**МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РОССИЙСКОЙ ФЕДЕРАЦИИ**

**Новосибирский государственный университет**

**Факультет естественных наук**

**УТВЕРЖДАЮ**

**\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

**"\_\_\_\_\_"\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_200\_\_ г.**

**Учебное пособие**

**«Программирование для биологов»**

**разработано в рамках реализации Программы развития НИУ-НГУ**

**Кафедра информационной биологии**

**Новосибирск 2012**

**АННОТАЦИЯ**

Данное учебное пособие посвящено изучению методов и алгоритмов анализа биологических данных, таких как геномные и протеомные последовательности, данные о воздействии различных веществ (включая лекарственные) на протекание биохимических процессов в организме и т.д. Особое внимание в пособии уделено изучению форматов представления данных генетики, молекулярной биологии и биомедицины. Учебное пособие разработано в рамках реализации Программы развития НИУ-НГУ.

Пособие состоит из четырёх основных разделов. Первый раздел посвящен базовым алгоритмическим конструкциям и современным принципам построения программных систем. Рассматриваются методы работы с файлами, синтаксический анализ XML-конструкций и конструирование собственных типов данных. В качестве базового языка программирования используется язык Java. Во втором разделе рассматриваются основные форматы описания биологических данных: форматы представления данных о нуклеотидных последовательностях (FASTA, GenBank и др.) и математических моделей биохимических реакций (SBML). Третий раздел пособия посвящен разработке и реализации методов анализа биологических данных и расчёту на их основе биологических характеристик. Рассматривается построение программ конвейерной обработки (pipe-line) биологических данных, в частности сырых экспериментальных данных, а также средства для полуавтоматической сборки таких конвейерных программ (в частности, система Taverna), а также программные интерфейсы (API) для написания встраиваемых модулей (plugins). В четвёртом разделе приведены примеры программ анализа биологических данных.

Пособие будет интересно молекулярным биологам, математикам, физикам, химикам, информатикам – сотрудникам университета, работникам научно-исследовательских институтов и студентам, интересующимся компьютерными инструментами обработки биологических данных. Актуальность курса обусловлена тем, что в настоящее время не существует пособий, дающих представление о методах и алгоритмах компьютерного анализа экспериментальных данных по молекулярной биологии, особенно в части построения программных систем конвейерной обработки таких данных.

**Сведения об авторах:** Сергей Александрович Лашин, научный сотрудник лаборатории молекулярно-генетических систем ИЦиГ СО РАН, старший преподаватель кафедры информационной биологии ФЕН НГУ, кандидат биологических наук. Афонников Дмитрий Аркадьевич, заведующий лабораторией эволюционной биоинформатики и теоретической генетики ИЦиГ СО РАН, доцент кафедры информационной биологии ФЕН НГУ, кандидат биологических наук.