

УДК 595.771:004
ББК 28.691.892.49
К 14

Казakov М.А.

Младший научный сотрудник отдела Интеллектуализации информационных и управляющих систем Института прикладной математики и автоматизации, Нальчик, e-mail: f_wolfgang@mail.ru

Кармоков М.Х.

Кандидат биологических наук, старший научный сотрудник лаборатории разнообразия беспозвоночных Института экологии горных территорий им. А.К. Темботова Кабардино-Балкарского научного центра РАН, Нальчик, e-mail: lacedemon@rambler.ru

Способ автоматизации подсчета генетического сходства и генетической дистанции по критерию Нея у комаров-звонцов рода *Chironomus* (Рецензирована)

Аннотация. Представлены результаты исследования по разработке способа автоматизации вычисления генетического сходства и генетической дистанции по критерию Нея у комаров-звонцов рода *Chironomus*. Описан разработанный программный продукт с простым, интуитивно понятным интерфейсом, позволяющим легко и быстро вычислять генетическое сходство и дистанции.

Ключевые слова: автоматизация вычисления, генетическая дистанция, генетическое сходство, хирономиды, *Chironomus*, политенные хромосомы.

Kazakov M.A.

Junior Researcher at the Department of Intellectualization of Information and Control Systems of the Institute of Applied Mathematics and Automation, Nalchik, e-mail: f_wolfgang@mail.ru

Karmokov M.Kh.

Candidate of Biology, Senior Researcher of Laboratory of Invertebrate Diversity, A.K. Tembotov Institute of Ecology of Mountain Territories of Kabardino-Balkarian Scientific Center of the Russian Academy of Sciences, Nalchik, e-mail: lacedemon@rambler.ru

The method to automate calculation of genetic similarity and genetic distance in chironomids of the genus *Chironomus* by the Nei criterion

Abstract. The results of studies on the development of a way to automate the calculation of genetic similarity and genetic distance in chironomids of the genus *Chironomus* by the Nei criteria are presented. A program product with a simple, intuitive interface allowing quick and easy calculation of the genetic distances and similarity is described.

Keywords: automation of computing, genetic distance, genetic similarity, chironomids, *Chironomus*, polytene chromosomes.

Введение

Виды комаров-звонцов рода *Chironomus* (Chironomidae, Diptera) являются удобной моделью для реконструкции закономерностей хромосомной дивергенции внутри популяций одного вида, а также близких видов одного рода в ходе эволюции. При этом особое значение имеет вычисление генетических расстояний и генетического сходства между популяциями одного вида. Наиболее популярным методом измерения этих величин у исследователей рода *Chironomus* является метод Нея [1]. Этот метод определения генетического расстояния и генетического сходства основан на предположении того, что генетическая разница возникает за счет мутаций и генетического дрейфа.

Согласно неodarвинской концепции, новый вид возникает в результате дифференциации любой популяции, относящейся к этому виду; иначе говоря, происходит постепенная дивергенция на уровне генофонда в этой популяции и вида в целом. Постепенная замена одних аллелей определенных генов на другие и лежит в основе этой дивергенции. Во время дивергенции двух популяций все меньшее число генов имеет аллели, встречающиеся в обеих популяциях. На завершающей стадии процесса видообразования для всех генов остаются аллели, характерные только для одной из двух попу-

ляций. При этом генетическое сходство становится равным нулю, а новый вид становится отличным от начального.

На практике для расчетов генетического расстояния у *Chironomus* определяют частоты встречаемости в каждой популяции различных последовательностей дисков (аллелей) в семи плечах политенных хромосом (А, В, С, D, E, F и G). Представим себе, что для хромосомного плеча F существует i различных аллелей (1, 2, 3 и т.д.). В популяции **A** эти аллели встречаются с частотами a_i ; в популяции **B** частотам аллелей соответствуют b_i . Исходя из этого, можно вычислить степень генетического сходства данных популяций по данному плечу:

$$I = \frac{\sum a_i b_i}{\sqrt{\sum a_i^2 \sum b_i^2}}.$$

Само же генетическое расстояние D вычисляется по формуле:

$$D = -\ln I.$$

Значения генетического сходства могут варьировать от 0 (между популяциями нет общих аллелей) до 1 (частоты аллелей в популяции одинаковые), соответственно генетическая дистанция варьирует от 0 до ∞ .

