



The Elizabeth H.
and James S. McDonnell III

McDONNELL
GENOME INSTITUTE
at Washington University

GenViz Module 0: Introductions

Malachi Griffith, Obi Griffith, Zachary Skidmore
Genomic Data Visualization and Interpretation

April 8-12, 2019
Freie Universität Berlin



Attribution-ShareAlike 4.0 International (CC BY-SA 4.0)

This is a human-readable summary of (and not a substitute for) the [license](#). [Disclaimer](#).

You are free to:

Share — copy and redistribute the material in any medium or format

Adapt — remix, transform, and build upon the material for any purpose, even commercially.



The licensor cannot revoke these freedoms as long as you follow the license terms.

Under the following terms:



Attribution — You must give [appropriate credit](#), provide a link to the license, and [indicate if changes were made](#). You may do so in any reasonable manner, but not in any way that suggests the licensor endorses you or your use.



ShareAlike — If you remix, transform, or build upon the material, you must distribute your contributions under the [same license](#) as the original.

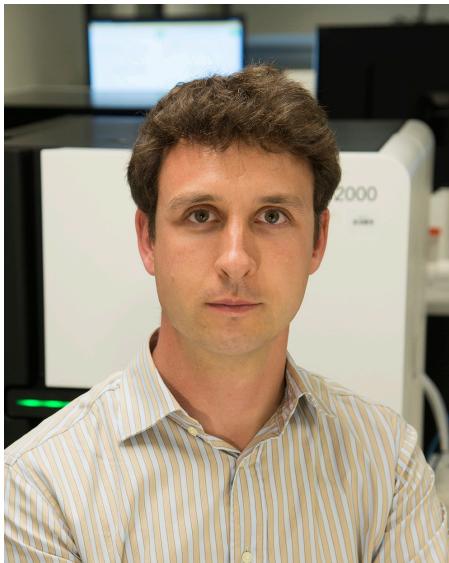
No additional restrictions — You may not apply legal terms or [technological measures](#) that legally restrict others from doing anything the license permits.

Introductions to course instructors



Malachi Griffith, PhD

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



Zachary Skidmore, MSc

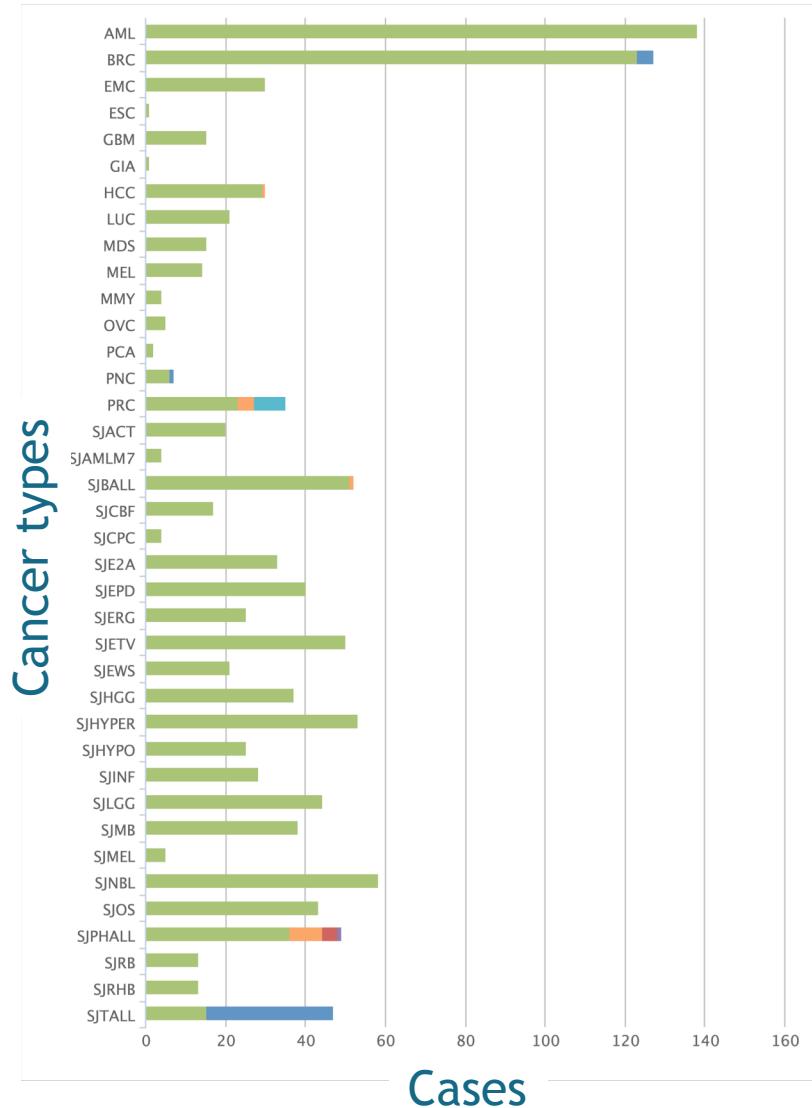
Staff Scientist, MGI

GenVisR creator

McDonnell Genome Institute, Washington University School of Medicine

griffithlab.org

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



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

- 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

Overview of lab's research (griffithlab.org/research)

- **Cancer genome analysis**
 - Breast cancer, Liver cancer, Lung cancer, Head and neck cancer, etc.
 - Variant Interpretation
 - Immunogenomics
- **Precision medicine for cancer**
 - [Genomics Tumor Board](#)
 - [Case Reports](#)
 - [Clinical Trials](#)
 - [Personalized Cancer Vaccines](#)
- **Education projects**
 - [RNA-seq analysis and cloud computing \(CBW, Toronto\)](#)
 - [Advanced Sequencing Technologies and Applications \(CSHL, New York\)](#)
 - [Genomic Data Visualization/Interpretation \(Physalia Courses, Berlin\)](#)
 - [Precision Medicine Bioinformatics \(PR Informatics, Glasgow\)](#)
 - [High-Throughput Biology: From Sequence to Networks \(CSHL / CBW, New York\)](#)
 - [Workshop on Genomics \(Evomics, Český Krumlov\)](#)
- **Tool development**

Where tools/resources do not exist we build them



www.dgidb.org

Search genes for known
and potentially druggable
interactions



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

Personalize vaccine design



www.docm.info

Filter against highly
curated set of mutations
known to cause cancer



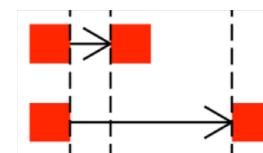
<https://github.com/griffithlab/GenVisR>

Create genomic visualizations



www.civicdb.org

Identify highly curated summaries
of clinical interpretations for
variants in cancer

