lab3_Biology

Batuhan Seyhan S20343

TASK 1

list 3 methods used in genealogy 1-Traditional genealogical methods 2-Genealogical DNA testing 3-Molecular evolutionary studies.

TASK 2

HVR2, HVR1, and control region (CR)

TASK 3

There are 6 markers found in the HVR2 region: 73G, 182T, 185A, 228A, 263G, 295T (315.1C), 462T, and 489C.

There are 19 markers found in the CR region: 750G, 2706G, 3010A, 4216C, 4769G, 6554T, 7028T, 8860G, 10398G, 11251G, 11719A, 12127A, 12612G, 13708A, 14452G, 14766T, 14798C, 15326G.

There are 4 markers found in the HVR1 region: 16069T, 16092C, 16126C, and 16261T.

TASK 4

A genetic marker is a DNA sequence that varies among individuals and can be used to identify unique genetic characteristics. These variations can be inherited and passed down through generations, making them useful in genetic studies, such as determining relationships among individuals, mapping the location of genes, and identifying genetic mutations associated with diseases. Examples of genetic markers include single nucleotide polymorphisms (SNPs), insertions, deletions, and tandem repeats. By analyzing genetic markers, researchers can gain insights into the genetic makeup and history of individuals and populations.

TASK 5

1. J1c7
2. J1c7a
3. J1c(C16261T)

TASK 6

rCRS stands for revised Cambridge Reference Sequence. It is a reference sequence of the human mitochondrial DNA (mtDNA) that was established in the early 1980s at the University of Cam-

bridge. The rCRS sequence represents a consensus sequence of the human mtDNA and is widely used as a reference for mtDNA analysis. It is used as a standard for comparison to determine sequence variations and to identify haplogroups. The rCRS sequence contains 16,569 base pairs and was revised in 1999 to improve accuracy and completeness.

TASK 7

H2a2a1

TASK 8

EU151466

TASK 9

The mutation that occurred between positions 12120 to 12130 is a transition from "G" to "A" at position 12126. This mutation is also known as the mtDNA mutation 12126G>A.

TASK 10

J1c7

TASK 11

vcf,fasta,hsd

TASK 12

50

TASK 13

""SampleID" "Haplogroup" "Rank" "Quality" "Range" "HG00140" "H4a1a4b" "1" "0.9660" "1-16569""HG00178" "H13a1a1d1" "1" "1.0000" "1-16569" "HG00356" "T2b4+152" "1" "0.9614" "1-16569""HG00365" "H1c" "1" "1.0000" "1-16569" "HG00513" "R11b1a" "1" "0.9294" "1-16569" "HG00599""B4a1c4" "1" "0.9890" "1-16569" "HG00629" "D4g2a1" "1" "0.9368" "1-16569" "HG01119" "A2w" "1""0.9786" "1-16569" "HG01284" "A2w" "1" "0.9850" "1-16569" "HG01372" "A2q" "1" "1.0000" "1-16569" "HG01597" "F1+16189" "1" "0.8444" "1-16569" "HG01630" "H7" "1" "0.9531" "1-16569""HG01631" "H1q" "1" "0.8298" "1-16569" "HG01844" "D4a7" "1" "0.9822" "1-16569" "HG01866""B4b1c1" "1" "1.0000" "1-16569" "HG01871" "M7b1a1b" "1" "0.9759" "1-16569" "HG02008" "D1" "1""0.9795" "1-16569" "HG02275" "C" "1" "0.9435" "1-16569" "HG02508" "L2a1c4a1" "1" "0.9505" "1- 16569" "HG02775" "U7b" "1" "0.9377" "1-16569" "HG02808" "L2a1c" "1" "0.9511" "1-16569""HG03160" "L3e2b8" "1" "0.9816" "1-16569" "HG03352" "L3e2b1a1" "1" "0.9961" "1-16569" "HG03432""L2a1i" "1" "0.9937" "1-16569" "HG03461" "L2a1i" "1" "0.9594" "1-16569" "HG03520" "L3e1" "1""0.9654" "1-16569" "HG03611" "M68" "1" "0.7805" "1-16569" "HG03698" "N1a2" "1" "0.9647" "1-16569" "HG03771" "M40a1" "1" "0.9636" "1-16569" "HG03817" "U2c1a" "1" "0.8820" "1-16569""HG04001"

