An Introduction

- Post-doc with Tim Ley
  - Acute Myeloid Leukemia
  - transcriptome, genome (copy number variants, whole genome/exome sequencing)

- Co-direct Barnes Molecular Diagnostics Lab

- My lab
  - Epigenomics of B cell lymphoma
  - Development of targeted epigenetic modifiers
Defining and Targeting the Malignant Epigenome in Non-Hodgkin Lymphoma

Jacqueline E. Payton, MD, PhD
Dept. of Pathology & Immunology
What do we mean by “cancer genetics”?

- the role of the genome in cancer
  - susceptibility
  - genesis
  - survival
  - progression
  - resistance
  - recurrence
How can the genome play a role in cancer?

• mutations
• “changes” in the genome sequence
• what are these changes?
• how do they occur?
• how do they influence the process that is cancer?
• what (if anything) can we do about them?
mutations

- sequence difference compared to the “reference”

- Human Genome project = reference genome
  - few individuals
  - not ethnically diverse
Human Genome Project

Completed 2001
$1 billion
10 years
Human Genomes
Output per instrument run

Platforms

Projects and publications

Mardis, EM Nature Feb 2011
Impact on (non-cancer) disease discovery

- common variants in neurodevelopmental syndromes
  - autism, schizophrenia, ADHD, bipolar disorder, depression
  - Lancet Feb 2013

- rare disorders
  - severe neurodevelopmental delays and immunodeficiencies
  - small families or single individuals
  - case of severe, intractable IBD in a child
Benefits

• Diagnostic testing
• Pre-symptomatic testing
• Identification of new loci for research
Cancer Genomes

- Breast
- Colon
- Lung
- Prostate
- Melanoma
- Pancreatic
- Renal
- Brain
- Leukemia
- Lymphoma
Benefits

- Diagnostic testing
  - differentiate subtypes

- Prognostic testing
  - different subtypes or presence of mutation impact outcome

- Theranostic testing
  - some mutant proteins are drug targets

- Identification of new loci for research
Mutated genes and their drug targets

- A detected mutation may suggest a drug:
  - ALK (fusion): Crizotinib
  - BRAF: Vemurafenib
  - EGFR: Gefitinib
  - JAK2: Ruxolitinib
  - KIT: Imatinib
  - ...

- A detected mutation may suggest not to use a drug:
  - KRAS, PTEN, PIK3CA, ...

- A detected mutation may suggest clinical trials:
  - FLT3, PIK3CA, …
Activation of epidermal growth factor signaling pathways

KRAS mutant

CCND1, gene encoding cyclin D1; CDKN1A, gene encoding p21; JAK, Janus kinase; TGFα, transforming growth factor-α.

“mutations”

- polymorphisms/germline variants
  - susceptibility
  - cancer syndromes

- somatic variants
  - sporadic cancers
  - therapy-related/environmental cancers
Inherited cancer syndromes

- Lynch syndrome (colon, endometrial)
  - mismatch repair genes
- Li-Fraumeni (sarcoma, breast, glioma, leukemia)
  - TP53
- Neurofibromatoses
  - NF1
- Inherited breast cancer
Inherited breast cancer

- Female carriers: >80% lifetime risk of breast cancer and 20-65% risk of ovarian cancer
- BRCA1 locus = 100kb, ORF 5.5kb
- BRCA2 locus = 84kb, ORF 10kb
- All mutation types reported (truncating, frameshift, missense, duplications, deletions, rearrangements)
- Clinical testing a challenge
BRCA1 and 2 clinical testing
Gene Patents

They're our breast cancer genes—we identified them.

It's kind of you to let us have the disease for free.
Gene patents

- > 40,000 patents on DNA molecules
- many patents filed on 15mer sequences
- court rulings have supported IP rights on these
- average gene matches 364 other genes as 15mers
- 41% of genes covered by at least one patent

Figure 1. Increasing number of DNA-based patents. We searched for any US patent that has DNA sequences present in the claims since 1983 and observed a continual increase of DNA patents, with the greatest increase around the time of the completion of the human genome draft sequence in 2001.

