The Sentieon Genomics Tools – A fast and accurate solution to variant calling from next-generation sequence data

Donald Freed[1*], Rafael Aldana[1], Jessica A. Weber[2], Jeremy S. Edwards[3,4,5]

[1]Sentieon Inc, Mountain View, CA.
[2]Department of Biology, University of New Mexico, Albuquerque, NM.
[3]Department of Chemistry & Chemical Biology, University of New Mexico, Albuquerque, NM.
[4]Department of Molecular Genetics and Microbiology, University of New Mexico, Albuquerque, NM.
[5]Department of Chemical & Nuclear Engineering, University of New Mexico, Albuquerque, NM.
*Correspondence:
don.freed@sentieon.com

Abstract

In the past six years worldwide capacity for human genome sequencing has grown by more than five orders of magnitude, with costs falling by nearly two orders of magnitude over the same period [1], [2]. The rapid expansion in the production of next-generation sequence data and the use of these data in a wide range of new applications has created a need for improved computational tools for data processing. The Sentieon Genomics tools provide an optimized reimplementation of the most accurate pipelines for calling variants from next-generation sequence data, resulting in more than a 10-fold increase in processing speed while providing identical results to best practices pipelines. Here we demonstrate the consistency and improved performance of Sentieon's tools relative to BWA, GATK, MuTect, and MuTect2 through analysis of publically available human exome, low-coverage genome, and high-depth genome sequence data.

Keywords: Variant calling; Next-generation sequencing; Genomics; Bioinformatics

Introduction

The cost of sequencing a human genome has fallen rapidly over the last decade and is now near $1000 for a single high-depth human genome, with annual increases in both sequencing efficiency in Gb per dollar and worldwide sequencing capacity consistently surpassing Moore's Law [1], [2]. Importantly, data quality has also improved due to advances in the underlying technology and chemistry of the sequencing machines. Read lengths of the Illumina's flagship X10 sequencer are now 150bp and data quality remains high across the entire read [3].

As the amount of available sequence data increases, efficient and accurate data analysis is becoming increasingly important. Next-generation sequence data are frequently being used to help inform economic and clinical decisions through applications such as non-invasive fetal testing [4], personalized therapy including cancer immunotherapy [5]–[10], genetic diagnosis [11], disease gene discovery [12], [13], discovery of contributory mutations in complex disease [14], and discovery of important genetic traits in agriculture [15]. Due to the increased reliance upon next-generation sequence data for informing these decisions, the pipelines for analyzing these data are understandably under increasing regulatory scrutiny [16].

While early human genome projects necessarily relied upon whole-genome *de novo* assembly, the short-read data produced by second-generation sequencers are not readily amenable to this method of assembly. As a result, mapping reads to a reference genome and identifying alleles present in the sample and absent from the reference genome has historically been the prominent paradigm for analysis of these data [17]–[20]. Initially these methods relied upon Bayesian approaches to evaluate the likelihoods of mutations occurring at single-base pairs. However, newer haplotype-based approaches have improved accuracy and have become the industry standard [21]–[23].

Two of the most popular tools for variant detection are the GATK and MuTect [23]–[25]. The publications describing these tools have been cited nearly 10,000 times and both have been used in many high-profile research projects [14], [26], [27]. Part of the reason for their wide adoption is their high accuracy; these tools perform well in many benchmarks including the precisionFDA and ICGC-TCGA DREAM challenges [28], [29]. However, these tools do have drawbacks. The improvements that produce higher accuracy have also resulted in long runtimes and high memory usage for some tasks. Downsampling is implemented in some algorithms to help mitigate these issues, but is not an ideal solution as it discards valuable data, leads to increased run-to-run variation, and possibly results in erroneous or missing variant calls. In some steps, intermediate file merging is also necessary to reduce computational load.

Due to the projected increase in the amount of available next-generation sequence data and the use of these data for increasingly important applications, there is a pressing need for new and

improved software solutions for next-generation sequence data analysis. We have developed the Sentieon Genomics tools to address this need. The Sentieon tools provide a complete rewrite of the mathematical models of the Best Practices GATK, Picard, MuTect, and MuTect 2 in the Sentieon DNAseq, TNseq and TNHaplotyper pipelines with a focus on computational efficiency, accuracy, and consistency.

Materials and Methods

Benchmarking comparisons were completed using BWA 0.7.12-r1039 [30], [31], Picard tools 1.112, GATK 3.5 [23], [24], SAMtools 1.2 [32], MuTect 1.1.5 [25], and version 201611 of the Sentieon Genomics tools. All commands were run on a 32 core 2.4 GHz Intel Xeon server with 64 GB memory and 2TB dual stripped SSDs for intermediate file storage. The server was dedicated to benchmarking and had no other running jobs. All samples used in benchmarking are publically available and are listed in Supplementary Appendix 1. Data from multiple lanes that were obtained from the Baylor Human Genome Sequencing Center were concatenated into a single file before analysis. Samples HG001 and HG002 were obtained from the precisionFDA platform.

For all samples, sequence reads were aligned to the Human reference genome (UCSC hg19) with BWA MEM followed by sorting and indexing using samtools or the Sentieon utility. Alignment summary, GC bias, base quality by sequencing cycle, base quality score distribution, and insert size metrics were collected and duplicate reads were removed with either Picard tools or the Sentieon driver. Indels were realigned and base quality was recalibrated using either the GATK or the Sentieon driver. For tumor-normal paired samples, joint indel realignment was performed using the GATK or the Sentieon driver. For germline samples, variants were called using the GATK UnifiedGenotyper and the GATK HaplotypeCaller or Sentieon's Haplotyper and Genotyper algorithms while variants were called in paired samples using MuTect and MuTect2 or Sentieon's TNsnv and TNHaplotyper algorithms.

For the benchmarking of the joint calling, gVCF files were generated using the Sentieon pipeline described in Supplementary Appendix 1 with the --emit_mode gvcf option added during variant calling. gVCF files were then genotyped by the GATK GenotypeGVCFs and Sentieon's GVCFtyper across chromosome 1. Genotyping was not performed across the entire genome as the estimated runtime for the whole genome analysis with the GATK was over two weeks. Since Sentieon's GVCFtyper does not drop alternate alleles, GenotypeGVCFs was run with the option -maxAltAlleles 100 to provide comparable results.

RTG Tools' vcfeval 3.5.1 was used to compare variants called by Sentieon relative to variants called by the other tools. Detailed commands used during the data processing are listed in Supplementary Appendix 1.

Results

The Sentieon Genomics Pipeline Tools

The Sention Genomics pipeline provides a suite of tools for secondary analysis of next-generation sequence data. Currently supported pipelines are composed of optimized implementations of the mathematical models of the most accurate variant calling pipelines. Improvements in performance are achieved through optimization of the algorithms and improved resource management. The tools run on both Linux (RedHat/CentOS, Debian, OpenSUSE and Ubuntu) and OSX distributions and require no specialized hardware, additional libraries or complex installation procedure. Sentieon's variant calling tools do not perform data down-sampling and are deterministic, providing perfect run-to-run consistency [28].

The Sentieon Tools Provide a 10-fold Improvement in Runtime Over the GATK, MuTect and MuTect2

To test the performance of the Sentieon Genomics pipeline tools, we evaluated the runtime of Sentieon's tools and the GATK, MuTect, and MuTect2 in a consistent computing environment using publically available data (a complete set of samples is listed in Supplementary Table 1). While advanced users may use sophisticated parallelization methods to improve performance (such as performing operations on small intervals and concatenating the results), optimization, validation and testing of these methods require expertise, expensive human capital investment, and vary in their effectiveness from user-to-user. Due to these costs only built-in parallelization methods are tested here. Multi-threaded processing of the Sentieon algorithms was accomplished using the "-t" option while the GATK, and MuTect2 were parallelized with the "-nct" option. Unfortunately, multi-threading is not supported in MuTect. The full set of commands used in all analyses is shown in Supplementary Appendix 1.

Prior to variant calling, sequence data were aligned to the human reference genome (UCSC hg19) using BWA MEM. BWA is one of the most popular aligners for alignment of next-generation sequence reads given its accuracy and ability to produce correct alignments to structural variants. Sentieon provides an optimized implementation of BWA providing an

average 1.9x speedup during alignment (range 1.0 to 3.9x, Figure 1, Supplementary Table 2), while producing identical alignments.
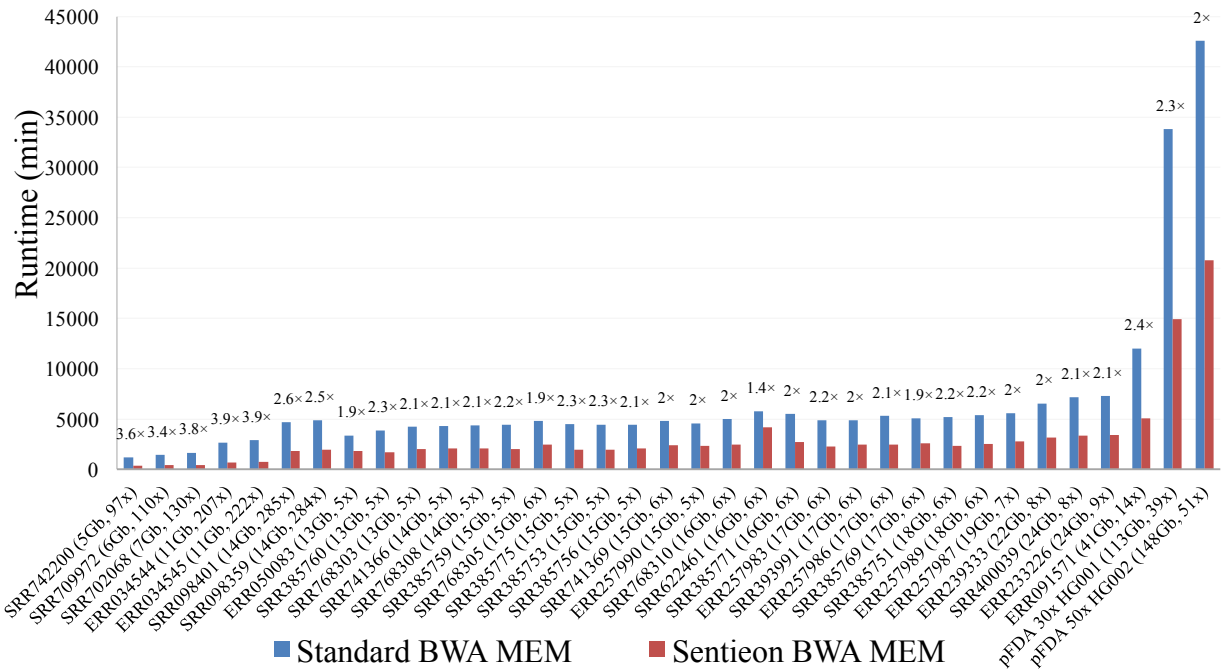
Figure 1. Standard BWA MEM and Sentieon BWA MEM runtime comparison.
Runtimes of the standard BWA with SAMtools sort and Sentieon BWA and sort on whole-exome and whole-genome samples. Labels indicate fold improvement in runtime provided by the Sentieon implementation.

Sention DNAseq and the GATK Best Practices Pipeline were run on seven whole-exome, 77 low-coverage, and two whole-genome samples. Among all samples, the Sentieon DNAseq pipeline resulted in an average 36x (range 16 to 55x) improvement in the pipeline runtime relative to the GATK Best Practices pipeline (Figure 2, Supplementary Table 2, Supplementary Table 3). The performance improvements were most notable in the indel realignment, base-quality score recalibration, and HaplotypeCaller variant calling stages where Sentieon's tools improved runtimes by an average of 56x, 46x, and 30x, respectively.
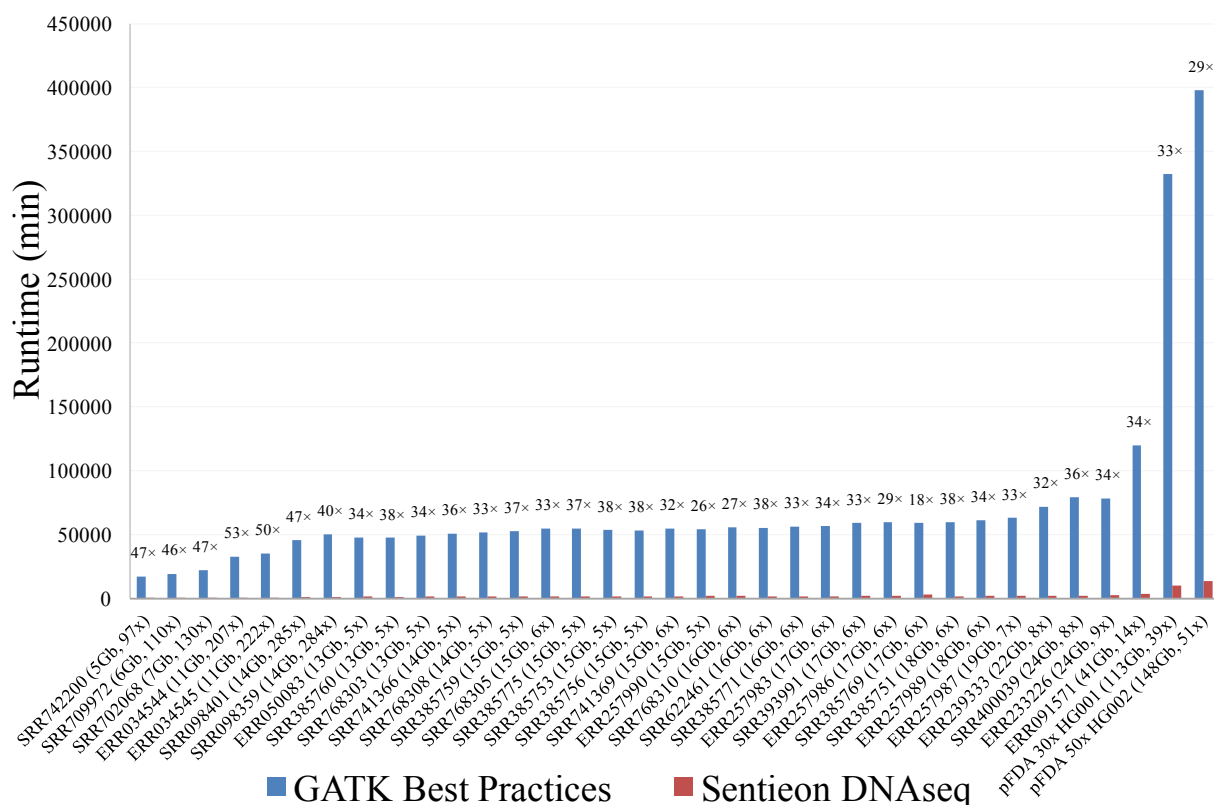
Figure 2. DNAseq pipeline runtime comparison.
Runtimes of the Sentieon DNAseq and GATK Best Practices pipelines on whole-exome, low-coverage whole genome, and high-coverage whole-genome samples for the metrics calculation through variant calling stages. Samples were sorted by their total number of sequenced bases. Labels indicate the fold improvement in runtime provided by the Sentieon tools over the GATK. The runtime improvement of Sentieon DNAseq over GATK ranges from 18-53x.

To better understand the performance differences between the GATK and Sentieon tools for joint genotyping, we ran the GATK's GenotypeGVCFs and Sention's Genotyper on chromosome 1 of 74 gVCF files from the 1000 Genomes Project preprocessed with the Sentieon DNAseq pipeline. Sentieon's tools provided a 183x improvement in runtime relative to the GATK for joint genotyping (216 and 39,555 seconds for Sentieon and the GATK, respectively).

We also measured the performance of the Sentieon TNseq and TNHaplotyper pipelines relative to MuTect and MuTect2 with three tumor-normal paired samples. Sentieon's tools resulted in an average speedup of 19x (range 10x to 24x) relative to MuTect2 and an average speedup of 42x (range 38 to 48x) relative to the single-threaded MuTect (Figure 3, Supplementary Table 4). Overall, the Sentieon tools provide an average 19x to 42x speedup over earlier implementations of the variant calling pipelines in the GATK, MuTect and MuTect2.
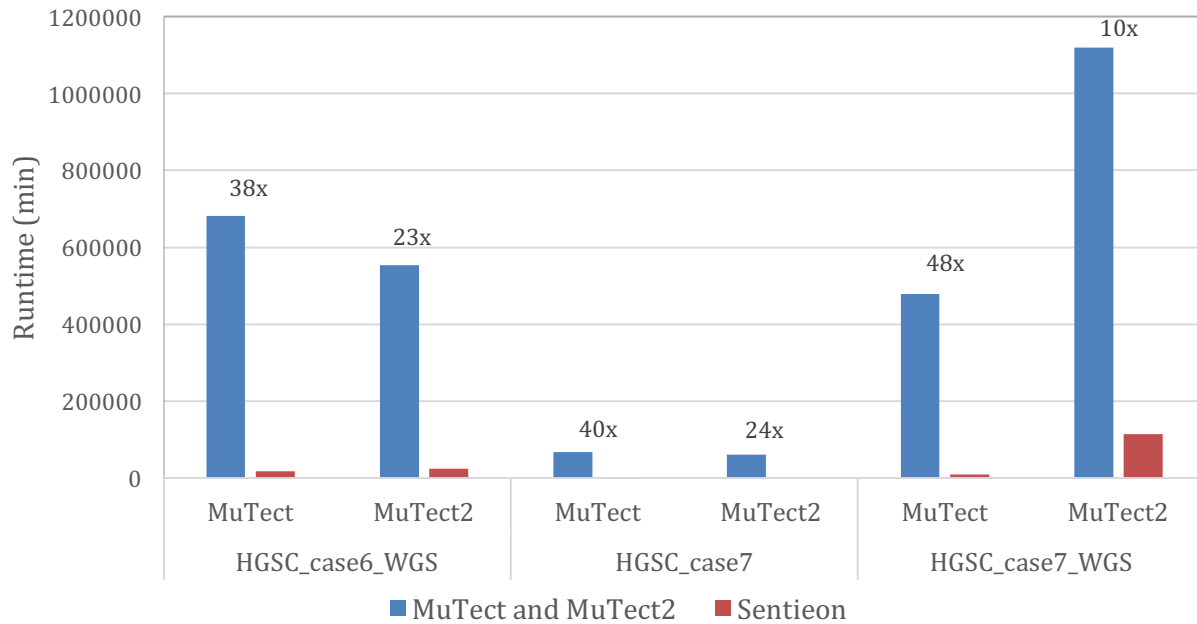
Figure 3. TNseq and TNhaplotyper runtime comparison.
Runtimes of the Sentieon TNseq and TNhaplotyper pipelines relative to MuTect and MuTect2 on whole-exome and whole-genome samples. Steps included in the plot are metrics calculation through variant calling. Labels indicate fold improvement in runtime provided by the Sentieon tools relative to MuTect and MuTect2.

The Sentieon Tools Produce Results Consistent with the GATK, MuTect and MuTect2

Using the VCF files produced in the runtime benchmark, we set out to evaluate the consistency of Sentieon's tools with the GATK, MuTect, and MuTect2 using RTG Tools vcfeval. When interpreting the results of these experiments, it is important to note the run-to-run variation of the tools. While the Sentieon tools are deterministic, repeated runs of multithreaded GATK and MuTect2 produce slightly different results, limiting the consistency that can be achieved between Sentieon and these other tools [33]. MuTect downsamples sequence data deterministically while GATK and MuTect2 downsample sequence data randomly when running with multiple threads. However, even deterministic downsampling may cause additional variants to be called or lost relative to an identical implementation without downsampling. Here we call the results identical if the measured difference between the tools is on the order of the intrinsic run-to-run variation of the GATK observed by Weber *et al*. [33]. The Sentieon DNAseq pipeline produced results identical to the GATK best practices pipeline (average F-score 0.9996; range 0.9974 to 1.0000; Figure 4; Supplementary Table 5). In the joint genotyping comparison, the Sentieon tools produced results identical to the GATK (Supplementary Table 6). While the Sentieon TNseq and TNHaplotyper pipelines produced results identical to MuTect and MuTect2 (Average F-score 0.9994 and 0.9920 for MuTect and MuTect2, respectively; Supplementary Table 7). These

results indicate that the Sentieon DNAseq, TNseq and TNHaplotyper pipelines provide a drop-in replacement for the best-practices pipelines with results within the run-to-run variation of the tools.
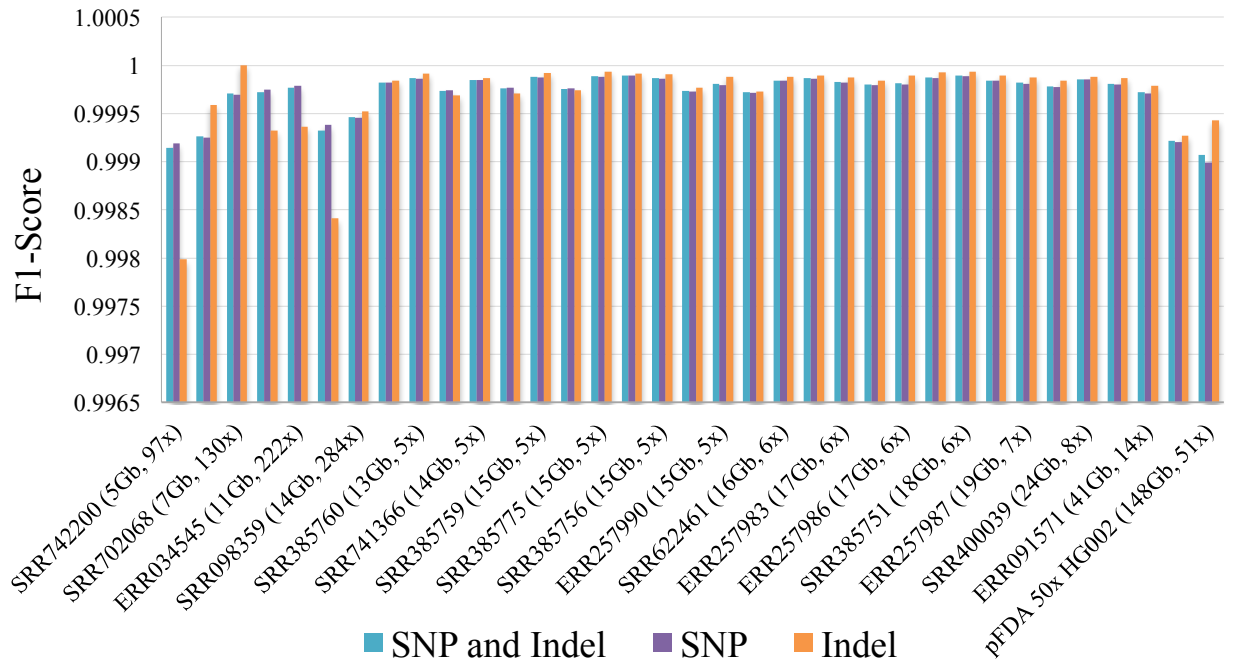
Figure 4. Consistency of Sentieon DNAseq and the GATK Best Practices.
F1-Score between the Sentieon DNAseq and GATK Best Practices variant calls is shown for combined SNPs and Indels and each separately. F1-score was measured using RTG tools vcfeval.

Discussion

In this manuscript, we present the Sentieon DNAseq, TNseq and TNHaplotyper pipelines. The Sentieon tools provide computationally efficient, multithreaded deterministic variant calling from germline samples and tumor-normal pairs with results identical to the GATK, MuTect and MuTect2 while providing a greater than 10-fold improvement in total runtime. As it provides identical results, the Sentieon tools can function as a drop-in replacement for the GATK, MuTect or MuTect2 resulting in cost-savings for researchers and clinicians. The lack of down-sampling enables high-depth sequencing for increased accuracy, the robust implementation enables joint genotyping of 100,000s of files simultaneously without intermediate file merging, and the algorithmic determinism provides the consistency required for medical applications. As next-generation sequence data is produced at an increasing rate and more frequently finds use in

important economic and clinical applications, the Sentieon Genomics pipeline tools provide a means to process this data with accuracy, efficiency and consistency.

Acknowledgements

References

[1] Z. D. Stephens *et al.*, "Big Data: Astronomical or Genomical?," *PLoS Biol.*, vol. 13, no. 7, p. e1002195, Jul. 2015.

[2] NHGRI, "The Cost of Sequencing a Human Genome." [Online]. Available: https://www.genome.gov/sequencingcosts/. [Accessed: 01-Feb-2017].

[3] Illumina, "HiSeq X Series of Sequencing Systems." [Online]. Available: https://www.illumina.com/content/dam/illumina-marketing/documents/products/datasheets/datasheet-hiseq-x-ten.pdf. [Accessed: 30-Jan-2017].

[4] D. W. Bianchi *et al.*, "DNA Sequencing versus Standard Prenatal Aneuploidy Screening," *N. Engl. J. Med.*, vol. 370, no. 9, pp. 799–808, Feb. 2014.

[5] E. A. Ashley, "Towards precision medicine," *Nat. Rev. Genet.*, vol. 17, no. 9, pp. 507–522, Aug. 2016.

[6] N. Hacohen, E. F. Fritsch, T. A. Carter, E. S. Lander, and C. J. Wu, "Getting Personal with Neoantigen-Based Therapeutic Cancer Vaccines," *Cancer Immunol. Res.*, vol. 1, no. 1, pp. 11–15, Jul. 2013.

[7] J. P. Finnigan, A. Rubinsteyn, J. Hammerbacher, and N. Bhardwaj, "Mutation-Derived Tumor Antigens: Novel Targets in Cancer Immunotherapy," *Oncol. Williston Park N*, vol. 29, no. 12, Dec. 2015.

[8] M. A. Hamburg and F. S. Collins, "The Path to Personalized Medicine," *N. Engl. J. Med.*, vol. 363, no. 4, pp. 301–304, Jul. 2010.

[9] T. N. Schumacher and R. D. Schreiber, "Neoantigens in cancer immunotherapy," *Science*, vol. 348, no. 6230, pp. 69–74, Apr. 2015.

[10]   S. J. Aronson and H. L. Rehm, "Building the foundation for genomics in precision medicine," *Nature*, vol. 526, no. 7573, pp. 336–342, Oct. 2015.

[11]   J. Thevenon *et al.*, "Diagnostic odyssey in severe neurodevelopmental disorders: toward clinical whole-exome sequencing as a first-line diagnostic test: Diagnostic odyssey in severe neurodevelopmental disorders," *Clin. Genet.*, vol. 89, no. 6, pp. 700–707, Jun. 2016.

[12]   S. B. Ng *et al.*, "Exome sequencing identifies the cause of a mendelian disorder," *Nat. Genet.*, vol. 42, no. 1, pp. 30–35, Jan. 2010.

[13]   K. M. Boycott, M. R. Vanstone, D. E. Bulman, and A. E. MacKenzie, "Rare-disease genetics in the era of next-generation sequencing: discovery to translation," *Nat. Rev. Genet.*, vol. 14, no. 10, pp. 681–691, Sep. 2013.

[14]   I. Iossifov *et al.*, "The contribution of de novo coding mutations to autism spectrum disorder," *Nature*, vol. 515, no. 7526, pp. 216–221, Nov. 2014.

[15]   H. D. Daetwyler *et al.*, "Whole-genome sequencing of 234 bulls facilitates mapping of monogenic and complex traits in cattle," *Nat. Genet.*, vol. 46, no. 8, pp. 858–865, Jul. 2014.

[16]   R. B. Altman *et al.*, "A research roadmap for next-generation sequencing informatics," *Sci. Transl. Med.*, vol. 8, no. 335, p. 335ps10, Apr. 2016.

[17]   H. Li, J. Ruan, and R. Durbin, "Mapping short DNA sequencing reads and calling variants using mapping quality scores," *Genome Res.*, vol. 18, no. 11, pp. 1851–1858, Nov. 2008.

[18]   R. Li *et al.*, "SNP detection for massively parallel whole-genome resequencing," *Genome Res.*, vol. 19, no. 6, pp. 1124–1132, Jun. 2009.

