



Cold  
Spring  
Harbor  
Laboratory

# Advanced Sequencing Technologies & Applications

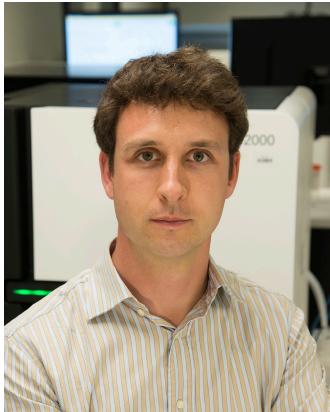
<http://meetings.cshl.edu/courses.html>

# Introductions to MGI/WUSTL instructors



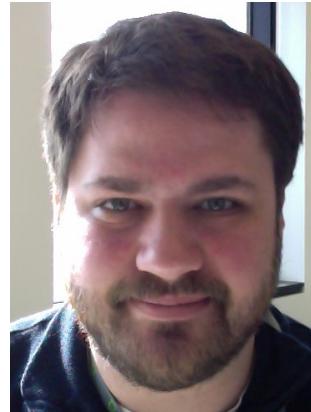
**Malachi Griffith**

Assistant Professor of Medicine  
Assistant Professor of Genetics  
Assistant Director, MGI



**Obi Griffith, PhD**

Assistant Professor of Medicine  
Assistant Professor of Genetics  
Assistant Director, MGI



