Practical Introduction

Variant Calling and Filtering for INDELs

December 9, 2014

Mary Kate Wing Adrian Tan Hyun Min Kang

Goals of This Session

- Aligned sequences -> indel calls
- Examine INDELs at particular genomic positions
- Evaluate quality of INDEL calls

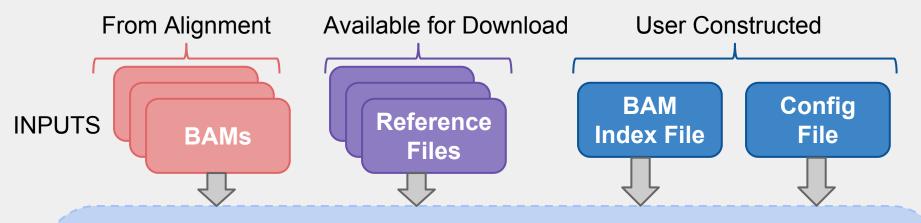
High Quality INDEL Calls from BAMs

- Many tools & best practices to choose from
- Our solution:

Genomes on the Cloud (GotCloud)

Yes, it has an indel pipeline too.

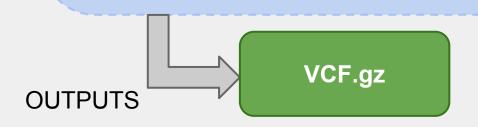
GotCloud INDEL Pipeline Overview



gotcloud indel --conf file.config

- 1) Discover variants per sample
- 2) Merge candidate variants across samples
- 3) Construct probes for genotyping
- 4) Genotype variants per sample
- 5) Merge and annotate

Makefile



GotCloud indel Input Files

- Same inputs as GotCloud snpcall
 - BAMs->INDELs instead of BAMs->SNPs

Looking at INDELs

- Use vt
 - http://genome.sph.umich.edu/wiki/Vt
 - view look at VCF
 - peek summarize VCF

- Your favorite VCF viewer/analyzer
 - tabix

Try it yourself

http://genome.sph.umich.edu/wiki/SeqShop:
_Variant_Calling_and_Filtering_for_INDELs_Pr
actical