[19]   J. Wang *et al.*, "The diploid genome sequence of an Asian individual," *Nature*, vol. 456, no. 7218, pp. 60–65, Nov. 2008.

[20]   T. J. Ley *et al.*, "DNA sequencing of a cytogenetically normal acute myeloid leukaemia genome," *Nature*, vol. 456, no. 7218, pp. 66–72, Nov. 2008.

[21]   E. Garrison and G. Marth, "Haplotype-based variant detection from short-read sequencing."

[22]   A. Rimmer *et al.*, "Integrating mapping, assembly and haplotype-based approaches for calling variants in clinical sequencing applications," *Nat. Genet.*, vol. 46, no. 8, pp. 912–918, Aug. 2014.

[23]   M. A. DePristo *et al.*, "A framework for variation discovery and genotyping using next-generation DNA sequencing data," *Nat. Genet.*, vol. 43, no. 5, pp. 491–498, May 2011.

[24]   A. McKenna *et al.*, "The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data," *Genome Res.*, vol. 20, no. 9, pp. 1297–1303, Sep. 2010.

[25]   K. Cibulskis *et al.*, "Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples," *Nat. Biotechnol.*, vol. 31, no. 3, pp. 213–219, Mar. 2013.

[26]   A. Auton *et al.*, "A global reference for human genetic variation," *Nature*, vol. 526, no. 7571, pp. 68–74, Sep. 2015.

[27]   N. Stransky *et al.*, "The Mutational Landscape of Head and Neck Squamous Cell Carcinoma," *Science*, vol. 333, no. 6046, pp. 1157–1160, Aug. 2011.

[28]   "precisionFDA." [Online]. Available: https://precision.fda.gov/. [Accessed: 27-Jan-2017].

[29]   A. D. Ewing *et al.*, "Combining tumor genome simulation with crowdsourcing to benchmark somatic single-nucleotide-variant detection," *Nat. Methods*, vol. 12, no. 7, pp. 623–630, May 2015.

[30]   H. Li and R. Durbin, "Fast and accurate short read alignment with Burrows-Wheeler transform," *Bioinformatics*, vol. 25, no. 14, pp. 1754–1760, Jul. 2009.

[31]   H. Li, "Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM," *ArXiv Prepr. ArXiv13033997*, 2013.

[32]   H. Li *et al.*, "The Sequence Alignment/Map format and SAMtools," *Bioinformatics*, vol. 25, no. 16, pp. 2078–2079, Aug. 2009.

[33]   Jessica A. Weber, Rafael Aldana, Brendan D. Gallagher, and Jeremy S. Edwards, "Sentieon DNA pipeline for variant detection - Software-only solution, over 20× faster than GATK 3.3 with identical results," 2016.

Supplementary Material

Supplementary Appendix 1.

DNAseq Pipeline – Single Sample

Analysis Step	GATK Best Practices Pipeline	Sentieon DNAseq
Read alignment	bwa mem -M -R \"@RG\tID:$group\tSM:$sample\tPL:$pl\" -t 32 $fasta $fastq_1 $fastq_2 | samtools view -Sb - >align.bam samtools sort -@ 32 align.bam sorted	bwa mem -M -R \"@RG\tID:$group\tSM:$sample\tPL:$pl\" -t 32 $fasta $fastq_1 $fastq_2 | sentieon util sort -o sorted.bam -t 32 --sam2bam -i –
Sample metrics collection	java -jar ${picard}/CollectAlignmentSummaryMetrics.jar INPUT=sorted.bam OUTPUT=aln_metrics.txt REFERENCE_SEQUENCE=$fasta ADAPTER_SEQUENCE=null VALIDATION_STRINGENCY=SILENT java -jar ${picard}/CollectGcBiasMetrics.jar INPUT=sorted.bam OUTPUT=gc_metrics.txt SUMMARY_OUTPUT=gc_summary.txt CHART_OUTPUT=gcbias.pdf REFERENCE_SEQUENCE=$fasta ASSUME_SORTED=true VALIDATION_STRINGENCY=SILENT java -jar ${picard}/MeanQualityByCycle.jar INPUT=sorted.bam OUTPUT=mq_metrics.txt CHART_OUTPUT=meanq_cycle.pdf REFERENCE_SEQUENCE=$fasta VALIDATION_STRINGENCY=SILENT PF_READS_ONLY=true java -jar ${picard_folder}/QualityScoreDistribution.jar INPUT=sorted.bam OUTPUT=qd_metrics.txt CHART_OUTPUT=qscore_dist.pdf REFERENCE_SEQUENCE=$fasta VALIDATION_STRINGENCY=SILENT PF_READS_ONLY=true java -jar ${picard_folder}/CollectInsertSizeMetrics.jar INPUT=sorted.bam OUTPUT=is_metrics.txt REFERENCE_SEQUENCE=$fasta HISTOGRAM_FILE=is_histogram.txt	sentieon driver -r $fasta -t 32 -i sorted.bam --algo MeanQualityByCycle mq_metrics.txt --algo QualDistribution qd_metrics.txt --algo GCBias --summary gc_summary.txt gc_metrics.txt --algo AlignmentStat aln_metrics.txt --algo InsertSizeMetricAlgo is_metrics.txt sentieon plot metrics -o metrics-report.pdf gc=gc_metrics.txt qd=qd_metrics.txt mq=mq_metrics.txt isize=is_metrics.txt
Duplicate read removal	java -jar ${picard}/MarkDuplicates.jar M=dup_reads I=sorted.bam O=deduped.bam samtools index deduped.bam	sentieon driver -t 32 -i sorted.bam --algo LocusCollector --fun score_info score.txt sentieon driver -t 32 -i sorted.bam --algo Dedup --rmdup --score_info score.txt deduped.bam
Indel realignment	java -jar GenomeAnalysisTK.jar -T RealignerTargetCreator -R $fasta -I deduped.bam -known $Mills -o realigner.intervals java -jar GenomeAnalysisTK.jar -T IndelRealigner -R $fasta -I deduped.bam -known $Mills -targetIntervals realigner.intervals -o realigned.bam	sentieon driver -r $fasta -t 32 -i deduped.bam --algo Realigner -k $Mills realigned.bam
Base quality score recalibration	java -jar GenomeAnalysisTK.jar -T BaseRecalibrator -nct 32 -R $fasta -I realigned.bam -knownSites $dbsnp -knownSites $Mills -o recal.table java -jar GenomeAnalysisTK.jar -T PrintReads -nct 32 -R $fasta -I realigned.bam -BQSR recal.table -o recaled.bam java -jar GenomeAnalysisTK.jar -T BaseRecalibrator -nct 32 -R $fasta -I realigned.bam -knownSites $dbsnp -knownSites $Mills -BQSR recal.table -o after_recal.table java -jar GenomeAnalysisTK.jar -T AnalyzeCovariates -R $fasta -before recal.table -after after_recal.table -plots recal_plots.pdf	sentieon driver -r $fasta -t 32 -i realigned.bam --algo QualCal -k $dbsnp -k $Mills recal_data.table sentieon driver -r $fasta -t 32 -i realigned.bam -q recal_data.table --algo QualCal -k $dbsnp -k $Mills recal_data.table.post --algo ReadWriter recaled.bam sentieon driver -t 32 --algo QualCal --plot --before recal_data.table --after recal_data.table.post recal.csv sentieon plot bqsr -o recal_plots.pdf recal.csv

Analysis Step	GATK Best Practices Pipeline	Sentieon DNAseq
Variant calling - UnifiedGenotyper	java -jar GenomeAnalysisTK.jar -T UnifiedGenotyper [-L $bed] -nt 32 -R $fasta -I recaled.bam -o UG.vcf	sentieon driver -r $fasta [--interval $bed] -t 32 -i realigned.bam -q recal_data.table --algo Genotyper output-ug.vcf
Variant calling - HaplotypeCaller	java -jar GenomeAnalysisTK.jar -T HaplotypeCaller [-L $bed] -nct 32 -R $fasta -I recaled.bam -o HC.vcf	senteion driver -r $fasta [--interval $bed] -t 32 -i realigned.bam -q recal_data.table --algo Haplotyper output-hc.vcf
Comparison	java -jar RTG.jar vcfeval --baseline=HC.vcf --calls=output-hc.vcf --output=$out_dir --template=$sdf_file [--bed-region $bed] --sample=$sample	

Commands used for testing the DNAseq pipeline on single samples are listed. Arguments supplying the bed file were used with whole-exome sequenced samples. The supplied bed file contained Illumina TruSeq capture targets.

DNAseq Pipeline – Joint Calling

Analysis Step	GATK Best Practices	Sentieon DNAseq
Variant calling – Joint Genotyping	java -jar GenomeAnalysisTK.jar -T GenotypeGVCFs -L chr1 -nt 32 -R $fasta $gvcf_list -o GATK_jointcalling_chr1.vcf.gz -maxAltAlleles 100	sentieon driver -r $fasta --interval chr1 -t 32 --algo GVCFtyper $gvcf_list sentieon_jointcalling_chr1.vcf.gz
Comparison for joint calling per sample	java -jar RTG.jar vcfeval --baseline=GATK_jointcalling_chr1.vcf.gz --calls=sentieon_jointcalling_chr1.vcf.gz --output=$out_dir --template=$sdf_file [--bed-region $bed] --sample=$sample	

Commands used for testing the joint calling with the DNAseq pipeline. For the consistency check, the comparison command was run for all 75 genotyped samples.

Tumor-Normal Pipeline

Analysis Step	MuTect and MuTect2	Sentieon TNseq and TNhaplotyper
Read alignment	bwa mem -R '@RG\tID:$group\tSM:$sample\tPL:$pl' -t 32 -M -K 10000000 $fasta '<bzip2 -dc $fastq_1.bz2' '<bzip2 -dc $fastq_2.bz2' | samtools sort -@ 32 -o sorted.bam -T sorted samtools index sorted.bam	bwa mem -R '@RG\tID:$group\tSM:$sample\tPL:$pl' -t 32 -M -K 10000000 $fasta '<bzip2 -dc $fastq_1.bz2' '<bzip2 -dc $fastq_2.bz2' | sentieon util sort -i - -r $fasta -t 32 -o sorted.bam --sam2bam
Sample metrics collection	java -jar ${picard}/CollectAlignmentSummaryMetrics.jar VALIDATION_STRINGENCY=SILENT INPUT=sorted.bam OUTPUT=aln_metrics.txt REFERENCE_SEQUENCE=$fasta ADAPTER_SEQUENCE=null java -jar ${picard}/CollectGcBiasMetrics.jar VALIDATION_STRINGENCY=SILENT INPUT=sorted.bam OUTPUT=gc_metrics.txt REFERENCE_SEQUENCE=$fasta SUMMARY_OUTPUT=gc_summary.txt CHART_OUTPUT=gc-metrics-report.pdf ASSUME_SORTED=true java -jar ${picard}/MeanQualityByCycle.jar VALIDATION_STRINGENCY=SILENT INPUT=sorted.bam OUTPUT=mq_metrics.txt REFERENCE_SEQUENCE=$fasta CHART_OUTPUT=mq-metrics-report.pdf PF_READS_ONLY=true java -jar ${picard}/QualityScoreDistribution.jar VALIDATION_STRINGENCY=SILENT INPUT=sorted.bam OUTPUT=qd_metrics.txt REFERENCE_SEQUENCE=$fasta CHART_OUTPUT=qd-metrics-report.pdf PF_READS_ONLY=true java -jar ${picard}/CollectInsertSizeMetrics.jar VALIDATION_STRINGENCY=SILENT INPUT=sorted.bam OUTPUT=is_metrics.txt REFERENCE_SEQUENCE=$fasta HISTOGRAM_FILE=metrics-report.pdf	sentieon driver -t 32 -r $fasta -i sorted.bam --algo MeanQualityByCycle mq_metrics.txt --algo InsertSizeMetricAlgo is_metrics.txt --algo QualDistribution qd_metrics.txt --algo GCBias --summary gc_summary.txt gc_metrics.txt --algo AlignmentStat --adapter_seq " aln_metrics.txt sentieon plot metrics -o metrics-report.pdf gc=gc_metrics.txt qd=qd_metrics.txt mq=mq_metrics.txt isize=is_metrics.txt
Duplicate read removal	java -jar ${picard}/MarkDuplicates.jar VALIDATION_STRINGENCY=SILENT INPUT=sorted.bam OUTPUT=deduped.bam METRICS_FILE=dedup_metrics.txt REMOVE_DUPLICATES=true MAX_FILE_HANDLES_FOR_READ_ENDS_MAP=65536 java -jar ${picard}/BuildBamIndex.jar VALIDATION_STRINGENCY=SILENT INPUT=deduped.bam	sentieon driver -t 32 -r $fasta -i sorted.bam --algo LocusCollector --fun score_info deduped.bam.score.gz sentieon driver -t 32 -r $fasta -i sorted.bam --algo Dedup --score_info deduped.bam.score.gz --rmdup --metrics dedup_metrics.txt deduped.bam
Indel realignment	java -jar GenomeAnalysisTK.jar -T RealignerTargetCreator -R $fasta -I deduped.bam -nt 32 -known $Mills -known $onekg_indels -o realigner.intervals java -jar GenomeAnalysisTK.jar -T IndelRealigner -R $fasta -I deduped.bam -known $Mills -known $onekg_indels --targetIntervals realigner.intervals -o realigned.bam	sentieon driver -t 32 -r $fasta -i deduped.bam --algo Realigner -k $Mills -k $onekg_indels realigned.bam
Base quality score recalibration	java -jar GenomeAnalysisTK.jar -T BaseRecalibrator -R $fasta -I realigned.bam -nct 32 --knownSites $dbsnp --knownSites $Mills --knownSites $onekg_indels -o recal_data.table java -jar GenomeAnalysisTK.jar -T PrintReads -R $fasta -I realigned.bam --BQSR recal_data.table -o recaled.bam java -jar GenomeAnalysisTK.jar -T BaseRecalibrator -R $fasta -I realigned.bam --BQSR recal_data.table -nct 32 --knownSites $dbsnp --knownSites $Mills --knownSites $onekg_indels -o recal_data.table.post java -jar GenomeAnalysisTK.jar -T AnalyzeCovariates -R $fasta -csv recal.csv -before recal_data.table -after recal_data.table.post -plots recal_plots.pdf	sentieon driver -t 32 -r $fasta -i realigned.bam --algo QualCal -k $dbsnp -k $Mills -k $onekg_indels recal_data.table sentieon driver -t 32 -r $fasta -i realigned.bam -q recal_data.table --algo QualCal -k $dbsnp -k $Mills -k $onekg_indels recal_data.table.post sentieon driver --passthru --algo QualCal -k $dbsnp -k $Mills -k $onekg_indels --before recal_data.table --after recal_data.table.post --plot recal.csv sentieon plot bqsr -o recal_plots.pdf recal.csv
Indel joint realignment	java -jar GenomeAnalysisTK.jar -T RealignerTargetCreator -R $fasta -I tumor_recaled.bam -I normal_recaled.bam -nt 32 -known $Mills -known $onekg_indels -o interval.list java -jar GenomeAnalysisTK.jar -T IndelRealigner -R $fasta -I tumor_recaled.bam -I normal_recaled.bam -known $Mills -known $onekg_indels --targetIntervals interval.list -nWayOut _corealigned.bam	sentieon driver -t 32 -r $fasta -i tumor_realigned.bam -i normal_realigned.bam -q tumor_recal_data.table -q normal_recal_data.table --algo Realigner -k $Mills -k $onekg_indels tn_corealigned.bam
Variant	java -jar mutect-1.1.5.jar -T MuTect -R $fasta --dbsnp $dbsnp -o output-	sentieon driver -t 32 -r $fasta -i tn_corealigned.bam --

Analysis Step	MuTect and MuTect2	Sentieon TNseq and TNhaplotyper
calling - MuTect	call.stats -vcf mutect.vcf.gz -I:tumor tumor_recaled_corealigned.bam -I:normal normal_recaled_corealigned.bam	algo TNsnv --tumor_sample $tumor_sample_name --normal_sample $normal_sample_name --dbsnp $dbsnp --call_stats_out output-call.stats output-tnsnv.vcf.gz
Variant calling - MuTect2	java -jar GenomeAnalysisTK.jar -T MuTect2 -R $fasta --dbsnp $dbsnp -o mutect2.vcf.gz -I:tumor tumor_recaled_corealigned.bam -I:normal normal_recaled_corealigned.bam -nct 32	sentieon driver -t 32 -r $fasta -i tn_corealigned.bam --algo TNhaplotyper --tumor_sample tumor_sample_name --normal_sample normal_sample_name --dbsnp $dbsnp output-tnhaplotyper.vcf.gz
Comparison	java -jar RTG.jar vcfeval --baseline=mutect2.vcf.gz --calls=tnhaplotyper.vcf.gz --output=$out_dir --template=$sdf_file [--bed-region $bed]	

Commands used in the tumor-normal benchmarking pipelines. The --bed-region argument was used for analysis of the whole-exome sequenced sample. The supplied bed file contained Illumina TruSeq capture targets.

Supplementary Table 1.

Sample Name	Sequencing Type	Run Name	Pipeline	Germline or Tumor-Normal	Source
NA06984	LowCov	SRR035025	DNAseq	Germline	ftp://ftp.sra.ebi.ac.uk/vol1/fastq/SRR035/SRR035025/SRR035025_1.fastq.gz,ftp://ftp.sra.ebi.ac.uk/vol1/fastq/SRR035/SRR035025/SRR035025_2.fastq.gz
NA06985	Exome	SRR709972	DNAseq	Germline	ftp://ftp.sra.ebi.ac.uk/vol1/fastq/SRR709/SRR709972/SRR709972_1.fastq.gz,ftp://ftp.sra.ebi.ac.uk/vol1/fastq/SRR709/SRR709972/SRR709972_2.fastq.gz
NA06985	LowCov	SRR400039	DNAseq	Germline	ftp://ftp.sra.ebi.ac.uk/vol1/fastq/SRR400/SRR400039/SRR400039_1.fastq.gz,ftp://ftp.sra.ebi.ac.uk/vol1/fastq/SRR400/SRR400039/SRR400039_2.fastq.gz
NA06986	LowCov	SRR027520	DNAseq	Germline	ftp://ftp.sra.ebi.ac.uk/vol1/fastq/SRR027/SRR027520/SRR027520_1.fastq.gz,ftp://ftp.sra.ebi.ac.uk/vol1/fastq/SRR027/SRR027520/SRR027520_2.fastq.gz
NA06989	LowCov	SRR035028	DNAseq	Germline	ftp://ftp.sra.ebi.ac.uk/vol1/fastq/SRR035/SRR035028/SRR035028_1.fastq.gz,ftp://ftp.sra.ebi.ac.uk/vol1/fastq/SRR035/SRR035028/SRR035028_2.fastq.gz
NA06994	LowCov	SRR385751	DNAseq	Germline	ftp://ftp.sra.ebi.ac.uk/vol1/fastq/SRR385/SRR385751/SRR385751_1.fastq.gz,ftp://ftp.sra.ebi.ac.uk/vol1/fastq/SRR385/SRR385751/SRR385751_2.fastq.gz
NA07000	LowCov	ERR162816	DNAseq	Germline	ftp://ftp.sra.ebi.ac.uk/vol1/fastq/ERR162/ERR162816/ERR162816_1.fastq.gz,ftp://ftp.sra.ebi.ac.uk/vol1/fastq/ERR162/ERR162816/ERR162816_2.fastq.gz
NA07037	LowCov	ERR257983	DNAseq	Germline	ftp://ftp.sra.ebi.ac.uk/vol1/fastq/ERR257/ERR257983/ERR257983_1.fastq.gz,ftp://ftp.sra.ebi.ac.uk/vol1/fastq/ERR257/ERR257983/ERR257983_2.fastq.gz
NA07048	LowCov	SRR032636	DNAseq	Germline	ftp://ftp.sra.ebi.ac.uk/vol1/fastq/SRR032/SRR032636/SRR032636_1.fastq.gz,ftp://ftp.sra.ebi.ac.uk/vol1/fastq/SRR032/SRR032636/SRR032636_2.fastq.gz
NA07051	LowCov	SRR026644	DNAseq	Germline	ftp://ftp.sra.ebi.ac.uk/vol1/fastq/SRR026/SRR026644/SRR026644_1.fastq.gz,ftp://ftp.sra.ebi.ac.uk/vol1/fastq/SRR026/SRR026644/SRR026644_2.fastq.gz
NA07056	LowCov	ERR003038	DNAseq	Germline	ftp://ftp.sra.ebi.ac.uk/vol1/fastq/ERR003/ERR003038/ERR003038_1.fastq.gz,ftp://ftp.sra.ebi.ac.uk/vol1/fastq/ERR003/ERR003038/ERR003038_2.fastq.gz
NA07347	LowCov	SRR029852	DNAseq	Germline	ftp://ftp.sra.ebi.ac.uk/vol1/fastq/SRR029/SRR029852/SRR029852_1.fastq.gz,ftp://ftp.sra.ebi.ac.uk/vol1/fastq/SRR029/SRR029852/SRR029852_2.fastq.gz
NA07357	LowCov	ERR001509	DNAseq	Germline	ftp://ftp.sra.ebi.ac.uk/vol1/fastq/ERR001/ERR001509/ERR001509_1.fastq.gz,ftp://ftp.sra.ebi.ac.uk/vol1/fastq/ERR001/ERR001509/ERR001509_2.fastq.gz
NA10851	LowCov	ERR001502	DNAseq	Germline	ftp://ftp.sra.ebi.ac.uk/vol1/fastq/ERR001/ERR001502/ERR001502_1.fastq.gz,ftp://ftp.sra.ebi.ac.uk/vol1/fastq/ERR001/ERR001502/ERR001502_2.fastq.gz
NA11831	LowCov	SRR400037	DNAseq	Germline	ftp://ftp.sra.ebi.ac.uk/vol1/fastq/SRR400/SRR400037/SRR400037_1.fastq.gz,ftp://ftp.sra.ebi.ac.uk/vol1/fastq/SRR400/SRR400037/SRR400037_2.fastq.gz
NA11832	LowCov	SRR385763	DNAseq	Germline	ftp://ftp.sra.ebi.ac.uk/vol1/fastq/SRR385/SRR385763/SRR385763_1.fastq.gz,ftp://ftp.sra.ebi.ac.uk/vol1/fastq/SRR385/SRR385763/SRR385763_2.fastq.gz
NA11840	LowCov	SRR393988	DNAseq	Germline	ftp://ftp.sra.ebi.ac.uk/vol1/fastq/SRR393/SRR393988/SRR393988_1.fastq.gz,ftp://ftp.sra.ebi.ac.uk/vol1/fastq/SRR393/SRR393988/SRR393988_2.fastq.gz
NA11843	LowCov	SRR035016	DNAseq	Germline	ftp://ftp.sra.ebi.ac.uk/vol1/fastq/SRR035/SRR035016/SRR035016_1.fastq.gz,ftp://ftp.sra.ebi.ac.uk/vol1/fastq/SRR035/SRR035016/SRR035016_2.fastq.gz
NA11881	LowCov	ERR050083	DNAseq	Germline	ftp://ftp.sra.ebi.ac.uk/vol1/fastq/ERR050/ERR050083/ERR050083_1.fastq.gz,ftp://ftp.sra.ebi.ac.uk/vol1/fastq/ERR050/ERR050083/ERR050083_2.fastq.gz
NA11892	LowCov	ERR015477	DNAseq	Germline	ftp://ftp.sra.ebi.ac.uk/vol1/fastq/ERR015/ERR015477/ERR015477_1.fastq.gz,ftp://ftp.sra.ebi.ac.uk/vol1/fastq/ERR015/ERR015477/ERR015477_2.fastq.gz
NA11893	LowCov	ERR009416	DNAseq	Germline	ftp://ftp.sra.ebi.ac.uk/vol1/fastq/ERR009/ERR009416/ERR009416_1.fastq.gz,ftp://ftp.sra.ebi.ac.uk/vol1/fastq/ERR009/ERR009416/ERR009416_2.fastq.gz
NA11894	LowCov	ERR000876	DNAseq	Germline	ftp://ftp.sra.ebi.ac.uk/vol1/fastq/ERR000/ERR000876/ERR000876_1.fastq.gz,ftp://ftp.sra.ebi.ac.uk/vol1/fastq/ERR000/ERR000876/ERR000876_2.fastq.gz
NA11918	LowCov	SRR006107	DNAseq	Germline	ftp://ftp.sra.ebi.ac.uk/vol1/fastq/SRR006/SRR006107/SRR006107_1.fastq.gz,ftp://ftp.sra.ebi.ac.uk/vol1/fastq/SRR006/SRR006107/SRR006107_2.fastq.gz
NA11919	LowCov	SRR027518	DNAseq	Germline	ftp://ftp.sra.ebi.ac.uk/vol1/fastq/SRR027/SRR027518/SRR027518_1.fastq.gz,ftp://ftp.sra.ebi.ac.uk/vol1/fastq/SRR027/SRR027518/SRR027518_2.fastq.gz
NA11920	LowCov	SRR024100	DNAseq	Germline	ftp://ftp.sra.ebi.ac.uk/vol1/fastq/SRR024/SRR024100/SRR024100_1.fastq.gz,ftp://ftp.sra.ebi.ac.uk/vol1/fastq/SRR024/SRR024100/SRR024100_2.fastq.gz
NA11930	Exome	SRR098416	DNAseq	Germline	ftp://ftp.sra.ebi.ac.uk/vol1/fastq/SRR098/SRR098416/SRR098416_1.fastq.gz,ftp://ftp.sra.ebi.ac.uk/vol1/fastq/SRR098/SRR098416/SRR098416_2.fastq.gz
NA11930	LowCov	ERR233226	DNAseq	Germline	ftp://ftp.sra.ebi.ac.uk/vol1/fastq/ERR233/ERR233226/ERR233226_1.fastq.gz,ftp://ftp.sra.ebi.ac.uk/vol1/fastq/ERR233/ERR233226/ERR233226_2.fastq.gz
NA11931	LowCov	ERR000902	DNAseq	Germline	ftp://ftp.sra.ebi.ac.uk/vol1/fastq/ERR000/ERR000902/ERR000902_1.fastq.gz,ftp://ftp.sra.ebi.ac.uk/vol1/fastq/ERR000/ERR000902/ERR000902_2.fastq.gz
NA11932	Exome	ERR034544	DNAseq	Germline	ftp://ftp.sra.ebi.ac.uk/vol1/fastq/ERR034/ERR034544/ERR034544_1.fastq.gz,ftp://ftp.sra.ebi.ac.uk/vol1/fastq/ERR034/ERR034544/ERR034544_2.fastq.gz
NA11933	Exome	ERR034545	DNAseq	Germline	ftp://ftp.sra.ebi.ac.uk/vol1/fastq/ERR034/ERR034545/ERR034545_1.fastq.gz,ftp://ftp.sra.ebi.ac.uk/vol1/fastq/ERR034/ERR034545/ERR034545_2.fastq.gz
NA11992	LowCov	SRR385759	DNAseq	Germline	ftp://ftp.sra.ebi.ac.uk/vol1/fastq/SRR385/SRR385759/SRR385759_1.fastq.gz,ftp://ftp.sra.ebi.ac.uk/vol1/fastq/SRR385/SRR385759/SRR385759_2.fastq.gz
NA11994	LowCov	SRR385753	DNAseq	Germline	ftp://ftp.sra.ebi.ac.uk/vol1/fastq/SRR385/SRR385753/SRR385753_1.fastq.gz,ftp://ftp.sra.ebi.ac.uk/vol1/fastq/SRR385/SRR385753/SRR385753_2.fastq.gz
NA11995	LowCov	SRR385754	DNAseq	Germline	ftp://ftp.sra.ebi.ac.uk/vol1/fastq/SRR385/SRR385754/SRR385754_1.fastq.gz,ftp://ftp.sra.ebi.ac.uk/vol1/fastq/SRR385/SRR385754/SRR385754_2.fastq.gz
NA12003	LowCov	SRR385756	DNAseq	Germline	ftp://ftp.sra.ebi.ac.uk/vol1/fastq/SRR385/SRR385756/SRR385756_1.fastq.gz,ftp://ftp.sra.ebi.ac.uk/vol1/fastq/SRR385/SRR385756/SRR385756_2.fastq.gz