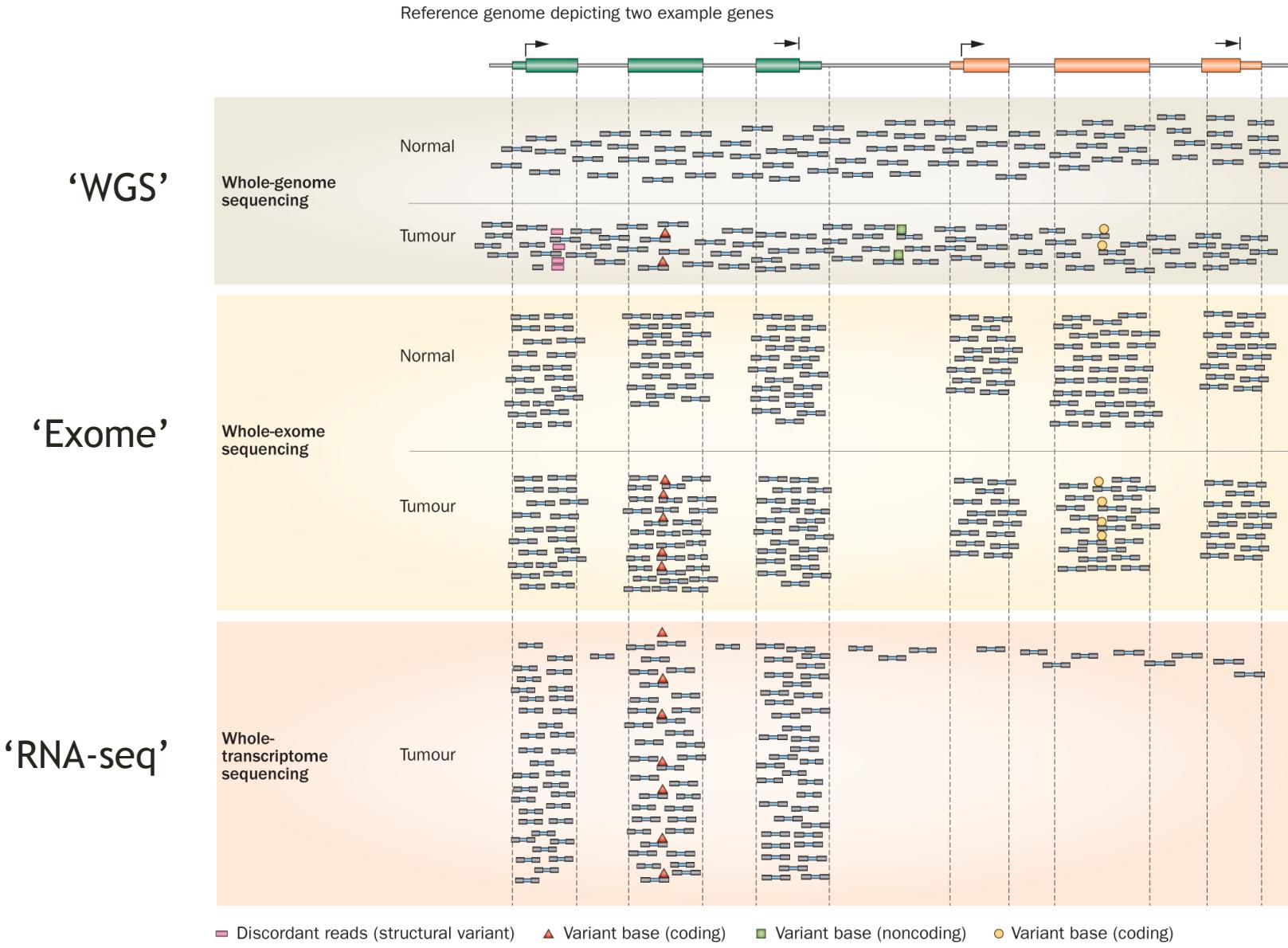


regtools

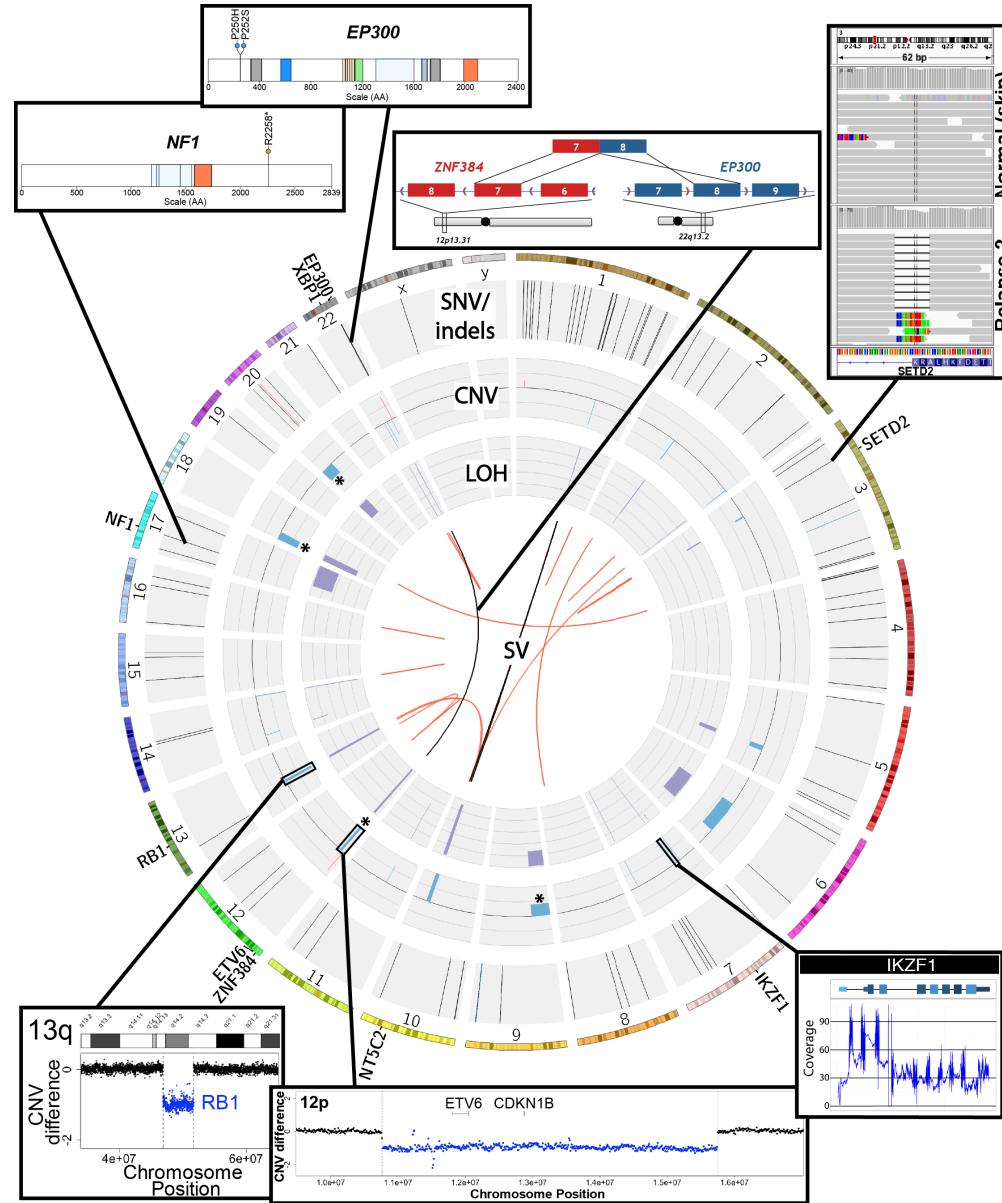
<https://github.com/griffithlab/regtools>

Identify regulatory variants

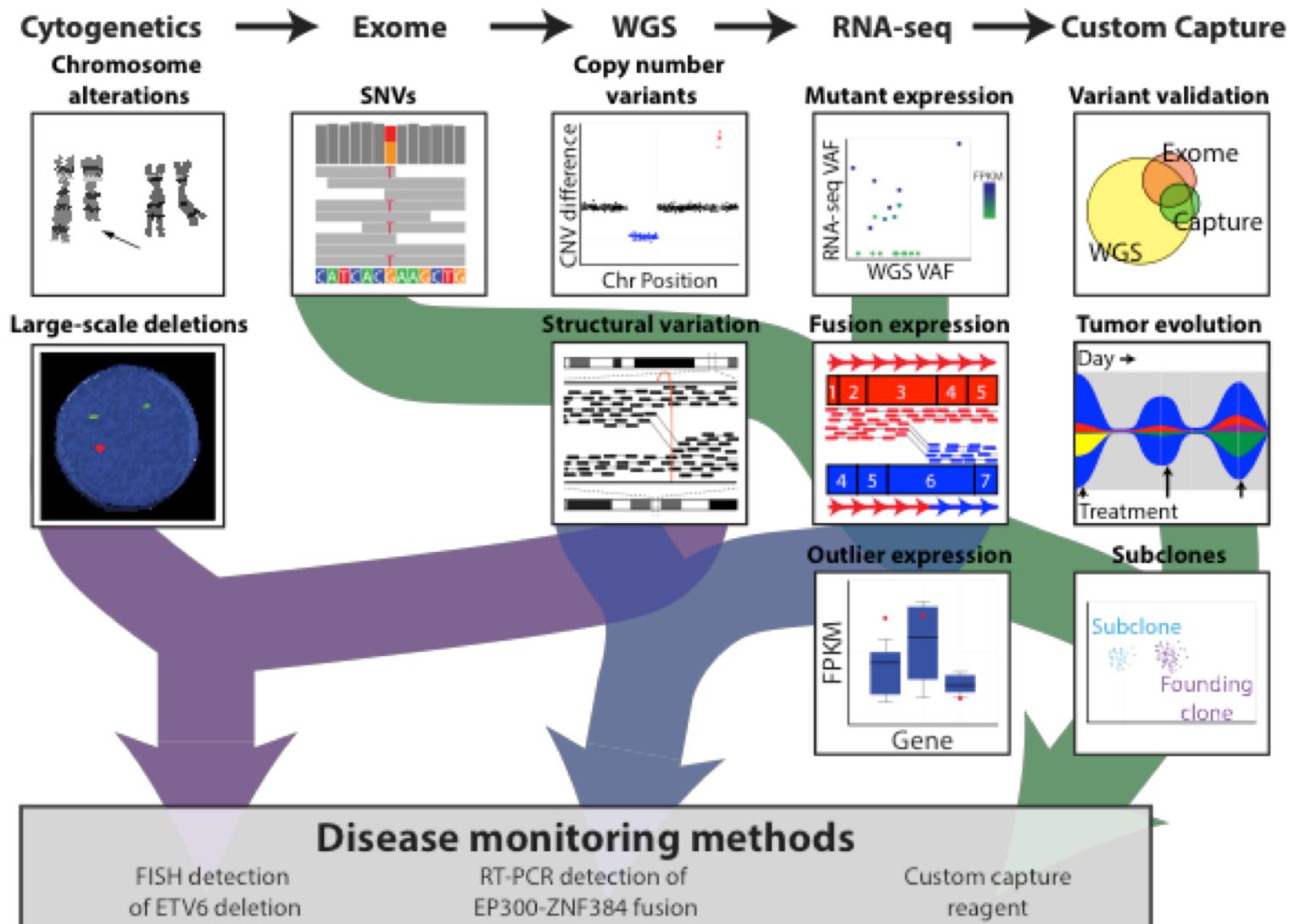
Whole genome, exome, transcriptome and other ‘omic’ sequencing allows us to detect and confirm many different variant types



