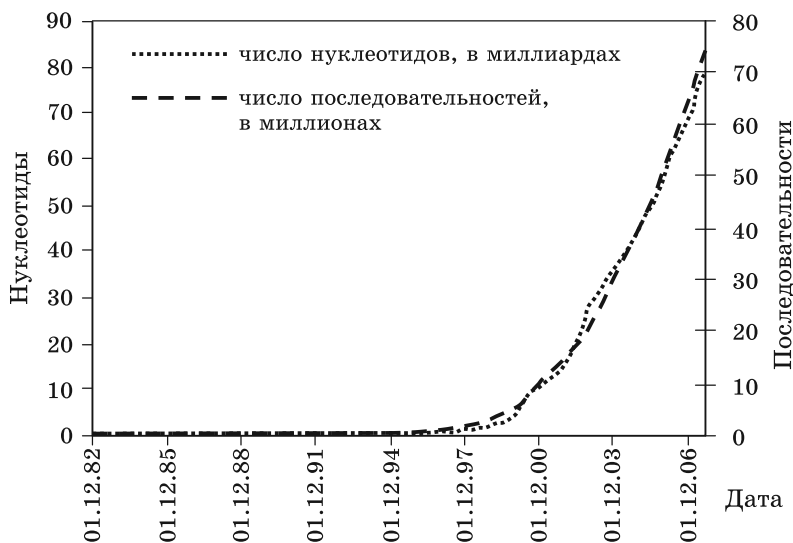


ПРЕДИСЛОВИЕ

Развитие методов полимеразной цепной реакции и автоматического определения нуклеотидных последовательностей сделало возможным получение генетической информации в не представимом ранее масштабе. Число нуклеотидных последовательностей в международной базе генетических данных GenBank превышает 70 миллионов, а их общая длина — 80 миллиардов нуклеотидов (см. рисунок). Объем этой информации удваивается в среднем каждые 18 месяцев.

Одновременное развитие компьютерных технологий привело к тому, что анализ генетической информации, бывший ранее темой работы относительно небольшого числа специалистов, становится повседневной задачей многих научных и практических лабораторий.

В связи с этим, актуально написание книги, освещающей принципы эволюционного анализа генетической информации.



Рост объема информации в генетической базе данных GenBank

Эволюционный анализ генетической информации является составной частью более общего раздела науки — биоинформатики. Спектр задач, которыми занимается биоинформатика, очень широк: к этой области науки относятся изучение происхождения наследственной информации, анализ регуляции синтеза макромолекул — нуклеиновых кислот и белков, предсказание структуры макромолекул, изучение взаимодействий между ними и многие другие вопросы [*Ратнер с соавт.*, 1985; *Гельфанд*, 2005].

Одной из ключевых, основополагающих задач биоинформатики является установление законов изменения макромолекул в ходе их эволюции и, на основании этих данных, — установление родственных отношений между формами жизни, от которых эти макромолекулы получены. Этими вопросами занимается молекулярная филогенетика, возникшая на стыке традиционных наук об эволюции, передаче наследственной информации и математики [*Грант*, 1980; *Иорданский*, 2001].

Наша книга является первой отечественной книгой, освещающей современное состояние вопросов, связанных с филогенетическим анализом в молекулярной эволюции. При работе над книгой, автор ставил целью ознакомить читателя с теоретическими основами и практическими подходами к решению ключевых задач эволюционного и филогенетического анализа. Эта книга предназначена для студентов, аспирантов и научных сотрудников биологических, медицинских и математических специальностей, занимающихся вопросами эволюционного и филогенетического анализа, а также медицинских сотрудников, интересующихся этими вопросами.

