

VU Learning HUB

BIF604 Quiz File

Chapter-wise Solved MCQs

vulearninghub.xyz

Module 1: Workshop Information & Document Structure

1. Which software viewer is specifically recommended in the handout for viewing the PDF document to ensure proper copy-and-paste functionality of commands and code blocks?

- A) Google Chrome
- B) Acrobat Reader
- C) Mozilla Firefox
- D) Evince PDF Viewer

2. According to the handout, what is the primary error type associated with Illumina high-throughput sequencing technology?

- A) Insertion errors
- B) Deletion errors
- C) Substitution errors
- D) Transversion errors

3. Under which Creative Commons license are the workshop documentation materials released?

- A) Creative Commons Attribution-NonCommercial 4.0
- B) Creative Commons Attribution 3.0 Unported License
- C) Creative Commons Zero (CC0)
- D) Creative Commons ShareAlike 2.0

4. Which quality control tool is used in the practicals to assess overall sequence quality and generate summary metrics across bases?

- A) Bowtie
- B) FastQC
- C) Velvet
- D) TopHat

Answer Key

Q1: B — Acrobat Reader is specifically recommended because it properly supports special characters and line skipping required for accurate command copy-pasting

Q2: C — Illumina 'sequence by synthesis' technology primarily produces substitution errors, which increase towards the 3' ends of reads

Q3: B — The workshop handout documentation is made available under a Creative Commons Attribution 3.0 Unported License

Q4: B — FastQC evaluates sequence read quality metrics and generates per-base quality boxplots and statistics

Module 2: Data Quality

1. In Phred quality score encoding, what quality encoding standard does the command option -Q 33 specify?

- A) Phred+64 quality encoding
- B) Solexa quality encoding
- C) Sanger / Phred+33 quality encoding
- D) Space-separated integer quality encoding

2. What is the effect of running fastq_quality_trimmer with parameters -t 20 and -l 50?

- A) Trims 20 bases from the 5' end and retains reads over 50 bp
- B) Filters bases at a quality cutoff of 20 and discards reads shorter than 50 bp
- C) Filters out reads with a GC content lower than 20%
- D) Converts FASTQ files to FASTA format if read length exceeds 50 bp

3. Which tool suite is highlighted for identifying duplicate reads based on identical start and end positions in SAM/BAM alignment files?

- A) FASTX-Toolkit
- B) Picard (MarkDuplicates)
- C) BEDTools
- D) SAMtools flagstat

4. In Illumina high-throughput sequencing data, where do base calling error rates typically increase in frequency?

- A) At the exact middle of each read
- B) Exclusively at the 5' end of reads
- C) Towards the 3' end of reads
- D) Evenly across all base positions

Answer Key

Q1: C — The -Q 33 option indicates that the input FASTQ quality scores are Sanger / Phred+33 encoded

Q2: B — The -t 20 option sets the quality score threshold to 20, and -l 50 discards any reads trimmed below 50 bp in length

Q3: B — Picard's MarkDuplicates tool detects duplicate reads by checking for identical alignment start and end coordinates in BAM files

Q4: C — Sequencing error rates in Illumina reads predominantly rise in frequency towards the 3' ends

Module 3: Read Alignment**1. Which Bowtie alignment flag instructs the aligner to write its output directly in SAM format?**

- A) -b
- B) -S
- C) -bam
- D) -sam

2. Which SAMtools command converts a text SAM file into a binary BAM file?

- A) samtools convert -b
- B) samtools view -bSo
- C) samtools sort -b
- D) samtools index -bam

3. In a SAM alignment record, what does a CIGAR string of 36M represent?

- A) 36 Mismatches
- B) 36 Matches or mismatches with no gaps or indels
- C) 36 Insertions
- D) 36 Mutated bases

4. Before generating an index file (.bai) using samtools index, what step must be performed on the BAM file?

- A) It must be trimmed using fastx_trimmer
- B) It must be sorted according to chromosomal coordinates
- C) It must be converted back to FASTQ format
- D) It must be filtered with Picard MarkDuplicates

Answer Key

Q1: B — Passing -S to Bowtie causes it to write alignment outputs in SAM format

Q2: B — samtools view -bSo specifies BAM output (-b), SAM input (-S), and output file destination (-o)

