

Promoting platform interoperability with portable bcbio workflows

Brad Chapman
Bioinformatics Core, Harvard Chan School

<https://bcb.io>

<http://j.mp/bcbiolinks>

8 July 2016



bcbio

<http://bcb.io>

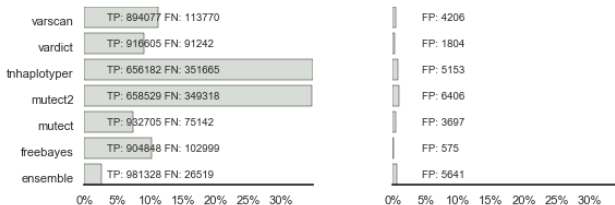
Good quality high throughput sequencing analyses

- Variant calling: SNPs and indels
- Structural variation
- Cancer heterogeneity
- RNA-seq
- small RNA
- Quality control

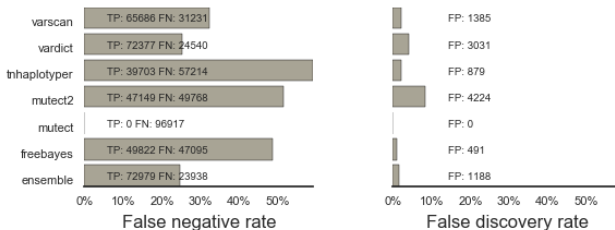
Validate and move quickly

NA12878/NA24385 mixture (GiaB 3.2.2) somatic validation

SNPs: NA12878/NA24385 mixture



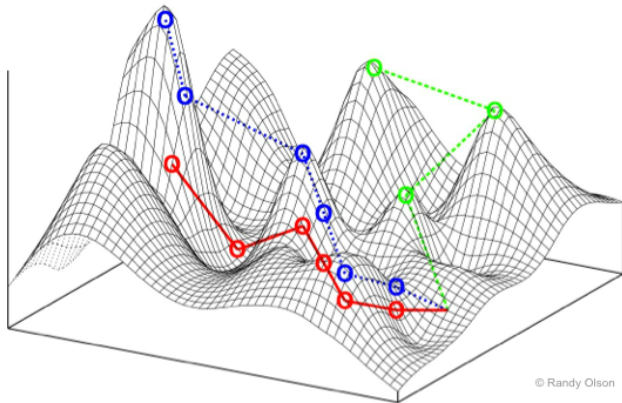
Indels: NA12878/NA24385 mixture



Infrastructure Goals

- Free, open source, community developed
- Welcoming to contributions
- Local machines
- Clusters: SLURM, SGE, Torque, PBS, LSF
- Clouds: Amazon, Google, Azure
- Clinical environments
- User interface for researchers
- Integrate with LIMS
- Accessible to the general public

Open source communities not yet optimal



https://en.wikipedia.org/wiki/Fitness_landscape

Better abstractions = more interoperability



COMMON
WORKFLOW
LANGUAGE



<https://bcbio-nextgen.readthedocs.io/en/latest/contents/cwl.html>