

Answer the questions / fill in missing words.

The technology that directly uses pH changes to detect sequence is

- ☐ Oxford NanoPore
- ☒ Ion semiconductor sequencing ✓
- ☐ Illumina
- ☐ PacBio

Which of the technologies may produce different length reads?

IonTorrent ⇅ ✓

This technology uses emulsion PCR on beads:

- ☐ Sanger
- ☐ Illumina
- ☐ Oxford Nanopore
- ☐ PacBio
- ☒ IonTorrent ✓

You may have some issues with homopolymers (more frequently than in other technologies).

Ion torrent ⇅ ✓

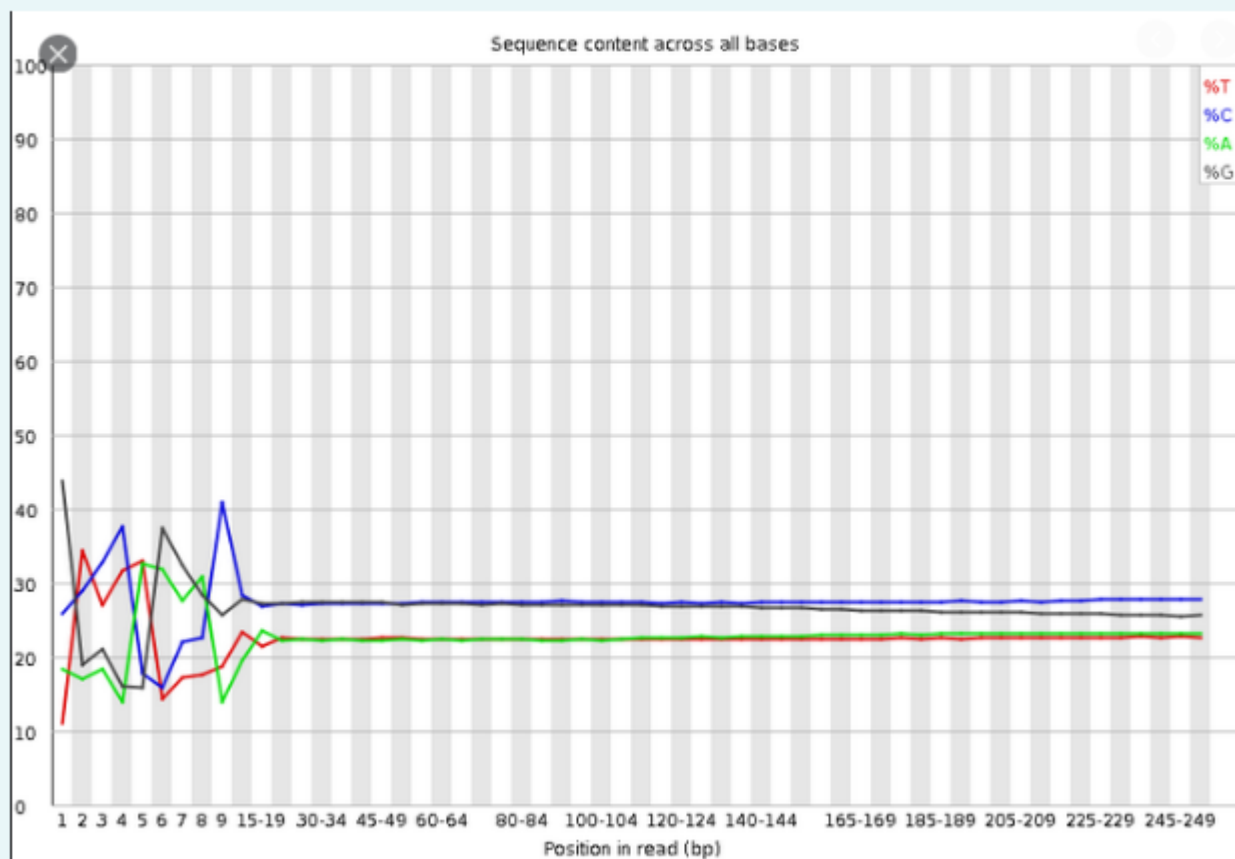
Which is the most useful (from the list) when detecting genomic rearrangements

