

**ПРИМЕНЕНИЕ МАТЕМАТИЧЕСКОГО МОДЕЛИРОВАНИЯ
ДЛЯ ОЦЕНКИ ВЛИЯНИЯ ИЗМЕНЕНИЯ ЧИСЛЕННОСТИ НА
ГЕНЕТИЧЕСКОЕ РАЗНООБРАЗИЕ ВОДНЫХ ОРГАНИЗМОВ**

В.В. Горбачев

*Институт биологических проблем севера ДВО РАН, ул. Портовая, 18, Магадан,
685000, Россия. E-mail: genetic2@yandex.ru*

В работе проводится верификация возможности применения математического моделирования для оценивания молекулярно-генетических параметров популяций. Оценено влияние демографических параметров на общую топологию медианных графов (дендрограмм). Было показано, что даже применение сильно отличных друг от друга демографических сценариев может приводить к получению сходных значений и, как следствие, к появлению возможных статистических ошибок при оценивании результатов.

**APPLICATION OF MATHEMATICAL MODELING FOR THE
ESTIMATION OF EFFECTS OF VARIATION IN NUMBER ON
GENETIC DIVERSITY OF AQUATIC ORGANISMS**

V.V. Gorbachev

*Institution of Biological Problems of the North FEB RAS, 18 Portovaya Str., Magadan,
685000, Russia. E-mail: genetic2@yandex.ru*

An attempt to verify the potential usage of mathematical simulation for the estimation of molecular-genetic parameters of population was undertaken. The influence of demographic parameters on general topology of median graphs (dendrograms) was evaluated. It was demonstrated that the employment of different demographical scenarios lead to similar results and, as consequence, to occurrence of statistical errors in the evaluation of data obtained.

Математическое моделирование является одним из самых распространенных методов научного анализа, в настоящее время преобладающая часть исследователей (около 60%) применяет его в своих работах (Otto, Day, 2007). Несмотря на то, что данный метод находит все больше применение во многих биологических дисциплинах (Othmer et al., 1997), все же, к настоящему времени, научный потенциал данного подхода во многом недооценен. Несмотря на критику математического моделирования «можно утверждать, что информация даже об относительно небольшом числе переменных служит достаточной основой для построения эффективных моделей» (Одум, 1975).

На сегодняшний день разработка математического аппарата позволила с успехом применять моделирование для экологических и генетических исследований при изучении природных и искусственных популяций организмов. Редукционизм применяемых моделей позволяет проследить зависимости между параметрами описанными математически а, следовательно, и оценивать влияние переменных друг от друга.

Одно из ведущих позиций в исследовании генетических параметров популяций приобрели методы коалесценции. Они основаны на ретроспективном подходе в изучении исторического развития популяций, позволяющего отслеживать изменение молекулярно-генетических параметров и их взаимосвязь с историческим развитием популяции (Wakeley, 2009). В противоположность коалесцентным моделям, так же разработаны методы моделирования в будущее – «forward simulations». Между собой эти модели, различающиеся скорее не математическим подходом, а направленностью вектора моделирования, позволяя делать прогнозирование популяционно-генетических параметров, что может быть применено в целях мониторинга промысловых популяций.

Цель, которая ставилась перед данным исследованием – оценить влияние демографических параметров популяций на ее молекулярно-генетическое разнообразие.

Моделирование выполнялось нами с применением программы Network 4.5.1.6 (Bandelt et al., 1999) всего нами верифицировались 4 варианта: модель «горлышка бутылки» (вариант 1), модель демографической экспансии (вариант 2), модель, предполагающая константную численность (вариант 3), модель флуктуации численности (вариант 4). Графики, получаемые при моделировании, представлены на рис. 1.

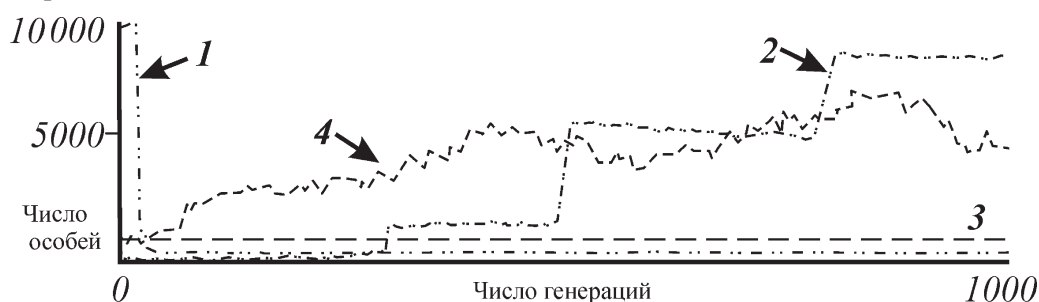


Рис.1. График изменения численности моделируемых популяций.
Примечание: цифрами обозначены модели (см. в тексте)

