732A51 Bioinformatics Lab 1

Raymond Sseguya, Martin Smelik, Duc Duong

2018 M11 7

Task 1

Task 1.1

diploid population = Aa,Aa = 2N (parents)

After random mating for first generation: children = AA, Aa, aA, aa

Total new first gneration population = 4

Number of AA homozygtes = 1

Number of aa homozygotes = 1

Number of Aa or aA hereterozygotes = 2

Proportion of AA homozygtes = 1/4 = 0.25

Proportion of aa homozygotes = 1/4 = 0.25

Proportion of Aa or aA hereterozygotes = 2/4 = 0.5

The proportion AA Homozygotes : Aa Heterozygotes : aa homozygotes is

0.25: 0.5 : 0.25

Again after random mating for first generation:

Probability of getting A allele = 3/6 = 0.5 = p

Probability of getting a allele = 3/6 = 0.5 = q

The proportion AA Homozygotes : Aa Heterozygotes : aa homozygotes is also square of p : 2 times p times q : square of q

(0.5)^2 : 2 times 0.5 times 0.5 : (0.5)^2

0.25: 0.5 : 0.25

This satisfies the Hardy Weinberg equilibrium.

Task 1.1.b

The probabibility of getting A allele and a allele will remain the same with continued random mating therefore the Hardy Weinberg equilibrium will always hold.

Task 3

3.1

According to Wikipedia, C. elegans is being extensively used as a model organism.It was the first multicellular organism to have its whole genome sequenced, and as of 2012, is the only organism to have its connectome (neuronal “wiring diagram”) completed. The C. elegans genome contains an estimated 20,470 protein-coding genes.[95] About 35% of C. elegans genes have human homologs. Remarkably, human genes have been shown repeatedly to replace their C. elegans homologs when introduced into C. elegans. Conversely, many C. elegans genes can function similarly to mammalian genes.

3.2

3.3

Lab01_Ex4_seq Reversed: C.elegans TTATTGTTTTCCAAGCTTTAATATCAATT-
TATTGTGCCCGATGTTACCAATTACACTTGA AAAATCTAAAAAGCTTGGAAAC-
TAGCCGAAAATGTGCAGTAAAACAAAATTTCCTATAAA ATCCGAGTTATTTGAAC-
CAAATTCATACTCTTCTCTATTTTATCGTTTTCCGAGCTCTAA TCGTATATAATAT-
TACCTATTTTCAGCTAAATGAGCACATCCGTAGCGGAAAACAAAGCA TTGTCAGCTTC-
CGGCGATGTGAATGCGTCCGATGCTTCAGTTCCTCCAGAGCTTCTCACC AGACAC-
CCCCTCCAGAATCGCTGGGCTCTCTGGTACTTGAAAGCTGACCGTAACAAGGAA
TGGGAGGATTGTCTGAAGGTAGAAGATTTTTAAATACGTCTTTTATCGATTTTTTCCAGA
TGGTTTCACTTTTCGACACTGTCGAGGACTTCTGGTCGCTGTACAATCACATTCAGTCTG
CCGGAGGATTGAACTGGGGATCCGATTATTACTTGTTCAAGGAAGGAATCAAGC-
CAATGT GGGAGGACGTCAACAACGTTCAAGGTGGACGTTGGTTGGTTGTTGTC-
GATAAGCAAGTAC GTTTTGAGAAATATATTTTATTCAATGAATCATAGAAGCTTCA-
GAGAAGAACGCAATTGC TCGATCACTACTGGTTGGAGCTGTTGATGGCTATTGTTG-
GAGAGCAATTCGACGAGTACG GAGACTACATCTGCGGAGCTGTCGTGAATGTTCGT-
CAAAAGGGTGACAAGGTTTCCTTGT GGACTCGTGATGCTACTCGCGATGATGT-
CAATCTTCGCATCGGACAGGTTTTGAAGCAGA AATTGAGCATTCCGGATACTGA-
GATTTTGAGGTAATTTTACAATTTTAGTATTTGCTATC TAAGTAAAATATTTCCA-
GATACGAAGTTCACAAGGACTCGTCGGCTCGCACCTCATCGAC TGTCAAGCCACG-
CATATGTCTTCCAGCCAAGGATCCAGCACCAGTGAAGGAAAAGGGACC AGCCG-
CAACGACTTCTCCATCGAATCCCGGCACGGAGGCTACAGGAACTTCTCCAGCCAC
CCCAACTCCTTAAGCATATTCTAAAGATCTCACCAATTCCTCTCACCGTAAAT-
GAGCTTC CCCGTACTCCCAGTCTCAATGTTGTCTTGAAAAATGAACTGTTTTTCG-
GACACGATCATC GCTTTAACTATTCGAAAATCAGCTCATTTTTCAAGTCGTACCC-
CCCACCTAATGTATTGG TGCTTCCCCTCCAATTTGTACCTACTGTTTCGCTTCCC-
CCTATTGATTTACCGGTTTTCG TATTGCTCTCTTGTTGTTACTAGATTCGAGACT-
GATCGACGCCTGTAGCCGAATTCGTTT GTTCTTCAGGTTAATTGATGAATATATATT-
TATTCGGTAAATATAAATAGATATGTTAGT TATTATTCTTCTTCACACACATGATTTG-
TAGGGCGTTTGATTTTGTACATTTTTAAAAAT

3.4

The query sequence is found on the 11th position

3.5