**Alex Wagner**

Postdoc fellow



**Jason Walker**

Director of Informatics



**Kelsy Cotto**

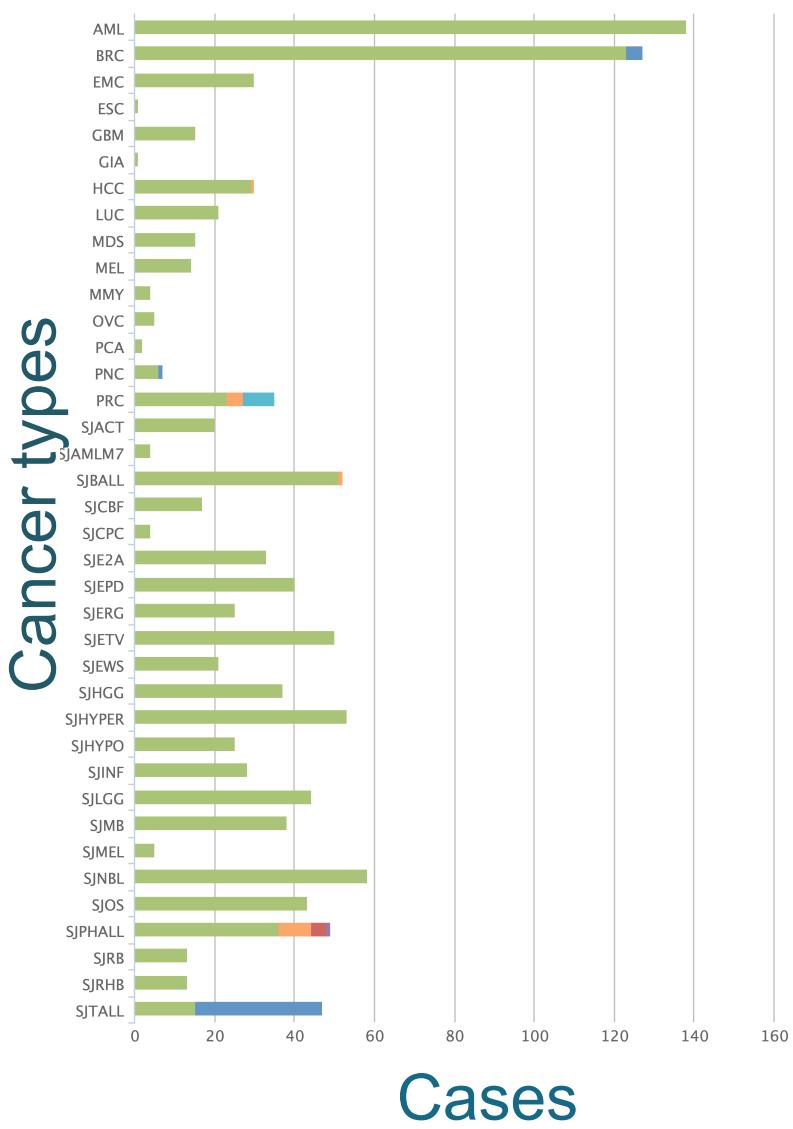
PhD candidate

**McDonnell Genome Institute, Washington University School of Medicine**

**Advanced Sequencing Technologies & Applications**

<http://meetings.cshl.edu/>

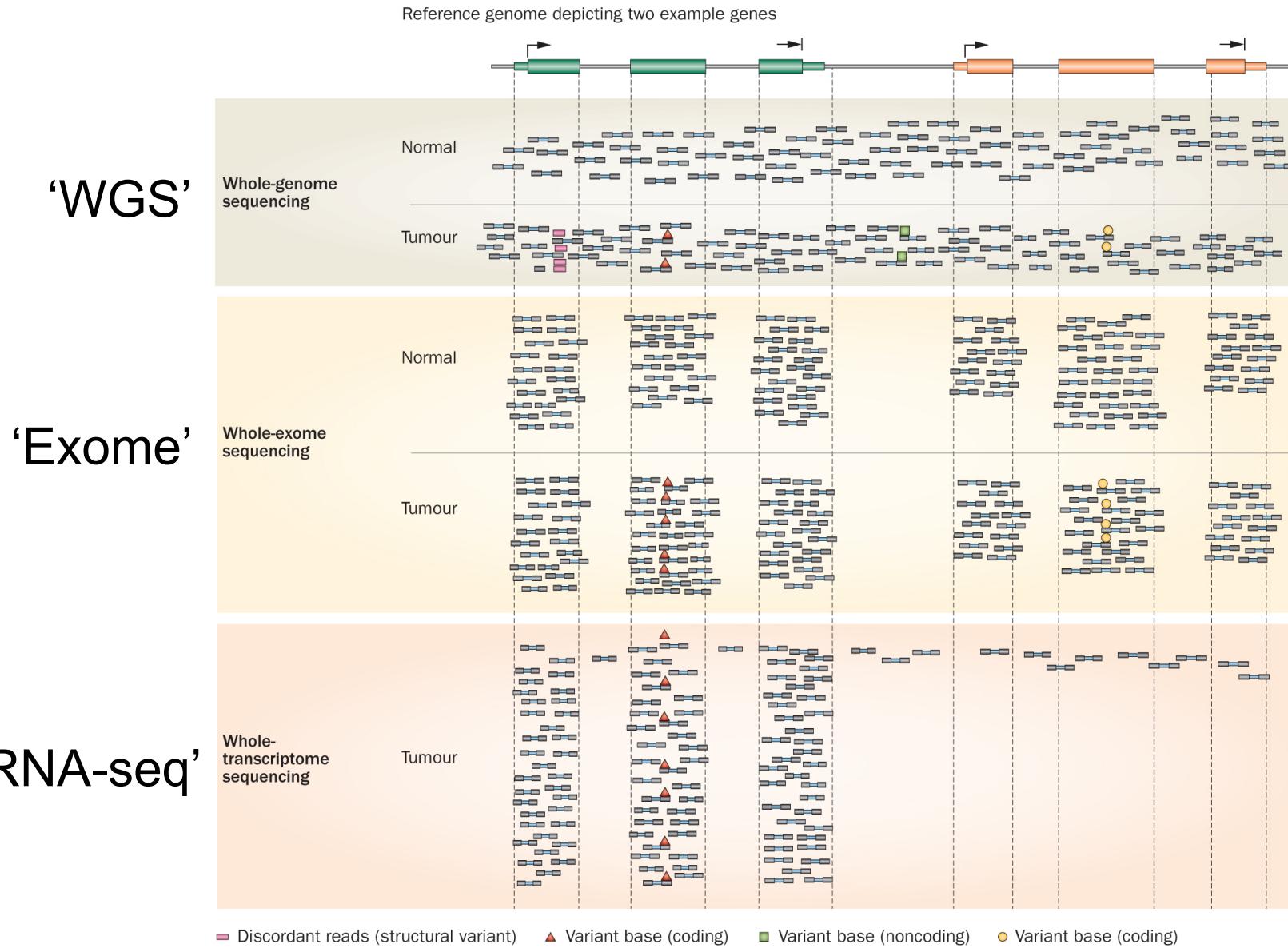
The McDonnell Genome Institute has pursued the field of genomics since inception: >1000 whole genomes, >5000 exomes, >1000 transcriptomes for dozens of tumor types