Illumina ⇅ ✗

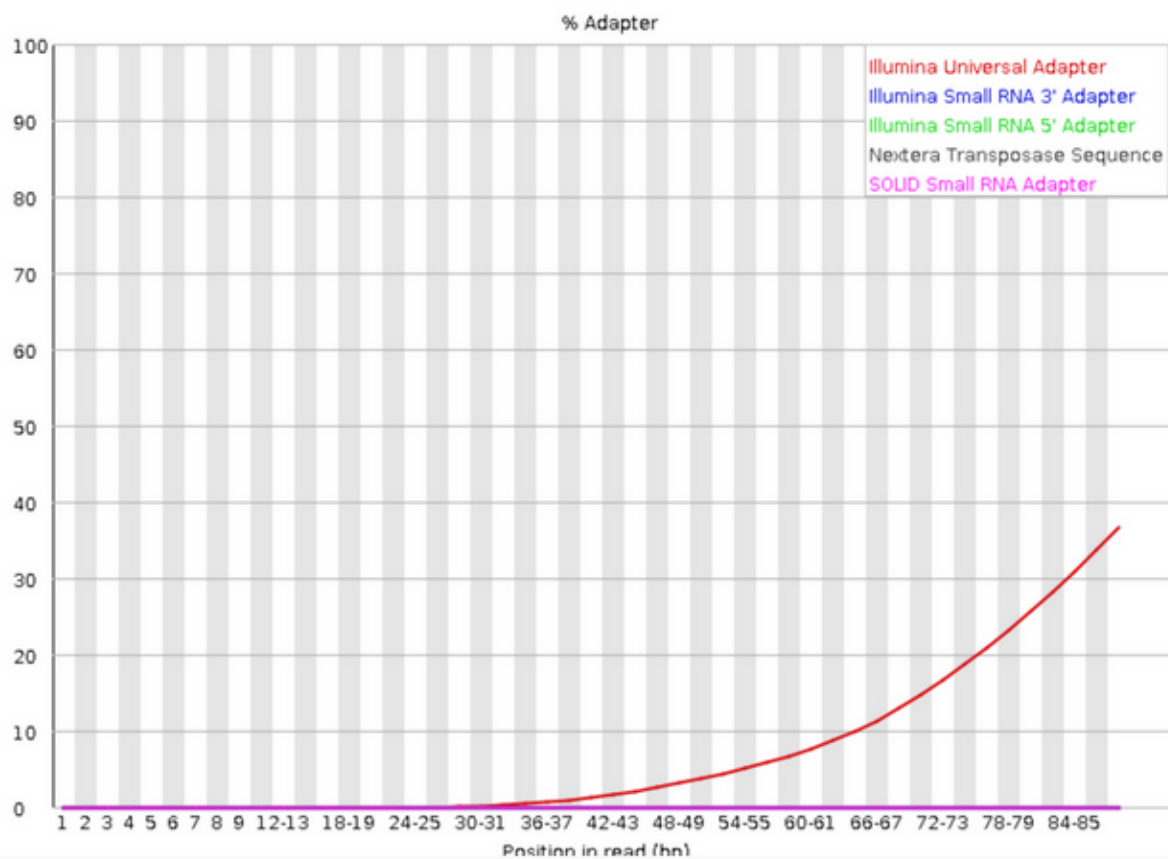
Most of the sequencing data comes in the format of

- ☒ FASTQ ✓
- ☐ FASTA
- ☐ GTF
- ☐ SAM
- ☐ VCF

You have two pictures (below). Which statement is correct?



Adapter Content



Your answer is incorrect.

Reads orientation is 5' -> 3'.

The correct answer is:

There is Illumina Universal Adapter at the 3' end of the sequence.

Your answer is incorrect.

Reads orientation is 5' -> 3'.

The correct answer is:

There is Illumina Universal Adapter at the 3' end of the sequence.

Your answer is incorrect.

Reads orientation is 5' -> 3'.

The correct answer is:

There is Illumina Universal Adapter at the 3' end of the sequence.

Your answer is incorrect.

Reads orientation is 5' -> 3'.

The correct answer is:

There is Illumina Universal Adapter at the 3' end of the sequence.

Which picture (pictures) shows FASTQ file format?

