

PROVA

1

VERSIONE

A

- 1. Un individuo si definisce eterozigote quando:**
  - a. ha due cromosomi sessuali diversi
  - b. ha una quantità maggiore di eterocromatina rispetto alla norma
  - c. possiede alleli diversi ad un determinato locus
  - d. possiede specie diverse di DNA mitocondriale
  
- 2. La trascrizione:**
  - a. avviene sui ribosomi
  - b. è quel processo attraverso cui l'informazione genetica viene tradotta in proteine
  - c. è quel processo attraverso cui l'informazione genetica passa dal DNA all'RNA
  - d. avviene durante il crossing-over
  
- 3. Quali enzimi metilano il DNA?**
  - a. DNA polimerasi RNA dipendente
  - b. DNA acetilasi
  - c. DNA metiltransferasi
  - d. DNA topoisomerasi
  
- 4. Il "crossing over" è:**
  - a. il processo di maturazione dei trascritti
  - b. il processo di ricombinazione tra cromosomi omologhi
  - c. il processo di mancata disgiunzione dei cromosomi durante la meiosi
  - d. il sistema di sequenziamento dei trascritti
  
- 5. Il codice genetico del DNA nucleare e mitocondriale è interscambiabile?**
  - a. sì, ma solo per i geni delle subunità del complesso III
  - b. sì, ma solo per i codoni di stop
  - c. no, alcuni codoni sono diversi
  - d. no, solo i codoni di stop sono diversi
  
- 6. Un polimorfismo è:**
  - a. un cambio nucleotidico senza effetti funzionali
  - b. una variante cromosomica estremamente rara a livello di sequenze ripetute
  - c. una variante rispetto alla sequenza di riferimento con una frequenza >1%
  - d. una mutazione patogenetica validata sperimentalmente
  
- 7. In un saggio quantitativo di PCR real-time, qual è la differenza tra l'utilizzo di intercalanti fluorescenti e le sonde fluorescenti?**
  - a. non è necessario disegnare i primers per il saggio
  - b. non serve ottimizzare le condizioni della reazione
  - c. la possibilità di discriminare tra prodotti di amplificazione specifici e aspecifici
  - d. la quantità di DNA stampo non deve essere dosata
  
- 8. Geni con lo stesso ancestore ma appartenenti a specie diverse sono detti:**
  - a. omeotici
  - b. paraloghi
  - c. ortologhi
  - d. protooncogeni
  
- 9. Se la probabilità di malattia è di 0,2 in ciascuna delle due persone in osservazione, qual è la probabilità che entrambe si ammalinano?**
  - a. 0,2
  - b. 0,4
  - c. 0,64
  - d. 0,04

- 10. Quale tra le seguenti può essere considerata una variabile quantitativa discreta:**
- età
  - glicemia
  - colesterolemia
  - numero di figli
- 11. Il sequenziamento dell'intero esoma (Whole Exome Sequencing) prevede l'analisi di:**
- sequenza delle regioni codificanti del genoma
  - intera sequenza del cromosoma X
  - sequenza delle regioni del genoma associate a patologia umana
  - circa 5000 esoni (corrispondenti a circa 300 geni)
- 12. L'analisi trascrittomica:**
- permette di individuare trascritti aberranti e splicing alternativi
  - è indipendente dal campione biologico analizzato
  - valuta l'effetto di varianti missenso sulla stabilità della proteina
  - permette di identificare varianti introniche
- 13. La metabolomica:**
- è un'analisi esclusivamente quantitativa
  - si basa sul sequenziamento di nuova generazione
  - può misurare nucleotidi e lipidi
  - valuta la metilazione del DNA
- 14. Nell'analisi dei dati NGS le varie fasi da eseguire in sequenza sono:**
- chiamata varianti - allineamento - annotazione
  - annotazione - allineamento - chiamata varianti
  - allineamento - annotazione - chiamata varianti
  - allineamento - chiamata varianti - annotazione
- 15. Per quale di questi studi risulta più adatto un sequenziamento NGS basato su short-read:**
- identificare e quantificare espansioni da triplette
  - identificare inserzioni/delezioni genomiche di piccole dimensioni (1-50 bp)
  - sequenziare *de novo* il genoma di un organismo
  - identificare inversioni cromosomiche
- 16. Il sequenziamento whole exome sequencing (WES) basato su short-read permette di ottenere una copertura di sequenziamento (coverage):**
- omogenea lungo tutte le regioni cromosomiche
  - omogenea solo all'interno delle regioni geniche
  - elevata solo in corrispondenza degli esoni all'interno dei vari geni
  - elevata solo nelle regioni telomeriche dei cromosomi
- 17. Clustal Omega è un programma per:**
- identificare domini funzionali
  - predire l'effetto di una variante di splicing
  - predire l'effetto deleterio di una variante missenso
  - allineare sequenze multiple
- 18. Phred è un programma per:**
- chiamare e valutare la qualità di sequenze di amino acidi
  - chiamare e valutare la qualità di sequenze di nucleotidi
  - identificare i siti di taglio di enzimi di restrizione
  - allineare sequenze multiple
- 19. In un sequenziamento NGS, i file ".fastq" contengono:**
- file di testo di sequenze nucleotidiche
  - elenco delle varianti presenti nel campione
  - annotazioni delle varianti presenti nel campione
  - sequenze allineate al genoma di riferimento
- 20. I criteri ACMG (American College of Medical Genetics and Genomics):**
- classificano le varianti in 5 categorie
  - indicano gli studi funzionali da compiere
  - danno indicazioni sulla diagnosi prenatale
  - definiscono i geni-malattia