NA12004	LowCov	SRR385757	DNAseq	Germline	ftp://ftp.sra.ebi.ac.uk/vol1/fastq/SRR385/SRR385757/SRR385757_1.fastq.gz,ftp://ftp.sra.ebi.ac.uk/vol1/fastq/SRR385/SRR385757/SRR385757_2.fastq.gz
NA12005	LowCov	SRR385767	DNAseq	Germline	ftp://ftp.sra.ebi.ac.uk/vol1/fastq/SRR385/SRR385767/SRR385767_1.fastq.gz,ftp://ftp.sra.ebi.ac.uk/vol1/fastq/SRR385/SRR385767/SRR385767_2.fastq.gz
NA12006	LowCov	SRR385760	DNAseq	Germline	ftp://ftp.sra.ebi.ac.uk/vol1/fastq/SRR385/SRR385760/SRR385760_1.fastq.gz,ftp://ftp.sra.ebi.ac.uk/vol1/fastq/SRR385/SRR385760/SRR385760_2.fastq.gz
NA12043	LowCov	ERR000228	DNAseq	Germline	ftp://ftp.sra.ebi.ac.uk/vol1/fastq/ERR000/ERR000228/ERR000228_1.fastq.gz,ftp://ftp.sra.ebi.ac.uk/vol1/fastq/ERR000/ERR000228/ERR000228_2.fastq.gz
NA12044	LowCov	SRR393991	DNAseq	Germline	ftp://ftp.sra.ebi.ac.uk/vol1/fastq/SRR393/SRR393991/SRR393991_1.fastq.gz,ftp://ftp.sra.ebi.ac.uk/vol1/fastq/SRR393/SRR393991/SRR393991_2.fastq.gz
NA12045	LowCov	SRR027531	DNAseq	Germline	ftp://ftp.sra.ebi.ac.uk/vol1/fastq/SRR027/SRR027531/SRR027531_1.fastq.gz,ftp://ftp.sra.ebi.ac.uk/vol1/fastq/SRR027/SRR027531/SRR027531_2.fastq.gz
NA12046	Exome	SRR742200	DNAseq	Germline	ftp://ftp.sra.ebi.ac.uk/vol1/fastq/SRR742/SRR742200/SRR742200_1.fastq.gz,ftp://ftp.sra.ebi.ac.uk/vol1/fastq/SRR742/SRR742200/SRR742200_2.fastq.gz
NA12046	LowCov	ERR239333	DNAseq	Germline	ftp://ftp.sra.ebi.ac.uk/vol1/fastq/ERR239/ERR239333/ERR239333_1.fastq.gz,ftp://ftp.sra.ebi.ac.uk/vol1/fastq/ERR239/ERR239333/ERR239333_2.fastq.gz
NA12058	LowCov	SRR035020	DNAseq	Germline	ftp://ftp.sra.ebi.ac.uk/vol1/fastq/SRR035/SRR035020/SRR035020_1.fastq.gz,ftp://ftp.sra.ebi.ac.uk/vol1/fastq/SRR035/SRR035020/SRR035020_2.fastq.gz
NA12144	LowCov	ERR162832	DNAseq	Germline	ftp://ftp.sra.ebi.ac.uk/vol1/fastq/ERR162/ERR162832/ERR162832_1.fastq.gz,ftp://ftp.sra.ebi.ac.uk/vol1/fastq/ERR162/ERR162832/ERR162832_2.fastq.gz
NA12154	LowCov	SRR385769	DNAseq	Germline	ftp://ftp.sra.ebi.ac.uk/vol1/fastq/SRR385/SRR385769/SRR385769_1.fastq.gz,ftp://ftp.sra.ebi.ac.uk/vol1/fastq/SRR385/SRR385769/SRR385769_2.fastq.gz
NA12155	Exome	SRR702068	DNAseq	Germline	ftp://ftp.sra.ebi.ac.uk/vol1/fastq/SRR702/SRR702068/SRR702068_1.fastq.gz,ftp://ftp.sra.ebi.ac.uk/vol1/fastq/SRR702/SRR702068/SRR702068_2.fastq.gz
NA12156	LowCov	SRR393992	DNAseq	Germline	ftp://ftp.sra.ebi.ac.uk/vol1/fastq/SRR393/SRR393992/SRR393992_1.fastq.gz,ftp://ftp.sra.ebi.ac.uk/vol1/fastq/SRR393/SRR393992/SRR393992_2.fastq.gz
NA12249	LowCov	SRR211278	DNAseq	Germline	ftp://ftp.sra.ebi.ac.uk/vol1/fastq/SRR211/SRR211278/SRR211278_1.fastq.gz,ftp://ftp.sra.ebi.ac.uk/vol1/fastq/SRR211/SRR211278/SRR211278_2.fastq.gz
NA12272	LowCov	ERR019488	DNAseq	Germline	ftp://ftp.sra.ebi.ac.uk/vol1/fastq/ERR019/ERR019488/ERR019488_1.fastq.gz,ftp://ftp.sra.ebi.ac.uk/vol1/fastq/ERR019/ERR019488/ERR019488_2.fastq.gz
NA12273	LowCov	ERR162833	DNAseq	Germline	ftp://ftp.sra.ebi.ac.uk/vol1/fastq/ERR162/ERR162833/ERR162833_1.fastq.gz,ftp://ftp.sra.ebi.ac.uk/vol1/fastq/ERR162/ERR162833/ERR162833_2.fastq.gz
NA12275	LowCov	ERR009417	DNAseq	Germline	ftp://ftp.sra.ebi.ac.uk/vol1/fastq/ERR009/ERR009417/ERR009417_1.fastq.gz,ftp://ftp.sra.ebi.ac.uk/vol1/fastq/ERR009/ERR009417/ERR009417_2.fastq.gz
NA12287	LowCov	ERR000854	DNAseq	Germline	ftp://ftp.sra.ebi.ac.uk/vol1/fastq/ERR000/ERR000854/ERR000854_1.fastq.gz,ftp://ftp.sra.ebi.ac.uk/vol1/fastq/ERR000/ERR000854/ERR000854_2.fastq.gz
NA12340	LowCov	SRR075006	DNAseq	Germline	ftp://ftp.sra.ebi.ac.uk/vol1/fastq/SRR075/SRR075006/SRR075006_1.fastq.gz,ftp://ftp.sra.ebi.ac.uk/vol1/fastq/SRR075/SRR075006/SRR075006_2.fastq.gz
NA12341	LowCov	SRR034932	DNAseq	Germline	ftp://ftp.sra.ebi.ac.uk/vol1/fastq/SRR034/SRR034932/SRR034932_1.fastq.gz,ftp://ftp.sra.ebi.ac.uk/vol1/fastq/SRR034/SRR034932/SRR034932_2.fastq.gz
NA12347	LowCov	ERR009404	DNAseq	Germline	ftp://ftp.sra.ebi.ac.uk/vol1/fastq/ERR009/ERR009404/ERR009404_1.fastq.gz,ftp://ftp.sra.ebi.ac.uk/vol1/fastq/ERR009/ERR009404/ERR009404_2.fastq.gz
NA12348	LowCov	ERR009412	DNAseq	Germline	ftp://ftp.sra.ebi.ac.uk/vol1/fastq/ERR009/ERR009412/ERR009412_1.fastq.gz,ftp://ftp.sra.ebi.ac.uk/vol1/fastq/ERR009/ERR009412/ERR009412_2.fastq.gz
NA12383	LowCov	ERR013164	DNAseq	Germline	ftp://ftp.sra.ebi.ac.uk/vol1/fastq/ERR013/ERR013164/ERR013164_1.fastq.gz,ftp://ftp.sra.ebi.ac.uk/vol1/fastq/ERR013/ERR013164/ERR013164_2.fastq.gz
NA12399	LowCov	SRR029683	DNAseq	Germline	ftp://ftp.sra.ebi.ac.uk/vol1/fastq/SRR029/SRR029683/SRR029683_1.fastq.gz,ftp://ftp.sra.ebi.ac.uk/vol1/fastq/SRR029/SRR029683/SRR029683_2.fastq.gz
NA12400	LowCov	SRR029680	DNAseq	Germline	ftp://ftp.sra.ebi.ac.uk/vol1/fastq/SRR029/SRR029680/SRR029680_1.fastq.gz,ftp://ftp.sra.ebi.ac.uk/vol1/fastq/SRR029/SRR029680/SRR029680_2.fastq.gz
NA12413	LowCov	ERR003045	DNAseq	Germline	ftp://ftp.sra.ebi.ac.uk/vol1/fastq/ERR003/ERR003045/ERR003045_1.fastq.gz,ftp://ftp.sra.ebi.ac.uk/vol1/fastq/ERR003/ERR003045/ERR003045_2.fastq.gz
NA12414	LowCov	SRR385771	DNAseq	Germline	ftp://ftp.sra.ebi.ac.uk/vol1/fastq/SRR385/SRR385771/SRR385771_1.fastq.gz,ftp://ftp.sra.ebi.ac.uk/vol1/fastq/SRR385/SRR385771/SRR385771_2.fastq.gz
NA12489	LowCov	SRR393994	DNAseq	Germline	ftp://ftp.sra.ebi.ac.uk/vol1/fastq/SRR393/SRR393994/SRR393994_1.fastq.gz,ftp://ftp.sra.ebi.ac.uk/vol1/fastq/SRR393/SRR393994/SRR393994_2.fastq.gz
NA12716	LowCov	ERR257986	DNAseq	Germline	ftp://ftp.sra.ebi.ac.uk/vol1/fastq/ERR257/ERR257986/ERR257986_1.fastq.gz,ftp://ftp.sra.ebi.ac.uk/vol1/fastq/ERR257/ERR257986/ERR257986_2.fastq.gz
NA12717	LowCov	ERR257987	DNAseq	Germline	ftp://ftp.sra.ebi.ac.uk/vol1/fastq/ERR257/ERR257987/ERR257987_1.fastq.gz,ftp://ftp.sra.ebi.ac.uk/vol1/fastq/ERR257/ERR257987/ERR257987_2.fastq.gz
NA12718	LowCov	ERR019486	DNAseq	Germline	ftp://ftp.sra.ebi.ac.uk/vol1/fastq/ERR019/ERR019486/ERR019486_1.fastq.gz,ftp://ftp.sra.ebi.ac.uk/vol1/fastq/ERR019/ERR019486/ERR019486_2.fastq.gz
NA12749	LowCov	ERR162824	DNAseq	Germline	ftp://ftp.sra.ebi.ac.uk/vol1/fastq/ERR162/ERR162824/ERR162824_1.fastq.gz,ftp://ftp.sra.ebi.ac.uk/vol1/fastq/ERR162/ERR162824/ERR162824_2.fastq.gz
NA12760	LowCov	SRR385773	DNAseq	Germline	ftp://ftp.sra.ebi.ac.uk/vol1/fastq/SRR385/SRR385773/SRR385773_1.fastq.gz,ftp://ftp.sra.ebi.ac.uk/vol1/fastq/SRR385/SRR385773/SRR385773_2.fastq.gz
NA12761	LowCov	ERR257989	DNAseq	Germline	ftp://ftp.sra.ebi.ac.uk/vol1/fastq/ERR257/ERR257989/ERR257989_1.fastq.gz,ftp://ftp.sra.ebi.ac.uk/vol1/fastq/ERR257/ERR257989/ERR257989_2.fastq.gz
NA12762	LowCov	SRR385775	DNAseq	Germline	ftp://ftp.sra.ebi.ac.uk/vol1/fastq/SRR385/SRR385775/SRR385775_1.fastq.gz,ftp://ftp.sra.ebi.ac.uk/vol1/fastq/SRR385/SRR385775/SRR385775_2.fastq.gz
NA12763	LowCov	ERR257990	DNAseq	Germline	ftp://ftp.sra.ebi.ac.uk/vol1/fastq/ERR257/ERR257990/ERR257990_1.fastq.gz,ftp://ftp.sra.ebi.ac.uk/vol1/fastq/ERR257/ERR257990/ERR257990_2.fastq.gz
NA12775	LowCov	ERR009409	DNAseq	Germline	ftp://ftp.sra.ebi.ac.uk/vol1/fastq/ERR009/ERR009409/ERR009409_1.fastq.gz,ftp://ftp.sra.ebi.ac.uk/vol1/fastq/ERR009/ERR009409/ERR009409_2.fastq.gz
NA12776	LowCov	ERR000162	DNAseq	Germline	ftp://ftp.sra.ebi.ac.uk/vol1/fastq/ERR000/ERR000162/ERR000162_1.fastq.gz,ft

					p://ftp.sra.ebi.ac.uk/vol1/fastq/ERR000/ERR000162/ERR000162_2.fastq.gz
NA12777	LowCov	ERR003052	DNAseq	Germline	ftp://ftp.sra.ebi.ac.uk/vol1/fastq/ERR003/ERR003052/ERR003052_1.fastq.gz,ftp://ftp.sra.ebi.ac.uk/vol1/fastq/ERR003/ERR003052/ERR003052_2.fastq.gz
NA12778	LowCov	ERR015474	DNAseq	Germline	ftp://ftp.sra.ebi.ac.uk/vol1/fastq/ERR015/ERR015474/ERR015474_1.fastq.gz,ftp://ftp.sra.ebi.ac.uk/vol1/fastq/ERR015/ERR015474/ERR015474_2.fastq.gz
NA12812	LowCov	SRR741369	DNAseq	Germline	ftp://ftp.sra.ebi.ac.uk/vol1/fastq/SRR741/SRR741369/SRR741369_1.fastq.gz,ftp://ftp.sra.ebi.ac.uk/vol1/fastq/SRR741/SRR741369/SRR741369_2.fastq.gz
NA12813	LowCov	SRR385777	DNAseq	Germline	ftp://ftp.sra.ebi.ac.uk/vol1/fastq/SRR385/SRR385777/SRR385777_1.fastq.gz,ftp://ftp.sra.ebi.ac.uk/vol1/fastq/SRR385/SRR385777/SRR385777_2.fastq.gz
NA12814	LowCov	SRR768303	DNAseq	Germline	ftp://ftp.sra.ebi.ac.uk/vol1/fastq/SRR768/SRR768303/SRR768303_1.fastq.gz,ftp://ftp.sra.ebi.ac.uk/vol1/fastq/SRR768/SRR768303/SRR768303_2.fastq.gz
NA12815	LowCov	SRR768305	DNAseq	Germline	ftp://ftp.sra.ebi.ac.uk/vol1/fastq/SRR768/SRR768305/SRR768305_1.fastq.gz,ftp://ftp.sra.ebi.ac.uk/vol1/fastq/SRR768/SRR768305/SRR768305_2.fastq.gz
NA12828	LowCov	ERR001695	DNAseq	Germline	ftp://ftp.sra.ebi.ac.uk/vol1/fastq/ERR001/ERR001695/ERR001695_1.fastq.gz,ftp://ftp.sra.ebi.ac.uk/vol1/fastq/ERR001/ERR001695/ERR001695_2.fastq.gz
NA12872	LowCov	SRR741366	DNAseq	Germline	ftp://ftp.sra.ebi.ac.uk/vol1/fastq/SRR741/SRR741366/SRR741366_1.fastq.gz,ftp://ftp.sra.ebi.ac.uk/vol1/fastq/SRR741/SRR741366/SRR741366_2.fastq.gz
NA12873	LowCov	SRR768308	DNAseq	Germline	ftp://ftp.sra.ebi.ac.uk/vol1/fastq/SRR768/SRR768308/SRR768308_1.fastq.gz,ftp://ftp.sra.ebi.ac.uk/vol1/fastq/SRR768/SRR768308/SRR768308_2.fastq.gz
NA12874	LowCov	SRR768310	DNAseq	Germline	ftp://ftp.sra.ebi.ac.uk/vol1/fastq/SRR768/SRR768310/SRR768310_1.fastq.gz,ftp://ftp.sra.ebi.ac.uk/vol1/fastq/SRR768/SRR768310/SRR768310_2.fastq.gz
NA12878	Exome	SRR098401	DNAseq	Germline	ftp://ftp.sra.ebi.ac.uk/vol1/fastq/SRR098/SRR098401/SRR098401_1.fastq.gz,ftp://ftp.sra.ebi.ac.uk/vol1/fastq/SRR098/SRR098401/SRR098401_2.fastq.gz
NA12878	LowCov	ERR091571	DNAseq	Germline	ftp://ftp.sra.ebi.ac.uk/vol1/fastq/ERR091/ERR091571/ERR091571_1.fastq.gz,ftp://ftp.sra.ebi.ac.uk/vol1/fastq/ERR091/ERR091571/ERR091571_2.fastq.gz
NA12878	LowCov	SRR622461	DNAseq	Germline	ftp://ftp.sra.ebi.ac.uk/vol1/fastq/SRR622/SRR622461/SRR622461_1.fastq.gz,ftp://ftp.sra.ebi.ac.uk/vol1/fastq/SRR622/SRR622461/SRR622461_2.fastq.gz
NA12891	Exome	SRR098359	DNAseq	Germline	ftp://ftp.sra.ebi.ac.uk/vol1/fastq/SRR098/SRR098359/SRR098359_1.fastq.gz,ftp://ftp.sra.ebi.ac.uk/vol1/fastq/SRR098/SRR098359/SRR098359_2.fastq.gz
HG001	Whole-genome	NA	DNAseq	Germline	NA12878-Garvan-Vial1_R1.fastq.gz,NA12878-Garvan-Vial1_R2.fastq.gz
HG002	Whole-genome	NA	DNAseq	Germline	HG002-NA24385-50x_1.fastq.gz,HG002-NA24385-50x_2.fastq.gz
HGSC_case6_WGS	Whole-genome	NA	TNseq Tnhaplotyper	Tumor	hgsc-sftp1.hgsc.bcm.tmc.edu/case_006/FileTCRBOA6-T-WGS.lane1.read1.fastq.bz2,hgsc-sftp1.hgsc.bcm.tmc.edu/case_006/FileTCRBOA6-T-WGS.lane1.read2.fastq.bz2,hgsc-sftp1.hgsc.bcm.tmc.edu/case_006/FileTCRBOA6-T-WGS.lane2.read1.fastq.bz2,hgsc-sftp1.hgsc.bcm.tmc.edu/case_006/FileTCRBOA6-T-WGS.lane2.read2.fastq.bz2,hgsc-sftp1.hgsc.bcm.tmc.edu/case_006/FileTCRBOA6-T-WGS.lane3.read1.fastq.bz2,hgsc-sftp1.hgsc.bcm.tmc.edu/case_006/FileTCRBOA6-T-WGS.lane3.read2.fastq.bz2,hgsc-sftp1.hgsc.bcm.tmc.edu/case_006/FileTCRBOA6-T-WGS.lane4.read1.fastq.bz2,hgsc-sftp1.hgsc.bcm.tmc.edu/case_006/FileTCRBOA6-T-WGS.lane4.read2.fastq.bz2
HGSC_case6_WGS	Whole-genome	NA	TNseq Tnhaplotyper	Normal	hgsc-sftp1.hgsc.bcm.tmc.edu/case_006/FileTCRBOA6-N-WGS.lane1.read1.fastq.bz2,hgsc-sftp1.hgsc.bcm.tmc.edu/case_006/FileTCRBOA6-N-WGS.lane1.read2.fastq.bz2,hgsc-sftp1.hgsc.bcm.tmc.edu/case_006/FileTCRBOA6-N-WGS.lane2.read1.fastq.bz2,hgsc-sftp1.hgsc.bcm.tmc.edu/case_006/FileTCRBOA6-N-WGS.lane2.read2.fastq.bz2
HGSC_case7	Exome	NA	TNseq Tnhaplotyper	Tumor	hgsc-sftp1.hgsc.bcm.tmc.edu/case_007/FileTCRBOA7-T-WEX.lane1.read1.fastq.bz2,hgsc-sftp1.hgsc.bcm.tmc.edu/case_007/FileTCRBOA7-T-WEX.lane1.read2.fastq.bz2
HGSC_case7	Exome	NA	TNseq Tnhaplotyper	Normal	hgsc-sftp1.hgsc.bcm.tmc.edu/case_007/FileTCRBOA7-N-WEX.lane1.read1.fastq.bz2,hgsc-sftp1.hgsc.bcm.tmc.edu/case_007/FileTCRBOA7-N-WEX.lane1.read2.fastq.bz2
HGSC_case7_WGS	Whole-genome	NA	TNseq Tnhaplotyper	Tumor	hgsc-sftp1.hgsc.bcm.tmc.edu/case_007/FileTCRBOA7-T-WGS.lane1.read1.fastq.bz2,hgsc-sftp1.hgsc.bcm.tmc.edu/case_007/FileTCRBOA7-T-WGS.lane1.read2.fastq.bz2,hgsc-sftp1.hgsc.bcm.tmc.edu/case_007/FileTCRBOA7-T-WGS.lane2.read1.fastq.bz2,hgsc-sftp1.hgsc.bcm.tmc.edu/case_007/FileTCRBOA7-T-WGS.lane2.read2.fastq.bz2,hgsc-sftp1.hgsc.bcm.tmc.edu/case_007/FileTCRBOA7-T-WGS.lane3.read1.fastq.bz2,hgsc-sftp1.hgsc.bcm.tmc.edu/case_007/FileTCRBOA7-T-WGS.lane3.read2.fastq.bz2,hgsc-sftp1.hgsc.bcm.tmc.edu/case_007/FileTCRBOA7-T-WGS.lane4.read1.fastq.bz2,hgsc-

					sftp1.hgsc.bcm.tmc.edu/case_007/FileTCRBOA7-T-WGS.lane4.read2.fastq.bz2
HGSC_ca se7_WGS	Whole-genome	NA	TNseq Tnhaploty per	Normal	hgsc-sftp1.hgsc.bcm.tmc.edu/case_007/FileTCRBOA7-N-WGS.lane1.read1.fastq.bz2,hgsc-sftp1.hgsc.bcm.tmc.edu/case_007/FileTCRBOA7-N-WGS.lane1.read2.fastq.bz2,hgsc-sftp1.hgsc.bcm.tmc.edu/case_007/FileTCRBOA7-N-WGS.lane2.read1.fastq.bz2,hgsc-sftp1.hgsc.bcm.tmc.edu/case_007/FileTCRBOA7-N-WGS.lane2.read2.fastq.bz2

Supplementary Table 2.

Case	ID	Case name	Stage	Improvement	Runtime GATK	Runtime Sentieon
ERR009409	NA12775_1000G_LowCov_ERR009409	ERR009409 (3Gb, 2x)	1 - Read Alignment	1.2	1259	1068
ERR009409	NA12775_1000G_LowCov_ERR009409	ERR009409 (3Gb, 2x)	2 - Calculate Metrics	23.2	649	28
ERR009409	NA12775_1000G_LowCov_ERR009409	ERR009409 (3Gb, 2x)	3 - Remove Duplicates	17.3	814	47
ERR009409	NA12775_1000G_LowCov_ERR009409	ERR009409 (3Gb, 2x)	4 - Indel Realign	51	3723	73
ERR009409	NA12775_1000G_LowCov_ERR009409	ERR009409 (3Gb, 2x)	5 - Base Quality Score Recalibration	41.3	8186	198
ERR009409	NA12775_1000G_LowCov_ERR009409	ERR009409 (3Gb, 2x)	6 - Variant Calling (Unified Genotyper)	50.1	1001	20
ERR009409	NA12775_1000G_LowCov_ERR009409	ERR009409 (3Gb, 2x)	7 - Variant Calling (Haplotype Caller)	34.1	6280	184
ERR009417	NA12275_1000G_LowCov_ERR009417	ERR009417 (4Gb, 2x)	1 - Read Alignment	1.8	1468	827
ERR009417	NA12275_1000G_LowCov_ERR009417	ERR009417 (4Gb, 2x)	2 - Calculate Metrics	25.5	766	30
ERR009417	NA12275_1000G_LowCov_ERR009417	ERR009417 (4Gb, 2x)	3 - Remove Duplicates	18.9	909	48
ERR009417	NA12275_1000G_LowCov_ERR009417	ERR009417 (4Gb, 2x)	4 - Indel Realign	55.4	3493	63
ERR009417	NA12275_1000G_LowCov_ERR009417	ERR009417 (4Gb, 2x)	5 - Base Quality Score Recalibration	39.6	10245	259
ERR009417	NA12275_1000G_LowCov_ERR009417	ERR009417 (4Gb, 2x)	6 - Variant Calling (Unified Genotyper)	40.4	889	22
ERR009417	NA12275_1000G_LowCov_ERR009417	ERR009417 (4Gb, 2x)	7 - Variant Calling (Haplotype Caller)	29.4	6737	229
ERR009404	NA12347_1000G_LowCov_ERR009404	ERR009404 (3Gb, 2x)	1 - Read Alignment	1.5	1293	856
ERR009404	NA12347_1000G_LowCov_ERR009404	ERR009404 (3Gb, 2x)	2 - Calculate Metrics	21.7	673	31
ERR009404	NA12347_1000G_LowCov_ERR009404	ERR009404 (3Gb, 2x)	3 - Remove Duplicates	17.3	777	45
ERR009404	NA12347_1000G_LowCov_ERR009404	ERR009404 (3Gb, 2x)	4 - Indel Realign	55	3409	62
ERR009404	NA12347_1000G_LowCov_ERR009404	ERR009404 (3Gb, 2x)	5 - Base Quality Score Recalibration	36.5	7925	217
ERR009404	NA12347_1000G_LowCov_ERR009404	ERR009404 (3Gb, 2x)	6 - Variant Calling (Unified Genotyper)	47.9	957	20
ERR009404	NA12347_1000G_LowCov_ERR009404	ERR009404 (3Gb, 2x)	7 - Variant Calling (Haplotype Caller)	33.1	5692	172
ERR000876	NA11894_1000G_LowCov_ERR000876	ERR000876 (619Mb, 1x)	1 - Read Alignment	1.3	511	393
ERR000876	NA11894_1000G_LowCov_ERR000876	ERR000876 (619Mb, 1x)	2 - Calculate Metrics	8.7	261	30
ERR000876	NA11894_1000G_LowCov_ERR000876	ERR000876 (619Mb, 1x)	3 - Remove Duplicates	13	208	16
ERR000876	NA11894_1000G_LowCov_ERR000876	ERR000876 (619Mb, 1x)	4 - Indel Realign	119.1	1548	13
ERR000876	NA11894_1000G_LowCov_ERR000876	ERR000876 (619Mb, 1x)	5 - Base Quality Score Recalibration	31.7	3269	103
ERR000876	NA11894_1000G_LowCov_ERR000876	ERR000876 (619Mb, 1x)	6 - Variant Calling (Unified Genotyper)	96.1	961	10
ERR000876	NA11894_1000G_LowCov_ERR000876	ERR000876 (619Mb, 1x)	7 - Variant Calling (Haplotype Caller)	53.5	2731	51
SRR622461	NA12878_1000G_LowCov_SRR622461	SRR622461 (16Gb, 6x)	1 - Read Alignment	1.4	5761	4215