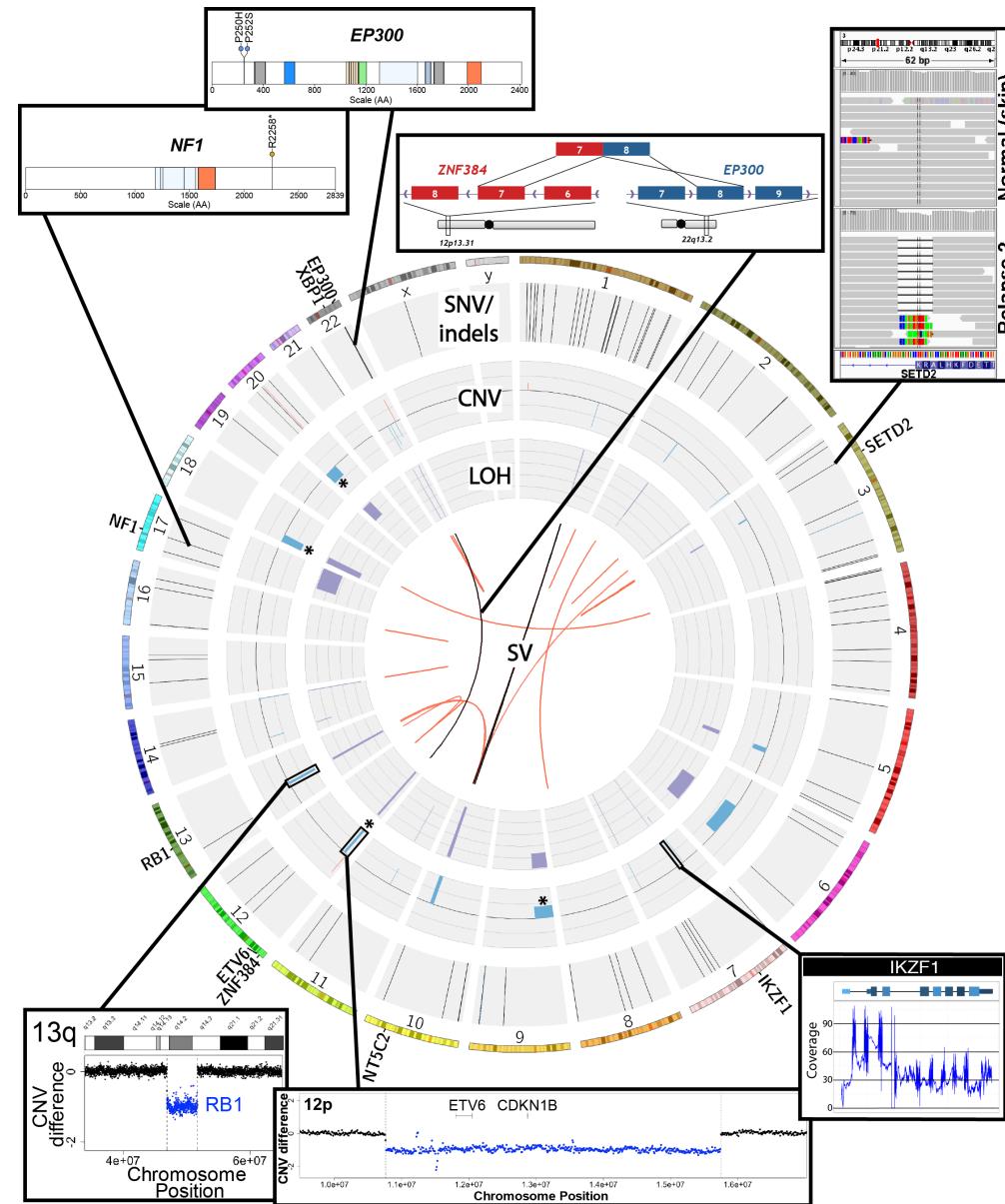
- Human Genome Project
- First to sequence and analyze a tumor whole genome sequence (Ley et al, 2008)
- Major contributor to TCGA, PCGP, etc.
- Most comprehensively sequenced single patient tumor ever published (Griffith et al, 2015)
- Early proof-of-principle for cancer precision medicine (Griffith et al, 2016)
- Analysis and tools for first personalized cancer vaccine design in humans (Carreno et al, 2015)
- Many other widely used tools

MG - 1000+ tumor whole genomes, many more exomes, X10 & NovaSeq will accelerate this!

