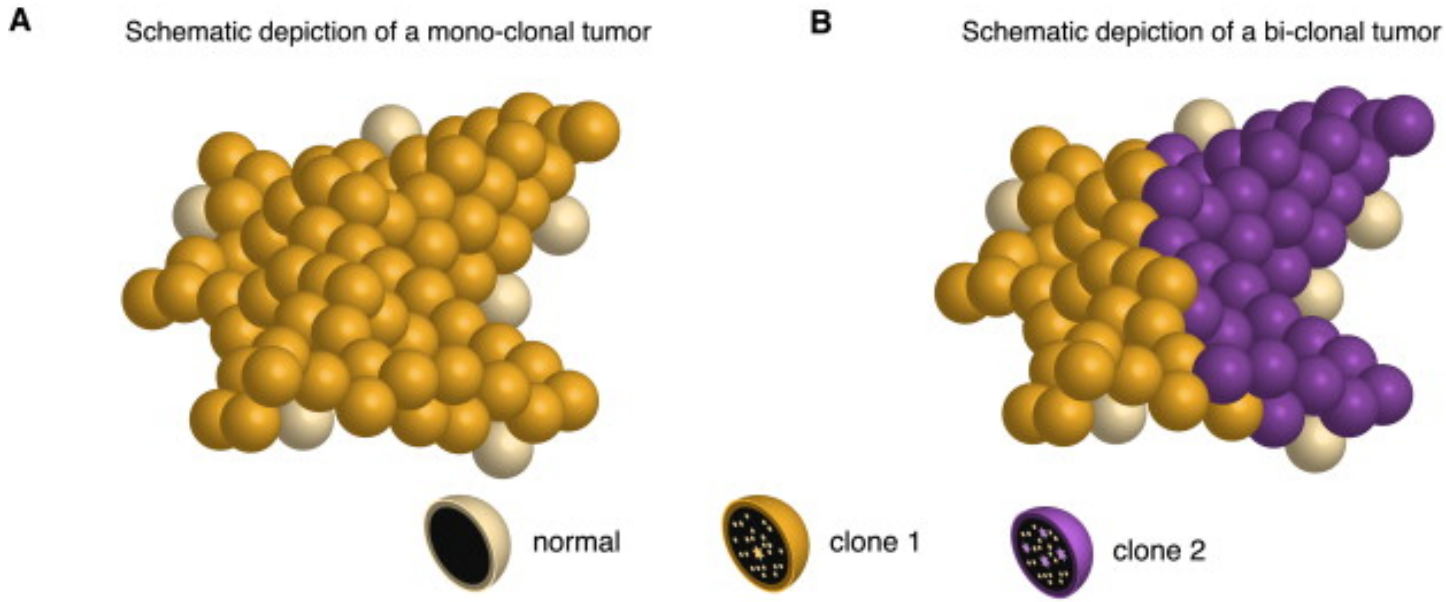
Assessing tumor heterogeneity and tracking clonal evolution using whole genome or exome sequencing

Chris Miller, PhD

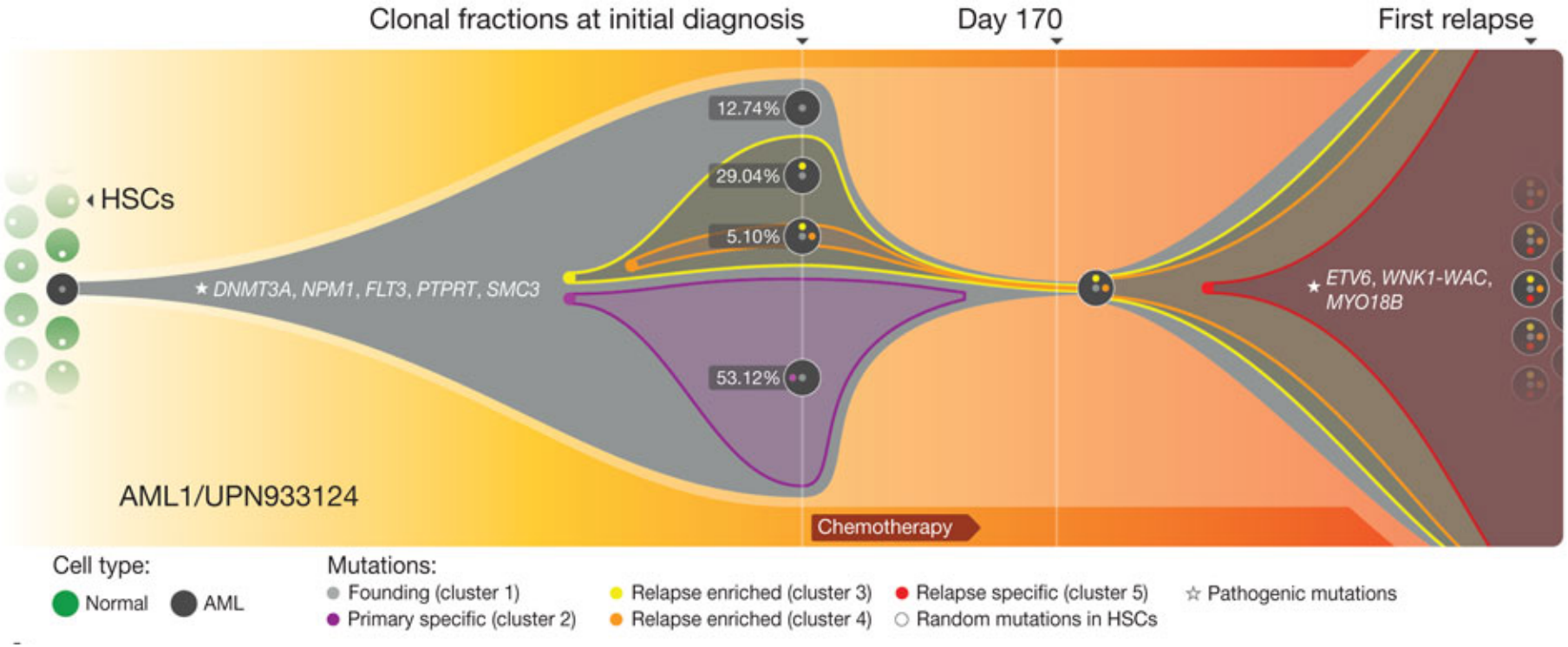
Tumors are heterogeneous



genetically diverse populations of cells

evolution occurs at the cellular level

Clonal evolution in relapsed AML

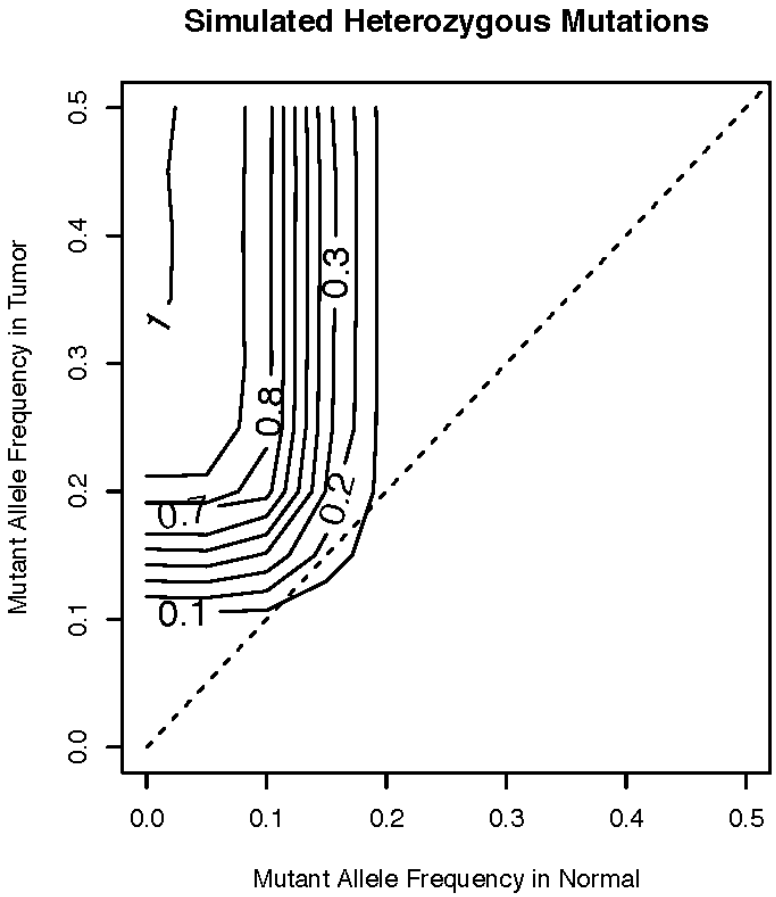


Nature 481, 506–510 doi:10.1038/nature10738

Challenges for detecting minor subclones

- Genomes are sequenced with low coverage
  - 30x not enough
- Algorithms aren’t designed to detect low-frequency events