**21. Gli approcci NGS basati su long-read sono ottimali per:**

- a. sequenziare gli esoni di piccoli gruppi di geni.
- b. misurare i livelli di espressione di uno specifico gene.
- c. sequenziare *de novo* il genoma di un organismo.
- d. analizzare la segregazione di una variante puntiforme nota nei vari individui di una famiglia.

**22. Quale tra questi sequenziatori commercialmente disponibili si basa su chimica long-read:**

- a. IonTorrent
- b. PacBio Sequel
- c. Illumina MiSeq
- d. Sequenziatore Sanger

**23. I files BED (Browser Extensible Data) sono composti dai seguenti campi obbligatori:**

- a. line name, chromosome, start position e stop position
- b. chromosome, start position e stop position
- c. chromosome, start position, stop position e strand
- d. chromosome, start position, stop position e track name

**24. Un Phred Quality Score di 30 (Q30) indica:**

- a. una probabilità di 1/1000 che la base chiamata sia sbagliata
- b. una probabilità di 30:1 che la base chiamata sia corretta
- c. un'accuratezza del 99.999% nella chiamata della base
- d. una probabilità di 0.003 che la base chiamata sia sbagliata

**25. Genome Analysis Toolkit (GATK): quale delle seguenti affermazioni è corretta?**

- a. è un tool che permette di generare un BCL file
- b. è un tool che permette di generare un VCF file
- c. è un tool che permette di generare un BAM file
- d. è un tool che permette di generare un FASTQ file

**26. I TAG-SNP:**

- a. sono SNP in linkage disequilibrium
- b. sono l'insieme degli SNP che costituiscono un aplotipo
- c. sono gli SNP che vengono imputati a partire dai dati di uno SNP-array
- d. sono SNP rappresentativi di un aplotipo

**27. In quali ambiti è possibile utilizzare tecniche NGS?**

- a. solo per sequenziamento DNA
- b. solo per sequenziamento DNA e RNA
- c. per sequenziamento DNA e RNA e per studi di metilazione sul DNA
- d. per sequenziamento e studi di metilazione sul DNA, ma non per analisi su RNA

**28. Il genome-wide association study (GWAS) è:**

- a. un metodo per identificare le varianti genetiche rare associate ad una specifica malattia o gruppo di individui
- b. un metodo per ottenere dati osservazionali associati ad una specifica malattia o gruppo di individui
- c. un metodo per ottenere dati di espressione genica associati ad una specifica popolazione
- d. un metodo per identificare varianti genetiche associate ad una specifica malattia o gruppo di individui

**29. Quale dei seguenti step richiede in input un file di coordinate genomiche**

- a. allineamento
- b. trimming
- c. identificazione dei duplicati
- d. chiamata delle varianti

**30. La moda:**

- a. corrisponde alla sommatoria di tutti i valori diviso il numero delle osservazioni
- b. corrisponde al 25° percentile
- c. è il valore che si manifesta con maggior frequenza in una serie di osservazioni
- d. è una misura di tendenza centrale