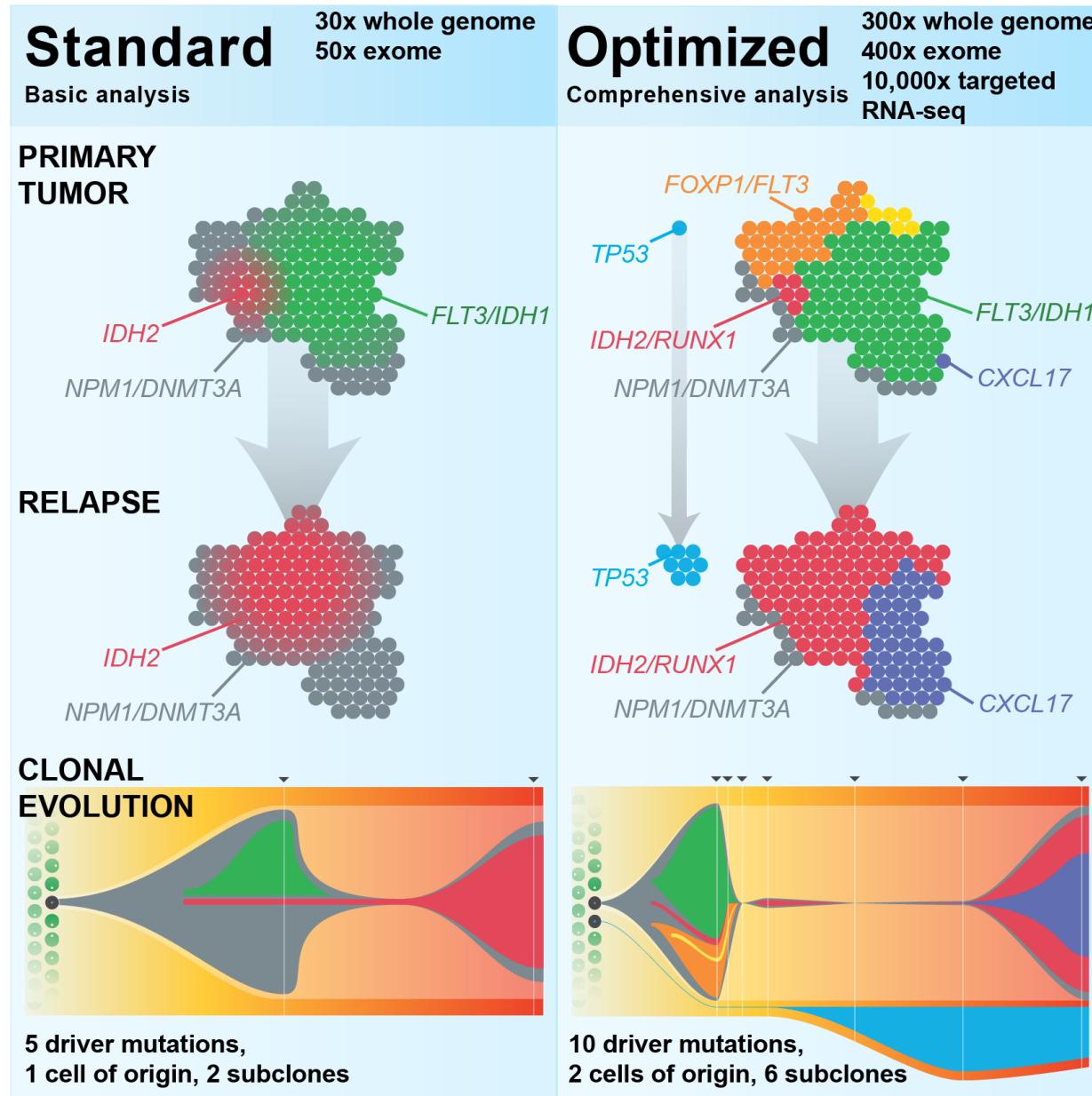
SNVs, Indels, CNVs, SVs, fusions, LOH, expression changes, methylation changes, and more



Comprehensive and integrative analysis methods are needed



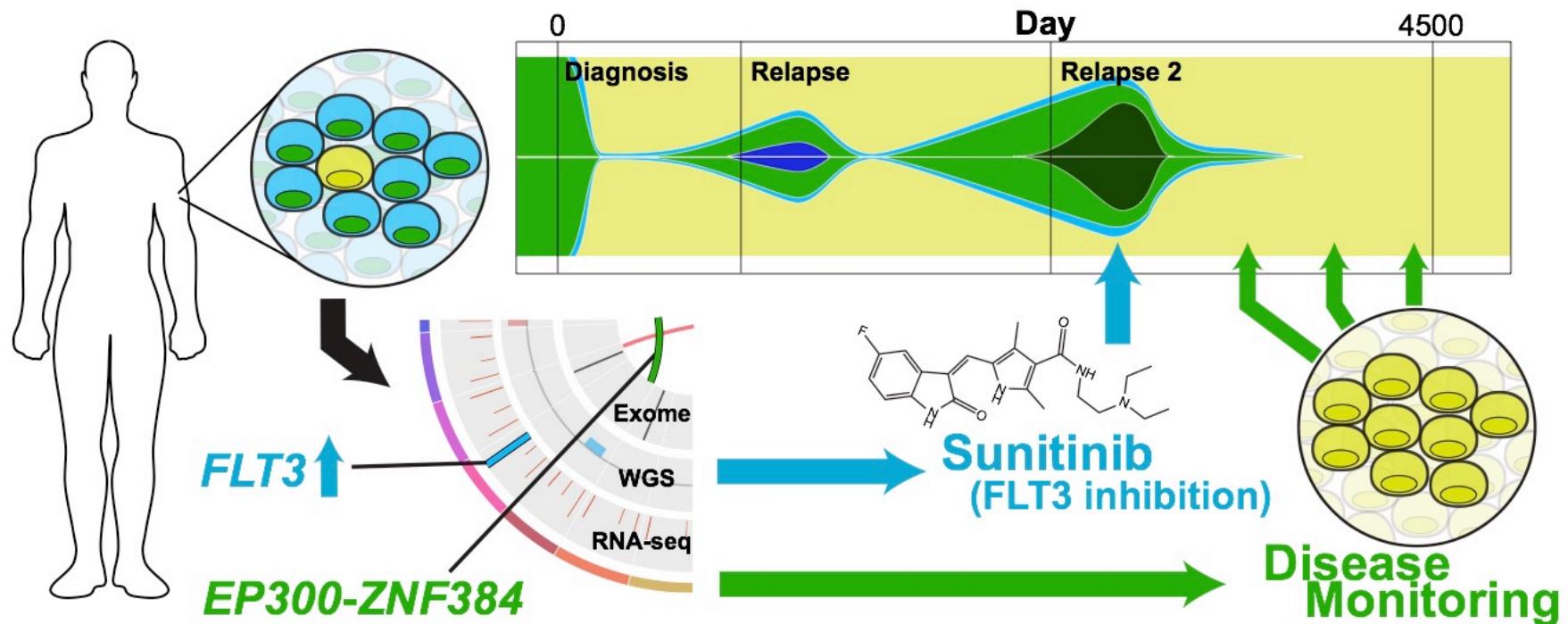
Each study often requires considerable customization



Griffith et al. 2015

Kilannin Krysiak

Personalized medicine requires personalized strategies



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

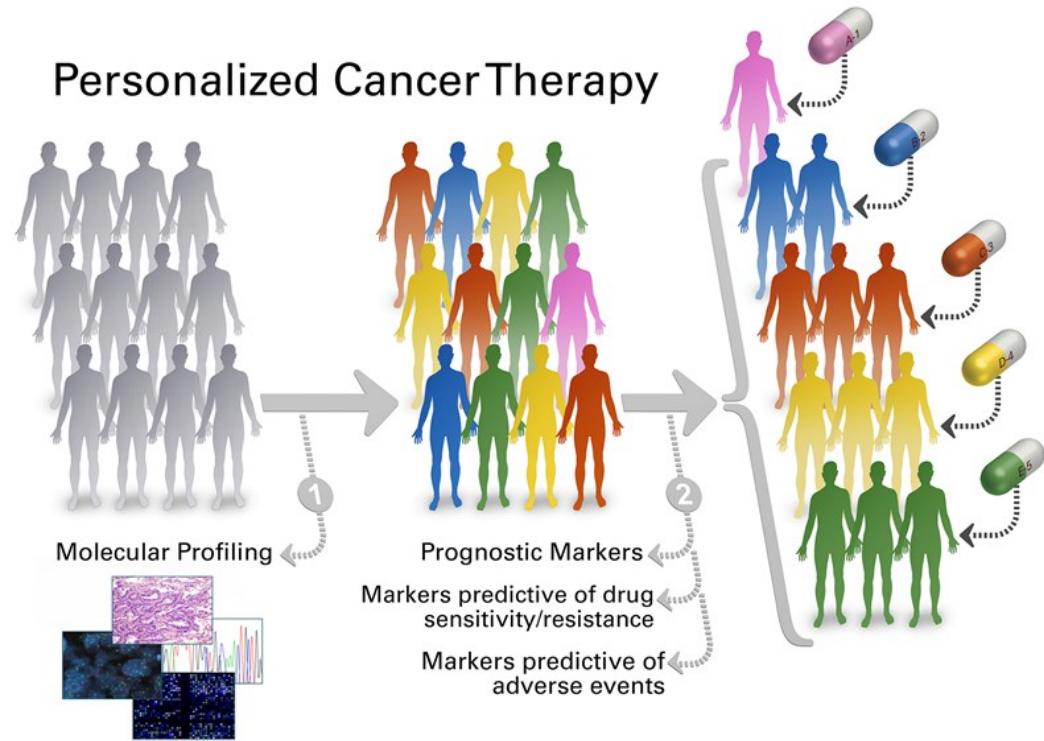
[Griffith et al. 2016](#)