SRR622461	NA12878_1000G_LowCov_SRR622461	SRR622461 (16Gb, 6x)	2 - Calculate Metrics	46.9	2533	54
SRR622461	NA12878_1000G_LowCov_SRR622461	SRR622461 (16Gb, 6x)	3 - Remove Duplicates	21.6	3009	139
SRR622461	NA12878_1000G_LowCov_SRR622461	SRR622461 (16Gb, 6x)	4 - Indel Realign	46.5	7389	159
SRR622461	NA12878_1000G_LowCov_SRR622461	SRR622461 (16Gb, 6x)	5 - Base Quality Score Recalibration	63.8	31017	486
SRR622461	NA12878_1000G_LowCov_SRR622461	SRR622461 (16Gb, 6x)	6 - Variant Calling (Unified Genotyper)	19.7	1043	53
SRR622461	NA12878_1000G_LowCov_SRR622461	SRR622461 (16Gb, 6x)	7 - Variant Calling (Haplotype Caller)	18.2	11342	623
SRR027518	NA11919_1000G_LowCov_SRR027518	SRR027518 (3Gb, 2x)	1 - Read Alignment	1.8	1158	656
SRR027518	NA11919_1000G_LowCov_SRR027518	SRR027518 (3Gb, 2x)	2 - Calculate Metrics	20.6	660	32
SRR027518	NA11919_1000G_LowCov_SRR027518	SRR027518 (3Gb, 2x)	3 - Remove Duplicates	18.7	765	41
SRR027518	NA11919_1000G_LowCov_SRR027518	SRR027518 (3Gb, 2x)	4 - Indel Realign	58.4	3210	55
SRR027518	NA11919_1000G_LowCov_SRR027518	SRR027518 (3Gb, 2x)	5 - Base Quality Score Recalibration	36	7813	217
SRR027518	NA11919_1000G_LowCov_SRR027518	SRR027518 (3Gb, 2x)	6 - Variant Calling (Unified Genotyper)	43.2	821	19
SRR027518	NA11919_1000G_LowCov_SRR027518	SRR027518 (3Gb, 2x)	7 - Variant Calling (Haplotype Caller)	36.1	5710	158
ERR257989	NA12761_1000G_LowCov_ERR257989	ERR257989 (18Gb, 6x)	1 - Read Alignment	2.2	5406	2508
ERR257989	NA12761_1000G_LowCov_ERR257989	ERR257989 (18Gb, 6x)	2 - Calculate Metrics	46	2484	54
ERR257989	NA12761_1000G_LowCov_ERR257989	ERR257989 (18Gb, 6x)	3 - Remove Duplicates	18.6	3514	189
ERR257989	NA12761_1000G_LowCov_ERR257989	ERR257989 (18Gb, 6x)	4 - Indel Realign	37.4	8500	227
ERR257989	NA12761_1000G_LowCov_ERR257989	ERR257989 (18Gb, 6x)	5 - Base Quality Score Recalibration	66.4	34536	520
ERR257989	NA12761_1000G_LowCov_ERR257989	ERR257989 (18Gb, 6x)	6 - Variant Calling (Unified Genotyper)	19.6	1234	63
ERR257989	NA12761_1000G_LowCov_ERR257989	ERR257989 (18Gb, 6x)	7 - Variant Calling (Haplotype Caller)	15	12190	813
ERR009412	NA12348_1000G_LowCov_ERR009412	ERR009412 (3Gb, 2x)	1 - Read Alignment	1.7	1358	784
ERR009412	NA12348_1000G_LowCov_ERR009412	ERR009412 (3Gb, 2x)	2 - Calculate Metrics	4.8	683	142
ERR009412	NA12348_1000G_LowCov_ERR009412	ERR009412 (3Gb, 2x)	3 - Remove Duplicates	18.4	829	45
ERR009412	NA12348_1000G_LowCov_ERR009412	ERR009412 (3Gb, 2x)	4 - Indel Realign	52.6	3524	67
ERR009412	NA12348_1000G_LowCov_ERR009412	ERR009412 (3Gb, 2x)	5 - Base Quality Score Recalibration	11.3	8014	708
ERR009412	NA12348_1000G_LowCov_ERR009412	ERR009412 (3Gb, 2x)	6 - Variant Calling (Unified Genotyper)	40.6	933	23
ERR009412	NA12348_1000G_LowCov_ERR009412	ERR009412 (3Gb, 2x)	7 - Variant Calling (Haplotype Caller)	31.3	5598	179
ERR257983	NA07037_1000G_LowCov_ERR257983	ERR257983 (17Gb, 6x)	1 - Read Alignment	2.2	4871	2247
ERR257983	NA07037_1000G_LowCov_ERR257983	ERR257983 (17Gb, 6x)	2 - Calculate Metrics	35.7	2283	64
ERR257983	NA07037_1000G_LowCov_ERR257983	ERR257983 (17Gb, 6x)	3 - Remove Duplicates	18.6	3170	170
ERR257983	NA07037_1000G_LowCov_ERR257983	ERR257983 (17Gb, 6x)	4 - Indel Realign	40.5	7940	196
ERR257983	NA07037_1000G_LowCov_ERR257983	ERR257983 (17Gb, 6x)	5 - Base Quality Score Recalibration	64.6	31416	486
ERR257983	NA07037_1000G_LowCov_ERR257983	ERR257983 (17Gb, 6x)	6 - Variant Calling (Unified Genotyper)	20.5	1187	58
ERR257983	NA07037_1000G_LowCov_ERR257983	ERR257983	7 - Variant Calling	15.4	11633	755

		wCov_ERR257983	(17Gb, 6x)	(Haplotype Caller)			
SRR768303	NA12814_1000G_LowCov_SRR768303	SRR768303 (13Gb, 5x)	1 - Read Alignment	2.1	4230	2013
SRR768303	NA12814_1000G_LowCov_SRR768303	SRR768303 (13Gb, 5x)	2 - Calculate Metrics	42.8	1883	44
SRR768303	NA12814_1000G_LowCov_SRR768303	SRR768303 (13Gb, 5x)	3 - Remove Duplicates	17.7	2667	151
SRR768303	NA12814_1000G_LowCov_SRR768303	SRR768303 (13Gb, 5x)	4 - Indel Realign	38.6	8109	210
SRR768303	NA12814_1000G_LowCov_SRR768303	SRR768303 (13Gb, 5x)	5 - Base Quality Score Recalibration	63	26031	413
SRR768303	NA12814_1000G_LowCov_SRR768303	SRR768303 (13Gb, 5x)	6 - Variant Calling (Unified Genotyper)	22	1123	51
SRR768303	NA12814_1000G_LowCov_SRR768303	SRR768303 (13Gb, 5x)	7 - Variant Calling (Haplotype Caller)	16.8	10580	629
ERR003045	NA12413_1000G_LowCov_ERR003045	ERR003045 (1Gb, 1x)	1 - Read Alignment	1.3	790	630
ERR003045	NA12413_1000G_LowCov_ERR003045	ERR003045 (1Gb, 1x)	2 - Calculate Metrics	11.2	416	37
ERR003045	NA12413_1000G_LowCov_ERR003045	ERR003045 (1Gb, 1x)	3 - Remove Duplicates	16	417	26
ERR003045	NA12413_1000G_LowCov_ERR003045	ERR003045 (1Gb, 1x)	4 - Indel Realign	66.8	2072	31
ERR003045	NA12413_1000G_LowCov_ERR003045	ERR003045 (1Gb, 1x)	5 - Base Quality Score Recalibration	29.2	4822	165
ERR003045	NA12413_1000G_LowCov_ERR003045	ERR003045 (1Gb, 1x)	6 - Variant Calling (Unified Genotyper)	72.8	946	13
ERR003045	NA12413_1000G_LowCov_ERR003045	ERR003045 (1Gb, 1x)	7 - Variant Calling (Haplotype Caller)	42.8	3938	92
ERR001695	NA12828_1000G_LowCov_ERR001695	ERR001695 (408Mb, 1x)	1 - Read Alignment	1.1	312	290
ERR001695	NA12828_1000G_LowCov_ERR001695	ERR001695 (408Mb, 1x)	2 - Calculate Metrics	14	182	13
ERR001695	NA12828_1000G_LowCov_ERR001695	ERR001695 (408Mb, 1x)	3 - Remove Duplicates	9.1	145	16
ERR001695	NA12828_1000G_LowCov_ERR001695	ERR001695 (408Mb, 1x)	4 - Indel Realign	142.4	1424	10
ERR001695	NA12828_1000G_LowCov_ERR001695	ERR001695 (408Mb, 1x)	5 - Base Quality Score Recalibration	61.1	2810	46
ERR001695	NA12828_1000G_LowCov_ERR001695	ERR001695 (408Mb, 1x)	6 - Variant Calling (Unified Genotyper)	125	1000	8
ERR001695	NA12828_1000G_LowCov_ERR001695	ERR001695 (408Mb, 1x)	7 - Variant Calling (Haplotype Caller)	58.6	2403	41
ERR257986	NA12716_1000G_LowCov_ERR257986	ERR257986 (17Gb, 6x)	1 - Read Alignment	2.1	5302	2493
ERR257986	NA12716_1000G_LowCov_ERR257986	ERR257986 (17Gb, 6x)	2 - Calculate Metrics	46.7	2476	53
ERR257986	NA12716_1000G_LowCov_ERR257986	ERR257986 (17Gb, 6x)	3 - Remove Duplicates	18.1	3384	187
ERR257986	NA12716_1000G_LowCov_ERR257986	ERR257986 (17Gb, 6x)	4 - Indel Realign	38.2	8397	220
ERR257986	NA12716_1000G_LowCov_ERR257986	ERR257986 (17Gb, 6x)	5 - Base Quality Score Recalibration	42.5	33399	785
ERR257986	NA12716_1000G_LowCov_ERR257986	ERR257986 (17Gb, 6x)	6 - Variant Calling (Unified Genotyper)	19.3	1195	62
ERR257986	NA12716_1000G_LowCov_ERR257986	ERR257986 (17Gb, 6x)	7 - Variant Calling (Haplotype Caller)	15	12237	814
pFDA 30x HG001	NA12878_30x_precisionFDA	pFDA 30x HG001 (113Gb, 39x)	1 - Read Alignment	2.3	33790	14960
pFDA 30x HG001	NA12878_30x_precisionFDA	pFDA 30x HG001 (113Gb, 39x)	2 - Calculate Metrics	47.1	14000	297
pFDA 30x HG001	NA12878_30x_precisionFDA	pFDA 30x HG001	3 - Remove Duplicates	18.6	19497	1051

		(113Gb, 39x)				
pFDA 30x HG001	NA12878_30x_precisionFDA	pFDA 30x HG001 (113Gb, 39x)	4 - Indel Realign	34.3	40233	1172
pFDA 30x HG001	NA12878_30x_precisionFDA	pFDA 30x HG001 (113Gb, 39x)	5 - Base Quality Score Recalibration	67.1	223313	3328
pFDA 30x HG001	NA12878_30x_precisionFDA	pFDA 30x HG001 (113Gb, 39x)	6 - Variant Calling (Unified Genotyper)	8.3	2494	301
pFDA 30x HG001	NA12878_30x_precisionFDA	pFDA 30x HG001 (113Gb, 39x)	7 - Variant Calling (Haplotype Caller)	8.3	35144	4219
SRR742200	NA12046_1000G_Exome_SRR742200	SRR742200 (5Gb, 97x)	1 - Read Alignment	3.6	1218	343
SRR742200	NA12046_1000G_Exome_SRR742200	SRR742200 (5Gb, 97x)	2 - Calculate Metrics	35.6	711	20
SRR742200	NA12046_1000G_Exome_SRR742200	SRR742200 (5Gb, 97x)	3 - Remove Duplicates	17.1	837	49
SRR742200	NA12046_1000G_Exome_SRR742200	SRR742200 (5Gb, 97x)	4 - Indel Realign	53.9	2588	48
SRR742200	NA12046_1000G_Exome_SRR742200	SRR742200 (5Gb, 97x)	5 - Base Quality Score Recalibration	49.8	10502	211
SRR742200	NA12046_1000G_Exome_SRR742200	SRR742200 (5Gb, 97x)	6 - Variant Calling (Unified Genotyper)	5.2	474	91
SRR742200	NA12046_1000G_Exome_SRR742200	SRR742200 (5Gb, 97x)	7 - Variant Calling (Haplotype Caller)	63.3	2343	37
ERR162833	NA12273_1000G_LowCov_ERR162833	ERR162833 (4Gb, 2x)	1 - Read Alignment	1.9	1208	638
ERR162833	NA12273_1000G_LowCov_ERR162833	ERR162833 (4Gb, 2x)	2 - Calculate Metrics	25.3	734	29
ERR162833	NA12273_1000G_LowCov_ERR162833	ERR162833 (4Gb, 2x)	3 - Remove Duplicates	18.9	850	45
ERR162833	NA12273_1000G_LowCov_ERR162833	ERR162833 (4Gb, 2x)	4 - Indel Realign	58.5	3571	61
ERR162833	NA12273_1000G_LowCov_ERR162833	ERR162833 (4Gb, 2x)	5 - Base Quality Score Recalibration	41.7	9718	233
ERR162833	NA12273_1000G_LowCov_ERR162833	ERR162833 (4Gb, 2x)	6 - Variant Calling (Unified Genotyper)	39.4	866	22
ERR162833	NA12273_1000G_LowCov_ERR162833	ERR162833 (4Gb, 2x)	7 - Variant Calling (Haplotype Caller)	27.8	6468	233
ERR162824	NA12749_1000G_LowCov_ERR162824	ERR162824 (4Gb, 2x)	1 - Read Alignment	1.8	1228	665
ERR162824	NA12749_1000G_LowCov_ERR162824	ERR162824 (4Gb, 2x)	2 - Calculate Metrics	18.1	725	40
ERR162824	NA12749_1000G_LowCov_ERR162824	ERR162824 (4Gb, 2x)	3 - Remove Duplicates	17.6	844	48
ERR162824	NA12749_1000G_LowCov_ERR162824	ERR162824 (4Gb, 2x)	4 - Indel Realign	57.9	3647	63
ERR162824	NA12749_1000G_LowCov_ERR162824	ERR162824 (4Gb, 2x)	5 - Base Quality Score Recalibration	37.9	9657	255
ERR162824	NA12749_1000G_LowCov_ERR162824	ERR162824 (4Gb, 2x)	6 - Variant Calling (Unified Genotyper)	43.8	963	22
ERR162824	NA12749_1000G_LowCov_ERR162824	ERR162824 (4Gb, 2x)	7 - Variant Calling (Haplotype Caller)	30.3	6524	215
SRR024100	NA11920_1000G_LowCov_SRR024100	SRR024100 (2Gb, 1x)	1 - Read Alignment	1.7	742	440
SRR024100	NA11920_1000G_LowCov_SRR024100	SRR024100 (2Gb, 1x)	2 - Calculate Metrics	14.4	417	29
SRR024100	NA11920_1000G_LowCov_SRR024100	SRR024100 (2Gb, 1x)	3 - Remove Duplicates	15.8	441	28
SRR024100	NA11920_1000G_LowCov	SRR024100	4 - Indel Realign	65.2	2414	37

	wCov_SRR024100	(2Gb, 1x)				
SRR024100	NA11920_1000G_LowCov_SRR024100	SRR024100 (2Gb, 1x)	5 - Base Quality Score Recalibration	30.1	5180	172
SRR024100	NA11920_1000G_LowCov_SRR024100	SRR024100 (2Gb, 1x)	6 - Variant Calling (Unified Genotyper)	73.3	953	13
SRR024100	NA11920_1000G_LowCov_SRR024100	SRR024100 (2Gb, 1x)	7 - Variant Calling (Haplotype Caller)	38.3	4441	116
SRR026644	NA07051_1000G_LowCov_SRR026644	SRR026644 (2Gb, 2x)	1 - Read Alignment	1.8	996	560
SRR026644	NA07051_1000G_LowCov_SRR026644	SRR026644 (2Gb, 2x)	2 - Calculate Metrics	17.6	546	31
SRR026644	NA07051_1000G_LowCov_SRR026644	SRR026644 (2Gb, 2x)	3 - Remove Duplicates	16.8	621	37
SRR026644	NA07051_1000G_LowCov_SRR026644	SRR026644 (2Gb, 2x)	4 - Indel Realign	57.2	2859	50
SRR026644	NA07051_1000G_LowCov_SRR026644	SRR026644 (2Gb, 2x)	5 - Base Quality Score Recalibration	31.8	6636	209
SRR026644	NA07051_1000G_LowCov_SRR026644	SRR026644 (2Gb, 2x)	6 - Variant Calling (Unified Genotyper)	55.8	949	17
SRR026644	NA07051_1000G_LowCov_SRR026644	SRR026644 (2Gb, 2x)	7 - Variant Calling (Haplotype Caller)	32.6	5052	155
ERR000902	NA11931_1000G_LowCov_ERR000902	ERR000902 (553Mb, 1x)	1 - Read Alignment	1.3	464	350
ERR000902	NA11931_1000G_LowCov_ERR000902	ERR000902 (553Mb, 1x)	2 - Calculate Metrics	13.6	245	18
ERR000902	NA11931_1000G_LowCov_ERR000902	ERR000902 (553Mb, 1x)	3 - Remove Duplicates	13.7	192	14
ERR000902	NA11931_1000G_LowCov_ERR000902	ERR000902 (553Mb, 1x)	4 - Indel Realign	115.1	1496	13
ERR000902	NA11931_1000G_LowCov_ERR000902	ERR000902 (553Mb, 1x)	5 - Base Quality Score Recalibration	41.8	3092	74
ERR000902	NA11931_1000G_LowCov_ERR000902	ERR000902 (553Mb, 1x)	6 - Variant Calling (Unified Genotyper)	82.5	907	11
ERR000902	NA11931_1000G_LowCov_ERR000902	ERR000902 (553Mb, 1x)	7 - Variant Calling (Haplotype Caller)	53.1	2441	46
SRR393992	NA12156_1000G_LowCov_SRR393992	SRR393992 (12Gb, 4x)	1 - Read Alignment	2.1	3536	1654
SRR393992	NA12156_1000G_LowCov_SRR393992	SRR393992 (12Gb, 4x)	2 - Calculate Metrics	32	1761	55
SRR393992	NA12156_1000G_LowCov_SRR393992	SRR393992 (12Gb, 4x)	3 - Remove Duplicates	18.7	2391	128
SRR393992	NA12156_1000G_LowCov_SRR393992	SRR393992 (12Gb, 4x)	4 - Indel Realign	45.1	6897	153
SRR393992	NA12156_1000G_LowCov_SRR393992	SRR393992 (12Gb, 4x)	5 - Base Quality Score Recalibration	38.8	23249	599
SRR393992	NA12156_1000G_LowCov_SRR393992	SRR393992 (12Gb, 4x)	6 - Variant Calling (Unified Genotyper)	24.8	1118	45
SRR393992	NA12156_1000G_LowCov_SRR393992	SRR393992 (12Gb, 4x)	7 - Variant Calling (Haplotype Caller)	21	9993	476
ERR257990	NA12763_1000G_LowCov_ERR257990	ERR257990 (15Gb, 5x)	1 - Read Alignment	2	4570	2315
ERR257990	NA12763_1000G_LowCov_ERR257990	ERR257990 (15Gb, 5x)	2 - Calculate Metrics	46.6	2191	47
ERR257990	NA12763_1000G_LowCov_ERR257990	ERR257990 (15Gb, 5x)	3 - Remove Duplicates	18.6	2995	161
ERR257990	NA12763_1000G_LowCov_ERR257990	ERR257990 (15Gb, 5x)	4 - Indel Realign	39.2	7992	204
ERR257990	NA12763_1000G_LowCov_ERR257990	ERR257990 (15Gb, 5x)	5 - Base Quality Score Recalibration	30.4	29455	970
ERR257990	NA12763_1000G_LowCov_ERR257990	ERR257990 (15Gb, 5x)	6 - Variant Calling (Unified Genotyper)	20.9	1148	55
ERR257990	NA12763_1000G_LowCov_ERR257990	ERR257990 (15Gb, 5x)	7 - Variant Calling (Haplotype Caller)	16	11419	712
SRR211278	NA12249_1000G_LowCov_SRR211278	SRR211278 (10Gb, 4x)	1 - Read Alignment	1.9	3046	1577
SRR211278	NA12249_1000G_LowCov_SRR211278	SRR211278 (10Gb, 4x)	2 - Calculate Metrics	8.3	1549	186

SRR211278	NA12249_1000G_LowCov_SRR211278	SRR211278 (10Gb, 4x)	3 - Remove Duplicates	19.3	2030	105
SRR211278	NA12249_1000G_LowCov_SRR211278	SRR211278 (10Gb, 4x)	4 - Indel Realign	44.5	6232	140
SRR211278	NA12249_1000G_LowCov_SRR211278	SRR211278 (10Gb, 4x)	5 - Base Quality Score Recalibration	15.5	19027	1224
SRR211278	NA12249_1000G_LowCov_SRR211278	SRR211278 (10Gb, 4x)	6 - Variant Calling (Unified Genotyper)	26.7	1096	41
SRR211278	NA12249_1000G_LowCov_SRR211278	SRR211278 (10Gb, 4x)	7 - Variant Calling (Haplotype Caller)	22.7	9438	415
SRR385775	NA12762_1000G_LowCov_SRR385775	SRR385775 (15Gb, 5x)	1 - Read Alignment	2.3	4512	1937
SRR385775	NA12762_1000G_LowCov_SRR385775	SRR385775 (15Gb, 5x)	2 - Calculate Metrics	46.5	2231	48
SRR385775	NA12762_1000G_LowCov_SRR385775	SRR385775 (15Gb, 5x)	3 - Remove Duplicates	20.2	3367	167
SRR385775	NA12762_1000G_LowCov_SRR385775	SRR385775 (15Gb, 5x)	4 - Indel Realign	39.5	7388	187
SRR385775	NA12762_1000G_LowCov_SRR385775	SRR385775 (15Gb, 5x)	5 - Base Quality Score Recalibration	56.9	30304	533
SRR385775	NA12762_1000G_LowCov_SRR385775	SRR385775 (15Gb, 5x)	6 - Variant Calling (Unified Genotyper)	20.7	1096	53
SRR385775	NA12762_1000G_LowCov_SRR385775	SRR385775 (15Gb, 5x)	7 - Variant Calling (Haplotype Caller)	20.2	11186	554
ERR050083	NA11881_1000G_LowCov_ERR050083	ERR050083 (13Gb, 5x)	1 - Read Alignment	1.9	3369	1802
ERR050083	NA11881_1000G_LowCov_ERR050083	ERR050083 (13Gb, 5x)	2 - Calculate Metrics	42.6	1919	45
ERR050083	NA11881_1000G_LowCov_ERR050083	ERR050083 (13Gb, 5x)	3 - Remove Duplicates	20.5	2459	120
ERR050083	NA11881_1000G_LowCov_ERR050083	ERR050083 (13Gb, 5x)	4 - Indel Realign	43.1	6943	161
ERR050083	NA11881_1000G_LowCov_ERR050083	ERR050083 (13Gb, 5x)	5 - Base Quality Score Recalibration	53.7	25325	472
ERR050083	NA11881_1000G_LowCov_ERR050083	ERR050083 (13Gb, 5x)	6 - Variant Calling (Unified Genotyper)	21.3	1023	48
ERR050083	NA11881_1000G_LowCov_ERR050083	ERR050083 (13Gb, 5x)	7 - Variant Calling (Haplotype Caller)	17.8	10842	609
ERR034545	NA11933_1000G_Exome_ERR034545	ERR034545 (11Gb, 222x)	1 - Read Alignment	3.9	2906	748
ERR034545	NA11933_1000G_Exome_ERR034545	ERR034545 (11Gb, 222x)	2 - Calculate Metrics	37.9	1593	42
ERR034545	NA11933_1000G_Exome_ERR034545	ERR034545 (11Gb, 222x)	3 - Remove Duplicates	18.7	2127	114
ERR034545	NA11933_1000G_Exome_ERR034545	ERR034545 (11Gb, 222x)	4 - Indel Realign	44.7	5235	117
ERR034545	NA11933_1000G_Exome_ERR034545	ERR034545 (11Gb, 222x)	5 - Base Quality Score Recalibration	62.3	21792	350
ERR034545	NA11933_1000G_Exome_ERR034545	ERR034545 (11Gb, 222x)	6 - Variant Calling (Unified Genotyper)	11.6	462	40
ERR034545	NA11933_1000G_Exome_ERR034545	ERR034545 (11Gb, 222x)	7 - Variant Calling (Haplotype Caller)	57.7	4270	74
SRR385769	NA12154_1000G_LowCov_SRR385769	SRR385769 (17Gb, 6x)	1 - Read Alignment	1.9	5044	2603
SRR385769	NA12154_1000G_LowCov_SRR385769	SRR385769 (17Gb, 6x)	2 - Calculate Metrics	18.5	2532	137
SRR385769	NA12154_1000G_LowCov_SRR385769	SRR385769 (17Gb, 6x)	3 - Remove Duplicates	20	3555	178
SRR385769	NA12154_1000G_LowCov_SRR385769	SRR385769 (17Gb, 6x)	4 - Indel Realign	39.2	8624	220

