

Bioinformatic pipeline to pooled data preparation

Kacper Żukowski

2016-02-02

Contents

Required software	1
Test dataset	2
Pipeline	2
Data trimming	2
Alignment to reference	2
Aligned data preparation and mpileup	2
Popoolation2 - the core	2
Variant calling and filtration	2
Data reduction and CMH testing as final step to genewiseR	2
References	2

Vignettes are long form documentation commonly included in packages. Because they are part of the distribution of the package, they need to be as compact as possible. The `html_vignette` output type provides a custom style sheet (and tweaks some options) to ensure that the resulting html is as small as possible. The `html_vignette` format:

- Never uses retina figures
- Has a smaller default figure size
- Uses a custom CSS stylesheet instead of the default Twitter Bootstrap style

Required software

1. to download test data from ncbi – **ncbi sra toolkit** [link](#)
2. to trimming raw sequences – **flexbar** [link](#)
3. to alignment – **bwa** [link](#)
4. to utilities – **picard tools** [link](#)
5. to utilities and mpileup – **samtools** [link](#)
6. to prepare sync data and more – **popoolation2** [link](#)
7. to variant calling – **freebayes** [link](#) and **gatk** [link](#)

8. to variant filtration – **snpsift** [link](#)

Note the various macros within the **vignette** section of the metadata block above. These are required in order to instruct R how to build the vignette. Note that you should change the **title** field and the `\VignetteIndexEntry` to match the title of your vignette.

Test dataset

Pipeline

Data trimming

Alignment to reference

Aligned data preparation and mpileup

Popoolation2 - the core

Variant calling and filtration

Data reduction and CMH testing as final step to genewiseR

References