УДК 004.9:550.84

УНИВЕРСАЛЬНЫЕ СИНОНИМИЧЕСКИЕ И АЛЬТЕРНАТИВНО-СИНОНИМИЧЕСКИЕ ИНФОРМАЦИОННЫЕ СИСТЕМЫ ПО БИОРАЗНООБРАЗИЮ И ВОЗМОЖНОСТИ ИХ ИСПОЛЬЗОВАНИЯ ДЛЯ КОНТРОЛЯ И МОНИТОРИНГА ОКРУЖАЮЩЕЙ СРЕДЫ

Ю.В. Агарков, А.Ю. Агарков*

Сочинский научно-исследовательский центр РАН
г. Сочи, ул. Театральная, 8а
E-mail: agarkov2011@yandex.ru
*Палеонтологический институт РАН
г. Москва, ул. Профсоюзная, 123

Приведена структура биологического блока универсальной информационной системы, позволяющей хранить любую информацию в области естествознания. Определен основной и сопутствующий объекты хранения информации. На основе имеющихся сведений по экологии организмов показана возможность проследить изменения среды обитания, как в современных экосистемах, так и в бассейнах геологического прошлого.

Введение. Познание видового богатства Земли в целом и ее отдельных регионов является одной из важнейших задач естествознания. Не менее важно знать качественные и количественные изменения для отдельных биологических комплексов, а также природных явлений. Только полные и достоверные данные о динамике этих процессов в совокупности со знаниями о краткосрочной и долгосрочной смене параметров среды обитания позволят понять влияние происходящих природных процессов по отдельности и в интегрированном виде и дать прогноз на будущее.

В связи с весьма коротким по времени периодом существующих наблюдений за природными явлениями (около 150 лет) для оценки возможных факторов риска и глобальных изменений необходимо учитывать сведения геологического прошлого, где природа уже поставила «экологические эксперименты», хотя их результаты и дошли до нас в не полном виде.

Для решения столь масштабных задач требуется инструментарий соответствующего уровня, которым могут стать мощные информационные системы. Однако на сегодня нет информационных систем или баз данных, содержащих относительно полные списки видов как по крупному биологическому таксону (тип, класс), так и по ряду различных таксонов для конкретного региона. Существующая наиболее полная база данных «Catalogue of life» содержит более полутора миллионов названий животных и растений, однако практически все типы или классы простейших, являющиеся начальным звеном трофических цепей, по сути, лишь обозначены [1]. Например, для типа Радиолярии приведено 70 названий, хотя только в современных океанах в настоящее время их известно более 1000 видов, а с учетом ископаемых их число достигает почти 15 тыс. Тоже и с диатомовыми водорослями, имеющими широкий спектр экологических ниш и являющимися индикаторами более десятка параметров среды обитания. Кроме того, по разным оценкам они составляют от 50 до 70% биомассы мирового океана.

Важнейшим фактором, влияющим на результаты анализа, является достоверность исходных данных. Для баз данных это в первую очередь полнота обработки имеющихся сведений в опубликованных и иных источниках и, во-вторых, отсутствие любых интерпретаций, даже если изменения всеми считаются общепризнанными. Если первый фактор можно исправить, добавляя со временем информацию, то операторская интерпретация при вводе авторских данных полностью уничтожает достоверность. Основными примерами интерпретации ввода являются синонимика и датировки геологического времени. Все известные нам базы данных, имеющие синонимику, приводят списки синонимов для общепризнанного вида в текстовом виде при описании или в виде таблицы синонимов. При этом вся информация пространственно-временной распространенности, экологии, морфологии и многом другом привязана к валидному виду. Однако всем хорошо известно, что синонимика и систематика не стабильны и ее

изменения через два-три года уже искажают фактические данные. Еще сложнее, когда определенные сведения принадлежат не виду, а подвидовому таксону (подвид, вариетете, форма, подформа) и они существенно отличаются от базового вида.

Основная часть. При палеонтологических исследованиях появляются проблемы, связанные с не полнотой организмов и появлением систематик, основанных на разных ведущих признаках морфологии скелета, раковины и др. В этом случае возникают абсолютно не сопоставимые системы, где один и тот же экземпляр одним автором трактуется, как представитель валидного вида, а в другом, относится к подформе совсем другого вида. Такие систематики существуют, например у силикофлагеллат.

