#load bowtie on nomachine

module load bowtie2

#create HXB2 index

bowtie2-build "//home/sdresche/Documents/K03455.fasta" HXB2

bowtie2-build "//home/sdresche/Documents/Q23.fa" Q23

#would build human index but this will take forever so load premade

## bowtie2-build "//home/sdresche/Documents/GCA\_000001405.27\_GRCh38.p12\_genomic.fna" Human

Export BOWTIE2\_INDEXES="/home/sdresche/Documents/GCA\_000001405.15\_GRCh38\_no\_alt\_analysis\_set.fna.bowtie\_index"

bowtie2-inspect -s GCA\_000001405.15\_GRCh38\_no\_alt\_analysis\_set.fna.bowtie\_index

#align to HXB2

bowtie2 -q --local -x HXB2 -1 /home/sdresche/Documents/nano/1\_S1\_L001\_R1\_001.fastq.gz -2 /home/sdresche/Documents/nano/1\_S1\_L001\_R2\_001.fastq.gz -S /home/sdresche/Documents/nano/SAMHXB2.sam

#align to Q23

bowtie2 -q --local -x Q23 -1 /home/sdresche/Documents/nano/1\_S1\_L001\_R1\_001.fastq.gz -2 /home/sdresche/Documents/nano/1\_S1\_L001\_R2\_001.fastq.gz -S /home/sdresche/Documents/nano/SAM\_Q23.sam

bowtie2 -q --local -x Q23 -1

/home/sdresche/Documents/nano/2\_S2\_L001\_R1\_001.fastq.gz -2 /home/sdresche/Documents/nano/2\_S2\_L001\_R2\_001.fastq.gz -S /home/sdresche/Documents/nano/SAM\_Q23\_02.sam

bowtie2 -q --local -x Q23 -1 /home/sdresche/Documents/nano/3\_S3\_L001\_R1\_001.fastq.gz -2 /home/sdresche/Documents/nano/3\_S3\_L001\_R2\_001.fastq.gz -S /home/sdresche/Documents/nano/SAM\_Q23\_03.sam

#align to human

bowtie2 -q --local --un-conc /home/sdresche/Documents/nano/unmapped\_to\_human --al-conc /home/sdresche/Documents/nano/mapped\_to\_human -x GCA\_000001405.15\_GRCh38\_no\_alt\_analysis\_set.fna.bowtie\_index -1 /home/sdresche/Documents/nano/1\_S1\_L001\_R1\_001.fastq.gz -2 /home/sdresche/Documents/nano/1\_S1\_L001\_R2\_001.fastq.gz -S /home/sdresche/Documents/nano/SAM01.sam

bowtie2 -q --local --un-conc /home/sdresche/Documents/nano/unmapped\_to\_human01 -x GCA\_000001405.15\_GRCh38\_no\_alt\_analysis\_set.fna.bowtie\_index -1 /home/sdresche/Documents/nano/1\_S1\_L001\_R1\_001.fastq.gz -2 /home/sdresche/Documents/nano/1\_S1\_L001\_R2\_001.fastq.gz -S /home/sdresche/Documents/nano/SAM\_to\_human\_01.sam

bowtie2 -q --local --un-conc /home/sdresche/Documents/nano/unmapped\_to\_human02 -x GCA\_000001405.15\_GRCh38\_no\_alt\_analysis\_set.fna.bowtie\_index -1 /home/sdresche/Documents/nano/2\_S2\_L001\_R1\_001.fastq.gz -2 /home/sdresche/Documents/nano/2\_S2\_L001\_R2\_001.fastq.gz -S /home/sdresche/Documents/nano/SAM\_to\_human\_02.sam

bowtie2 -q --local --un-conc /home/sdresche/Documents/nano/unmapped\_to\_human03 -x GCA\_000001405.15\_GRCh38\_no\_alt\_analysis\_set.fna.bowtie\_index -1 /home/sdresche/Documents/nano/3\_S3\_L001\_R1\_001.fastq.gz -2 /home/sdresche/Documents/nano/3\_S3\_L001\_R2\_001.fastq.gz -S /home/sdresche/Documents/nano/SAM\_to\_human\_03.sam

bowtie2 -q --local --un-conc /home/sdresche/Documents/nano/unmapped\_to\_human04 -x GCA\_000001405.15\_GRCh38\_no\_alt\_analysis\_set.fna.bowtie\_index -1 /home/sdresche/Documents/nano/4\_S4\_L001\_R1\_001.fastq.gz -2 /home/sdresche/Documents/nano/4\_S4\_L001\_R2\_001.fastq.gz -S /home/sdresche/Documents/nano/SAM\_to\_human\_04.sam

