

СЕТЕВОЕ ОПИСАНИЕ БИОСИСТЕМ: ПРОБЛЕМЫ И ПОДХОДЫ*

Н. А. КОЛЧАНОВ, Д. С. МИГИНСКИЙ, Н. Л. ПОДКОЛОДНЫЙ,
Д. А. РАССКАЗОВ, В. В. СУСЛОВ, В. С. ТИМОНОВ
Институт цитологии и генетики СО РАН, Новосибирск, Россия
e-mail: valya@bionet.nsc.ru

М. Г. СЕРГЕЕВ
Новосибирский государственный университет, Россия

The value of data on any biological system is greatly increased if formalized and represented in a structured computer database. Such representation provides a wide spectrum of capabilities to analyze and to share the information between experts all over the world and to integrate this data with other information resources of the same domain. Since ecology is an essentially synthetic science, multidisciplinary research is a necessity. However, what happens is that individuals working in different fields stick to their respective paradigms and terminology, which leads to confusion in terminology and hampers comparison. In this paper, a formal ontological description of the structural and functional organization of ecosystems (which all together comprise the knowledge domain) is considered. The approach allows dealing with possible dynamics/development/evolution of the ecosystems and with development of a software package aimed at collection and operation with bio-system models (genenetworks and econetworks).

Введение

Глобальный вызов XXI в. — сохранение биоразнообразия на фоне стремительного роста населения планеты и техногенного разрушения окружающей среды. Решение этой проблемы имеет важнейшее значение для устойчивого развития человечества, которое также является частью биосферы. Фактически это означает, что мы должны научиться оценивать биоразнообразие и прогнозировать его динамику в самом широком смысле. Представление об экосистеме складывается из обобщения первичной информации, получаемой при описании (наблюдении) в природе отдельных ее компонентов — растительных сообществ, животного населения, почв, а также элементов субвидового

*Работа выполнена при финансовой поддержке Российского фонда фундаментальных исследований (грант № 05-07-98011), интеграционного проекта СО РАН № 34, проекта “Происхождение и эволюция биосферы” программы президиума РАН (№ 18.3), Госконтракта № 02.467.11.1005 от 30.09.2005 г. “Разработка комплекса программных компонентов для компьютерного моделирования и дизайна в области постгеномной системной биологии (системная биология in silico)” и проекта № 10.4 “Компьютерное моделирование и экспериментальное конструирование генных сетей” Программы Президиума РАН по молекулярной и клеточной биологии.

© Институт вычислительных технологий Сибирского отделения Российской академии наук, 2007.

уровня биоразнообразия — от популяций до генов. Поэтому важнейшей задачей является построение информационной модели экосистемы, где разнообразная информация о свойствах ее компонентов была бы организована и связана так, чтобы максимально обеспечить представление об экосистеме как о целостном явлении. Разрешение этой проблемы с информационной точки зрения возможно посредством: 1) разработки иерархической системы понятий, связанных с экосистемным уровнем (разработкой онтологий экосистемного уровня); 2) определения отношений общих и частных понятий и характера связей между ними; 3) разработки программных сред, обеспечивающих единое информационное пространство на экосистемном уровне и возможности создания информационно-аналитических вычислительных систем. В этой статье мы рассмотрим в основном первую и третью задачи.

1. Взаимосвязь субвидового и надвидового уровней биоразнообразия

Несмотря на то что связь между разнообразием таксонов, эковиоморф и устойчивостью экосистем неоднозначна [1–4], можно утверждать, что между разнообразием сообществ, таксонов и сложностью организации их представителей¹ есть глубокая взаимосвязь, обуславливающая устойчивость экосистем и биосферы в целом. Данные палеонтологии говорят о совпадении глобальных трендов роста таксономического разнообразия, усложнения организмов и роста числа структурных типов экосистем. Если кризисы от ордовика до перми заставляли кривую таксономического разнообразия колебаться (и порой существенно) возле некоего среднего уровня (палеозойское плато), то послепермские кризисы идут на фоне роста биоразнообразия [6–9].

Снижение таксономического разнообразия — искусственное (формирование одновидовых фитоценозов [10]) или естественное (изменение условий внешней среды и/или вымирание многих видов [11, 12]) — приводит к сходным результатам: быстрому массовому формированию из фенотипически однородной популяции(й) различных эковиоморф (порой фенотипически однородных особенностей других видов). Экосистема восстанавливает биоразнообразие (но не исходный видовой состав!), хотя разнообразие таксонов и их ранг еще долго могут оставаться низкими.

Существование организма зависит от информации его генома, но для ее реализации в онтогенезе нужны определенные условия внешней среды, к которым организм приспособился в филогенезе. Воздействия среды не могут направленно изменить генетическую программу, но способны дать возможность реализации того или иного из уже существующих вариантов ее прочтения [13–16] как прямо (путем изменения регуляции экспрессии генов сигнальным воздействием через рецепторы генных сетей [17] или эпигенетической регуляции генома [18, 19]), так и косвенно (путем изменения уровня “онтогенетического шума” теми же механизмами [20]). Организмы, формируя сообщества, экосистемы и биогеоценозы, получают возможность канализировать колебания условий внешней среды, влиять на формирование ландшафтов и в масштабе биосферы — на климат планеты.

¹Под ростом сложности организма независимо от его положения на филогенетическом древе будем понимать увеличение количества элементов, связей между ними, уровней иерархии, а также количества элементов и связей, работающих в единицу времени и/или в единицу объема, и увеличение разнообразия режимов поведения [5].

Разные ценопопуляции вида могут обитать в разных экологических условиях, подвергаясь разнонаправленному отбору. Дисбаланс (контрбаланс) генетических и эколого-ценотических векторов эволюции мешает сужению нормы реакции, обеспечивающей широкую экологическую амплитуду видов, что в итоге препятствует как специализации [10, 21], так и дивергенции [21–23]. Норма реакции вида есть функция его генома. Таким образом, информация об эволюционной истории вида “записывается” в его геноме [24, 25], причем обширность видового ареала и разнообразие условий внутри него², обеспечивая дисбаланс векторов эволюции, служат гарантией стабильности такой “записи”, часть которой в противном случае пропала бы в ходе специализации. Записанная в геноме разнородная информация реализуется и тестируется только через фенотип, чем и можно объяснить адаптивный компонент при массовом формировании экобиоморфа.