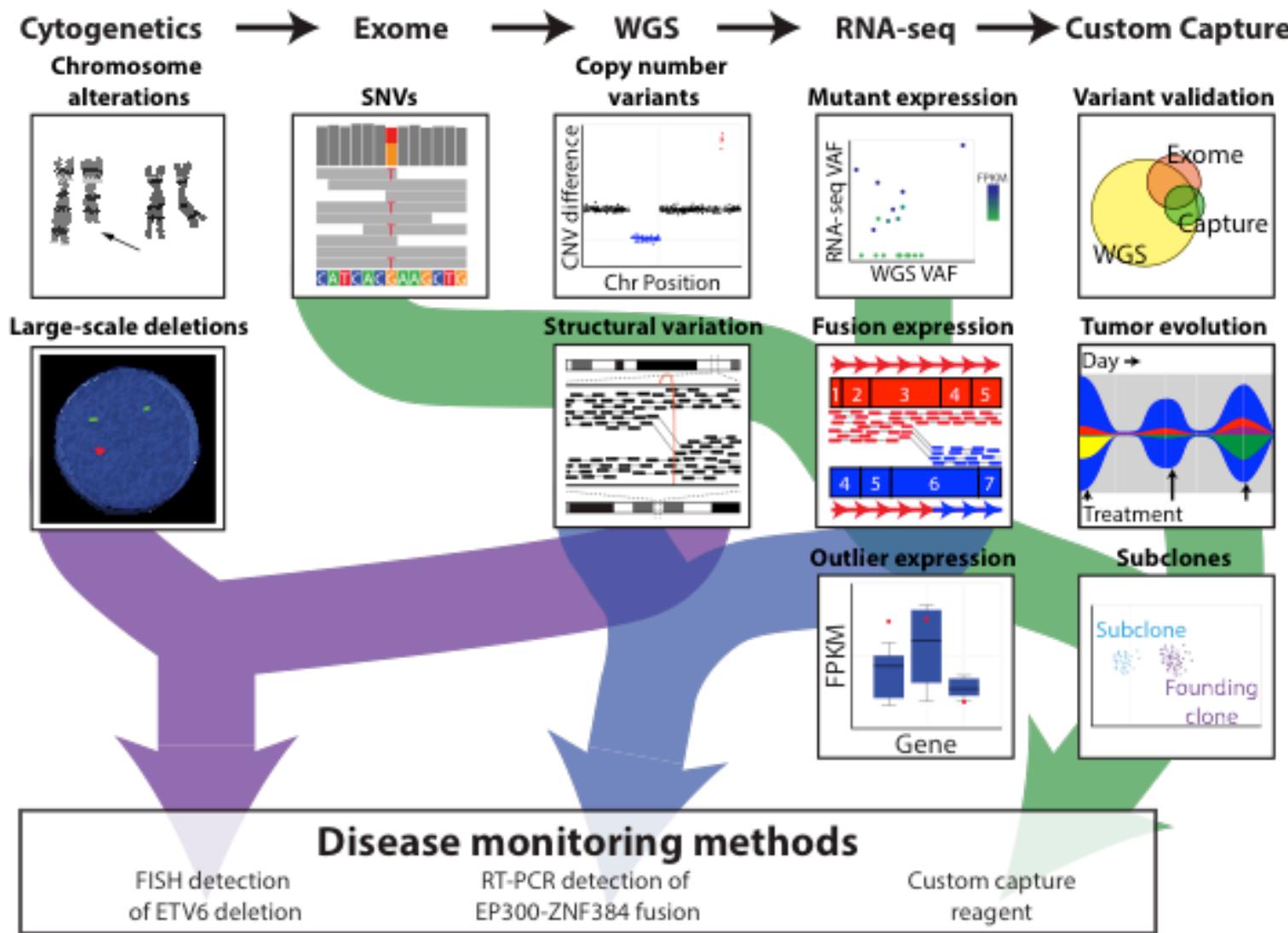
# Whole genome, exome, transcriptome and other ‘omic’ sequencing allows us to detect and confirm many different variant types