При анализе временных рамок распространения и при построении карт ареалов требуется однозначность определения возраста и часто при введении этих данных на основе корреляционных шкал операторы, при неправильно организованной структуре таблиц, занимаются интерпретацией при вводе данных. Однако даже в таких консервативных областях знаний как стратиграфия только в XX веке трижды менялось положение некоторых ярусов, не говоря о свитах, зонах и других стратонах.

Для устранения всех перечисленных и многих других проблем авторами разработана структура хранения практически всех данных, используемых в естественных науках, и позволяющая ввод информации из первоисточников, не интерпретируя сведения, отмеченные в них. Основой для нее послужила предложенная Ю.В. Агарковым в 1993 г. и в дальнейшем постоянно усовершенствующаяся многоцелевая информационная система (ИС) по радиоляриям (RADBASE), а затем и другим организмам (BIOBASE) [2 – 4].

База данных представляет собой комплексную геоинформационную систему, которая содержит и постоянно обновляет все сведения по современным и ископаемым организмам, как авторского, так и обширного опубликованного материала. Она состоит из нескольких взаимосвязанных блоков: собственно биологической базы, библиографической, базы кодировки геологического пространства и картографической основы. Каждый из этих блоков представляет собой отдельную информационную систему, которая может быть использована и для создания прочих баз данных (рис. 1).

Биологический блок является универсальным и позволяет хранить информацию по любой группе фауны и флоры, не зависимо от их морфологической характеристики, возрастного диапазона существования и разнообразия экологических обстановок существования. В настоящее время с целью комплексного анализа биологических данных он включает сведения по современным и ископаемым микроорганизмам (радиоляриям, диатомовым водорослям, силикофлагеллатам, эбриидеям, актиномицидам, кокколитофоридам, фораминиферам) и многоклеточным (губки, кораллы, моллюски, насекомые, высшие животные и растения). Всего на данный момент содержатся сведения более чем по 130 тыс. видам вышеперечисленных групп организмов, собранные со всех континентов и океанов отобранные авторами в различных экспедициях и из более чем 20 тыс. опубликованных и рукописных работ.

Единицей хранения биологической информации в ИС является не вид или подвидовой таксон, как принято во всех известных аналогичных разработках, а совокупность реально существующих экземпляров, одинаково называемых (в биноминальной ботанической, зоологической или палеонтологической номенклатуре) и, главное, одинаково понимаемых определенной группой биологов.

В целом биологический блок представляет серии реляционных таблиц для каждой из выше перечисленных групп микроорганизмов и связанных системами ключевых полей. Вследствие универсальности таблиц хранения информации и гибкости операционной обработки данных, система позволяет проводить анализ как для каждого класса или отдела, так и для всех таксонов высокого ранга вместе. Основными в описываемом блоке являются следующие таблицы:

Р и с. 1. Упрощенная модель данных биологического блока

1. Словарь видов.
2. Систематика видов.
3. Описание видов.
4. Изображение видов.
5. Распространенность видов.
6. Морфологические признаки видов.
7. Словарь морфологических признаков видов.
8. Параметры скелета видов.
9. Экология видов.
10. Возрастное распространение видов.
11. Словарь родов.
12. Описание родов.
13. Систематика родовых таксонов.
14. Словарь надродовых таксонов.
15. Систематика надродовых таксонов.
16. Описание надродовых таксонов.
17. Таблица синтеза данных для анализа распространения.

Основной словарь видовых и подвидовых таксонов содержит от нескольких десятков тысяч (радиолярии, диатомовые, высшие растения) до сотен и единиц (эбриидеи, актиномициды) валидных таксонов и их синонимов. Для каждого из них приведены все доступные описания на языке оригинала и данные о возрастной распространенности, рассчитанные на основании интервалов существования видов данного рода.