Strategies to bring genomics information to bear for as many cancer patients as possible

1. Precision medicine targeting of driver mutations
2. Leveraging passenger variants
 - a. Tracking minimal residual disease
 - b. Identifying neoepitopes
 - Predicting response to immunotherapy
 - Developing personalized vaccines

Precision medicine targeting of driver mutations

Personalized Cancer Therapy



1980s: Development of Targeted Therapies

2000: Human Genome Sequencing Project

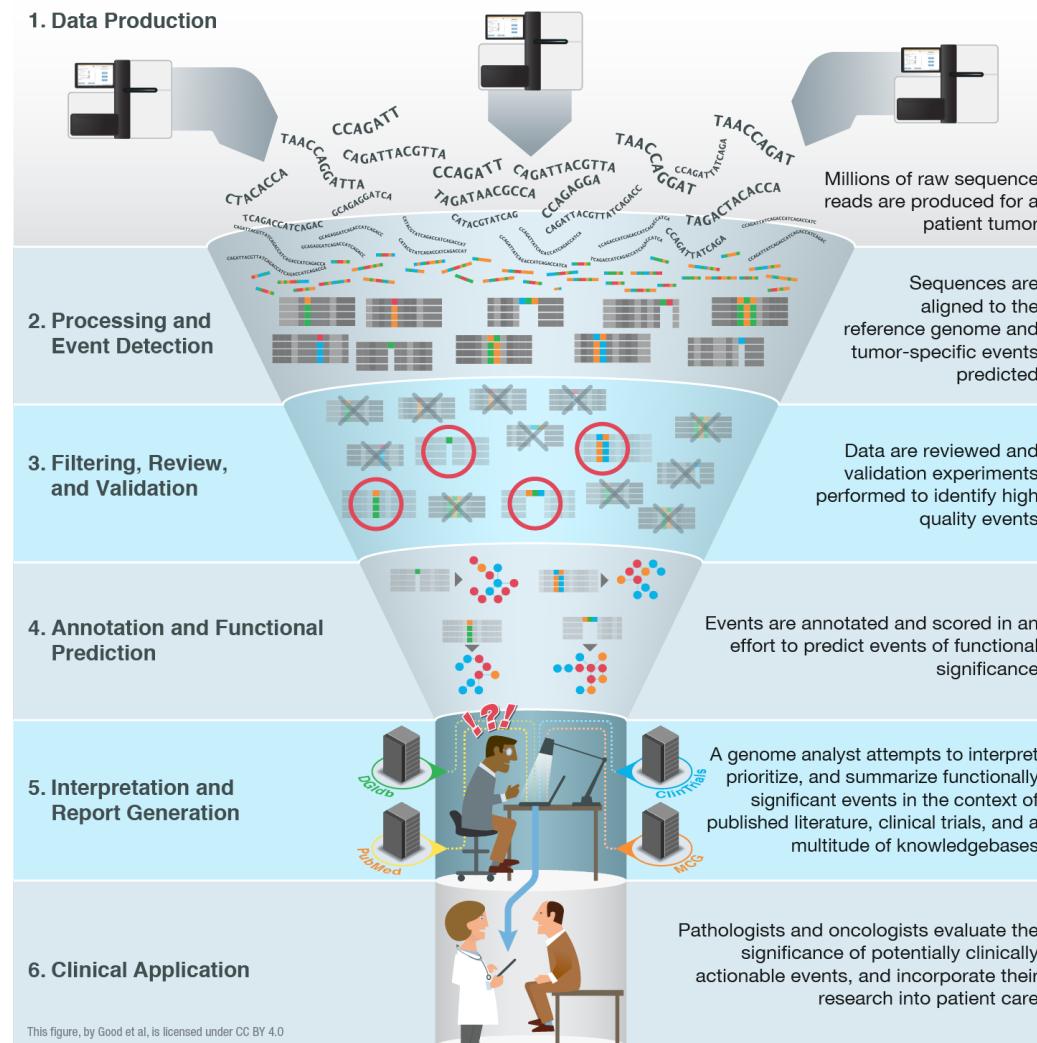
2001: FDA Approval for Imatinib in BCR-ABL1 CML

2014: FDA Approval for BRCA Testing

2017: FDA Approval of Pembrolizumab for MSI tumors

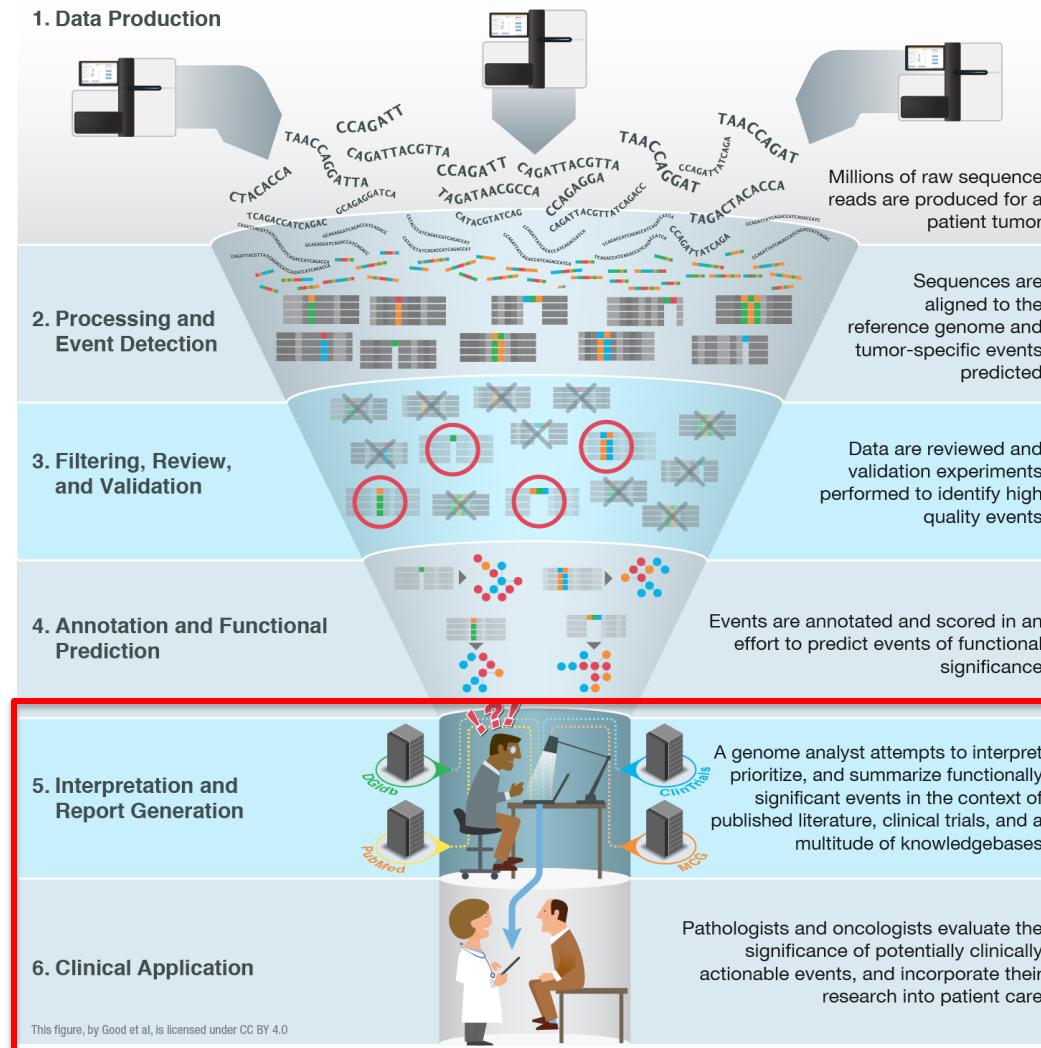
BRAF → V600E → Melanoma → Predictive → Vemurafenib
ERBB2 → Amplification → Breast → Predictive → Trastuzumab
EGFR → L858R → Lung → Predictive → Erlotinib
ALK → Fusions → Lung → Predictive → Crizotinib
EWSR1-FLI1 → Fusions → Ewing Sarcoma → Diagnostic
DNAJB1-PRKACA → Fusions → fHCC → Diagnostic
VHL → Loss of function mutations → Kidney → Predisposing
... an increasingly long tail of rare but clinically relevant variants

High-throughput sequencing has been largely automated allowing rapid identification of somatic and germline variants



