The Elizabeth H. and James S. McDonnell III
McDonnell Genome Institute
at Washington University

# GenViz Module 3: Introduction to GenVisR

Malachi Griffith, Obi Griffith, Zachary Skidmore

Genomic Data Visualization and Interpretation

September 11-15, 2017

Berlin

# Learning objectives of the course

- Module 1: Introduction to genomic data visualization and interpretation
- Module 2: Using R for genomic data visualization and interpretation
- **Module 3: Introduction to GenVisR**
- Module 4: Expression profiling, visualization, and interpretation
- Module 5: Variant annotation and interpretation
- Module 6: Q & A, discussion, integrated assignments, and working with your own data

- Tutorials
  - Provide working examples of data visualization and interpretation
  - Self contained, self explanatory, portable

# Learning objectives of module 3

- Introduction to GenVisR
  - Creation and interpretation of GenVisR plots
    - waterfall()
    - TvTi()
    - genCov()
    - cnFreq()
    - cnSpec()
    - lohSpec()
    - covBars()
  - Adding plot layers to GenVisR
  - Arranging viewports with gridExtra
  - Aligning viewports with gtable

# What is GenVisR?

- Provides a toolkit for visualizing Genomic data
- Implemented in R
- Emphasis on producing “publication quality” graphics with a minimal amount of user input.
- Part of bioconductor
  - https://www.bioconductor.org/packages/3.3/bioc/html/GenVisR.html
- Active development on github
  - https://github.com/griffithlab/GenVisR
- Functions focused in three areas “Small variants”, “Copy number alterations”, and “Data quality”

# Why GenVisR?

- The package is built upon ggplot2 and will allow us to leverage information we've learned in previous modules.
- The package is intended to be flexible supporting multiple file types, species, etc.
- The package is relatively popular; In the top 20% of bioconductor downloads.
- The package is regularly updated with improvements, bug fixes, etc.
- The package is maintained by the griffithlab

# GenVisR::waterfall()

- What does it do?
  - Visualize the types of mutations within a cohort
  - Visualize the mutation burden in a data set
  - Visualize the proportion of samples with a mutated gene
  - Visualize clinical data
- Why do we care?
  - Determine mutually exclusive or co-occurring genomic events
  - Recognize patterns within clinical data
  - See if the mutation burden conforms to expectations