Somatic Sniper power simulations



- 90x coverage
- Power to detect event at 20% VAF: 85%
- Power to detect event at 10% VAF: 10%

Larson, et al. Bioinformatics (2012) 28 (3): 311-317. doi: 10.1093/bioinformatics/btr665

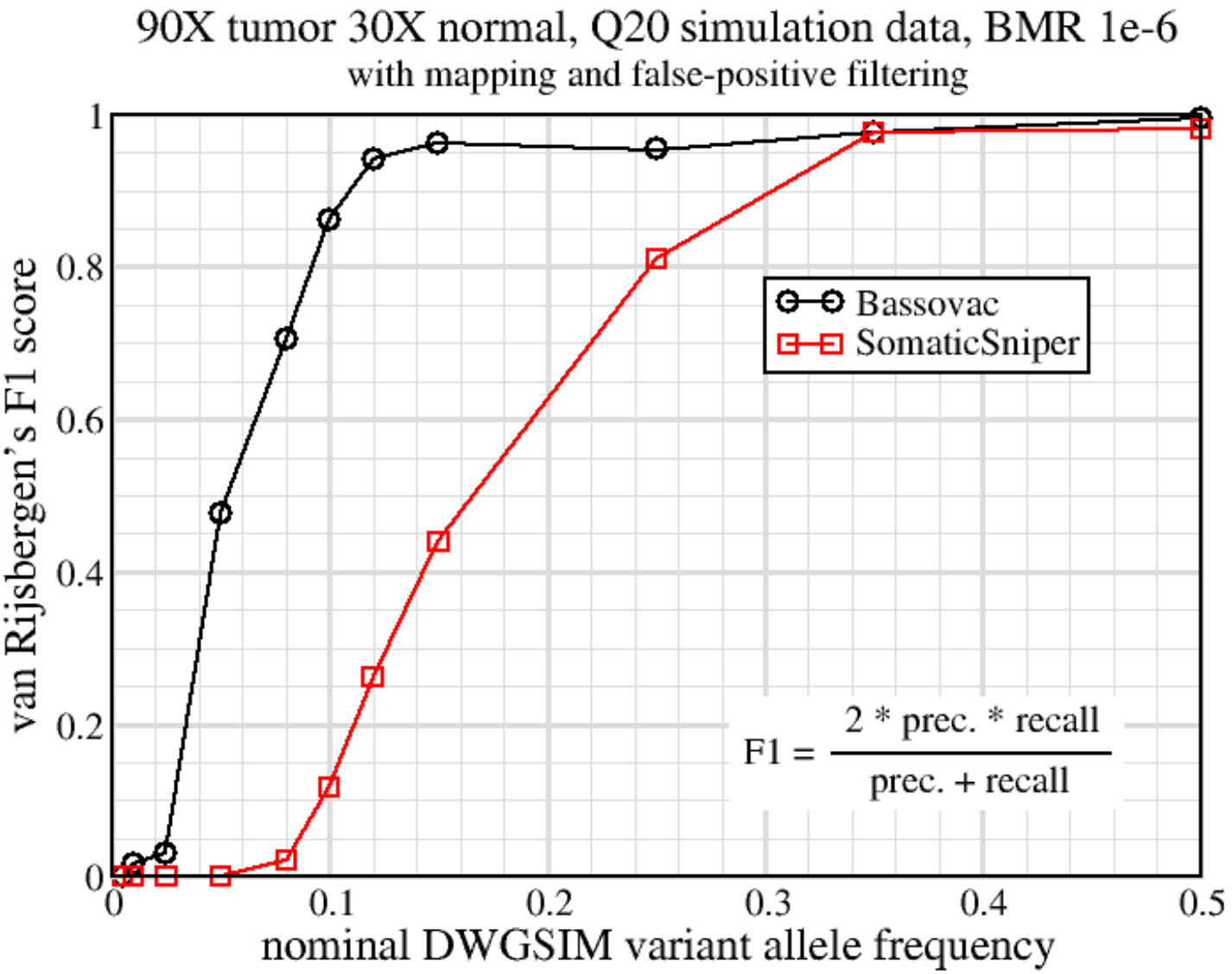
BASSOVAC

- **BA**yesian **S**coring of **S**omatic **VA**riant read **C**ounts

- Incorporates purity, ploidy, base quality, allele frequency, and overall mutation rate.

- Bayesian framework for inversion to obtain, probabilities of heterozygous and homozygous somatic events, given the data

Simulation


90X tumor 30X normal, Q20 simulation data, BMR 1e-6
with mapping and false-positive filtering
van Rijsbergen's F1 score
nominal DWGSIM variant allele frequency
Bassovac
SomaticSniper
F1 = 2 * prec. * recall / prec. + recall
0
0.2
0.4
0.6
0.8
1
0.1
0.3
0.5

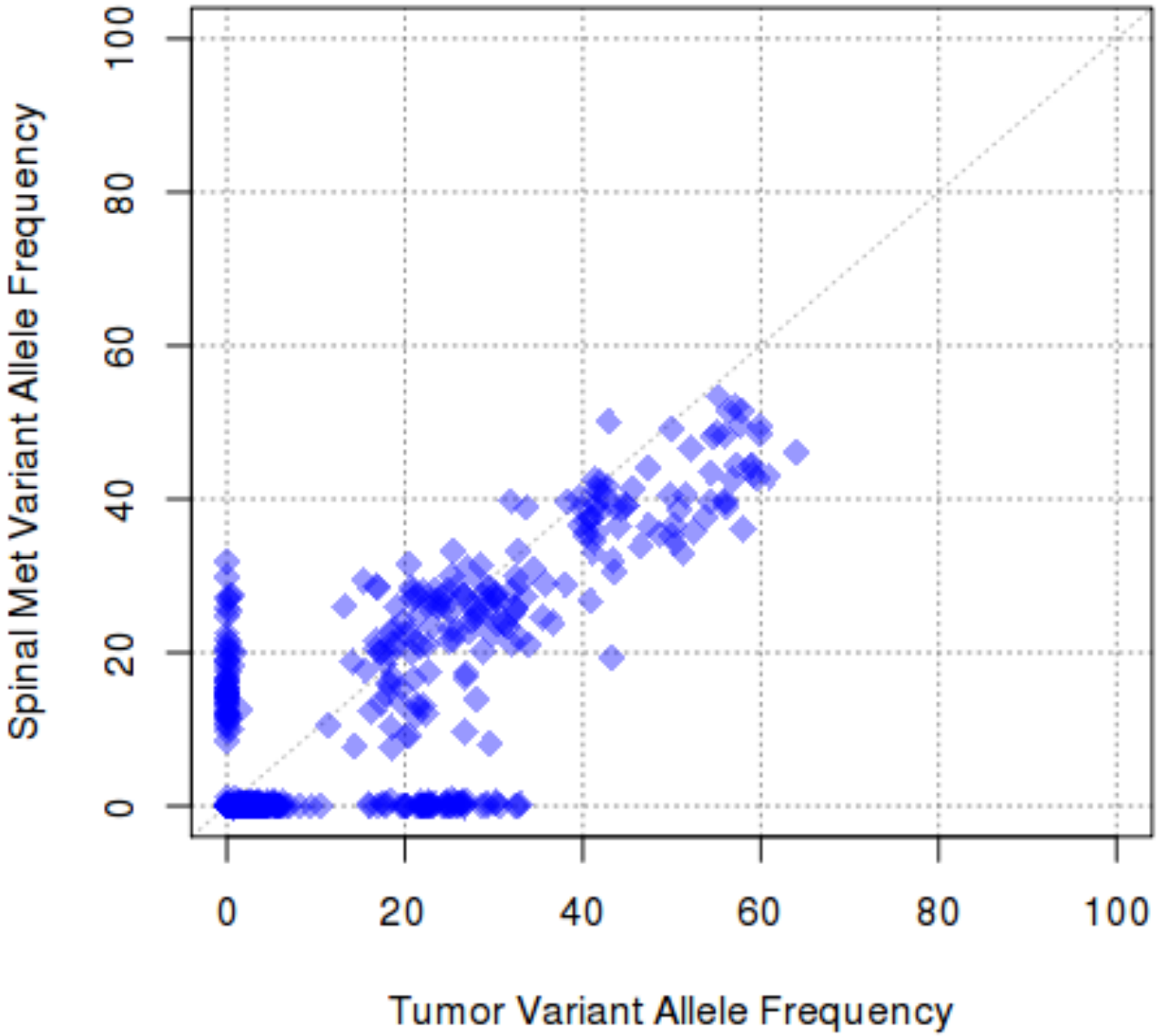
Real-world testing

- Primary breast tumor
- Matched normal
- 3 different metastases:
  - Spinal
  - Liver
  - Adrenal