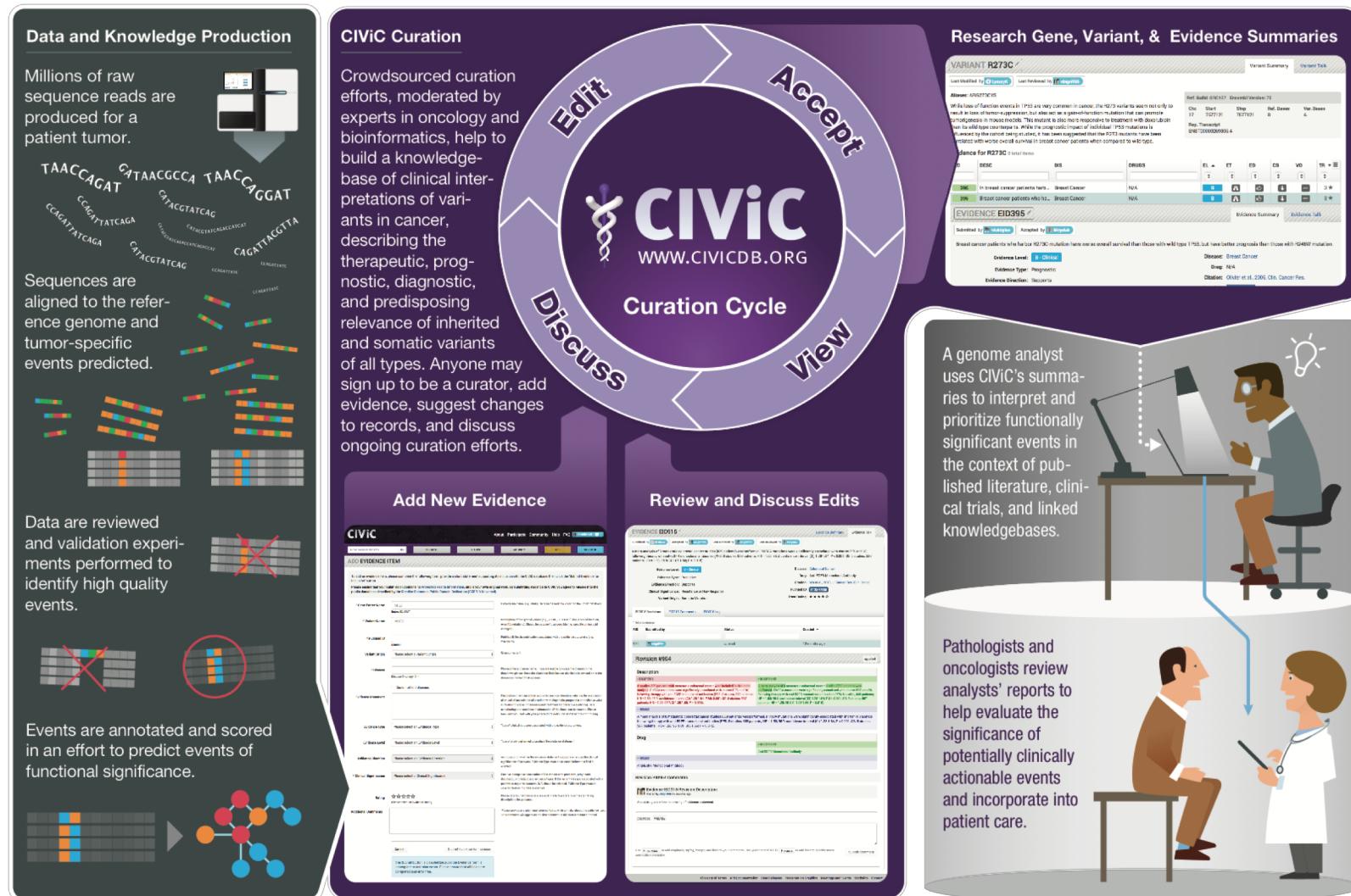
Good BM, Ainscough BJ, McMichael JF, Su AI, Griffith OL. 2014. Genome Biology. 15(8):438.

Interpretation and visualization of genomic alterations remains the bottleneck for realizing precision medicine



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

We created CIViC to address this need - an open knowledgebase and curation system for clinical interpretation of variants in cancer



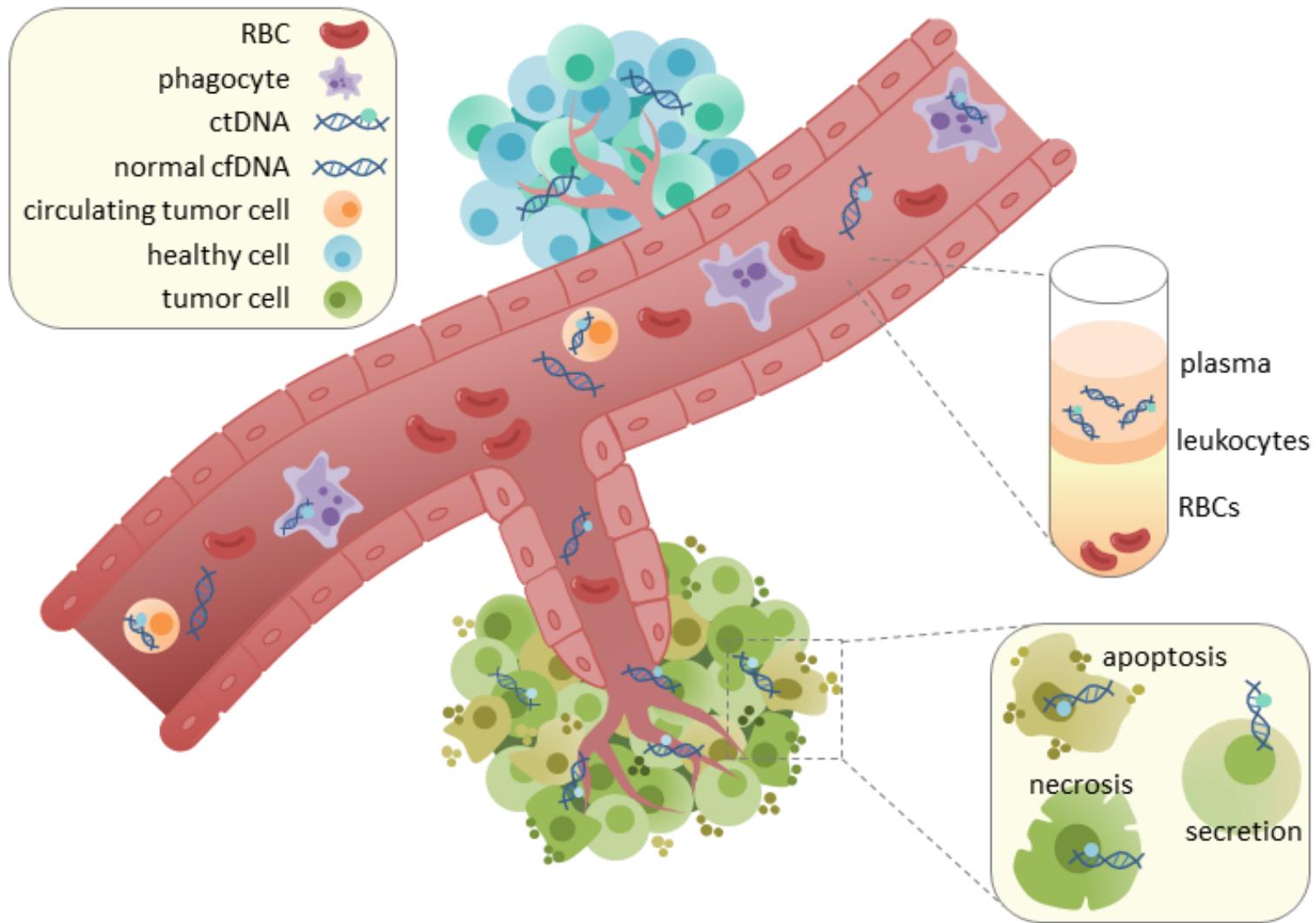
www.civicdb.org

Strategies to bring genomics information to bear for as many cancer patients as possible