Rosenfeld Genome Medicine 2013
Diagnostic gene patents are broader

- written to find any known or unknown variation of a gene
- \( > (2N) - 1 \) possible combinations of mutations for a gene
Claim #1. An isolated DNA coding for a BRCA1 polypeptide, said polypeptide having the amino acid sequence set forth in SEQ ID NO:2 (the BRCA1 cDNA).

Claim #2. The isolated DNA of claim 1, wherein said DNA has the nucleotide sequence set forth in SEQ ID NO:1 (the BRCA1 gene).

Claim #5. An isolated DNA having at least 15 nucleotides of the DNA of claim 1

Claim #6. An isolated DNA having at least 15 nucleotides of the DNA of claim 2.

http://www.google.com/patents/US5747282
BRCA1 and 2 clinical testing can only be performed by Myriad
Association for Molecular Pathology v. Myriad Genetics

ACLU Challenges Patents On Breast Cancer Genes

This case involves a challenge to patents on two human genes known as BRCA1 and BRCA2. Mutations in those genes correlate with an increased risk of breast and ovarian cancer. On behalf of a large coalition of research scientists, patients, and patient advocacy groups, we have argued that human genes cannot be patented because they are classic products of nature.

On May 12, 2009, the ACLU and the Public Patent Foundation (PUBPAT) filed a lawsuit charging that patents on two human genes associated with breast and ovarian cancer are unconstitutional and invalid. The suit charges that the patents stifle diagnostic testing and research that could lead to cures and that they limit women’s options regarding their medical care.

Watch the Video >>

SIGN A MESSAGE OF SUPPORT >>

Join the thousands of people who have voiced support for the BRCA gene patents challenge.
BRCA1 patent challenge

- In 2009 the ACLU + many other groups and individual patients challenged the BRCA patent


- researchers, genetic counselors, women patients, cancer survivors, breast cancer and women's health groups, and scientific associations representing 150,000 geneticists, pathologists, and laboratory professionals
BRCA1 patent challenge

• in 2010, a US District Court ruled that isolated DNA is “not patentable subject matter”

• in 2011, US Federal Court of Appeals overruled (2 to 1) stating that isolated DNA is “markedly different” from native genomic DNA and therefore patentable

Figure 2. Court-proposed molecular points of distinction that allow claims on isolated DNA sequences. On the basis of two molecular changes (small circles), specifically to a single phosphate and one hydroxyl group, the Federal Circuit suggested that a new DNA fragment (right) is patentable subject matter. However, such molecular structures rarely appear in gene patents, and they are not present in the BRCA patents.
On November 30, 2012, the Supreme Court agreed to hear argument during the current session on the patentability of human genes.
Somatic cancers
**Constitutional/Germline vs. Somatic Mutations**

**Constitutional/Germline Genomes**

*Unaffected "Carrier" Father* and *Unaffected "Carrier" Mother*:
- Constitutional/germline genomes present in every cell of the body, including cancer cells.

**Somatic Mutations**

- Mutation present only in cancer cells (cancer genome).

[Diagram showing genetic inheritance patterns between parents and offspring, illustrating the difference between constitutional/germline and somatic mutations.]
DNA sequencing of a cytogenetically normal acute myeloid leukaemia genome

Acute myeloid leukaemia is a highly malignant haematopoietic tumour that affects about 13,000 adults in the United States each year. The treatment of this disease has changed little in the past two decades, because most of the genetic events that initiate the disease remain undiscovered. Whole-genome sequencing is now possible at a reasonable cost and timeframe to use this approach for the unbiased discovery of tumour-specific somatic mutations that alter the protein-coding genes. Here we present the results obtained from sequencing a typical acute myeloid leukaemia genome, and its matched normal counterpart obtained from the same patient’s skin. We discovered ten genes with acquired mutations; two were previously described mutations that are thought to contribute to tumour progression, and eight were new mutations present in virtually all tumour cells at presentation and relapse, the function of which is not yet known. Our study establishes whole-genome sequencing as an unbiased method for discovering cancer-initiating mutations in previously unidentified genes that may respond to targeted therapies.
Recurrent mutations in the U2AF1 splicing factor in myelodysplastic syndromes