Q3: B — A CIGAR string of 36M means 36 alignment matches/mismatches without any insertions, deletions, or gaps

Q4: B — BAM files must be coordinate-sorted before samtools index can generate a valid .bai index file

Module 4: ChIP-Seq

1. How does MACS handle the issue of ChIP fragments being sequenced from both ends?

- A) By removing all reverse-strand reads
- B) By identifying bimodally mapped read patterns and shifting fragments inward
- C) By truncating read lengths to 10 bp
- D) By masking repetitive regions

2. What is the primary purpose of using a control dataset when calling peaks with MACS?

- A) To extend the read length
- B) To capture genomic biases and ensure that called peaks represent true ChIP enrichment
- C) To convert SAM files to BAM format
- D) To speed up alignment times

3. Which tool is used for motif discovery in peak regions during the ChIP-Seq practical?

- A) TopHat
- B) MEME
- C) Cufflinks
- D) Velvet

4. In MACS peak output files, what does the "number of tags" metric represent?

- A) The maximum peak height at the summit
- B) The read length
- C) The total number of reads across the entire peak region
- D) The Phred quality score

Answer Key

Q1: B — MACS detects bimodal peaks on forward and reverse strands and shifts fragments inward by half the peak distance

Q2: B — A control sample mirrors local background biases; peaks absent in the control confirm true binding enrichment

Q3: B — MEME is used to analyze FASTA peak sequences to discover enriched sequence motifs

Q4: C — Column 6 ("number of tags") in MACS peak files counts total reads within the peak region rather than summit height

Module 5: RNA-Seq

1. Which short read aligner is used in the RNA-Seq module specifically to align reads across splice junctions?

- A) Bowtie 1
- B) TopHat
- C) Velvet
- D) MACS

2. What expression metric is calculated by Cufflinks for quantifying transcript and isoform abundance?

- A) RPKM only
- B) FPKM (Fragments Per Kilobase of exon per Million fragments mapped)
- C) MAPQ score
- D) CIGAR count

3. Which tool within the Cufflinks suite compares transcript expression between conditions (e.g., 2cells vs. 6h) to identify differentially expressed genes?

- A) TopHat
- B) Cuffdiff
- C) FastQC
- D) Bowtie2

4. Which R-Bioconductor package is introduced in the tutorial to visualize and explore Cufflinks RNA-Seq results?

- A) ggplot2
- B) cummeRbund
- C) DESeq2
- D) edgeR-viz

Answer Key

Q1: B — TopHat is a specialized spliced aligner built on top of Bowtie that aligns RNA-Seq reads spanning exon-exon junctions

Q2: B — Cufflinks estimates transcript and gene abundances in FPKM

Q3: B — Cuffdiff measures differential expression at the gene and transcript levels between experimental samples

Q4: B — CummeRbund is the R-Bioconductor package used to visualize, filter, and explore Cuffdiff output data

Module 6: de novo Genome Assembly

1. What graph structure does the Velvet assembler use to perform de novo genome assembly?

- A) Overlap-Layout-Consensus graph
- B) de Bruijn graph
- C) Suffix tree
- D) Directed Acyclic Graph (DAG)

2. What is the primary role of the velveth command in the Velvet assembly workflow?

- A) To remove low-coverage contigs
- B) To break reads into k-mers and construct a hash table
- C) To map reads to a reference genome
- D) To visualize assemblies in Hawkeye

3. Which command flag in velvetg enables read-tracking, allowing detailed read information to be saved for AMOS visualization?

- A) -amos_file yes
- B) -read_trkg yes
- C) -cov_cutoff auto
- D) -exp_cov max

4. Which tool from the AMOS suite is used in the practical to view assembly scaffolds and contigs?

- A) IGV
- B) Hawkeye
- C) Ensembl Browser
- D) Cytoscape

Answer Key

Q1: B — Velvet is a de novo genome assembler based on the de Bruijn graph algorithm

Q2: B — velveth processes sequence reads, breaks them into k-mers, and constructs the dataset hash table

Q3: B — Setting -read_trkg yes enables read-tracking, which writes essential read details required for AMOS message files (.afg)

Q4: B — Hawkeye is the visual inspection tool in the AMOS suite used to examine assembly contigs and scaffolds

End of Quiz

BIF604 — 24 solved MCQs
Prepared by VU Learning HUB
vulearninghub.xyz