SRR385769	NA12154_1000G_LowCov_SRR385769	SRR385769 (17Gb, 6x)	5 - Base Quality Score Recalibration	15.4	32916	2138
SRR385769	NA12154_1000G_LowCov_SRR385769	SRR385769 (17Gb, 6x)	6 - Variant Calling (Unified Genotyper)	15.3	1089	71
SRR385769	NA12154_1000G_LowCov_SRR385769	SRR385769 (17Gb, 6x)	7 - Variant Calling (Haplotype Caller)	19.9	11489	576
ERR162832	NA12144_1000G_LowCov_ERR162832	ERR162832 (5Gb, 2x)	1 - Read Alignment	1.7	1474	877
ERR162832	NA12144_1000G_LowCov_ERR162832	ERR162832 (5Gb, 2x)	2 - Calculate Metrics	15.5	868	56
ERR162832	NA12144_1000G_LowCov_ERR162832	ERR162832 (5Gb, 2x)	3 - Remove Duplicates	19.2	1039	54
ERR162832	NA12144_1000G_LowCov_ERR162832	ERR162832 (5Gb, 2x)	4 - Indel Realign	55.5	4055	73
ERR162832	NA12144_1000G_LowCov_ERR162832	ERR162832 (5Gb, 2x)	5 - Base Quality Score Recalibration	21.2	11104	525
ERR162832	NA12144_1000G_LowCov_ERR162832	ERR162832 (5Gb, 2x)	6 - Variant Calling (Unified Genotyper)	33.1	893	27
ERR162832	NA12144_1000G_LowCov_ERR162832	ERR162832 (5Gb, 2x)	7 - Variant Calling (Haplotype Caller)	26.9	7226	269
SRR393988	NA11840_1000G_LowCov_SRR393988	SRR393988 (11Gb, 4x)	1 - Read Alignment	2.4	3305	1396
SRR393988	NA11840_1000G_LowCov_SRR393988	SRR393988 (11Gb, 4x)	2 - Calculate Metrics	40.3	1652	41
SRR393988	NA11840_1000G_LowCov_SRR393988	SRR393988 (11Gb, 4x)	3 - Remove Duplicates	20.2	2483	123
SRR393988	NA11840_1000G_LowCov_SRR393988	SRR393988 (11Gb, 4x)	4 - Indel Realign	44.8	6094	136
SRR393988	NA11840_1000G_LowCov_SRR393988	SRR393988 (11Gb, 4x)	5 - Base Quality Score Recalibration	58.7	22072	376
SRR393988	NA11840_1000G_LowCov_SRR393988	SRR393988 (11Gb, 4x)	6 - Variant Calling (Unified Genotyper)	26	1094	42
SRR393988	NA11840_1000G_LowCov_SRR393988	SRR393988 (11Gb, 4x)	7 - Variant Calling (Haplotype Caller)	22.6	9555	422
ERR015474	NA12778_1000G_LowCov_ERR015474	ERR015474 (3Gb, 2x)	1 - Read Alignment	1.3	1098	817
ERR015474	NA12778_1000G_LowCov_ERR015474	ERR015474 (3Gb, 2x)	2 - Calculate Metrics	27.4	602	22
ERR015474	NA12778_1000G_LowCov_ERR015474	ERR015474 (3Gb, 2x)	3 - Remove Duplicates	18.7	693	37
ERR015474	NA12778_1000G_LowCov_ERR015474	ERR015474 (3Gb, 2x)	4 - Indel Realign	62.9	3460	55
ERR015474	NA12778_1000G_LowCov_ERR015474	ERR015474 (3Gb, 2x)	5 - Base Quality Score Recalibration	61.4	8469	138
ERR015474	NA12778_1000G_LowCov_ERR015474	ERR015474 (3Gb, 2x)	6 - Variant Calling (Unified Genotyper)	50.6	962	19
ERR015474	NA12778_1000G_LowCov_ERR015474	ERR015474 (3Gb, 2x)	7 - Variant Calling (Haplotype Caller)	30.1	5660	188
ERR001509	NA07357_1000G_LowCov_ERR001509	ERR001509 (589Mb, 1x)	1 - Read Alignment	1.3	472	375
ERR001509	NA07357_1000G_LowCov_ERR001509	ERR001509 (589Mb, 1x)	2 - Calculate Metrics	8.4	245	29
ERR001509	NA07357_1000G_LowCov_ERR001509	ERR001509 (589Mb, 1x)	3 - Remove Duplicates	14.9	194	13
ERR001509	NA07357_1000G_LowCov_ERR001509	ERR001509 (589Mb, 1x)	4 - Indel Realign	125.4	1505	12
ERR001509	NA07357_1000G_LowCov_ERR001509	ERR001509 (589Mb, 1x)	5 - Base Quality Score Recalibration	27.3	3161	116
ERR001509	NA07357_1000G_LowCov_ERR001509	ERR001509 (589Mb, 1x)	6 - Variant Calling (Unified Genotyper)	84.8	933	11
ERR001509	NA07357_1000G_LowCov_ERR001509	ERR001509 (589Mb, 1x)	7 - Variant Calling (Haplotype Caller)	53.1	2812	53
SRR702068	NA12155_1000G_Exome_SRR702068	SRR702068 (7Gb, 130x)	1 - Read Alignment	3.8	1667	436
SRR702068	NA12155_1000G_Exome_SRR702068	SRR702068 (7Gb, 130x)	2 - Calculate Metrics	30.6	978	32
SRR702068	NA12155_1000G_Ex	SRR702068	3 - Remove Duplicates	18.9	1231	65

	ome_SRR702068	(7Gb, 130x)				
SRR702068	NA12155_1000G_Exome_SRR702068	SRR702068 (7Gb, 130x)	4 - Indel Realign	48.5	3247	67
SRR702068	NA12155_1000G_Exome_SRR702068	SRR702068 (7Gb, 130x)	5 - Base Quality Score Recalibration	53.8	13710	255
SRR702068	NA12155_1000G_Exome_SRR702068	SRR702068 (7Gb, 130x)	6 - Variant Calling (Unified Genotyper)	15.4	462	30
SRR702068	NA12155_1000G_Exome_SRR702068	SRR702068 (7Gb, 130x)	7 - Variant Calling (Haplotype Caller)	59.2	2840	48
ERR019488	NA12272_1000G_LowCov_ERR019488	ERR019488 (4Gb, 2x)	1 - Read Alignment	1.5	1398	934
ERR019488	NA12272_1000G_LowCov_ERR019488	ERR019488 (4Gb, 2x)	2 - Calculate Metrics	8	842	105
ERR019488	NA12272_1000G_LowCov_ERR019488	ERR019488 (4Gb, 2x)	3 - Remove Duplicates	19.1	896	47
ERR019488	NA12272_1000G_LowCov_ERR019488	ERR019488 (4Gb, 2x)	4 - Indel Realign	61.9	3531	57
ERR019488	NA12272_1000G_LowCov_ERR019488	ERR019488 (4Gb, 2x)	5 - Base Quality Score Recalibration	18.6	10765	579
ERR019488	NA12272_1000G_LowCov_ERR019488	ERR019488 (4Gb, 2x)	6 - Variant Calling (Unified Genotyper)	36	899	25
ERR019488	NA12272_1000G_LowCov_ERR019488	ERR019488 (4Gb, 2x)	7 - Variant Calling (Haplotype Caller)	29	6895	238
ERR239333	NA12046_1000G_LowCov_ERR239333	ERR239333 (22Gb, 8x)	1 - Read Alignment	2	6521	3194
ERR239333	NA12046_1000G_LowCov_ERR239333	ERR239333 (22Gb, 8x)	2 - Calculate Metrics	26.4	3084	117
ERR239333	NA12046_1000G_LowCov_ERR239333	ERR239333 (22Gb, 8x)	3 - Remove Duplicates	18.2	4345	239
ERR239333	NA12046_1000G_LowCov_ERR239333	ERR239333 (22Gb, 8x)	4 - Indel Realign	34.9	9377	269
ERR239333	NA12046_1000G_LowCov_ERR239333	ERR239333 (22Gb, 8x)	5 - Base Quality Score Recalibration	56.8	41659	734
ERR239333	NA12046_1000G_LowCov_ERR239333	ERR239333 (22Gb, 8x)	6 - Variant Calling (Unified Genotyper)	15.1	1150	76
ERR239333	NA12046_1000G_LowCov_ERR239333	ERR239333 (22Gb, 8x)	7 - Variant Calling (Haplotype Caller)	15.2	13400	879
SRR029683	NA12399_1000G_LowCov_SRR029683	SRR029683 (3Gb, 2x)	1 - Read Alignment	1.7	1014	587
SRR029683	NA12399_1000G_LowCov_SRR029683	SRR029683 (3Gb, 2x)	2 - Calculate Metrics	17.1	580	34
SRR029683	NA12399_1000G_LowCov_SRR029683	SRR029683 (3Gb, 2x)	3 - Remove Duplicates	18	665	37
SRR029683	NA12399_1000G_LowCov_SRR029683	SRR029683 (3Gb, 2x)	4 - Indel Realign	54	2862	53
SRR029683	NA12399_1000G_LowCov_SRR029683	SRR029683 (3Gb, 2x)	5 - Base Quality Score Recalibration	33.5	7011	209
SRR029683	NA12399_1000G_LowCov_SRR029683	SRR029683 (3Gb, 2x)	6 - Variant Calling (Unified Genotyper)	45.1	811	18
SRR029683	NA12399_1000G_LowCov_SRR029683	SRR029683 (3Gb, 2x)	7 - Variant Calling (Haplotype Caller)	35.8	5441	152
SRR098401	NA12878_1000G_Exome_SRR098401	SRR098401 (14Gb, 285x)	1 - Read Alignment	2.6	4681	1824
SRR098401	NA12878_1000G_Exome_SRR098401	SRR098401 (14Gb, 285x)	2 - Calculate Metrics	39.9	2515	63
SRR098401	NA12878_1000G_Exome_SRR098401	SRR098401 (14Gb, 285x)	3 - Remove Duplicates	19.6	3441	176
SRR098401	NA12878_1000G_Exome_SRR098401	SRR098401 (14Gb, 285x)	4 - Indel Realign	39.6	8085	204
SRR098401	NA12878_1000G_Exome_SRR098401	SRR098401 (14Gb, 285x)	5 - Base Quality Score Recalibration	61.3	28378	463

SRR098401	NA12878_1000G_Exome_SRR098401	SRR098401 (14Gb, 285x)	6 - Variant Calling (Unified Genotyper)	11.2	427	38
SRR098401	NA12878_1000G_Exome_SRR098401	SRR098401 (14Gb, 285x)	7 - Variant Calling (Haplotype Caller)	49.5	3367	68
SRR385759	NA11992_1000G_LowCov_SRR385759	SRR385759 (15Gb, 5x)	1 - Read Alignment	2.2	4417	2028
SRR385759	NA11992_1000G_LowCov_SRR385759	SRR385759 (15Gb, 5x)	2 - Calculate Metrics	44.7	2146	48
SRR385759	NA11992_1000G_LowCov_SRR385759	SRR385759 (15Gb, 5x)	3 - Remove Duplicates	19.6	3119	159
SRR385759	NA11992_1000G_LowCov_SRR385759	SRR385759 (15Gb, 5x)	4 - Indel Realign	40.9	7485	183
SRR385759	NA11992_1000G_LowCov_SRR385759	SRR385759 (15Gb, 5x)	5 - Base Quality Score Recalibration	61.8	29175	472
SRR385759	NA11992_1000G_LowCov_SRR385759	SRR385759 (15Gb, 5x)	6 - Variant Calling (Unified Genotyper)	20	1058	53
SRR385759	NA11992_1000G_LowCov_SRR385759	SRR385759 (15Gb, 5x)	7 - Variant Calling (Haplotype Caller)	19.2	10873	565
SRR385757	NA12004_1000G_LowCov_SRR385757	SRR385757 (12Gb, 4x)	1 - Read Alignment	2.2	3475	1576
SRR385757	NA12004_1000G_LowCov_SRR385757	SRR385757 (12Gb, 4x)	2 - Calculate Metrics	42	1764	42
SRR385757	NA12004_1000G_LowCov_SRR385757	SRR385757 (12Gb, 4x)	3 - Remove Duplicates	20.4	2594	127
SRR385757	NA12004_1000G_LowCov_SRR385757	SRR385757 (12Gb, 4x)	4 - Indel Realign	42.7	6453	151
SRR385757	NA12004_1000G_LowCov_SRR385757	SRR385757 (12Gb, 4x)	5 - Base Quality Score Recalibration	53.1	21445	404
SRR385757	NA12004_1000G_LowCov_SRR385757	SRR385757 (12Gb, 4x)	6 - Variant Calling (Unified Genotyper)	25.5	1120	44
SRR385757	NA12004_1000G_LowCov_SRR385757	SRR385757 (12Gb, 4x)	7 - Variant Calling (Haplotype Caller)	22.6	10482	464
SRR385763	NA11832_1000G_LowCov_SRR385763	SRR385763 (10Gb, 4x)	1 - Read Alignment	2.4	3098	1291
SRR385763	NA11832_1000G_LowCov_SRR385763	SRR385763 (10Gb, 4x)	2 - Calculate Metrics	38.5	1538	40
SRR385763	NA11832_1000G_LowCov_SRR385763	SRR385763 (10Gb, 4x)	3 - Remove Duplicates	19.6	2236	114
SRR385763	NA11832_1000G_LowCov_SRR385763	SRR385763 (10Gb, 4x)	4 - Indel Realign	44.6	6071	136
SRR385763	NA11832_1000G_LowCov_SRR385763	SRR385763 (10Gb, 4x)	5 - Base Quality Score Recalibration	56.8	20564	362
SRR385763	NA11832_1000G_LowCov_SRR385763	SRR385763 (10Gb, 4x)	6 - Variant Calling (Unified Genotyper)	26	1064	41
SRR385763	NA11832_1000G_LowCov_SRR385763	SRR385763 (10Gb, 4x)	7 - Variant Calling (Haplotype Caller)	23.6	9596	406
ERR162816	NA07000_1000G_LowCov_ERR162816	ERR162816 (4Gb, 2x)	1 - Read Alignment	1.9	1187	621
ERR162816	NA07000_1000G_LowCov_ERR162816	ERR162816 (4Gb, 2x)	2 - Calculate Metrics	23.4	701	30
ERR162816	NA07000_1000G_LowCov_ERR162816	ERR162816 (4Gb, 2x)	3 - Remove Duplicates	19.4	835	43
ERR162816	NA07000_1000G_LowCov_ERR162816	ERR162816 (4Gb, 2x)	4 - Indel Realign	59.8	3531	59
ERR162816	NA07000_1000G_LowCov_ERR162816	ERR162816 (4Gb, 2x)	5 - Base Quality Score Recalibration	40.2	9198	229
ERR162816	NA07000_1000G_LowCov_ERR162816	ERR162816 (4Gb, 2x)	6 - Variant Calling (Unified Genotyper)	44.2	928	21
ERR162816	NA07000_1000G_LowCov_ERR162816	ERR162816 (4Gb, 2x)	7 - Variant Calling (Haplotype Caller)	29.7	6378	215
SRR035028	NA06989_1000G_LowCov_SRR035028	SRR035028 (3Gb, 2x)	1 - Read Alignment	1.4	970	676
SRR035028	NA06989_1000G_LowCov_SRR035028	SRR035028 (3Gb, 2x)	2 - Calculate Metrics	18.8	565	30
SRR035028	NA06989_1000G_Lo	SRR035028	3 - Remove Duplicates	18.6	632	34

	wCov_SRR035028	(3Gb, 2x)				
SRR035028	NA06989_1000G_LowCov_SRR035028	SRR035028 (3Gb, 2x)	4 - Indel Realign	61	2807	46
SRR035028	NA06989_1000G_LowCov_SRR035028	SRR035028 (3Gb, 2x)	5 - Base Quality Score Recalibration	36.3	7265	200
SRR035028	NA06989_1000G_LowCov_SRR035028	SRR035028 (3Gb, 2x)	6 - Variant Calling (Unified Genotyper)	57	969	17
SRR035028	NA06989_1000G_LowCov_SRR035028	SRR035028 (3Gb, 2x)	7 - Variant Calling (Haplotype Caller)	33.8	5435	161
ERR003052	NA12777_1000G_LowCov_ERR003052	ERR003052 (2Gb, 1x)	1 - Read Alignment	1.3	863	677
ERR003052	NA12777_1000G_LowCov_ERR003052	ERR003052 (2Gb, 1x)	2 - Calculate Metrics	14.1	438	31
ERR003052	NA12777_1000G_LowCov_ERR003052	ERR003052 (2Gb, 1x)	3 - Remove Duplicates	16.7	484	29
ERR003052	NA12777_1000G_LowCov_ERR003052	ERR003052 (2Gb, 1x)	4 - Indel Realign	75.4	2262	30
ERR003052	NA12777_1000G_LowCov_ERR003052	ERR003052 (2Gb, 1x)	5 - Base Quality Score Recalibration	38.1	5216	137
ERR003052	NA12777_1000G_LowCov_ERR003052	ERR003052 (2Gb, 1x)	6 - Variant Calling (Unified Genotyper)	57.9	926	16
ERR003052	NA12777_1000G_LowCov_ERR003052	ERR003052 (2Gb, 1x)	7 - Variant Calling (Haplotype Caller)	42.3	4149	98
SRR385756	NA12003_1000G_LowCov_SRR385756	SRR385756 (15Gb, 5x)	1 - Read Alignment	2.1	4431	2073
SRR385756	NA12003_1000G_LowCov_SRR385756	SRR385756 (15Gb, 5x)	2 - Calculate Metrics	45.2	2216	49
SRR385756	NA12003_1000G_LowCov_SRR385756	SRR385756 (15Gb, 5x)	3 - Remove Duplicates	20.4	3271	160
SRR385756	NA12003_1000G_LowCov_SRR385756	SRR385756 (15Gb, 5x)	4 - Indel Realign	41.9	7378	176
SRR385756	NA12003_1000G_LowCov_SRR385756	SRR385756 (15Gb, 5x)	5 - Base Quality Score Recalibration	63.4	29206	461
SRR385756	NA12003_1000G_LowCov_SRR385756	SRR385756 (15Gb, 5x)	6 - Variant Calling (Unified Genotyper)	21.5	1142	53
SRR385756	NA12003_1000G_LowCov_SRR385756	SRR385756 (15Gb, 5x)	7 - Variant Calling (Haplotype Caller)	20	11082	553
SRR032636	NA07048_1000G_LowCov_SRR032636	SRR032636 (3Gb, 2x)	1 - Read Alignment	1.7	1002	588
SRR032636	NA07048_1000G_LowCov_SRR032636	SRR032636 (3Gb, 2x)	2 - Calculate Metrics	17.1	564	33
SRR032636	NA07048_1000G_LowCov_SRR032636	SRR032636 (3Gb, 2x)	3 - Remove Duplicates	17.9	681	38
SRR032636	NA07048_1000G_LowCov_SRR032636	SRR032636 (3Gb, 2x)	4 - Indel Realign	58.1	2787	48
SRR032636	NA07048_1000G_LowCov_SRR032636	SRR032636 (3Gb, 2x)	5 - Base Quality Score Recalibration	31.8	6744	212
SRR032636	NA07048_1000G_LowCov_SRR032636	SRR032636 (3Gb, 2x)	6 - Variant Calling (Unified Genotyper)	52.3	941	18
SRR032636	NA07048_1000G_LowCov_SRR032636	SRR032636 (3Gb, 2x)	7 - Variant Calling (Haplotype Caller)	35.4	5345	151
SRR035016	NA11843_1000G_LowCov_SRR035016	SRR035016 (3Gb, 2x)	1 - Read Alignment	1.8	1012	575
SRR035016	NA11843_1000G_LowCov_SRR035016	SRR035016 (3Gb, 2x)	2 - Calculate Metrics	15.3	581	38
SRR035016	NA11843_1000G_LowCov_SRR035016	SRR035016 (3Gb, 2x)	3 - Remove Duplicates	18	649	36
SRR035016	NA11843_1000G_LowCov_SRR035016	SRR035016 (3Gb, 2x)	4 - Indel Realign	67.2	3026	45
SRR035016	NA11843_1000G_LowCov_SRR035016	SRR035016 (3Gb, 2x)	5 - Base Quality Score Recalibration	37.1	7423	200
SRR035016	NA11843_1000G_LowCov_SRR035016	SRR035016 (3Gb, 2x)	6 - Variant Calling (Unified Genotyper)	55.9	1006	18
SRR035016	NA11843_1000G_LowCov_SRR035016	SRR035016 (3Gb, 2x)	7 - Variant Calling (Haplotype Caller)	33.2	5615	169
ERR233226	NA11930_1000G_LowCov_ERR233226	ERR233226 (24Gb, 9x)	1 - Read Alignment	2.1	7299	3403

ERR233226	NA11930_1000G_LowCov_ERR233226	ERR233226 (24Gb, 9x)	2 - Calculate Metrics	48.9	3469	71
ERR233226	NA11930_1000G_LowCov_ERR233226	ERR233226 (24Gb, 9x)	3 - Remove Duplicates	18.4	4845	264
ERR233226	NA11930_1000G_LowCov_ERR233226	ERR233226 (24Gb, 9x)	4 - Indel Realign	36.7	10104	275
ERR233226	NA11930_1000G_LowCov_ERR233226	ERR233226 (24Gb, 9x)	5 - Base Quality Score Recalibration	61	45536	747
ERR233226	NA11930_1000G_LowCov_ERR233226	ERR233226 (24Gb, 9x)	6 - Variant Calling (Unified Genotyper)	14.4	1194	83
ERR233226	NA11930_1000G_LowCov_ERR233226	ERR233226 (24Gb, 9x)	7 - Variant Calling (Haplotype Caller)	15	14068	941
SRR741366	NA12872_1000G_LowCov_SRR741366	SRR741366 (14Gb, 5x)	1 - Read Alignment	2.1	4289	2089
SRR741366	NA12872_1000G_LowCov_SRR741366	SRR741366 (14Gb, 5x)	2 - Calculate Metrics	45.3	1995	44
SRR741366	NA12872_1000G_LowCov_SRR741366	SRR741366 (14Gb, 5x)	3 - Remove Duplicates	19.1	2751	144
SRR741366	NA12872_1000G_LowCov_SRR741366	SRR741366 (14Gb, 5x)	4 - Indel Realign	40.7	7530	185
SRR741366	NA12872_1000G_LowCov_SRR741366	SRR741366 (14Gb, 5x)	5 - Base Quality Score Recalibration	65.6	27605	421
SRR741366	NA12872_1000G_LowCov_SRR741366	SRR741366 (14Gb, 5x)	6 - Variant Calling (Unified Genotyper)	22.2	1109	50
SRR741366	NA12872_1000G_LowCov_SRR741366	SRR741366 (14Gb, 5x)	7 - Variant Calling (Haplotype Caller)	17.5	10786	617
SRR393994	NA12489_1000G_LowCov_SRR393994	SRR393994 (11Gb, 4x)	1 - Read Alignment	2.3	3272	1422
SRR393994	NA12489_1000G_LowCov_SRR393994	SRR393994 (11Gb, 4x)	2 - Calculate Metrics	39.2	1645	42
SRR393994	NA12489_1000G_LowCov_SRR393994	SRR393994 (11Gb, 4x)	3 - Remove Duplicates	18.4	2283	124
SRR393994	NA12489_1000G_LowCov_SRR393994	SRR393994 (11Gb, 4x)	4 - Indel Realign	44.5	6538	147
SRR393994	NA12489_1000G_LowCov_SRR393994	SRR393994 (11Gb, 4x)	5 - Base Quality Score Recalibration	57.5	21789	379
SRR393994	NA12489_1000G_LowCov_SRR393994	SRR393994 (11Gb, 4x)	6 - Variant Calling (Unified Genotyper)	26.5	1113	42
SRR393994	NA12489_1000G_LowCov_SRR393994	SRR393994 (11Gb, 4x)	7 - Variant Calling (Haplotype Caller)	21.9	9700	442
SRR385751	NA06994_1000G_LowCov_SRR385751	SRR385751 (18Gb, 6x)	1 - Read Alignment	2.2	5183	2312
SRR385751	NA06994_1000G_LowCov_SRR385751	SRR385751 (18Gb, 6x)	2 - Calculate Metrics	44.2	2474	56
SRR385751	NA06994_1000G_LowCov_SRR385751	SRR385751 (18Gb, 6x)	3 - Remove Duplicates	20	3668	183
SRR385751	NA06994_1000G_LowCov_SRR385751	SRR385751 (18Gb, 6x)	4 - Indel Realign	40.2	8006	199
SRR385751	NA06994_1000G_LowCov_SRR385751	SRR385751 (18Gb, 6x)	5 - Base Quality Score Recalibration	62.3	33813	543
SRR385751	NA06994_1000G_LowCov_SRR385751	SRR385751 (18Gb, 6x)	6 - Variant Calling (Unified Genotyper)	17.8	1066	60
SRR385751	NA06994_1000G_LowCov_SRR385751	SRR385751 (18Gb, 6x)	7 - Variant Calling (Haplotype Caller)	19.3	11757	609
SRR741369	NA12812_1000G_LowCov_SRR741369	SRR741369 (15Gb, 6x)	1 - Read Alignment	2	4818	2374
SRR741369	NA12812_1000G_LowCov_SRR741369	SRR741369 (15Gb, 6x)	2 - Calculate Metrics	44.5	2180	49
SRR741369	NA12812_1000G_LowCov_SRR741369	SRR741369 (15Gb, 6x)	3 - Remove Duplicates	17.4	3052	175
SRR741369	NA12812_1000G_LowCov_SRR741369	SRR741369 (15Gb, 6x)	4 - Indel Realign	38.4	8335	217
SRR741369	NA12812_1000G_LowCov_SRR741369	SRR741369 (15Gb, 6x)	5 - Base Quality Score Recalibration	60	29716	495
SRR741369	NA12812_1000G_LowCov_SRR741369	SRR741369 (15Gb, 6x)	6 - Variant Calling (Unified Genotyper)	20.6	1197	58
SRR741369	NA12812_1000G_LowCov_SRR741369	SRR741369	7 - Variant Calling	15	11166	745

	wCov_SRR741369	(15Gb, 6x)	(Haplotype Caller)			
ERR015477	NA11892_1000G_LowCov_ERR015477	ERR015477 (4Gb, 2x)	1 - Read Alignment	1.7	1280	757
ERR015477	NA11892_1000G_LowCov_ERR015477	ERR015477 (4Gb, 2x)	2 - Calculate Metrics	23	714	31
ERR015477	NA11892_1000G_LowCov_ERR015477	ERR015477 (4Gb, 2x)	3 - Remove Duplicates	18.5	813	44
ERR015477	NA11892_1000G_LowCov_ERR015477	ERR015477 (4Gb, 2x)	4 - Indel Realign	55.6	3724	67
ERR015477	NA11892_1000G_LowCov_ERR015477	ERR015477 (4Gb, 2x)	5 - Base Quality Score Recalibration	41.8	9951	238
ERR015477	NA11892_1000G_LowCov_ERR015477	ERR015477 (4Gb, 2x)	6 - Variant Calling (Unified Genotyper)	42.9	900	21
ERR015477	NA11892_1000G_LowCov_ERR015477	ERR015477 (4Gb, 2x)	7 - Variant Calling (Haplotype Caller)	29.4	6635	226
ERR001502	NA10851_1000G_LowCov_ERR001502	ERR001502 (630Mb, 1x)	1 - Read Alignment	1.2	499	404
ERR001502	NA10851_1000G_LowCov_ERR001502	ERR001502 (630Mb, 1x)	2 - Calculate Metrics	8.6	257	30
ERR001502	NA10851_1000G_LowCov_ERR001502	ERR001502 (630Mb, 1x)	3 - Remove Duplicates	15.4	216	14
ERR001502	NA10851_1000G_LowCov_ERR001502	ERR001502 (630Mb, 1x)	4 - Indel Realign	118.7	1543	13
ERR001502	NA10851_1000G_LowCov_ERR001502	ERR001502 (630Mb, 1x)	5 - Base Quality Score Recalibration	27.5	3215	117
ERR001502	NA10851_1000G_LowCov_ERR001502	ERR001502 (630Mb, 1x)	6 - Variant Calling (Unified Genotyper)	84.6	931	11
ERR001502	NA10851_1000G_LowCov_ERR001502	ERR001502 (630Mb, 1x)	7 - Variant Calling (Haplotype Caller)	51.4	2827	55
SRR768305	NA12815_1000G_LowCov_SRR768305	SRR768305 (15Gb, 6x)	1 - Read Alignment	1.9	4803	2480
SRR768305	NA12815_1000G_LowCov_SRR768305	SRR768305 (15Gb, 6x)	2 - Calculate Metrics	44.3	2169	49
SRR768305	NA12815_1000G_LowCov_SRR768305	SRR768305 (15Gb, 6x)	3 - Remove Duplicates	17.6	3061	174
SRR768305	NA12815_1000G_LowCov_SRR768305	SRR768305 (15Gb, 6x)	4 - Indel Realign	38.5	8503	221
SRR768305	NA12815_1000G_LowCov_SRR768305	SRR768305 (15Gb, 6x)	5 - Base Quality Score Recalibration	62.5	29354	470
SRR768305	NA12815_1000G_LowCov_SRR768305	SRR768305 (15Gb, 6x)	6 - Variant Calling (Unified Genotyper)	20.4	1144	56
SRR768305	NA12815_1000G_LowCov_SRR768305	SRR768305 (15Gb, 6x)	7 - Variant Calling (Haplotype Caller)	15.4	11395	742
SRR027531	NA12045_1000G_LowCov_SRR027531	SRR027531 (3Gb, 2x)	1 - Read Alignment	1.5	985	640
SRR027531	NA12045_1000G_LowCov_SRR027531	SRR027531 (3Gb, 2x)	2 - Calculate Metrics	13.8	565	41
SRR027531	NA12045_1000G_LowCov_SRR027531	SRR027531 (3Gb, 2x)	3 - Remove Duplicates	17.8	642	36
SRR027531	NA12045_1000G_LowCov_SRR027531	SRR027531 (3Gb, 2x)	4 - Indel Realign	55.6	3000	54
SRR027531	NA12045_1000G_LowCov_SRR027531	SRR027531 (3Gb, 2x)	5 - Base Quality Score Recalibration	24.8	6899	278
SRR027531	NA12045_1000G_LowCov_SRR027531	SRR027531 (3Gb, 2x)	6 - Variant Calling (Unified Genotyper)	46	1013	22
SRR027531	NA12045_1000G_LowCov_SRR027531	SRR027531 (3Gb, 2x)	7 - Variant Calling (Haplotype Caller)	34.2	5578	163
ERR034544	NA11932_1000G_Exome_ERR034544	ERR034544 (11Gb, 207x)	1 - Read Alignment	3.9	2679	693
ERR034544	NA11932_1000G_Exome_ERR034544	ERR034544 (11Gb, 207x)	2 - Calculate Metrics	39.9	1475	37
ERR034544	NA11932_1000G_Exome_ERR034544	ERR034544 (11Gb, 207x)	3 - Remove Duplicates	19	1980	104