The correct answers are:

```
ACACATACGCACTGGGCATAAAGGGCGCGCAGGCGGTGAGGCGTCGGTGCTCAAAGTCCACCGCTTAACGGTGGGAGGCGTGI  
+HC9D00P01AN1VB  
FFFFFFFFFFFFG554A6911144442AAABDFFFIIIIIIIIIIHHHFFFFFFFA@CFDFFFDFDC??CFFFFFFFFI  
@HC9D00P01AWYAE rank=0000402 x=258.0 y=772.0 length=373  
ACACATACGCACTGGGCATAAAGGGCACGTAGGCGGATTGTAAAGTCAGGGGTGAATCCC GGCGCTCAACCTCGGAACTGCCT'  
+HC9D00P01AWYAE  
IIIIIIIIIIIIHHHII;666HHHIIIIIIIIIIICCCIIIIIFDC2//.<-//93.....-9?CCCCFECECCIIIIIDI  
@HC9D00P01A3C8R rank=0000675 x=331.0 y=1081.0 length=373  
ACACATACGCACTGGGTAAAGGGTGCGTAGGCGGGTCTTTAAGTCAGGGGTGAAATCCTGGAGCTCAACTCCAGA ACTGCCT'  
+HC9D00P01A3C8R  
IIIIIIIIIIII3../...-4AIIIECCE466GH974EEIAC@.0004.000>9@CEEIIIIIIIIIIIIIIIIHHI  
@HC9D00P01AW8TJ rank=0000926 x=261.5 y=2133.0 length=373  
ACACATACGCACTGGGTGTAAGGCGCACGTAGGCGGATTGCTAAGTCAGGGGTGAAATCCTGGAGCTCAACTCCAGA ACTGCCT'  
+HC9D00P01AW8TJ  
IIIIIIHHHIIIIIIHHHII;;IIIIIIIIIIIIIIIIIIIIIIII@000H4;;?IIIIIIIIIIIIIIIIIIIIIIHI  
@HC9D00P01AU1BY rank=0000952 x=230.0 y=2656.0 length=372  
ACACATACGCACTGGGCATAAAGAGCGCGTAGGCGGCCTTGTTAGTCGAGTGTGAAAGCCCTGGGCTTAACCCGGGAAGCGCGC'  
+HC9D00P01AU1BY  
IIIIIIIIIIIIHHHII;;IIIIIIIIIIIIIIIIIIIIIIIIIIII?666DHHHIHFEIIIIC;;555994?FIGI:  
@HC9D00P01AU1W rank=0000977 x=228.0 y=226.0 length=372
```

```
@Ocean.Water::80NGZ:04890:03499  
TACGGAGGGTGCAAGCGTTAATCGGAATTACTGGGCGTAAAGCGCGCGTAGGC  
GGCTCTTTAAGCCGGATGTGAAAGCCCCGGGCTCAACCTGGGA ACTGCATCCG  
ATACTGGAGGGCTAGAGTGTAGGAGAGGGAGGTAGAATTCATGTGTAGCGGT  
GGAATGCGTAGATATTATGAAGAACACCAGTGGCGGAAGGCCCTCTGGATC  
TGTA CTGACACTGAGGTGCGAAAGCGTGGGTAGCGAACAGGATTAGATACCCC  
TG TAGTCC  
  
+  
CCBA=AA@B9@;;2;AA@<@8BCE@C@F?BBCCC=ACBDC=D;  
9;5:9@@@;;29@AAA<B@CB?C@B@;@A::(4;44&4<3;;=D@C?  
BCC>C<DB=CCDCACCCCD C?AAF>CCCB@ @ @ @ BB?BA?  
<=&6;;0;ACA@C@B@B::CB?>?B>AA<?;999@BBBBBBC@BBAA=@???  
@A7AAAB@ABA@C?@ @ ;@::5:>:>@>>@?==>==>==::  
08=>::=0687252'25796=122639B=@@A>>=38;<?BBA;  
@9B::80NGZ:01200:07012  
ACGGAGGGTGCAAGCGTTGTCCGAATCACTGGGTGTAAAGGGTGTGTAGGCG  
GATATGTADGCTGAGGGGTGADAAGGCGACGGGCTDAGCGGTGADCTGCTTGA
```

The correct answers are:

```
ACACATACGCACTGGGCATAAAGGGCGCGCAGGCGGTGTCGAGGCGTCCAAAGTCCACCGCTTAACGGTGGGAGGCGTG  
+HC9D00P01AN1VB  
FFFFFFFFFFFFG554A6911144442AAABDFFFIIIIIIIIIHHHFFFFFFFA@CFDFDFDFDC??CFFFFFFFFI  
@HC9D00P01AWYAE rank=0000402 x=258.0 y=772.0 length=373  
ACACATACGCACTGGGCATAAAGGGCACGTAGGCGGATTGTAAGTCAGGGGTGAATCCCAGGGCGTCAACTCGGAACTGCCT'  
+HC9D00P01AWYAE  
IIIIIIIIIIIIHHHII;666HHHIIIIIIIIIIICCCIIIIFDC2//.<-//93.....-9?CCCCFECECCIIIIIDI  
@HC9D00P01A3C8R rank=0000675 x=331.0 y=1081.0 length=373  
ACACATACGCACTGGGTAAAGGGTGCGTAGGCGGGTCTTTAAGTCAGGGGTGAAATCCTGGAGCTCAACTCCAGAAGTGCCT'  
+HC9D00P01A3C8R  
IIIIIIIIIIII3../...-4AIIIECCE466GH974EEIAC@.0004.000>9@CEEIIIIIIIIIIIIIIIIHHI  
@HC9D00P01AW8TJ rank=0000926 x=261.5 y=2133.0 length=373  
ACACATACGCACTGGGTGTAAGGCGCACGTAGGCGGATTGCTAAGTCAGGGGTGAAATCCTGGAGCTCAACTCCAGAAGTGCCT'  
+HC9D00P01AW8TJ  
IIIIIIHHHIIIIIIHHHII;;IIIIIIIIIIIIIIIIIIIIIIII@000H4;;?IIIIIIIIIIIIIIIIIIIIHHI  
@HC9D00P01AU1BY rank=0000952 x=230.0 y=2656.0 length=372  
ACACATACGCACTGGGCATAAAGAGCGCGTAGGCGGCCTTGTTAGTCGAGTGTGAAAGCCCTGGGCTTAACCCGGGAAGCGCGC'  
+HC9D00P01AU1BY  
IIIIIIIIIIIIHHHII;;IIIIIIIIIIIIIIIIIIIIIIIIIIII?666DHHIHFEEIIIIIC;;555994?FIGI:  
@HC9D00P01AU1W rank=0000977 x=228.0 y=226.0 length=372
```

```
@Ocean.Water::80NGZ:04890:03499  
TACGGAGGGTGCAAGCGTTAATCGGAATTACTGGGCGTAAAGCGCGCGTAGGC  
GGCTCTTTAAGCCGGATGTGAAAGCCCCGGGCTCAACTGGGAAGTGCATCCG  
ATACTGGAGGGCTAGAGTGTAGGAGAGGGAGGTAGAATTCATGTGTAGCGGT  
GGAATGCGTAGATATTATGAAGAACACCAGTGGCGGAAGGCCCTCTGGATC  
TGTAAGTACACTGAGGTGCGAAAGCGTGGGTAGCGAACAGGATTAGATACCCC  
TGTAAGTCC  
+  
CCBA=AA@B9@;;2;AA@<@8BCE@C@F?BBCCC=ACBDC=D;  
9;5:9@@@;;29@AAA<B@CB?C@B@;@A::(4;44&4<3;;=D@C?  
BCC>C<DB=CCDCACCCCDCC?AAF>CCCB@&&&BB?BA?  
<=>;0;ACA@C@B@B::CB?>?B>AA<?;999@BBBBBBC@BBAA=@???  
@A7AAAB@ABA@C?@&@;@::5:>:>@>>>@?==>==>==::  
08=>::=0687252'25796=122639B=@@A>>=38;<?BBA;  
@9B::80NGZ:01200:07012  
ACGGAGGGTGCAAGCGTTGTCCGAATCACTGGGTGTAAAGGGTGTGTAGGCG  
GATATGTADGCTAGGCGGTGADAAGGCGAGGCGCTAGGCGGTGAGTGCCTTGA
```

