1. What procedure do we perform to read out an entire genome?

a. Reading

b. Sequencing

c. Determining

2. _____ is a procedure for analyzing the differences between two genomes of the same species.

a. Base calling

b. Variant Calling

c. Variant Determination

3. What molecular machine is responsible for unwinding and denaturing complementary strands upon transcription and replication?

a. DNA Helicase

b. DNAse

c. DNA unwinder

4. What was the name of the method that Fred Sanger used for sequencing DNA?

a. Sanger method

b. Chain Termination method

c. Shot gun sequencing

5. What is the name of the type of nucleotides used in Chain Termination seq?

a. dideoxy nucleotides or ddNTPs

b. deoxy nucleotides dNTPs

c. NTPs

6. When current flows accross an electrophoresis gel, larger/longer molecules travel faster than shorter ones.

a. True

b. False

7. The technique of isolating DNA fragments according to their length along an electrophorsis gel is referred to as....

a. Northern blotting

b. Southern blotting

c. blotting

8. What is the name of the class of sequencing technologies that is in use today?

a. Next Generation Sequencing

b. Modern Sequencing

c. Novel Genome Sequencing

9. NGS technologies heavily rely on the sequencing by _____ approach.

a. Synthesis

b. Construction

c. Denaturation

10. What does Sequencing by synthesis mean? What is being synthesized?

a. RNA

b. DNA

c. mRNA

11. What is the name of the chemical reaction that facilitates DNA synthesis?

The polymerase chain reaction (PCR) is a method widely used to rapidly make millions to billions of copies (complete or partial) of a specific DNA sample.

12. Thermo cycling is the thermal control procedure by which we control polymerase acitivity.

a. True

b. False

13. What happens to polymerase if we increase the temperature above 100C?

a. Disintegrates

b. Works faster

c. Works slower (not sure about this)

14. What is the name of the polymerase that we use today for DNA synthesis?

Taq polymerase is a thermostable DNA polymerase I named after the thermophilic eubacterial microorganism Thermus aquaticus, from which it was originally isolated by Chien et al. in 1976. Deoxyribonucleic acid polymerase from the extreme thermophile Thermus aquaticus, A Chien, D B Edgar, and J M Trela, J Bacteriol. 1976 Sep; 127(3): 1550–1557, doi:10.1128/jb.127.3.1550-1557.1976

15. Can polymerase work without a primer?

a. Yes

b. No

16. How do we anneal a primer to its complementary region on a single stranded DNA, without the use of polymerase?

a. Hybridization

b. Complementary union

c. Hydrolyzation

17. What kind of a reaction is a reaction in which two independent DNA molecules are appended together?

a. Hybridization reaction

b. Ligase reaction

c. Polymerase Chain Reaction

18. GC bias is the result of what?

a. strong hydrogen bonds between G and C

b. weak hydrogen bonds between G and C

c. systematic errors of the sequencer

19. What are the two most widely used NGS technologies today? Check all that apply.

a. CRT/TIRF

b. SMRT

c. ION Semiconductor

d. NANO pore

20. What are the differences between these two technologies? Check all that apply. Check all that apply.

a. molecular immobilization

b. cycles vs. realtime

c. template preparation

21. illumina is the biggest manufacturer of SMRT sequencing technologies.

a. True

b. False

22. What is the difference between a fragment, an insert, a template and a read?

DNA fragmentation is the separation or breaking of DNA strands into pieces. Those pieces are called fragments.

An insert is a fragment of DNA that is inserted into a larger DNA vector by a recombinant DNA technique, such as ligation or recombination, allowing it to be multiplied, selected or further manipulated.

Template is a polynucleotide that encodes the information from which another polynucleotide, of complementary sequence, is synthesized.

A read is an inferred sequence of base pairs (or base pair probabilities) corresponding to all or part of a single DNA fragment.

23. How is fragmentation of genomic DNA performed?

a. Chemically

b. Physically

c. Sonication

24. Adapters are primers and primer complements.

a. True

b. False

25. Why does illumina use cluster amplification of templates?

a. To increase photo-sensitivity

b. To decrease margin of error

c. Increase quality scores

26. After "x" thermo-cycles in the the c-bot machine during cluster prep, how many copies of each individual template does each cluster contain?

a. 2^x

b. x^2

c. x^2^2

27. What is the name of the glass slide used by illumina, for sequencing?

a. Cell

b. Flow Cell

c. Lane

28. How many lanes can a slide house?

a. Eight (but also 1 and 2 for some types of flow cell)

b. Twelve

c. Two

29. Lanes are divided into .....?

a. Boxes

b. Squares

c. Tiles

30. Typically how many clusters can we find per tile?

a. 30.000

b. 20.000

c. 3.000.000

31. What is the name of the amplification technology used by iLLumina?

a. CBA -

b. Clonal Bridge Amplification

c. Solid Phase Amplification

d. PCR free amplification

32. How many pictures are taken per cycle using illumina seq?

a. 1

b. 2

c. 3

d. 4

33. What is the biggest problem with SMRT sequencing?

a. Long reads

b. Bad cameras

c. Dark nucleotides

34. What is immobilized in SMRT?

a. Templates

b. Polymers

c. Inserts

35. SMRT uses clonal bridge amplification.

a. True

b. False

36. SMRT reads can be 10.000 bp long.

a. True

b. False

37. Reads are stored in text files.

a. True

b. False

38. How are these files encoded?

a. UTF

b. ASCII

c. RUSKI

39. What is the name of the text file which stores reads and read data?

a. FASTA

b. FASTQ

c. BAM

40. FASTQ files are comprised of _____ .

a. lines

b. quality scores

c. records

41. Each record contains two parts.... Check all that apply.

a. nucleotide sequence

b. primer sequence

c. quality score sequence

42. The quality scores of nucleotide sequences are based on.....

a. Systematic errors

b. Fluorescent intensities of base calls

c. Nucleotide context evaluation

43. The name of the algorithm that performs quality scoring is...

a. Sanger

b. Fred

c. Phred

44. The quality scores for each nucleotide are encoded according to ....

a. UTF table

b. HTML table

c. ASCII table

45. Quality scoring can use different scales. These scales are related to what in the ASCII table?

a. Start and stop intervals

b. ASCII table start offsets

c. ASCII table window

46. Which scales are typically used by GATK suite tools, when illumina reads are in question? Check all that apply.

a. Sanger

b. illumina 1.8

c. illumina 1.3

47. Analytically, quality scores are log transformed _____.

a. base calls

b. error rates

c. success rates

48. Why do we encode this transformation using ascii?

a. To save space in memory

b. Ease of reading

c. Faster

49. Typically how long are illumina DNA templates?

a. 100 bp

b. 1000bp

c. 500 bp

d. 200bp

50. Why do we perform paired-end sequencing?

Paired-end sequencing allows us to sequence both ends of a fragment and generate high-quality sequence data, in addition to producing twice the number of reads for the same time and effort in library preparation. Sequences aligned as read pairs enable more accurate read alignment and the ability to detect insertion-deletion (indel) variants, which is not possible with single-read data.