Последняя модель предполагала, помимо изменения величины размера популяции, наличия еще и более высоких темпов мутации (в 4 раза) тем самым предполагался сценарий наличия горячих сайтов. Оценка молекулярных параметров выполнялась с применением программы Arlequin 3.0 (Excoffier et al, 2005). При оценке влияния численности популяции на топологию медианного графа моделирование повторялись в трипликате. Для оценки зависимости появления вероятного числа гаплотипов от эффективной численности (далее N_e), моделирование проводилось по 106 раз с заданной величиной N_e . Средняя продолжительность жизни одного поколения нами была принята за 10 лет, при 1000 учитываемых генераций.

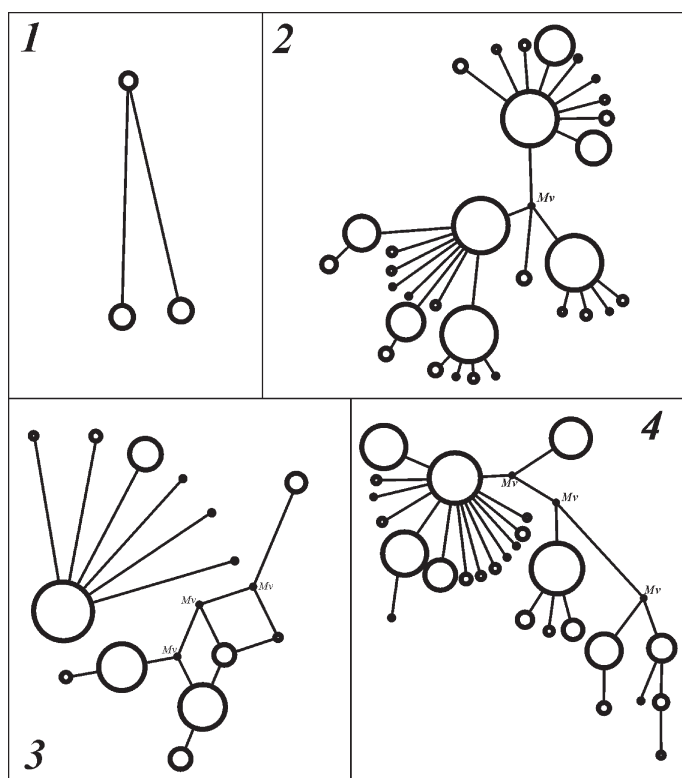


Рис. 2 Медианные графы, описывающие характер топологии от выбранных моделей.

Примечание: Mv- медианные векторы, описание моделей в тексте

Максимальное число N_e до 10000 особей. Получаемые частоты далее использовались для реконструкции общей топологии медианного графа (рис.2), расчета молекулярных параметров (таблица). Для определения вероятности появления гаплотипов в популяции использовалась 3-я модель, так как единственным фактором в ней, влияющим на частоты генов является дрейф генов, что позволило более точно верифицировать получаемые значения.

Первая модель предполагает, сценарий «горлышка бутылки», указывая, что, даже при наличии изначальной высокой численности в ранние периоды существования популяции, от

Таблица
Значение молекулярных параметров для популяций полученных в ходе моделирования

Параметр	Модель			
	1	2	3	4
θ_s (темпы появления единичных мутаций)	$0,38 \pm 0,2$	$1,5 \pm 0,4$	$0,69 \pm 0,3$	$1,19 \pm 0,4$
π (среднее нуклеотидное разнообразие)	$0,97 \pm 0,7$	$1,48 \pm 0,9$	$0,59 \pm 0,5$	$0,97 \pm 0,7$

Примечание. Молекулярные параметры рассчитывались по: Tajima, 1989.

накопленного генетического разнообразия остается считанное количество гаплотипических вариантов, что сильно обедняет генетическую структуру и сказывается на общей топологии графа (рис.2, под цифрой 1). Медианная сеть в данном случае склонна к сильной редукции, стоит отметить, что оцениваемые молекулярно-генетические параметры так же стремятся к уменьшению, при этом отклоне-

ние (далее σ), может, принимать величину сопоставимую с основным значением или превышать ее (таблица).

Вторая модель подразумевает несколько всплесков численности, что так же сказывается на изменении общей топологии медианного графа. Характерной чертой получаемых при этой модели данных является наличие выраженных «звездообразных» кластеров (star-like) (рис.2, под цифрой 2). Данный вид кластеров близок к политомии, при которой ветви отпочковываются из одного центра, при этом внутренние ветви имеют тенденцию к укорочению по сравнению с внешними (Лукашев, 2009). Молекулярные параметры, оценивающие темпы появления единичных сайтов (θ_s) в данной модели склонны к увеличению и принимает значение, сопоставимое со средним нуклеотидным разнообразием π (таблица).

