Surveying the genomic landscape of tumours and tumour models- the next frontier

Sunday 24 February 2019
Plenary 1, 9:30 - 10:30 AM
RCPA Pathology Update 2019, Melbourne, Australia

Obi Griffith <obigriffith@wustl.edu>
Cancer genomics research has exploded with the rapid advances in DNA sequencing technologies.
How does it work? Short read alignments are the fundamental currency of cancer genome analysis

- Alignment is about fitting individual pieces (reads) into the correct part of the puzzle.
- The human genome project gave us the picture on the box cover (the reference genome).
- Imperfections in how the pieces fit can indicate damage or variation in picture.

Reference: AGCCTGAGACCGTAAAAAAGTCAAG

A read sequence: GAGACCGTAAAAACGTC

A variant
Whole genome, exome and transcriptome sequencing allows us to detect and confirm many different ‘omic events types.
Tumor genome analysis will typically reveal dozens to thousands of alterations of multiple types per patient.
First cancer whole genomes sequenced in last 10 years

- 2008 - AML
  
  DNA sequencing of a cytogenetically normal acute myeloid leukaemia genome
  
  Timothy J. Ley, ..., Brian C. Druker

- 2009 - Breast
  
  Mutational evolution in a lobular breast tumour profiled at single nucleotide resolution
  
  Schwab P, Shi L, Ryt, Allen Delaney, Karen Mark, ... David Housman, M

- 2010 - SCLC
  
  A small-cell lung cancer genome with complex signatures
  
  Erin O. Poessner, Philip J. Stranges, Helen R. Davies, ... R. And

- 2010 - MEL
  
  A comprehensive catalog of mutations from a human
  
  Erin O. Poessner, Philip J. Stranges, Helen R. Davies, ... R. And

- 2011 - PRC
  
  The genomic complexity of primary human prostate cancer
  
  Michael J. Savage, Michael G. Lawrence, ... G. David DeSouza
Cancer genomics has experienced a decade of explosive growth with large international initiatives (TCGA and ICGC).
These global consortia have been highly successful

- 15,000+ tumors profiles across ~30 tumor types
- Many high-profile publications
  - 70+ official TCGA papers, hundreds more, thousands of citations
- Established methods and tools
- Formed valuable reference datasets
- Helped survey the landscape of most major tumor types
- Identified many novel cancer driver genes and pathways
Pan-cancer sequencing surveys have identified the predominant driver mutations for each major cancer type

Revised our understanding of cancer predisposition

- 871 predisposition variants/CNVs discovered in 8% of 10,389 cases of 33 cancers

Defined new molecular subtypes of disease that both span and further stratify tissue-of-origin classifications

- Clinically relevant molecular subtypes have been identified in multiple tumor types (breast, AML, HNSCC, DLBCL, etc)
- Some tumor types coalesce across traditional classifications
- In some cases specific molecular alterations may support rationale drug-repurposing
  - But, not always...
  - TRK inhibitors vs ERBB2 inhibitors vs BRAF inhibitors
Tumour genome sequencing mission accomplished?
Common tumor types

Primary tumors

Exomes + some WGS and RNAseq

European populations

Protein coding space

Preliminary analyses

Moderate sequence depth/breadth

Some clinical annotations
Rare and challenging tumor types

Metastatic, relapsed, resistant tumors

Comprehensive omic profiling

Ethnic diversity

Characterization of models (mouse, PDX)

Spatial/Temporal heterogeneity

High sequence depth and breadth

Deep clinical annotations

Non-coding space

Increased sample sizes (power)

Not a complete list...
Rare and challenging tumor types

Metastatic, relapsed, resistant tumors

Comprehensive omic profiling

Ethnic diversity

Characterization of models (mouse, PDX)

Spatial/Temporal heterogeneity

High sequence depth and breadth

Deep clinical annotations

Non-coding space

Increased sample sizes (power)
• Androgen receptor is affected by mutation or structural variation in 85% of mCRPC

Quigley et al. 2018. Cell
Rare and challenging tumor types

Metastatic, relapsed, resistant tumors

Comprehensive omic profiling

Ethnic diversity

Characterization of models (mouse, PDX)

Spatial/Temporal heterogeneity

High sequence depth and breadth

Deep clinical annotations

Non-coding space

Increased sample sizes (power)
Deeper sequencing reveals hidden tumor architecture

**Standard**
- Basic analysis
- 30x whole genome
- 50x exome

**Optimized**
- Comprehensive analysis
- 300x whole genome
- 400x exome
- 10,000x targeted RNA-seq

**Primary Tumor**
- IDH2
- FLT3/IDH1
- NPM1/DNMT3A

**Relapse**
- IDH2
- NPM1/DNMT3A

**Clonal Evolution**
- 5 driver mutations,
  1 cell of origin, 2 subclones

- 10 driver mutations,
  2 cells of origin, 6 subclones
Rare and challenging tumor types

Metastatic, relapsed, resistant tumors

Comprehensive omic profiling

Ethnic diversity

Characterization of models (mouse, PDX)

Spatial/Temporal heterogeneity

High sequence depth and breadth

Deep clinical annotations

Non-coding space

Increased sample sizes (power)
Sample collection in a neoadjuvant window treatment (Pembrolizumab) clinical trial (n = 35)

Pre-treatment  Pembrolizumab  Surgery (Post-treatment)
Multiple physical locations sampled from post-surgery specimens
Rare and challenging tumor types

Metastatic, relapsed, resistant tumors

Comprehensive omic profiling

Ethnic diversity

Characterization of models (mouse, PDX)

Spatial/Temoral heterogeneity

High sequence depth and breadth

Deep clinical annotations

Non-coding space

Increased sample sizes (power)
Recurrent TERT promoter mutations in cancer result in increased TERT expression

**TERT Promoter Mutations in Familial and Sporadic Melanoma**

Susanne Horn,1,2 Adina Figl,1,2 P. Sivaramakrishna Rachakonda,1 Christine Fischer,3 Antje Sucker,2 Andreas Gast,1,2 Stephanie Kadel,1,2 Iris Moll,2 Eduardo Nagore,4 Kari Hemminki,1,5 Dirk Schadendorf,2† Rajiv Kumar1†

**Highly Recurrent TERT Promoter Mutations in Human Melanoma**

Franklin W. Huang,1,2,3† Eran Hodis,1,3,4† Mary Jue Xu,1,3,4 Gregory V. Kryukov,1 Lynda Chin,5,6 Levi A. Garraway1,2,3†
TERT promoter mutations found in many other tumor types

| TERT promoter mutations extremely common |
| Occur in proximal promoter ~100-300bp upstream of TSS |
| Why missed until 2013? |
| GC content |
| Coding-space blinders |

<table>
<thead>
<tr>
<th>Tissue Type</th>
<th>n</th>
<th>Ref</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bladder</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Urothelial carcinoma in upper urinary epithelium</td>
<td>52</td>
<td>[108]</td>
</tr>
<tr>
<td>Non-invasive and flat papillary urothelial carcinoma</td>
<td>59</td>
<td>[35]</td>
</tr>
<tr>
<td>Urothelial carcinoma with glandular differentiation</td>
<td>29</td>
<td>[107]</td>
</tr>
<tr>
<td>Urothelial carcinoma</td>
<td>21</td>
<td>[34]</td>
</tr>
<tr>
<td>Non-invasive and flat papillary urothelial carcinoma</td>
<td>82</td>
<td>[43]</td>
</tr>
<tr>
<td>Urothelial carcinoma with glandular differentiation</td>
<td>327</td>
<td>[58]</td>
</tr>
<tr>
<td>Urothelial carcinoma</td>
<td>52</td>
<td>[116]</td>
</tr>
<tr>
<td>Non-invasive and flat papillary urothelial carcinoma</td>
<td>252</td>
<td>[74]</td>
</tr>
<tr>
<td>Urothelial carcinoma with glandular differentiation</td>
<td>468</td>
<td>[118]</td>
</tr>
<tr>
<td>Urothelial carcinoma</td>
<td>240</td>
<td>[108]</td>
</tr>
<tr>
<td>Non-invasive and flat papillary urothelial carcinoma</td>
<td>21</td>
<td>[139]</td>
</tr>
<tr>
<td>Atypical fibroxanthomas</td>
<td>27</td>
<td>[119]</td>
</tr>
<tr>
<td>Pleomorphic dermal sarcomas</td>
<td>34</td>
<td>[119]</td>
</tr>
<tr>
<td>Myxoid liposarcoma</td>
<td>24</td>
<td>[34]</td>
</tr>
<tr>
<td>Solitary fibrous tumor</td>
<td>39</td>
<td>[109]</td>
</tr>
<tr>
<td>Malignant peripheral nerve sheath tumor</td>
<td>31</td>
<td>[109]</td>
</tr>
<tr>
<td>Synovial sarcoma</td>
<td>35</td>
<td>[109]</td>
</tr>
<tr>
<td>Metastatic melanoma</td>
<td>25</td>
<td>[109]</td>
</tr>
<tr>
<td>Basal cell carcinoma</td>
<td>53</td>
<td>[32]</td>
</tr>
<tr>
<td>Squamous cell carcinoma</td>
<td>145</td>
<td>[69]</td>
</tr>
<tr>
<td>Metastatic melanoma</td>
<td>42</td>
<td>[49]</td>
</tr>
<tr>
<td>Basal cell carcinoma</td>
<td>32</td>
<td>[126]</td>
</tr>
<tr>
<td>Squamous cell carcinoma</td>
<td>94</td>
<td>[70]</td>
</tr>
<tr>
<td>Metastatic melanoma</td>
<td>26</td>
<td>[49]</td>
</tr>
<tr>
<td>Basal cell carcinoma</td>
<td>27</td>
<td>[77]</td>
</tr>
<tr>
<td>Squamous cell carcinoma</td>
<td>34</td>
<td>[126]</td>
</tr>
<tr>
<td>Metastatic melanoma</td>
<td>70</td>
<td>[33]</td>
</tr>
<tr>
<td>Melanomas</td>
<td>56</td>
<td>[43]</td>
</tr>
<tr>
<td>Melanomas</td>
<td>116</td>
<td>[70]</td>
</tr>
<tr>
<td>Melanomas</td>
<td>38</td>
<td>[139]</td>
</tr>
<tr>
<td>Primary melanoma</td>
<td>77</td>
<td>[32]</td>
</tr>
<tr>
<td>Primary melanoma</td>
<td>145</td>
<td>[69]</td>
</tr>
<tr>
<td>Primary melanoma</td>
<td>207</td>
<td>[48]</td>
</tr>
<tr>
<td>Primary melanoma</td>
<td>53</td>
<td>[134]</td>
</tr>
<tr>
<td>Primary melanoma</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
We have designed a Custom Regulome Capture sequencing assay - explore non-coding space of breast cancers
Rare and challenging tumor types

Metastatic, relapsed, resistant tumors

Comprehensive omic profiling

Ethnic diversity

Characterization of models (mouse, PDX)

Spatial/Temporal heterogeneity

High sequence depth and breadth

Deep clinical annotations

Non-coding space

Increased sample sizes (power)
Rare tumor types can be informative - Primary HCC with three large tumor nodules - mixed conventional and fibrolamellar HCC

- ~25 year old female
- At least 8 surgeries, two rounds of chemo, five rounds of radiation, died after 44 months of treatment
- Deep exome coverage of tumor (352X), normal (113X), RNA-seq data
- Very few point mutations or copy-number alterations detected

RNAseq revealed Intra-chromosomal gene fusion *DNAJB1-PRKACA*

Exon 1 of DNAJB1 fused to exons 2-10 of PRKACA in head-to-tail fashion
Detection of a Recurrent **DNAJB1-PRKACA** Chimeric Transcript in Fibrolamellar Hepatocellular Carcinoma

Joshua N. Honeyman,1,2,* Elana P. Simon,1,3,* Nicolas Robine,4* Rachel Chiaroni-Clarke,1 David G. Darcy,1,2 Irene Isabel P. Lim,1,2 Caroline E. Gleason,4 Jennifer M. Murphy,1,2 Brad R. Rosenberg,2 Lydia Teegan,2 Constantin N. Takacs,1 Sergio Botero,1 Rachel Belote,1 Soren Germer,4 Anne-Katrin Emde,1 Vladimir Vacic,4 Umesh Bhanot,6 Michael P. LaQuaglia,2 Sanford M. Simon1†

Fibrolamellar hepatocellular carcinoma (FL-HCC) is a rare liver tumor affecting adolescents and young adults with no history of primary liver disease or cirrhosis. We identified a chimeric transcript that is expressed in FL-HCC but not in adjacent normal liver and that arises as the result of a ~400-kilobase deletion on chromosome 19. The chimeric RNA is predicted to code for a protein containing the amino-terminal domain of DNAJ1, a homolog of the molecular chaperone DNAJ, fused in frame with PRKACA, the catalytic domain of protein kinase A. Immunoprecipitation and Western blot analyses confirmed that the chimeric protein is expressed in tumor tissue, and a cell culture assay indicated that it retains kinase activity. Evidence supporting the presence of the **DNAJB1-PRKACA** chimeric transcript in 100% of the FL-HCCs examined (15/15) suggests that this genetic alteration contributes to tumor pathogenesis.

- Since shown to be present in 100% of FL-HCC cases and mixed FL-HCC (our study)
- Potentially perfectly sensitive and specific diagnostic tool
- PRKACA kinase a clear potential drug target
  - existing inhibitors non-specific
Increased capacity and new technologies are allowing study of previously inaccessible tumor types

- Hodgkins Lymphoma - has remained a challenge because HRS cells account for ~1-5% of cells in the tumor tissue.
  - Isolate HRS cells and then apply genomic techniques
    - Flow sorting - very challenging cell type
    - Laser capture microdissection - extremely low inputs for sequencing
  - Handful of cell lines have been profiled
  - Almost no genome-wide sequencing data exists

Felicia Gomez

Fehniger
Brute force strategies now possible - Ultra Deep Exome sequencing of 31 Hodgkins tumor/normal pairs

- IDT Exome capture reagent
- Three KAPA libraries were constructed/sample
- Libraries were sequenced across eight lanes of an Illumina HiSeq
- Somatic Variant Calling
  - SNV were called using 5 variant callers
  - Indels were called using 4 variant callers
- Target depth ~1000x
- Validation with Haloplex technology
Deep Exome Sequencing: ~1000X coverage achieved

AUC\(_{\geq 1000x} = 0.44\)
Median Coverage = 926X

AUC\(_{\geq 1000x} = 0.50\)
Median Coverage = 1009X
Extensive filtering and manual review required to go from >1M candidates to thousands of high quality variant candidates

<table>
<thead>
<tr>
<th>Remaining</th>
<th>Filtered</th>
<th>Total Filtered</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Variants</td>
<td>1226791</td>
<td>0</td>
</tr>
<tr>
<td>Include Chromosome 1-22, X, and MT</td>
<td>1192606</td>
<td>34185</td>
</tr>
<tr>
<td>Include only Coding Sites</td>
<td>7630</td>
<td>1184976</td>
</tr>
<tr>
<td>Exclude sites with normal VAF &gt;= 5%</td>
<td>7563</td>
<td>67</td>
</tr>
<tr>
<td>Exclude sites with tumor VAF &lt;= 0.5%</td>
<td>7543</td>
<td>20</td>
</tr>
<tr>
<td>Exclude sites with &lt;= 5 variant supporting reads in the tumor</td>
<td>6862</td>
<td>681</td>
</tr>
<tr>
<td>Exclude sites with &lt;= 50x in normal</td>
<td>6596</td>
<td>266</td>
</tr>
<tr>
<td>Exclude sites with &lt;= 50x in tumor</td>
<td>6591</td>
<td>5</td>
</tr>
<tr>
<td>Exclude sites with &lt;= 0.1% Allele Frequency in ExAC</td>
<td>5383</td>
<td>1208</td>
</tr>
<tr>
<td>Exclude Sites that Failed Normals Filter</td>
<td>4930</td>
<td>453</td>
</tr>
<tr>
<td>Exclude Sites that Failed Per Library Filter</td>
<td>4788</td>
<td>142</td>
</tr>
<tr>
<td>Sites Added From Re-Calling Variants</td>
<td>4822</td>
<td>34</td>
</tr>
<tr>
<td>Sites the Passed Manual Review</td>
<td>4702</td>
<td>120</td>
</tr>
</tbody>
</table>

- 3689 genes are represented in the sites that passed manual review
- 282 genes are mutated in two or more patients
• 37 genes have mutations in 3 or more samples
• IGLL5 - 22.6%
• TNFAIP3 & SOCS1 - 19.3%
• ITPKB & B2M - 12.9%
• BTG1, ACTB, RNASE7 & PCDHB6 - 9.6%

Represents one of the first comprehensive surveys of HL - mutation landscape reveals known and novel genes
Recurrently Mutated Genes - *TNFAIP3*

- 7 mutations in 7 patients (~23%)
- All mutations are frame shift, splice site, or nonsense mutations
  - All potential truncating mutations
- *TNFAIP3* is part of the NF-κB pathway
  - Constitutive expression of NF-κB
  - Codes for A20 - known tumor suppressor non-Hodgkin lymphoma
- Known to be recurrently mutated in cHL
• 10 mutations in 7 patients (~23%)
• Little is known about the function of this gene
  • Mutated in DLBCL, HL (Reichel et al.), and CLL
• In CLL (Kasar et al., 2015, Nature Communications) \textit{IGLL5} was the most frequently mutated gene
  • Showed a pattern of AID off target activity
Recurrently Mutated Genes - BTG1

- 6 mutations in 3 patients (~10%)
- B-cell translocation gene 1, belongs to the BTG/TOB family of anti-proliferative genes that regulate cell growth and differentiation
- Known to be mutated in ALL, DLBCL, other lymphomas and leukemias
- Functional analyses show that in interact with transcription factors (transcription co-regulator)
Haloplex method provides an ultra-sensitive method for orthogonal validation

- Genomic DNA is fragmented using restriction enzymes
- Custom haloplex probe library is hybridized to target DNA
  - Includes Unique Molecular Index (UMI)
- Capture of target DNA is achieved by biotinylation of probe DNA using magnetic beads.
- PCR is used to amplify targeted fragments
- Fragments are sequenced

Kock, et al., Journal of Molecular Genetics, 2015
Incorporation of unique molecular bar codes allows error correction and ultra-sensitive variant calling

- **Molecular barcodes (UMI)** are degenerate oligonucleotide sequences (10-16bp) attached to individual DNA molecules

- UMIs allow for the correction of sequencer and PCR errors in high coverage NGS data
  - Effectively reduces error rate and increases sensitivity for low VAF variants

http://archerdx.com/immunoverse
Results indicate extremely high coverage and very good signal to noise ratio (tumor vs normal)

- We designed a Haloplex reagent with 1875 sites and 25 genes
- So far we have piloted 8 samples (4 tumors, 4 normals)

Most tumor VAFs 1-5%

- 200,000-400,000X Raw coverage
- 5,000-10,000X Barcode coverage
- Sensitivity down to 0.01%
Comparison to WES results reveals >95% validation rate, VAFs well-correlated

\[ R^2 = 0.61 \]
Rare and challenging tumor types

Metastatic, relapsed, resistant tumors

Comprehensive omic profiling

Ethnic diversity

Characterization of models (mouse, PDX)

Spatial/Temoral heterogeneity

High sequence depth and breadth

Deep clinical annotations

Non-coding space

Increased sample sizes (power)
In some cancers specific driver genes show different frequencies across ethnic groups

- In non-small cell lung cancer mutations in \textit{EGFR} show differences in frequencies across global human populations
- This has of \textbf{clinical significance} because specific \textit{EGFR} mutations are associated with response to specific cancer treatments

Kohno et al. Translational Lung Cancer Research 2015
TCGA is underpowered to detect driver genes, at even 10% mutation rate, for any but those of European descent.

Spratt et al. JAMA Oncol 2016
Rare and challenging tumor types

Metastatic, relapsed, resistant tumors

Comprehensive omic profiling

Ethnic diversity

Characterization of models (mouse, PDX)

Spatial/Temporal heterogeneity

High sequence depth and breadth

Deep clinical annotations

Non-coding space

Increased sample sizes (power)
100,000 samples to achieve 80% power to detect rare variants (MAF 0.05%), even in 95% penetrant

Similar power limitations apply to clinical associations

Rare and challenging tumor types
Metastatic, relapsed, resistant tumors
Comprehensive omic profiling
Ethnic diversity
Characterization of models (mouse, PDX)

Spatial/Temoral heterogeneity
High sequence depth and breadth
Deep clinical annotations
Non-coding space
Increased sample sizes (power)
New drug target and prognostic markers waiting to be discovered with larger, disease-focused efforts

Study underway of 2000+ additional ER+ breast cancers, consistently treated, long-term follow-up, deep targeted sequencing

Activating missense

Loss-of-function truncating
Sequencing larger cohorts has also allowed identification of rare but clinically significant alterations. Bose. et al, 2012 (Cancer Discovery) showed that these predict response to HER2 inhibitors similar to HER2amp.

ERBB2 kinase domain mutations in ~1-2% samples

*Known Activating*

*Bose. et al, 2012 (Cancer Discovery) showed that these predict response to HER2 inhibitors similar to HER2amp*
Rare but significant mutations have important prognostic implications

**DDR1 - BCSS**

- WT (605)
- MT (17)
- \( p = 0.000137 \)
- \( HR = 3.41 \)

**NF1 - RFS**

- WT (617)
- MT (8)
- \( p = 0.00282 \)
- \( HR = 2.98 \)

Novel splice site mutation hotspot discovered in CBFB

- At least 7 genomic alterations (SNV/INS/DEL) result in 14 exon2 splice site mutations
- Truncation and likely loss-of-function mutation predicted in 3.6% of samples

Beta subunit of a TF – master regulator of hematopoiesis and osteogenesis genes (RUNX1/RUNX2) – CBFB fusions -> AML
TCGA exomes had very low coverage of exons 1 and 2
Low coverage possibly explained by higher GC content

**GC content: CBFB exons**

**hg19 GC content: all CDS exons**
Rare and challenging tumor types

Metastatic, relapsed, resistant tumors

Comprehensive omic profiling

Ethnic diversity

Characterization of models (mouse, PDX)

Spatial/Temporal heterogeneity

High sequence depth and breadth

Deep clinical annotations

Non-coding space

Increased sample sizes (power)
Model #1: STAT1 loss is implicated in ER+ breast cancer

- The transcription factor STAT1 functions as a tumor suppressor in mammary gland epithelial cells.
- Selective loss of STAT1 expression is associated with a significant percentage of human estrogen receptor alpha-positive (ERα+) luminal breast tumors.
- Stat1-/- mice are predisposed to mammary adenocarcinoma development.
- These tumors develop in a manner similar to ERα+/PgR+ invasive ductal carcinoma in humans.
- These tumors also display transcript expression profiles that cluster more closely with human ERα+ luminal breast cancers than other mammary tumor models.
Whole genome sequencing (WGS) used to identify co-operating genomic events in Stat1-/- mice

- Goal: Identify genomic event(s) in Stat1-/- mice by WGS that fully transform the phenotype of mammary gland epithelial cells into cancer cells
We characterized the genetic changes in 106 tissues from 49 Stat1−/− and wild-type mice by whole genome and targeted sequencing.

Discovery Samples
33 mice / 52 samples

- 10 Wild type mammary glands
- 5 Ovarian hormone independent tumors
- 3 Mammary tumor-derived cell lines
- 14 Primary mammary tumors
- 5 Tumor-free mammary glands

Extension Samples
16 mice / 54 samples

- 9 DCIS (FFPE)
- 1 Ovarian hormone independent tumor
- 25 Tumor-free organs
- 10 Primary mammary tumors
- * Second tumor/DCIS
- +/+ Stat1+/+ mouse
- +/- Stat1−/− mouse
- Mouse with normal tail sample
- Mouse without normal tail sample

Griffith et al. Cell Reports.
~150 recurrently mutated genes (SNVs and Indels) identified

<table>
<thead>
<tr>
<th>Gene</th>
<th>#</th>
<th>Gene</th>
<th>#</th>
<th>Gene</th>
<th>#</th>
<th>Gene</th>
<th>#</th>
<th>Gene</th>
<th>#</th>
<th>Gene</th>
<th>#</th>
<th>Gene</th>
<th>#</th>
<th>Gene</th>
<th>#</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prlr</td>
<td>17</td>
<td>Sbsn</td>
<td>3</td>
<td>Brca1</td>
<td>2</td>
<td>Gdf9</td>
<td>2</td>
<td>Kcnj14</td>
<td>2</td>
<td>Nell1</td>
<td>2</td>
<td>Slc1a1</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Arid4a</td>
<td>6</td>
<td>Slc39a12</td>
<td>3</td>
<td>Cacna1e</td>
<td>2</td>
<td>Gm10710</td>
<td>2</td>
<td>Kcnn3</td>
<td>2</td>
<td>Pappa</td>
<td>2</td>
<td>Slc47a1</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gm10750</td>
<td>4</td>
<td>Taar7e</td>
<td>3</td>
<td>Cacna1g</td>
<td>2</td>
<td>Gm12025</td>
<td>2</td>
<td>Kdm6a</td>
<td>2</td>
<td>Pappa2</td>
<td>2</td>
<td>Slit3</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gm5277</td>
<td>4</td>
<td>Tgoln1</td>
<td>3</td>
<td>Ccdc36</td>
<td>2</td>
<td>Gm12696</td>
<td>2</td>
<td>Kif3b</td>
<td>2</td>
<td>Phlpp1</td>
<td>2</td>
<td>Sos2</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mll3</td>
<td>4</td>
<td>Tmem181b-ps</td>
<td>3</td>
<td>Cd209c</td>
<td>2</td>
<td>Gm14373</td>
<td>2</td>
<td>Klhdc8b</td>
<td>2</td>
<td>Pik3r2</td>
<td>2</td>
<td>Spag17</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BC006779</td>
<td>3</td>
<td>Trp53</td>
<td>3</td>
<td>Cdh15</td>
<td>2</td>
<td>Gm15283</td>
<td>2</td>
<td>Lama5</td>
<td>2</td>
<td>Pclb2</td>
<td>2</td>
<td>Src</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dnaic1</td>
<td>3</td>
<td>Vmn2r90</td>
<td>3</td>
<td>Chd2</td>
<td>2</td>
<td>Gm1614</td>
<td>2</td>
<td>Lrp1b</td>
<td>2</td>
<td>Plcd1</td>
<td>2</td>
<td>Stk11</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Esrrg</td>
<td>3</td>
<td>Zfp335</td>
<td>3</td>
<td>Clec16a</td>
<td>2</td>
<td>Gm16880</td>
<td>2</td>
<td>Lrrc24</td>
<td>2</td>
<td>Plexb1</td>
<td>2</td>
<td>Syne2</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fbxl7</td>
<td>3</td>
<td>March11</td>
<td>2</td>
<td>Crb1</td>
<td>2</td>
<td>Gm16950</td>
<td>2</td>
<td>Mapk7</td>
<td>2</td>
<td>Poln</td>
<td>2</td>
<td>Tiam1</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Galt5</td>
<td>3</td>
<td>Abca15</td>
<td>2</td>
<td>Csmd1</td>
<td>2</td>
<td>Gm2058</td>
<td>2</td>
<td>Mapt</td>
<td>2</td>
<td>Polr2a</td>
<td>2</td>
<td>Trp4m</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gatsl3</td>
<td>3</td>
<td>Abca2</td>
<td>2</td>
<td>Cyp3a44</td>
<td>2</td>
<td>Gm2940</td>
<td>2</td>
<td>Mast2</td>
<td>2</td>
<td>Ppp1r13b</td>
<td>2</td>
<td>Ttc6</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gfap</td>
<td>3</td>
<td>Acpl2</td>
<td>2</td>
<td>Daam1</td>
<td>2</td>
<td>Gm3513</td>
<td>2</td>
<td>Mgate</td>
<td>2</td>
<td>Prkg1</td>
<td>2</td>
<td>Ttn</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gm10601</td>
<td>3</td>
<td>Actn4</td>
<td>2</td>
<td>Daglb</td>
<td>2</td>
<td>Gm5151</td>
<td>2</td>
<td>Mogs</td>
<td>2</td>
<td>Ptprb</td>
<td>2</td>
<td>Ush2a</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gm11867</td>
<td>3</td>
<td>Adamts14</td>
<td>2</td>
<td>Dedd2</td>
<td>2</td>
<td>Gm6706</td>
<td>2</td>
<td>Mrc1</td>
<td>2</td>
<td>Ptpfr</td>
<td>2</td>
<td>Vmn2r84</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gm16372</td>
<td>3</td>
<td>Akap12</td>
<td>2</td>
<td>Dennd5b</td>
<td>2</td>
<td>Gm7353</td>
<td>2</td>
<td>Mrc2</td>
<td>2</td>
<td>Qrich2</td>
<td>2</td>
<td>Vps13c</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gm16957</td>
<td>3</td>
<td>Ank1</td>
<td>2</td>
<td>Dhx57</td>
<td>2</td>
<td>Grf1</td>
<td>2</td>
<td>Mrpl15</td>
<td>2</td>
<td>Rad50</td>
<td>2</td>
<td>Zfp523</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gm16972</td>
<td>3</td>
<td>Ankrd36</td>
<td>2</td>
<td>Dnahc2</td>
<td>2</td>
<td>Gse1</td>
<td>2</td>
<td>Mst1r</td>
<td>2</td>
<td>Rangap1</td>
<td>2</td>
<td>Zfp119a</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gm6369</td>
<td>3</td>
<td>Anp32b</td>
<td>2</td>
<td>Dnm3os</td>
<td>2</td>
<td>Hcfc1</td>
<td>2</td>
<td>Musk</td>
<td>2</td>
<td>Rbm15</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gne</td>
<td>3</td>
<td>Apobec4</td>
<td>2</td>
<td>Dusp8</td>
<td>2</td>
<td>Hps5</td>
<td>2</td>
<td>Myh9</td>
<td>2</td>
<td>Reep6</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gnl1</td>
<td>3</td>
<td>Arid1b</td>
<td>2</td>
<td>Fancm</td>
<td>2</td>
<td>Htr1f</td>
<td>2</td>
<td>Myl4</td>
<td>2</td>
<td>Rfc2</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Krt15</td>
<td>3</td>
<td>Atad2</td>
<td>2</td>
<td>Fnip2</td>
<td>2</td>
<td>Igsf3</td>
<td>2</td>
<td>Myo9b</td>
<td>2</td>
<td>Rgl1</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rbbp6</td>
<td>3</td>
<td>Bap1</td>
<td>2</td>
<td>Fsip2</td>
<td>2</td>
<td>Ip6k2</td>
<td>2</td>
<td>Ncapd2</td>
<td>2</td>
<td>Slc12a4</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Many known breast cancer players (P53, ARID4A/1B, MLL3, BRCA1, MAPK, SRC, RAD50)
Manual review of WGS data confirmed high-quality somatic variants and identified several missed by callers.

4/5 remaining tumors have WGS evidence of similar Prlr mutations that were not called. **Total frequency of 21/22 (95.5%).** Only TAC246 has no evidence of mutation in Prlr.
Ultimately, using WGS, Sanger and targeted MiSeq we confirmed Prlr mutations in 32/32 tumors, 7/9 DCIS, and 0/40 controls examined to date.
Prlr SNVs/indels were all truncating mutations

<table>
<thead>
<tr>
<th>chr</th>
<th>start</th>
<th>stop</th>
<th>Ref/var</th>
<th>type</th>
<th>Trv_type</th>
<th>aa change</th>
<th>Het/Hom</th>
<th>sample</th>
</tr>
</thead>
<tbody>
<tr>
<td>15</td>
<td>10258182</td>
<td>10258182</td>
<td>G/-</td>
<td>DEL</td>
<td>frame_shift_del</td>
<td>G330fs</td>
<td>Het</td>
<td>B3R1.5-B3R1.5_tumor</td>
</tr>
<tr>
<td>15</td>
<td>10258180</td>
<td>10258180</td>
<td>C/-</td>
<td>DEL</td>
<td>frame_shift_del</td>
<td>P329fs</td>
<td>Het</td>
<td>B3R1R2L1-B3R1R2L1_tumor</td>
</tr>
<tr>
<td>15</td>
<td>10258151</td>
<td>10258151</td>
<td>G/-</td>
<td>DEL</td>
<td>frame_shift_del</td>
<td>L320fs</td>
<td>Hom?</td>
<td>OVX3_L2-OVX3_L2_tumor</td>
</tr>
<tr>
<td>15</td>
<td>10258151</td>
<td>10258151</td>
<td>G/-</td>
<td>DEL</td>
<td>frame_shift_del</td>
<td>L320fs</td>
<td>Het</td>
<td>OVX6_R2-OVX6_R2_tumor</td>
</tr>
<tr>
<td>15</td>
<td>10258147</td>
<td>10258147</td>
<td>G/-</td>
<td>DEL</td>
<td>frame_shift_del</td>
<td>E318fs</td>
<td>Het</td>
<td>SSM1-SSM1</td>
</tr>
<tr>
<td>15</td>
<td>10258139</td>
<td>10258140</td>
<td>C/-</td>
<td>DEL</td>
<td>frame_shift_del</td>
<td>E319fs</td>
<td>Het</td>
<td>SSM2-SSM2</td>
</tr>
<tr>
<td>15</td>
<td>10258195</td>
<td>10258195</td>
<td>A/-</td>
<td>DEL</td>
<td>frame_shift_del</td>
<td>K334fs</td>
<td>Het</td>
<td>SSM3-SSM3</td>
</tr>
<tr>
<td>15</td>
<td>10258180</td>
<td>10258180</td>
<td>C/-</td>
<td>DEL</td>
<td>frame_shift_del</td>
<td>P329fs</td>
<td>Het</td>
<td>TAC171-TAC171_tumor</td>
</tr>
<tr>
<td>15</td>
<td>10258147</td>
<td>10258147</td>
<td>G/-</td>
<td>DEL</td>
<td>frame_shift_del</td>
<td>E318fs</td>
<td>Het</td>
<td>TAC183-TAC183_tumor</td>
</tr>
<tr>
<td>15</td>
<td>10258147</td>
<td>10258148</td>
<td>G/-</td>
<td>DEL</td>
<td>frame_shift_del</td>
<td>E318fs</td>
<td>Het</td>
<td>TAC186-TAC186_tumor</td>
</tr>
<tr>
<td>15</td>
<td>10258153</td>
<td>10258168</td>
<td>G/-</td>
<td>DEL</td>
<td>frame_shift_del</td>
<td>L320fs</td>
<td>Het</td>
<td>TAC247-TAC247_tumor</td>
</tr>
<tr>
<td>15</td>
<td>10258182</td>
<td>10258182</td>
<td>G/-</td>
<td>DEL</td>
<td>frame_shift_del</td>
<td>G330fs</td>
<td>Het</td>
<td>TAC266-TAC266_tumor</td>
</tr>
<tr>
<td>15</td>
<td>10258151</td>
<td>10258154</td>
<td>G/GCT/-</td>
<td>DEL</td>
<td>frame_shift_del</td>
<td>R319fs</td>
<td>Het</td>
<td>TAC268-TAC268_tumor</td>
</tr>
<tr>
<td>15</td>
<td>10258147</td>
<td>10258147</td>
<td>G/-</td>
<td>DEL</td>
<td>frame_shift_del</td>
<td>E318fs</td>
<td>Het</td>
<td>TAC269-TAC269_tumor</td>
</tr>
<tr>
<td>15</td>
<td>10258180</td>
<td>10258180</td>
<td>C/-</td>
<td>DEL</td>
<td>frame_shift_del</td>
<td>P329fs</td>
<td>Het</td>
<td>TAC270-TAC270_L. cerv</td>
</tr>
<tr>
<td>15</td>
<td>10258147</td>
<td>10258148</td>
<td>G/-</td>
<td>DEL</td>
<td>frame_shift_del</td>
<td>E318fs</td>
<td>Het</td>
<td>TAC270-TAC270_L. ing</td>
</tr>
<tr>
<td>15</td>
<td>10258179</td>
<td>10258179</td>
<td>T/A</td>
<td>SNP</td>
<td>nonsense</td>
<td>Y328*</td>
<td>Het</td>
<td>TAC271-TAC271_tumor</td>
</tr>
<tr>
<td>15</td>
<td>10258146</td>
<td>10258168</td>
<td>CGAGCGGCTAATGCCATCCCATT/-</td>
<td>DEL</td>
<td>frame_shift_del</td>
<td>E318fs</td>
<td>Het</td>
<td>TAC272-TAC272_tumor</td>
</tr>
<tr>
<td>15</td>
<td>10258195</td>
<td>10258195</td>
<td>A/-</td>
<td>DEL</td>
<td>frame_shift_del</td>
<td>K334fs</td>
<td>Het</td>
<td>TAC273-TAC273_L. tho. tumor</td>
</tr>
<tr>
<td>15</td>
<td>10258146</td>
<td>10258149</td>
<td>C/GAG/-</td>
<td>DEL</td>
<td>frame_shift_del</td>
<td>E318fs</td>
<td>Het</td>
<td>TAC273-TAC273_R. ing. tumor</td>
</tr>
<tr>
<td>15</td>
<td>10258147</td>
<td>10258147</td>
<td>G/T</td>
<td>SNP</td>
<td>nonsense</td>
<td>E318*</td>
<td>Het</td>
<td>TAC274-TAC274_R. ing. tumor</td>
</tr>
</tbody>
</table>