# SNVs, Indels, CNVs, SVs, fusions, LOH, expression changes, methylation changes, and more

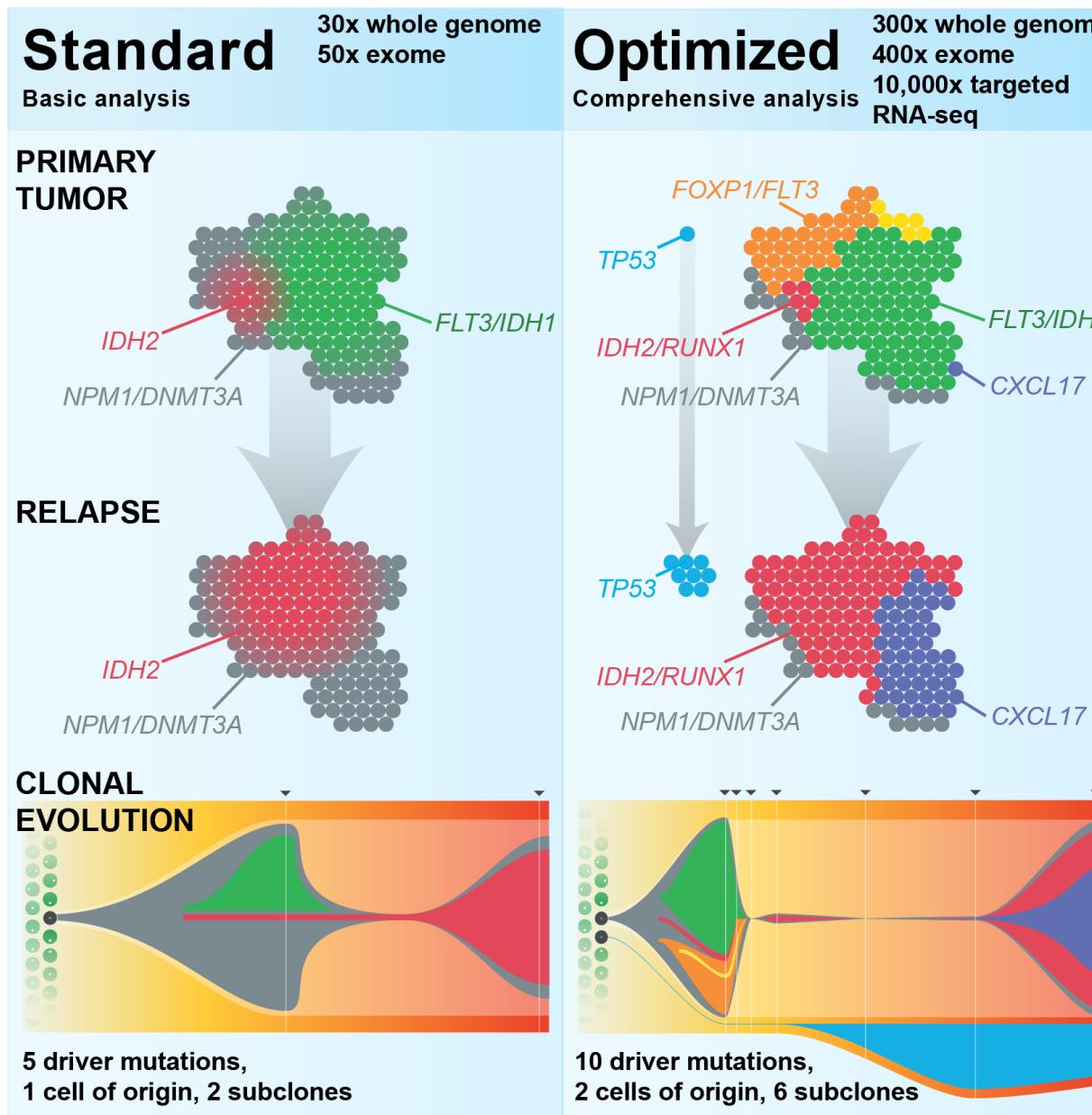


# Comprehensive and integrative analysis methods are needed



Kilannin Krysiak

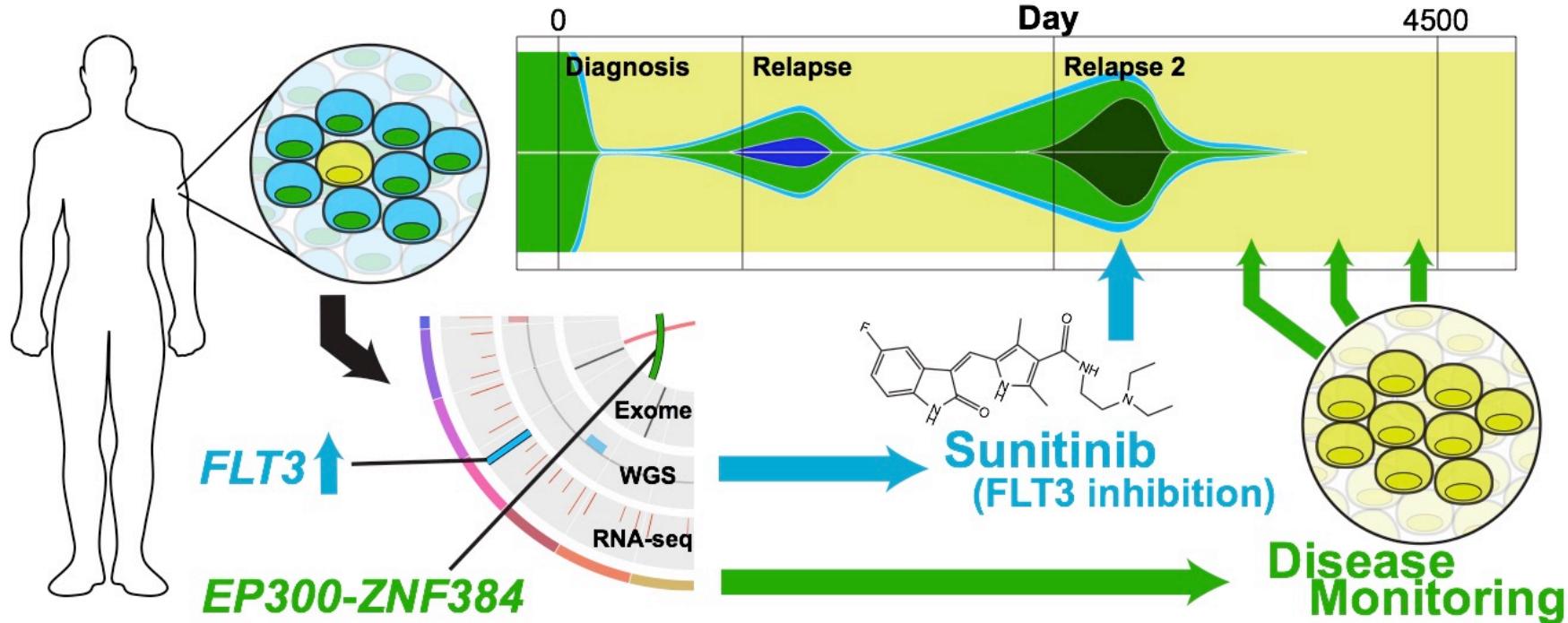
# Each study often requires considerable customization