ERR034544	NA11932_1000G_Exome_ERR034544	ERR034544 (11Gb, 207x)	4 - Indel Realign	43.1	4699	109
ERR034544	NA11932_1000G_Exome_ERR034544	ERR034544 (11Gb, 207x)	5 - Base Quality Score Recalibration	69.2	20407	295
ERR034544	NA11932_1000G_Exome_ERR034544	ERR034544 (11Gb, 207x)	6 - Variant Calling (Unified Genotyper)	11.7	444	38
ERR034544	NA11932_1000G_Exome_ERR034544	ERR034544 (11Gb, 207x)	7 - Variant Calling (Haplotype Caller)	58.2	4131	71
ERR000162	NA12776_1000G_LowCov_ERR000162	ERR000162 (593Mb, 1x)	1 - Read Alignment	1.4	416	296
ERR000162	NA12776_1000G_LowCov_ERR000162	ERR000162 (593Mb, 1x)	2 - Calculate Metrics	12	228	19
ERR000162	NA12776_1000G_LowCov_ERR000162	ERR000162 (593Mb, 1x)	3 - Remove Duplicates	12.4	199	16
ERR000162	NA12776_1000G_LowCov_ERR000162	ERR000162 (593Mb, 1x)	4 - Indel Realign	123.5	1606	13
ERR000162	NA12776_1000G_LowCov_ERR000162	ERR000162 (593Mb, 1x)	5 - Base Quality Score Recalibration	32.8	3215	98
ERR000162	NA12776_1000G_LowCov_ERR000162	ERR000162 (593Mb, 1x)	6 - Variant Calling (Unified Genotyper)	110.9	998	9
ERR000162	NA12776_1000G_LowCov_ERR000162	ERR000162 (593Mb, 1x)	7 - Variant Calling (Haplotype Caller)	49.1	2601	53
SRR385777	NA12813_1000G_LowCov_SRR385777	SRR385777 (9Gb, 3x)	1 - Read Alignment	1.3	2708	2009
SRR385777	NA12813_1000G_LowCov_SRR385777	SRR385777 (9Gb, 3x)	2 - Calculate Metrics	43	1419	33
SRR385777	NA12813_1000G_LowCov_SRR385777	SRR385777 (9Gb, 3x)	3 - Remove Duplicates	18.9	1838	97
SRR385777	NA12813_1000G_LowCov_SRR385777	SRR385777 (9Gb, 3x)	4 - Indel Realign	49.2	5554	113
SRR385777	NA12813_1000G_LowCov_SRR385777	SRR385777 (9Gb, 3x)	5 - Base Quality Score Recalibration	51.3	18882	368
SRR385777	NA12813_1000G_LowCov_SRR385777	SRR385777 (9Gb, 3x)	6 - Variant Calling (Unified Genotyper)	29	1074	37
SRR385777	NA12813_1000G_LowCov_SRR385777	SRR385777 (9Gb, 3x)	7 - Variant Calling (Haplotype Caller)	23.8	9129	384
SRR400039	NA06985_1000G_LowCov_SRR400039	SRR400039 (24Gb, 8x)	1 - Read Alignment	2.1	7180	3354
SRR400039	NA06985_1000G_LowCov_SRR400039	SRR400039 (24Gb, 8x)	2 - Calculate Metrics	37.8	3481	92
SRR400039	NA06985_1000G_LowCov_SRR400039	SRR400039 (24Gb, 8x)	3 - Remove Duplicates	23.1	5711	247
SRR400039	NA06985_1000G_LowCov_SRR400039	SRR400039 (24Gb, 8x)	4 - Indel Realign	41.1	11430	278
SRR400039	NA06985_1000G_LowCov_SRR400039	SRR400039 (24Gb, 8x)	5 - Base Quality Score Recalibration	52.7	44927	853
SRR400039	NA06985_1000G_LowCov_SRR400039	SRR400039 (24Gb, 8x)	6 - Variant Calling (Unified Genotyper)	15.1	1174	78
SRR400039	NA06985_1000G_LowCov_SRR400039	SRR400039 (24Gb, 8x)	7 - Variant Calling (Haplotype Caller)	18.6	13395	722
SRR385773	NA12760_1000G_LowCov_SRR385773	SRR385773 (10Gb, 4x)	1 - Read Alignment	2	3361	1662
SRR385773	NA12760_1000G_LowCov_SRR385773	SRR385773 (10Gb, 4x)	2 - Calculate Metrics	41.7	1460	35
SRR385773	NA12760_1000G_LowCov_SRR385773	SRR385773 (10Gb, 4x)	3 - Remove Duplicates	18.6	1988	107
SRR385773	NA12760_1000G_LowCov_SRR385773	SRR385773 (10Gb, 4x)	4 - Indel Realign	46	6205	135
SRR385773	NA12760_1000G_LowCov_SRR385773	SRR385773 (10Gb, 4x)	5 - Base Quality Score Recalibration	54.7	20360	372
SRR385773	NA12760_1000G_LowCov_SRR385773	SRR385773 (10Gb, 4x)	6 - Variant Calling (Unified Genotyper)	27.1	1031	38
SRR385773	NA12760_1000G_Lo	SRR385773	7 - Variant Calling	22.8	9467	416

	wCov_SRR385773	(10Gb, 4x)	(Haplotype Caller)			
SRR385767	NA12005_1000G_LowCov_SRR385767	SRR385767 (12Gb, 4x)	1 - Read Alignment	1.9	3649	1908
SRR385767	NA12005_1000G_LowCov_SRR385767	SRR385767 (12Gb, 4x)	2 - Calculate Metrics	45.2	1942	43
SRR385767	NA12005_1000G_LowCov_SRR385767	SRR385767 (12Gb, 4x)	3 - Remove Duplicates	19.4	2578	133
SRR385767	NA12005_1000G_LowCov_SRR385767	SRR385767 (12Gb, 4x)	4 - Indel Realign	38.6	7100	184
SRR385767	NA12005_1000G_LowCov_SRR385767	SRR385767 (12Gb, 4x)	5 - Base Quality Score Recalibration	53.6	22124	413
SRR385767	NA12005_1000G_LowCov_SRR385767	SRR385767 (12Gb, 4x)	6 - Variant Calling (Unified Genotyper)	23.6	1085	46
SRR385767	NA12005_1000G_LowCov_SRR385767	SRR385767 (12Gb, 4x)	7 - Variant Calling (Haplotype Caller)	18.6	10474	562
ERR257987	NA12717_1000G_LowCov_ERR257987	ERR257987 (19Gb, 7x)	1 - Read Alignment	2	5612	2755
ERR257987	NA12717_1000G_LowCov_ERR257987	ERR257987 (19Gb, 7x)	2 - Calculate Metrics	47.3	2599	55
ERR257987	NA12717_1000G_LowCov_ERR257987	ERR257987 (19Gb, 7x)	3 - Remove Duplicates	18.3	3618	198
ERR257987	NA12717_1000G_LowCov_ERR257987	ERR257987 (19Gb, 7x)	4 - Indel Realign	37	8548	231
ERR257987	NA12717_1000G_LowCov_ERR257987	ERR257987 (19Gb, 7x)	5 - Base Quality Score Recalibration	64	35520	555
ERR257987	NA12717_1000G_LowCov_ERR257987	ERR257987 (19Gb, 7x)	6 - Variant Calling (Unified Genotyper)	16.7	1116	67
ERR257987	NA12717_1000G_LowCov_ERR257987	ERR257987 (19Gb, 7x)	7 - Variant Calling (Haplotype Caller)	14.8	12737	858
ERR009416	NA11893_1000G_LowCov_ERR009416	ERR009416 (3Gb, 2x)	1 - Read Alignment	1.6	1231	765
ERR009416	NA11893_1000G_LowCov_ERR009416	ERR009416 (3Gb, 2x)	2 - Calculate Metrics	22.1	685	31
ERR009416	NA11893_1000G_LowCov_ERR009416	ERR009416 (3Gb, 2x)	3 - Remove Duplicates	17	832	49
ERR009416	NA11893_1000G_LowCov_ERR009416	ERR009416 (3Gb, 2x)	4 - Indel Realign	52	3430	66
ERR009416	NA11893_1000G_LowCov_ERR009416	ERR009416 (3Gb, 2x)	5 - Base Quality Score Recalibration	37.2	8328	224
ERR009416	NA11893_1000G_LowCov_ERR009416	ERR009416 (3Gb, 2x)	6 - Variant Calling (Unified Genotyper)	40.1	843	21
ERR009416	NA11893_1000G_LowCov_ERR009416	ERR009416 (3Gb, 2x)	7 - Variant Calling (Haplotype Caller)	31.9	5925	186
SRR098359	NA12891_1000G_Exome_SRR098359	SRR098359 (14Gb, 284x)	1 - Read Alignment	2.5	4903	1987
SRR098359	NA12891_1000G_Exome_SRR098359	SRR098359 (14Gb, 284x)	2 - Calculate Metrics	43.2	2637	61
SRR098359	NA12891_1000G_Exome_SRR098359	SRR098359 (14Gb, 284x)	3 - Remove Duplicates	19.9	3580	180
SRR098359	NA12891_1000G_Exome_SRR098359	SRR098359 (14Gb, 284x)	4 - Indel Realign	40.9	10396	254
SRR098359	NA12891_1000G_Exome_SRR098359	SRR098359 (14Gb, 284x)	5 - Base Quality Score Recalibration	45.3	30284	668
SRR098359	NA12891_1000G_Exome_SRR098359	SRR098359 (14Gb, 284x)	6 - Variant Calling (Unified Genotyper)	4.5	409	91
SRR098359	NA12891_1000G_Exome_SRR098359	SRR098359 (14Gb, 284x)	7 - Variant Calling (Haplotype Caller)	33.1	3447	104
ERR000228	NA12043_1000G_LowCov_ERR000228	ERR000228 (495Mb, 1x)	1 - Read Alignment	1.4	404	291

ERR000228	NA12043_1000G_LowCov_ERR000228	ERR000228 (495Mb, 1x)	2 - Calculate Metrics	18.9	283	15
ERR000228	NA12043_1000G_LowCov_ERR000228	ERR000228 (495Mb, 1x)	3 - Remove Duplicates	11.7	176	15
ERR000228	NA12043_1000G_LowCov_ERR000228	ERR000228 (495Mb, 1x)	4 - Indel Realign	126.4	1517	12
ERR000228	NA12043_1000G_LowCov_ERR000228	ERR000228 (495Mb, 1x)	5 - Base Quality Score Recalibration	53.3	3094	58
ERR000228	NA12043_1000G_LowCov_ERR000228	ERR000228 (495Mb, 1x)	6 - Variant Calling (Unified Genotyper)	77.3	1005	13
ERR000228	NA12043_1000G_LowCov_ERR000228	ERR000228 (495Mb, 1x)	7 - Variant Calling (Haplotype Caller)	51.1	2503	49
SRR709972	NA06985_1000G_Exome_SRR709972	SRR709972 (6Gb, 110x)	1 - Read Alignment	3.4	1440	419
SRR709972	NA06985_1000G_Exome_SRR709972	SRR709972 (6Gb, 110x)	2 - Calculate Metrics	27.7	832	30
SRR709972	NA06985_1000G_Exome_SRR709972	SRR709972 (6Gb, 110x)	3 - Remove Duplicates	18.3	1025	56
SRR709972	NA06985_1000G_Exome_SRR709972	SRR709972 (6Gb, 110x)	4 - Indel Realign	50.2	2811	56
SRR709972	NA06985_1000G_Exome_SRR709972	SRR709972 (6Gb, 110x)	5 - Base Quality Score Recalibration	50.3	11759	234
SRR709972	NA06985_1000G_Exome_SRR709972	SRR709972 (6Gb, 110x)	6 - Variant Calling (Unified Genotyper)	17.1	461	27
SRR709972	NA06985_1000G_Exome_SRR709972	SRR709972 (6Gb, 110x)	7 - Variant Calling (Haplotype Caller)	60.7	2792	46
ERR019486	NA12718_1000G_LowCov_ERR019486	ERR019486 (5Gb, 2x)	1 - Read Alignment	1.9	1704	911
ERR019486	NA12718_1000G_LowCov_ERR019486	ERR019486 (5Gb, 2x)	2 - Calculate Metrics	25.4	837	33
ERR019486	NA12718_1000G_LowCov_ERR019486	ERR019486 (5Gb, 2x)	3 - Remove Duplicates	16.9	1029	61
ERR019486	NA12718_1000G_LowCov_ERR019486	ERR019486 (5Gb, 2x)	4 - Indel Realign	47.4	4128	87
ERR019486	NA12718_1000G_LowCov_ERR019486	ERR019486 (5Gb, 2x)	5 - Base Quality Score Recalibration	45.2	11423	253
ERR019486	NA12718_1000G_LowCov_ERR019486	ERR019486 (5Gb, 2x)	6 - Variant Calling (Unified Genotyper)	34	986	29
ERR019486	NA12718_1000G_LowCov_ERR019486	ERR019486 (5Gb, 2x)	7 - Variant Calling (Haplotype Caller)	26.5	6893	260
ERR003038	NA07056_1000G_LowCov_ERR003038	ERR003038 (1Gb, 1x)	1 - Read Alignment	1.3	797	594
ERR003038	NA07056_1000G_LowCov_ERR003038	ERR003038 (1Gb, 1x)	2 - Calculate Metrics	13.7	425	31
ERR003038	NA07056_1000G_LowCov_ERR003038	ERR003038 (1Gb, 1x)	3 - Remove Duplicates	16.2	422	26
ERR003038	NA07056_1000G_LowCov_ERR003038	ERR003038 (1Gb, 1x)	4 - Indel Realign	78.7	2124	27
ERR003038	NA07056_1000G_LowCov_ERR003038	ERR003038 (1Gb, 1x)	5 - Base Quality Score Recalibration	31	4867	157
ERR003038	NA07056_1000G_LowCov_ERR003038	ERR003038 (1Gb, 1x)	6 - Variant Calling (Unified Genotyper)	70	980	14
ERR003038	NA07056_1000G_LowCov_ERR003038	ERR003038 (1Gb, 1x)	7 - Variant Calling (Haplotype Caller)	46.1	3732	81
ERR000854	NA12287_1000G_LowCov_ERR000854	ERR000854 (503Mb, 1x)	1 - Read Alignment	1.3	414	318
ERR000854	NA12287_1000G_LowCov_ERR000854	ERR000854 (503Mb, 1x)	2 - Calculate Metrics	9.4	282	30
ERR000854	NA12287_1000G_LowCov_ERR000854	ERR000854 (503Mb, 1x)	3 - Remove Duplicates	14.8	178	12
ERR000854	NA12287_1000G_LowCov_ERR000854	ERR000854 (503Mb, 1x)	4 - Indel Realign	126.2	1514	12
ERR000854	NA12287_1000G_LowCov_ERR000854	ERR000854 (503Mb, 1x)	5 - Base Quality Score Recalibration	48	3022	63
ERR000854	NA12287_1000G_LowCov_ERR000854	ERR000854 (503Mb, 1x)	6 - Variant Calling (Unified Genotyper)	89.5	984	11
ERR000854	NA12287_1000G_Lo	ERR000854	7 - Variant Calling	53.7	2362	44

	wCov_ERR000854	(503Mb, 1x)	(Haplotype Caller)			
ERR091571	NA12878_1000G_LowCov_ERR091571	ERR091571 (41Gb, 14x)	1 - Read Alignment	2.4	11989	5089
ERR091571	NA12878_1000G_LowCov_ERR091571	ERR091571 (41Gb, 14x)	2 - Calculate Metrics	51.3	5389	105
ERR091571	NA12878_1000G_LowCov_ERR091571	ERR091571 (41Gb, 14x)	3 - Remove Duplicates	18.1	7497	415
ERR091571	NA12878_1000G_LowCov_ERR091571	ERR091571 (41Gb, 14x)	4 - Indel Realign	34.6	15599	451
ERR091571	NA12878_1000G_LowCov_ERR091571	ERR091571 (41Gb, 14x)	5 - Base Quality Score Recalibration	67.5	73285	1085
ERR091571	NA12878_1000G_LowCov_ERR091571	ERR091571 (41Gb, 14x)	6 - Variant Calling (Unified Genotyper)	11.6	1460	126
ERR091571	NA12878_1000G_LowCov_ERR091571	ERR091571 (41Gb, 14x)	7 - Variant Calling (Haplotype Caller)	12.4	18253	1469
SRR393991	NA12044_1000G_LowCov_SRR393991	SRR393991 (17Gb, 6x)	1 - Read Alignment	2	4900	2456
SRR393991	NA12044_1000G_LowCov_SRR393991	SRR393991 (17Gb, 6x)	2 - Calculate Metrics	33.2	2521	76
SRR393991	NA12044_1000G_LowCov_SRR393991	SRR393991 (17Gb, 6x)	3 - Remove Duplicates	20.2	3447	171
SRR393991	NA12044_1000G_LowCov_SRR393991	SRR393991 (17Gb, 6x)	4 - Indel Realign	38.8	8450	218
SRR393991	NA12044_1000G_LowCov_SRR393991	SRR393991 (17Gb, 6x)	5 - Base Quality Score Recalibration	53.4	32356	606
SRR393991	NA12044_1000G_LowCov_SRR393991	SRR393991 (17Gb, 6x)	6 - Variant Calling (Unified Genotyper)	18.2	1094	60
SRR393991	NA12044_1000G_LowCov_SRR393991	SRR393991 (17Gb, 6x)	7 - Variant Calling (Haplotype Caller)	16.8	12162	725
SRR035025	NA06984_1000G_LowCov_SRR035025	SRR035025 (4Gb, 2x)	1 - Read Alignment	1.9	1394	716
SRR035025	NA06984_1000G_LowCov_SRR035025	SRR035025 (4Gb, 2x)	2 - Calculate Metrics	21.5	752	35
SRR035025	NA06984_1000G_LowCov_SRR035025	SRR035025 (4Gb, 2x)	3 - Remove Duplicates	18.8	900	48
SRR035025	NA06984_1000G_LowCov_SRR035025	SRR035025 (4Gb, 2x)	4 - Indel Realign	56.8	3409	60
SRR035025	NA06984_1000G_LowCov_SRR035025	SRR035025 (4Gb, 2x)	5 - Base Quality Score Recalibration	33.7	9061	269
SRR035025	NA06984_1000G_LowCov_SRR035025	SRR035025 (4Gb, 2x)	6 - Variant Calling (Unified Genotyper)	44.5	980	22
SRR035025	NA06984_1000G_LowCov_SRR035025	SRR035025 (4Gb, 2x)	7 - Variant Calling (Haplotype Caller)	31.9	6443	202
SRR035020	NA12058_1000G_LowCov_SRR035020	SRR035020 (5Gb, 2x)	1 - Read Alignment	1.5	1551	1022
SRR035020	NA12058_1000G_LowCov_SRR035020	SRR035020 (5Gb, 2x)	2 - Calculate Metrics	15.2	849	56
SRR035020	NA12058_1000G_LowCov_SRR035020	SRR035020 (5Gb, 2x)	3 - Remove Duplicates	19	1007	53
SRR035020	NA12058_1000G_LowCov_SRR035020	SRR035020 (5Gb, 2x)	4 - Indel Realign	54.6	3985	73
SRR035020	NA12058_1000G_LowCov_SRR035020	SRR035020 (5Gb, 2x)	5 - Base Quality Score Recalibration	26	10325	397
SRR035020	NA12058_1000G_LowCov_SRR035020	SRR035020 (5Gb, 2x)	6 - Variant Calling (Unified Genotyper)	38.4	921	24
SRR035020	NA12058_1000G_LowCov_SRR035020	SRR035020 (5Gb, 2x)	7 - Variant Calling (Haplotype Caller)	32.8	7045	215
SRR385771	NA12414_1000G_LowCov_SRR385771	SRR385771 (16Gb, 6x)	1 - Read Alignment	2	5517	2698
SRR385771	NA12414_1000G_LowCov_SRR385771	SRR385771 (16Gb, 6x)	2 - Calculate Metrics	43.8	2322	53
SRR385771	NA12414_1000G_LowCov_SRR385771	SRR385771 (16Gb, 6x)	3 - Remove Duplicates	18.8	3436	183
SRR385771	NA12414_1000G_LowCov_SRR385771	SRR385771 (16Gb, 6x)	4 - Indel Realign	34.2	7801	228
SRR385771	NA12414_1000G_LowCov_SRR385771	SRR385771 (16Gb, 6x)	5 - Base Quality Score Recalibration	51.3	31369	612

SRR385771	NA12414_1000G_LowCov_SRR385771	SRR385771 (16Gb, 6x)	6 - Variant Calling (Unified Genotyper)	20.9	1172	56
SRR385771	NA12414_1000G_LowCov_SRR385771	SRR385771 (16Gb, 6x)	7 - Variant Calling (Haplotype Caller)	18.6	11274	605
pFDA 50x HG002	NA24385_50x_precisionFDA	pFDA 50x HG002 (148Gb, 51x)	1 - Read Alignment	2	42559	20778
pFDA 50x HG002	NA24385_50x_precisionFDA	pFDA 50x HG002 (148Gb, 51x)	2 - Calculate Metrics	50.4	16877	335
pFDA 50x HG002	NA24385_50x_precisionFDA	pFDA 50x HG002 (148Gb, 51x)	3 - Remove Duplicates	17	24580	1448
pFDA 50x HG002	NA24385_50x_precisionFDA	pFDA 50x HG002 (148Gb, 51x)	4 - Indel Realign	30.2	47375	1570
pFDA 50x HG002	NA24385_50x_precisionFDA	pFDA 50x HG002 (148Gb, 51x)	5 - Base Quality Score Recalibration	65.9	266654	4045
pFDA 50x HG002	NA24385_50x_precisionFDA	pFDA 50x HG002 (148Gb, 51x)	6 - Variant Calling (Unified Genotyper)	7.1	3097	437
pFDA 50x HG002	NA24385_50x_precisionFDA	pFDA 50x HG002 (148Gb, 51x)	7 - Variant Calling (Haplotype Caller)	6.7	42500	6349
SRR006107	NA11918_1000G_LowCov_SRR006107	SRR006107 (1Gb, 1x)	1 - Read Alignment	1	583	583
SRR006107	NA11918_1000G_LowCov_SRR006107	SRR006107 (1Gb, 1x)	2 - Calculate Metrics	12.3	392	32
SRR006107	NA11918_1000G_LowCov_SRR006107	SRR006107 (1Gb, 1x)	3 - Remove Duplicates	17.4	331	19
SRR006107	NA11918_1000G_LowCov_SRR006107	SRR006107 (1Gb, 1x)	4 - Indel Realign	102.8	1851	18
SRR006107	NA11918_1000G_LowCov_SRR006107	SRR006107 (1Gb, 1x)	5 - Base Quality Score Recalibration	30.4	4136	136
SRR006107	NA11918_1000G_LowCov_SRR006107	SRR006107 (1Gb, 1x)	6 - Variant Calling (Unified Genotyper)	78.8	946	12
SRR006107	NA11918_1000G_LowCov_SRR006107	SRR006107 (1Gb, 1x)	7 - Variant Calling (Haplotype Caller)	27.3	3769	138
SRR385760	NA12006_1000G_LowCov_SRR385760	SRR385760 (13Gb, 5x)	1 - Read Alignment	2.3	3873	1692
SRR385760	NA12006_1000G_LowCov_SRR385760	SRR385760 (13Gb, 5x)	2 - Calculate Metrics	43.2	1899	44
SRR385760	NA12006_1000G_LowCov_SRR385760	SRR385760 (13Gb, 5x)	3 - Remove Duplicates	19.5	2716	139
SRR385760	NA12006_1000G_LowCov_SRR385760	SRR385760 (13Gb, 5x)	4 - Indel Realign	43.7	7163	164
SRR385760	NA12006_1000G_LowCov_SRR385760	SRR385760 (13Gb, 5x)	5 - Base Quality Score Recalibration	60.3	25082	416
SRR385760	NA12006_1000G_LowCov_SRR385760	SRR385760 (13Gb, 5x)	6 - Variant Calling (Unified Genotyper)	22.3	1046	47
SRR385760	NA12006_1000G_LowCov_SRR385760	SRR385760 (13Gb, 5x)	7 - Variant Calling (Haplotype Caller)	21.8	10523	482
SRR075006	NA12340_1000G_LowCov_SRR075006	SRR075006 (11Gb, 4x)	1 - Read Alignment	1.9	3389	1817
SRR075006	NA12340_1000G_LowCov_SRR075006	SRR075006 (11Gb, 4x)	2 - Calculate Metrics	39.7	1707	43
SRR075006	NA12340_1000G_LowCov_SRR075006	SRR075006 (11Gb, 4x)	3 - Remove Duplicates	19.2	2252	117
SRR075006	NA12340_1000G_Lo	SRR075006	4 - Indel Realign	46.8	6509	139

	wCov_SRR075006	(11Gb, 4x)				
SRR075006	NA12340_1000G_LowCov_SRR075006	SRR075006 (11Gb, 4x)	5 - Base Quality Score Recalibration	55	22179	403
SRR075006	NA12340_1000G_LowCov_SRR075006	SRR075006 (11Gb, 4x)	6 - Variant Calling (Unified Genotyper)	26.4	1110	42
SRR075006	NA12340_1000G_LowCov_SRR075006	SRR075006 (11Gb, 4x)	7 - Variant Calling (Haplotype Caller)	20.2	9740	483
SRR029680	NA12400_1000G_LowCov_SRR029680	SRR029680 (3Gb, 2x)	1 - Read Alignment	1.7	1216	712
SRR029680	NA12400_1000G_LowCov_SRR029680	SRR029680 (3Gb, 2x)	2 - Calculate Metrics	21.1	655	31
SRR029680	NA12400_1000G_LowCov_SRR029680	SRR029680 (3Gb, 2x)	3 - Remove Duplicates	15.6	778	50
SRR029680	NA12400_1000G_LowCov_SRR029680	SRR029680 (3Gb, 2x)	4 - Indel Realign	57.6	3055	53
SRR029680	NA12400_1000G_LowCov_SRR029680	SRR029680 (3Gb, 2x)	5 - Base Quality Score Recalibration	35.7	7927	222
SRR029680	NA12400_1000G_LowCov_SRR029680	SRR029680 (3Gb, 2x)	6 - Variant Calling (Unified Genotyper)	42.9	944	22
SRR029680	NA12400_1000G_LowCov_SRR029680	SRR029680 (3Gb, 2x)	7 - Variant Calling (Haplotype Caller)	33.5	5727	171
SRR768310	NA12874_1000G_LowCov_SRR768310	SRR768310 (16Gb, 6x)	1 - Read Alignment	2	4989	2494
SRR768310	NA12874_1000G_LowCov_SRR768310	SRR768310 (16Gb, 6x)	2 - Calculate Metrics	46	2302	50
SRR768310	NA12874_1000G_LowCov_SRR768310	SRR768310 (16Gb, 6x)	3 - Remove Duplicates	17.3	3106	180
SRR768310	NA12874_1000G_LowCov_SRR768310	SRR768310 (16Gb, 6x)	4 - Indel Realign	35.9	8610	240
SRR768310	NA12874_1000G_LowCov_SRR768310	SRR768310 (16Gb, 6x)	5 - Base Quality Score Recalibration	35.7	30329	850
SRR768310	NA12874_1000G_LowCov_SRR768310	SRR768310 (16Gb, 6x)	6 - Variant Calling (Unified Genotyper)	20	1159	58
SRR768310	NA12874_1000G_LowCov_SRR768310	SRR768310 (16Gb, 6x)	7 - Variant Calling (Haplotype Caller)	15.6	11433	734
SRR768308	NA12873_1000G_LowCov_SRR768308	SRR768308 (14Gb, 5x)	1 - Read Alignment	2.1	4375	2069
SRR768308	NA12873_1000G_LowCov_SRR768308	SRR768308 (14Gb, 5x)	2 - Calculate Metrics	43.4	2041	47
SRR768308	NA12873_1000G_LowCov_SRR768308	SRR768308 (14Gb, 5x)	3 - Remove Duplicates	17.8	2870	161
SRR768308	NA12873_1000G_LowCov_SRR768308	SRR768308 (14Gb, 5x)	4 - Indel Realign	36.6	7989	218
SRR768308	NA12873_1000G_LowCov_SRR768308	SRR768308 (14Gb, 5x)	5 - Base Quality Score Recalibration	62.5	27828	445
SRR768308	NA12873_1000G_LowCov_SRR768308	SRR768308 (14Gb, 5x)	6 - Variant Calling (Unified Genotyper)	20.2	1149	57
SRR768308	NA12873_1000G_LowCov_SRR768308	SRR768308 (14Gb, 5x)	7 - Variant Calling (Haplotype Caller)	15.7	10841	692
SRR029852	NA07347_1000G_LowCov_SRR029852	SRR029852 (2Gb, 1x)	1 - Read Alignment	1.7	803	486
SRR029852	NA07347_1000G_LowCov_SRR029852	SRR029852 (2Gb, 1x)	2 - Calculate Metrics	14.3	443	31
SRR029852	NA07347_1000G_LowCov_SRR029852	SRR029852 (2Gb, 1x)	3 - Remove Duplicates	16.5	479	29
SRR029852	NA07347_1000G_LowCov_SRR029852	SRR029852 (2Gb, 1x)	4 - Indel Realign	70.9	2340	33
SRR029852	NA07347_1000G_LowCov_SRR029852	SRR029852 (2Gb, 1x)	5 - Base Quality Score Recalibration	29.3	5471	187
SRR029852	NA07347_1000G_LowCov_SRR029852	SRR029852 (2Gb, 1x)	6 - Variant Calling (Unified Genotyper)	63.1	883	14
SRR029852	NA07347_1000G_LowCov_SRR029852	SRR029852 (2Gb, 1x)	7 - Variant Calling (Haplotype Caller)	39.4	4569	116
SRR034932	NA12341_1000G_LowCov_SRR034932	SRR034932 (4Gb, 2x)	1 - Read Alignment	1.6	1508	916
SRR034932	NA12341_1000G_LowCov_SRR034932	SRR034932 (4Gb, 2x)	2 - Calculate Metrics	26.2	786	30

