

The Government of the Russian Federation Federal State Autonomous Educational Institution of
Higher Professional Education “National Research University “Higher School of Economics”
Faculty of Computer Science

**Syllabus for the course
“Molecular Evolution”**

for Master’s Program “Data Analysis in Biology and Medicine”

Field of Study 01.04.02 Applied Mathematics and Informatics

Author of the program: Georgii Bazykin

Approved at the meeting of the Department of Technologies for Complex System Modeling

Approved by the Academic Supervisor of master’s program
«Data Analysis in Biology and Medicine»

Academic Supervisor, Mikhail Gelfand _____

Moscow, 2017

This program cannot be used by other divisions of the university and other institutions of higher education without the permission of the department - the developer of the program.

1. Course description

Nothing in biology makes sense except in the light of evolution. This course introduces the fundamentals of evolutionary science as applied to genomics. It will allow to see how the basic population genetics processes create, maintain and affect variability in populations and lead to their changes with time. The focus will be on molecular evolution, i.e., the manifestation of these processes in genomes. As humans, we will be particularly interested in evolutionary aspects of medicine. The course assumes no prior familiarity with evolutionary biology, although knowledge of the basics of molecular biology and genetics is expected. The themes covered will include basic concepts in evolutionary biology and generalizations in evolutionary genomics; population genetics and factors of microevolution; and basics of quantitative genetics. The course will alternate lectures with reading and student-guided discussion of recent literature in the field. There will be more lecturing near the beginning of the course, and more discussing near its end.

2. Course location within the Structure of the educational program

Prerequisite (recommended): introductory course on molecular biology

3. Course learning outcomes

According to the course learning outcomes, after completion of the Course a student should exhibit the following learning outcomes (classified in the three categories of knowledge, skill and experience):

Compliance of course learning outcomes to program learning outcomes

№	Course Learning Outcomes	Результаты обучения по дисциплине
	Overall, the discipline	В целом по дисциплине
	Knowledge	Знать
1.	Understanding of the basic aspects of evolutionary theory and population genetics as applied to genomics.	Понимать основные аспекты эволюционной теории и популяционной генетики в применении к геномике.
	Skill	Уметь
2.	Ability to apply evolutionary reasoning to a wide range of biological problems, for example, explaining how the evolutionary origin of a particular biological system constrains its structure and functioning;	Уметь применять эволюционное мышление к широкому спектру биологических задач, например, объясняя, как эволюционное происхождение определенной биологической системы ограничивает ее структуру и функцию;
3.	Ability to correctly interpret the patterns observed in genomic data in terms of contributing factors such as mutation, drift and selection;	Уметь верно интерпретировать паттерны, наблюдаемые в геномных данных, в терминах создающих их факторов, в т.ч. мутаций, дрейфа и отбора;
4.	Ability to apply major mathematical and statistical models of population genetics to genomic data;	Уметь применять основные математические и статистические модели популяционной генетики к геномным данным;
5.	Ability to make oral and written presentations.	Способность делать устные и письменные презентации.
	Experience	Владеть
6.	Students are expected to be able to understand the population and evolutionary genomics aspects of modern biological literature, infer signals of	Студенты должны быть способны понять популяционно-геномные и эволюционно-геномные аспекты современной биологической

mutation and selection from genomic data, interpret genome patterns in terms of the underlying forces, and present their own results as well as the results of others in evolutionary terms.	литературы, выявлять сигнала мутагенеза и отбора из геномных данных, интерпретировать геномные паттерны через силы, лежащие в основе них, и представлять собственные и чужие результаты в эволюционных терминах.
--	--

4. Structure and Content of the Course

Total discipline workload is 5 credits, 190 academic hours.

Course structure

Topic #	Topic	Instructor-based learning activities, hours		Independent student learning activities	TOTAL
		Lectures	Seminars		
1.	Basic concepts of evolutionary biology	4	4	20	28
2.	Evidence for past evolution	2	4	20	26
3.	Evolution observed	2	4	20	26
4.	Phylogenetics	2	4	20	26
5.	Population genetics and microevolution	4	4	20	28
6.	Quantitative genetics	2	4	20	26
7.	Basics of evolutionary medicine	4	4	22	30
	TOTAL	20	28	142	190

Course content

	Topic	Annotated Topic content
1.	Basic concepts of evolutionary biology	Evolution, common ancestry, genotype, phenotype, trait, fitness, adaptation, mutation, variation, population, selection, allele replacement, fitness landscape, clade, species, microevolution, macroevolution.
2.	Evidence for past evolution	Homology, unforced hierarchy, unforced patterns in spatial distributions, scenario-based and theory-based evidence. Examples.
3.	Evolution observed	Brief review of history of life on Earth from the viewpoint of genomics. Observations of ongoing evolution: natural populations, domestication, controlled populations, pathogens, cancer cells.
4.	Phylogenetics	Evolution as movement in trait space. Approaches to reconstruction of phylogenies. Distance-based vs. model-based methods. Tree rooting. Concerns: homoplasy, variable rates of evolution, non tree-like phylogenies.
5.	Population genetics and microevolution	Factors of microevolution. Mutation, types, distribution, mechanisms and rates of mutations, mutational equilibrium in two-allele model, role of mutation in sequence evolution. Selection, Fisher's fundamental theorem of natural selection, dynamics of allele frequency under selection, types of selection. Joint action of mutation and selection, genetic load, mutational load. Sexual reproduction, segregation, recombination, nonrandom mating and population structure. Selection in a sexual population, evolution of dominance. Genetic drift, Wright-Fisher model, loss of variation due to drift. Coalescent theory. Mutation-drift equilibrium, effective population size. Neutral theory of evolution, molecular clock, distinguishing between drift and selection-driven evolution. Tests for positive and negative selection. Molecular mechanisms of