The correct answers are:

```
ACACATACGCACTGGGCGTAAAGGGCGCGCAGGCGGTGAGGCGGTGGTGCTCAAAGTCCACCGCTTAACGGTGGGAGGCGTGI
+HC9D00P01AN1VB
FFFFFFFFFFFFG0554A6911144442AAABDFFFIIIIIIIIIIHHHFFFFFFFA@CFDFFDFDC??CFFFFFFFFI
@HC9D00P01AWYAE rank=0000402 x=258.0 y=772.0 length=373
ACACATACGCACTGGGCATAAAGGGCAGCTAGGCGGATTGTAAAGTCAGGGGTGAATCCCGGGCGTCAACCTCGGAACTGCCT'
+HC9D00P01AWYAE
IIIIIIIIIIIIHHHII;666HHHIIIIIIIIIIICCCIIIIIFDC2//.<-//93.....-9?CCCCFECCIIIIIDI
@HC9D00P01A3C8R rank=0000675 x=331.0 y=1081.0 length=373
ACACATACGCACTGGGTAAAGGGTGCGTAGGCGGGTCTTTAAGTCAGGGGTGAAATCCTGGAGCTCAACTCCAGAAGTGCCTT'
+HC9D00P01A3C8R
IIIIIIIIIIII3../...-4AIIIECCE466GH974EEIAC@.0004.000>9@CEEEIIIIIIIIIIIIIIIIHHI
@HC9D00P01AW8TJ rank=0000926 x=261.5 y=2133.0 length=373
ACACATACGCACTGGGTGTAAGGCGCAGCTAGGCGGATTGCTAAGTCAGGGGTGAAATCCTGGAGCTCAACTCCAGAAGTGCCT'
+HC9D00P01AW8TJ
IIIIHHHIIIIIIHHHII;;IIIIIIIIIIIIIIIIIIIIIIII@000H4;;?IIIIIIIIIIIIIIIIIIIIIIH
@HC9D00P01AU18Y rank=0000952 x=230.0 y=2656.0 length=372
ACACATACGCACTGGGCATAAAGAGCGCGTAGGCGGCCTTGTTAGTCGAGTGTGAAAGCCCTGGGCTTAACCCGGGAAGCGCGC'
+HC9D00P01AU18Y
IIIIIIIIIIIIHHHII;;IIIIIIIIIIIIIIIIIIIIIIIIIIII?666DHHIHFEEIIIIIC;;555994?FIGI
@HC9D00P01AU1W rank=0000977 x=228.0 y=226.0 length=372

@Ocean.Water::08NGZ:04890:03499
TACGGAGGGTGCAAGCGTTAATCGGAATTACTGGGCGTAAAGCGCGCGTAGGC
GGCTCTTTAAGCCGGATGTGAAAGCCCCGGGCTCAACCTGGGAAGTGCATCCG
ATACTGGAGGGCTAGAGTGTAGGAGAGGGAGGTAGAATTCATGTGTAGCGGT
GGAATGCGTAGATATTATGAAGAACACAGTGCGCGAAGGCGGCCTCTGGATC
TGTAAGTACTGACACTGAGGTGCGAAAGCGTGGGTAGCGAACAGGATTAGATACCC
TGTAAGTCC
+
CCBA=AA@B9@;;2;AA@<@8BCE@C@F?BBCCC=ACBDC=D;
9;5:9@e@;;:29@AAA<B@CB?C@B@;@A::(4;44&4<3;;=D@C?
BCC>C<DB=CCDCACCCCDCC?AAF>CCCB@e@e@BB?BA?
<=6;;0;ACA@C@B@B::CB?>?B>AA<?;999@ABBBBBC@BBAA=@???
@A7AAAB@ABA@C?@e@;@::5>:>@>>>@?=>=>=>=>==::
08=>::=0687252'25796=122639B=@@A>=>=38;<?BBA;
@9B::08NGZ:01200:07012
ACGGAGGGTGCAAGCGTTGTCCGAATCACTGGGTGTAAAGGTGTGTAGGCG
GATATGTATAGCTAGGCGGTGAAAGGCGCGGCGCTTAAAGGCGGTGAGTGCCTTGA
```

Kuris teiginys apie techninių replikų koreliacijas yra teisingas?

