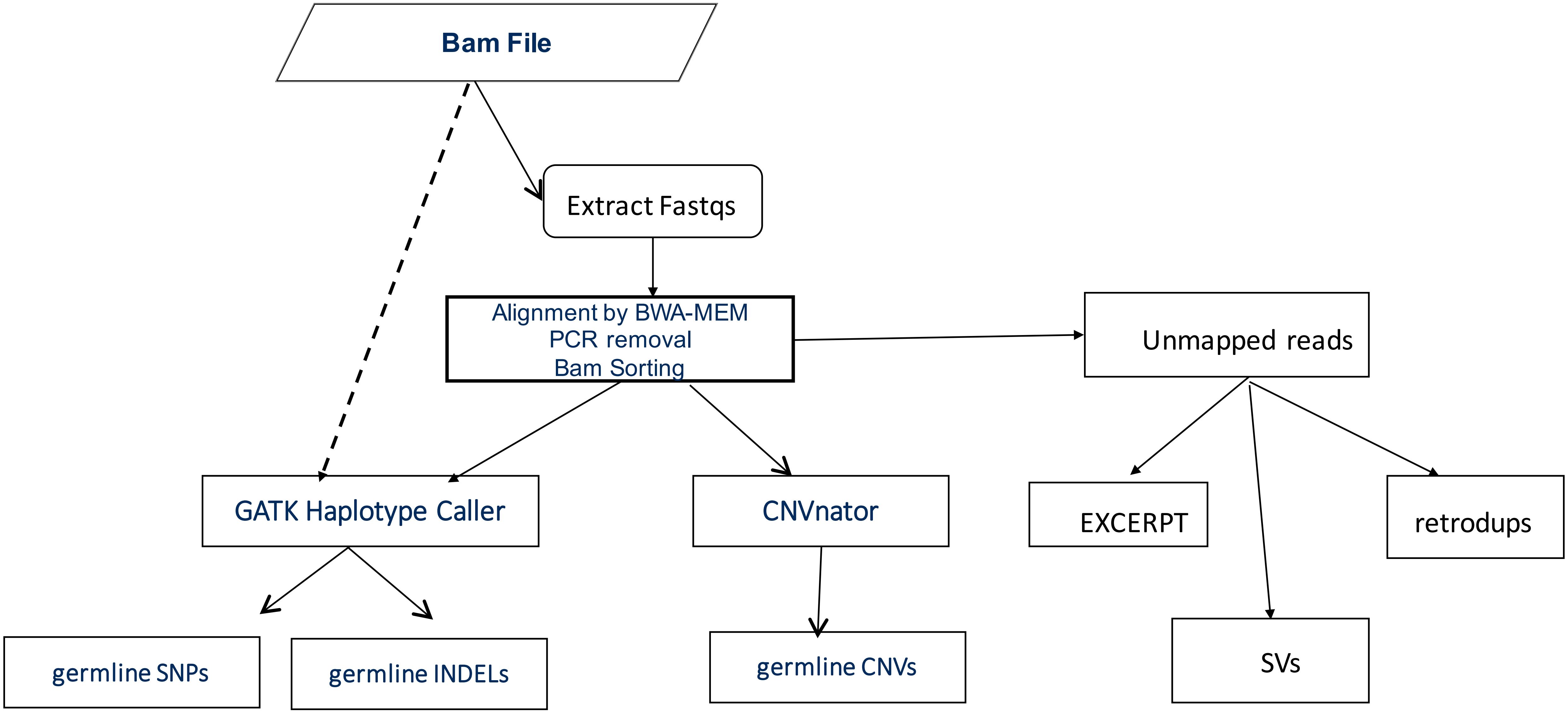
Personal Genome Analysis

P2-Tech



Bam File

Extract Fastqs

Alignment by BWA-MEM
PCR removal
Bam Sorting

Unmapped reads

GATK Haplotype Caller

CNVnator

EXCERPT

retrodups

germline SNPs

germline INDELs

germline CNVs

SVs