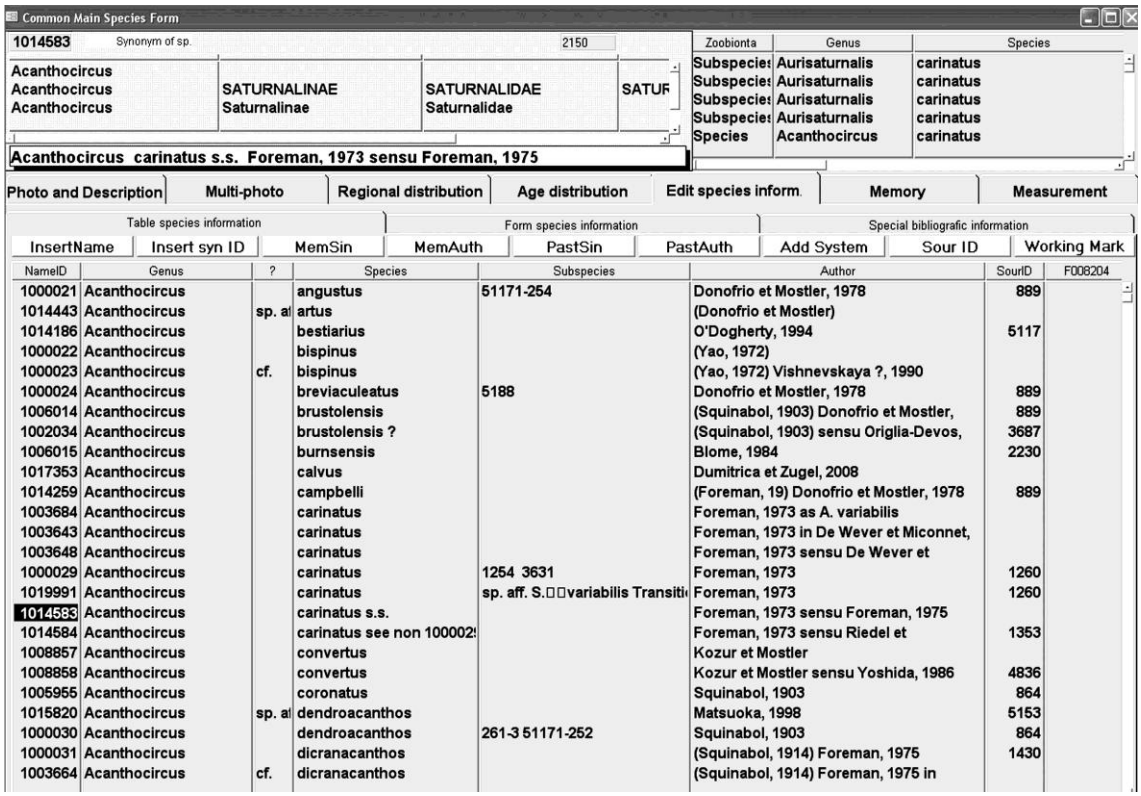
Главной особенностью программы является возможность, помимо существовавшей синонимической кодировки, перехода на альтернативную кодировку таксономических синонимов и выбора вариантов надродовой систематики.

Вышеперечисленные таблицы в соответствии с разработанной структурой сведены в единую форму (рис. 2). Помимо основных таблиц в ИС работает более сорока различных словарей и справочных таблиц, которые определяют тип записи в информационной системе (вид, базиноним, синоним, сомнительный вид, подвид, вариетет, форма, подформа и др.). Учтены методики количественного подсчета, авторство определения таксона в пробе, характер обобщения материала в публикации (данные по образцу, группе образцов, группе обнажений или скважин, общая сводка по региону), уровень достоверности определений экземпляра, точность и метод получения координат при отборе проб, погодные условия, литотипы пород и их различные характеристики и многие другие параметры.