SRR034932	NA12341_1000G_LowCov_SRR034932	SRR034932 (4Gb, 2x)	3 - Remove Duplicates	21	924	44
SRR034932	NA12341_1000G_LowCov_SRR034932	SRR034932 (4Gb, 2x)	4 - Indel Realign	60.8	3649	60
SRR034932	NA12341_1000G_LowCov_SRR034932	SRR034932 (4Gb, 2x)	5 - Base Quality Score Recalibration	43.6	9471	217
SRR034932	NA12341_1000G_LowCov_SRR034932	SRR034932 (4Gb, 2x)	6 - Variant Calling (Unified Genotyper)	45.4	954	21
SRR034932	NA12341_1000G_LowCov_SRR034932	SRR034932 (4Gb, 2x)	7 - Variant Calling (Haplotype Caller)	31.9	6121	192
SRR385754	NA11995_1000G_LowCov_SRR385754	SRR385754 (10Gb, 4x)	1 - Read Alignment	2.3	2883	1274
SRR385754	NA11995_1000G_LowCov_SRR385754	SRR385754 (10Gb, 4x)	2 - Calculate Metrics	39.3	1455	37
SRR385754	NA11995_1000G_LowCov_SRR385754	SRR385754 (10Gb, 4x)	3 - Remove Duplicates	19.9	2146	108
SRR385754	NA11995_1000G_LowCov_SRR385754	SRR385754 (10Gb, 4x)	4 - Indel Realign	44.6	5757	129
SRR385754	NA11995_1000G_LowCov_SRR385754	SRR385754 (10Gb, 4x)	5 - Base Quality Score Recalibration	56.7	19453	343
SRR385754	NA11995_1000G_LowCov_SRR385754	SRR385754 (10Gb, 4x)	6 - Variant Calling (Unified Genotyper)	26.4	976	37
SRR385754	NA11995_1000G_LowCov_SRR385754	SRR385754 (10Gb, 4x)	7 - Variant Calling (Haplotype Caller)	24.1	9503	395
SRR400037	NA11831_1000G_LowCov_SRR400037	SRR400037 (9Gb, 3x)	1 - Read Alignment	2	2731	1390
SRR400037	NA11831_1000G_LowCov_SRR400037	SRR400037 (9Gb, 3x)	2 - Calculate Metrics	40.1	1444	36
SRR400037	NA11831_1000G_LowCov_SRR400037	SRR400037 (9Gb, 3x)	3 - Remove Duplicates	19.9	1929	97
SRR400037	NA11831_1000G_LowCov_SRR400037	SRR400037 (9Gb, 3x)	4 - Indel Realign	47	5829	124
SRR400037	NA11831_1000G_LowCov_SRR400037	SRR400037 (9Gb, 3x)	5 - Base Quality Score Recalibration	53	18827	355
SRR400037	NA11831_1000G_LowCov_SRR400037	SRR400037 (9Gb, 3x)	6 - Variant Calling (Unified Genotyper)	26.1	992	38
SRR400037	NA11831_1000G_LowCov_SRR400037	SRR400037 (9Gb, 3x)	7 - Variant Calling (Haplotype Caller)	25.2	9237	367
ERR013164	NA12383_1000G_LowCov_ERR013164	ERR013164 (5Gb, 2x)	1 - Read Alignment	2	1543	778
ERR013164	NA12383_1000G_LowCov_ERR013164	ERR013164 (5Gb, 2x)	2 - Calculate Metrics	25.6	793	31
ERR013164	NA12383_1000G_LowCov_ERR013164	ERR013164 (5Gb, 2x)	3 - Remove Duplicates	17.5	928	53
ERR013164	NA12383_1000G_LowCov_ERR013164	ERR013164 (5Gb, 2x)	4 - Indel Realign	52.6	4054	77
ERR013164	NA12383_1000G_LowCov_ERR013164	ERR013164 (5Gb, 2x)	5 - Base Quality Score Recalibration	36.8	11036	300
ERR013164	NA12383_1000G_LowCov_ERR013164	ERR013164 (5Gb, 2x)	6 - Variant Calling (Unified Genotyper)	36.3	908	25
ERR013164	NA12383_1000G_LowCov_ERR013164	ERR013164 (5Gb, 2x)	7 - Variant Calling (Haplotype Caller)	28.7	6626	231
SRR385753	NA11994_1000G_LowCov_SRR385753	SRR385753 (15Gb, 5x)	1 - Read Alignment	2.3	4470	1951
SRR385753	NA11994_1000G_LowCov_SRR385753	SRR385753 (15Gb, 5x)	2 - Calculate Metrics	45.4	2180	48
SRR385753	NA11994_1000G_LowCov_SRR385753	SRR385753 (15Gb, 5x)	3 - Remove Duplicates	19.9	3162	159
SRR385753	NA11994_1000G_LowCov_SRR385753	SRR385753 (15Gb, 5x)	4 - Indel Realign	40.3	7495	186
SRR385753	NA11994_1000G_LowCov_SRR385753	SRR385753 (15Gb, 5x)	5 - Base Quality Score Recalibration	63.7	29426	462
SRR385753	NA11994_1000G_LowCov_SRR385753	SRR385753 (15Gb, 5x)	6 - Variant Calling (Unified Genotyper)	22	1165	53
SRR385753	NA11994_1000G_LowCov_SRR385753	SRR385753 (15Gb, 5x)	7 - Variant Calling (Haplotype Caller)	20.2	11228	556

Supplementary Table 3.

Case	ID	Case name	Stage	Improvement	Runtime GATK	Runtime Sentieon
ERR009409	NA12775_1000G_LowCov_ERR009409	ERR009409 (3Gb, 2x)	0 - Overall (With alignment and HC only)	13.1	20911	1598
ERR009409	NA12775_1000G_LowCov_ERR009409	ERR009409 (3Gb, 2x)	0 - Overall (Without alignment and HC only)	37.1	19652	530
ERR009417	NA12275_1000G_LowCov_ERR009417	ERR009417 (4Gb, 2x)	0 - Overall (With alignment and HC only)	16.2	23618	1456
ERR009417	NA12275_1000G_LowCov_ERR009417	ERR009417 (4Gb, 2x)	0 - Overall (Without alignment and HC only)	35.2	22150	629
ERR009404	NA12347_1000G_LowCov_ERR009404	ERR009404 (3Gb, 2x)	0 - Overall (With alignment and HC only)	14.3	19769	1383
ERR009404	NA12347_1000G_LowCov_ERR009404	ERR009404 (3Gb, 2x)	0 - Overall (Without alignment and HC only)	35.1	18476	527
ERR000876	NA11894_1000G_LowCov_ERR000876	ERR000876 (619Mb, 1x)	0 - Overall (With alignment and HC only)	14.1	8528	606
ERR000876	NA11894_1000G_LowCov_ERR000876	ERR000876 (619Mb, 1x)	0 - Overall (Without alignment and HC only)	37.6	8017	213
SRR622461	NA12878_1000G_LowCov_SRR622461	SRR622461 (16Gb, 6x)	0 - Overall (With alignment and HC only)	10.8	61051	5676
SRR622461	NA12878_1000G_LowCov_SRR622461	SRR622461 (16Gb, 6x)	0 - Overall (Without alignment and HC only)	37.8	55290	1461
SRR027518	NA11919_1000G_LowCov_SRR027518	SRR027518 (3Gb, 2x)	0 - Overall (With alignment and HC only)	16.7	19316	1159
SRR027518	NA11919_1000G_LowCov_SRR027518	SRR027518 (3Gb, 2x)	0 - Overall (Without alignment and HC only)	36.1	18158	503
ERR257989	NA12761_1000G_LowCov_ERR257989	ERR257989 (18Gb, 6x)	0 - Overall (With alignment and HC only)	15.5	66630	4311
ERR257989	NA12761_1000G_LowCov_ERR257989	ERR257989 (18Gb, 6x)	0 - Overall (Without alignment and HC only)	34	61224	1803
ERR009412	NA12348_1000G_LowCov_ERR009412	ERR009412 (3Gb, 2x)	0 - Overall (With alignment and HC only)	10.4	20006	1925
ERR009412	NA12348_1000G_LowCov_ERR009412	ERR009412 (3Gb, 2x)	0 - Overall (Without alignment and HC only)	16.3	18648	1141
ERR257983	NA07037_1000G_LowCov_ERR257983	ERR257983 (17Gb, 6x)	0 - Overall (With alignment and HC only)	15.6	61313	3918
ERR257983	NA07037_1000G_LowCov_ERR257983	ERR257983 (17Gb, 6x)	0 - Overall (Without alignment and HC only)	33.8	56442	1671
SRR768303	NA12814_1000G_LowCov_SRR768303	SRR768303 (13Gb, 5x)	0 - Overall (With alignment and HC only)	15.5	53500	3460
SRR768303	NA12814_1000G_LowCov_SRR768303	SRR768303 (13Gb, 5x)	0 - Overall (Without alignment and HC only)	34	49270	1447
ERR003045	NA12413_1000G_LowCov_ERR003045	ERR003045 (1Gb, 1x)	0 - Overall (With alignment and HC	12.7	12455	981

			only)			
ERR003045	NA12413_1000G_LowCov_ERR003045	ERR003045 (1Gb, 1x)	0 - Overall (Without alignment and HC only)	33.2	11665	351
ERR001695	NA12828_1000G_LowCov_ERR001695	ERR001695 (408Mb, 1x)	0 - Overall (With alignment and HC only)	17.5	7276	416
ERR001695	NA12828_1000G_LowCov_ERR001695	ERR001695 (408Mb, 1x)	0 - Overall (Without alignment and HC only)	55.3	6964	126
ERR257986	NA12716_1000G_LowCov_ERR257986	ERR257986 (17Gb, 6x)	0 - Overall (With alignment and HC only)	14.3	65195	4552
ERR257986	NA12716_1000G_LowCov_ERR257986	ERR257986 (17Gb, 6x)	0 - Overall (Without alignment and HC only)	29.1	59893	2059
pFDA 30x HG001	NA12878_30x_precisionFDA	pFDA 30x HG001 (113Gb, 39x)	0 - Overall (With alignment and HC only)	14.6	365977	25027
pFDA 30x HG001	NA12878_30x_precisionFDA	pFDA 30x HG001 (113Gb, 39x)	0 - Overall (Without alignment and HC only)	33	332187	10067
SRR742200	NA12046_1000G_Exome_SRR742200	SRR742200 (5Gb, 97x)	0 - Overall (With alignment and HC only)	25.7	18199	708
SRR742200	NA12046_1000G_Exome_SRR742200	SRR742200 (5Gb, 97x)	0 - Overall (Without alignment and HC only)	46.5	16981	365
ERR162833	NA12273_1000G_LowCov_ERR162833	ERR162833 (4Gb, 2x)	0 - Overall (With alignment and HC only)	18.2	22549	1239
ERR162833	NA12273_1000G_LowCov_ERR162833	ERR162833 (4Gb, 2x)	0 - Overall (Without alignment and HC only)	35.5	21341	601
ERR162824	NA12749_1000G_LowCov_ERR162824	ERR162824 (4Gb, 2x)	0 - Overall (With alignment and HC only)	17.6	22625	1286
ERR162824	NA12749_1000G_LowCov_ERR162824	ERR162824 (4Gb, 2x)	0 - Overall (Without alignment and HC only)	34.5	21397	621
SRR024100	NA11920_1000G_LowCov_SRR024100	SRR024100 (2Gb, 1x)	0 - Overall (With alignment and HC only)	16.6	13635	822
SRR024100	NA11920_1000G_LowCov_SRR024100	SRR024100 (2Gb, 1x)	0 - Overall (Without alignment and HC only)	33.8	12893	382
SRR026644	NA07051_1000G_LowCov_SRR026644	SRR026644 (2Gb, 2x)	0 - Overall (With alignment and HC only)	16	16710	1042
SRR026644	NA07051_1000G_LowCov_SRR026644	SRR026644 (2Gb, 2x)	0 - Overall (Without alignment and HC only)	32.6	15714	482
ERR000902	NA11931_1000G_LowCov_ERR000902	ERR000902 (553Mb, 1x)	0 - Overall (With alignment and HC only)	15.4	7930	515
ERR000902	NA11931_1000G_LowCov_ERR000902	ERR000902 (553Mb, 1x)	0 - Overall (Without alignment and HC only)	45.2	7466	165
SRR393992	NA12156_1000G_LowCov_SRR393992	SRR393992 (12Gb, 4x)	0 - Overall (With alignment and HC only)	15.6	47827	3065
SRR393992	NA12156_1000G_LowCov_SRR393992	SRR393992 (12Gb, 4x)	0 - Overall (Without alignment and HC only)	31.4	44291	1411
ERR257990	NA12763_1000G_LowCov_ERR257990	ERR257990 (15Gb, 5x)	0 - Overall (With alignment and HC only)	13.3	58622	4409

ERR257990	NA12763_1000G_LowCov_ERR257990	ERR257990 (15Gb, 5x)	0 - Overall (Without alignment and HC only)	25.8	54052	2094
SRR211278	NA12249_1000G_LowCov_SRR211278	SRR211278 (10Gb, 4x)	0 - Overall (With alignment and HC only)	11.3	41322	3647
SRR211278	NA12249_1000G_LowCov_SRR211278	SRR211278 (10Gb, 4x)	0 - Overall (Without alignment and HC only)	18.5	38276	2070
SRR385775	NA12762_1000G_LowCov_SRR385775	SRR385775 (15Gb, 5x)	0 - Overall (With alignment and HC only)	17.2	58988	3426
SRR385775	NA12762_1000G_LowCov_SRR385775	SRR385775 (15Gb, 5x)	0 - Overall (Without alignment and HC only)	36.6	54476	1489
ERR050083	NA11881_1000G_LowCov_ERR050083	ERR050083 (13Gb, 5x)	0 - Overall (With alignment and HC only)	15.8	50857	3209
ERR050083	NA11881_1000G_LowCov_ERR050083	ERR050083 (13Gb, 5x)	0 - Overall (Without alignment and HC only)	33.8	47488	1407
ERR034545	NA11933_1000G_Exome_ERR034545	ERR034545 (11Gb, 222x)	0 - Overall (With alignment and HC only)	26.2	37923	1445
ERR034545	NA11933_1000G_Exome_ERR034545	ERR034545 (11Gb, 222x)	0 - Overall (Without alignment and HC only)	50.2	35017	697
SRR385769	NA12154_1000G_LowCov_SRR385769	SRR385769 (17Gb, 6x)	0 - Overall (With alignment and HC only)	11	64160	5852
SRR385769	NA12154_1000G_LowCov_SRR385769	SRR385769 (17Gb, 6x)	0 - Overall (Without alignment and HC only)	18.2	59116	3249
ERR162832	NA12144_1000G_LowCov_ERR162832	ERR162832 (5Gb, 2x)	0 - Overall (With alignment and HC only)	13.9	25766	1854
ERR162832	NA12144_1000G_LowCov_ERR162832	ERR162832 (5Gb, 2x)	0 - Overall (Without alignment and HC only)	24.9	24292	977
SRR393988	NA11840_1000G_LowCov_SRR393988	SRR393988 (11Gb, 4x)	0 - Overall (With alignment and HC only)	18.1	45161	2494
SRR393988	NA11840_1000G_LowCov_SRR393988	SRR393988 (11Gb, 4x)	0 - Overall (Without alignment and HC only)	38.1	41856	1098
ERR015474	NA12778_1000G_LowCov_ERR015474	ERR015474 (3Gb, 2x)	0 - Overall (With alignment and HC only)	15.9	19982	1257
ERR015474	NA12778_1000G_LowCov_ERR015474	ERR015474 (3Gb, 2x)	0 - Overall (Without alignment and HC only)	42.9	18884	440
ERR001509	NA07357_1000G_LowCov_ERR001509	ERR001509 (589Mb, 1x)	0 - Overall (With alignment and HC only)	14	8389	598
ERR001509	NA07357_1000G_LowCov_ERR001509	ERR001509 (589Mb, 1x)	0 - Overall (Without alignment and HC only)	35.5	7917	223
SRR702068	NA12155_1000G_Exome_SRR702068	SRR702068 (7Gb, 130x)	0 - Overall (With alignment and HC only)	26.2	23673	903
SRR702068	NA12155_1000G_Exome_SRR702068	SRR702068 (7Gb, 130x)	0 - Overall (Without alignment and HC only)	47.1	22006	467
ERR019488	NA12272_1000G_LowCov_ERR019488	ERR019488 (4Gb, 2x)	0 - Overall (With alignment and HC only)	12.4	24327	1960
ERR019488	NA12272_1000G_LowCov_ERR019488	ERR019488 (4Gb, 2x)	0 - Overall (Without alignment and HC	22.3	22929	1026

			only)			
ERR239333	NA12046_1000G_LowCov_ERR239333	ERR239333 (22Gb, 8x)	0 - Overall (With alignment and HC only)	14.4	78386	5432
ERR239333	NA12046_1000G_LowCov_ERR239333	ERR239333 (22Gb, 8x)	0 - Overall (Without alignment and HC only)	32.1	71865	2238
SRR029683	NA12399_1000G_LowCov_SRR029683	SRR029683 (3Gb, 2x)	0 - Overall (With alignment and HC only)	16.4	17573	1072
SRR029683	NA12399_1000G_LowCov_SRR029683	SRR029683 (3Gb, 2x)	0 - Overall (Without alignment and HC only)	34.1	16559	485
SRR098401	NA12878_1000G_Exome_SRR098401	SRR098401 (14Gb, 285x)	0 - Overall (With alignment and HC only)	18	50467	2798
SRR098401	NA12878_1000G_Exome_SRR098401	SRR098401 (14Gb, 285x)	0 - Overall (Without alignment and HC only)	47	45786	974
SRR385759	NA11992_1000G_LowCov_SRR385759	SRR385759 (15Gb, 5x)	0 - Overall (With alignment and HC only)	16.6	57215	3455
SRR385759	NA11992_1000G_LowCov_SRR385759	SRR385759 (15Gb, 5x)	0 - Overall (Without alignment and HC only)	37	52798	1427
SRR385757	NA12004_1000G_LowCov_SRR385757	SRR385757 (12Gb, 4x)	0 - Overall (With alignment and HC only)	16.7	46213	2764
SRR385757	NA12004_1000G_LowCov_SRR385757	SRR385757 (12Gb, 4x)	0 - Overall (Without alignment and HC only)	36	42738	1188
SRR385763	NA11832_1000G_LowCov_SRR385763	SRR385763 (10Gb, 4x)	0 - Overall (With alignment and HC only)	18.3	43103	2349
SRR385763	NA11832_1000G_LowCov_SRR385763	SRR385763 (10Gb, 4x)	0 - Overall (Without alignment and HC only)	37.8	40005	1058
ERR162816	NA07000_1000G_LowCov_ERR162816	ERR162816 (4Gb, 2x)	0 - Overall (With alignment and HC only)	18.2	21830	1197
ERR162816	NA07000_1000G_LowCov_ERR162816	ERR162816 (4Gb, 2x)	0 - Overall (Without alignment and HC only)	35.8	20643	576
SRR035028	NA06989_1000G_LowCov_SRR035028	SRR035028 (3Gb, 2x)	0 - Overall (With alignment and HC only)	15.4	17674	1147
SRR035028	NA06989_1000G_LowCov_SRR035028	SRR035028 (3Gb, 2x)	0 - Overall (Without alignment and HC only)	35.5	16704	471
ERR003052	NA12777_1000G_LowCov_ERR003052	ERR003052 (2Gb, 1x)	0 - Overall (With alignment and HC only)	13.4	13412	1002
ERR003052	NA12777_1000G_LowCov_ERR003052	ERR003052 (2Gb, 1x)	0 - Overall (Without alignment and HC only)	38.6	12549	325
SRR385756	NA12003_1000G_LowCov_SRR385756	SRR385756 (15Gb, 5x)	0 - Overall (With alignment and HC only)	16.6	57584	3472
SRR385756	NA12003_1000G_LowCov_SRR385756	SRR385756 (15Gb, 5x)	0 - Overall (Without alignment and HC only)	38	53153	1399
SRR032636	NA07048_1000G_LowCov_SRR032636	SRR032636 (3Gb, 2x)	0 - Overall (With alignment and HC only)	16	17123	1070
SRR032636	NA07048_1000G_LowCov_SRR032636	SRR032636 (3Gb, 2x)	0 - Overall (Without alignment and HC only)	33.4	16121	482

SRR035016	NA11843_1000G_LowCov_SRR035016	SRR035016 (3Gb, 2x)	0 - Overall (With alignment and HC only)	17.2	18306	1063
SRR035016	NA11843_1000G_LowCov_SRR035016	SRR035016 (3Gb, 2x)	0 - Overall (Without alignment and HC only)	35.4	17294	488
ERR233226	NA11930_1000G_LowCov_ERR233226	ERR233226 (24Gb, 9x)	0 - Overall (With alignment and HC only)	15	85321	5701
ERR233226	NA11930_1000G_LowCov_ERR233226	ERR233226 (24Gb, 9x)	0 - Overall (Without alignment and HC only)	34	78022	2298
SRR741366	NA12872_1000G_LowCov_SRR741366	SRR741366 (14Gb, 5x)	0 - Overall (With alignment and HC only)	15.7	54956	3500
SRR741366	NA12872_1000G_LowCov_SRR741366	SRR741366 (14Gb, 5x)	0 - Overall (Without alignment and HC only)	35.9	50667	1411
SRR393994	NA12489_1000G_LowCov_SRR393994	SRR393994 (11Gb, 4x)	0 - Overall (With alignment and HC only)	17.7	45227	2556
SRR393994	NA12489_1000G_LowCov_SRR393994	SRR393994 (11Gb, 4x)	0 - Overall (Without alignment and HC only)	37	41955	1134
SRR385751	NA06994_1000G_LowCov_SRR385751	SRR385751 (18Gb, 6x)	0 - Overall (With alignment and HC only)	16.6	64901	3902
SRR385751	NA06994_1000G_LowCov_SRR385751	SRR385751 (18Gb, 6x)	0 - Overall (Without alignment and HC only)	37.6	59718	1590
SRR741369	NA12812_1000G_LowCov_SRR741369	SRR741369 (15Gb, 6x)	0 - Overall (With alignment and HC only)	14.6	59267	4055
SRR741369	NA12812_1000G_LowCov_SRR741369	SRR741369 (15Gb, 6x)	0 - Overall (Without alignment and HC only)	32.4	54449	1681
ERR015477	NA11892_1000G_LowCov_ERR015477	ERR015477 (4Gb, 2x)	0 - Overall (With alignment and HC only)	17	23117	1363
ERR015477	NA11892_1000G_LowCov_ERR015477	ERR015477 (4Gb, 2x)	0 - Overall (Without alignment and HC only)	36	21837	606
ERR001502	NA10851_1000G_LowCov_ERR001502	ERR001502 (630Mb, 1x)	0 - Overall (With alignment and HC only)	13.5	8557	633
ERR001502	NA10851_1000G_LowCov_ERR001502	ERR001502 (630Mb, 1x)	0 - Overall (Without alignment and HC only)	35.2	8058	229
SRR768305	NA12815_1000G_LowCov_SRR768305	SRR768305 (15Gb, 6x)	0 - Overall (With alignment and HC only)	14.3	59285	4136
SRR768305	NA12815_1000G_LowCov_SRR768305	SRR768305 (15Gb, 6x)	0 - Overall (Without alignment and HC only)	32.9	54482	1656
SRR027531	NA12045_1000G_LowCov_SRR027531	SRR027531 (3Gb, 2x)	0 - Overall (With alignment and HC only)	14.6	17669	1212
SRR027531	NA12045_1000G_LowCov_SRR027531	SRR027531 (3Gb, 2x)	0 - Overall (Without alignment and HC only)	29.2	16684	572
ERR034544	NA11932_1000G_Exome_ERR034544	ERR034544 (11Gb, 207x)	0 - Overall (With alignment and HC only)	27	35371	1309
ERR034544	NA11932_1000G_Exome_ERR034544	ERR034544 (11Gb, 207x)	0 - Overall (Without alignment and HC only)	53.1	32692	616
ERR000162	NA12776_1000G_LowCov_ERR000162	ERR000162 (593Mb, 1x)	0 - Overall (With alignment and HC	16.7	8265	495

			only)			
ERR000162	NA12776_1000G_LowCov_ERR000162	ERR000162 (593Mb, 1x)	0 - Overall (Without alignment and HC only)	39.4	7849	199
SRR385777	NA12813_1000G_LowCov_SRR385777	SRR385777 (9Gb, 3x)	0 - Overall (With alignment and HC only)	13.2	39530	3004
SRR385777	NA12813_1000G_LowCov_SRR385777	SRR385777 (9Gb, 3x)	0 - Overall (Without alignment and HC only)	37	36822	995
SRR400039	NA06985_1000G_LowCov_SRR400039	SRR400039 (24Gb, 8x)	0 - Overall (With alignment and HC only)	15.5	86124	5546
SRR400039	NA06985_1000G_LowCov_SRR400039	SRR400039 (24Gb, 8x)	0 - Overall (Without alignment and HC only)	36	78944	2192
SRR385773	NA12760_1000G_LowCov_SRR385773	SRR385773 (10Gb, 4x)	0 - Overall (With alignment and HC only)	15.7	42841	2727
SRR385773	NA12760_1000G_LowCov_SRR385773	SRR385773 (10Gb, 4x)	0 - Overall (Without alignment and HC only)	37.1	39480	1065
SRR385767	NA12005_1000G_LowCov_SRR385767	SRR385767 (12Gb, 4x)	0 - Overall (With alignment and HC only)	14.8	47867	3243
SRR385767	NA12005_1000G_LowCov_SRR385767	SRR385767 (12Gb, 4x)	0 - Overall (Without alignment and HC only)	33.1	44218	1335
ERR257987	NA12717_1000G_LowCov_ERR257987	ERR257987 (19Gb, 7x)	0 - Overall (With alignment and HC only)	14.8	68634	4652
ERR257987	NA12717_1000G_LowCov_ERR257987	ERR257987 (19Gb, 7x)	0 - Overall (Without alignment and HC only)	33.2	63022	1897
ERR009416	NA11893_1000G_LowCov_ERR009416	ERR009416 (3Gb, 2x)	0 - Overall (With alignment and HC only)	15.5	20431	1321
ERR009416	NA11893_1000G_LowCov_ERR009416	ERR009416 (3Gb, 2x)	0 - Overall (Without alignment and HC only)	34.5	19200	556
SRR098359	NA12891_1000G_Exome_SRR098359	SRR098359 (14Gb, 284x)	0 - Overall (With alignment and HC only)	17	55247	3254
SRR098359	NA12891_1000G_Exome_SRR098359	SRR098359 (14Gb, 284x)	0 - Overall (Without alignment and HC only)	39.7	50344	1267
ERR000228	NA12043_1000G_LowCov_ERR000228	ERR000228 (495Mb, 1x)	0 - Overall (With alignment and HC only)	18.1	7977	440
ERR000228	NA12043_1000G_LowCov_ERR000228	ERR000228 (495Mb, 1x)	0 - Overall (Without alignment and HC only)	50.8	7573	149
SRR709972	NA06985_1000G_Exome_SRR709972	SRR709972 (6Gb, 110x)	0 - Overall (With alignment and HC only)	24.6	20659	841
SRR709972	NA06985_1000G_Exome_SRR709972	SRR709972 (6Gb, 110x)	0 - Overall (Without alignment and HC only)	45.5	19219	422
ERR019486	NA12718_1000G_LowCov_ERR019486	ERR019486 (5Gb, 2x)	0 - Overall (With alignment and HC only)	16.2	26014	1605
ERR019486	NA12718_1000G_LowCov_ERR019486	ERR019486 (5Gb, 2x)	0 - Overall (Without alignment and HC only)	35	24310	694
ERR003038	NA07056_1000G_LowCov_ERR003038	ERR003038 (1Gb, 1x)	0 - Overall (With alignment and HC only)	13.5	12367	916