Timothy A Graubert1,3,9, Dong Shen4,9, Li Ding4,5,9, Theresa Okeyo-Owuo1, Cara L Lunn1, Jin Shao1, Kilannin Krysiak1, Christopher C Harris4, Daniel C Koboldt1, David E Larson1, Michael D McLellan4, David J Dooling4, Rachel M Abbott4, Robert S Fulton4, Heather Schmidt4, Joelle Kalicki-Weizer4, Michelle O’Laughlin4, Marcus Grillot1, Jack Baty6, Sharon Heath1, John L Frater3, Talat Nasim7,8, Daniel C Link1,2, Michael H Tomasson1,2, Peter Westervelt1,2, John F DiPersio1,2, Elaine R Mardis2,4,5, Timothy J Ley1,2,4, Richard K Wilson2,4,5 & Matthew J Walker1,2,5

Identification of a Novel TP53 Cancer Susceptibility Mutation Through Whole-Genome Sequencing of a Patient With Therapy-Related AML

Daniel C Link, MD
Laura C. Schattell, MD, PhD
Dong Shen, MD, PhD
Feiling Wang, MD
Matthew J Walker, MD
Shashikant Kulkarni, MD
Jacqueline E. Payton, MD, PhD

Context The identification of patients with inherited cancer susceptibility syndromes facilitates early diagnosis, prevention, and treatment. However, in many cases of suspected cancer susceptibility, the family history is unclear and genetic testing of common cancer susceptibility genes is unwieldy.

Objective To apply whole-genome sequencing to a patient with no significant family history of cancer but with suspected increased cancer susceptibility because of multiple primary tumors to identify rare or novel germline variants in cancer susceptibility genes.

Design, Setting, and Participants Skin normal and bone marrow (leukemia) DNA were obtained from a patient with early-stage breast and ovarian cancer (positive for BRCA1 and BRCA2 mutations) and therapy-related acute myeloid leukemia (AML) and analyzed with the following: whole-genome sequencing using paired-end reads, single-nucleotide polymorphism (SNP) genotyping, RNA expression profiling, and spectral karyotyping.

Main Outcome Measures Structural variants, copy number alterations, single-nucleotide variants, and small insertions and deletions (indels) were detected and validated using the described platforms.

Results Whole-genome sequencing revealed novel, heterogeneous 3-kilobase deletion removing exons 7-9 of TP53 in the patient’s normal skin DNA, which was homozygous in the leukemia DNA as a result of uniparental disomy. In addition, a total of 26 validated somatic single-nucleotide variations or indels in coding genes, 8 somatic structural variants, and 12 somatic copy number alterations were detected in the patient’s leukemia genome.

Conclusion Whole-genome sequencing can identify novel, cryptic variants in cancer susceptibility genes in addition to providing unbiased information on the spectrum of mutations in a cancer genome.

JAMA. 2011;295(15):1689-1703

LETTER

Clonal evolution in relapsed acute myeloid leukaemia revealed by whole-genome sequencing

Li Ding1, Timothy J. Ley4,5, David E Larson1, Christopher A Miller1, Daniel C Koboldt1, John S Welch6, Julie K Ritchey4, Margaret A. Young1, Tamara Lamprecht1, Michael D McLellan4, Joshua F. McMichael4, Michael H Tomasson4, John W Walli2,2, Charles Lu1, Dong Shen1, Christopher C Harris1, David J Dooling1, Robert S Fulton1,2, Lucinda L Fulton1,2, Ken Chen1,2, Heather Schmidt4, Joelle Kalicki-Weizer4, Vincent J. Magrini1,2, Lisa Cook1, David M. Grimsby6, Tamimi L. Vickery1, Michael C. Wendel1,2, Sharon Heath1, Mark A. Watson5, Daniel C Link1,4, Michael H Tomasson4, William D. Shannon4, Jacqueline E Payton4, Shashikant Kulkarni2,4,5, Peter Westervelt1,2, Matthew J Walker4,3, Timothy A Graubert1,4, Elaine R Mardis1,2,4,5, Richard K Wilson2,4,5 & John F DiPersio1,2,4