Любой фенотипический признак — следствие работы определенной генной сети — группы координированно экспрессирующихся генов [17]. Степень координации в сложных генных сетях организмов может быть разной: от прямой регуляции транскрипции и трансляции до различных накладывающихся на них опосредованных функциональных, градиентзависимых и даже стохастических механизмов [17, 28–31]. Организация генной сети морфологически проявляется в скоррелированности признаков, образовании корреляционных плеяд [32, 33], радикалов [33], а высокая степень координации — в узкой норме реакции, что в генетике соответствует понятию “качественный признак” [13, 14, 34]. В эволюции корреляционные плеяды формируются из независимых признаков вида, когда на них воздействует общий фактор отбора [32, 33]. В ходе такой коэволюции стабилизирующий отбор будет поддерживать смену функциональной связи на генетически закрепленную, т. е. усиливать координацию генной сети: начнется автономизация плеяды признаков от внешней среды [24]. Однако высокая координация имеет обратную сторону — мутация основного центрального регулятора затрагивает множество генов и отражается на морфологии всей плеяды (плейотропный эффект) [35, 36]. В итоге стабилизирующего отбора эволюционная пластичность вида снижается. При изменении среды он либо вымрет, не сумев приспособиться, либо может найти адаптивный компромисс [37] между требованием среды и морфологией плеяды, либо сформирует собственный вектор эволюции, заданный морфологией плеяды³, запустив филоценогенетические процессы образования собственной экосистемы [39]. Следовательно, в ходе эволюции может наблюдаться своеобразная эстафета: первоначально сложившись на уровне экосистемы как устойчивое сочетание факторов среды, стабилизирующий отбор формирует плеяду — генную сеть, которая, в свою очередь, начинает выступать как фактор стабилизирующего отбора.

Таким образом, по мнению М.М. Камшилова [15], экосистемы работают как многократно дублированные информационные каналы, связывающие прошлое и будущее, онтогенез и филогенез, причем метаинформация (код) представлена в виде взаимодействий между видами и надвидовыми сообществами [16, 40]. Все биосистемы от клетки до биосферы можно представить как саморегулирующиеся иерархические системы из

²Эту гипотезу подтверждает хорошо изученный на современном и ископаемом материале факт высокой устойчивости и экспансивности фаун, сформировавшихся на обширных территориях, а также входящих в них видов [21, 26]. Напротив, фауны небольших изолятов (острова, озера), несмотря на крайнее своеобразие адаптаций их представителей, как правило, мало способны к экспансии (хотя могут довольно успешно противостоят ей (например, Новая Зеландия [27])) и часто деградируют при нарушении изоляции [12, 27].

³Палеонтология знает примеры поразительно устойчивых тенденций появления в фенотипе таксонов конкретных морфологических плеяд, несмотря на изменения среды [38].

элементов, соединенных регуляторными контурами в сеть. Все биосистемы подвержены динамике (обратимые изменения), развитию (закономерные необратимые изменения) и эволюции (изменения в закономерностях развития), в ходе которых значительно варьируют число и состав элементов и связей между ними [41]. Наконец, регуляторные контуры могут замыкаться на разных уровнях иерархии от биогеоценотического и экосистемного [1–3] до молекулярно-генетического [42–45]. Значит, программная среда, описывающая различные биосистемы (генные и белковые сети, экосистемы), должна обеспечить не только анализ их определенного статического состояния, но дать возможность накопления данных и расчета динамики поведения практически любых вариантов систем, которые могут возникнуть на базе такого эталона.

2. Подходы к описанию экосистем

Онтология предметной области — набор утверждений, истинных в любой возможной ситуации из предметной области (в нашем случае — организации экосистем). Поскольку разнообразие сообществ, таксономическое разнообразие и сложность организмов взаимосвязаны, онтология должна иметь возможность расширяться на субвидовые уровни биоразнообразия (организменный, клеточный, субклеточный, молекулярно-генетический). Эколого-ценотическая связь между онтогенезом и филогенезом по-разному отражена в традиционных подходах к описанию организации экосистем, которые можно разбить на три группы:

1) функциональные подходы, восходящие к классическим работам А. Тэнсли [46] и формализованные в виде трофических сетей [47];

2) геоэкологические подходы, формализацию которых предпринял В.Б. Сочава [48], отталкивающиеся от близких понятий биогеоценоза и геосистемы — элементарных динамических, коэволюционирующих систем живых, косных и биокосных элементов [40, 49];

3) парагенетические подходы [50], базирующиеся на корреляции программ онтогенеза организмов — членов сообщества [3] и впервые разработанные в рамках паразитологии [51, 52].

Функциональный подход, уделяющий основное внимание самому факту взаимодействия, часто доминирует в экологии водных экосистем, где гидродинамическая подвижность среды облегчает перенос вещества и энергии, позволяя в значительной мере абстрагироваться от фактора размерности [53]. В отличие от экосистемы биогеоценоз — это система вполне определенного размерного класса. В наземных экосистемах с ростом размерности обмен веществом и энергией между удаленными компонентами все более затруднен, поэтому основными факторами консолидации сообществ становятся общность потребных условий среды (факторов) и в меньшей степени скорость передачи сигнала [54]. В обоих случаях исследователя интересуют не столько виды, сколько жизненные формы, объединенные в различные ценоэлементы либо фактом взаимодействия [2, 10], либо общностью факторов среды [54, 55].

Парагенетический подход, напротив, таксонспецифичен: сукцессионная система является гомеостатированным сообществом из особей определенных видов, жизнедеятельность которых определяет характер сукцессии в данных условиях [3]. Хотя в настоящее время количество экосистем, чья степень изученности позволяет спуститься до молекулярного уровня, сравнительно невелико (симбиоз между растениями и азотфик-

сирующими бактериями, паразитарные системы шистосомы и моллюска, вольбахии и дрозофилы), их число постепенно растет⁴ [42–45].

Все три подхода взаимодополняют друг друга, отражая различные аспекты функционирования экосистем. Общее в трех подходах — представление об элементарном объекте (ЭО) — минимальном объекте, неделимом без потери характерных свойств, однако принципы выделения таких объектов различны. Элементарные объекты могут объединяться в группы определенной размерности или масштаба, образующие разные иерархии в зависимости от физической природы ЭО и примененного подхода [41]. Между собой ЭО могут взаимодействовать посредством элементарных связей (ЭС), обеспечивающих обмен вещества, энергии и информации. В литературе классификация элементарных связей разработана слабо по сравнению с элементарными объектами. Наиболее детальная классификация Н.В. Беклемишева [57] малоформализована: выделяя различные классы элементарных связей, автор считал, что одну и ту же ЭС можно включать в разные классы или переводить из класса в класс при смене размерности/масштаба. Наконец, ЭО и их совокупности могут объединяться по принципу часть — целое.