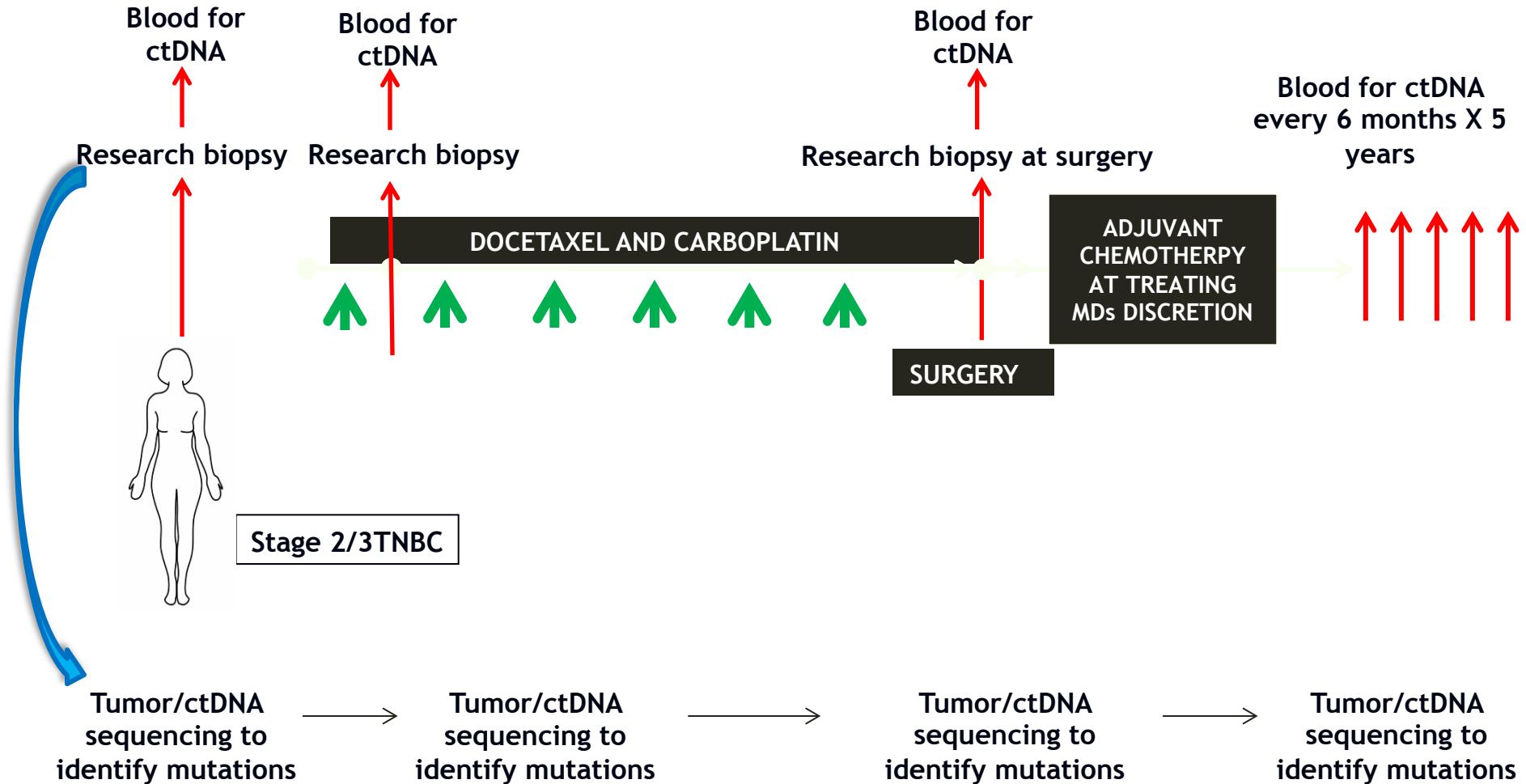
1. Precision medicine targeting of driver mutations

2. Leveraging passenger variants
 - a. Tracking minimal residual disease
 - b. Identifying neoepitopes
 - Predicting response to immunotherapy
 - Developing personalized vaccines

Circulating tumor DNA (ctDNA) could allow generalized tracking in any cancer type



ctDNA tracking in triple negative breast cancer



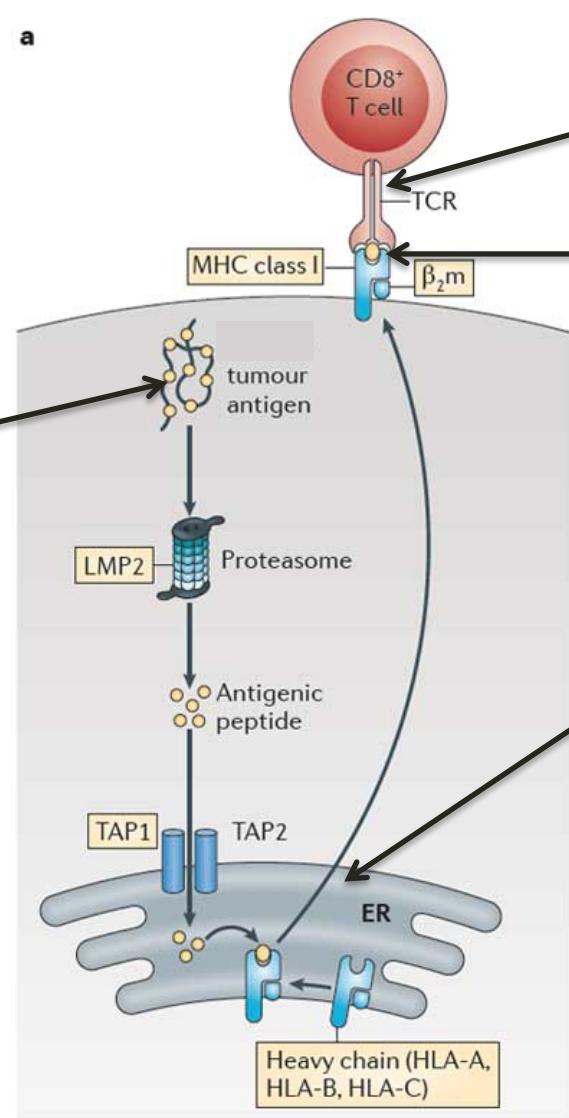
Strategies to bring genomics information to bear for as many cancer patients as possible

1. Precision medicine targeting of driver mutations

2. Leveraging passenger variants
 - a. Tracking minimal residual disease
 - b. Identifying neoepitopes
 - Predicting response to immunotherapy
 - Developing personalized vaccines

Invoking an adaptive immune response against the tumor (focus on CD8+ T cells)

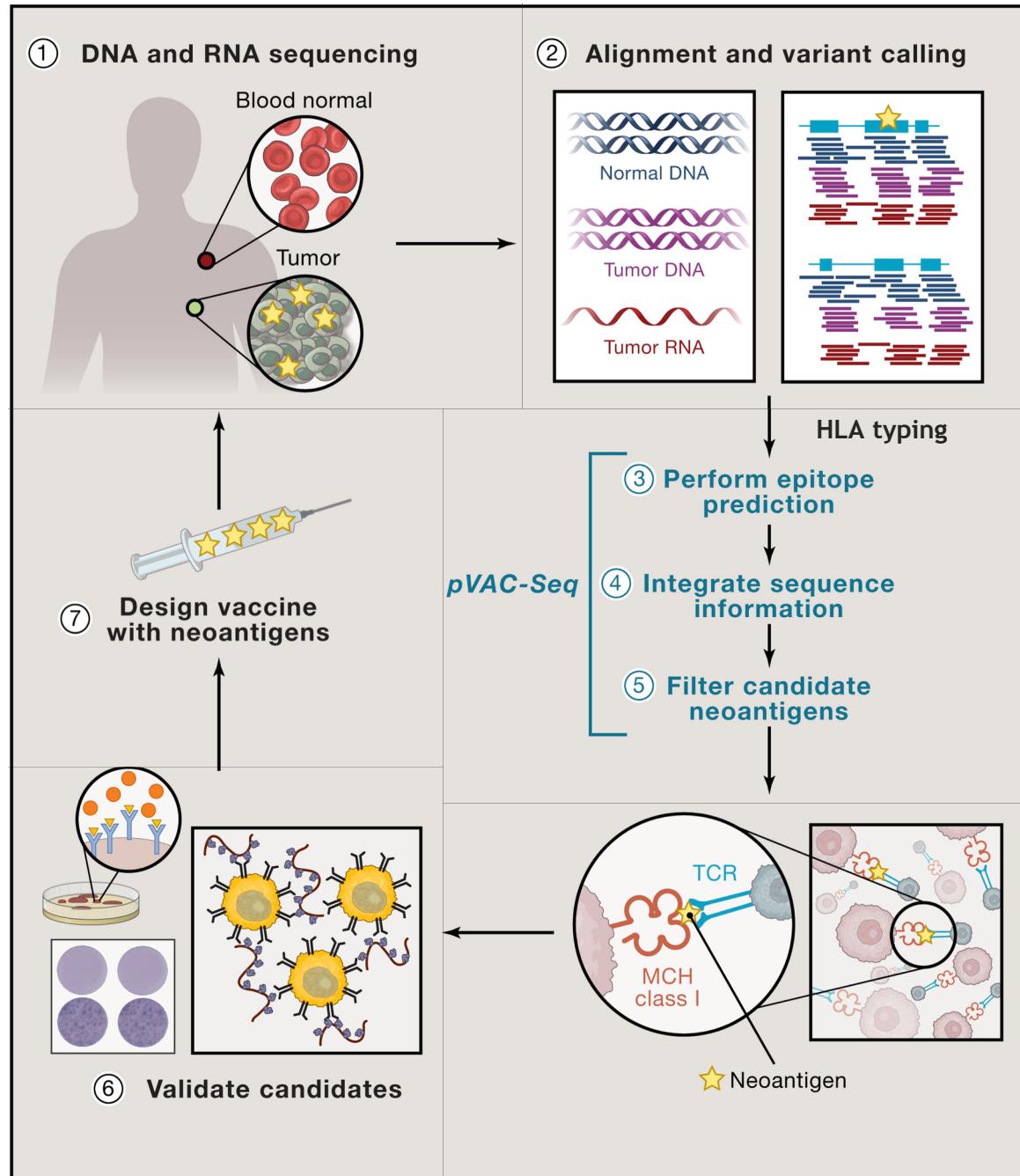
1. Tumor produces a unique peptide corresponding to a somatic mutation