Pipeline to generate variants for personal genome analysis

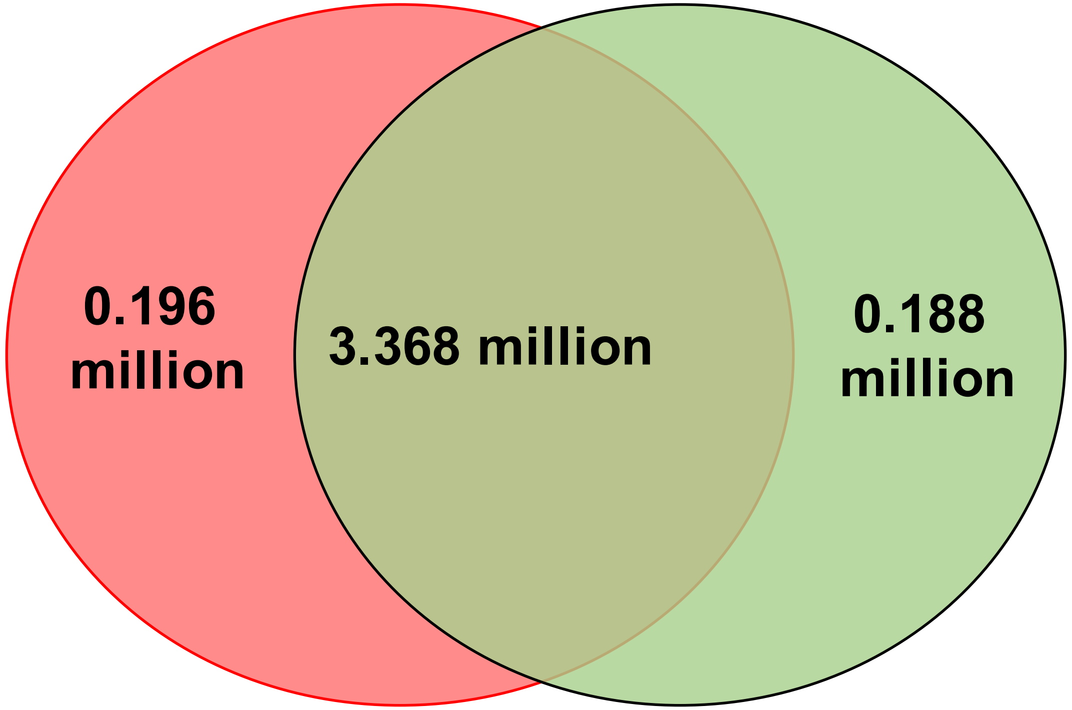
Bam statistics for zimmerome

Paramters	Zimmerome illumina Bam	Zimmerome Gerstein Bam
Mapper read %	94.47 %	98.99 %
Unmapped read %	5.53%	1.01%
Properly paired read%	92.87%	98.49
Average Mapping quality	40.2	49.52

