Datasets, Doctors, and Disease: Bridging the gap from genomics analysis to clinical change.

February 13, 2014

W. Kimryn Rathmell, MD, PhD
Estimated New Cancer Cases* in the US in 2013

**Men** 854,790  **Women** 805,500

- Prostate  28%
- Lung & bronchus  14%
- Colon & rectum  9%
- Urinary bladder  6%
- Melanoma of skin  5%
- Kidney & renal pelvis  5%
- Non-Hodgkin lymphoma  4%
- Oral cavity  3%
- Leukemia  3%
- Pancreas  3%
- All Other Sites  20%
- Breast  29%
- Lung & bronchus  14%
- Colon & rectum  9%
- Uterine corpus  6%
- Thyroid  6%
- Non-Hodgkin lymphoma  4%
- Melanoma of skin  4%
- Kidney & renal pelvis  3%
- Pancreas  3%
- Ovary  3%
- All Other Sites  19%

*Excludes basal cell and squamous cell skin cancers and in situ carcinoma except urinary bladder.
Renal Cell Carcinoma (RCC)

- Originates in the renal cortex
- Most common solid lesion occurring in the kidney (80-85% of all primary renal neoplasms)
Outline

• Appreciating differences in similar tumors.
• Using biological signatures to improve prognosis.
• The problem/opportunity of heterogeneity.
• Integrating epigenetic programs into the clinical and biological picture.
**Renal Cell Carcinoma—not one disease**

<table>
<thead>
<tr>
<th>Subtype</th>
<th>Prevalence</th>
<th>Tumor Features</th>
<th>Microscopic Features</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clear cell carcinoma ccRCC</td>
<td>75–85%</td>
<td>Multinodular; large clear cells with prominent nucleoli, organized in nests surrounded by vessel bundles</td>
<td><img src="image1.png" alt="Image" /></td>
</tr>
<tr>
<td>Chromophilic (papillary) carcinomas pRCC</td>
<td>10–15%</td>
<td>Ball-shaped outline, trabecular pattern, foamy macrophages, commonly multifocal</td>
<td><img src="image2.png" alt="Image" /></td>
</tr>
<tr>
<td>Chromophobic carcinomas chRCC</td>
<td>5–10%</td>
<td>Bland nucleus, eosinophilic cytoplasm with central clearing.</td>
<td><img src="image3.png" alt="Image" /></td>
</tr>
</tbody>
</table>
KIRC, KICH, KIRP

A tale of three kidney cancer genomes
TCGA: What’s in a Core Data Set?

Data from Tissue Source Sites
- Complete path report
- Paired metastatic samples
- Double normals
- Treatment data

Core Data Set
- Synoptic path report
- Histology images
- Required clinical data
- Whole exome
- SNP 6.0 array
- mRNAseq
- miRNAseq
- Methylation array

Data Generated by GCCs & GSCs
- 50X WGS
- 8X WGS
- Methylseq
- RPPA
Meet ccRCC:
Metabolic Network: All glycolysis
Now, meet Chromophobe RCC

(a) CCRCC (n=437) and ChRCC (n=66). Bars represent individual cases, colored according to copy number changes: blue for copy loss, red for copy gain, and green for genome doubling cases. Cases are categorized into classic (n=47) and eosinophilic (n=19) subtypes.

(b) Frequency of nonsilent mutation (%) across different genes in 66 ChRCC cases.

Genes: TP53, PTEN, FLI1, AICDA, ATM, CDKN1A, MLL3, MTOR, RB1, SDHA, TSC1, TSC2, USP14, NRAS, PDK3, SETD2, SMARCB1, CDKN1B.
Different methylation, expression, and origin in the nephron.
A different biology-focus on mitochondria
Comparing Copy Number Variation:

- KIRP, KICH, and KIRC display very different SCNA patterns
  - Chromosome 17 is one good example.
Summary

• The renal cell carcinomas represent highly distinct, *unrelated* diseases.
• The cancer genome atlas provides a framework for defining a cancer.
ccA + ccB = ccRCC

Molecular stratification of clear cell Renal Cell Carcinoma
Most cancers of the kidney and renal pelvis are diagnosed when the disease is still localized to the primary site.