В гл. 1 излагаются понятия, используемые в молекулярной эволюции, включая относящиеся к строению макромолекул, генетическому коду, видам эволюционных событий и методам их изучения. При этом подразумевается, что читатель знаком с принципами организации, воспроизводства и наследования генетической информации — строением генов и организацией генома, процессами транскрипции и трансляции, механизмом наследования признаков, законами естественного отбора — из курсов биологии, биохимии и генетики. Глава 2 рассказывает о принципах выравнивания генетических последовательностей. Эволюционные модели и методы установления эволюционных дистанций между последовательностями описываются в гл. 3. Глава 4 посвящена принципам филогенетического анализа и методам построения филогенетических деревьев. В гл. 5 описыва-

ется ряд специальных вопросов, возникающих при изучении молекулярной эволюции, включая рекомбинационный анализ, исследование нуклеотидного состава и использования кодонов, анализ митохондриальной ДНК и молекулярных часов. Специальное внимание уделено использованию генетического анализа при изучении эпидемиологии инфекционных заболеваний — молекулярной эпидемиологии. Все эти вопросы рассмотрены с использованием многочисленных примеров, на конкретных экспериментальных данных. Наконец, гл. 6 посвящена описанию компьютерных программ, используемых в эволюционном анализе.

Написание этой книги было бы невозможным без участия моих коллег. Прежде всего благодарю Эдуарда Карамова и Япа Годсмита (bedankt, Jaap!) за многолетнюю поддержку моих научных интересов. Серьезную помощь при написании книги оказали ее рецензенты: Михаил Гельфанд (Институт проблем передачи информации РАН и Московский государственный университет им. М. В. Ломоносова) и Александр Гултяев (Университет Лейдена, Нидерланды). Александр Пастернак (Университет Амстердама, Нидерланды) помогал в редактировании книги. Владимир Муронец (НИИ физико-химической биологии им. А. Н. Белозерского, Московский государственный университет им. М. В. Ломоносова) и Герман Шипулин (ЦНИИ эпидемиологии Роспотребнадзора, Москва) оказали помощь при публикации книги.

Автор будет признателен за отзывы, замечания, предложения касательно этой книги, которые следует направлять по адресу: Lukashov_book@mail.ru.

В заключение предисловия отметим, что

любой математический алгоритм, любая компьютерная программа, используемая при анализе генетической информации, проанализирует заложенные исследователем данные, используя выбранные исследователем параметры, — и выдаст результаты анализа. Вопрос о том, насколько заложенные в программу данные и выбранные при анализе параметры правильны, соответственно — насколько результаты анализа отражают истинные эволюционные события, — является вопросом, на который должен ответить сам исследователь. Среди прочих методов — и с помощью компьютерных программ.

*Владимир Лукашов,
Амстердам, Нидерланды,
5 апреля 2009 г.*

ENGLISH ANNOTATION

Rapid developments in the areas of PCR and automatic sequencing have made it possible to obtain unprecedented volumes of genetic information. Due to simultaneous progress in computer technologies, evolutionary analysis of genetic information is becoming an everyday task of many scientific and practical laboratories.

This book is entitled «Molecular evolution and phylogenetic analysis» and written by Vladimir Lukashov (Academic Medical Center, University of Amsterdam, The Netherlands and D.I. Ivanovsky Institute of Virology, Moscow, Russia). The book is aimed at pre- and post-graduate students as well as researchers and describes a broad spectrum of theoretical and practical issues arising in the evolutionary analysis.

Chapter 1 describes the principles, aims, and terminology used in molecular evolution and phylogenetic analysis. Chapter 2 deals with sequence alignment. Evolutionary models and genetic distances are described in Chapter 3. Chapter 4 describes phylogenetic tree building methods. Chapter 5 discusses a number of specific issues arising during the evolutionary analysis, including recombination analysis, nucleotide and amino acid composition, codon usage, the analysis of mitochondrial DNA and molecular clock. Specific attention is given to describing the role of evolutionary analysis in studying the epidemiology of infectious disease — molecular epidemiology. All these issues are discussed using actual experimental data. Finally, Chapter 6 describes computer programmes used in the evolutionary analysis.