- ☒ a. Techninių replikų koreliacija turėtų būti artima 1 (0.9-1) ✓
- ☐ b. Techninių replikų koreliacija turėtų būti artima -1.
- ☐ c. Techninių replikų tarpusavio koreliacija turėtų būti labai aukšta ir žemesnė negu (arba lygi) biologinių replikų tarpusavio koreliacija.
- ☐ d. Techninių replikų koreliacija turėtų būti artima 0.
- ☐ e. Techninių replikų tarpusavio koreliacija turėtų būti gana žema ir žemesnė negu (arba lygi) biologinių replikų tarpusavio koreliacija.
- ☐ f. Biologinių replikų tarpusavio koreliacija turėtų būti aukštesnė negu techninių replikų tarpusavio koreliacija

Your answer is correct.

The correct answer is:

Techninių replikų koreliacija turėtų būti artima 1 (0.9-1)

Suporuokite, kuri programa kam skirta.

Raiškos skaičiavimas	deeptools	✗
Kartografavimas (angl. map)	samtools	✗
FASTQ failų tvarkymas	cutadapt	✓

Your answer is partially correct.

You have correctly selected 1.

The correct answer is:

Raiškos skaičiavimas → featureCounts,

Kartografavimas (angl. map) → STAR,

FASTQ failų tvarkymas → cutadapt

Kuris teiginys apie biologines replikas teisingas?

- ☐ a. Biologinės replikos leidžia įvertinti, kokios variacijos yra sutinkamos/gaunamos dėl techninių priežasčių.
- ☐ b. Biologinės replikos neleidžia įvertinti jokios variacijos ir yra naudojamos tik tam, kad padidinti duomenų kiekį.
- ☒ c. Biologinės replikos leidžia įvertinti, kokios variacijos yra sutinkamos natūraliai, t.y. dėl biologinių priežasčių. ✓
- ☐ d. Biologinės replikos nėra reikalingos ir jų galima nedaryti.

Your answer is correct.

The correct answer is:

Biologinės replikos leidžia įvertinti, kokios variacijos yra sutinkamos natūraliai, t.y. dėl biologinių priežasčių.

Kodėl reikia normalizuoti raw-counts duomenis?

Pažymėkite visus teisingus atsakymus

- ☐ a. Dėl to, kad skirtinguose organizmuose genai yra skirtingo ilgio ir nenormalizavus neįmanoma jų sulyginti.
- ☒ b. Kai kurie genai gali turėti daugiau egzonų ir dėl to gauti daugiau nuskaitymų. ✗ Pats egzonų skaičius nėra svarbu. Svarbu ilgis.
- ☒ c. Kadangi skiriasi geno ilgis, ilgesni transkriptai gali būti dėl ilgio turėti daugiau nuskaitymų. ✓
- ☐ d. Priežasčių nėra - tai standartinė procedūra, kurią galima praleisti.
- ☒ e. Patogiau ir lengviau lyginti tada, kai visų genų raiškos varijuoja aplink tuos pačius įverčius. ✗
- ☒ f. Skirtingo dydžio bibliotekos nulems tai, kad genų raiška didesniame mėginyje atrodys didesnė nei iš tiesų yra. ✓
- ☒ g. Gali būti situacijų, kai kažkokie labai gausūs transkriptai sunaudoja labai daug nuskaitymų ir dėl to likusių genų raiška atrodo sumažėjusi. ✓

Your answer is partially correct.

You have selected too many options.

Match abbreviations and their meaning.

Transcripts Per Kilobase Million

TPM ✓

Reads Per Kilobase of transcript, per Million mapped reads

RPKM ✓

Fragments Per Kilobase of transcript, per Million mapped reads

FPKM ✓

Your answer is correct.

Which of the following is a key factor affecting the quality of genome assembly?

- ☒ a. The size of the genome being assembled. ✓
- ☐ b. The type of DNA sequencing technology used.
- ☐ c. The weather conditions during the time of DNA extraction and sequencing.
- ☐ d. The number of reads generated during sequencing.

Your answer is partially correct.

You have correctly selected 1.

The correct answers are:

The size of the genome being assembled.,

The type of DNA sequencing technology used.,

The number of reads generated during sequencing.

Which of the following best describes a contig in the context of genome assembly?

- ☐ a. A small fragment of DNA that is used as a template for PCR amplification.
- ☐ b. A complete set of chromosomes from a single organism that has been sequenced.
- ☒ c. A contiguous sequence of overlapping DNA fragments that are assembled into a longer sequence. ✓
- ☐ d. A specific location on a chromosome that has been targeted for genetic engineering.

Your answer is correct.

The correct answer is: A contiguous sequence of overlapping DNA fragments that are assembled into a longer sequence.

Which of the following is true regarding the De Bruijn graph in genome assembly?

- ☐ a. It is a graph-based approach that is only used in hybrid genome assembly.
- ☒ b. It is a data structure that represents overlaps between short reads in a genome. ✓
- ☐ c. It is a tool for mapping reads to a reference genome.
- ☐ d. It is a method for ordering and orienting contigs based on mate-pair information.

Your answer is correct.