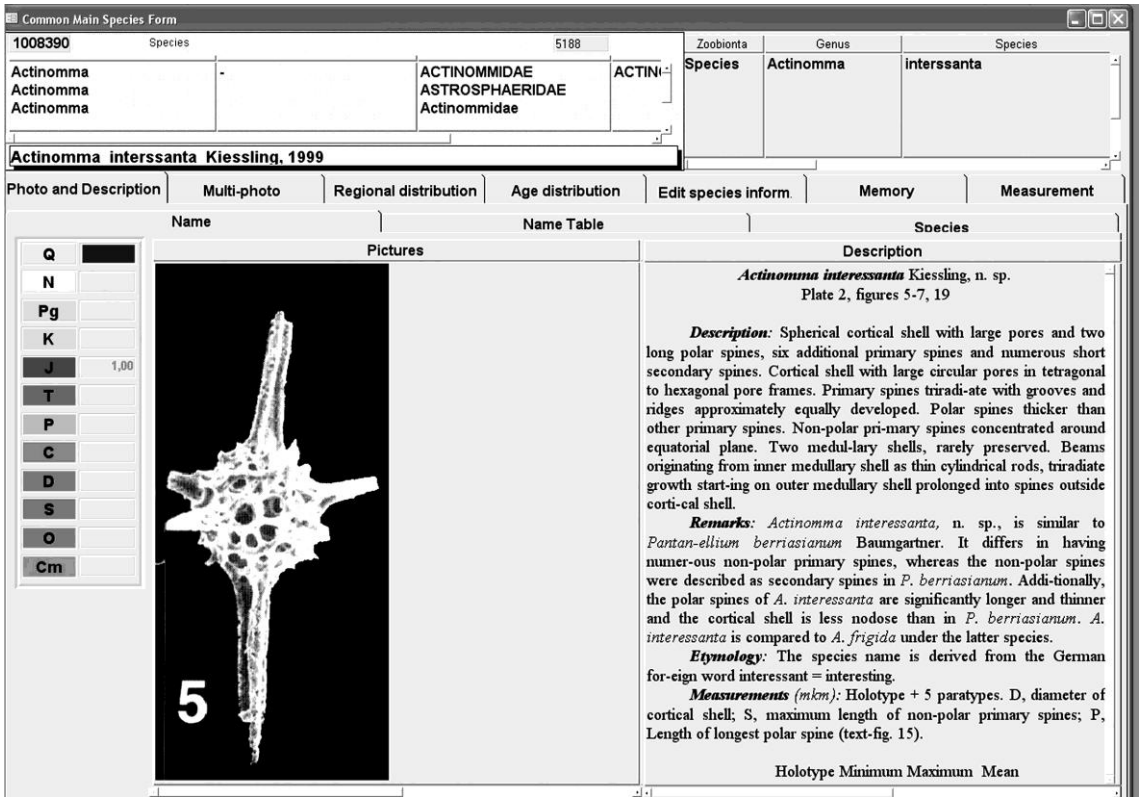
Основная форма содержит серии закладок, отражающих определенный тип информации по данному названию (записи) (рис.2). На представленном фрагменте списка названий (Edit species information) видно, что устанавливая курсор на определенном названии (виде) в данном списке помимо основной атрибутики (название, автор, код источника информации, старое название рода) в верхней общей части формы появляется полное наименование таксона, а в правой части вся его синонимика. Серия кнопок над списком является вынесенными пунктами меню, позволяют быстро внести новую запись, определить ее синонимическое положение, внести информацию о систематическом положении вида и при желании получить обобщенные данные по всем синонимам данного названия.

На странице «Photo and Description» на вкладыше «Name» можно пролистать по отдельности все имеющиеся в информационной системе отсканированные и уменьшенные под единый формат изображения и все известные описания, относящиеся к данному названию (рис. 3). Здесь же в обобщенном виде дан интервал распространения всего вида. В серии глубже расположенных вкладышей “Name table” расположены все списки изображений с пояснениями к ним и все описания с указаниями типа данных (описания голотипа, переописание, примечания и т.д.). При необходимости эти же сведения, но целиком по всему виду с вариантом выбора альтернативной синонимик можно просмотреть на вкладыше “Species”.