3. Ранговость элементарных объектов

Ранг — это определенная характеристика элементарного объекта, изменение которой существенно влияет на его взаимодействия с другими объектами экосистемы. Набор рангов зависит от природы объекта и может включать, например, яркость окраски, рост, физиологический статус, возраст, численность, половозрастные характеристики или репродуктивный статус у популяций, различные физические и химические характеристики элементов педо- и мортмасс. С помощью рангов можно описать изменение характеристик одного и того же объекта экосистемы в процессе ее функционирования. Примеры такой ранговой динамики объектов: рост организма, смена стадий генезиса косных или биокосных элементов (образование почв, разложение опада), различные этапы образования сообществ (популяций, стад, фитоценозов, паразитоценозов). Любой объект экосистемы может быть охарактеризован несколькими ранговыми шкалами с различными градациями. Рангами можно описать обратимые и необратимые процессы [58].

4. Онтология экосетей

Формально экосистема может быть представлена в виде графа, чьи вершины — элементарные объекты (независимо от их природы) или их группы — соединены ребрами — элементарными связями. Такое представление экосистемы назовем *экосетью*. Оно позволяет описывать все вышеперечисленные особенности сообществ: разнородность элементарных объектов и элементарных связей, иерархическую организацию экосистем, ранговость образующих их объектов. В экосетях могут быть изменения двух родов:

1) *микродинамика* — изменения мощности ребер и вершин в графе из-за изменения величин переменных, характеризующих элементарные объекты и связи;

⁴Как свежий пример можно отметить симбиопаразитарные отношения между налимом, цестодами и бактериями кишечника [56].

2) *макродинамика* — качественные изменения графа из-за появления новых или исчезновения старых элементарных объектов и связей [58].

Примером микродинамики может служить динамика популяций в климаксом сообществе, одно из наиболее ярких проявлений макродинамики — изменение разнообразия видов (динамика биоразнообразия) в ходе сукцессии. В первом случае практически полное использование ресурса, обеспеченное биогеохимическим круговоротом и пространственно-временными ограничениями, препятствует изменению схемы взаимодействий, а плотная упаковка экологических ниш гарантирует стабильность схемы даже при исчезновении того или иного вида [59]. Во втором случае дисбаланс биогеохимических циклов ведет к появлению (исчезновению) ресурсов, что делает сообщество нестабильным по составу, а следовательно, и по схеме взаимодействий [3, 59].

Онтология событий микродинамики (рис. 1, *a*) выделяет элементарные объекты (m1) и связи (m2) [58]. Элементарные объекты делим на живые (m1.1) (организмы и их группировки — популяции, гильдии, синузии и т. д.), косные (m1.2) (элементы ландшафта, химические вещества, физические факторы) и биокосные (m1.3). Это почвы, илы, морт-массы и биологически активные вещества, выделяемые организмами во внешнюю среду (феромоны, витамины, метаболиты). Элементарная связь микродинамики⁵ (m2) может действовать на ЭО, изменяя его характеристики или локализацию (m2.2 — реакция) или на другую ЭС (m2.1 — регуляторный эффект⁶). Элементарные связи микродинамики делятся на массоэнергетические (m2.2.1), топические (m2.2.2), фабрические (m2.2.3), форические (m2.2.4)⁷ и связи развития (m2.2.5).

Если эффект прямо пропорционален количеству потребленной субстанции (энергии) — это трофические связи (m2.2.1.1), если есть сильный нелинейный эффект — информационные (m2.2.1.2) (рис. 1, *a*), к которым отнесем осязательные, зрительные, иммунные и другие взаимодействия.

Топические элементарные связи (m2.2.2) ограничим конкуренцией за пространство местообитания (например, конкуренция гельминтов в узких полостях хозяина), кондиционированием элементами биогеоценоза характеристик биотопа⁸ и физическими ограничениями на доступность компонентов биогеоценоза. Форические элементарные связи (m2.2.4) — перемещения ЭО из структуры в структуру без его особого изменения: миграционные ЭС (m2.2.4.1) — активные и пассивные перемещения живых элементарных объектов; транспортные (m2.2.4.2) — то же для косных элементарных объектов.

Фабрические ЭС (m2.2.3) — деятельность организмов по физическому преобразованию биотопа или биоценоза (например, рытье нор жителями почвы и ила⁹).

⁵Элементарные связи микро- и макродинамики могут быть прямые (без промежуточных стадий) и не прямые (для простоты или из-за недостатка информации промежуточные стадии опущены). Это разделение на рис. 1, *a* и *b* для простоты не показано.

⁶Для регуляторных эффектов микро- и макродинамики принята классификация, аналогичная GeneNet [60]: усиление (increase), ослабление (decrease), включение (switch on), выключение (switch off), за одним исключением. Регуляторные эффекты макродинамики делятся на простые (M2.1.1), с классификацией, аналогичной GeneNet, и эффекты с оптимумом (M2.1.2), которые воздействуют в зависимости от “мощности” или “количества” того объекта в экосети, который оказывает эффект. Так, оптимальное число паразитов способствует гомеостазу экосистемы, тогда как отклонение от него в любую сторону приводит к ее разбалансировке (рис. 1).

⁷Термины заимствованы у Беклемишева [57], но интерпретация отличается.

⁸Яркий пример кондиционирования — затенение. Будучи пойкилотермными, саранчовые часто перемещаются между открытыми, хорошо прогреваемыми микростациями и микростациями, покрытыми растениями, где и кормятся [58].

⁹Если подобная деятельность ведет к построению нового биотопа, следует прибегнуть к онтологии