The correct answer is:

It is a data structure that represents overlaps between short reads in a genome.

Which of the following best describes a scaffold in the context of genome assembly?

- ☒ a. A set of contigs that have been ordered and oriented based on paired-end reads or mate-pair libraries. ✓
- ☐ b. A contiguous sequence of overlapping DNA fragments that are assembled into a longer sequence.
- ☐ c. A complete set of chromosomes from a single organism that has been sequenced and annotated.
- ☐ d. A type of RNA molecule that carries genetic information from the DNA to the ribosome.

Your answer is correct.

The correct answer is:

A set of contigs that have been ordered and oriented based on paired-end reads or mate-pair libraries.

~/.DB/get_fails yra sąrašas, raudas, tai rodo, kur yra failai. Kokiu formatu (bloku informacija) yra pateikta duomenų bazėje?

FASTA



. Tai yra standartinis formatas, kuriuo yra platinamos sekos. Dabar su komandine eilute nueikite į ~/DB.

Tam, kad galėtumėte atlikti BLAST užklausas savo kompiutere, reikia duomenų bazę paruošti specialiu būdu, t.y. paleisti specialią komandą. ši komanda yra makeblastdb. Kokiu būdu šiai komandai reikės pateikti mūsų input failo vardą. Kokia opcija bus naudojama tam, kad nurodyti input failo vardą? (pasiziūrėkite makeblastdb -help).

-in



Taip pat šiai komandai reikia nurodyti, kokio tipo duomenų bazę turite - nukleotidų ar baltymų. Kokio tipo duomenų bazę yra pas jus?

Baltymų



. Šį argumentą mes turėsime pateikti opcijai -dbtype. Deja, bet programai negalime nurodyti lietuviško pavadinimo, tad kokį argumentą naudosite -dbtype opcijai?

prot



(vėlgi, pasižiūrėkite help'ą). Mes taip pat dar panaudosime opciją -parse_seqids. Tiesiog parašykite šią opciją be argumentų.

Taigi, kaip atrodys jūsų komanda? Laikykitės, kad duomenų bazės kūrimą atliekate darbinėje direktorijoje, o opcijos surašomos tokia eilės tvarka: input failas, duomenų bazės tipas ir opcijas -parse_seqids.

makeblastdb -in TAIR10_pep_20101214_updated.fa -dbtype prot -parse (įrašykite komandą).

Komanda neturėtų veikti ilgai - turėtumėte gauti keletą eilučių teksto. Kiek viso buvo sekų mūsų faile?

35386



Paieškos duomenų bazėje atliekamos naudojantis specialiomis komandomis. Komandų pavadinimai atitinka internetinio blast pavadinimus (su jais susipažinome per paskaitą). Konkrečią blast versiją reikia pasirinkti pačiam pagal tai, kokio tipo užklausą ir kokio tipo duomenų bazę turite. Šioje užduotyje naudosime seką **B1Q3J6**. Kokios duomenų bazės ID jums čia pateiktas?

UniProt



. O kaip naudodamiesi efetch įrankiu atsisiųsite šią seką FASTA formatu?

efetch -db protein -id B1Q3J6 -format fasta > ~/Desktop/B1Q3J6.fasta



. Praeito klausimo atsakymas

yra komanda, kuri seką parašo į ekraną - čia ne visai tai, ko norime.. pasimodifikuokite komandą taip, kad ji savo rezultatus išsaugotų faile B1Q3J6.fasta ant darbastalio (galite pasidėti ir į logiškesnę direktoriją, t.y. pvz. konkrečios paskaitos direktoriją).

Dabar mes turime du daiktus - turime duomenų bazę, kurioje ieškosime ir turime seką, kurios homologų ieškosime. Galime atlikti paiešką. Remdamiesi duomenų bazės ir užklausos sekos (ta seka, kurią atsisiuntėte ką tik yra vadinama užklausos seka, query) nurodykite, kokią blast versiją naudosite?

blastp



(pvz. blastn, blastp, tblastn ir t.t.).

Galime formuoti blast užklausą/komandą. Komanda turi daugybę opcijų (visas galite pasižiūrėti su -h opcija), bet kelios opcijos yra esminės, t.y. būtinos. Konkrečiai mums būtinai reikės nurodyti duomenų bazę, kurioje ieškosime ir seką, kurios ieškosime. Tai yra nurodoma per -db ir -query opcijas. Suformuokite blastp užklausą ir ją įvykdykite (kaip rašyta - reikia tik dviejų opcijų ir jų argumentų) ir rezultatus išsaugokite. Toliau jums bus pateikta keletas klausimų apie rezultatus - pabandykite išsiaiškinti ir rasti atsakymus rezultatų faile:

Klausimas 7

Baigta

Balas 30.00 iš 30.00

Pažymėti klausimą

Paaiškinkite, kas yra nespecifinis filtravimas (non-specific filtering): ką jis daro ir kada gali padėti.

Kai kuriais atvejais turime daug duomenų lygių, tarp kurių yra mažos raiškos duomenų, todėl nespecifinis filtravimas leidžia atsikratyti zondu, kurie, mes esame tikri, neduos rezultatų (kurių raiška yra ganėtinai maža), arba nėra viarabilumo (t.y. nors ir didelė raiška, bet tas pats rezultatas).

