1. Describe genetic variation within natural populations, covering the following subjects: What factors are responsible for creation, maintenance, and elimination of genetic variation? What parameters can be used to characterize genetic variation quantitatively? What levels of genetic variation are encountered in natural populations? How can patterns in genetic variation be used to detect negative, positive, and balancing selection?

2. Define effective size of a population, N_e. Under what conditions a population would have an infinitely large effective size? Why are effective sizes of natural populations almost always lower than their actual sizes? What methods can be used to estimate effective sizes of natural populations?

3. Formulate, [derive,] and explain Fisher's Fundamental Theorem of natural selection.

4. Какими будут равновесные частоты двух аллелей, A и a, под действием мутаций (и только мутаций)? Является ли мутационное равновесие устойчивым?

5. Опишите несколько возможных подходов к измерению скорости мутирования.

6. For a diploid population with 2 alleles at a locus, *A* and *a*, derive the equation for the Hardy-Weinberg equilibrium. What assumptions does the Hardy-Weinberg equation make? If a population is in Hardy-Weinberg equilibrium, how is it evolving?

7. For a diploid population with two alleles at Hardy-Weinberg equilibrium, at what frequency of first allele *p* is the frequency of heterozygotes maximized? Explain your reasoning.

8. What factors are necessary for Darwinian evolution? Briefly explain each factor.

9. [Derive] and explain Fisher's Fundamental Theorem. What does this equation describe? What are the inherent assumptions of FFT? What does FFT imply?

10. Fisher's fundamental theorem of natural selection states that *the rate of increase in fitness of any organism at any time is equal to its genetic variance in fitness at that time*. **(a)** What is "genetic variance in fitness"? **(b)** Assuming that there is genetic variance in fitness, does Fisher's theorem imply that the mean fitness of a population is always increasing? Why or why not?

11. You have a population of 1000 cockroaches, of which 450 are black, and 550 are gold. You have found that the black variety produces 8 offspring per individual, and the gold produces 5. What is the frequency of the gold genotype in the next generation?

12. Приведите несколько примеров положительного и отрицательного отбора.

13. Что такое балансирующий отбор? Можно ли что-нибудь сказать о том, насколько он распространен в природе? Какую роль он играет в молекулярной эволюции? Обоснуйте ответ.

14. What evidence would suggest that a genomic region under investigation was subject to negative selection? Are there alternative explanations for the patterns of data you have described?

15. What evidence would suggest that a genomic region under investigation was subject to positive selection? Are there alternative explanations for the patterns of data you have described?

16. Consider one locus A with two alleles, A and a, in a large population of haploid organisms. Only mutations from A to a occur, with rate m. Relative fitnesses of genotypes A and a are 1 and 1-s, respectively. Assuming that selection against genotype a is strong enough to keep it rare, show that if the population is at mutation-selection equilibrium, the frequency of a is m/s.

17. Какова скорость потери виртуальной гетерозиготности в результате генетического дрейфа? Объясните.

18. Present several reasons why the effective size of a natural population (the size of equivalent Fisher-Wright population) can be below its actual (head-count) size.

19. Nucleotide diversity ("virtual heterozygosity") H has been measured in two closely related species. In one of them, H = 0.01. In another species, H = 0.05. Consider possible reasons for the difference. What data are needed to determine the true reason?

20. Nucleotide heterozygosity, H, in human is 0.001, and the mutation rate is roughly 2 x 10-8. (a) Estimate the effective population size. Show your work. (b) The current population size in human is about 7 billion. Discuss two possible reasons why your estimation of the effective population size is so small compared to the "head count". (c) Using above information, calculate the minimum value of selection coefficient required for selection to have the dominant effect in the rate of evolution.

21. In a population of 1000 individuals, replacing Gly with Ala at site 25 of a particular protein would increase fitness by 1%. A nonsynonymous nucleotide substitution that is necessary for this replacement occurs with rate 10^{-8}. Estimate roughly the expected number of generations until the new, favorable allele containing Ala at site 25 will be fixed in a population.

22. Нейтральная теория молекулярной эволюции.

23. Consider a stretch of noncoding, nonfunctional DNA that is 1,000 nucleotides long. Assume that the mutation rate is high — 2 changes per site per million generations. (a) In a diploid population of 10,000 individuals, how many new alleles will be created in the population each generation? (b) What fraction of these new alleles will ultimately be fixed? What is the rate of molecular evolution of this sequence? (c) In a population of 50 individuals with the same mutation rate, how many new alleles will be created per generation? What fraction of these will ultimately be fixed? What is the rate of molecular evolution of this sequence?

24. In 90% of nonsynonymous polymorphisms found in human populations both alleles (amino acid variants of the protein) confer essentially the same fitness, and only in 10% of polymorphisms 1 of the alleles is clearly inferior. In contrast, if we introduce a random nonsynonymous mutation into a human protein-coding gene, it does not affect fitness with probability only 10-15%, and reduces fitness substantially in 85-90% of cases. Why mutations are subject to much stronger negative selection than polymorphisms?

25. Как доминантность аллеля влияет на скорость его фиксации под действием положительного отбора?

26. Describe 2 methods that can be used to detect positive selection.

27. Under what (unrealistic) assumption genetic drift would be absent in an asexual population of a finite size? Why the same assumption does not lead to complete absence of drift in a sexual population?

28. Понятие мутационного груза для одного и многих локусов.

29. Соматические мутации и мутации зародышевой линии. Оценка de novo скорости мутирования в человеке: современные оценки, оценки непрямыми способами. Ко-репликационное накопление ошибок, влияние возраста родителей на число de novo мутаций. Локальные свойства ДНК и скорость мутирования. Время репликации и скорость мутирования.

30. Соматические мутации. Влияние эпигенетики на распределение соматических мутаций по геному. Эффект клетки происхождения (cell of origin). MMR и локальная изменчивость скорости мутирования, MMR и время репликации.

31. Репликация: структура репликационной вилки. Специфичность основных репликативных полимераз по лидирующей/отстающей цепи. MMR: молекулярные механизмы работы. Эффективность MMR на лидирующей и отстающей цепи. Как MMR меняет спектр мутирования? Как MMR различает материнскую и дочернюю цепи ДНК?

32. NER: молекулярные основы работы GG-NER и TC-NER. Спектр репарируемых повреждений. К каким последствиям приводит поломка в генах каждой из ветвей NER? XR-seq. Зависимость TC-NER и GG-NER от уровня экспрессии гена. Ассоциированная с репарацией транскрипционная асимметрия. NER и транскрипционные факторы. NER и нуклеосомы.

33. Мутационные подписи. Основные мутационные подписи и их свойства (подписи, ассоциированные с астрономическим временем, подпись UV, подпись курения, подпись APOBEC).