**31. La correlazione:**

- a. Misura la direzione e la forza della relazione lineare fra due variabili qualitative
- b. Misura la direzione e la forza della relazione lineare fra due variabili quantitative
- c. Misura unicamente la direzione della relazione lineare fra due variabili quantitative
- d. Può assumere unicamente valori positivi

**32. L'obiettivo dell'inferenza statistica è:**

- a. Rispondere a domande specifiche poste prima di avere a disposizione i dati
- b. Rispondere a domande che sorgono una volta osservati i risultati di una sperimentazione
- c. Effettuare il maggior numero di test possibili al fine di osservare una significatività
- d. Verificare ipotesi non generalizzabili

**33. Aumentando la numerosità campionaria mantenendo fisso il livello di significatività:**

- a. Non si ottiene alcun effetto sulla potenza di un test di significatività
- b. La potenza di un test di significatività aumenta
- c. La potenza di un test di significatività diminuisce
- d. La probabilità di avere falsi positivi aumenta

**34. Se la distribuzione di una popolazione si allontana drasticamente dalla distribuzione normale, allora il test t di Student:**

- a. È inappropriato
- b. È indubbiamente il test più appropriato
- c. È equivalente ad un test non parametrico
- d. Permette di verificare se la distribuzione della popolazione in esame è simmetrica

**35. Quale tra questi non è un modello probabilistico:**

- a. Bernulli
- b. Gay-Lussac
- c. Poisson
- d. Gauss

PROVA  
2

VERSIONE  
A

**1. L'allele che in un organismo eterozigote NON manifesta la sua azione viene detto:**

- a. dominante
- b. minore
- c. recessivo
- d. co-dominante

**2. Il telomero è:**

- a. la parte terminale di un cromosoma
- b. la parte centrale di un cromosoma
- c. una regione regolativa sul braccio lungo dei cromosomi
- d. il punto di formazione del cinetocore

**3. Che termine indica "un carattere misurabile posseduto da un organismo, visibile o deducibile con ricerche compiute per la sua identificazione"?**

- a. genotipo
- b. fenotipo
- c. cariotipo
- d. aplotipo

**4. Cos'è il DNA mitocondriale:**

- a. tratti di DNA trasferiti dal nucleo ai mitocondri
- b. DNA non codificante residuo del trasferimento genico durante l'endosimbiosi
- c. DNA circolare presente all'interno dei mitocondri codificante proteine della catena respiratoria mitocondriale
- d. DNA circolare intra-mitocondriale che sintetizza trascritti esportati nel citoplasma

**5. Il processo di splicing:**

- a. coinvolge sequenze di DNA
- b. avviene a livello degli RNA messaggeri
- c. è una modificazione post-traduzionale
- d. porta alla formazione di microRNA

**6. La tecnica della PCR consente di:**

- a. isolare frammenti di DNA
- b. confrontare sequenze di DNA
- c. amplificare frammenti di DNA
- d. inserire frammenti di DNA esogeno in una cellula

**7. L'epigenetica studia:**

- a. cambi irreversibili a livello del DNA
- b. mutazioni genetiche a carico degli istoni
- c. cambi nella cromatina che regolano l'espressione genica
- d. i legami idrogeno tra basi azotate complementari

**8. Quale tipo di mutazione non è possibile diagnosticare mediante sequenziamento di Sanger?**

- a. Mutazione di splicing eterozigote
- b. Delezione esonica eterozigote
- c. Mutazione missenso eterozigote
- d. Mutazione missenso omozigote

**9. La curva di Kaplan-Meier serve per rappresentare:**

- a. La crescita cellulare in funzione del numero di passaggi in coltura
- b. La distribuzione di un parametro all'interno della popolazione
- c. La sopravvivenza in funzione del tempo
- d. La risposta ad un farmaco in funzione alla sua concentrazione

**10. Geni omologhi con diversa funzione nello stesso genoma sono detti:**

- a. omeotici
- b. paraloghi
- c. ortologhi
- d. protooncogeni

**11. La regione target sequenziata mediante un approccio di Whole Exome Sequencing è:**

- a. di qualche decina di kb
- b. di qualche decina di Mb
- c. di qualche decina di Gb
- d. di qualche decina di Tb

**12. Nell'analisi dei dati NGS le varie fasi da eseguire in sequenza sono:**

- a. filtraggio - allineamento - annotazione
- b. annotazione - filtraggio - chiamata varianti
- c. allineamento - annotazione - chiamata varianti
- d. allineamento - chiamata varianti - annotazione

**13. BlastN è un algoritmo per:**

- a. cercare regioni di similarità tra sequenze nucleotidiche
- b. trovare i promotori
- c. predire l'effetto deleterio di una variante nonsense
- d. identificare i siti di legame dei miRNA

**14. Nella chiamata delle varianti, il quality score Q30 corrisponde ad una probabilità di base non corretta pari a:**

- a. 10%
- b. 0,3%
- c. 0,1%
- d. 0.01%

**15. In un sequenziamento NGS, i file ".vcf" corrispondono a:**

- a. sequenze allineate al genoma di riferimento
- b. elenco delle varianti presenti nel campione
- c. annotazioni delle varianti presenti nel campione
- d. valori di intensità di fluorescenza dei nucleotidi marcati

**16. I criteri ACMG (American College of Medical Genetics and Genomics):**

- a. integrano dati di frequenza di popolazione e segregazione ma non gli studi funzionali
- b. sono basati esclusivamente su studi funzionali
- c. considerano precedenti pubblicazioni sulla variante
- d. definiscono il fenotipo associato ad un determinato gene

**17. L'analisi dell'espressione differenziale basata sul sequenziamento dell'RNA:**

- a. valuta solo l'espressione dei singoli esoni
- b. è solo univariata
- c. può essere sia univariata che multivariata
- d. valuta solo l'espressione dei geni housekeeping

**18. Nei vari approcci NGS basati su short-read la lunghezza media delle singole sequenze ottenute è compresa tra:**

- a. 500-1.000 bp
- b. 75-300 bp
- c. 15-30 bp
- d. 10.000-11.000 bp

**19. Il sequenziamento Whole Exome Sequencing (WES) basato su short-read è ottimale per:**

- a. identificare delezioni/inserzioni di grandi dimensioni (>1.000 bp)
- b. identificare traslocazioni cromosomiche
- c. studiare i cambiamenti epigenetici del DNA nucleare
- d. identificare mutazioni puntiformi e inserzioni/delezioni di piccole dimensioni (1-50 bp)

- 20. Il sequenziamento whole genome basato su long-read permette di ottenere una copertura di sequenziamento (coverage):**
- omogenea lungo tutte le regioni cromosomiche.
  - elevata solo in prossimità degli esoni.
  - bassa nelle regioni intergeniche e alta nelle regioni ricche di geni.
  - variabile a seconda di quanto sono espressi i geni sequenziati.
- 21. Gli approcci NGS basati su long-read sono ottimali per:**
- identificare e quantificare trascritti generati da splicing alternativi.
  - identificare su DNA genomico inserzioni/delezioni di piccole dimensioni (1-50 bp).
  - identificare varianti malattia puntiformi in geni malattia noti.
  - analizzare la segregazione di una variante malattia nei genitori di un probando.
- 22. Quale dei seguenti non è un linguaggio di programmazione ad oggetti:**
- C
  - N
  - R
  - S
- 23. Il formato FASTQ è composto normalmente da 4 linee per sequenza che indicano rispettivamente:**
- L'identificativo della sequenza preceduto da carattere "#", la sequenza nucleotidica, la descrizione della sequenza preceduta dal carattere "+", I valori di qualità per ogni base in carattere binario.
  - L'identificativo della sequenza preceduto da carattere "@", la sequenza nucleotidica, la descrizione della sequenza preceduta dal carattere "-", I valori di qualità per ogni base espressi in numero da 0 a 50.
  - L'identificativo della sequenza preceduto da carattere "#", la sequenza nucleotidica, la descrizione della sequenza preceduta dal carattere "+", I valori di qualità per ogni base in carattere alfa-numerico.
  - L'identificativo della sequenza preceduto da carattere "@", la sequenza nucleotidica, la descrizione della sequenza preceduta dal carattere "+", I valori di qualità per ogni base in carattere ASCII.
- 24. In una analisi di target resequencing mediante NGS il controllo di qualità è effettuato in quale delle seguenti fasi:**
- nelle fasi pre-analitiche sul DNA di input
  - nelle fasi analitiche sulla libreria genomica
  - nelle fasi di analisi bioinformatica sui dati ottenuti
  - in tutte le fasi elencate
- 25. L' algoritmo di chiamata delle varianti più utilizzato di GATK è:**
- UnifiedGenotyper
  - GenotypeCaller
  - HaplotypeCaller
  - VariantHunter
- 26. Quante varianti mediamente vengono identificate da un'analisi di Whole Exome Sequencing:**
- qualche centinaia di varianti
  - qualche migliaia di varianti
  - qualche decina di migliaia di varianti
  - qualche milione di varianti
- 27. Quale dei seguenti non è un tool di allineamento:**
- LINE
  - STAR
  - LAST
  - BWA
- 28. Il sequenziamento NGS con tecnologia Illumina è basata su:**
- sequencing by ligation
  - pyrosequencing
  - shotgun sequencing
  - sequencing by synthesis
- 29. Quale dei seguenti step di analisi non richiede in input una sequenza di riferimento:**
- allineamento
  - chiamata delle varianti
  - identificazione delle reads duplicate
  - analisi del coverage

**30. La media aritmetica può essere calcolata:**

- a. Per variabili misurate su scala nominale
- b. Per variabili quantitative
- c. Per variabili misurate su scala ordinale
- d. Per qualsiasi variabile basta che la modalità sia un numero

**31. Il test del chi quadrato:**

- a. si utilizza per confrontare due proporzioni
- b. si utilizza per confrontare due medie
- c. si utilizza per calcolare i limiti di confidenza di una media
- d. si utilizza per calcolare l'errore standard di una media

**32. In uno studio osservazionale:**

- a. Tutte le unità vengono deliberatamente sottoposte a trattamenti al fine di osservare le reazioni provocate
- b. È possibile ricorrere alla randomizzazione delle unità
- c. Si osservano le unità e si misurano le variabili di interesse, ma non si cerca di influenzare le risposte
- d. Il trattamento viene somministrato in maniera casuale

**33. A quale delle seguenti domande risponde un test di significatività:**

- a. Il campione dell'esperimento è organizzato in modo appropriato?
- b. L'effetto osservato è dovuto al caso?
- c. L'effetto osservato è importante?
- d. L'effetto osservato è generalizzabile a popolazioni con caratteristiche diverse dal campione estratto?

**34. La correzione di Bonferroni:**

- a. È un metodo utilizzato in statistica per contrastare il problema dei confronti multipli
- b. Non si applica se le variabili sono di tipo dicotomico
- c. È un metodo utilizzato per il calcolo della dimensione campionaria
- d. È un metodo utilizzato in statistica per contrastare il problema del confondimento

**35. Il modello probabilistico di Gauss si applica a:**

- a. variabili binarie
- b. variabili continue
- c. variabili qualitative
- d. variabili discrete

PROVA  
3

VERSIONE  
A

**1. Il DNA mitocondriale umano contiene introni?**

- a. sì
- b. no
- c. sì, ma solo nei geni che codificano subunità del complesso I
- d. sì, ma solo nei tRNA

**2. Nella classica doppia elica del DNA la citosina è appaiata con**

- a. la timina
- b. l'adenina
- c. la guanina
- d. l'uracile

**3. Qual è il principale effetto dei siRNA?**

- a. attivazione genica
- b. inibizione dell'espressione genica
- c. inattivazione cromosomica
- d. inattivazione del cromosoma X

**4. Cosa si intende per Replisoma?**

- a. un insieme di proteine che super avvolgono il DNA
- b. un sistema di proteine coinvolte nel processo di replicazione del DNA
- c. un insieme di proteine che trascrivono il DNA
- d. un insieme di enzimi che consentono l'appaiamento dei cromosomi omologhi

**5. Cos'è un RNA policistronico?**

- a. Un RNA che può essere tradotto sia nel citoplasma che nei mitocondri
- b. Un RNA non codificante che porta l'informazione per produrre miRNA
- c. Un RNA che codifica per più proteine
- d. Un RNA a doppia o tripla elica

**6. Per individuare una piccola delezione subcromosomica l'analisi più indicata è:**

- a. analisi del cariotipo con bandeggio G
- b. ibridizzazione genomica comparativa
- c. sequenziamento dell'esoma
- d. analisi di sequenza di pannelli genici

**7. Nelle malattie complesse/multifattoriali:**

- a. i fattori genetici sono sempre più importanti di quelli ambientali
- b. i fattori ambientali sono sempre più importanti di quelli genetici
- c. il contributo genetico può essere valutato con studi sui gemelli
- d. la concordanza fenotipica nei gemelli monozigoti è sempre >50%

**8. Data una sequenza generica di DNA, in quanti modi è possibile tradurla:**

- a. 1
- b. 3
- c. 4
- d. 6

**9. Geni omologhi ma appartenenti a specie diverse sono detti:**

- a. ortologhi
- b. specifici
- c. paraloghi
- d. oncogeni

**10. Polygenic risk score: quale delle seguenti affermazioni è corretta?**

- a. è il rischio di sviluppare una malattia essendo portatori di mutazioni patogeniche in numerosi geni-malattia
- b. è il rischio di sviluppare una malattia se si è portatori di importanti fattori di rischio genetico
- c. è il rischio di sviluppare una malattia tenuto conto di un numero ampio di polimorfismi genetici ognuno dei quali dà un contributo molto piccolo alla patologia
- d. tutte le precedenti

**11. Il sequenziamento dell'intero esoma (Whole Exome Sequencing) prevede l'analisi:**

- a. della sequenza degli RNA messaggeri
- b. dell'intera sequenza dei cromosomi sessuali
- c. dell'intera sequenza dei cromosomi autosomici
- d. della sequenza delle regioni codificanti del genoma

**12. Dall'integrazione dei dati di trascrittomica con quelli di genomica si può:**

- a. Avere indicazione sulla stabilità di una proteina
- b. Stabilire l'effetto di una variante intronica
- c. identificare le modifiche epigenetiche del promotore
- d. dedurre l'origine parentale delle varianti nel DNA

**13. Quali sono le principali differenze tra il sequenziamento di Sanger e i sistemi di next-generation sequencing (NGS)?**

- a. nel sequenziamento di Sanger si sequenzia un singolo frammento di DNA alla volta
- b. si utilizzano enzimi diversi
- c. si analizzano solo gli esoni
- d. si analizzano solo i promotori dei geni

**14. Tra le sequenze nucleotidiche, EST indica:**

- a. Expression Silencing Timepoint
- b. Eastern Standard Time
- c. Expressed Sequence Tag
- d. Exome Sequence Tail

**15. Nell'analisi dei dati NGS le varie fasi da eseguire in sequenza sono:**

- a. chiamata varianti - allineamento - annotazione
- b. allineamento - chiamata varianti - annotazione
- c. allineamento - annotazione - chiamata varianti
- d. annotazione - chiamata varianti - allineamento

**16. BlastP è un programma che confronta:**

- a. sequenze nucleotidiche
- b. sequenze numeriche
- c. sequenze amino acidiche
- d. sequenze palindromiche

**17. In un sequenziamento NGS, i file ".bam" corrispondono a:**

- a. sequenze allineate al genoma di riferimento
- b. varianti presenti nel campione
- c. annotazioni delle varianti presenti nel campione
- d. valori di intensità di fluorescenza dei nucleotidi marcati

**18. I criteri ACMG (American College of Medical Genetics and Genomics):**

- a. valutano dati di frequenza di popolazione ma non gli studi funzionali
- b. sono basati esclusivamente su studi funzionali
- c. definiscono il fenotipo associato ad un determinato gene
- d. tengono in considerazione se la variante è de novo

**19. RPKM e FPKM sono metodi di normalizzazione di dati:**

- a. di espressione genica (sequenziamento RNA)
- b. di metilazione (epigenetica)
- c. di espressione proteica (proteomica)
- d. di frequenza di polimorfismi (studi GWAS)