4. T cell receptor that uniquely matches the tumor peptide
3. Neoantigen peptide presented by MHC

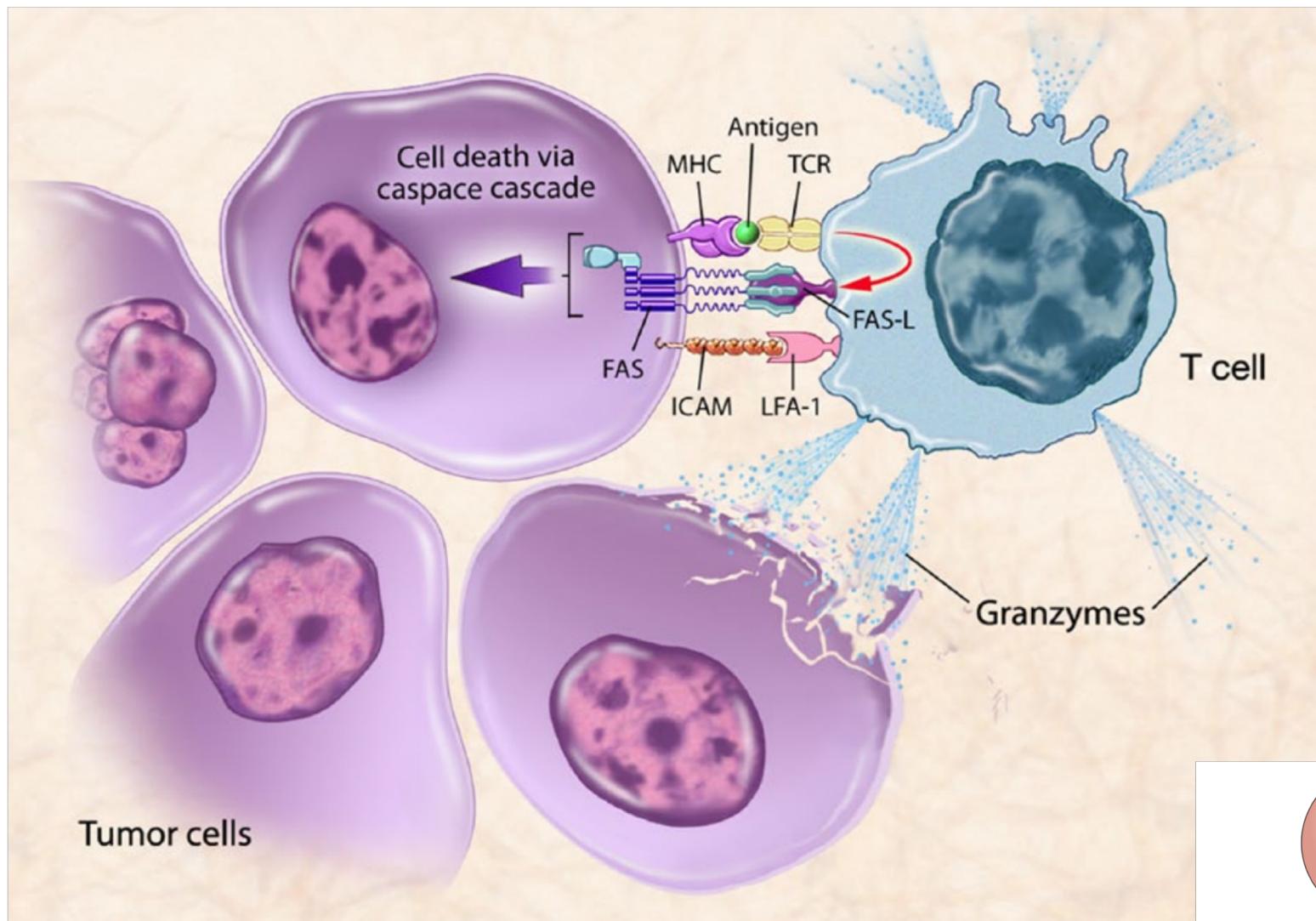
2. Processing and presentation of the tumor specific peptide

Neoepitope characterization workflow



Jasreet Hundal & Katie Campbell
Shirley X. Liu, Elaine R. Mardis.
Applications of Immunogenomics
to Cancer. Cell Press. 2017.

T cell mediated cell death



Personalized cancer vaccine trials

- Kidney Cancer (n = 15 patients)
 - PolyImmune {Durvalumab (MEDI4736) and Tremelimumab} & Vaccine Orchestrated Treatment for Patients With Advanced/Metastatic Renal Cell Carcinoma ([NCT03598816](#)). Collaboration with MedImmune.
- Lung Cancer (n = 15 patients)
 - A Personal Cancer Vaccine (NEO-PV-01) With Pembrolizumab and Chemotherapy for Patients With Lung Cancer ([NCT03380871](#)). Collaboration with Neon Therapeutics and Merck.
- Glioblastoma (n = 30 patients)
 - Neoantigen-based Personalized Vaccine Combined With Immune Checkpoint Blockade Therapy in Patients With Newly Diagnosed, Unmethylated Glioblastoma ([NCT03422094](#))
- Breast Cancer (n = 54 patients)
 - Neoantigen DNA Vaccine Alone vs. Neoantigen DNA Vaccine Plus Durvalumab in Triple Negative Breast Cancer Patients Following Standard of Care Therapy ([NCT03199040](#)). Collaboration with MedImmune.
 - Safety and Immunogenicity of a Personalized Polyepitope DNA Vaccine Strategy in Breast Cancer Patients With Persistent Triple-Negative Disease Following Neoadjuvant Chemotherapy ([NCT02348320](#))
- Follicular Lymphoma (n = 20 patients)
 - Personalized Tumor Vaccine Strategy and PD-1 Blockade in Patients With Follicular Lymphoma ([NCT03121677](#)). Collaboration with Bristol-Myers Squibb.
- Prostate Cancer (n = 20 patients)
 - Neoantigen DNA Vaccine in Combination With Nivolumab/Ipilimumab and PROSTVAC in Metastatic Hormone-Sensitive Prostate Cancer ([NCT03532217](#)). Collaboration with Bristol-Myers Squibb.
- Pancreatic Cancer (n = 15 patients)
 - Neoantigen DNA Vaccine in Pancreatic Cancer Patients Following Surgical Resection and Adjuvant Chemotherapy ([NCT03122106](#)).
- Melanoma (n = 12 patients)
 - Dendritic Cell Vaccination in Patients With Advanced Melanoma ([NCT03092453](#)). Collaboration with UPenn/Parker ICI.

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 search bar, 'Pull requests', 'Issues', 'Gist', and a user profile icon. Below the header, the organization's name 'The Griffith Lab' is displayed along with a photo of two men and a brief description: 'Academic Lab of Obi and Malachi Griffith'. A link to their website is provided. The main content area shows three repository cards: 'dgi-db', 'pVAC-Seq', and 'GenVisR'. Each card includes the repository name, a brief description, a green waveform graphic, and a list of metrics: stars, forks, and last updated time. To the right of the repositories, there are sections for 'Top languages' (R, Python, Ruby, Perl, HTML) and 'People' (a grid of 25 user profiles). A 'New' button is also visible.

- 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>

Advanced Sequencing Technologies and Applications at CSHL

 Cold Spring Harbor Laboratory [ABOUT US](#) [RESEARCH](#) [EDUCATION](#) [PUBLIC EVENTS](#) [NEWSSTAND](#) [PARTNER WITH US](#) [GIVING](#)



Meetings & Courses Program
Crossroads of Biology

[Home](#) [Meetings](#) [Courses](#)

[WELCOME](#) [INFO](#) [APPLICATION](#) [TRAVEL](#) [SPONSORS](#) [PAYMENTS](#) [POLICIES](#)

Advanced Sequencing Technologies & Applications

November 5 - 17, 2019
Application Deadline: July 15, 2019

Instructors:

Obi Griffith, Washington University School of Medicine
Malachi Griffith, Washington University School of Medicine
Elaine Mardis, Nationwide Children's Hospital Research Institute
W. Richard McCombie, Cold Spring Harbor Laboratory
Aaron Quinlan, University of Utah

<https://meetings.cshl.edu/courses.aspx?course=C-SEQTEC&year=19>

Canadian Bioinformatics Workshop Series

2019 Workshop Program

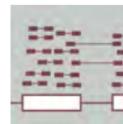
Early Application Fee Deadline for All Workshops: April 12th, 2019



Introduction to R

May 13 - 14, 2019 - Toronto, ON

INSTRUCTORS: Boris Steipe and Lauren Erdman



Informatics for RNA-Seq Analysis

June 11 - 13, 2019 - Toronto, ON

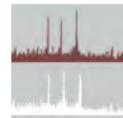
INSTRUCTORS: Obi Griffith, Malachi Griffith, Florence Cavalli, and Brian Haas



Exploratory Analysis of Biological Data using R

May 15 - 16, 2019 - Toronto, ON

INSTRUCTORS: Boris Steipe and Lauren Erdman



Epigenomic Data Analysis

June 18 - 19, 2019 - Montreal, QC

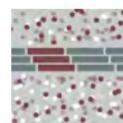
INSTRUCTORS: Guillaume Bourque, Martin Hirst, David Bujold, and Misha Bilenky