Comparison of SNP calls for zimmerome

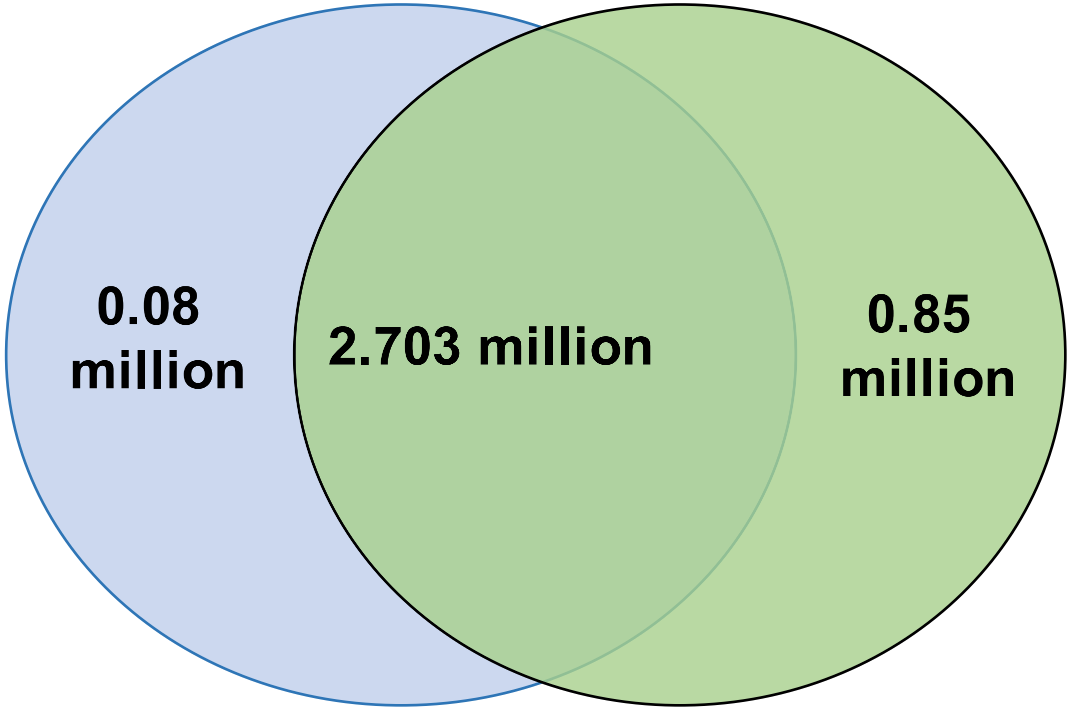
GATK illumina bam ~ 3.56 million

Isaac illumina bam ~ 3.55 million



GATK GS illumina bam ~ 2.78million

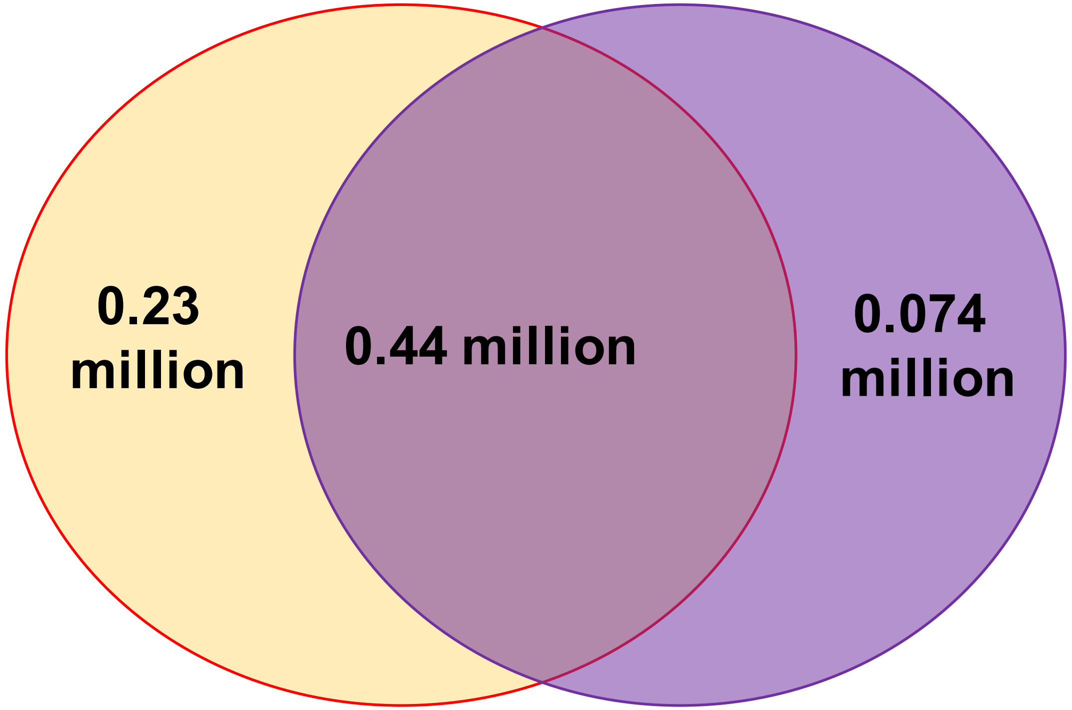
Isaac illumina bam ~ 3.55 million



Comparison of INDEL calls for zimmerome

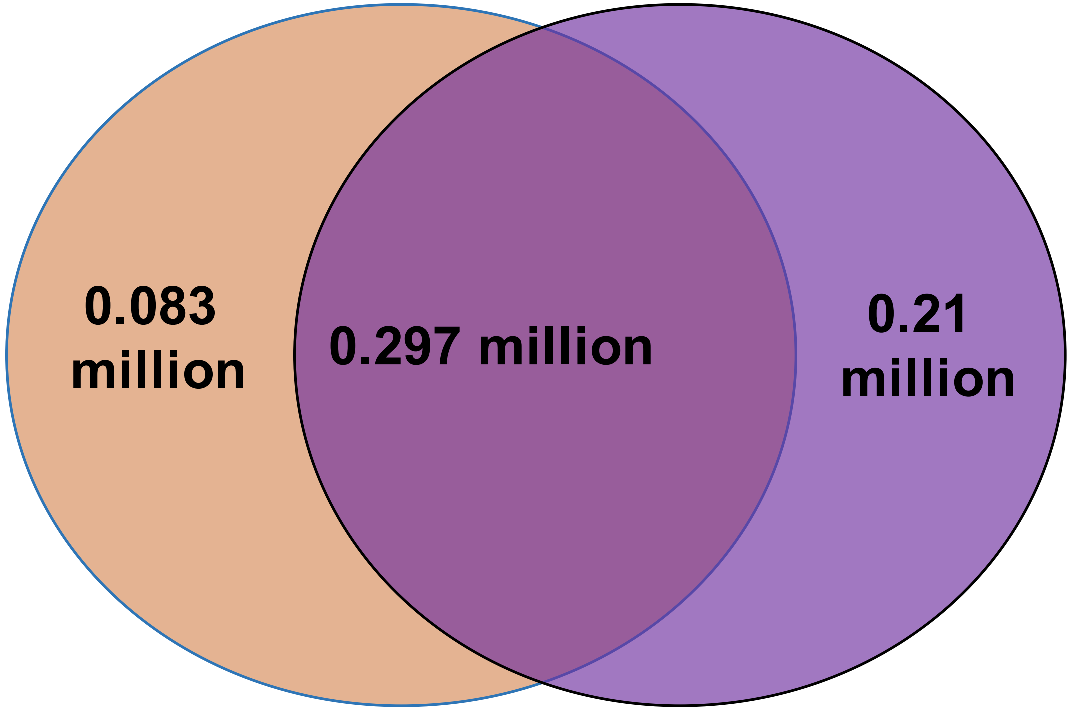
GATK illumina bam ~ 0.67 million

Isaac illumina bam ~ 0.51 million



GATK GS illumina bam ~ 0.38million

Isaac illumina bam ~ 0.51 million



To-do list

Finish NA12878 germline SNV and Indel calling

Finish SV and CNV calling for snyderome and NA12878. Perform overlap with genomic elements

Provide germline calls to YF for Aloft run. SV and germline calls to TG for personal genome construction.

Map SNVs onto three dimensional protein structure and provide input to DC for STRESS calculation.

Rest of SNV call set for FunSeq analysis (XL +SKL)

NA12878 unmapped read to RK for exogenous element analysis