Kilannin Krysiak

[Griffith et al. 2015](#)

# Personalized medicine requires personalized strategies



ELSEVIER



CrossMark

Experimental Hematology 2016;44:603–613

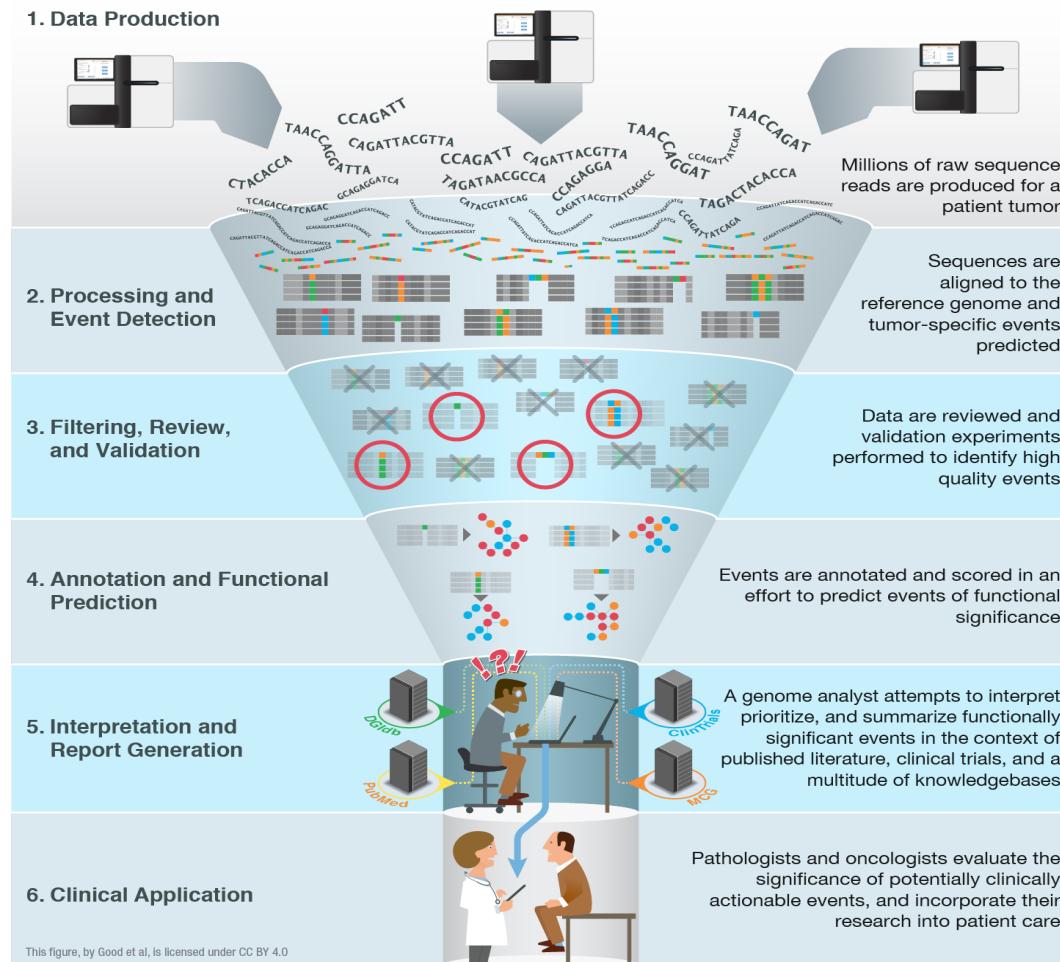
Experimental  
Hematology

Comprehensive genomic analysis reveals *FLT3* activation and a therapeutic strategy for a patient with relapsed adult B-lymphoblastic leukemia