Третья модель подразумевает наличие константной численности, а, так как гаплотипический состав при соблюдении данной модели не может измениться за счет увеличения \ уменьшения числа особей, то основным фактором изменчивости, как отмечалось выше, являются случайные процессы. Получаемые топологии при данной модели могут сильно отличаться, принимая различный вид: от «звездообразных» до топологий изобилующих большим числом медианных векторов (Mv – указывающих на не экономный характер топологии) и многомерных кубов (наличие которых можно расценивать как присутствие гомоплазий) (рис.2, под цифрой 3). Подобная топология появляется вследствие исчезновения промежуточных гаплотипических вариантов как результат влияния дрейфа генов. Подобные предположения находят свое подтверждение и в молекулярных параметрах, которые стремятся, как и в 1 модели, к низкому значению, при наличии, как и в 1 модели, большого значения σ (стандартного отклонения), которое получается при достаточно сильных флуктуациях частот признаков, что согласуется с ранее обоснованной теорией нейтральности (Кимура, 1985).

Четвертый вариант характеризуется наличием кластеров с большой длиной ветвей и наличием «горячих сайтов», т.е. нуклеотидов, мутирующих в несколько раз чаще (рис.2, под цифрой 4). И хотя в данной модели, как и во 3-ем вариан-

те, могут присутствовать медианные векторы, все же влияние дрейфа генов при образовании многомерных структур несопоставимо выше.

При оценке влияния демографических параметров на генетическую структуру (гаплотипическое разнообразие) (рис.3) можно увидеть, что число гаплотипов не находится в прямой зависимости от параметра N_e . Эффективная численность скорее влияет на разброс значений, нежели на абсолютное число возникающих гаплотипов.

Таким образом, при оценке влияния численности на генетическую изменчивость популяции необ-

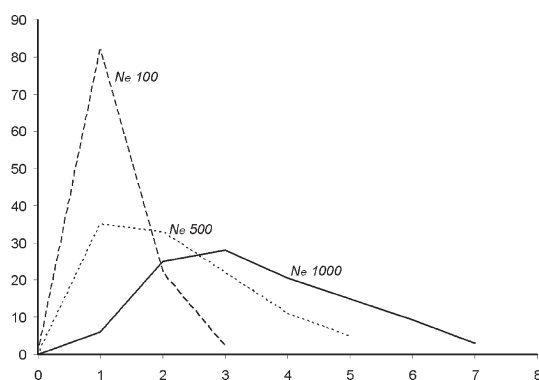


Рис. 3. Ожидаемое число гаплотипических вариантов для популяций с заданной величиной эффективной численности (N_e).

Примечание: на оси абсцисс – число гаплотипов, на оси ординат – вероятность появления.

ходимо учитывать несколько параметров, так как сама по себе топология медианного графа (или дендрограммы) или численные соотношения гаплотипических вариантов (оценка генетической структуры популяции) могут способствовать появлению статистических ошибок как I, так и II рода, за счет близости числовых значений. В связи с этим можно предположить, что проведение даже нескольких статистических тестов не гарантирует правильность оценки наблюдаемых результатов и всякий раз оставляет «место для сомнений».

ЛИТЕРАТУРА

- Кимура М. 1985.** Молекулярная эволюция: теория нейтральности. М.: Мир. 394 с.
- Лукашев В.В. 2009.** Молекулярная эволюция и филогенетический анализ. М.: БИНОМ. Лаборатория знаний. 256 с.
- Одум Е. 1975.** Основы экологии. М.: Высш. шк. 740 с.
- Bandelt H.J., Forster P., Röhl A. 1999.** Median-joining networks for inferring intraspecific phylogenies // *Mol. Biol. Evol.* V. 16. P. 37–48.
- Excoffier L., Laval G., Schneider S. 2005.** Arlequin ver. 3.0: An integrated software package for population genetics data analysis // *Evol. Bioinform. Online.* V. 1. P. 47–50.
- Othmer H. G. Adler F. R. Lewis M. A. Dallon J. C. 1997.** Case Studies in Mathematical Modeling—Ecology, Physiology, and Cell Biology. Utah: Prentice-Hall. 415 p.
- Otto S. P., Day T. 2007.** A Biologist's Guide to Mathematical Modeling in Ecology and Evolution. Princeton University Press. 727 p.
- Tajima F. 1989.** Statistical method for testing the neutral mutation hypothesis by DNA polymorphism // *Genetics.* V. 123. P. 585–595.
- Wakeley J. 2009.** Coalescent theory: An introduction. Colorado: Roberts & Co. Publishers, Greenwood Village. 326 p.