.5
- 1
- 1.5
- 2
- 2.5
- 3
- 3.5
- 4
- 4.5

**Gene Locations:**
- Intergenic
- Gene body
- Promoter

**Gene Body Parts:**
- Cytosine methylation differences correlate with transcript level changes
Are there differences in histone tail modifications in CKD?

<table>
<thead>
<tr>
<th>Term</th>
<th>%</th>
<th>P-Value</th>
<th>Corrected</th>
</tr>
</thead>
<tbody>
<tr>
<td>developmental process</td>
<td>23.4</td>
<td>3.50E-20</td>
<td>7.60E-19</td>
</tr>
<tr>
<td>biological adhesion</td>
<td>6.4</td>
<td>1.80E-11</td>
<td>2.00E-10</td>
</tr>
<tr>
<td>multicellular organismal process</td>
<td>27.2</td>
<td>3.50E-08</td>
<td>2.50E-07</td>
</tr>
<tr>
<td>cellular process</td>
<td>59</td>
<td>1.70E-04</td>
<td>9.60E-04</td>
</tr>
<tr>
<td>reproduction</td>
<td>5</td>
<td>1.90E-02</td>
<td>8.20E-02</td>
</tr>
<tr>
<td>reproductive process</td>
<td>4.9</td>
<td>2.10E-02</td>
<td>7.60E-02</td>
</tr>
<tr>
<td>cellular component organization</td>
<td>14.7</td>
<td>2.60E-02</td>
<td>8.00E-02</td>
</tr>
</tbody>
</table>
Maternal calorie restriction in rats causes low nephron number, HTN, albuminuria and distinct epigenetic changes.

Poster by
The Epigenetics of Kidneys Is Altered in Offspring of Maternal Caloric Restriction
Howard Slomko, DO¹, Hye Heo, MD², Fabien Delahaye, PhD², Yongmei Zhao², Zhongfang Du MD¹ Kimberly J Reidy, MD¹ and Francine H Einstein, MD²
Conclusion
CONCLUSION 2.

Small but highly consistent cytosine methylation changes in CKD tubule samples

Methylation changes are enriched on kidney specific enhancer regions

Fibrosis and developmental genes are more affected by methylation changes

Kidney disease might have a “developmental” origin
Acknowledgment

Susztak Lab

Yi-An Ko (Graduate student)
Huigang Yi PhD
Hyun Mi Kang, PhD
Kimberly Reidy, MD
Esther Park, MD
Mariya Sweetwyne PhD
Seon Yeok Han MD PhD
Pazit Beckerman, MD
Nora Ledo MD
Frank Chinga
Mendy Liang
Kriti Gaur, PhD
Laura Malaga MD, PhD

Collaborators:
John Stam, UW
Casey Brown, Penn
Hongzhe Li, Penn
John Greally, Einstein
Shanon Fisher, Penn
Mike Pack, Penn
Anna Kottgen, Freiburg

NIDDK Pilot Award
www.diaomp.org

Lilly
Penn
NIH
Boehringer Ingelheim
American Diabetes Association
JDRF