- **Localized Disease**: 56%
- **Loco-regional Spread**: 19%
- **Metastatic Spread**: 20%
- **Unknown**: 5%

Determining Prognosis: Anatomic Extent of Disease

• Most consistent factor used to determine RCC prognosis

5-year Cancer-specific Survival Based on TNM Stage

<table>
<thead>
<tr>
<th>TNM Stage</th>
<th>5-year Cancer-specific Survival</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stage I</td>
<td>91 ± 2.5%</td>
</tr>
<tr>
<td>Stage II</td>
<td>74 ± 6.9%</td>
</tr>
<tr>
<td>Stage III</td>
<td>67 ± 6.1%</td>
</tr>
<tr>
<td>Stage IV</td>
<td>32 ± 3.2%</td>
</tr>
</tbody>
</table>

External Validation of the Mayo Clinic Stage, Size, Grade, and Necrosis (SSIGN) Score for Clear-Cell Renal Cell Carcinoma in a Single European Centre Applying Routine Pathology

Richard Zigeuner, Georg Hutterer, Thomas Chromecki, Arvin Imamovic, Karin Kampel-Kettner, Peter Rehak, Cord Langner, Karl Pummer

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Score points</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pathologic tumour category</td>
<td></td>
</tr>
<tr>
<td>pT1</td>
<td>0</td>
</tr>
<tr>
<td>pT2</td>
<td>1</td>
</tr>
<tr>
<td>pT3a</td>
<td>2</td>
</tr>
<tr>
<td>pT3b</td>
<td>2</td>
</tr>
<tr>
<td>pT3c</td>
<td>2</td>
</tr>
<tr>
<td>pT4</td>
<td>0</td>
</tr>
<tr>
<td>Nodal status</td>
<td></td>
</tr>
<tr>
<td>Nx</td>
<td>0</td>
</tr>
<tr>
<td>pN0</td>
<td>2</td>
</tr>
<tr>
<td>pN1</td>
<td>2</td>
</tr>
<tr>
<td>pN2</td>
<td>2</td>
</tr>
<tr>
<td>Metastasis category</td>
<td></td>
</tr>
<tr>
<td>M0</td>
<td>0</td>
</tr>
<tr>
<td>M1</td>
<td>4</td>
</tr>
<tr>
<td>Tumour size</td>
<td></td>
</tr>
<tr>
<td>&lt;5 cm</td>
<td>0</td>
</tr>
<tr>
<td>≥5 cm</td>
<td>2</td>
</tr>
<tr>
<td>Tumour grade</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>Tumour necrosis</td>
<td></td>
</tr>
<tr>
<td>Absent</td>
<td>0</td>
</tr>
<tr>
<td>Present</td>
<td>2</td>
</tr>
</tbody>
</table>

Fig. 1. Kaplan-Meier curves showing estimated 10-yr cancer-specific survival probabilities of the present study population related to respective SSIGN-score categories (p < 0.001, log-rank test).
Consensus clusters permit refined analysis

Brannon, et al, Genes and Cancer, 2010
ccA overexpresses RCC pathways

Angiogenesis

ccA

ccB

FLT4
FLT1
VEGFB
ENG
KDR
BAI1

Beta-oxidation

ccA

ccB

ACADL
ACADS
ACADM
ACAT1
HADHA

Fatty Acid Metabolism

ccA

ccB

ACSL4
ACSL1
ACADL
DCI
ACADSB
ACADS
ACAA2
ACAT1
ACADM
EHHAADH
ALDH3A2
ADH5
HADHB
PECI
ALDH9A1
HADHA
ALDH7A1
**ccB overexpresses aggressive genes**

![Gene expression heatmap showing overexpression of aggressive genes in ccB compared to ccA.](image-url)

- **Epithelial-to-Mesenchymal Transition:**
  - ccA
  - ccB
  - Genes: CYP1B1, MMP2, B2M, COL6A1, ADA, TNXB, CDH15, SLPI, Ti11, PCOLCE, DCN, PDGFRA, TNC, MMP12, UPP1, CDH2, FMO1, DAB2, IFIT3, PROCR, VIM, SDC2