Informatics and Statistics for Metabolomics

May 27 - 28, 2019 - Edmonton, AB

INSTRUCTORS: David Wishart and Jeff Xia



Informatics on High-Throughput Sequencing Data

June 20 - 21, 2019 - Montreal, QC

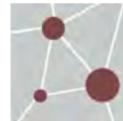
INSTRUCTORS: Mathieu Bourgey and Jared Simpson



Bioinformatics for Cancer Genomics

June 3 - 7, 2019 - Toronto, ON

INSTRUCTORS: Jared Simpson, Trevor Pugh, Francis Ouellette, Sorana Morrissey, Juri Reimand, Lincoln Stein, Robin Haw, Anna Goldenberg, Florence Cavalli, and Mark Phillips



Pathway and Network Analysis of -omics Data

June 26 - 28, 2019 - Toronto, ON

INSTRUCTORS: Gary Bader, Michael Hoffman, Sara Mostafavi, Lincoln Stein, Robin Haw, and Veronique Voisin

<https://bioinformatics.ca/>

Online courses available (griffithlab.org/teaching)

RNA-seq Bioinformatics

Introduction to bioinformatics for RNA sequence analysis

rnabio.org

Genomic Data Visualization and Interpretation

Tutorial series for visualizing and interpreting omic data

genviz.org

Precision Medicine Bioinformatics

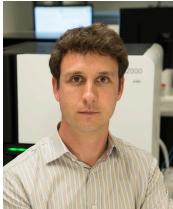
Introduction to bioinformatics for DNA and RNA sequence analysis

pmbio.org

Acknowledgements (griffithlab.org/team/)



Malachi
Griffith



Obi
Griffith



Erica
Barnell



Katie
Campbell



Kaitlin
Clark



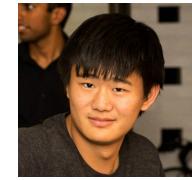
Adam
Coffman



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



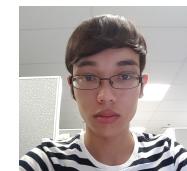
Lee
Trani



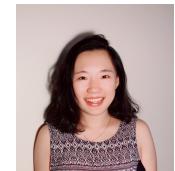
Alex
Wagner



Jason
Walker



Alex
Wollam



Huiming
Xia

Acknowledgements: McDonnell Genome Institute

McDonnell Genome Institute @ Washington University School of Medicine

We are The Genome Institute. Scott Abbott, Travis E. Abbott, Derek C. Albracht, Mark A. Ames, Lucinda L. Antonacci-Fulton, Elizabeth L. Appelbaum, Kelly R. Becker, Michael C. Becker, Edward A. Belter Jr., Paul M. Bender, Angela K. Bentlage, Veena Bhonagiri, Tamberlyn Bieri, Thomas D. Biewald, Andrew P. Bohannon, Kirk A. Brege, Rose L. Brockhouse, Anthony M. Brummett, Mark M. Burnett, Christopher R. Cabanski, Theresa A. Caesar, Matthew B. Callaway, Francine R. Camacho, Edgar A. Campbell, Darren Casimere, Wan-Ching Chan, Lei Chen, Mari Jo Clark, Adam C. Coffman, Lisa L. Cook, Matthew Cordes, Laura P. Courtney, Amy K. D'Albora, John S. Dalbora, Indranil Das, Sara E. Decarlo, Nathan D. Dees, Kimberley D. Delehaunty, Brandon J. Delgado, Anthony M. DeLuca, Tracie L DeLuca, Ryan T. Demeter, Ricardo A. DeMoya, Brian R. Derickson, Li Ding, Kara R. Dix, Gregory P. Dooling, Feiyu Du, James M. Eldred, Efrem Exum, Jason M. Falcone, Candace N. Farmer, Ian T. Ferguson, Grace M. Freeman, Catrina C. Fronick, William L. Fronick, Bradley T. Fulton, Robert S. Fulton, Kristi N. Futhey, Xin Gao, Barbara Gillam, Chakravarthy Girda, Jennifer Godfrey, Jenna Christine Goeckner, Laura Jane Gottschalk, Susan D. Grasso, Malachi Griffith, Obi Griffith, Jeanne M. Grigsby, David S. Gudermann, Priscilla Hale, Terri M. Hall, Kevin V. Haub, Amy E. Hawkins, Todd G. Hepler, Brandi N. Herter, Jennifer S. Hodges, Kimberly Hughes, Jasreet Hundal, Stephanie M. Jackson, Rodney L. Jones, Wendall B. Jones, Krishna L. Kanchi, Cyriac Kandoth, Kimberly Diane Keen, Kyung H. Kim, Michael J. Kiwala, Daniel Koboldt, Sara Kohlberg, Colin L. Kremitzki, Milinn R. Kremitzki, Sheila R. Lakanen, David E. Larson, Sai Harn Lek, Shawn M. Leonard, Shin F. Leong, Andrew D. Levy, Shannon M. Lewis, Timothy J. Ley, Ling Lin, Tina A. Lindsay, Charles Lu, Kaelyn A. Lundry, Amy L. Ly, Sandra K. MacMillan, Vincent J. Magrini, Lenon G. Maguire, Christopher A Maher, Jennifer Maher, Rondy Janjak Malik, Elaine R. Mardis, Christopher M. Markovic, John C. Martin, Judith M. Mc Cart, Suzanne E. McGaugh, Sean D McGrath, Michael D. McLellan II, Joshua F. McMichael, Samantha N. McNulty, Brian Meininger, Karyn N. Meltz Steinberg, Kathie A. Mihindukulasuriya, Lenka Mikalova, Christopher A. Miller, Walter Miller, Patrick J. Minx, Makedonka Mitreva, Deborah S. Moeller, Michael J. Montague, Thomas B. Mooney, Andrew Z. Morrison, David L. Morton, Jennifer B. Murphy, Pamela A. Nangle, William E. Nash, Maze Bi Ndukum Ndonwi, Joanne O. Nelson, Nham Nhan, Beifang Niu, Nathaniel G. Nutter, Benjamin J. Oberkfell, David J O'Brien, Kerri R. Ochoa, Gretchen A. O'Donnell, Michelle D. O'Laughlin, Philip Jeffrey Ozersky, Stephanie J. Parish, Josh Peck, Kymberlie A. Pepin, Stephanie K. Pleasant, Craig S. Pohl, Eric Ponce, Allison Ann Regier, Amy D. Reily, Gaia A. Remerowski, Natalia Rivera, Susan M. Rock, Ronald E. Rodriguez-Santiago, Irina R. Ronko, Bruce A. Rosa, Kelsi M. Rotter, Ryan J. Rupp, Gabriel Eugene Sanderson, Amy P. Sansone, Kyriena L. Schatzkamer, Debra E. Scheer, Heather K. Schmidt, Michael T. Schmidt, Zachary L. Schools, William E. Schroeder, N'Desha T'vorie Scott, Harry E. Senaldi, Cheri-lynn R. Shadding, Nicholas T. Sheehan, Dong Shen, Susanna Siebert, Jessica M. Silva-Fisher, Julie T. Smith, Scott M. Smith, Erica J. Sodergren Weinstock, Steven S. Spargur, John G. Spieth, Gary P. Stiehr, Dennis B. Striegel, Cynthia L. Strong, Dawn K. Sutter, Kenneth B. Swanson, Yat T. Tang, Andrea R. Taylor, Thynn K. Thane, Brenda A Theising, Audra S Thomas-Monti, Brianne M. Thomeczek, Valerie L. Thompson, Chad M. Tomlinson, Lee M. Trani, Nikki L. Trapp, John L. Trevaskis III, Evanne Trevaskis, Richard Trippeer, Rusudan B. Turabelidze, Rahul Tyagi, Joelle M. Veizer, Tammi L. Vickery, Jason E. Waligorski, Jason R. Walker, Patricia E. Wallace, John W. Wallis, Qi Wang, Wesley C. Warren, Brian H. Watts, James V. Weible, Matthew R. Weil, George M. Weinstock, Michael C. Wendt, Richard K. Wilson, Roxanne M. Wilson, Aye M. Wollam, Dee Wu, Kristine M. Wylie, Todd N. Wylie, Kai Ye, Xu Zhang, Yanjiao Zhou

Genomic Data Visualization and Interpretation

- Course goals - by the end of the course we hope you will:
 - Have a solid understanding of basic computational biology techniques for interpreting data
 - Be familiar with a variety of web resources (e.g., genome browsers and portals) to aid in the interpretation of omic data
 - Be able to use R graphical programs for data visualization (e.g., ggplot2, GenVisR) and interactive visualization (e.g., shiny)
 - Be able to perform, visualize, and interpret expression, differential expression and pathway analysis
 - Be able to assess a variant's biological and clinical significance (e.g., VEP, ClinVar, and CIViC)
- What this course does not cover:
 - Upstream genome analysis (alignment, variant calling, assembly, etc, etc)
 - Linux command line and other core bioinformatics skills
- This is a relatively new workshop for us - feedback welcome!

Student introductions

- Please introduce yourself: Name, affiliation and area of research
- 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?
 - Other?
 - What organism do you work with?
 - Did you bring data?

Student poll continued

- 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?