Kilannin Krysiak

[Griffith et al. 2016](#)

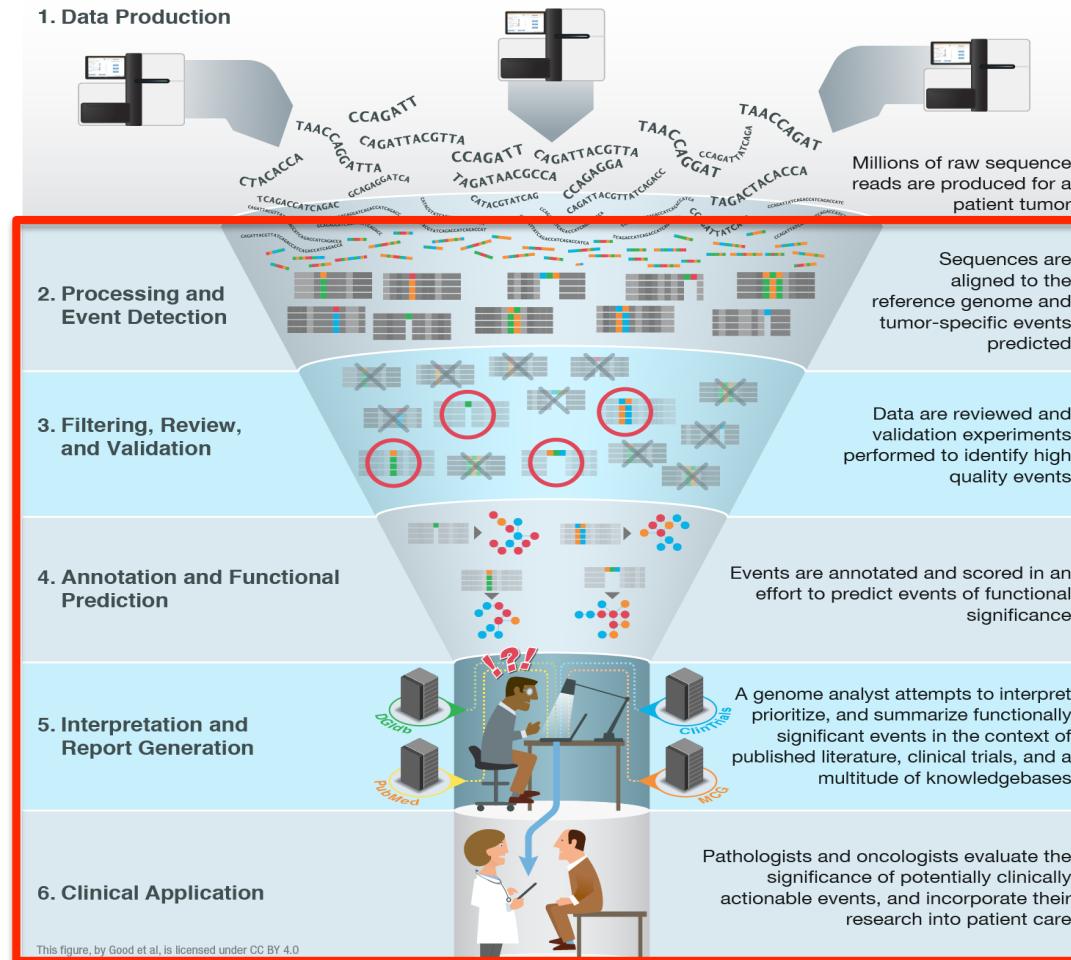
# High-throughput sequencing has been largely automated



Good BM, Ainscough BJ, McMichael JF, Su AI, Griffith OL. 2014.  
Genome Biology. 15(8):438.

Joshua McMichael

The rest of this workshop will focus on the methods and tools needed to take raw sequence data to interpretation and application – still the bottleneck!



Good BM, Ainscough BJ, McMichael JF, Su AI, Griffith OL. 2014.  
Genome Biology. 15(8):438.