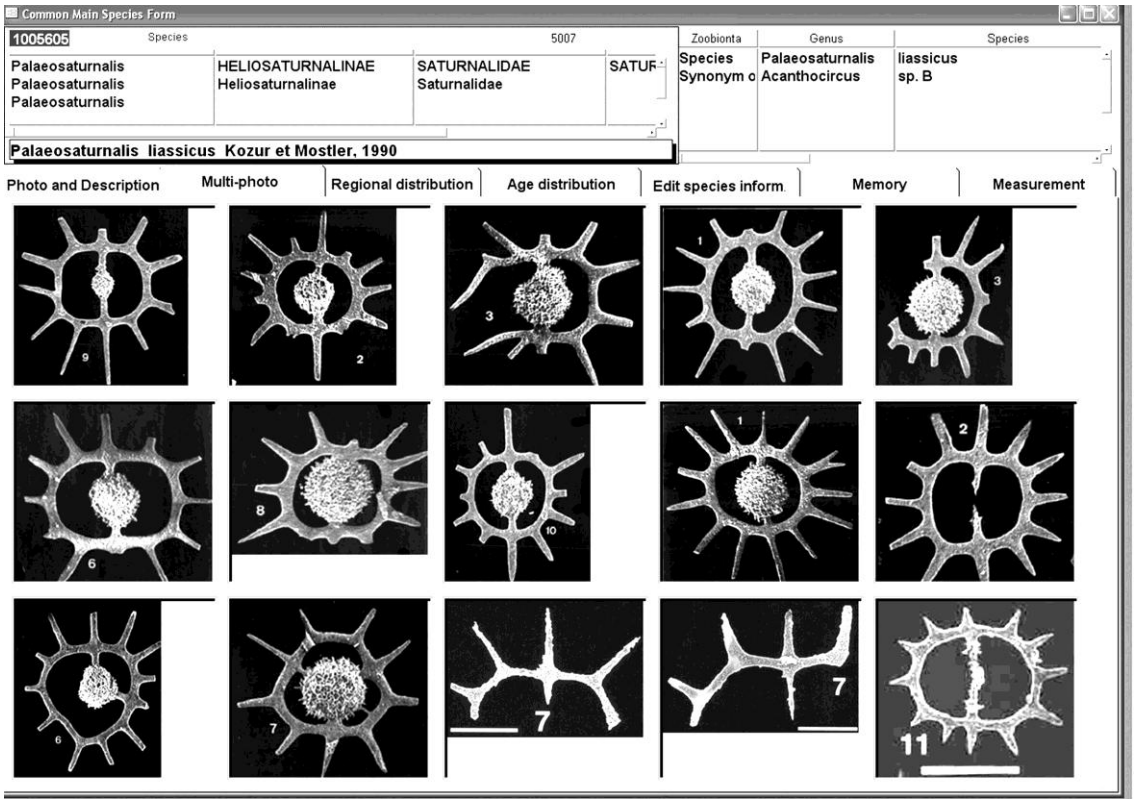
Для обзора нескольких изображений синонима или вида существует страница «Multiphoto» (рис. 4). На ней можно просмотреть одновременно по восемь изображений или перенастроить ее на любое количество одновременно просматриваемых изображений в зависимости от характеристики используемых аппаратных средств. На описываемой странице, как и на предыдущей все



Р и с. 2. Основная форма. Страница Edit species information



Р и с. 3. Основная форма. Страница Edit species information



Р и с. 4. Основная форма. Страница Multiphoto

просматриваемые изображения приведены к одному формату. В специальных полях таблицы изображений имеются эти же изображения, приведенные в зависимости от качества информации в первоисточнике к двум стандартизированным авторами масштабам, позволяющим просматривать мелкие и крупные формы микроорганизмов. Подобная система предоставления графической информации позволяет оценить не только морфологические особенности экземпляров, но и визуально оценить характер изменения их размеров особенно при анализе географической распространенности и вынесении изображений на картографическую основу.

Экологический блок существенно различается по набору анализируемых факторов в зависимости от уровня лабильности организмов различных таксонов высокого ранга к среде обитания. Наиболее широко представлены сведения по диатомовым водорослям, распространенным во всех биотопах и имеющих широкий спектр организмов по солености, температуры и др. Главными экологическими параметрами, учитываемыми при внесении в ИС являются: среда обитания (водные, почвенные, аэральные и др.); подвижность (планктон, бентос с детализацией на 3 уровня, определяющей объекты обрастания, субстрат); соленость; галобность; температурный режим; региональные флористические и зоологические провинции (широты, области); отношение к течениям (реофильные, лимнофильные); апвеллинговые формы; приливные формы; сопробность (нутриентность); Ph-среды; вегетационный период.