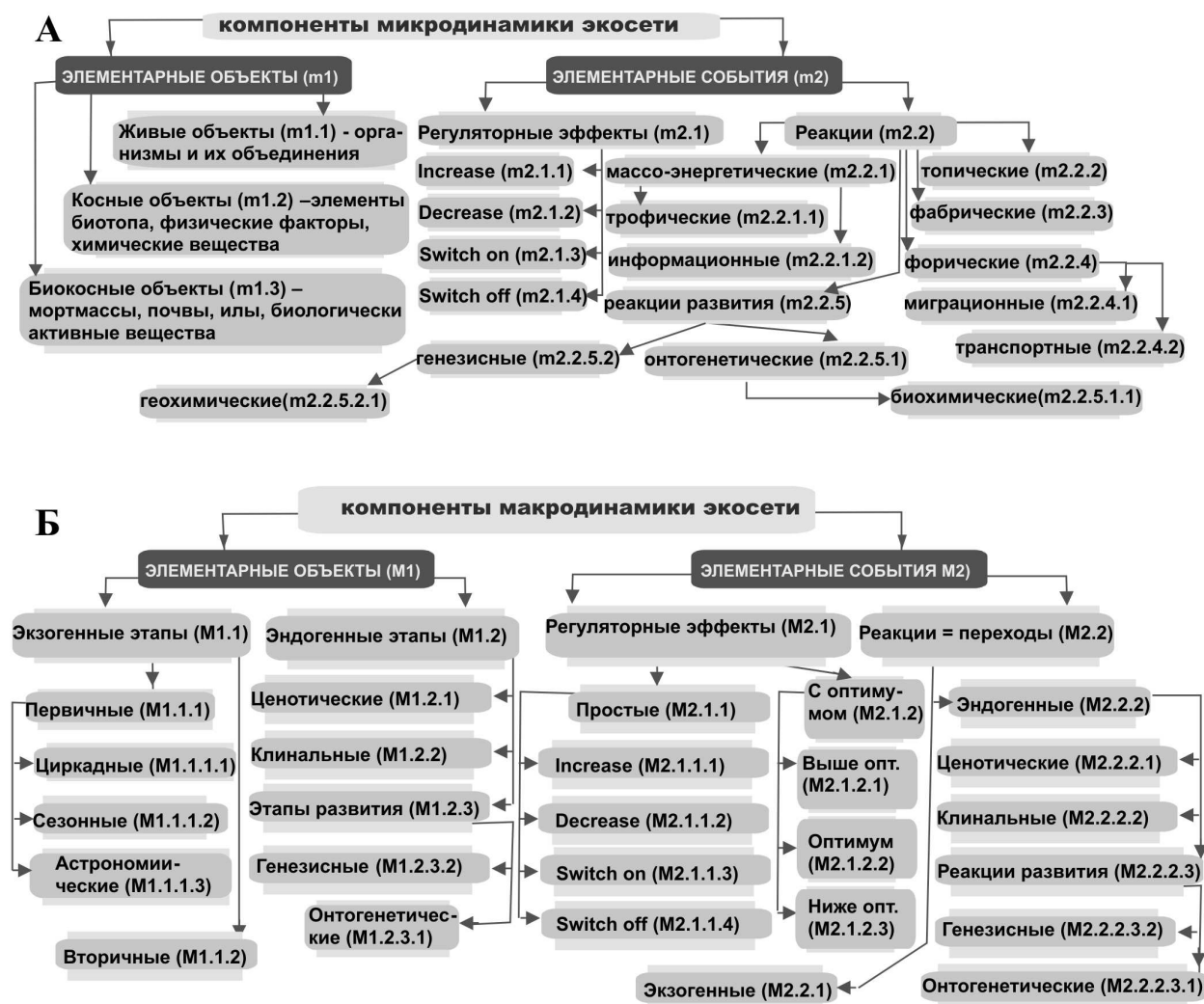


Рис. 1. Компоненты микродинамики (а) и макродинамики (б) экосетей

Элементарные связи развития (m2.2.5) — изменение элементарного объекта, обусловленное его внутренними закономерностями [41]; онтогенетические ЭС для живого объекта (m2.2.5.1); генезисные ЭС — для косных и биокосных ЭО¹⁰ (m2.2.5.2).

Биохимические онтогенетические реакции (m2.2.5.1.1) — связующее звено между экосетями и генными сетями — дают выход на гены онтогенеза. Связь с генными сетями можно прописать и для других элементарных связей живых элементарных объектов по мере накопления информации.

Элементарные объекты макродинамики — это этапы (M1) существования экосетей (рис. 1, б). На каждом этапе экосистема характеризуется стабильностью параметров (числом видов, наличием компонентов биотопа, температурой, влажностью и т. п.). Именно стабильность и определяет дискретность этапов, позволяя выявить устойчивый граф взаимодействий [58]. Этапы делим на экзогенные (M1.1), которые разделены на

событий макродинамики.

¹⁰Примеры онтогенетических ЭС: рост, онтогенез, жизненный цикл, инстинктивные поведенческие акты. Примеры генезисных ЭС — разложение мортмасс, генезис почв, осадочных пород. Совокупностью онтогенетических и генезисных ЭС можно описать смену стадий сукцессии.

первичные (M1.1.1), вторичные (M1.1.2)¹¹, и эндогенные (M1.2). Механизм смены эндогенных этапов лежит внутри экосистемы (дисбаланс биогеохимических циклов, генные сети видов-эдификаторов). Изменение графа экосети из-за появления или выпадения какого-либо компонента сообщества описывает ценотический этап (M1.2.1), изменение графа без изменения состава экосети — клинальный этап (M1.2.2). Изменение графа экосети из-за онтогенеза живых или генезиса косных ЭО микродинамики описывает этапы развития (M1.2.3): онтогенетический (M1.2.3.1) и генезисный (M1.2.3.2).

Элементарные связи макродинамики (M2) включают в себя регуляторные эффекты (M2.1) и реакции (M2.2) = переходы между этапами, которые делятся на экзо- (M2.2.1)¹² и эндогенные (M2.2.2). Последние включают ценотические (M2.2.2.1), клинальные (M2.2.2.2) и переходы развития (M2.2.2.3) — онтогенетические (M2.2.3.1) и генезисные (M2.2.3.2) (рис. 1, б).

5. Программный инструментарий для описания биосистем — экосетей и генных сетей

Описанная онтология, обеспечивающая единообразие описания и сравнения, послужила основой для формирования требований к программному инструментарию для комплексного описания биосистем. За основу была взята база данных GeneNet, разработанная в лаборатории теоретической генетики Института цитологии и генетики СО РАН для описания распределенных в пространстве взаимодействий между объектами генных сетей [60]. База данных GeneNet, реализованная в рамках объектно-ориентированного подхода, позволила описывать простейшие экосистемы как гибридные генные сети, в которых гены одного организма регулируют экспрессию генов другого, оформлять подсети и осуществлять декомпозицию сетей (фильтрация). Однако правила фильтрации, ассортимент и иерархия элементарных объектов в GeneNet были жестко заданы молекулярно-биологической интерпретацией и связаны с системой хранения данных [60]. Это оказалось неприемлемым для описания сложных экосетей с их различными критериями выделения элементарных объектов и их композиции в этапы¹³ (рис. 2).

Ключевыми требованиями при разработке нового инструментария стали:

— независимость структуры базы данных от биологической интерпретации элементарных объектов. Пользователь должен самостоятельно определять объекты в соответствии с применяемым им подходом и терминологией, в идеале пользователь должен получить возможность самому определять набор правил для своей предметной области, т. е. строить собственные онтологии;

— каждый объект может включать в себя другие объекты по принципу часть — целое (если онтология позволяет это, впоследствии она используется как ключ для декомпозиции).

¹¹Первичные этапы дифференцируются по характеру астрономических событий, их вызывающих [61]. Для простоты показаны только циркадные (M1.1.1.1) и сезонные (M1.1.1.2) этапы, а остальные сведены в группу астрономических (M1.1.1.3) (рис. 1, б). Вторичные экзогенные этапы чрезвычайно гетерогенны по своей природе, их классификация выходит за рамки данной онтологии.