- **TGFbeta:**
  - ccA
  - ccB
  - Genes: MTA1, MMP17, TNC, COL16A1, THBS2

- **Wnt Targets:**
  - ccA
  - ccB
  - Genes: CCND1, NRCAM, CD44, LEF1, BIRC5, FST, PPARD

*Brannon, et al., Genes and Cancer, 2010*
Marked survival differences between subtypes in validation set

Brannon, et al, Genes and Cancer, 2010
Creating a predictive tool

- 95 clear cell tumors
- LAD and ConsensusCluster analysis
- Set aside arrays with non-concordant assignments
- 72 arrays (microarray standard set) (69 tumors, 3 replicates)
  - 43 ccA arrays (42 tumors, 1 replicate)
  - 29 ccB arrays (27 tumors, 2 replicates)
- Prediction Analysis for Microarrays (PAM)
- Predictive biomarkers: ClearCode34
Prognostic value of ClearCode34 evaluated in TCGA

- **Recurrence-free Survival (probability)**
  - ccA: (n=205), No. of events: 50, Median RFS, months: 91, HR: 2.3; 95% CI: 1.6 to 3.3; P=4.3e-06
  - ccB: (n=175), No. of events: 75, Median RFS, months: 53

- **Cancer-Specific Survival (probability)**
  - ccA: (n=205), No. of events: 14, Median CSS, months: --
  - ccB: (n=175), No. of events: 29, Median CSS, months: --
  - HR: 2.9; 95% CI: 1.6 to 5.6; P=0.0005

- **Overall Survival (probability)**
  - ccA: (n=205), No. of events: 35, Median OS, months: 94
  - ccB: (n=175), No. of events: 58, Median OS, months: 65
  - HR: 2.4; 95% CI: 1.6 to 3.7; P=2.3e-05
Prognostic value of ClearCode34 validated in UNC cohort

Recurrence-Free Survival (probability)

<table>
<thead>
<tr>
<th>Time (months)</th>
<th>ccA</th>
<th>ccB</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>50</td>
<td>0.7</td>
<td>0.7</td>
</tr>
<tr>
<td>100</td>
<td>0.4</td>
<td>0.4</td>
</tr>
<tr>
<td>150</td>
<td>0.2</td>
<td>0.2</td>
</tr>
<tr>
<td>200</td>
<td>0.1</td>
<td>0.1</td>
</tr>
</tbody>
</table>

No. of events ccA: 26, ccB: 59
Median RFS, months ccA: 88, ccB: 52
HR, 2.1; 95% CI, 1.3 to 3.4; P=0.001
Prognostic value of ClearCode34 validated in TCGA

<table>
<thead>
<tr>
<th>Variable</th>
<th>Univariate Model</th>
<th>Multivariate Model</th>
<th>Final Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subtype*</td>
<td>2.2, &lt;.001</td>
<td>1.7, &lt;.001</td>
<td>.0009</td>
</tr>
<tr>
<td>Stage$</td>
<td>&lt;.001</td>
<td></td>
<td>.0007</td>
</tr>
<tr>
<td>II</td>
<td>1.4, .121</td>
<td>1.3, .307</td>
<td></td>
</tr>
<tr>
<td>III/IV</td>
<td>2.4, &lt;.001</td>
<td>1.8, &lt;.001</td>
<td></td>
</tr>
<tr>
<td>Grade</td>
<td></td>
<td>&lt;.001</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>1.6, .002</td>
<td>1.3, .138</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>5.1, &lt;.001</td>
<td>3.1, &lt;.001</td>
<td></td>
</tr>
</tbody>
</table>

Abbreviation: HR, hazard ratio
Subtype ccA was used as reference in univariate and multivariate analysis.
$ Stage I was used as reference in univariate and multivariate analysis. Stage was encoded as an ordinal variable with three levels.
|| Grade 1 and 2 were combined and used as reference in univariate and multivariate analysis. Grade was encoded as an ordinal variable with three levels.
Integrated prognostic models can evaluate risk outcomes