# GenVisR::waterfall()

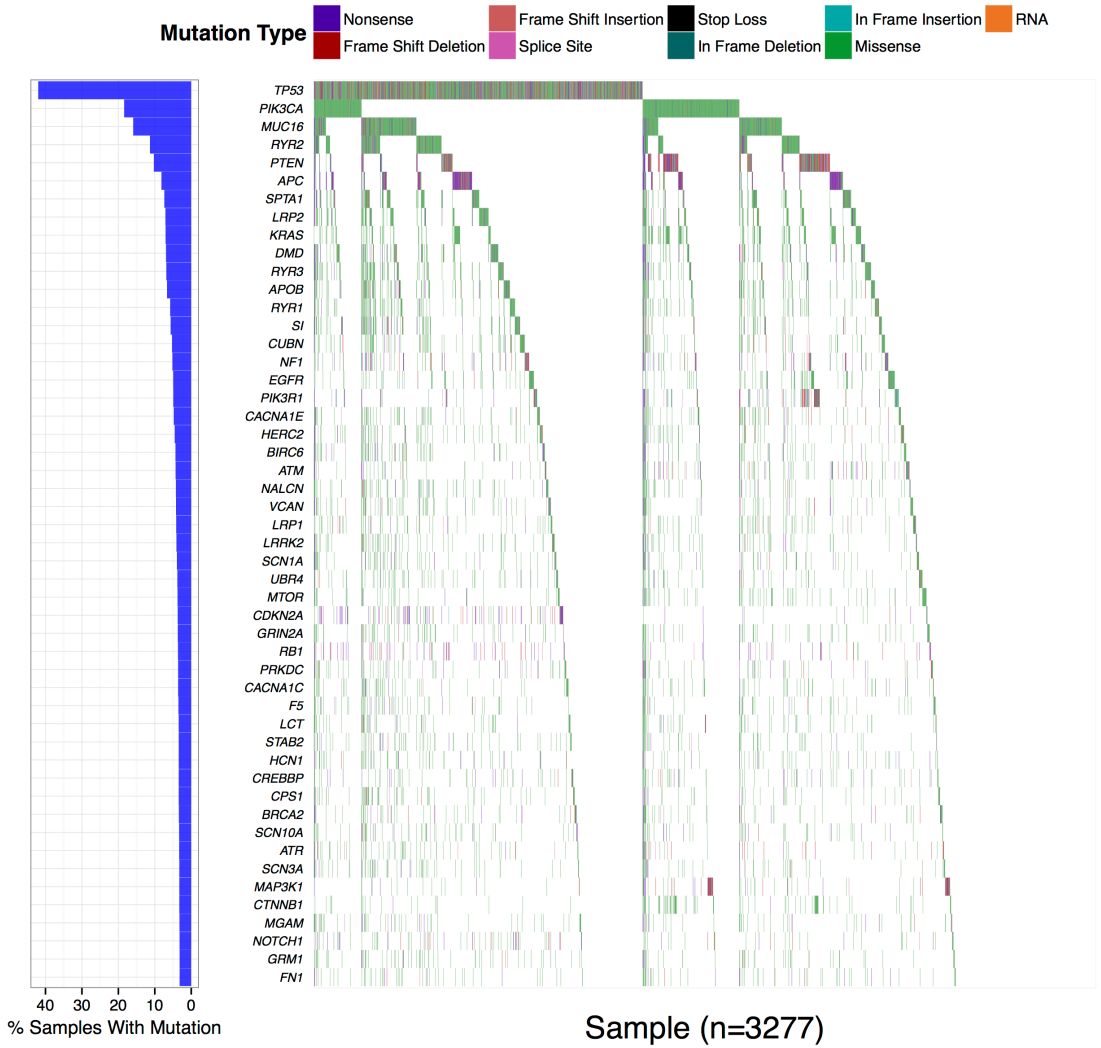



Fig. 2: “DGIdb 2.0: mining clinically relevant drug-gene interactions”

# GenVisR::TvTi()

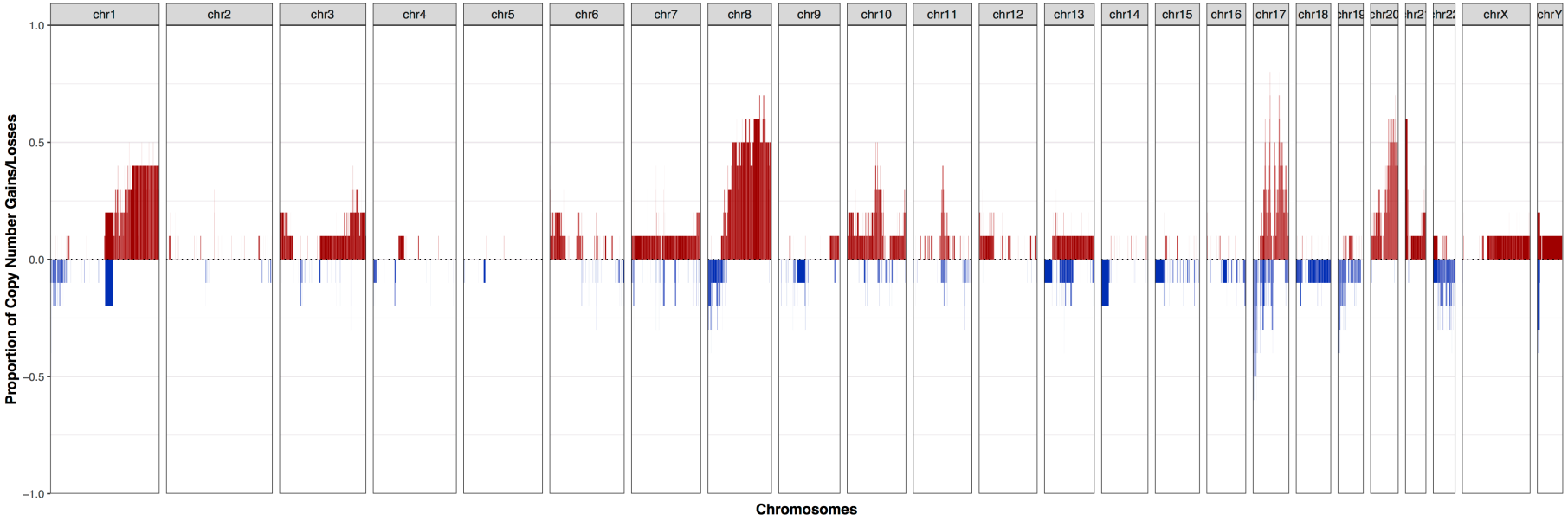
- What does it do?
  - Visualize Transition and Transversion proportions
  - Visualize Transition and Transversion frequencies
  - Compare these proportions with expectations
- Why do we care?
  - Visualize mutation profile patterns
    - Ex. Smoking tends to increase G -> T/C -> A transversions due to oxidative damage
    - Practical application verify smoking status of patients with lung carcinomas

# GenVisR::TvTi()

# GenVisR::cnFreq()

- What does it do?
  - Another view displayed from cnSpec
  - Summarize copy number alterations across a cohort
  - Plots segmented copy number alterations for a cohort
- Why do we care?
  - Easily view recurrently amplified/deleted regions

# GenVisR::cnFreq()

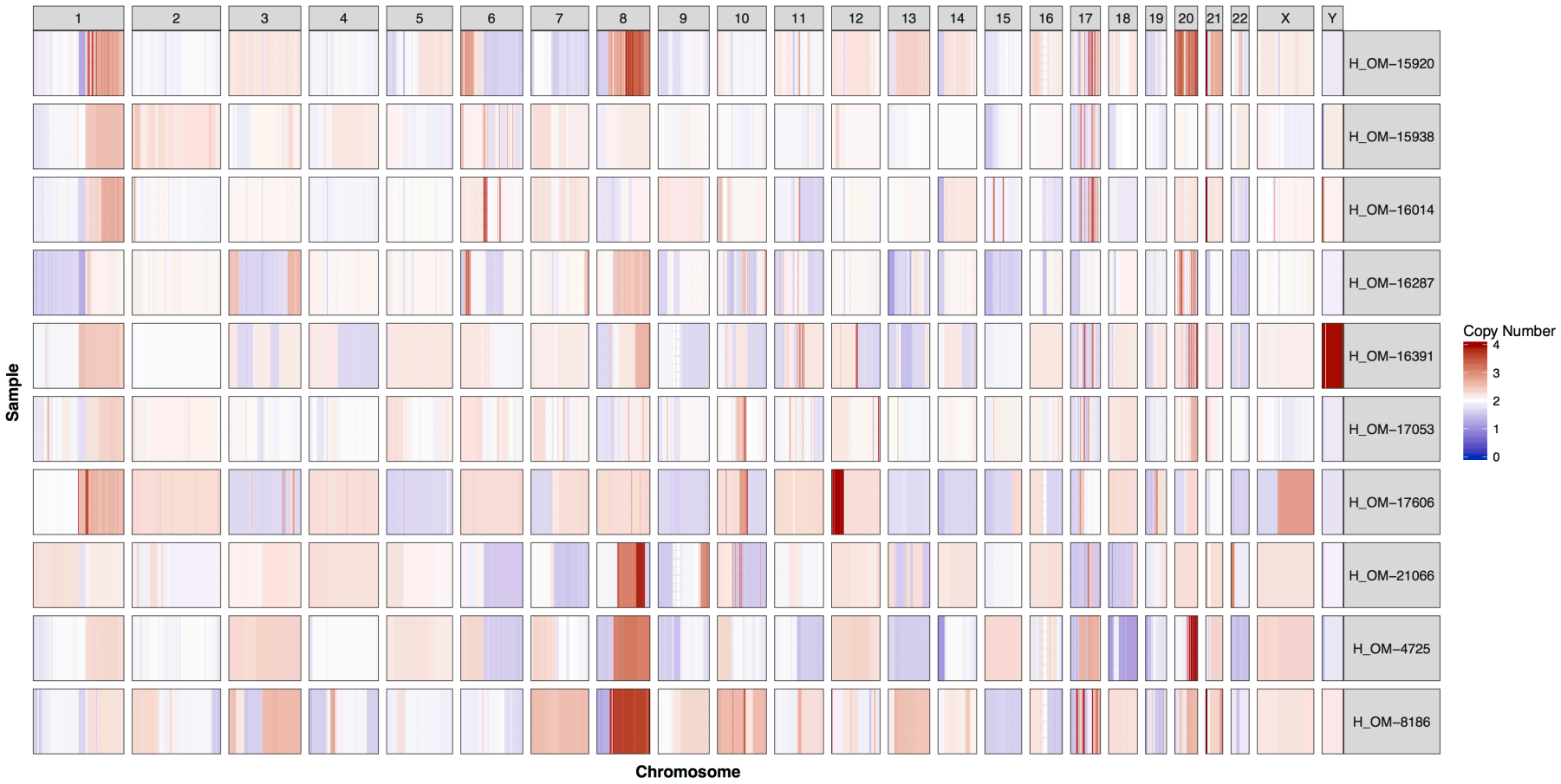



Data from: "Genomic characterization of HER2-positive breast cancer and response to neoadjuvant trastuzumab and chemotherapy-results from the ACOSOG Z1041 (Alliance) trial."

# GenVisR::cnSpec()

- What does it do?
  - Display recurrent copy number alterations
- Why do we care?
  - Recognize patterns of amplifications and deletions while maintaining a sample information

# GenVisR::cnSpec()

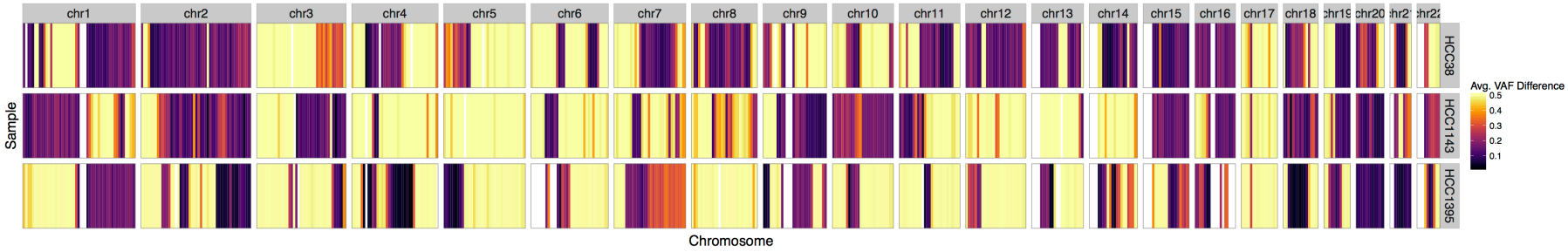



Data from: "Genomic characterization of HER2-positive breast cancer and response to neoadjuvant trastuzumab and chemotherapy-results from the ACOSOG Z1041 (Alliance) trial."

# GenVisR::lohSpec()

- What does it do?
  - Visualize loss of heterozygosity (LOH) for multiple samples
- Why we care
  - Copy-neutral LOH could be missed with a plot produced with functions such as cnSpec however these events are potentially interesting
  - lohSpec allows these events to be viewed and summarized for multiple samples allowing viewing of recurrent LOH patterns

# GenVisR::lohSpec() for 3 breast cancer cell lines



# GenVisR::genCov()

- What does it do?
  - Visualize coverage in a region of interest
- Why do we care?
  - Infer the impact of a structural deletion
  - Verify gene knockouts

# GenVisR::genCov()

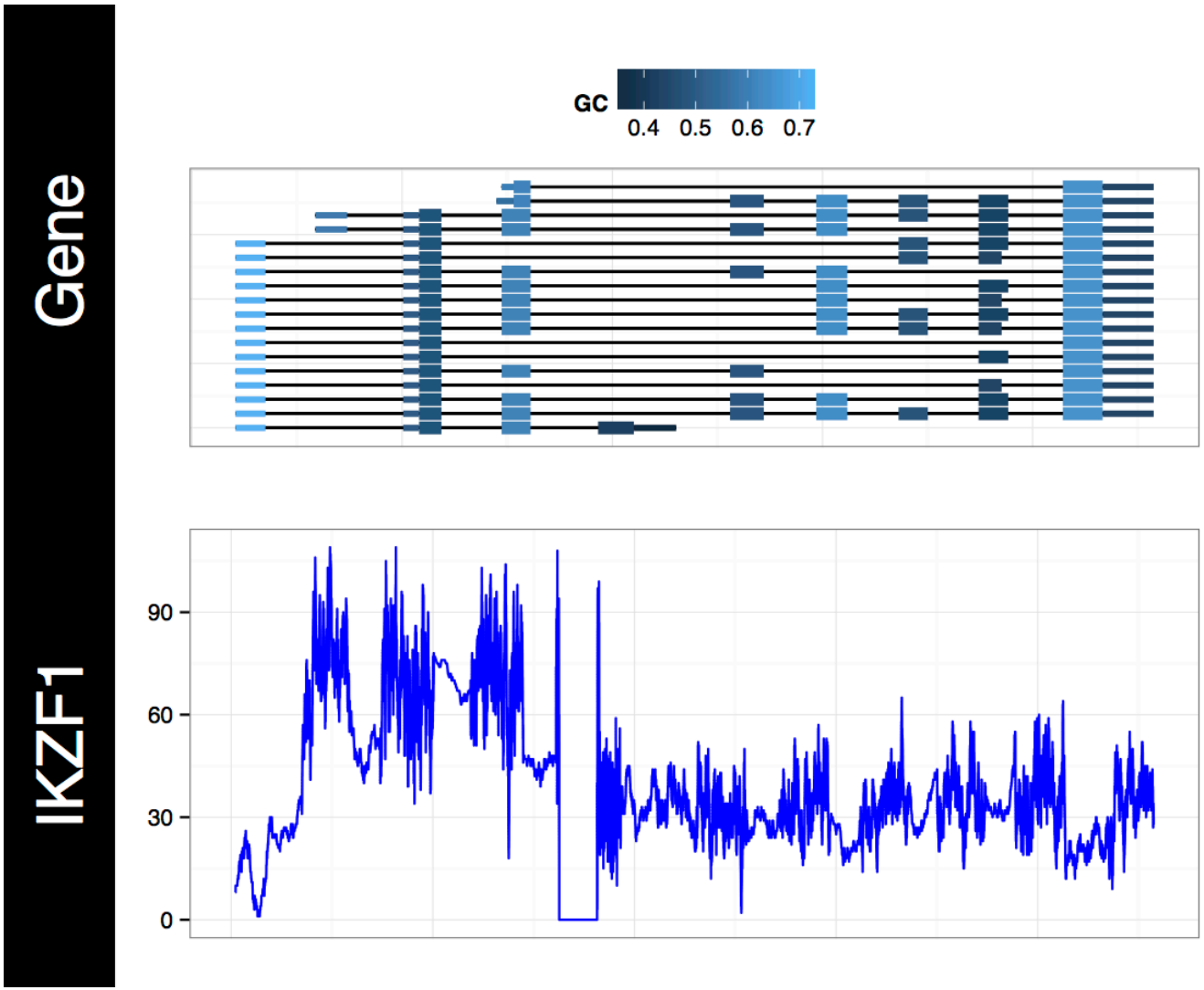
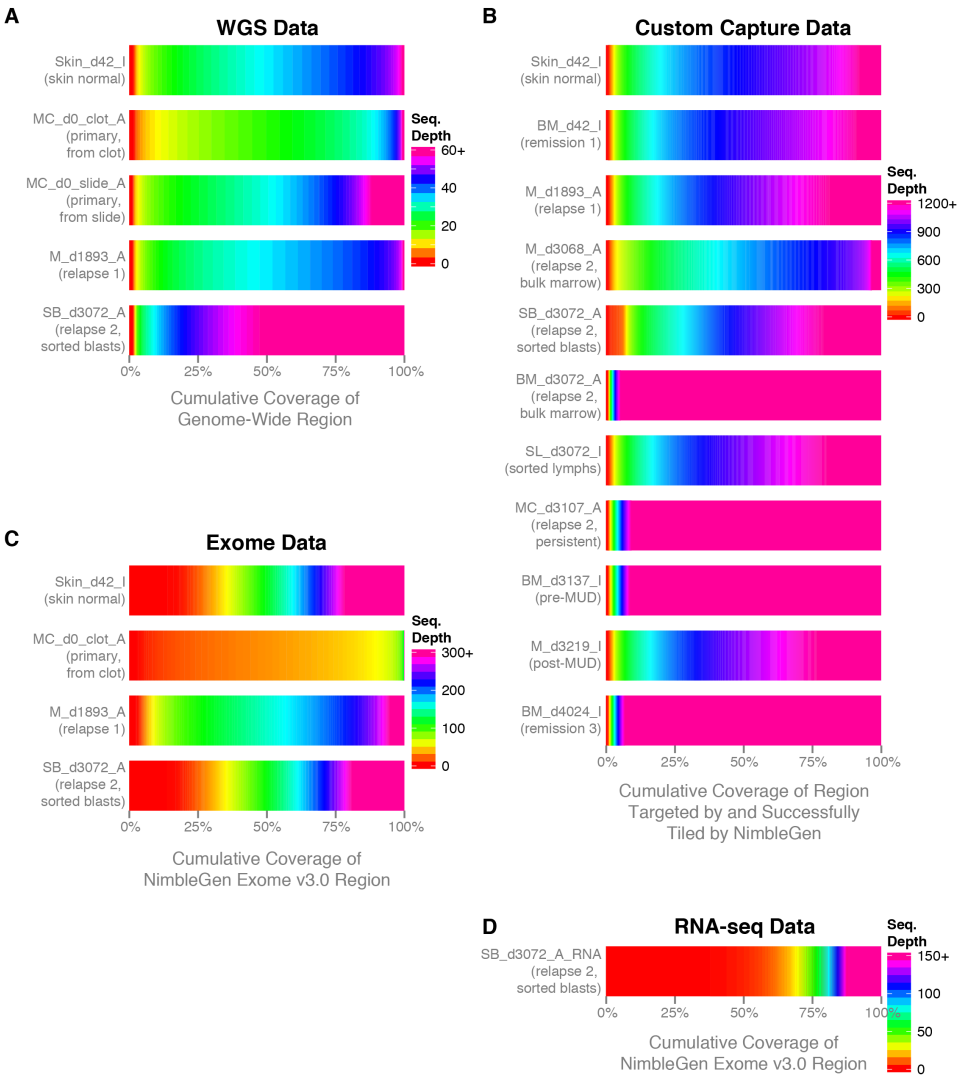



Fig. 2: “Comprehensive genomic analysis reveals FLT3 activation and a therapeutic strategy for a patient with relapsed adult B-lymphoblastic leukemia”

# GenVisR::covBars()

- What does it do?
  - Visualize sequencing coverage achieved across samples
- Why do we care?
  - Immediately identify if a sample has failed sequencing or has not achieved the expected sequencing depth

# GenVisR::covBars()




Supp. Fig. 3: “Comprehensive genomic analysis reveals FLT3 activation and a therapeutic strategy for a patient with relapsed adult B-lymphoblastic leukemia”