¹²Подразделение на экзогенные первичные и вторичные на рис. 1, б для простоты не показано.

¹³Аналогичные проблемы позже выявились при описании генных сетей онтогенеза, которым также свойственна этапность функционирования.

Главные достоинства базы данных, опирающейся на эту архитектуру, — расширяемость и гибкость набора понятий [62, 63]. Архитектура модели данных для комплексного описания биосистем делится на три основных уровня описания (рис. 3): 1) уровень онтологий; 2) уровень моделей; 3) уровень схем и слоев [62, 63].

На первом уровне описываются *онтологии* в виде набора прикладных *понятий* — элементарных объектов, составляющих биосистему. Характеристики понятий описываются в виде наборов *атрибутов*. Атрибуты могут быть простые — любые числовые

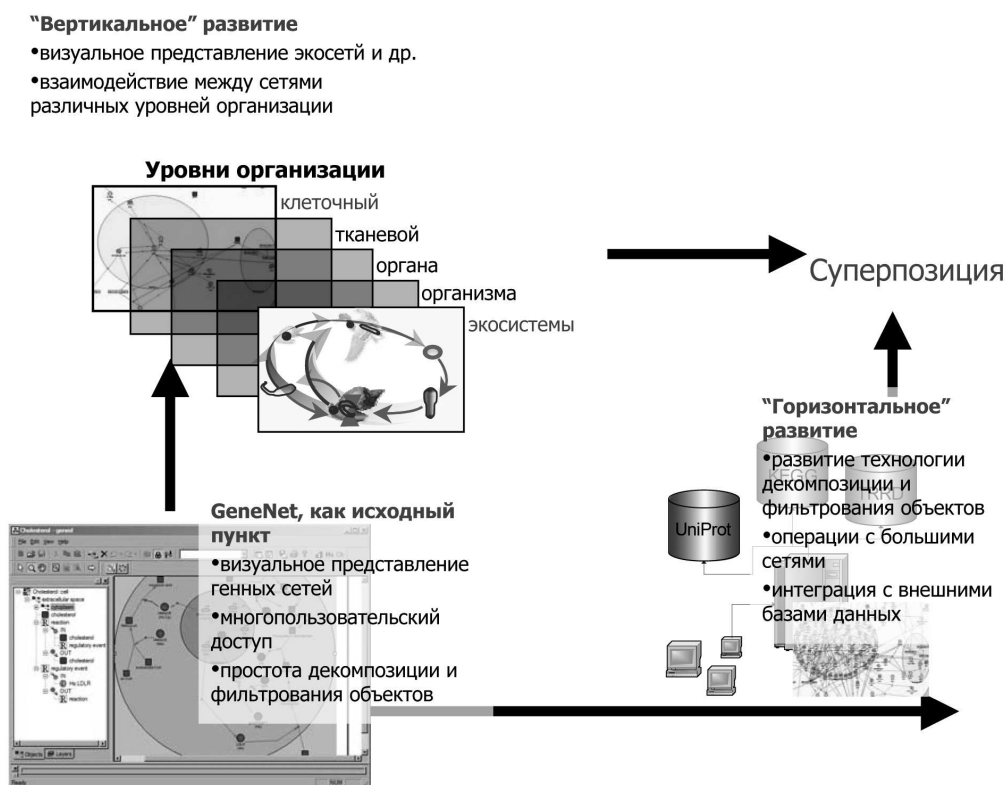


Рис. 2. Схема развития программного инструментария для комплексного описания биосистем



Рис. 3. Архитектура предметной области базы данных для комплексного описания биосистем

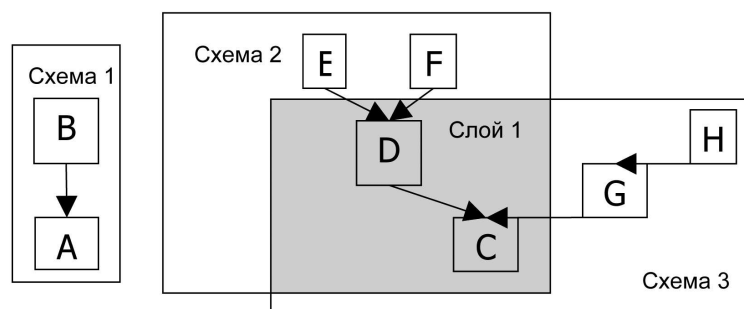


Рис. 4. Организация слоев и схем

характеристики (например, ранги [58]), строковые параметры, такие как имена (например, названия ЭС и т. д.), или ссылочные — входы-выходы взаимодействий между объектами, ссылки на источник информации и т. д. Каждый из атрибутов получает значение в *экземплярах понятия* на нижележащих уровнях (объекты модели, схемы) [58, 62, 63].

Второй уровень предметной области содержит *модели*, каждая из которых представляет собой полное обобщенное описание какой-либо биосистемы в контексте определенной онтологии [58, 62, 63]. Это описание рассматривается как эталонное. Так или иначе, модифицируя его на следующем, третьем, уровне предметной области, мы получаем возможность описывать практически любые траектории развития биосистем, а также их динамику и эволюцию. На втором уровне модель содержит *объекты-модели* (= экземпляры понятий), представляющие собой обобщенные описания конкретных объектов данной биосистемы. Другими словами, это тот материал, при помощи которого мы будем описывать конкретную биосистему.

Третий уровень предметной области позволяет описать конкретное состояние моделируемой биосистемы при помощи *объектов-схем*, являющихся экземплярами объектов модели и группирующихся в *схемы* и *слои*. Слой — многократно используемая часть схемы, содержащая несколько взаимосвязанных объектов. Например, если в ходе изменения биогеоценоза в нем меняется только видовой состав, то объекты, формирующие биотоп, можно описать в виде слоя. Схема может состоять из нескольких слоев, при этом каждый из них может входить в несколько схем (рис. 4). Слои могут быть зависимы между собой (например, биоценоз и соответствующий ему биотоп, орган и составляющие его ткани, ядро и цитоплазма и т. д.) [58, 62, 63]. Во избежание сложностей при декомпозиции взаимозависимые слои организованы в иерархическую древовидную структуру.

Переходя к вопросу реализации программного комплекса, предназначенного для описания, анализа, моделирования и визуального представления биосистем, обозначим ряд критериев, которым он должен удовлетворять. Кроме расширяемости и возможности создания онтологий самим пользователем без привлечения программиста необходимы распределенная клиент-серверная архитектура, благодаря которой обеспечивается конкурентный (многопользовательский) доступ к данным, выделение бизнес-логики в отдельный уровень, что снижает требования к клиентским АРМ¹⁴. Важными требованиями являются также кроссплатформенность (способность работать на различных аппаратных платформах и операционных системах), масштабируемость (способность

¹⁴АРМ — автоматизированное рабочее место.