<table>
<thead>
<tr>
<th>Group</th>
<th>Risk Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low</td>
<td>0-0.5</td>
</tr>
<tr>
<td>Intermediate</td>
<td>0.5-1.5</td>
</tr>
<tr>
<td>High</td>
<td>&gt;1.5</td>
</tr>
</tbody>
</table>

- **Recurrence-free Survival**
  - Low Risk (N=98)
  - Intermediate Risk (N=140)
  - High Risk (N=28)
  - Log-rank $P=3.04 \times 10^{-9}$

- **Cancer-Specific Survival**
  - Low Risk
  - Intermediate Risk
  - High Risk
  - Log-rank $P=2.03 \times 10^{-8}$
ClearCode34 Model outperforms established algorithms
Summary

• Clear cell RCC can be divided based on gene expression into two groups.
• ccA and ccB tumors can be discriminated with 34 genes.
• A nanostring probeset allows assignment in fixed clinical specimens.
• Biomarkers add to clinical data in predicting risk assignments.
Heterogeneity-Hype, Hysteria, or Headache
RCC tumors-heterogeneity, with convergent evolution
Images of renal tumors:
Imaging, another look at heterogeneity

The same tumor can appear homogeneous by one method, and heterogeneous by another.
Measuring classes: Small renal masses

Low degree of gene expression heterogeneity in small tumors

<table>
<thead>
<tr>
<th>Sample</th>
<th>Biopsies</th>
<th>Sublocations</th>
<th>Classification</th>
</tr>
</thead>
<tbody>
<tr>
<td>160212</td>
<td>3</td>
<td>12</td>
<td>ccA</td>
</tr>
<tr>
<td>170212</td>
<td>3</td>
<td>12</td>
<td>ccA</td>
</tr>
<tr>
<td>240212</td>
<td>2</td>
<td>15</td>
<td>ccA</td>
</tr>
<tr>
<td>50312</td>
<td>3</td>
<td>15</td>
<td>ccA</td>
</tr>
<tr>
<td>80312</td>
<td>3</td>
<td>15</td>
<td>ccA</td>
</tr>
<tr>
<td>190412</td>
<td>3</td>
<td>15</td>
<td>ccA</td>
</tr>
</tbody>
</table>
A tale of one tumor.
DNA Heterogeneity

A.

<table>
<thead>
<tr>
<th>Tumor</th>
<th>% reads mapped</th>
<th>% perfect pair</th>
</tr>
</thead>
<tbody>
<tr>
<td>L1-A</td>
<td>93.4</td>
<td>80.5</td>
</tr>
<tr>
<td>L1-B</td>
<td>91.1</td>
<td>66.8</td>
</tr>
<tr>
<td>L1-C</td>
<td>88.4</td>
<td>49.5</td>
</tr>
<tr>
<td>L1-D</td>
<td>90.5</td>
<td>65.1</td>
</tr>
<tr>
<td>L1-E</td>
<td>89.2</td>
<td>56.7</td>
</tr>
</tbody>
</table>

B.

C.
Effects of tumor purity.
Heterogeneity of gene expression
Variance by gene expression
And the winner for most stable platform is...
Summary

- Heterogeneity is the new normal, with pathway homogeneity (convergent mutations).
- Renal cell carcinomas can be heterogeneous, but this likely emerges with stage.
- Heterogeneity: DNA > RNA Biomarkers > RNA > enhancers
- Can imaging enable us to take a holistic view?
SETD2, transcription, and the histone code

“Unraveling” the cancer genome
SETD2: a required H3K36 methyltransferase
SETD2 is mutated in ccRCC
**SETD2: a required H3K36 methyltransferase**

Loss of SETD2 has been shown to be associated with:

- Decreased global H3K36me3 levels
- Differential exon inclusion for individual genes (Luco et al., *Science* 2010)
Tumors with SETD2 mutations display altered chromatin organization