Joshua McMichael

# The Griffith lab is focused on developing methods to address this bottleneck for precision medicine in cancer

- **Cancer genome analysis**
  - AML
    - [Optimizing cancer genome analysis](#)
  - Breast cancer
    - [Immunotherapy](#) (cancer vaccines)
    - STAT1-/ mouse model
    - Targeted sequencing of 625 ER+ with long-term follow-up
  - Liver cancer
  - Small cell lung cancer
  - OSCC...
- **Precision Medicine for Cancer**
  - [Genomics Tumor Board](#)
    - [Case Reports](#) and [Clinical Trials](#)
- **Education projects**
  - [RNA-seq analysis and cloud computing](#)
  - [CSHL](#) and [CBW](#)
  - [Genomic Data Visualization/Interpretation](#)
- **Tool development**

# Where tools/resources do not exist we build them



[www.dgidb.org](http://www.dgidb.org)

Search genes for known and potentially druggable interactions



<https://github.com/griffithlab/pVAC-Seq>

Personalize vaccine design



[www.docm.info](http://www.docm.info)

Filter against highly curated set of mutations known to cause cancer



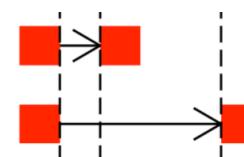
<https://github.com/griffithlab/GenVisR>

Create genomic visualizations



[www.civicdb.org](http://www.civicdb.org)

Identify highly curated summaries of clinical interpretations for variants in cancer



**regtools**

<https://github.com/griffithlab/regtools>

Identify regulatory variants

# Encourage best practices for software development

The screenshot shows the GitHub organization page for 'The Griffith Lab'. At the top, there's a header with a GitHub icon, 'This organization', 'Search', 'Pull requests', 'Issues', 'Gist', and a notification bell. Below the header is a profile picture of two men and the text 'The Griffith Lab' and 'Academic Lab of Obi and Malachi Griffith'. A link to their website is provided. The main navigation bar includes 'Repositories', 'People 25', 'Teams 21', 'Projects 0', and 'Settings'. Below the navigation is a search bar, filters for 'Type: All' and 'Language: All', and a 'Customize pinned repositories' section with a 'New' button. The main content area lists three repositories: 'dgi-db' (Rails frontend to The Genome Institute's drug gene interaction database), 'pVAC-Seq' (A cancer immunotherapy pipeline), and 'GenVisR' (Genome data visualizations). Each repository has a brief description, a language badge, and a commit history. To the right, there's a 'Top languages' section showing R, Python, Ruby, Perl, and HTML, and a 'People' section showing a grid of 25 team members.

- All code deposited in Github
  - 35 repos for different software/projects
  - Virtually all public and open-source
  - Test-driven, code review, pull requests, etc

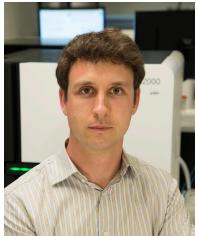
<https://github.com/griffithlab>

<https://github.com/genome>

# WUSTL – MGI: Group members



**Malachi  
Griffith**



**Obi  
Griffith**



**Benjamin  
Ainscough**



**Erica  
Barnell**



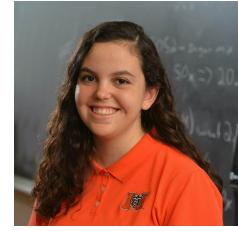
**Katie  
Campbell**



**Kaitlin  
Clark**



**Adam  
Coffman**



**Kelsy  
Cotto**



**Arpad  
Danos**



**Yang-Yang  
Feng**



**Felicia  
Gomez**



**Jasreet  
Hundal**



**Susann  
a Kiwala**



**Kilanin  
Krysiak**



**Lynzey  
Kujan**



**Jason  
Kunisaki**



**Josh  
McMichael**



**Cody  
Ramirez**



**Zachary  
Skidmore**



**Nick  
Spies**



**Lee  
Trani**



**Alex  
Wagner**



**Jason  
Walker**

# Informatics background poll

- Student poll
  - Are you doing genomics research?
  - Do you consider yourself a bioinformatician?
  - What kind of NGS data are you working with?
    - WGS?
    - Exome?
    - RNAseq?
    - Epigenome?
    - Single cell?
    - Other?
  - What organism do you work with?
  - Did you bring data?

# Student poll continued

Not counting the pre-requisites and materials for this course:

- Are you familiar with linux/command line?
  - Intermediate?
  - Expert?
- Do you sometimes write code?
  - What language?
- Are you familiar with R?
  - Intermediate?
  - Expert?
- Are you familiar with ggplot?
  - Intermediate?
  - Expert?
- Do you use git/github?