системы увеличивать свою производительность при повышенной нагрузке и добавлении ресурсов), возможность представления данных в различных видах (графы, плоское текстовое представление, XML и др.) и минимизация стоимости комплекса за счет использования свободного программного обеспечения (Open Source) [62]. Наиболее полно приведенным выше критериям удовлетворяет Java EE Platform. Данная архитектура в полной мере поддерживает расширяемость и масштабируемость на уровне сервера и является свободно распространяемым программным продуктом.

Говоря о кроссплатформенности, нужно заметить, что необходимо обеспечить не просто функционирование комплекса на различных аппаратных платформах и операционных системах, но и способность интеграции с модулями, написанными на разных языках программирования. В настоящее время в биоинформатике существует огромное количество алгоритмов, реализованных на низкоуровневых языках программирования. Было бы глупо отказываться от такого богатого наследства или реализовывать его заново на высокоуровневых языках. Собрав существующие реализации в модули и разместив их на J2EE серверах, можно работать с ними через механизм web-сервисов.

Клиент-серверная архитектура позволяет наилучшим образом обеспечить хранение и изменение разделяемых данных, реализуя механизмы конкурентного доступа. Она также позволяет реализовать многоуровневую распределенную программную модель, разделяя логику приложения на разные компоненты по их функциональному назначению, а именно компоненты клиентского уровня, работающие на АРМ, компоненты web-уровня и компоненты бизнес-уровня, работающие на J2EE серверах, и EIS¹⁵-уровень — программное обеспечение (СУБД и т. п.), работающее на EIS-серверах [64].

Опираясь на такие архитектурные решения, нам удалось разработать программный комплекс для визуального моделирования генных и экологических сетей. Архитектура программного комплекса представлена тремя основными функциональными уровнями: клиентским, уровнем сервера приложений и уровнем баз данных (рис. 5).

Клиентский уровень предназначен для визуального моделирования биосистем и их анализа (декомпозиция, фильтрация и др.). Архитектура данного комплекса представлена stand-alone Java-приложением, построенным на основе технологии Eclipse RCP. Благодаря этой технологии клиентское приложение можно разделить на отдельные модули — плагины (plug-in's) по функциональному назначению. Плагины представляют собой java-архивы (jar) определенной структуры, которые подключаются простым копированием в определенное место файловой структуры приложения. Таким образом, комбинируя различные плагины между собой, пользователь получает приложение с различным функциональным наполнением: либо простой вьюер, либо редактор биосетей, либо редактор онтологий и др. Взаимодействие пользовательского приложения с компонентами серверного уровня (EJB) происходит через набор внешних интерфейсов (remote interfaces), либо через механизм RMI (для Java-приложений), либо через web-интерфейсы для приложений, базирующихся на других программных платформах.

Серверный уровень производит основные операции с данными, обслуживает пользовательские запросы и обеспечивает связь с внешними базами данных. На серверном уровне инкапсулирована бизнес-логика приложения. Бизнес-код находится в компонентах Enterprise JavaBeans (EJB), которые размещены в специализированном контейнере на J2EE-сервере приложений (Application Server). Задача этих компонентов состоит в том, чтобы получать запросы от клиентских приложений, обрабатывать их, если необ-

¹⁵EIS — Enterprise Information System.

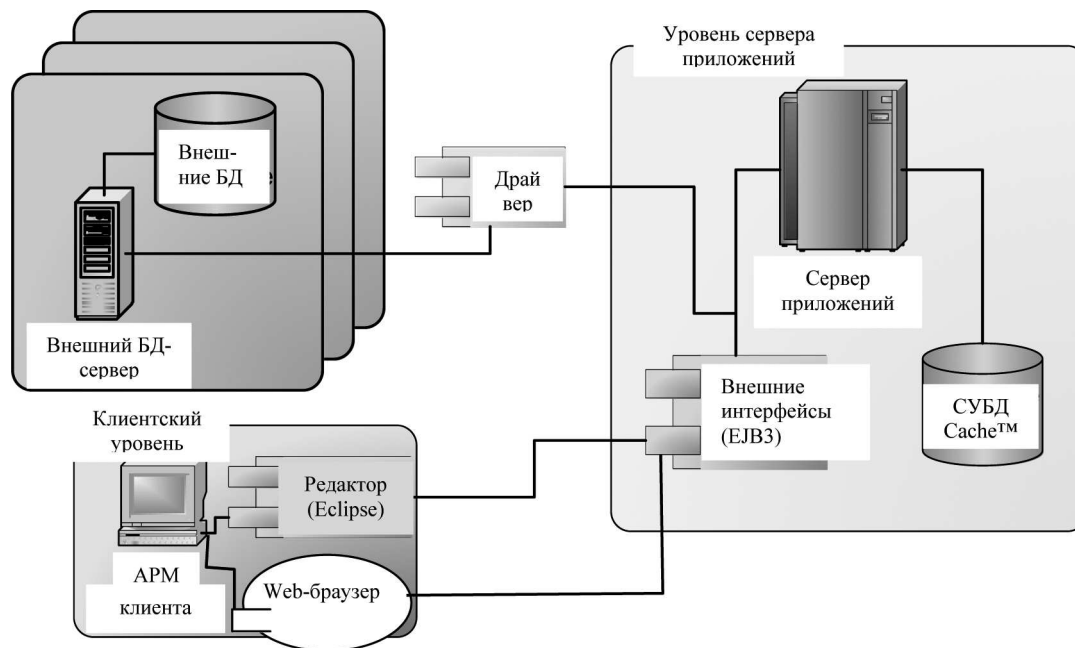


Рис. 5. Архитектура программного комплекса для визуального моделирования генных и экологических сетей

ходимо, и посылать на EIS-уровень. Также EJB получает данные из базы данных, обрабатывает их, если необходимо, и возвращает клиенту.

EIS-уровень предназначен для долговременного хранения пользовательских данных. Он включает в себя различные СУБД¹⁶ и средства для взаимодействия с ними. В качестве основного хранилища данных выбрана постреляционная СУБД Cache™. Данная СУБД поддерживает: организацию персистентности (долговременного хранения) и поиска объектов; встроенные языки для описания классов, методов и хранимых запросов; ограниченное отображение на реляционную модель; двустороннюю навигацию по ссылкам; прямую проекцию пользовательских классов на языки C++, Java; встроенный инструментарий для разработки, а также поддерживает двустороннюю сериализацию в XML.

Источниками данных могут быть внешние удаленные базы данных, взаимодействие с которыми осуществляется через специальные драйверы, при этом внешние базы данных могут быть как полноценными СУБД, так и флет-файлами.

Авторы выражают благодарность в.н.с. Н.И. Юрловой (ИСиЭЖ СО РАН) за плодотворное обсуждение результатов работы.

Список литературы

- [1] БИГОН М., ХАРПЕР ДЖ., ТАУНСЕНД К. Экология. В 2 т. М.: Мир, 1989.
- [2] ОДУМ Ю. Экология. М.: Прогресс, 1986. 620 с.
- [3] РАЗУМОВСКИЙ С.М. Закономерности динамики биогеоценозов. М.: Наука, 1981. 231 с.

¹⁶СУБД — система управления базами данных.

- [4] КАЛАНДАДЗЕ Н.Н., РАУТИАН А.С. Симптоматика экологических кризисов // Стратиграфия. Геологическая корреляция. 1993. Т. 1, № 5. С. 3–8.
- [5] СУСЛОВ В.В., ГУНБИН К.В., КОЛЧАНОВ Н.А. Генетические механизмы кодирования биологической сложности // Экологическая генетика. 2004. Т. 2, № 1. С. 13–26.
- [6] SERKOSKI J.J. Limits to randomness in paleobiologic models: the case of Phanerozoic species diversity // Acta Palaeontol. Polon. 1994. Vol. 38, N 3–4. P. 175–198.
- [7] SERKOSKI J.J. Patterns of Phanerozoic extinction: a perspective from global data bases // Global Events and Event Stratigraphy. Berlin: Springer, 1996. P. 35–51.
- [8] НЕВЕССКАЯ Л.А. Пермо-триасовый и мел-палеогеновый кризисы биоты: различие и сходство // Экосистемные перестройки и эволюция биосферы: Сб. науч. тр. М.: ПИН РАН. 2004. Вып. 6. С. 51–56.
- [9] МАРКОВ А.В. Новый подход к оценке динамики разнообразия фанерозойской морской биоты // Палеонтолог. журн. 2001. № 1. С. 3–12.
- [10] ГАЛАНИН А.В. Флора и ландшафтно-экологическая структура растительного покрова. Владивосток: ДВО АН СССР, 1991. 272 с.
- [11] АНДРЕЕВ Н.И., АНДРЕЕВА С.И. Эволюционные преобразования двустворчатых моллюсков Аральского моря в условиях экологического кризиса. Омск: Изд-во Омского гос. педагог. ун-та, 2003. 382 с.
- [12] ГОЛУБЦОВ А.С. Эволюционные процессы в сообществах с низким видовым разнообразием на примере формообразования у рыб в горных озерах // Междунар. рабочее совещание “Происхождение и эволюция биосферы”: Тез. докл. Новосибирск: ИК СО РАН, 2005. С. 105–107.
- [13] КАМШИЛОВ М.М. Отбор, как фактор, меняющий зависимость признака от изменений внешних условий // Докл. АН СССР. 1939. Т. 23, № 4. С. 361–364.
- [14] КАМШИЛОВ М.М. Проявление признака и изменчивость. Эффект порога проявления // Докл. АН СССР. 1940. Т. 26, № 6. С. 605–608.
- [15] КАМШИЛОВ М.М. Эволюция биосферы. 2-е изд. М.: Наука, 1974. 256 с.
- [16] МАРГАЛЕФ Р. Облик биосферы. М.: Наука, 1992. 214 с.
- [17] КОЛЧАНОВ Н.А., ПОДКОЛОДНАЯ О.А., ИГНАТЬЕВА Е.В. и др. Интеграция генных сетей, контролирующих физиологические функции организма // Информ. вест. ВОГиС. 2005. Т. 9, № 2. С. 179–198.
- [18] SRINIVASAN M., LAYCHOCK S.G., HILL D.J., PATEL M.S. Neonatal nutrition: metabolic programming of pancreatic islets and obesity // Exp. Biol. Med. (Maywood). 2003. Vol. 228, Is. 1. P. 15–23.
- [19] VADLAMUDI S., KALHAN S.C., PATEL M.S. Persistence of metabolic consequences in the progeny of rats fed a HC formula in their early postnatal life // Am. J. Physiol. 1995. Vol. 269, Is. 4. Pt 1. P. E731–E738.

- [20] СТРУННИКОВ В.А., ВЫШИНСКИЙ И.Л. Реализационная изменчивость у тутового шелкопряда // Проблемы генетики и теории эволюции: Сб. науч. тр. Новосибирск: Наука, 1991. С. 214–228.
- [21] ШВАРЦ С.С. Экологические закономерности эволюции. М.: Наука, 1980. 278 с.
- [22] СЕВЕРЦОВ А.С. О причинах эволюционной стабильности популяций видов в природе // Журн. общ. биол. 1986. Т. 4, № 6. С. 723–734.
- [23] СЕВЕРЦОВ А.С. Внутривидовое разнообразие как причина эволюционной стабильности // Журн. общ. биол. 1990. Т. 51, № 5. С. 79–89.
- [24] ШМАЛЬГАУЗЕН И.И. Факторы эволюции (теория стабилизирующего отбора). М.: Наука, 1968. 452 с.
- [25] БЕРДНИКОВ В.А. Основные факторы макроэволюции. Новосибирск: Наука, 1990. 253 с.
- [26] СИМПСОН ДЖ. Великолепная изоляция. М.: Мир, 1983. 256 с.
- [27] ЖЕРИХИН В.В. Биоценотическая регуляция эволюции // Палеонтолог. журн. 1986. № 1. С. 3–12.
- [28] ЧУРИКОВ Н.А. Молекулярные механизмы эпигенетики // Биохимия. 2005. Т. 70, № 4. С. 493–513.
- [29] FINNEGAN E., MATZKE M. The small RNA world // J. Cell. Sci. 2003. Vol. 116. P. 4689–4693.
- [30] HELD L.I. Imaginal Discs: the Genetic and Cellular Logic of Pattern Formation. Cambridge: Cambridge Univ. Press, 2002. 476 p.
- [31] MOUSE Development: Patterning, Morphogenesis and Organogenesis / J. Rossant, P. Tam, (Eds). San Diego: Acad. Press, 2002. 736 p.
- [32] БЕРГ Р.Л. Корреляционные плеяды и стабилизирующий отбор // Генетика и эволюция. Избранные труды: Сб. науч. тр. Новосибирск: Наука, 1993. С. 137–178.
- [33] ВАВИЛОВ Н.И. Избранные произведения в двух томах. Л.: Наука, 1967. Т. 1. С. 7–61.
- [34] КАМШИЛОВ М.М. Изменчивость и проявление. Проблема нормального фенотипа // Докл. АН СССР. 1940. Т. 29, № 3. С. 239–243.
- [35] ГУНБИН К.В., СУСЛОВ В.В., ОМЕЛЬЯНЧУК Н.А., КОЛЧАНОВ Н.А. Генетические механизмы морфологической эволюции. Ч. 1 // Сиб. эколог. журн. 2004. Т. 11, № 5. С. 599–610.
- [36] ГУНБИН К.В., СУСЛОВ В.В., ОМЕЛЬЯНЧУК Н.А., КОЛЧАНОВ Н.А. Генетические механизмы морфологической эволюции. Ч. 2 // Сиб. эколог. журн. 2004. Т. 11, № 5. С. 611–621.
- [37] РАСНИЦЫН А.П. Темпы эволюции и эволюционная теория (гипотеза адаптивного компромисса) // Эволюция и биоценотические кризисы: Сб. науч. тр. М.: Наука, 1987. С. 46–64.
- [38] РОЗАНОВ А.Ю. Закономерности морфологической эволюции археоциат и вопросы ярусного расчленения нижнего кембрия. М.: Наука, 1973. 164 с.
- [39] ЖЕРИХИН В.В. Основные закономерности филогенетических процессов (на примере неморских сообществ мезозоя и кайнозоя): Дис. ... д-ра биол. наук в форме научного доклада. 1997. <http://macroevolution.narod.ru>