“M18a” “1” “0.9865” “1-16569” “HG04006” “M5a” “1” “0.9247” “1-16569” “HG04026”“W3a1b” “1” “0.9606” “1-16569” “NA12282” “J1b1a1a” “1” “1.0000” “1-16569” “NA12815” “H27” “1”“0.9658” “1-16569” “NA12874” “J1c8a” “1” “0.9842” “1-16569” “NA18561” “F1d” “1” “0.8996” “1-16569” “NA18648” “D4a” “1” “0.9401” “1-16569” “NA19210” “L1b1a3” “1” “0.9890” “1-16569”“NA19225” “L2a1+143+16189+(16192)” “1” “0.9648” “1-16569” “NA19315” “L3e3b1” “1” “0.9844” “1-16569” “NA19462” “L1b1a15” “1” “0.9911” “1-16569” “NA19712” “L1c2b1c” “1” “0.9687” “1-16569”“NA19747” “D1i1” “1” “0.9796” “1-16569” “NA19780” “B2s” “1” “0.9918” “1-16569” “NA20530”“K1a4c11” “1” “0.9627” “1-16569” “NA20797” “V2” “1” “0.9797” “1-16569” “NA20827” “U5a1g” “1”“0.9607” “1-16569” “NA20870” “H7b” “1” “0.9421” “1-16569” “NA21097” “U1a1c1d” “1” “0.9867” “1-16569”“‘

TASK 14

https://ibb.co/XkVkjMW

TASK 15

DYS456 is a genetic marker used in genetic genealogy and forensic DNA analysis. It is part of the Y-chromosome DNA (Y-DNA) testing panel, specifically the short tandem repeat (STR) markers used to identify male-specific DNA.

DYS456 is an STR located on the Y-chromosome and consists of a repeating sequence of nucleotides. The number of repeats in the sequence can vary from person to person, making it useful for identifying individuals and tracing paternal lineages.

In genetic genealogy, DYS456 is one of the markers commonly used in Y-DNA testing to identify genetic matches and establish relationships between individuals who share a common paternal ancestor. In forensic DNA analysis, DYS456 is used in conjunction with other Y-STR markers to create a DNA profile that can be used to identify suspects or match DNA evidence to individuals.

TASK 16

DYS456 is a short tandem repeat (STR) marker, meaning it consists of a repeating sequence of nucleotides. The specific sequence repeated in DYS456 is AGAGG.

TASK 17

The number of repeats in the DYS456 sequence can vary between individuals. The range of the number of repeats for DYS456 is typically 10 to 24, which means there are at least 15 known alleles of this marker.

TASK 18

Y-DNA haplogroup prediction, mtDNA haplogroup prediction, Y-STR haplotype analysis, Autosomal DNA ancestry analysis.

TASK 19

The yhrd.org database is a public repository of Y-chromosome short tandem repeat (Y-STR) data. It is a collaborative project that aims to collect, curate, and share Y-STR data from populations around the world. The database provides a valuable resource for genetic research, forensic investigations, and other applications that require Y-STR data.

TASK 20

According to the Database statistics on the yhrd.org website, the following datasets are published in the database:

Minimal Y12 Y17 Y23 Y27 Ymax Each dataset includes a different set of Y-STR markers and has different numbers of haplotypes, population samples, national databases, and metapopulations

TASK 21

6, 9, 10, 11, 12, 13, 13.2, 13.3, 14, 14.1, 14.2, 14.3, 15, 15.2, 16, 16.2, 17, 18, 19, 19.1, 20

TASK 22

12&2000

TASK 23

https://ibb.co/MVwRX9d

TASK 24

kit “powerflex y” is deteced

TASK 25

https://ibb.co/rvJHMx5

TASK 26

sequence_a = input(“Enter the DNA sequence 1:”) sequence_b = input(“Enter the DNA sequence 2:”)

if len(sequence_a) != len(sequence_b): print(“Sequences of different lengths”) else: # Create an empty list to store the differences differences = []

```
# Iterate over the sequences and compare each nucleotide
for i in range(len(sequence_a)):
    if sequence_a[i] != sequence_b[i]:
        differences.append((i+1, sequence_a[i], sequence_b[i]))

if len(differences) == 0:
```

```
    print("Sequences are identical")
else:
    print(differences)
```

TASK 27

def complement(nucleotide): if nucleotide == ‘A’: return ‘T’ elif nucleotide == ‘T’: return ‘A’ elif nucleotide == ‘C’: return ‘G’ elif nucleotide == ‘G’: return ‘C’ else: return None

sequence = input(“Enter the DNA sequence:”)

max_pairs = 0

Iterate over possible hairpin lengths

for i in range(1, len(sequence)//2 + 1): # Check if the hairpin is complementary on both sides pairs = 0 for j in range(i): if sequence[j] == complement(sequence[-(i-j)]) and (i-j) > pairs: pairs = i-j for k in range(1, pairs): if sequence[j+k] != complement(sequence[-(i-j-k)]): break else: continue break max_pairs = max(max_pairs, pairs // 2)

print(max_pairs)