Блок пространственно-возрастной распространенности позволяет хранить сведения о таксонах в любой удобной пользователю системе привязки к местности (административная тектоническая и др.), включая географические координаты, связывающие ИС с электронными картами. Организация информации построена по принципу биологический вид – образец, где для образца приведена вся исчерпывающая информация, включая уровни достоверности определения вида, возраста, географического положения точки. Возрастная кодировка производится в различных вариантах (ярусы,

свиты, зоны) и не ограничивает пользователя (возможна кодировка возрастной распространенности группы по зонам других ископаемых групп, многократное альтернативное кодирование и др.). Стратиграфический словарь включает более 5 тыс. стратиграфических терминов, а сам блок данных более 200 тыс. записей. Для приведения возрастных данных к единой системе времени разработан алгоритм пересчета на стандартные шкалы времени или на различные системы по стратиграфическим зонам.

Библиографический блок в биологическом и палеонтологическом анализе в основном используется для передачи данных об источниках информации, однако он может работать как самостоятельная библиографическая информационная система, насчитывающая более 11 тыс. работ и представляющая наиболее полную русскоязычную библиографию по радиоляриям (более 5,5 тыс. работ). Данный блок помимо стандартного описания источника информации на языке оригинала содержит реферат источника, а иногда и сам полнотекстовый источник со всеми имеющимися в нем иллюстрациями. Поисковая система данного блока позволяет находить публикации и рукописные источники практически по любому параметру (трехуровневая тематика публикации, автор и год издания, вид издания, язык, описываемые в работе классы организмов, геологический возраст, четырехуровневая рубрикация географического охвата работы, петрографическая и минералогическая характеристика объектов и многое другое.

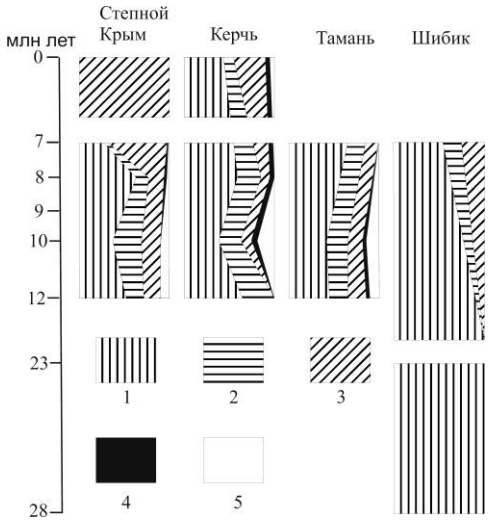
Использование ИС дает возможность автоматически строить послойные карты ареалов распространенности видов на любые возрастные интервалы и любого масштаба с последующим определением путей миграции и зон дивергенции. Кроме того, возможно оценивать геологический возраст образцов по комплексу видов, проводить диагностику таксонов по совокупности морфологических признаков, анализировать экологические условия среды обитания, анализировать видовое богатство таксонов любого уровня с детальностью от эратемы до подвека с учетом любых дополнительных условий (морфологических элементов и параметров скелетов, экологических особенностей и др.). В расчетах обрабатывается информация, основанная на анализе более 4,5 тыс. источников и более 200 тыс. определений (вид/образец), и постоянно обновляется новыми сведениями, публикуемыми по радиоляриям и другим, доступным в ИС классам организмов. При анализе скорости видо- и родообразования используются данные различных, на усмотрение пользователя, глобальных и региональных хроностратиграфических схем о длительности основных подразделений [5 – 7].

Имеющиеся в ИС сведения позволяют проследить изменения факторов среды в Паратетисе на протяжении от позднего олигоцена до современности. Анализ проведен, как по конкретным разрезам в районах Степного Крыма, Керчи, Таманского полуострова и бассейна р. Шибик, где нами обнаружены выходы пород, содержащих богатый комплекс диатомовых водорослей в отделениях хатского яруса, среднесарматского, позднесарматского и мэотического региоярусов, так и по бассейну в целом.