- [40] АРМАНД А.Д. Самоорганизация и саморегулирование географических систем. М.: Наука, 1988. 264 с.
- [41] ЖЕРИХИН В.В. Эволюционная биоценология. Проблема выбора моделей // Экосистемные перестройки и эволюция биосферы: Сб. науч. тр. / М.: Недра, 1994. Вып. 1. С. 13–20.
- [42] ГОРЯЧЕВА И.И. Бактерии рода *Wolbachia* — репродуктивные паразиты членистоногих // Успехи современной биологии. 2004. Т. 124, № 3. С. 246–259.
- [43] SCHULTZE M., KONDOROSI A. Regulation of symbiotic root nodule development // Ann. Rev. Genet. 1998. Vol. 32. P. 33–57.
- [44] DE JONG-BRINK M. How schistosomes profit from the stress responses they elicit in their hosts // Adv. Parasitol. 1995. Vol. 35. P. 177–256.
- [45] ADAMO S.A. Modulating the modulators: parasites, neuromodulators and host behavioral change // Brain Behav. Evol. 2002. Vol. 60, N 6. P. 370–377.
- [46] TANSLEY A. The use and abuse of vegetational concepts and terms // Ecology. 1935. Vol. 16. P. 284–307.
- [47] LUCZKOVICH J.J., BORGATTI S.P., JOHNSON J.C. ET AL. Defining and measuring trophic role similarity in food webs using regular equivalence // J. Theor. Biol. 2003. Vol. 220, N 3. P. 303–321.
- [48] СОЧАВА В.Б. Введение в учение о геосистемах. Новосибирск: Наука, 1978. 319 с.
- [49] СУКАЧЕВ В.Н. Основы лесной типологии и биогеоценологии. Л.: Наука, 1972. 418 с.
- [50] КУРКИН К.А. Системные исследования динамики лугов. М.: Наука, 1976. 284 с.
- [51] ПАВЛОВСКИЙ Е.Н. Организм как среда обитания // Природа. 1934. № 1. С. 80–91.
- [52] БЕКЛЕМИШЕВ Н.В. Популяции и микропопуляции паразитов и нидиколов // Биоценологические основы сравнительной паразитологии: Сб. науч. тр. / М.: Наука, 1970. С. 215–226.
- [53] ДЕГЕРМЕНДЖИ А.Г., АДАМОВИЧ В.В., РОГОЗИН Д.Ю. и др. Модели динамики и биоразнообразия экосистем континентальных водоемов // Биоразнообразие и динамика экосистем: информационные технологии и моделирование: Коллективная монография. Новосибирск: Изд-во СО РАН, 2006. С. 419–456.
- [54] БЕКЛЕМИШЕВ Н.В. Пространственная и функциональная структура популяций // Биоценологические основы сравнительной паразитологии: Сб. науч. тр. М.: Наука, 1970. С. 226–239.
- [55] НАЗИМОВА Д.И., АНДРЕЕВА Н.М., КОФМАН Г.Б. и др. Портретные модели структурного биоразнообразия лесного покрова // Биоразнообразие и динамика экосистем: информационные технологии и моделирование: Коллективная монография. Новосибирск: Изд-во СО РАН, 2006. С. 517–535.
- [56] ИЗВЕКОВА Г.И. Трофические отношения в системе хозяин (налим *Lota lota*) — паразит (*Eubothrium rugosum*) — симбионтная микрофлора при гидролизе углеводных компонентов пищи // Журн. эволюционной биохимии и физиологии. 2006. Т. 42, № 5. С. 472–478.

- [57] БЕКЛЕМИШЕВ Н.В. О классификации биоценологических связей // Биоценологические основы сравнительной паразитологии: Сб. науч. тр. М.: Наука. 1970. С. 90–138.
- [58] СЕРГЕЕВ М.Г., СУСЛОВ В.В., МИГИНСКИЙ Д.С. Опыт создания моделей на основе онтологии экосистем // Биоразнообразие и динамика экосистем: информационные технологии и моделирование: Коллективная монография. Новосибирск: Изд-во СО РАН, 2006. С. 95–118.
- [59] КРАСИЛОВ В.А. Нерешенные проблемы теории эволюции. Владивосток: ДВНЦ АН СССР, 1986. 138 с.
- [60] ANANKO E.A., PODKOLODNY N.L., STEPANENKO I.L. ET AL. GeneNet: a database on structure and functional organisation of gene networks // Nucleic Acids Res. 2002. Vol. 30. P. 398–401.
- [61] ДАЖО Р. Основы экологии. М.: Прогресс, 1976. 450 с.
- [62] MIGINSKY D.S., SOKOLOV S.A., LABUZHNSKY V.V. ET AL. Object-oriented approach to bioinformatics software resources integration // Proc. of the Fifth Intern. Conf. on Bioinformatics of Genome Regulation and Structure (BGRS'2006). Novosibirsk: ICG RAS, 2006. Vol. 3. P. 288–291.
- [63] MIGINSKY D.S., SUSLOV V.V., RASSKAZOV D.A. ET AL. Architecture of software toolkit for storing and operating with biosystems models // Ibid. P. 292–295.
- [64] SUN Developer Network(SDN). <http://java.sun.com/javaee/reference/>

Поступила в редакцию 11 мая 2007 г.