		adaptation.
6.	Quantitative genetics	Correlation between relatives, heritability, components of variance of a quantitative trait. Response to selection, breeder's equation. Genetic associations, GWAS. Evolutionary quantitative genetics.
7.	Basics of evolutionary medicine	Mendelian and multifactorial disease. Evolutionary causes of disease: constraints, trade-offs and conflicts, resolvable and irresolvable suboptimalities. Ageing. Host-pathogen evolution, pathogen dynamics.
8.	Final exam	

5. Instructions and guidelines for Independent student learning activities

Independent student work on the discipline includes homework assignments, preparation for lectures, seminars and other learning activities, and preparation for exams, as well as preparation of papers and their presentation.

Educational and methodical support of Independent student work presented by topics of all kinds of tasks and guidelines for their implementation.

7.1. Examples of assessment tools for formative assessment

Example:

I. Independent work

1. In 1978, an experimental population of threespine sticklebacks has been established in the freshwater quarry Goluboy near the White Sea consisting of 20 individuals with freshwater *EDA* allele and 20 individuals with marine *EDA* allele. In 2012, when this population was sampled, the frequency of the freshwater *EDA* allele was found to be 72%. (i) Assuming the generation time of two years, constancy of selection, and intermediate dominance of freshwater alleles ($h=0.5$), estimate the selection coefficient s in favor of the freshwater allele. (ii) No such increase has been observed in Quarry Malysh, in which the founder population consisted of just one individual with freshwater *EDA* allele and one individual with marine *EDA* allele. Suggest a reason for this difference between populations.
2. Briefly describe your own two examples of positive selection and two examples of negative selection. Please try to come up with your own examples, not the ones used in class.

7.2 Examples of assessment tools of summative assessment - exam tasks.

II. Paper presentation. Each student will choose one of the suggested recent papers in evolutionary genomics. A ~10 minutes powerpoint presentation is expected, followed by answers to questions and a brief class discussion facilitated by the presenter, for a total of up to 20 minutes. For multidisciplinary papers, the focus should be on the aspects of evolutionary genomics. For maximum grade, the student is expected to present the background, the basics of the methods, main results, and their significance, and to understand the details of the methods and results.

III. Final exam (free response and multiple choice): up to 15 questions each weighting 3 to 15 points, for a total of maximum 100 points.

Examples:

1. (12 pts) (a) How do the different phylogenetic methods we discussed in class differ in terms of their speed of operation on a computer, ability to sample many possible trees, and appropriateness for large data sets? (b) What is the meaning of a bootstrap value of 70%?
2. Essay (20 pts): As a distinguished SkolTech student, you have been called to testify in court in the debate over whether evolution is “just a theory” or whether it has been sufficiently proven. What do you think are three of the best examples that document the evolution by common descent of all life on Earth?

8. Final Grading

№	Grade Component	% of the Final Grade Derived from Each Component
1.	Independent work	45%
2.	Presentations of papers and classwork	25%
3.	Written final exam (during exam week)	30%

9. Grading system

- 10: >95%
9: 85.01 – 95%
8: 75.01 – 85%
7: 65.01 – 75%
6: 55.01 – 65%
5: 45.01 – 55%
4: 35.01 – 45%
3: 25.01 – 35%
2: 15.01 – 25%
1: 0 – 15%

10. Information and Methodological Resources Required for the Course Study

There is no single textbook to cover all material in this course. Main references:

1. B. Charlesworth and D. Charlesworth. Elements of evolutionary genetics. Roberts and Company Publishers, 2010.

Additional references:

2. J.H. Gillespie. Population genetics: a concise guide. The Johns Hopkins University Press, 1998.

3. D. Graur, W.-H.Li. Fundamentals of molecular evolution. 2000.

4. D.L. Hartl, A.G.Clark. Principles of population genetics. Sinauer Associates, Inc., 1997.

5. M. Lynch. The origins of genome architecture. Sinauer Associates, Inc., 2007.

6. J. Wakeley. Coalescent theory: an introduction. Roberts and Company Publishers, 2009.

7. M. Kimura. The neutral theory of molecular evolution. Any edition.

8. M. Lynch and B. Walsh. Genetics and Analysis of Quantitative Traits. Sinauer, Sunderland, MA, 1998.

10. Materials and Equipment Required for the Course Study

Equipment, facilities, materials, etc.:

Lecture Room.

Syllabus documents and materials on the topics of discipline.

The library, including electronic publications.

Access to the Internet through a computer class and Wi-Fi network provided by HSE.