Recurrent mutations in the U2AF1 splicing factor in myelodysplastic syndromes

Timothy A Graubert1,3,9, Dong Shen4,9, Li Ding4,5,9, Theresa Okeyo-Owuo1, Cara L Lunn1, Jin Shao1, Kilannin Krysiak1, Christopher C Harris4, Daniel C Koboldt1, David E Larson1, Michael D McLellan4, David J Dooling4, Rachel M Abbott4, Robert S Fulton4, Heather Schmidt4, Joelle Kalicki-Weizer4, Michelle O’Laughlin4, Marcus Grillot1, Jack Baty6, Sharon Heath1, John L Frater3, Talat Nasim7,8, Daniel C Link1,2, Michael H Tomasson1,2, Peter Westervelt1,2, John F DiPersio1,2, Elaine R Mardis2,4,5, Timothy J Ley1,2,4, Richard K Wilson2,4,5 & Matthew J Walker1,2,5
AML clonality
AML evolution

Clonal Evolution from MDS to sAML

Cell Type
- Green: Normal
- Gray: MDS or sAML

Mutations
- Germline
- Cluster 1 SNVs
- Cluster 2 SNVs
- Cluster 3 SNVs
- Cluster 4 SNVs
- Cluster 5 SNVs

Percentage of mutations:
- 26%
- 52%
- 22%
- 6%
- 4%
- 33%
- 10%
- 14%
- 33%
The Origin and Evolution of Mutations in Acute Myeloid Leukemia

M1 initiating mutations
(NPM1, DNMT3A, IDH1, TET2 & others)

M3 initiating mutation
(PML-RARA)

M1 initiating mutations
(cooperating mutations
(FLT3 and others)

HSPC

X: age-dependent passenger mutations pre-existing in HSPC
Y: passenger mutations gained between initiating and cooperating mutations
Z: passenger mutations gained during progression to subclones

subclone 1

founding AML clone

subclone 2
Benefits of cancer mutation discovery

- Diagnostic testing
  - differentiate subtypes

- Prognostic testing
  - different subtypes or presence of mutation impact outcome

- Theranostic testing
  - some mutant proteins are drug targets

- Identification of new loci for research
Comprehensive Cancer Gene Set

Washington University Genomics and Pathology Services
Clinical Next-Generation Sequencing (NGS)

Our *comprehensive gene set tests* allow for cost effective and efficient analysis of clinically actionable biomarkers spanning a wide range of diseases. The analysis of the *disease-specific gene targets*, when considered with other pathology findings, aids in the stratification of disease and elucidates *optimal treatment strategies*.

We currently offer analysis of the following genes:

<table>
<thead>
<tr>
<th>ALK*</th>
<th>BRAF</th>
<th>CHIC2</th>
<th>CSF1R</th>
<th>CTNNB1</th>
<th>DNMT3A</th>
<th>EGFR**</th>
<th>FLT3</th>
<th>IDH1</th>
</tr>
</thead>
<tbody>
<tr>
<td>IDH2</td>
<td>JAK2</td>
<td>KIT</td>
<td>KRAS</td>
<td>MAPK1 (ERK)</td>
<td>MAPK2 (MEK)</td>
<td>MET</td>
<td>MLL*</td>
<td>NRAS</td>
</tr>
<tr>
<td>NPM1</td>
<td>PDGFRA</td>
<td>PIK3CA</td>
<td>PTEN</td>
<td>PTPN11</td>
<td>RET</td>
<td>RUNX1</td>
<td>TP53</td>
<td>WF1</td>
</tr>
</tbody>
</table>

* FISH  
** SNV-NGS, Amplification-FISH

---

**Comprehensive Cancer Set**  
*(Next generation sequencing of all coding regions of ABL1, ALK, APC, ASXL1, ATM, BRAF, CEBPA, CTNNB1, DNMT3A, ERBB2, EGFR, ESR1, FGFR4, FLT3, IDH1, IDH2, JAK2, KIT, KRAS, MAPK1(ERK), MAP2K2(MEK), MET, MLL, MPL, MYC, MYD88, NOTCH1, NPM1, NRAS, PDGFRA, PIK3CA, PTEN, PTPN11, RB1, RET, RUNX1, TET2, TP53, VHL, WT1)*

**Version 2**

**PART 1: Genes to Report**

- [ ] Report All Genes
- [ ] Do Not Report All Genes; Report Only the Following Gene(s) (select all that apply):

| ABL1 | ALK | APC | ASXL1 | ATM | BRAF | CEBPA | CTNNB1 | DNMT3A* | ERBB2 | EGFR | ESR1 | FGFR4 | FLT3* | IDH1 | IDH2 | JAK2* | KIT | MAPK1(ERK) | MAP2K2(MEK) | MET | MLL | MPL | MYC | MYD88 | NOTCH1 | NPM1 | NRAS | PDGFRA | PIK3CA | PTEN | PTPN11 | RB1 | RET | RUNX1 | TET2 | TP53 | VHL | WT1 |
|------|-----|-----|-------|-----|------|-------|--------|--------|-------|------|------|-------|-------|-------|------|------|-------|-----|------------|------------|-----|-----|-----|-----|-------|-------|-----|------|-------|------|------|-------|-----|-----|------|-----|-------|-----|-----|-----|-----|

* cannot be ordered as single genes

**PART 2: Select FISH Testing**

- [ ] ALK – 2p23
- [ ] EGFR – 7p12
- [ ] HER2-NEU
- [ ] MLL – 11q23
- [ ] MYC – 8q24 Rearrangement
- [ ] MYC – 8q24 Amplification
- [ ] MYCN – 2p24.3 Amplification
- [ ] Other:
- [ ] ROS1 – 6q22.1
- [ ] RET – 10q11.21

**Sanger Sequencing Assays**

- [ ] HER2 (ERBB2) TKD
- [ ] TREX1
- [ ] BRAF – V600E Genotyping
- [ ] WFS1 Full Coding Region
Vision

‘one assay to rule them all’

slide courtesy of Eric Duncavage
Overview

DNA Capture Sequence Align SNV Indels Validate Report

Translocations

Variant Detection

slide courtesy of Eric Duncavage
Data Analysis Pipeline

---

**SNV Calls (GATK)**

**Flag SNPs (GATK)**

**Merge Data Files**

**Data Output**

---

**Read Alignment (Novoalign)**

**Indel Calls (Pindel)**

**Translocation Calls (Breakdancer)**

**Translocation Validation (Slope)**

---

*slide courtesy of Eric Duncavage*
Patient case report

- 55 y.o. F presented with lower abdominal pain
- CT scan and FDG PET noted 2 lesions in liver and 1 in the mediastinum
- Liver biopsy consistent with metastatic, poorly-differentiated carcinoma
- Pathological diagnosis: metastatic thymic carcinoma.
Treatments…


2. Cycle #5 of paclitaxel and carboplatin dose reduced to 175 mg/m² on 12/9/2010; due to neuropathy.


4. Gemcitabine 1000 mg/m² started on 8/25/2011, with 6 cycles through 1/5/2012.

5. Chemoembolization of the liver tumors in segments 5 and 8 with 50 mg of doxorubicin on 2/3/2012.
NGS sequencing results

IGV (Integrated Genome Viewer)
Three base pair deletion resulting in p.D579del

IGV (Integrated Genome Viewer)
Sequencing Results

- “possible pathogenic mutation” in the KIT gene.
- p.D579del (c.1735_1737delGAT) KIT mutation
- occurs in the juxtamembrane domain (exon 11) of the protein
- was Sanger sequenced to confirm
Sanger sequencing to confirm deletion
Sanger sequencing to confirm deletion
p.D579del

- This variant has been observed previously in the setting of gastrointestinal stromal tumors (GIST).

- KIT mutations have been documented in a small proportion ~7% of thymic carcinomas and the spectrum of mutations is thought to overlap those detected in GISTs (Clin Cancer Res. 2009 Nov 15;15(22):6790-9).
KIT

• v-kit Hardy-Zuckerman 4 feline sarcoma viral oncogene homolog of proto-oncogene c-kit

• type 3 transmembrane receptor for MGF (mast cell growth factor)

• Mutations associated with gastrointestinal stromal tumors, mast cell disease, acute myelogenous leukemia, and piebaldism
KIT mutations

p.D579del
Targeted therapy

- Imatinib (Gleevec)
- tyrosine kinase inhibitor
- first targeted cancer therapy
- inhibits bcr-abl fusion kinase and others
Imatinib (Gleevec) for GIST patients:

- ~85% of GIST patients have a KIT mutation
- 2 yr survival: from 26% before Imatinib, to 70% with imatinib
  - Blanke, CD et al, JCO;26,4, Feb 1 2008; 626-632
Treatments history for this patient

1. Paclitaxel and carboplatin, 4 cycles
2. Cycle #5 of paclitaxel and carboplatin dose reduced to 175 mg/m2 on --------; due to neuropathy.
3. Pemetrexed 500 mg/m2, two cycles, 9 months later
4. Gemcitabine 1000 mg/m2 started on ------, 6 cycles
5. Chemoembolization of the liver tumors in segments 5 and 8 with 50 mg of doxorubicin on -------
6. Imatinib therapy started based on the presence of mutation in exon 11 of the KIT gene
The patient returns today for scheduled follow up visit. Her most recent CT scan shows essentially stable with disease with no new or worsening metastatic lesions. She is continuing to tolerate the Gleevec quite well.
Cancer genetics summary

- really big topic
- a small introduction today
- the genome is just a platform for the real action....
**SNPs per individual genome**

**Table 1: Summary of 1000 Genomes Project phase I data**

An integrated map of genetic variation from 1,092 human genomes

The 1000 Genomes Project Consortium

*Nature* 491, 56–65 (01 November 2012)  doi:10.1038/nature11632

<table>
<thead>
<tr>
<th></th>
<th>Autosomes</th>
<th>Chromosome X</th>
<th>GENCODE regions*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Samples</td>
<td>1,092</td>
<td>1,092</td>
<td>1,092</td>
</tr>
<tr>
<td>Total raw bases (Gb)</td>
<td>19,049</td>
<td>804</td>
<td>327</td>
</tr>
<tr>
<td>Mean mapped depth (×)</td>
<td>5.1</td>
<td>3.9</td>
<td>80.3</td>
</tr>
<tr>
<td>SNPs</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No. sites overall</td>
<td>36.7 M</td>
<td>1.3 M</td>
<td>498 K</td>
</tr>
<tr>
<td>Novelty rate†</td>
<td>58%</td>
<td>77%</td>
<td>50%</td>
</tr>
<tr>
<td>No. synonymous/non-synonymous/nonsense</td>
<td>NA</td>
<td>4.7/6.5/0.097 K</td>
<td>199/293/6.3 K</td>
</tr>
<tr>
<td>Average no. SNPs per sample</td>
<td>3.60 M</td>
<td>105 K</td>
<td>24.0 K</td>
</tr>
<tr>
<td>Indels</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No. sites overall</td>
<td>1.38 M</td>
<td>59 K</td>
<td>1,867</td>
</tr>
<tr>
<td>Novelty rate†</td>
<td>62%</td>
<td>73%</td>
<td>54%</td>
</tr>
<tr>
<td>No. inframe/frameshift</td>
<td>NA</td>
<td>19/14</td>
<td>719/1,066</td>
</tr>
<tr>
<td>Average no. indels per sample</td>
<td>344 K</td>
<td>13 K</td>
<td>440</td>
</tr>
<tr>
<td>Genotyped large deletions</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No. sites overall</td>
<td>13.8 K</td>
<td>432</td>
<td>847</td>
</tr>
<tr>
<td>Novelty rate†</td>
<td>54%</td>
<td>54%</td>
<td>50%</td>
</tr>
<tr>
<td>Average no. variants per sample</td>
<td>717</td>
<td>26</td>
<td>39</td>
</tr>
</tbody>
</table>

NA, not applicable.

* Autosomal genes only.

† Compared with dbSNP release 135 (Oct 2011), excluding contribution from phase I 1000 Genomes Project (or equivalent data for large deletions).