#convert SAM to BAM

#HXB2

samtools view -b -S -o /home/sdresche/Documents/nano/SAMtoHXB2.bam /home/sdresche/Documents/nano/SAMHXB2.sam

#Q23

samtools view -b -S -o /home/sdresche/Documents/nano/SAMtoQ23.bam /home/sdresche/Documents/nano/SAM\_Q23.sam

samtools view -b -S -o /home/sdresche/Documents/nano/BAMtoQ23\_03.bam /home/sdresche/Documents/nano/SAM\_Q23\_03.sam

samtools view -b -S -o /home/sdresche/Documents/nano/SAM\_to\_human\_03.bam /home/sdresche/Documents/nano/SAM\_to\_human\_03.sam

#sort

samtools sort /home/sdresche/Documents/nano/SAM\_to\_human\_03.bam /home/sdresche/Documents/nano/BAM\_to\_human\_03.bam.sorted

samtools sort /home/sdresche/Documents/nano/BAMtoQ23\_03.bam /home/sdresche/Documents/nano/BAMtoQ23\_03.sorted.bam

samtools sort /home/sdresche/Documents/nano/SAMtoHXB2.bam /home/sdresche/Documents/nano/BAMtoHXB2.sorted

samtools sort /home/sdresche/Documents/nano/SAMtoQ23.bam /home/sdresche/Documents/nano/BAMtoQ23.sorted

#make Igv index

samtools index /home/sdresche/Documents/nano/BAMtoHXB2.sorted.bam

samtools index /home/sdresche/Documents/nano/BAMtoQ23.sorted.bam

samtools index /home/sdresche/Documents/nano/BAM\_to\_human\_03.bam.sorted

#need to make an alias file for IGV and then can use it

#see:

http://software.broadinstitute.org/software/igv/loaddata

#Now I'll try to make a sample specific consensus with bcf tools

module load bcftools

#faidx

#samtools faidx <ref.fasta> [region1 [...]]

#Index reference sequence in the FASTA format or extract subsequence from indexed reference #sequence. If no region is specified, faidx will index the file and create <ref.fasta>.fai on #the disk. If regions are specified, the subsequences will be retrieved and printed to stdout in #the FASTA format. The input file can be compressed in the BGZF format.

#need to have faidx with same name in same directory of the samtools mpileup

samtools faidx //home/sdresche/Documents/K03455.fasta

samtools mpileup -uf ref.fa aln.bam | bcftools call -c | vcfutils.pl vcf2fq > cns.fq

# call variants

#which mpileup? The bcf one didnt' work so I'll try samtools

#Ok this one works:

samtools mpileup -Ou -f K03455.fasta nano/BAMtoHXB2.sorted.bam | bcftools call -mv -Oz -o nano/calls.vcf.gz

samtools mpileup -f K03455.fasta nano/BAMtoHXB2.sorted.bam | bcftools call -mv -Oz > calls.vcf.gz

samtools mpileup -Ou -f K03455.fasta nano/BAMtoHXB2.sorted.bam | bcftools call -mv -Oz -o nano/calls.vcf.gz

bcftools norm -f K03455.fasta nano/calls.vcf.gz -Ob -o calls.norm.bcf

# filter adjacent indels within 5bp

bcftools filter --IndelGap 5 calls.norm.bcf -Ob -o calls.norm.flt-indels.bcf

#now do consensus?

cat K03455.fasta | bcftools consensus nano/calls.vcf.gz > consensus.fa

#| bcftools call -mv -Oz -o /home/sdresche/Documents/nano/calls.vcf.gz

bcftools index /home/sdresche/Documents/nano/calls.vcf.gz

#bcftools mpileup -Ou -f "//home/sdresche/Documents/K03455.fasta" /home/sdresche/Documents/nano/BAMtoHXB2.sorted.bam

#| bcftools call -mv -Oz -o /home/sdresche/Documents/nano/calls.vcf.gz

bcftools index /home/sdresche/Documents/nano/calls.vcf.gz

bowtie2-inspect -s GCA\_000001405.15\_GRCh38\_no\_alt\_analysis\_set

export BOWTIE2\_INDEXES="/home/sdresche/Documents/GCA\_000001405.15\_GRCh38\_no\_alt\_analysis\_set.fna.bowtie\_index"

bowtie2-inspect -s GCA\_000001405.15\_GRCh38\_no\_alt\_analysis\_set.fna.bowtie\_index

samtools view -H /home/sdresche/Documents/nano/BAMtoHXB2.sorted.bam

samtools view -b -S -o alignments/sim\_reads\_aligned.bam alignments/sim\_reads\_aligned.sam