Komentaras:

Iš tiesų galima ir pagal daugiau kriterijų filtruoti - mes jūsų įvardintus turėjome per paskaitą, bet tai tiesiog keli pavyzdžiai.

Klausimas 6

Baigta

Balas 30.00 iš 30.00

Pažymėti klausimą

Paaiškinkite, kuo skiriasi kDNR (cDNA) gardelės nuo oligonukleotidinių gardelių (kartu paaiškinkite ir abudu terminus, kDNR ir oligonukleotidai).

Oligonukleotidas – trumpas (apie 20 nt) RNR ar DNR fragmentas, naudojamas komplementariai RNR ar DNR aptikti, o cDNA yra dvigrandė DNR susintetinta iš viengrandinės RNR (pavyzdžiui: iRNR), kurios šablonas katalizuoja fermento atvirkštinės transkriptazės.

cDNA mikrogardelė, dviejų kanalų gardelė, zondai gauti iš kDNR yra 500~5000 bazių, robotas juos sudėlioja, vadinami „DNA microarray“.

Pagrindinis bruožas oligonukleotidų mikrogardelės yra tai, kad kiekvienas genas paprastai atstovauja daugiau nei vieną zondu: skirtingų genų regionai kurie yra zondu žemėlapyje paprastai vadinami zondu rinkiniu.

Komentaras:

cDNA nėra dviejų spalvų "platforma". Tai iš viso nėra platforma. Spalvų kiekis priklauso nuo mėginių (o ne čipo) paruošimo.

Klausimas 5

Baigta

Balas 5.00 iš 30.00

Pažymėti klausimą

Paprastai paaiškinkite, kaip veikia mikrogardelės ir kodėl jų pagalba galime matuoti genų raišką. Įsivaizduokite, kad aiškinate žmogui, kuris nėra susijęs su bioinformatika/informatika. Paaiškinimas turėtų užimti bent 5-6 sakinius.

Mikrogardelė, leidžia detektuoti tūkstančių genų raišką vienu metu, tai didelio našumo technologija, jos gali būti DNR, Baltymų, Audinių mikrogardelės. Su mikrogardelėmis galima analizuoti daug genų iš karto, jos nedidelio dydžio (minimalizmas), daug pavyzdžių iš karto, reagentai ir čipai gaminami automatizuotai. Mikrogardelės apjungia tokius dalykus, kaip aiškinimąsi, kurių genų raiška skiriasi bėgant laikui, tarp audinių arba tarp ligos stadijų. Apjungiama taip pat ir medžiagų toksiškumas, genetinių ligų požymiai, polimorfizmai.

Komentaras:

Neatsakyta į klausimą - tiesiog bendrai aprašytos gardelės.

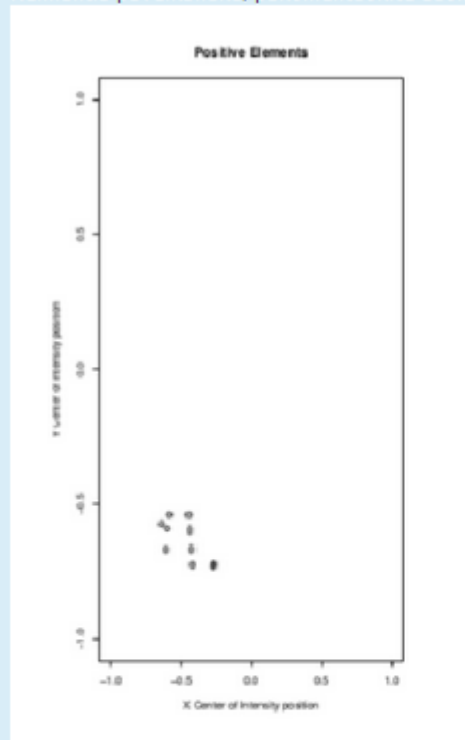
Klausimas 4

Baigta

Balas 25.00 iš 30.00

Pažymėti klausimą

Jums pateiktas paveikslukas iš AffyQCRReport (COI diagrama). Pateikta diagrama tik "positive" elementams. Remiantis paveiksluku, pakomentuokite duomenų kokybę.



Hibridizacija nėra tolygi, nes kontrolės centras nėra piešinuko viduryje. Todėl tai gali būti hibridizacijos problemos.

Komentaras:

Nėra komentaro apie pačią duomenų kokybę (netolygi hibridizacija nėra tas pats kas duomenų kokybė)

Klausimas 3

Baigta

Balas 30.00 iš 30.00

Pažymėti klausimą

Paašškinkite (3-4 sakiniais), kodėl atliekamas normalizavimas ir kokias problemas jis sutvarko.

Normalizavimas atliekant norint ištaisyti fono korekcijas, skalės korekcijas, agregacija, jo pagrindinis tikslas yra pašalinti nebiologinės kilmės elementus, signalus arba dalis. Jei nebus atliekamas normalizavimas, nelabai bus galima dirbti su duomenimis, nes jie bus "nešvarūs".

Komentaras:

O kam jūs dar norite taisyti korekcijas? ("ištaisyti fono korekcijas"),

Ir šiuo atveju nėra tinkamas terminas "nešvarūs" duomenys - tai daugiau sekoskaitoje naudojamas terminas.