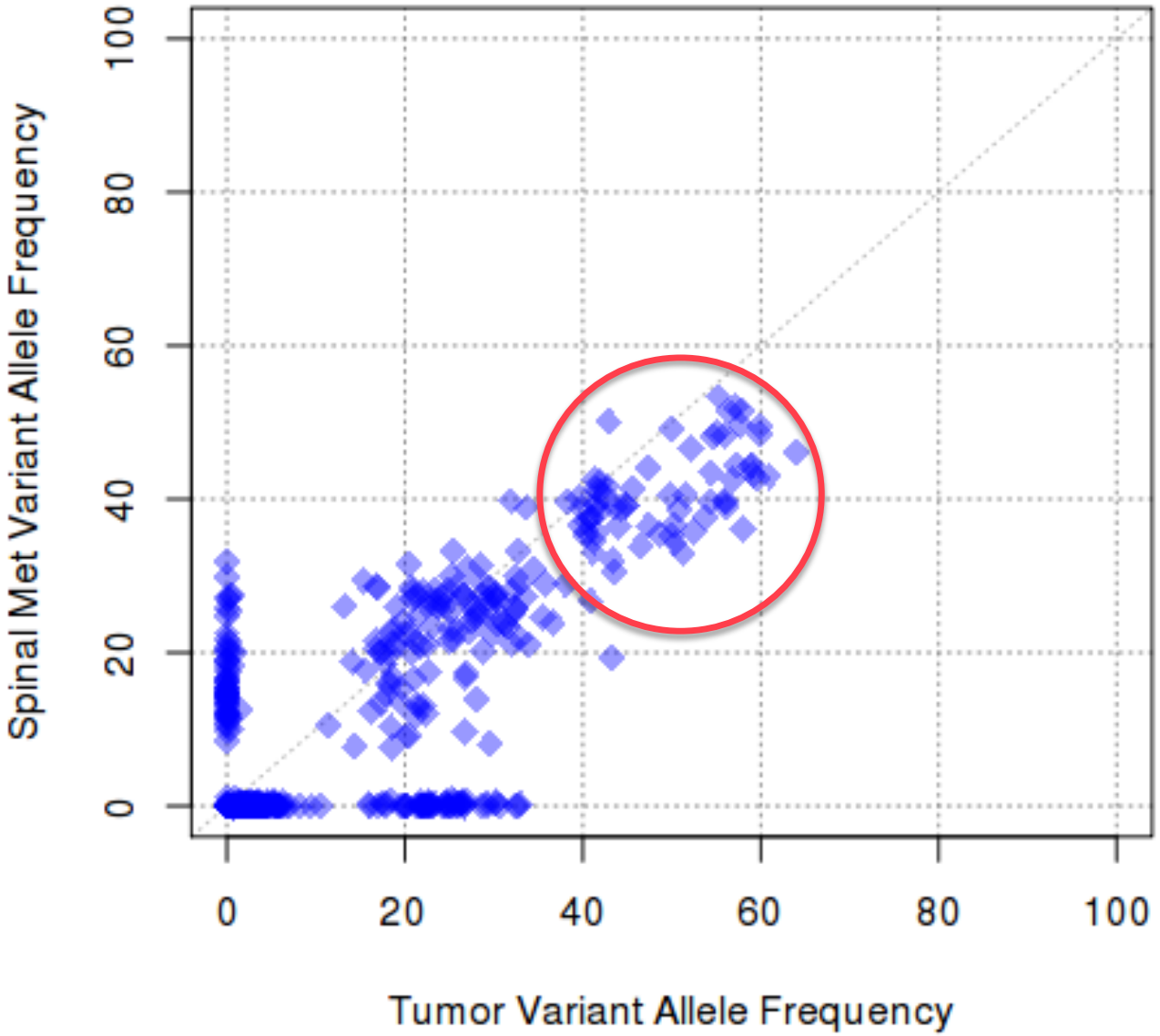
Real-world testing

- All whole-genome sequenced to 30x
- Mutation calls made with Somatic Sniper and Varscan
- Capture validation performed for all variants
- Deep readcounts obtained from validation sequencing for all variants in all samples

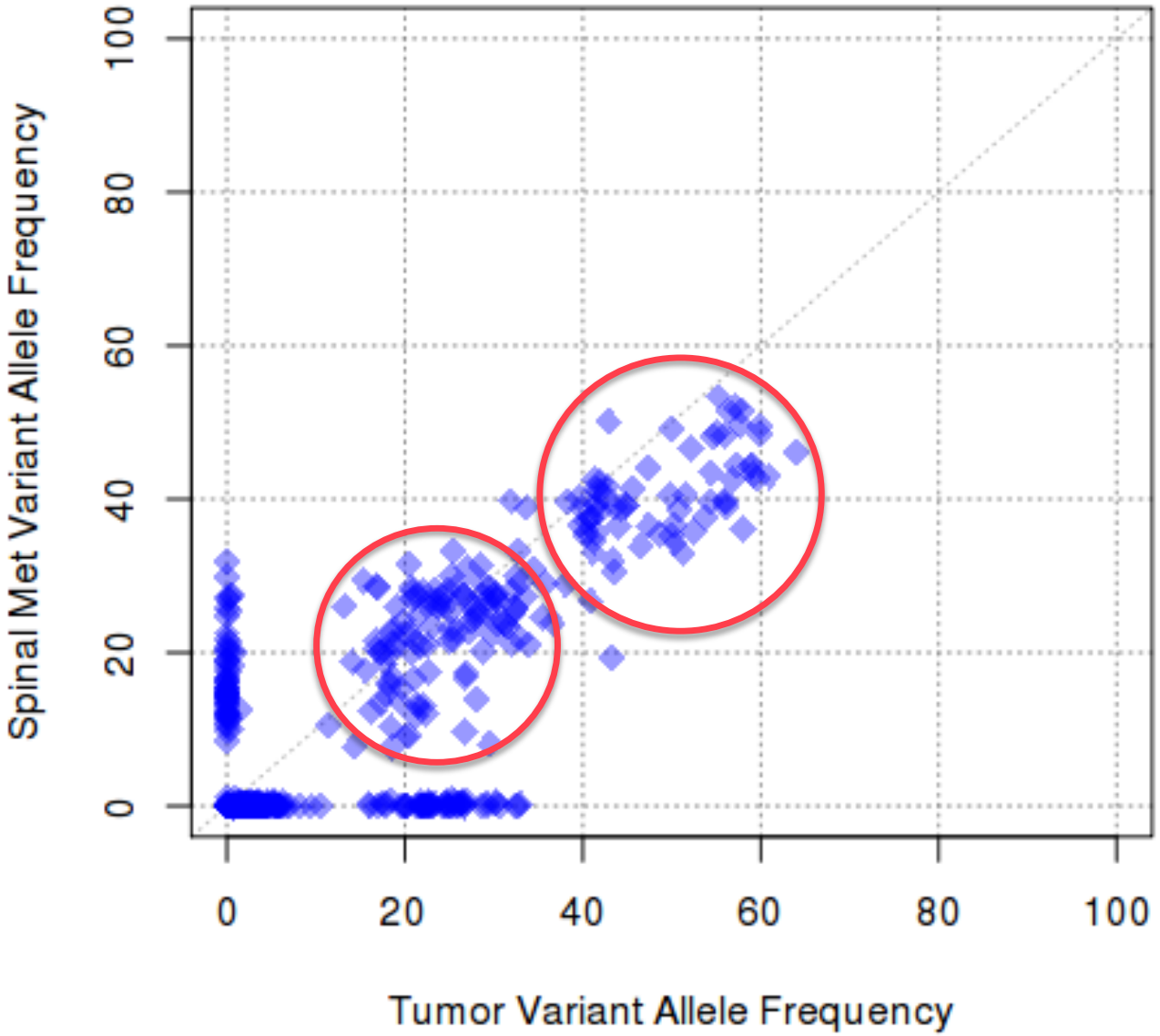
BRCA Tumor vs Spinal Metastasis



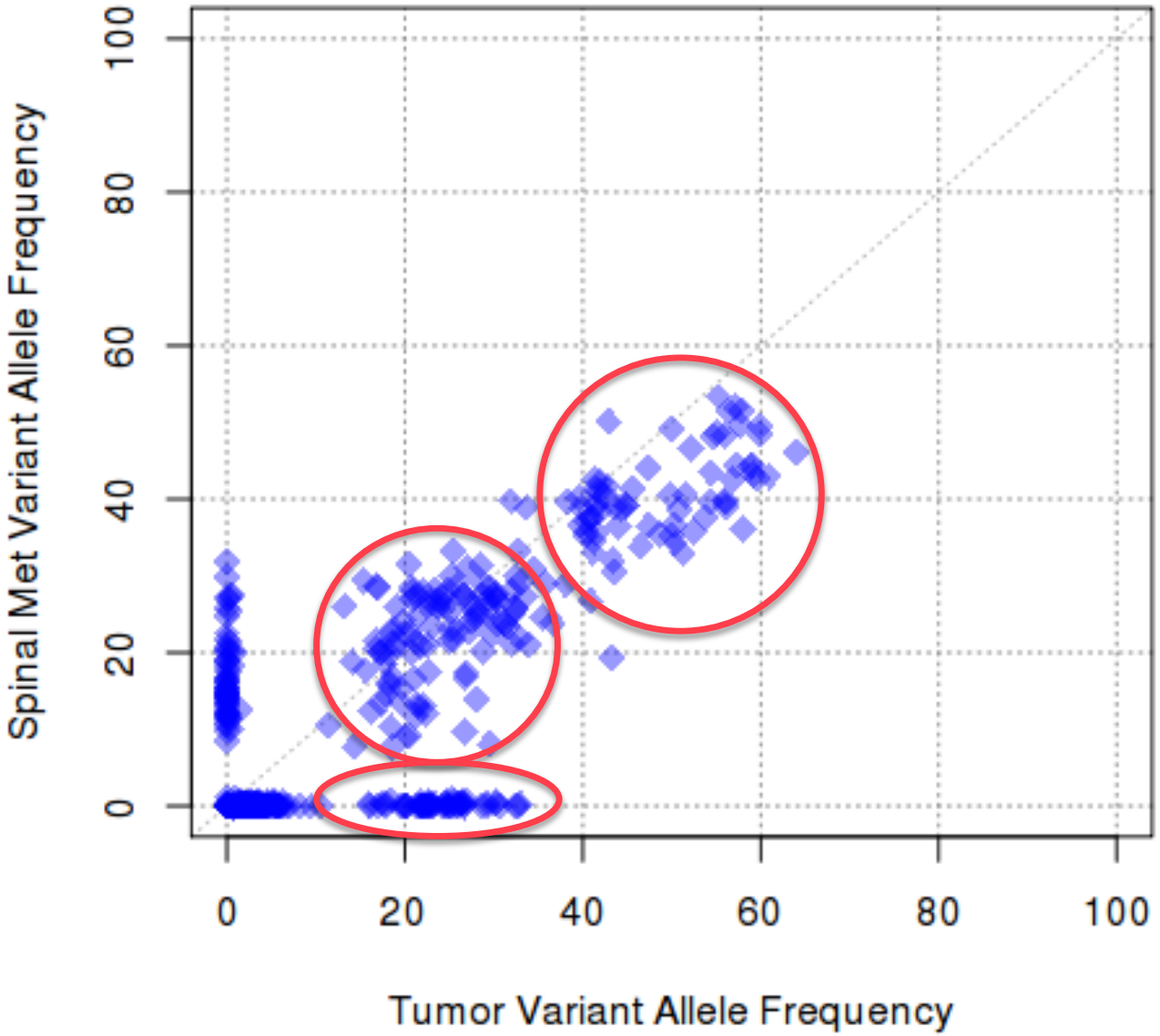
BRCA Tumor vs Spinal Metastasis



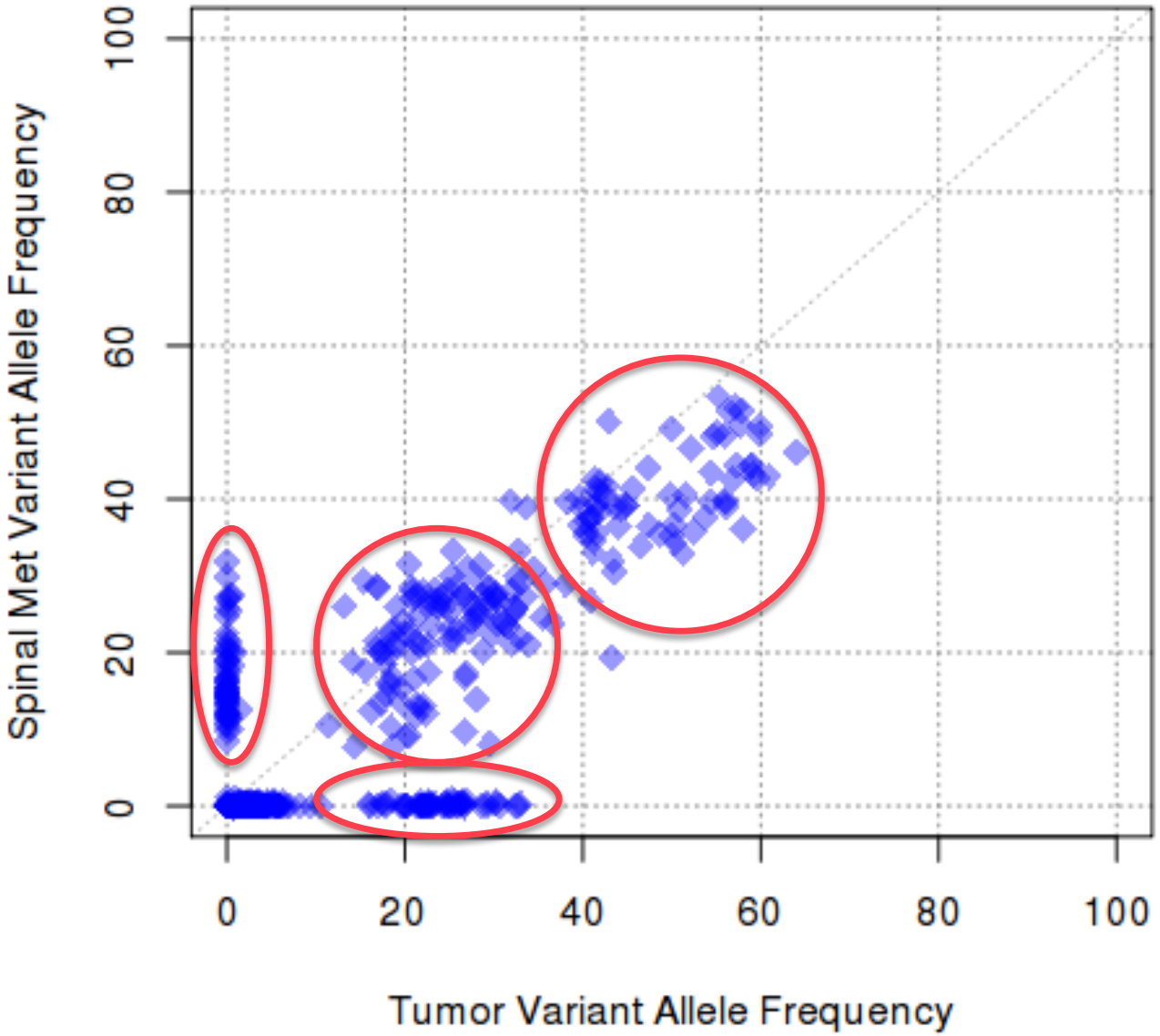
BRCA Tumor vs Spinal Metastasis



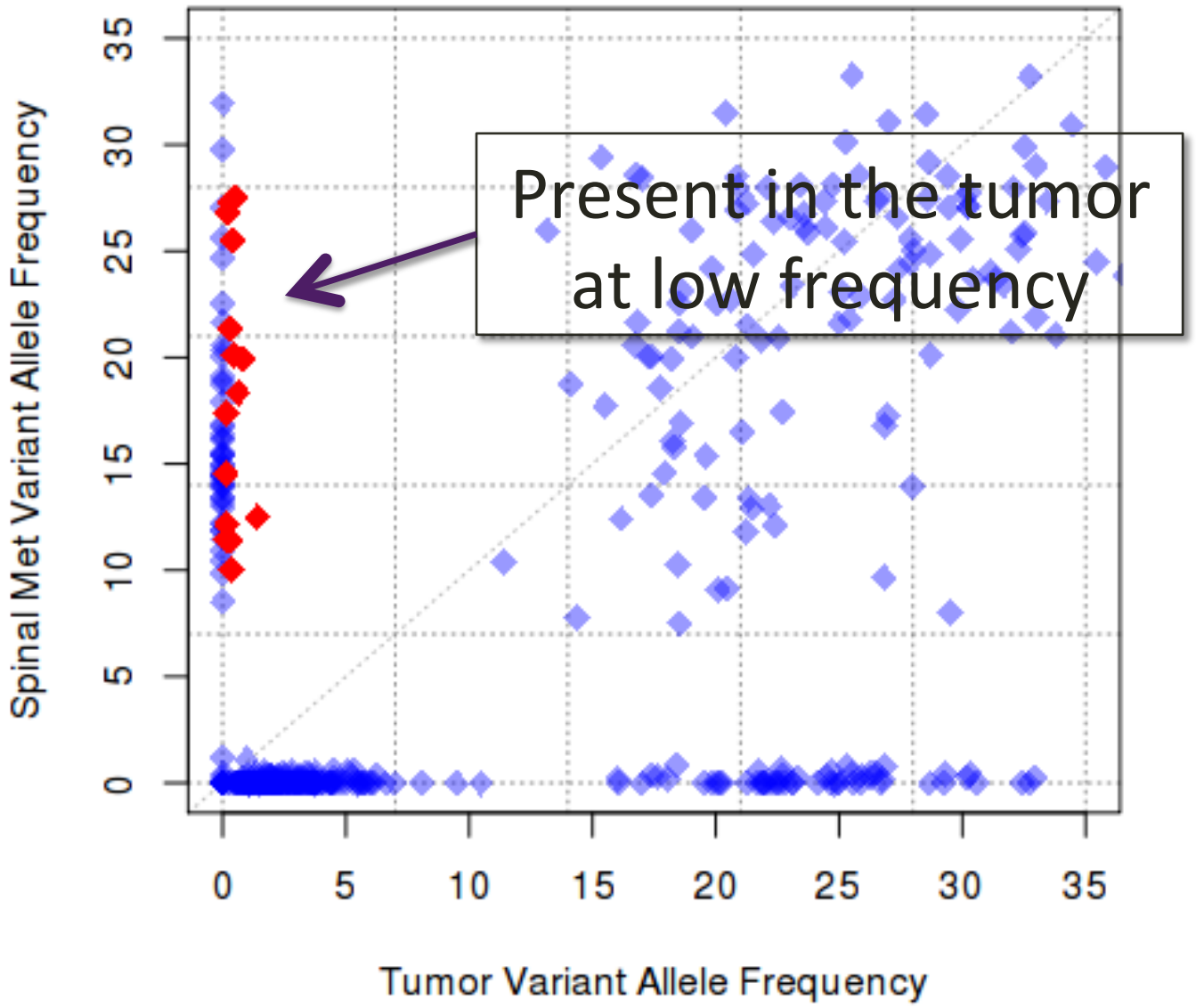
BRCA Tumor vs Spinal Metastasis



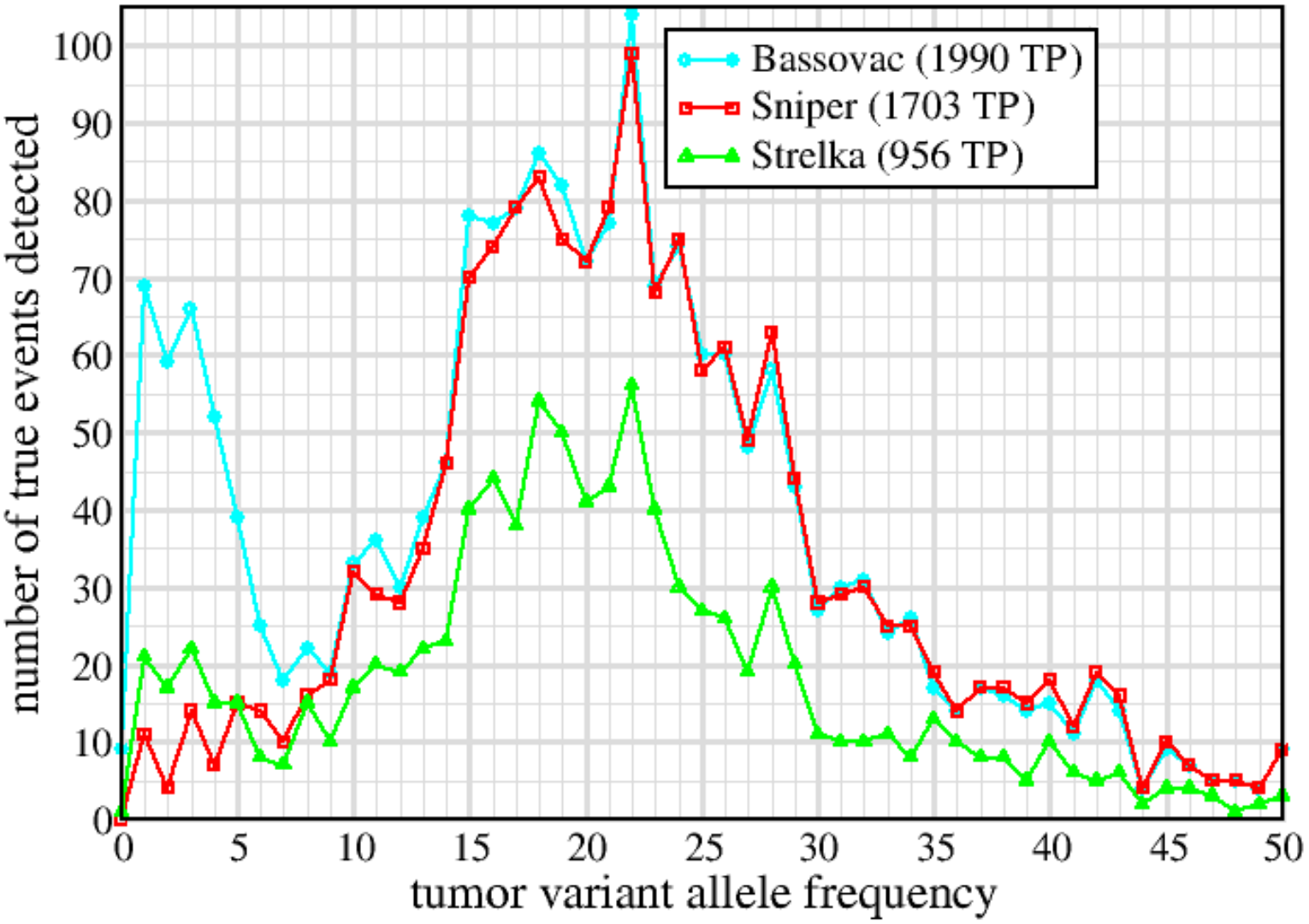
BRCA Tumor vs Spinal Metastasis



BRCA Tumor vs Spinal Metastasis



BASSOVAC sensitivity - BRCA met SNVs in tumor



BASSOVAC

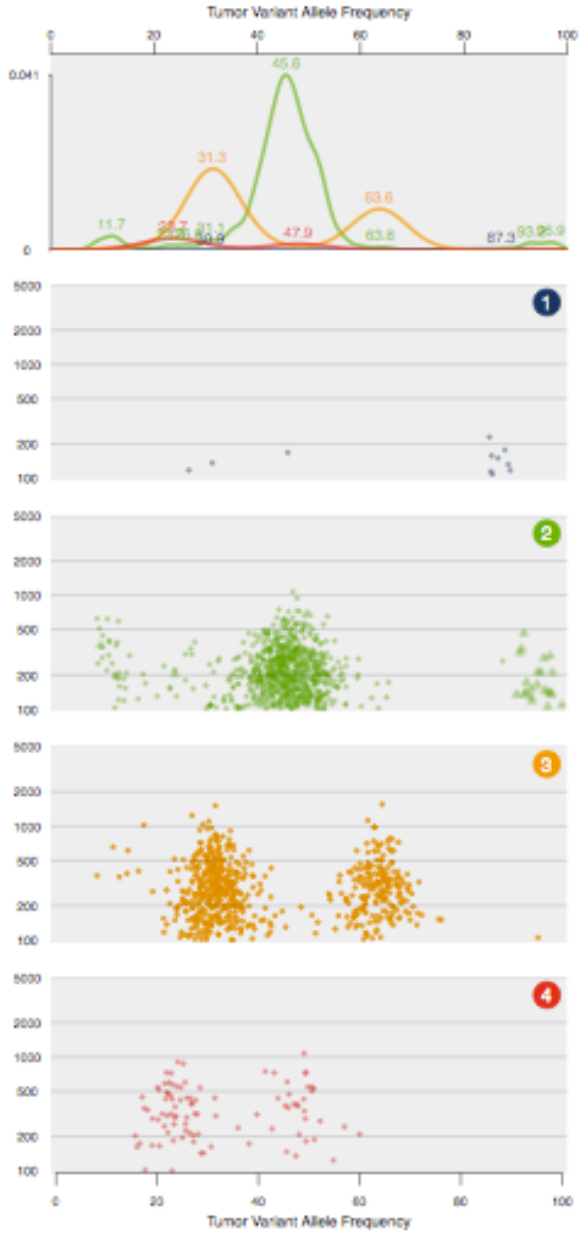
- Over 50% of the variants present in the metastases are present at a detectable level in the tumor

- We can use BASSOVAC to detect true variants at very low frequencies (< 2%)

Clonal inference

- Given information about a tumor, how many clones are present?
  - Which variants are present in different subclones?
- Requires integrative approach
  - Variant allele frequencies
  - Copy number calls
  - Purity and Ploidy information