- 20. Quale tra questi strumenti da laboratorio è un sequenziatore NGS che si basa su chimica con short-read:**
- Droplet Digital PCR System
  - PacBio Sequel
  - Illumina MiSeq
  - Nanopore MinION
- 21. Il sequenziamento RNAseq basato short-read è ottimale per:**
- sequenziare *de novo* il genoma di un organismo.
  - effettuare studi di differential gene expression.
  - analizzare la segregazione di una variante malattia nei genitori di un probando.
  - identificare ri-arrangiamenti del DNA mitocondriale.
- 22. Per quale di questi studi risulta più adatto un sequenziamento NGS basato su long-read:**
- identificazione di varianti genomiche puntiformi in geni malattia noti.
  - identificazione di inserzioni/delezioni genomiche di piccole dimensioni (1-50 bp).
  - identificazione di inserzioni/delezioni genomiche di grandi dimensioni (>1.000 bp).
  - quantificazione dei livelli di espressione di un gruppo di geni malattia noti.
- 23. Nei vari approcci NGS basati su long-read la lunghezza delle sequenze ottenute è generalmente:**
- compresa tra 75-300 bp
  - inferiore a 500 bp
  - compresa tra 300-800 bp
  - maggiore di 800 bp
- 24. Quale di queste affermazioni riguardo al Phred Quality Score è corretta:**
- il Phred Quality Score è inversamente proporzionale all'accuratezza della base chiamata
  - il Phred Quality Score è direttamente proporzionale alla probabilità di errore della base chiamata
  - il Phred Quality Score è inversamente proporzionale alla probabilità di errore della base chiamata
  - il Phred Quality Score è direttamente proporzionale alla stima dell'errore sulla base chiamata
- 25. Un file in formato FASTQ contiene le seguenti informazioni:**
- solo la sequenza nucleotidica
  - la sequenza nucleotidica e i rispettivi valori di qualità
  - L'identificativo della sequenza, la sequenza nucleotidica e i valori di qualità
  - L'identificativo della sequenza preceduto da carattere ">" e la sequenza nucleotidica
- 26. Quali di questi non è un tool di annotazione:**
- VEP
  - Coffalyzer
  - SNPeff
  - Annovar
- 27. Quale dei seguenti approcci è più indicato per l'identificazione di varianti rare e ultra-rare associate a patologia:**
- GWAS
  - SNParray
  - WGS
  - CGH-array
- 28. Quali dei seguenti non è un formato di allineamento:**
- sam
  - bam
  - fam
  - cram
- 29. Quali sono i vantaggi dell'approccio di RNA-seq rispetto ai microarrays:**
- maggior throughput
  - maggior riproducibilità
  - minor bias legato all'acquisizione e alla normalizzazione del dato grezzo
  - tutte le precedenti

**30. Quale delle seguenti è considerata una misura di dispersione?**

- a. media
- b. moda
- c. mediana
- d. deviazione standard

**31. In fase di verifica delle ipotesi, che cosa è il p-value?**

- a. La probabilità di rifiutare erroneamente l'ipotesi nulla quando questa è vera.
- b. La probabilità di non rifiutare erroneamente l'ipotesi nulla quando questa è vera.
- c. Un indicatore della plausibilità della relazione lineare tra due variabili.
- d. La probabilità di rifiutare giustamente l'ipotesi nulla quando questa è vera.

**32. Lo studio epidemiologico più adatto per valutare i fattori di rischio di malattie rare è:**

- a. studio a coorte
- b. studio sperimentale
- c. studio caso-controllo
- d. studio trasversale

**33. Un'associazione tra due variabili:**

- a. Se molto forte implica necessariamente una relazione causa-effetto tra le due variabili
- b. Anche se molto forte non comporta necessariamente che esista una relazione causa effetto tra le due variabili
- c. È sempre del tutto casuale
- d. Può esistere solo se le due variabili sono qualitative

**34. Si definisce statisticamente significativo:**

- a. Qualsiasi effetto osservato di qualsiasi entità purché clinicamente rilevante
- b. Qualsiasi effetto osservato di entità sufficientemente grande da poter essere definito clinicamente rilevante
- c. Un effetto che può essere attribuito al caso
- d. Qualsiasi effetto osservato di entità sufficientemente grande da non poter essere spiegato facilmente ricorrendo al caso

**35. La misura d'associazione comunemente utilizzata negli studi caso-controllo è:**

- a. Il rischio relativo (RR)
- b. La differenza assoluta tra rischi (ARR)
- c. L'odds ratio (OR)
- d. Il numero di pazienti da trattare per prevenire un evento (NNT)