With Jeremy Simon

[Diagram showing FAIRE signal, H3K36me3 ChIP, and FAIRE signal for Normal Kidney, SETD2 normal tumor, and SETD2 mutant tumor.]
Tumors with SETD2 mutations display aberrant mRNA processing in poly(A)+ RNA

Types of RNA Processing Defects:

- Altered exon utilization: 66%
- Intron retention: 12%
- Alternate TSS/TTS: 22%

3929 transcripts with p < 0.01

~25% SETD2 mutant vs. wild-type

Darshan Singh, Jeremy Simon, Kate Hacke
The Cancer Genome Atlas ccRCC
Tumors with SETD2 mutations display aberrant mRNA processing in poly(A)+ RNA
Sites of altered splicing display an increase in chromatin accessibility.

Diagram showing misspliced exon start with enrichment peaks for H3K36me3 and loss of H3K36me3 at exon start and intron. FAIRE enrichment and ChIP data comparing H3K36me3 deficient and normal conditions.
All three SETD2 alleles in 7860 cells are targeted

SETD2 Wild-type

Left Target  Spacer  Right Target

5′ - TCATGTAACATCCAGGCCACGCTGGCTACTACCACAGCAGTAGCATCTCCA - 3′

Allele: Representative SETD2 Inactivation

#1  5′ - TCATGTAACATCCAGGCCACGCTGG----T----CAGCAGTAGCATCTCCA - 3′
#2  5′ - TCATGTAACATCCAGGCCACGCT----ACTACCACAGCAGTAGCATCTCCA - 3′
#3  5′ - TCATGTAACATCCAGGCCACGCTGGGC-ACTACCACAGCAGTAGCATCTCCA - 3′
Single cell sorting isolates SETD2 inactivated, H3K36me3-negative clones
SETD2 loss increases cell proliferation

* p < 0.05
Shown = HKCs – human renal cell line with SV40 transformation
Confirmed in: 293Ts and 7860s (human ccRCC cell line)
SETD2 loss increases anchorage-independent growth

Cell line = 7860s

*p<0.05
SETD2 loss is sufficient for anchorage-independent growth

Cell line = HKCs
**Isolated SETD2 loss results in widespread RNA processing defects**

- 1269 aberrantly processed transcripts $\rightarrow$ 1010 individual genes
- $\sim$32% overlap with aberrant transcripts in SETD2 mutant tumors
- Overlapping transcripts affect a wide variety of cellular processes
Isolated SETD2 loss results in widespread RNA processing defects

- Chromosome organization
- Ubl conjugation
- Nucleoplasm
- Chromatin regulator
- Cytoskeletal organization
- Protein-lysine N-methyltransferase activity
- Cell cycle
- DNA recombination
- Actin binding
- Histone modification
- Chromatin remodeling complex
- Transcription from RNAPII promoter
- Telomere maintenance
- Regulation of cell motion
- Transcription regulation

\( p = 0.05 \)
Summary

• SETD2 mutation is associated with changes in chromatin pattern and RNA processing.
• Association with loss of nucleosome at misspliced exon starts.
• SETD2 loss confers a proliferative and survival advantage.
Acknowledgments

Rathmell Lab
Kate Hacker
Alex Arreola
Samira Brooks
Zufan Debebe, PhD
Catherine Fahey
Sudarshan Mohan
Neal Rasmussen, PhD
Oishee Sen
Adam Sendor

Rathmell Lab Past Members
Rose Brannon, PhD
Shufen Chen, MD, PhD
Lance Cowey, MD
Caroline Martz Lee, MD, PhD
Courtney McGuire, MS
Tricia Wright, PhD

Translational Pathology Laboratory
Clinical Cytogenetics, Genomics Core

Davis Lab
Ian Davis, MD, PhD
Jeremy Simon, PhD

Jordan Shavit, MD, PhD (University of Michigan)

Strahl Lab
Brian Strahl, PhD
Deepak Jha

Bhanot Lab
Gyan Bhanot, PhD
Michael Seiler, PhD
Anupama Reddy, PhD

Joel Parker, PhD

The Cancer Genome Atlas
Particularly: Chad Creighton, Marston Linehan, Richard Gibbs, Kenna Shaw