В результате полученных путем обработки данных установлено, что в олигоценовое время преобладают (66%) планктонные формы и незначительно встречаются бентосные (9%) и эпифитные (25%). Данное соотношение говорит о шельфовом бассейне с глубинами до 200 м. В сарматское время происходит резкое сокращение планктона и основная масса диатомей представлена бентосными (40 – 57%) и эпифитными (31 – 38%) жизненными формами, планктонные всего 5 – 15%. Такое соотношение видов как среди единично представленных особей, так и более часто встречающихся колониальных форм, указывает на мелководный бассейн с богатой водной растительностью. В мэотическом бассейне роль бентоса и эпифитов остается прежней (89%), что позволяет сделать вывод о сохранении пышной водной растительности небольших глубин (до 15 м).

Изменение солености можно прослеживать по соотношению галобиотопов (рис. 5). Они показывают, что мор-

ской бассейн на ранней стадии имел нормальную соленость. С начала среднего миоцена происходит опреснение водоема и существенную роль начинают играть солоноватоводно-морские диатомовые, а затем и солоноватоводные, характерные для солености порядка 10 ‰. Аналогичное соотношение существует и сегодня в открытой части Азовского моря.



Р и с. 5. Изменение соотношения галобиотопов диатомовых водорослей в Паратетисе. Условные обозначения: 1 – морской, 2 – солоноватоводно-морской, 3 – солоноватоводный, 4 – пресноводно-солоноватоводный, 5 – пресноводный

Температурный режим бассейна на протяжении всего времени оставался практически стабильным. В позднем олигоцене преобладали умеренные формы (80%) и лишь 20% – тепловодные. В миоцене сокращается количество тепловодных видов и появляются виды холодноводные (10%). Такое соотношение указывает на температурный режим оптимальный для бурых водорослей со среднегодовыми температурами порядка 14ºС. Это подтверждается и независимыми данными по изотопам кислорода в створках диатомовых водорослей, извлеченных из диатомитов Таманского полуострова.

Характер течений во время описываемого периода изменялся в сторону увеличения приливных явлений. На ранних этапах приливофобы составляли до 5%. При этом наблюдается и постепенное увеличение индикаторов апвеллинга, если в начале описываемого периода они отсутствовали, то в последующее время начинают появляться единичные представители, которые далее становятся типичными. На этом фоне происходит и увеличение лимнофильных форм. Это вероятно связано не с изоляцией бассейна, а с формированием извилистой береговой линии при активном тектоническом режиме.

Заключение. Разработанная информационная система позволяет хранить широкий спектр сведений и использовать их при анализе изменений экологических факторов не только для крупных интервалов времени, но и при сезонных и даже суточных изменениях в зависимости от интервалов наблюдений.

СПИСОК ЛИТЕРАТУРЫ

1. *Catalogue* of life:http://www.catalogue-oflife.org (дата обращения: 01.10.2015)
2. *Агарков Ю.В.* Информационная система по радиоляриям // Тез. Докл. Десятого семинара по радиоляриям. – М.: Ин-т литосферы РАН, 1996. – С. 6.
3. *Агарков Ю.В.* Новые данные о видовом разнообразии радиолярий мезозоя // Матер. 11 Семинара по радиоляриям (19–24 июня 2000 г.). – СПб. –М., 2000. – С. 9–10.
4. *Agarkov Yu.V.* Universal Biological Information System “Biobase’ // Proceedings of the Sixth International Conference Environmental Micropaleontology, Microbiology and Meiobenthology. – Moscow, 2011. – P. 18–21.
5. *Афанасьев С.Л.* Изотопная геохронологическая шкала венда – фанерозоя // Геол. и геофиз. – 1993. – 34, № 3. – С. 3–9.
6. *Афанасьев С.Л.* Наноциклитная геохронологическая шкала венда – фанерозоя // Бюл. Моск. о-ва испыт. природы. Отд. геол. 1993. – Т. 68, № 2. – С. 24–30.
7. *The Geologic* Time Scale 2012 / ed. F.M.. Gradstein, J.G. Ogg, M. D. Schmitz, G.M. Ogg. – Amsterdam, London, New York, Oxford. Tokyo. 2012. – 1160 p.