Klausimas 2

Baigta

Balas 30.00 iš 30.00

Pažymėti klausimą

Kuris apibrėžimas **tiksliausiai** apibūdina MM tipo zondus?

Pasirinkite:

- ☒ a. Tai zondas, kuriame seka nuo analizuojamo geno skiriasi vienu nukleotidu
- ☐ b. Tai zondas, kuriame seka vidutiniškai atitinka analizuojamą geną.
- ☐ c. Tai zondas, kuriame seka visiškai neatitinka analizuojamo geno.
- ☐ d. Tai zondas, kuriame seka idealiai atitinka analizuojamą geną.

Your answer is correct.

Klausimas 1

Baigta

Balas 20.00 iš 30.00

Pažymėti klausimą

Kurso metu susipažinote su dviem būdais, kaip galima matuoti genų raišką. Kokie šių metodų privalumai/trūkumai vienas kito atžvilgiu? Įvardinkite bent du dalykus, kur vienas metodas yra geresnis negu kitas.

Sekvenavimas kooperuojasi į genomo surinkimą, tam, kad nustatyti, kai kurias genomo variacijas, problema yra ta, kad mes turime neapibrėžtą readų ilgį jis gali varijuoti, taip pat gali pasitaikyti sisteminių klaidų, visokios biologinės priežastys, pavyzdžiui, diploidija, taip pat mes turime tik 4 nukleotidus, todėl kartotinių fragmentų kiekis yra milžiniškas. Mikrogardelėse taip pat gali atsitiesti sisteminių klaidų, neužkaista gerai aparatas, ar atsiranda tam tikros problemos su gardelės kraštais, bet pagrindinis mikrogardelių tikslas - detektuoti tūkstančių jau žinomų genų raišką vienu metu.

Komentaras:

Read'o ilgis priklauso nuo platformos ir dažnai galime naudoti gana fiksuotą ilgį.

Dėl sekvenavimo ir genomo surinkimo - tai tik vienas iš pritaikymų.

Klausimas 8

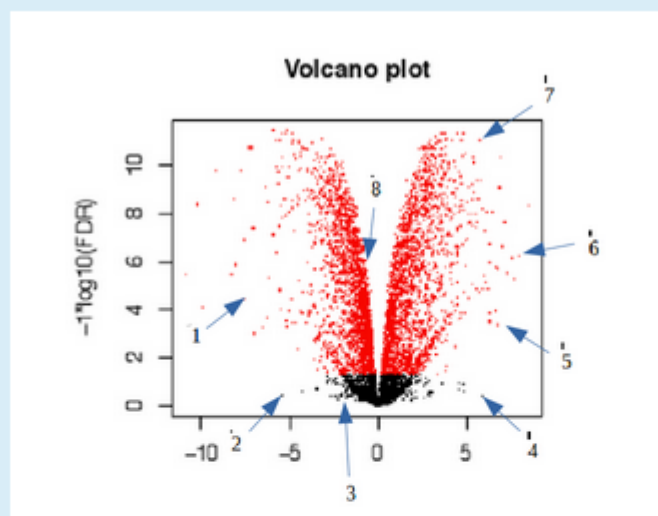
Baigta

Balais 40.00 iš 40.00

Pažymuo klausimą

Jums patelkita volcano plot. Aprašykite:

1. ką bendrai atvaizduoja volcano plot
2. kokie taškai pažymėti numerukais 1-8 (geri/negeri; tikėtina, kad statistiškai patikimi ir su dideliais reiškos pokyčiais ar ne).



1. Kai mes atliekame statistinius eksperimentus, norime kad genų reiška būtų pakitusi, kontrolė maža, didelio reiškos pokyčio, kuo mažesnis statistinio [verčio (kad jis būtų kuo geresnis), volcano plot apjungia abu dalykus, reiškos pokytis ir statistinį patikimumą (mažesnis statistinis [vertis]), tam tikras taškas yra zondas, mūsų duomenys pasiskirsto [vieną arba [kitą pusę, kuo pokytis yra didesnis, tuo jis nukeliauja [vieną arba [kitą pusę, centre grupuojasi kuo mažesnis pokytis. Mus domina tik tai kas yra šonuose, ten yra patys geriausi taškai ir mus domina geras pvalue, kuo geresnis pvalue tuo taškai yra aukščiau, taškai žemiau jie nėra labai patikimi, o tie viršuje patikimi turi gerą pvalue. Apjungus šiuos du kriterijus geriausi duomenys yra tie kurie yra nutolę šonuose ir yra nutolę [viršus, o kas yra centre jų reiška nelabai keltėsi.

2.

1- genėtinai geras taškas, didelė reiška, statistiškai beveik patikimas, yra ir geresnių taškų.

2- biogas taškas, statistiškai nepatikimas, maža reiška

3- biogas taškas, statistiškai nepatikimas, maža reiška

4- biogas taškas, statistiškai nepatikimas, maža reiška

5- genėtinai biogas taškas, nėra statistiškai patikimas, nors reiška gera.

6- statistiškai patikimas, geras taškas, didelė reiška

7- statistiškai patikimas, geras taškas, didelė reiška

8- statistiškai nepatikimas, reiška maža, biogas taškas.

Komentaras: