

Validated low frequency somatic variants and copy number heterogeneity with bcbio

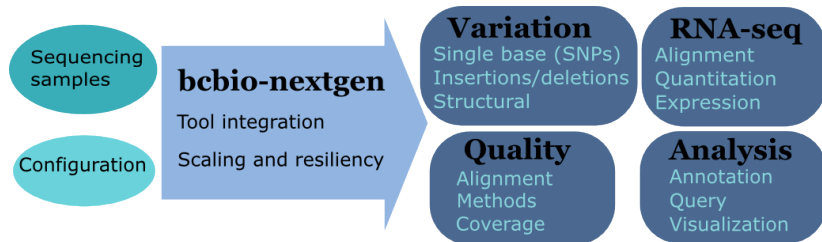
Brad Chapman
Bioinformatics Core, Harvard Chan School

<https://bcb.io>

<http://j.mp/bcbiolinks>

9 April 2018

Interoperable, community built workflows

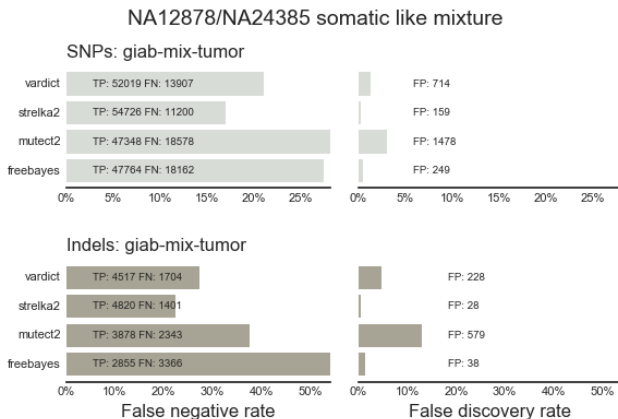


<https://github.com/bcbio/bcbio-nextgen>

- Common Workflow Language (CWL) descriptions
- Multiple platforms (HPC, DNAnexus, Arvados, SevenBridges)
- Multiple analysis areas:
 - Germline calling; single and joint
 - Somatic low frequency
 - Structural variants

https://github.com/bcbio/bcbio_validation_workflows

Somatic low frequency variants



DeepVariant (<https://github.com/google/deepvariant>)

https://github.com/bcbio/bcbio_validations/tree/master/strelka2

Tumor heterogeneity

- Difficult inputs: ctDNA, tumor-only, FFPE, panels
- Baseline calls, variants + CNVs with CNVkit
- Mixture of methods
 - TitanCNA – tumor/normal; purity + subclonal CNVs
 - ichorCNA – circulating tumor; purity estimation
 - PureCN – tumor only; purity + subclonal CNVs
 - PhyloWGS – evolutionary influence from subclonal
- Need standard validation sets