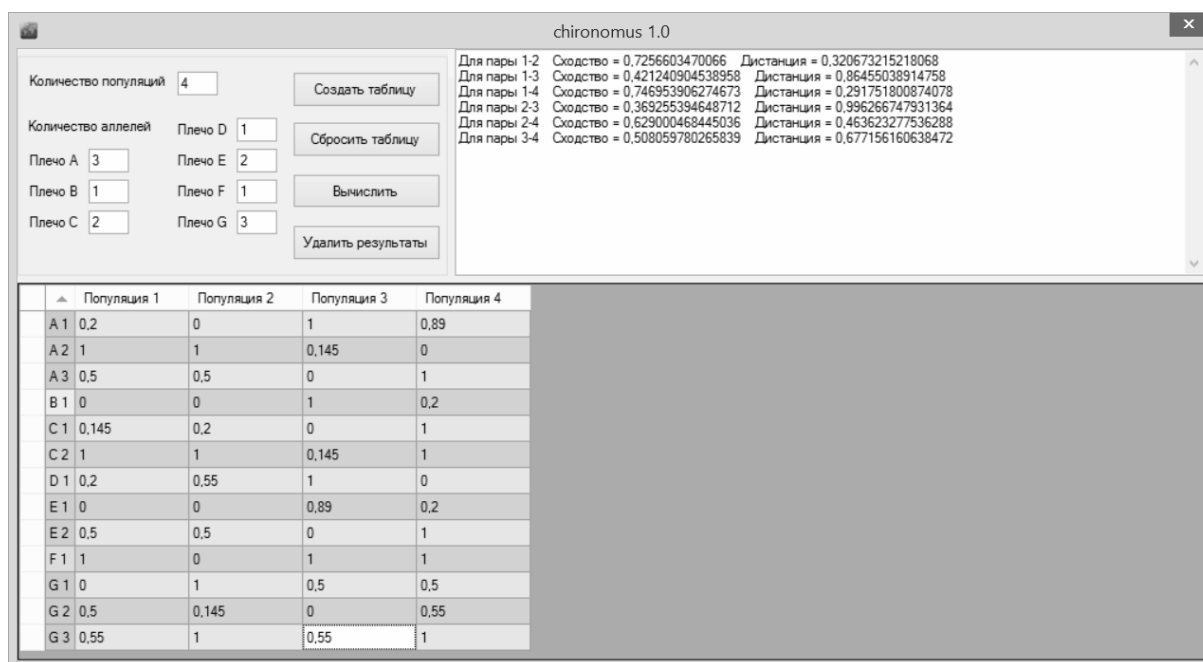
Существует несколько программных продуктов, позволяющих кроме всего прочего вычислять также цитогенетическое расстояние и цитогенетическое сходство по критерию Нея. Наиболее популярные из них это пакеты программ PHYLIP [2] и MEGA 6 [3-5]. В первую очередь данные продукты предназначены для работы с нуклеотидными последовательностями ДНК и последовательностями аминокислот белков.

Вычисление генетических расстояний по критерию Нея у рода *Chironomus* возможно с помощью выше указанных пакетов программ. Однако в силу сложности процесса предварительной подготовки исходных данных для ввода и сложного, не интуитивного интерфейса рассмотренные выше программные продукты не являются достаточно удобными вычислительными средствами для решения поставленной задачи. В силу чего возникает потребность в создании программного продукта, отвечающего следующим критериям:

- простота ввода исходных данных, не требующая предварительной многошаговой обработки;
- интерфейс программы должен быть достаточно простым и интуитивно понятным;
- наглядность получаемых в результате вычислений;
- повышенная точность вычислений;
- возможность при необходимости добавления дополнительных методов вычислений генетической дистанции и генетического сходства;
- возможность одновременного сравнения большого числа популяций.

Результаты исследования

В результате исследования был создан специализированный программный продукт *Chironomus 1.0* (рис. 1), позволяющий быстро и наглядно вычислять генетические расстояния и генетическое сходство у рода *Chironomus* по методу Нея. Программа имеет простой русскоязычный интерфейс и позволяет одновременно сравнивать до 14 различных популяций, написана на языке программирования C# и имеет размер 120 Кбайт. Программа содержит 46 циклов и работает как приложение для платформы .NET Framework.