10 WT and 5 tumor-free STAT1-/- mammary glands were reviewed and no evidence of PRLR mutations found. None in matched tails either.
Truncating Prlr mutations highly clustered in an ~85bp (28aa) window

Discovery (22 tumors)

- Truncating mutations cluster tightly around amino acid position: 316-343

Extension (10 tumors; 9 DCIS)

- Truncating mutations cluster tightly around amino acid position: 316-343
PRLR truncations predicted to produce a protein between intermediate and short natural isoforms in length

- Jak2 site preserved
- Downstream Tyrosine phosphorylation sites lost

Truncation window
Expression of FL+T Prlr resulted in activation of the oncogenic Stat3 and Stat5, and promoted anchorage-independent growth

Ruby Chan and Bob Schreiber

- Hypothesis: heterodimers of FL and T Prlr cause constitutive Prlr activation and tumorigenic phenotype of Stat1-/- mammary epithelial cells
- pStat3 and pStat5 detected in MEFs expressing FL/T heterodimer but neither for FL or T homodimers
- MEFs with FL/T also show significantly increased colony formation in soft agar anchorage-independent growth assay
Nude mice transplanted with Stat1-/- MEFs expressing FL/T Prlr + Jak2 develop tumors much more readily than FL/FL, T/T, or Jak2 alone.
Mutational hotspot analysis of PRLR in human TCGA breast cancer exome data identifies rare truncating mutations
Model #2: Prolactin expression is implicated in ER+ breast cancer

• Epidemiologic data link prolactin (PRL) exposure to development of ER+ metastatic breast cancer

• Schuler lab has developed a murine model which overexpresses PRL in mammary epithelial cells (NRL-PRL) leading to:
  • Mammary epithelial cells with increased progenitor and stem cell activity
  • Development of spontaneous metastatic ER+ carcinomas with characteristics of luminal B human cancers
We characterized genetic changes in 41 tissues from NRL-PRL mice by whole genome and targeted sequencing.
Extension sequencing identified Ras alterations in $\frac{23}{29}$ (79%) of tumors examined.

<table>
<thead>
<tr>
<th>Mutation Type</th>
<th>Sample Source</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Cell Line</td>
<td>Tumor</td>
</tr>
<tr>
<td>WT</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>Kras Amplification</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Kras G12</td>
<td>1*</td>
<td>12</td>
</tr>
<tr>
<td></td>
<td>G12C</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>G12D</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>G12V</td>
<td>1*</td>
</tr>
<tr>
<td>Kras G13</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>G13A</td>
<td>1</td>
</tr>
<tr>
<td>Kras Q61</td>
<td>1*</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>Q61H</td>
<td>1*</td>
</tr>
<tr>
<td></td>
<td>Q61L</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Q61R</td>
<td>3</td>
</tr>
<tr>
<td>Nras A59</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>A59D</td>
<td>1</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>2</td>
<td>27</td>
</tr>
</tbody>
</table>

*Two cell lines independently derived from 1 tumor*
KRAS activating mutations exist but are rare in human breast ER+ breast cancers

- Investigating Kras activation expression signature in human ER+ breast cancer RNA data
Summary

• TCGA/ICGC have transformed understanding of cancer genomics
• These efforts were just the beginning
• Advanced disease states remain to be characterized
• More comprehensive sequencing will identify missed drivers, tumor architecture, non-coding regions, etc
• Rare/challenging tumor type remain understudied
• More diverse study populations needed
• Larger, well-annotated cohorts will allow omic events to be associated with clinical outcomes
• Tumor models must be genomically characterized
Tumour genome sequencing mission accomplished?
Tumour genome sequencing not mission accomplished
Griffith lab members did all the work!

Malachi Griffith
Obi Griffith
Erica Barnell
Katie Campbell
Kaitlin Clark
Adam Coffman
Kelsy Cotto
Arpad Danos
Yang Yang Feng
Felicia Gomez
Jasreet Hundal
Susanna Kiwalla
Kilannin Krysiak
Lynzey Kujan
Josh McMichael
Shahil Pema
Cody Ramirez
Megan Richters
Peter Ronning
Lana Sheta
Zachary Skidmore
Nick Spies
Lee Trani
Alex Wagner
Jason Walker
Alex Wollam
Huiming Xia
Nearly all projects are built on collaboration with physician scientists.
Thank you!