Lab II

1. Import data from IIHG course sample 1 -> 1.bam, 1 R1.fastq, 1 R2.fastq otoscope.bed
2. Run flagstat on 1 R1.fastq
3. Run fastqc on 1 R1.fastq
4. Run NGSRich on 1.bam and otoscope.bed
5. Create histogram of target coverage (bedtools)
6. Select NOT matching '^all'
7. Cut c1,c2,c3,c4 (really keeps the first 4 columns)
8. Submit to either local or UCSC genomic browser
9. Evaluate coverage of TPRN, compare to RDX
10. Align 1 R1.fastq and 1 R2.fastq to hg19 using bowtie
11. Run flagstat on output bam file
    a. Compare 1.bam to output bam
12. Extract workflow from history
13. Edit workflow to just include flagstat and fastqc steps
14. Import sample 6 PGM-7.fastq sample
15. Run new workflow on PGM-7.fastq and 1 R2.fastq
    a. Compare fastqc outputs
    b. We will pause and then discuss
        i. Bowtie output vs 1.bam
        ii. PGM7 fastqc vs 1 R1 fastq
16. Start new history
17. Go to galaxy main
18. Find published workflow "Sureselect Find Uncovered Target Intervals"
19. Download and import into your current workflow
20. Try running
21. Discuss
    a. Why didn't this work
    b. Two (or more) ways you can get this to work
22. Start new history
23. Go to our galaxy and import "training workflow"
24. Spend 20-30 minutes making this workflow look like CLCG Illumina Paired-End Workflow
    a. Get used to connecting the spaghetti
    b. Examine how to set parameters
    c. Examine how to chain outputs
25. Start new history
26. Import either CLCG Illumina Paired-End Workflow or your modified training workflow (if brave)
27. Import sample 2 data as well as bed file and dbsnp.vcf file
28. Run workflow

29. Share or publish history -> make history accessible via link and email url to your neighbor
30. Open your neighbors history