Рис 1. Интерфейс программного продукта *Chironomus 1.0*

Процесс ввода данных и получение результата состоит из следующих шагов:

1. Вводится количество сравниваемых популяций и число аллелей для каждого хромосомного плеча. Эти значения являются первичными входными данными алгоритма программы;

2. В соответствии с введенными значениями создается наглядная таблица для ввода частот аллелей популяций. Заполнение этой таблицы определяет вторичные входные данные алгоритма программы;

3. После заполнения всей таблицы производится вычисление генетического сходства и генетической дистанции для каждой пары популяций, после чего производится автоматический вывод в поле результатов.

Обсуждение результатов

Программа была протестирована на данных по шести реально существующим популяциям вида *Ch. bernensis* Центрального Кавказа и Предкавказья [6]. Независимые расчеты генетических расстояний и генетических дистанции по тем же исходным данным подтвердили точность вычислений требуемых значений созданным нами программным продуктом.

Таким образом, в результате проведенной работы удалось создать небольшой специализированный для исследователей рода *Chironomus* программный продукт с простым, интуитивно понятным интерфейсом, позволяющим легко и быстро вычислять генетическое сходство и генетические дистанции по критерию Нея. В ближайшем будущем нами планируется расширение программных возможностей нашего продукта и его размещение в открытом доступе в сети Интернет, для того чтобы все желающие могли воспользоваться разработанным программным продуктом.

Примечания:

1. Nei M. The genetic distance between populations // American Naturalist. 1972. Vol. 106. P. 283-290.

References:

1. Nei M. The genetic distance between populations // American Naturalist. 1972. Vol. 106. P. 283-290.

2. Felsenstein J. Phylogeny Inference Package (PHYLIP) // Version 3.5. 1993. University of Washington, Seattle.
 3. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods / K. Tamura, D. Peterson, N. Peterson, G. Stecher // *Molecular Biology and Evolution* 2011. Vol. 28. P. 2731-2739.
 4. MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0 / K. Tamura, G. Stecher, D. Peterson, A. Filipski // *Molecular Biology and Evolution*. 2013. Vol. 30. P. 2725-2729.
 5. MEGA-CC: computing core of molecular evolutionary genetics analysis program for automated and iterative data analysis / S. Kumar, G. Stecher, D. Peterson, K. Tamura // *Bioinformatics*. 2012. Vol. 28. P. 2685-2686.
 6. Karmokov M.Kh., Polukonova N.V., Sinichkina O.V. Karyotype characteristics and polymorphism peculiarities of *Chironomus bernensis* Wülker & Klötzli, 1973 (Diptera, Chironomidae) from the Central Caucasus and Ciscaucasia // *Comparative Cytogenetics*. 2015. Vol. 9 (3). P. 281-297. doi: 10.3897/CompCytogen.v9i3.4519
2. Felsenstein J. Phylogeny Inference Package (PHYLIP) // Version 3.5. 1993. University of Washington, Seattle.
 3. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods / K. Tamura, D. Peterson, N. Peterson, G. Stecher // *Molecular Biology and Evolution* 2011. Vol. 28. P. 2731-2739.
 4. MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0 / K. Tamura, G. Stecher, D. Peterson, A. Filipski // *Molecular Biology and Evolution*. 2013. Vol. 30. P. 2725-2729.
 5. MEGA-CC: computing core of molecular evolutionary genetics analysis program for automated and iterative data analysis / S. Kumar, G. Stecher, D. Peterson, K. Tamura // *Bioinformatics*. 2012. Vol. 28. P. 2685-2686.
 6. Karmokov M.Kh., Polukonova N.V., Sinichkina O.V. Karyotype characteristics and polymorphism peculiarities of *Chironomus bernensis* Wülker & Klötzli, 1973 (Diptera, Chironomidae) from the Central Caucasus and Ciscaucasia // *Comparative Cytogenetics*. 2015. Vol. 9 (3). P. 281-297. doi: 10.3897/CompCytogen.v9i3.4519