Clonality Plot

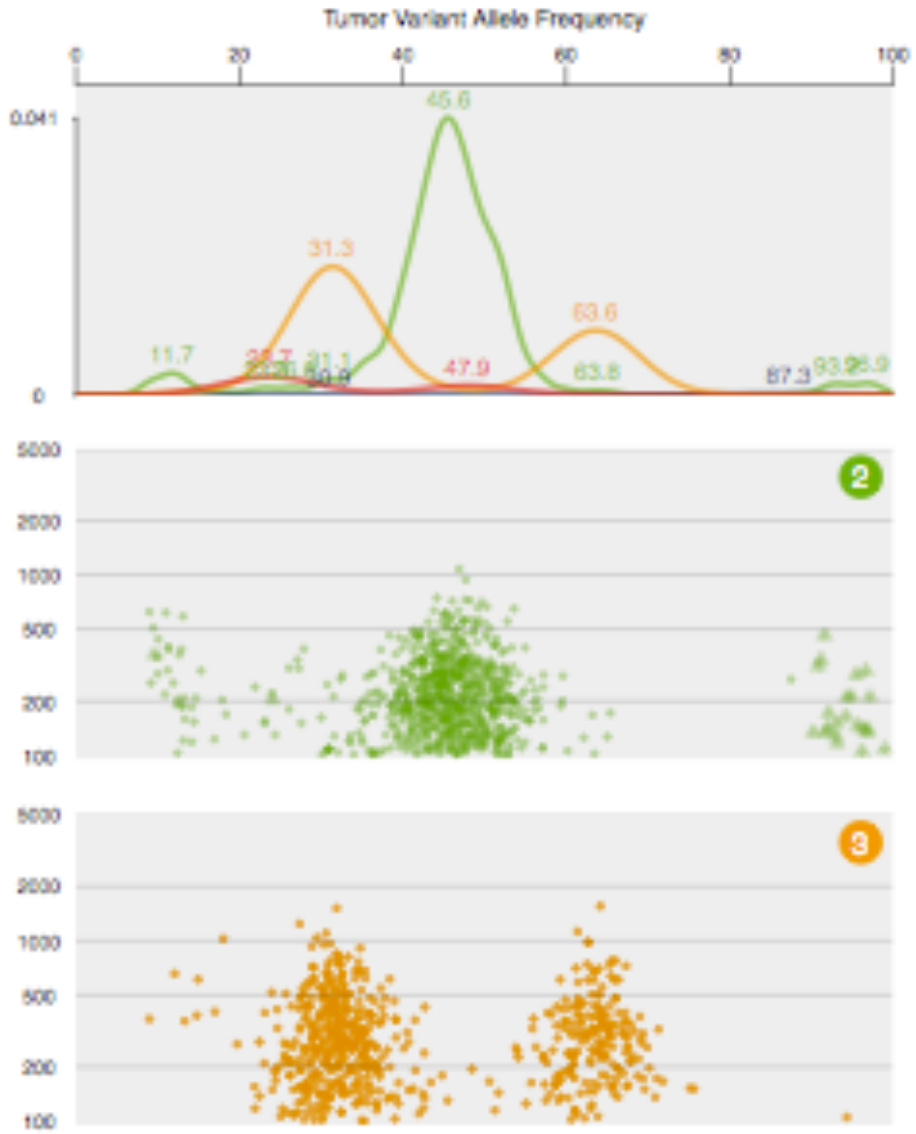


Clonality Plot

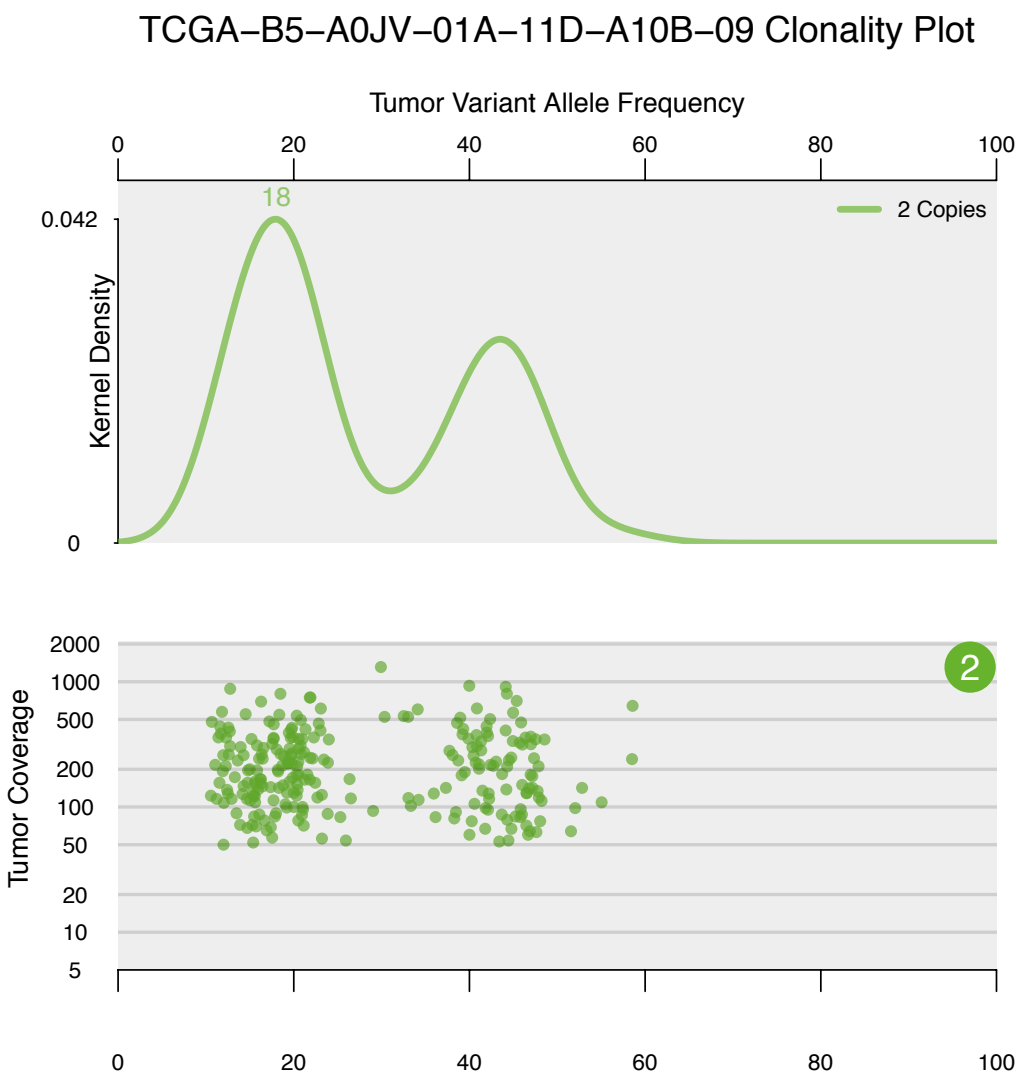
Kernel Density

SNVs in CN2 regions

SNVs in CN3 regions

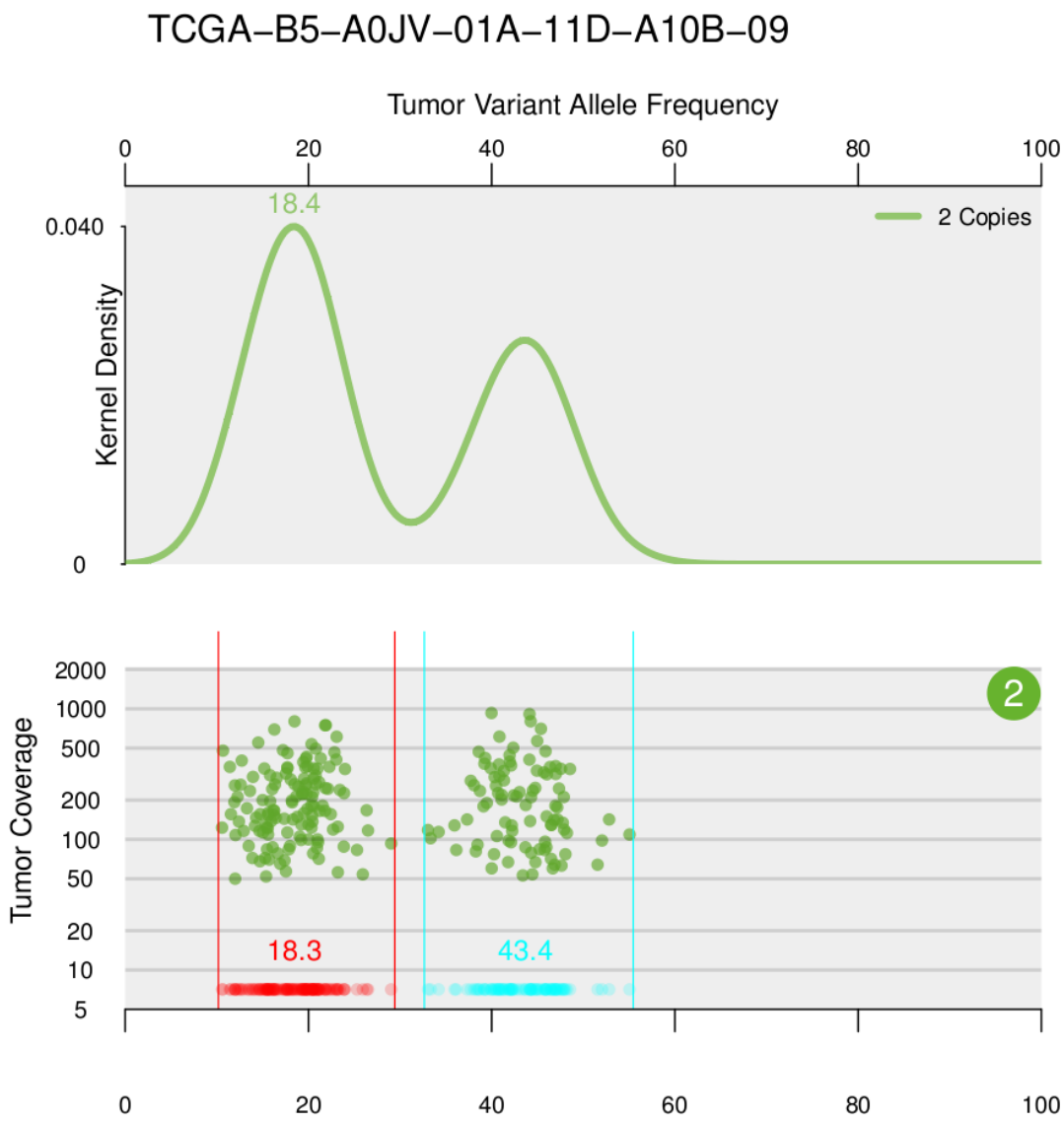


Clonality Plot

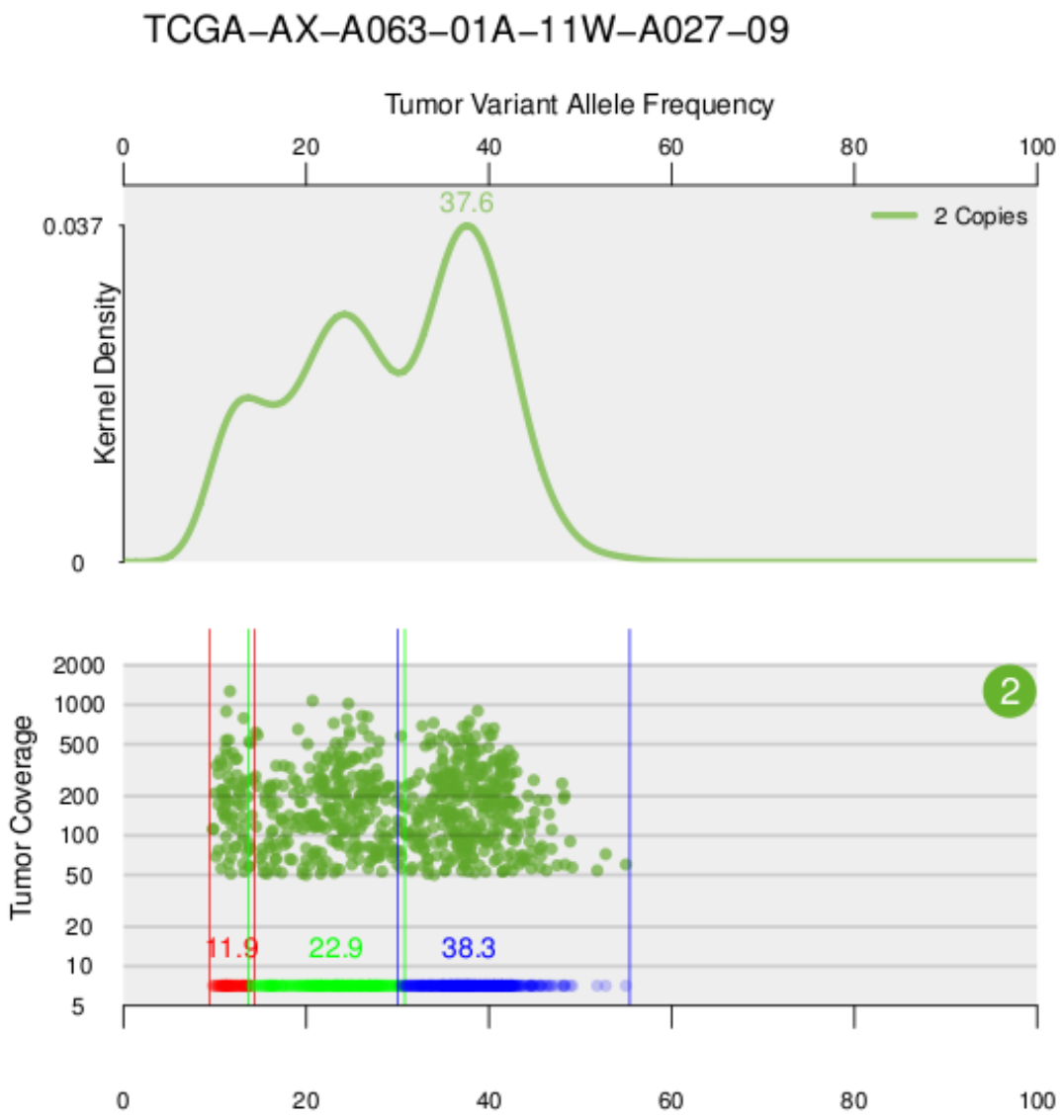


Infer clones in an automated, unbiased manner

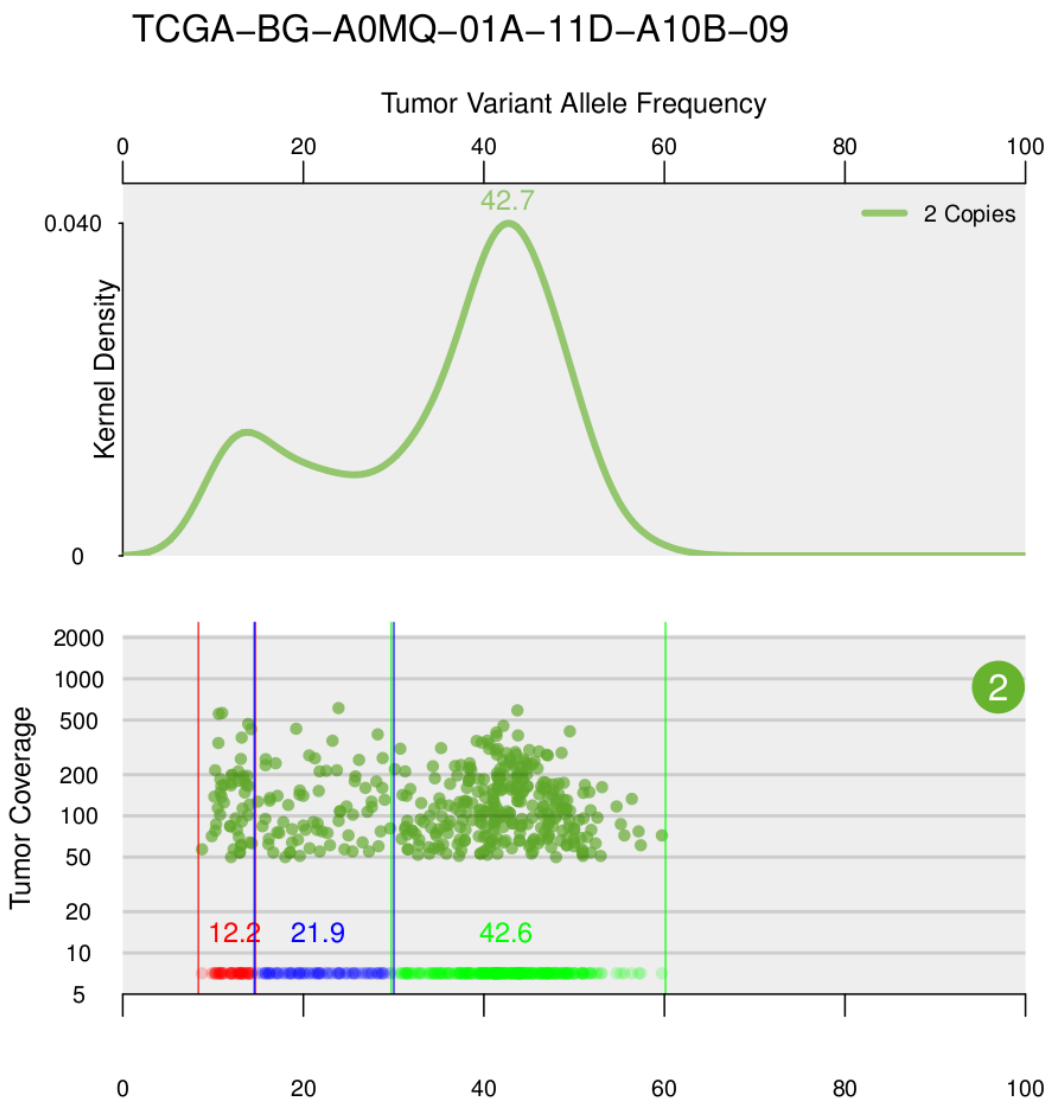
Biclonal sample



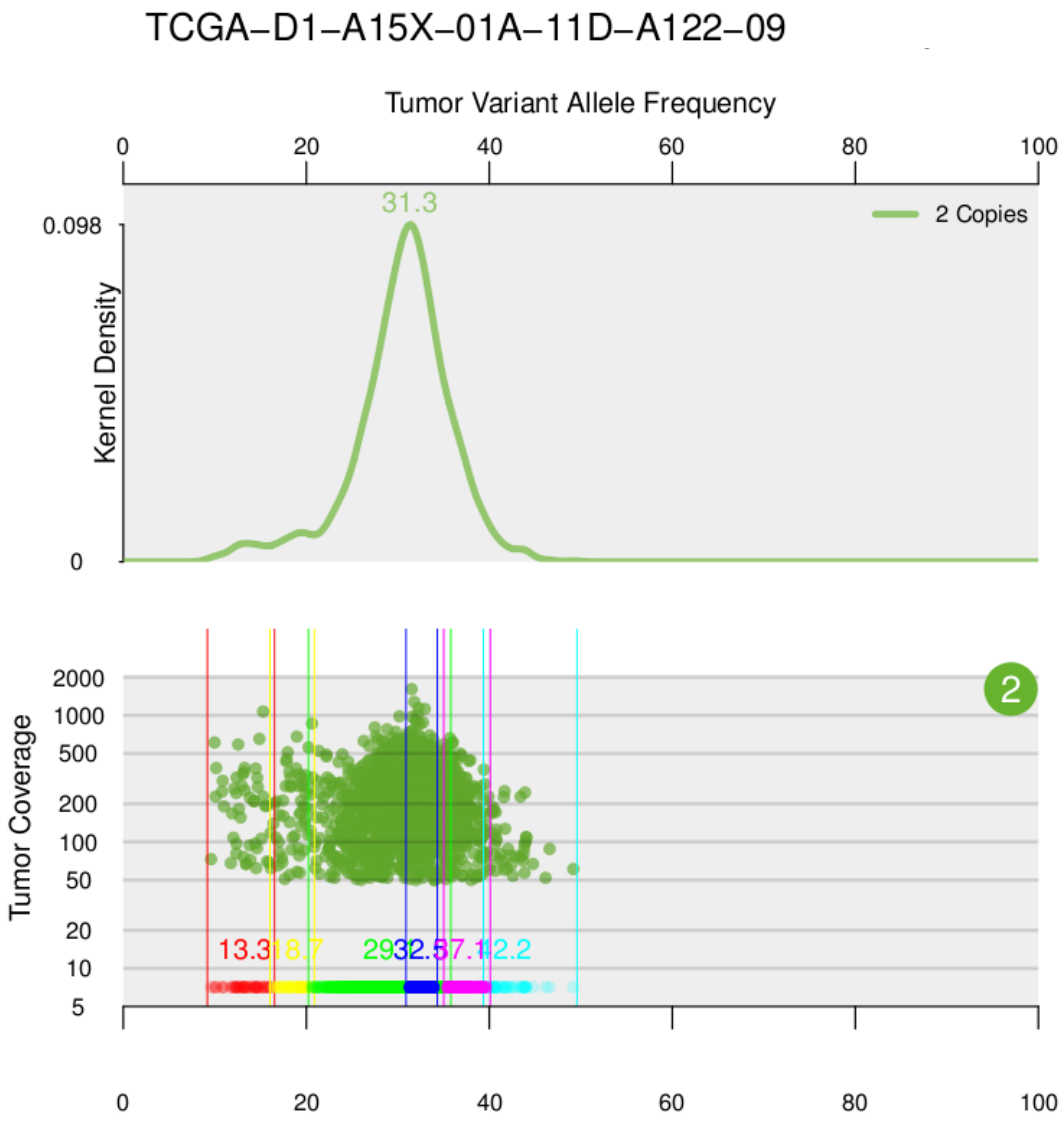
Triclonal sample



Non-intuitive sample



Multi-clonal Sample



Clonal inference

- Most tumors have a founding clone and one or more subclones (LAML, BRCA, UCEC)

- Lower bound on number of clones

Conclusions

- We can detect somatic mutations at very low frequencies using BASSOVAC
- We have developed robust automatic methods for inferring details about the subclonal architecture of a tumor
- Goal: characterizing minor subclones at diagnosis, rather than discovering their presence at relapse

Acknowledgements

- **Mike Wendl**
- **Nathan Dees**

- Dave Larson
- Travis Abbott
- Beifang Niu
- Brian White
- Will Schierding
- Josh McMichael
- Charles Lu
- Krishna Kanchi

- Tim Ley
- Michael Tomasson
- Ramaswamy Govindan
- Matthew Ellis
- Chuck Perou

- **Elaine Mardis**
- **Rick Wilson**
- **Li Ding**

- NHGRI
- NCI

The Cancer Genome Atlas

Understanding genomics to improve cancer care