ERR003038	NA07056_1000G_Lo wCov_ERR003038	ERR003038 (1Gb, 1x)	0 - Overall (Without alignment and HC only)	35.9	11570	322
ERR000854	NA12287_1000G_Lo wCov_ERR000854	ERR000854 (503Mb, 1x)	0 - Overall (With alignment and HC only)	16.2	7772	479
ERR000854	NA12287_1000G_Lo wCov_ERR000854	ERR000854 (503Mb, 1x)	0 - Overall (Without alignment and HC only)	45.7	7358	161
ERR091571	NA12878_1000G_Lo wCov_ERR091571	ERR091571 (41Gb, 14x)	0 - Overall (With alignment and HC only)	15.3	132012	8614
ERR091571	NA12878_1000G_Lo wCov_ERR091571	ERR091571 (41Gb, 14x)	0 - Overall (Without alignment and HC only)	34	120023	3525
SRR393991	NA12044_1000G_Lo wCov_SRR393991	SRR393991 (17Gb, 6x)	0 - Overall (With alignment and HC only)	15	63836	4252
SRR393991	NA12044_1000G_Lo wCov_SRR393991	SRR393991 (17Gb, 6x)	0 - Overall (Without alignment and HC only)	32.8	58936	1796
SRR035025	NA06984_1000G_Lo wCov_SRR035025	SRR035025 (4Gb, 2x)	0 - Overall (With alignment and HC only)	16.5	21959	1330
SRR035025	NA06984_1000G_Lo wCov_SRR035025	SRR035025 (4Gb, 2x)	0 - Overall (Without alignment and HC only)	33.5	20565	614
SRR035020	NA12058_1000G_Lo wCov_SRR035020	SRR035020 (5Gb, 2x)	0 - Overall (With alignment and HC only)	13.6	24762	1816
SRR035020	NA12058_1000G_Lo wCov_SRR035020	SRR035020 (5Gb, 2x)	0 - Overall (Without alignment and HC only)	29.2	23211	794
SRR385771	NA12414_1000G_Lo wCov_SRR385771	SRR385771 (16Gb, 6x)	0 - Overall (With alignment and HC only)	14.1	61719	4379
SRR385771	NA12414_1000G_Lo wCov_SRR385771	SRR385771 (16Gb, 6x)	0 - Overall (Without alignment and HC only)	33.4	56202	1681
pFDA 50x HG002	NA24385_50x_precisi onFDA	pFDA 50x HG002 (148Gb, 51x)	0 - Overall (With alignment and HC only)	12.8	440545	34525
pFDA 50x HG002	NA24385_50x_precisi onFDA	pFDA 50x HG002 (148Gb, 51x)	0 - Overall (Without alignment and HC only)	29	397986	13747
SRR006107	NA11918_1000G_Lo wCov_SRR006107	SRR006107 (1Gb, 1x)	0 - Overall (With alignment and HC only)	11.9	11062	926
SRR006107	NA11918_1000G_Lo wCov_SRR006107	SRR006107 (1Gb, 1x)	0 - Overall (Without alignment and HC only)	30.6	10479	343
SRR385760	NA12006_1000G_Lo wCov_SRR385760	SRR385760 (13Gb, 5x)	0 - Overall (With alignment and HC only)	17.5	51256	2937
SRR385760	NA12006_1000G_Lo wCov_SRR385760	SRR385760 (13Gb, 5x)	0 - Overall (Without alignment and HC only)	38.1	47383	1245
SRR075006	NA12340_1000G_Lo wCov_SRR075006	SRR075006 (11Gb, 4x)	0 - Overall (With alignment and HC only)	15.2	45776	3002
SRR075006	NA12340_1000G_Lo wCov_SRR075006	SRR075006 (11Gb, 4x)	0 - Overall (Without alignment and HC only)	35.8	42387	1185
SRR029680	NA12400_1000G_Lo wCov_SRR029680	SRR029680 (3Gb, 2x)	0 - Overall (With alignment and HC only)	15.6	19358	1239
SRR029680	NA12400_1000G_Lo wCov_SRR029680	SRR029680 (3Gb, 2x)	0 - Overall (Without alignment and HC	34.4	18142	527

			only)			
SRR768310	NA12874_1000G_LowCov_SRR768310	SRR768310 (16Gb, 6x)	0 - Overall (With alignment and HC only)	13.4	60769	4548
SRR768310	NA12874_1000G_LowCov_SRR768310	SRR768310 (16Gb, 6x)	0 - Overall (Without alignment and HC only)	27.2	55780	2054
SRR768308	NA12873_1000G_LowCov_SRR768308	SRR768308 (14Gb, 5x)	0 - Overall (With alignment and HC only)	15.4	55944	3632
SRR768308	NA12873_1000G_LowCov_SRR768308	SRR768308 (14Gb, 5x)	0 - Overall (Without alignment and HC only)	33	51569	1563
SRR029852	NA07347_1000G_LowCov_SRR029852	SRR029852 (2Gb, 1x)	0 - Overall (With alignment and HC only)	16	14105	882
SRR029852	NA07347_1000G_LowCov_SRR029852	SRR029852 (2Gb, 1x)	0 - Overall (Without alignment and HC only)	33.6	13302	396
SRR034932	NA12341_1000G_LowCov_SRR034932	SRR034932 (4Gb, 2x)	0 - Overall (With alignment and HC only)	15.4	22459	1459
SRR034932	NA12341_1000G_LowCov_SRR034932	SRR034932 (4Gb, 2x)	0 - Overall (Without alignment and HC only)	38.6	20951	543
SRR385754	NA11995_1000G_LowCov_SRR385754	SRR385754 (10Gb, 4x)	0 - Overall (With alignment and HC only)	18	41197	2286
SRR385754	NA11995_1000G_LowCov_SRR385754	SRR385754 (10Gb, 4x)	0 - Overall (Without alignment and HC only)	37.9	38314	1012
SRR400037	NA11831_1000G_LowCov_SRR400037	SRR400037 (9Gb, 3x)	0 - Overall (With alignment and HC only)	16.9	39997	2369
SRR400037	NA11831_1000G_LowCov_SRR400037	SRR400037 (9Gb, 3x)	0 - Overall (Without alignment and HC only)	38.1	37266	979
ERR013164	NA12383_1000G_LowCov_ERR013164	ERR013164 (5Gb, 2x)	0 - Overall (With alignment and HC only)	17	24980	1470
ERR013164	NA12383_1000G_LowCov_ERR013164	ERR013164 (5Gb, 2x)	0 - Overall (Without alignment and HC only)	33.9	23437	692
SRR385753	NA11994_1000G_LowCov_SRR385753	SRR385753 (15Gb, 5x)	0 - Overall (With alignment and HC only)	17.2	57961	3362
SRR385753	NA11994_1000G_LowCov_SRR385753	SRR385753 (15Gb, 5x)	0 - Overall (Without alignment and HC only)	37.9	53491	1411

Supplementary Table 4.

Stage	Case	Runtime MuTect/MuTect2	Runtime Sentieon	Improvement
Alignment-Tumor	HGSC_case6_WGS	46576	29408	1.6
Calculate Metrics-Tumor	HGSC_case6_WGS	8151	408	20
Remove Duplicates-Tumor	HGSC_case6_WGS	35234	1877	18.8
Indel Realignment-Tumor	HGSC_case6_WGS	42507	2057	20.7
Base Quality Score Recalibration-Tumor	HGSC_case6_WGS	158679	3092	51.3
Alignment-Normal	HGSC_case6_WGS	22868	14255	1.6
Calculate Metrics-Normal	HGSC_case6_WGS	4016	212	19
Remove Duplicates-Normal	HGSC_case6_WGS	17611	938	18.8
Indel Realignment-Normal	HGSC_case6_WGS	21534	1034	20.8
Base Quality Score Recalibration-Normal	HGSC_case6_WGS	82959	1634	50.8
Corealignment of Indels	HGSC_case6_WGS	96626	4839	20
TNhaplotyper	HGSC_case6_WGS	86323	7934	10.9
TNsnv	HGSC_case6_WGS	213840	1734	123.3
Alignment-Tumor	HGSC_case7	2170	1652	1.3
Calculate Metrics-Tumor	HGSC_case7	580	39	14.8
Remove Duplicates-Tumor	HGSC_case7	2209	132	16.7
Indel Realignment-Tumor	HGSC_case7	3872	211	18.3
Base Quality Score Recalibration-Tumor	HGSC_case7	11750	188	62.7
Alignment-Normal	HGSC_case7	2158	1526	1.4
Calculate Metrics-Normal	HGSC_case7	529	38	13.9
Remove Duplicates-Normal	HGSC_case7	2176	130	16.8
Indel Realignment-Normal	HGSC_case7	3703	195	18.9
Base Quality Score Recalibration-Normal	HGSC_case7	11364	180	63.1
Corealignment of Indels	HGSC_case7	9198	479	19.2
TNhaplotyper	HGSC_case7	14989	908	16.5
TNsnv	HGSC_case7	22916	119	193.3
Alignment-Tumor	HGSC_case7_WGS	29609	18434	1.6
Calculate Metrics-Tumor	HGSC_case7_WGS	6977	364	19.2
Remove Duplicates-Tumor	HGSC_case7_WGS	25082	1304	19.2
Indel Realignment-Tumor	HGSC_case7_WGS	23854	1077	22.1
Base Quality Score Recalibration-Tumor	HGSC_case7_WGS	126111	1937	65.1
Alignment-Normal	HGSC_case7_WGS	14717	8632	1.7
Calculate Metrics-Normal	HGSC_case7_WGS	3353	195	17.2
Remove Duplicates-Normal	HGSC_case7_WGS	12569	632	19.9
Indel Realignment-Normal	HGSC_case7_WGS	11403	553	20.6
Base Quality Score Recalibration-Normal	HGSC_case7_WGS	66012	1006	65.6
Corealignment of Indels	HGSC_case7_WGS	52459	2415	21.7
TNhaplotyper	HGSC_case7_WGS	791142	104963	7.5
TNsnv	HGSC_case7_WGS	151693	467	324.7

Supplementary Table 5.

Case name	F1-score	F1-score SNP	F1-score Indel	INDELs in both GATK and Sentieon	INDELs present in GATK but missed by Sentieon	INDELs present in Sentieon but not in GATK	SNPs in both GATK and Sentieon	SNPs present in GATK but missed by Sentieon	SNPs present in Sentieon but not in GATK
SRR742200 (5Gb, 97x)	0.999	0.999	0.998	1240	3	2	27268	18	26
SRR709972 (6Gb, 110x)	0.999	0.999	1.000	1215	0	1	27372	14	27
SRR702068 (7Gb, 130x)	1.000	1.000	1.000	1339	0	0	27799	4	13
ERR034544 (11Gb, 207x)	1.000	1.000	0.999	1471	0	2	25816	4	9
ERR034545 (11Gb, 222x)	1.000	1.000	0.999	1567	2	0	26276	5	6
SRR098401 (14Gb, 285x)	0.999	0.999	0.998	2200	3	4	36325	24	21
SRR098359 (14Gb, 284x)	0.999	0.999	1.000	2099	0	2	35902	17	22
ERR001695 (408Mb, 1x)	1.000	1.000	0.997	352	0	2	8627	1	4
ERR000228 (495Mb, 1x)	1.000	1.000	1.000	636	0	0	12551	3	2
ERR000854 (503Mb, 1x)	1.000	1.000	1.000	633	0	0	13452	0	0
ERR000902 (553Mb, 1x)	1.000	1.000	1.000	751	0	0	16155	0	1
ERR000162 (593Mb, 1x)	0.999	0.999	0.999	858	0	1	18416	6	19
ERR001509 (589Mb, 1x)	1.000	1.000	1.000	800	0	0	20171	1	2
ERR000876 (619Mb, 1x)	0.997	0.997	0.998	860	2	1	19791	10	93
ERR001502 (630Mb, 1x)	1.000	1.000	1.000	943	0	0	20998	1	10
SRR006107 (1Gb, 1x)	1.000	1.000	0.999	882	0	2	16327	4	10
SRR024100 (2Gb, 1x)	0.999	0.999	1.000	11402	2	4	150956	49	149
SRR029852 (2Gb, 1x)	0.998	0.998	0.999	12438	2	14	174251	102	594
ERR003045 (1Gb, 1x)	1.000	1.000	1.000	6098	3	0	80939	11	7
ERR003038 (1Gb, 1x)	1.000	1.000	0.999	5359	4	4	58489	18	24
ERR015474 (3Gb, 2x)	1.000	1.000	1.000	44190	3	10	465489	88	175
ERR003052 (2Gb, 1x)	1.000	1.000	1.000	6208	1	2	74745	24	21
SRR026644 (2Gb, 2x)	1.000	1.000	1.000	21657	2	1	291000	28	57
SRR027531 (3Gb, 2x)	1.000	1.000	1.000	24302	3	5	300330	113	101
SRR032636 (3Gb, 2x)	1.000	1.000	1.000	26968	3	12	312909	81	157
SRR029683 (3Gb, 2x)	1.000	1.000	1.000	26728	8	10	317084	97	123
SRR035028 (3Gb, 2x)	1.000	1.000	1.000	33799	4	4	389303	66	32
ERR015477 (4Gb, 2x)	1.000	1.000	1.000	68508	10	19	651526	103	133
ERR162816	1.000	1.000	1.000	64017	18	13	716255	178	128

(4Gb, 2x)									
ERR009417 (4Gb, 2x)	1.000	1.000	1.000	32502	10	11	461495	156	168
SRR035016 (3Gb, 2x)	0.999	0.999	1.000	37615	7	24	440867	118	482
ERR162833 (4Gb, 2x)	1.000	0.999	1.000	64860	16	13	731947	163	577
ERR162824 (4Gb, 2x)	1.000	1.000	1.000	64412	5	10	717928	124	127
ERR009404 (3Gb, 2x)	1.000	1.000	1.000	32117	0	4	412303	43	61
ERR019488 (4Gb, 2x)	1.000	1.000	1.000	73109	14	12	705588	134	150
SRR027518 (3Gb, 2x)	1.000	1.000	1.000	24599	5	1	313791	90	96
ERR009412 (3Gb, 2x)	1.000	1.000	1.000	32355	10	8	401492	86	113
ERR009416 (3Gb, 2x)	1.000	1.000	1.000	30688	2	5	375231	76	102
ERR009409 (3Gb, 2x)	1.000	1.000	1.000	32502	10	11	402040	109	104
SRR029680 (3Gb, 2x)	1.000	1.000	1.000	28770	6	5	359670	100	114
ERR019486 (5Gb, 2x)	1.000	1.000	1.000	105924	26	24	864256	145	126
ERR013164 (5Gb, 2x)	1.000	1.000	1.000	91573	18	7	845434	136	135
SRR034932 (4Gb, 2x)	1.000	1.000	1.000	44416	7	8	501883	160	108
ERR162832 (5Gb, 2x)	1.000	1.000	1.000	87757	16	11	944020	206	168
SRR035025 (4Gb, 2x)	1.000	1.000	1.000	49490	10	2	559106	101	129
SRR035020 (5Gb, 2x)	1.000	1.000	1.000	67465	14	8	736102	117	138
SRR400037 (9Gb, 3x)	1.000	1.000	1.000	173339	17	21	1590606	188	254
SRR385777 (9Gb, 3x)	1.000	1.000	1.000	193602	16	27	1762271	301	293
SRR385754 (10Gb, 4x)	1.000	1.000	1.000	203293	17	27	1742834	196	273
SRR385773 (10Gb, 4x)	1.000	1.000	1.000	224139	22	19	1878420	223	321
SRR211278 (10Gb, 4x)	1.000	1.000	1.000	230885	15	11	1921535	238	208
SRR385763 (10Gb, 4x)	1.000	1.000	1.000	224618	16	22	1903679	290	270
SRR393994 (11Gb, 4x)	1.000	1.000	1.000	220017	40	17	1818111	219	283
SRR075006 (11Gb, 4x)	1.000	1.000	1.000	226010	34	31	1917784	404	340
SRR393988 (11Gb, 4x)	1.000	1.000	1.000	236942	15	13	1975378	250	197
SRR393992 (12Gb, 4x)	1.000	1.000	1.000	237121	28	26	2155376	287	300
SRR385757 (12Gb, 4x)	1.000	1.000	1.000	237557	20	22	2077890	212	264
SRR385767 (12Gb, 4x)	1.000	1.000	1.000	246183	45	37	2044204	417	532
ERR050083 (13Gb, 5x)	1.000	1.000	1.000	269677	38	48	2229393	426	377
SRR385760 (13Gb, 5x)	1.000	1.000	1.000	275177	25	23	2249087	266	340
SRR768303 (13Gb, 5x)	1.000	1.000	1.000	240573	83	68	2194489	543	576
SRR741366 (14Gb, 5x)	1.000	1.000	1.000	325553	48	36	2474779	430	333

SRR768308 (14Gb, 5x)	1.000	1.000	1.000	257072	86	64	2338201	549	546
SRR385759 (15Gb, 5x)	1.000	1.000	1.000	325221	28	23	2470357	296	317
SRR768305 (15Gb, 6x)	1.000	1.000	1.000	275377	77	64	2423528	569	600
SRR385775 (15Gb, 5x)	1.000	1.000	1.000	319081	22	18	2476771	313	258
SRR385753 (15Gb, 5x)	1.000	1.000	1.000	321739	26	27	2483230	275	245
SRR385756 (15Gb, 5x)	1.000	1.000	1.000	319160	38	21	2477392	356	330
SRR741369 (15Gb, 6x)	1.000	1.000	1.000	292213	74	61	2479144	572	755
ERR257990 (15Gb, 5x)	1.000	1.000	1.000	366403	54	35	2620432	559	512
SRR768310 (16Gb, 6x)	1.000	1.000	1.000	275869	88	60	2412148	724	631
SRR622461 (16Gb, 6x)	1.000	1.000	1.000	349772	39	44	2554085	422	398
SRR385771 (16Gb, 6x)	1.000	1.000	1.000	343268	34	39	2573401	333	365
ERR257983 (17Gb, 6x)	1.000	1.000	1.000	386294	48	48	2728667	481	506
SRR393991 (17Gb, 6x)	1.000	1.000	1.000	332988	46	61	2517664	434	587
ERR257986 (17Gb, 6x)	1.000	1.000	1.000	406318	52	31	2795236	502	586
SRR385769 (17Gb, 6x)	1.000	1.000	1.000	317855	23	22	2488653	305	346
SRR385751 (18Gb, 6x)	1.000	1.000	1.000	363944	30	19	2697751	321	289
ERR257989 (18Gb, 6x)	1.000	1.000	1.000	422192	53	38	2883154	476	455
ERR257987 (19Gb, 7x)	1.000	1.000	1.000	440191	63	47	2943646	478	634
ERR239333 (22Gb, 8x)	1.000	1.000	1.000	451250	69	74	3095689	700	703
SRR400039 (24Gb, 8x)	1.000	1.000	1.000	363975	54	34	2856564	431	396
ERR233226 (24Gb, 9x)	1.000	1.000	1.000	467242	60	64	3181411	612	641
ERR091571 (41Gb, 14x)	1.000	1.000	1.000	646237	123	151	3769534	985	1215
pFDA 30x HG001 (113Gb, 39x)	0.999	0.999	0.999	831887	656	556	4018088	3248	3149
pFDA 50x HG002 (148Gb, 51x)	0.999	0.999	0.999	885955	509	506	4083018	3924	4329

Supplementary Table 6.

Sample	Shared Variants	Different Variants
NA06984_1000G_LowCov_SRR035025	53095	0
NA06985_1000G_LowCov_SRR400039	288008	1
NA06986_1000G_LowCov_SRR027520	38816	0
NA06989_1000G_LowCov_SRR035028	38016	0
NA07000_1000G_LowCov_ERR162816	71752	0
NA07037_1000G_LowCov_ERR257983	272277	0
NA07048_1000G_LowCov_SRR032636	33133	0
NA07051_1000G_LowCov_SRR026644	31601	0
NA07056_1000G_LowCov_ERR003038	8639	0
NA07347_1000G_LowCov_SRR029852	17764	0
NA07357_1000G_LowCov_ERR001509	1882	0
NA10851_1000G_LowCov_ERR001502	1947	0
NA11831_1000G_LowCov_SRR400037	164686	0
NA11832_1000G_LowCov_SRR385763	196813	0
NA11840_1000G_LowCov_SRR393988	210061	0
NA11843_1000G_LowCov_SRR035016	42760	0
NA11881_1000G_LowCov_ERR050083	226323	0
NA11892_1000G_LowCov_ERR015477	65522	0
NA11893_1000G_LowCov_ERR009416	43415	0
NA11894_1000G_LowCov_ERR000876	1932	0
NA11918_1000G_LowCov_SRR006107	1516	0
NA11919_1000G_LowCov_SRR027518	32140	0
NA11920_1000G_LowCov_SRR024100	16598	0
NA11930_1000G_LowCov_ERR233226	307631	2
NA11931_1000G_LowCov_ERR000902	1475	0
NA11992_1000G_LowCov_SRR385759	256070	0
NA11994_1000G_LowCov_SRR385753	256623	0
NA11995_1000G_LowCov_SRR385754	187497	0
NA12003_1000G_LowCov_SRR385756	255379	0
NA12004_1000G_LowCov_SRR385757	215670	0
NA12005_1000G_LowCov_SRR385767	213016	0
NA12043_1000G_LowCov_ERR000228	1110	0
NA12044_1000G_LowCov_SRR393991	256679	0
NA12045_1000G_LowCov_SRR027531	31114	0
NA12046_1000G_LowCov_ERR239333	297494	3
NA12058_1000G_LowCov_SRR035020	69302	0
NA12144_1000G_LowCov_ERR162832	97169	0
NA12154_1000G_LowCov_SRR385769	255036	0
NA12156_1000G_LowCov_SRR393992	210808	0

NA12249_1000G_LowCov_SRR211278	187413	0
NA12272_1000G_LowCov_ERR019488	75323	0
NA12273_1000G_LowCov_ERR162833	75293	0
NA12275_1000G_LowCov_ERR009417	54496	0
NA12287_1000G_LowCov_ERR000854	1306	0
NA12340_1000G_LowCov_SRR075006	195348	0
NA12341_1000G_LowCov_SRR034932	46737	0
NA12347_1000G_LowCov_ERR009404	44976	0
NA12348_1000G_LowCov_ERR009412	45624	0
NA12383_1000G_LowCov_ERR013164	79074	0
NA12399_1000G_LowCov_SRR029683	33485	0
NA12400_1000G_LowCov_SRR029680	37864	0
NA12413_1000G_LowCov_ERR003045	10210	0
NA12414_1000G_LowCov_SRR385771	260766	0
NA12489_1000G_LowCov_SRR393994	189253	0
NA12716_1000G_LowCov_ERR257986	279494	0
NA12717_1000G_LowCov_ERR257987	290987	0
NA12718_1000G_LowCov_ERR019486	93841	0
NA12749_1000G_LowCov_ERR162824	74167	0
NA12760_1000G_LowCov_SRR385773	195899	0
NA12761_1000G_LowCov_ERR257989	286276	0
NA12762_1000G_LowCov_SRR385775	252249	0
NA12763_1000G_LowCov_ERR257990	259352	0
NA12775_1000G_LowCov_ERR009409	44290	0
NA12776_1000G_LowCov_ERR000162	1625	0
NA12777_1000G_LowCov_ERR003052	10279	0
NA12778_1000G_LowCov_ERR015474	46592	0
NA12812_1000G_LowCov_SRR741369	250266	0
NA12813_1000G_LowCov_SRR385777	176725	0
NA12814_1000G_LowCov_SRR768303	220853	0
NA12815_1000G_LowCov_SRR768305	243769	0
NA12828_1000G_LowCov_ERR001695	767	0
NA12872_1000G_LowCov_SRR741366	249896	2
NA12873_1000G_LowCov_SRR768308	226836	0
NA12874_1000G_LowCov_SRR768310	229457	0

Supplementary Table 7.

Test name	Caller	Variant	Called by MuTect(2) and Sentieon	Called by MuTect(2) but not Sentieon	Called by Sentieon but not by MuTect(2)	Precision	Sensitivity	F-measure
Mutect1_HGSC_case7_mutect1	TNseq	Both	607	0	1	1	0.998355	0.999177
Mutect1_HGSC_case7_mutect1	TNseq	SNPs	607	0	1	1	0.998355	0.999177
Mutect1_HGSC_case7_mutect1	TNseq	INDELs	0	0	0	0	0	0
Mutect2_HGSC_case7_MuTect2	TNhaplotyper	Both	279	1	3	0.996429	0.989362	0.992883
Mutect2_HGSC_case7_MuTect2	TNhaplotyper	SNPs	252	1	3	0.996047	0.988235	0.992126
Mutect2_HGSC_case7_MuTect2	TNhaplotyper	INDELs	27	0	0	1	1	1
Mutect1_HGSC_case7_WGS_MuTect1	TNseq	Both	1832	2	0	0.998909	1	0.999454
Mutect1_HGSC_case7_WGS_MuTect1	TNseq	SNPs	1832	2	0	0.998909	1	0.999454
Mutect1_HGSC_case7_WGS_MuTect1	TNseq	INDELs	0	0	0	0	0	0
Mutect2_HGSC_case7_WGS_MuTect2	TNhaplotyper	Both	1265	6	3	0.995279	0.997634	0.996455
Mutect2_HGSC_case7_WGS_MuTect2	TNhaplotyper	SNPs	1140	4	3	0.996503	0.997375	0.996939
Mutect2_HGSC_case7_WGS_MuTect2	TNhaplotyper	INDELs	125	2	0	0.984252	1	0.992063
Mutect1_HGSC_case6_WGS_MuTect1	TNseq	Both	6559	2	0	0.999695	1	0.999848
Mutect1_HGSC_case6_WGS_MuTect1	TNseq	SNPs	6559	2	0	0.999695	1	0.999848
Mutect1_HGSC_case6_WGS_MuTect1	TNseq	INDELs	0:00:00	0	0	0	0	0
Mutect2_HGSC_case6_WGS_MuTect2	TNhaplotyper	Both	2078	33	23	0.984368	0.989053	0.986705
Mutect2_HGSC_case6_WGS_MuTect2	TNhaplotyper	SNPs	1725	26	18	0.985151	0.989673	0.987407
Mutect2_HGSC_case6_WGS_MuTect2	TNhaplotyper	INDELs	353	7	5	0.980556	0.986034	0.983287