

*Симонов Сергей Александрович,  
Макунин Алексей Игоревич,  
Stephen J. O'Brien*

УДК 57.087:004.42

## БРАУЗЕР ДЛЯ ГЕНОМА<sup>1</sup>

### Аннотация

В статье кратко описываются задачи, которые возможно решать с помощью компьютерного инструмента – геномного браузера как средства для визуализации последовательностей нуклеотидов в геномах и расположения в них геномных элементов. Приводится краткий обзор имеющихся подходов и проводится аналогия с браузерами в Интернете, открывшими новую эру в его развитии.

На примере распространенного геномного браузера – GBrowse – описывается как подготовить и увидеть элементы в геноме, как самостоятельно провести компьютерные эксперименты.

**Ключевые слова:** геном, генетические элементы, визуализация, браузер.

### ВВЕДЕНИЕ

Расшифровка нуклеотидных последовательностей геномов организмов в конце XX века открыла новую эру в генетике. В действительности исследователи получили возможность взглянуть на последовательность букв, так называемых «пар оснований» или «нуклеотидных пар», кодирующих саму жизнь. Но что они означают? Прочтение этих последовательностей, основанное на гигантских достижениях науки и техники современности можно сравнить с обнаружением древнего письма на незнакомом языке. Расшифровка этого языка, попытки понять смысл, а также правила и грамматика этого языка до сих пор остаются неясны в полной мере.

Многие функции геномных элементов уже известны. Но далеко не все. А самое главное, в геноме взаимодействие различных

элементов может играть решающую роль. Так работа некоторых генов по воспроизводству себе подобных и синтезу белков – самых важных элементов жизни (а функции белков – это не только строительный материал для организма, это еще и регуляционные механизмы, включая гормоны и ферменты, а также многие другие), зачастую может зависеть от ближайшего окружения – такого окружения, которое может располагаться на удалении тысяч и тысяч нуклеотидных пар от элемента, на который они оказывают влияние.

Возможность посмотреть на огромные массивы данных – часто возникающая потребность в науке. Зачастую найти какие-то закономерности возможно только при ком-

<sup>1</sup> Работа выполнена при поддержке гранта Правительства Российской Федерации для государственной поддержки научных исследований, проводимых под руководством ведущих учёных в российских образовательных учреждениях высшего профессионального образования. Номер гранта 11.G34.31.0068 под руководством Стефена О'Брайена.

© Симонов С.А., Макунин А.И.,  
Stephen J. O'Brien, 2013

бинировании взгляда с высоты «птичьего полета» и пристального рассмотрения деталей при большом увеличении. Подобные задачи решаются, например, при распознавании образов (изображений, звуков, других сигналов). Но, в отличие от традиционных задач, где можно «усреднить» некоторые значения без потери качества общей картины, в генетике каждый элемент играет важную роль, и замена только одной буквы в трехмиллиардной последовательности генома может сыграть решающую, даже фатальную роль.

Попытки построения образов геномов, карт в виде разноцветных последовательностей и меток приводят к картинкам необъятных размеров и практической невозможности охватить взглядом всю сущность.

Возможности современных вычислительных средств и средств компьютерной графики позволяют увидеть то, что скрыто за деталями, и в то же время не пропустить мельчайших подробностей. Такие средства получили название геномные браузеры, по аналогии со средствами просмотра информации в Интернете.

В статье кратко описываются задачи, которые можно решать с помощью этого инструмента – геномного браузера. Приводится небольшой обзор имеющихся подходов и проводится аналогия с браузерами в Интернете, открывшими новую эру в его развитии.

На примере распространенного геномного браузера – GBrowse – описывается как подготовить и увидеть новые элементы в геноме. Приводятся примеры и описание, что и как каждый человек может увидеть и посмотреть, пользуясь абсолютно открытыми и доступными средствами и доступом в Интернет.

Статья ориентирована на неподготовленного читателя, имеющего базовые (школьные) знания в биологии и элементарные навыки работы в Интернет.

В жизни часто приходится сталкиваться с большим количеством данных. Структурирование и представление данных в виде, наиболее приемлемом для человеческого восприятия, – задача программных средств под общим названием «БРАУЗЕР».

Изначально английское слово *to browse* означало осматривание (витрин). И, соответственно, глазающие на витрины покупатели именовались браузерами.

С появлением Интернета появилась возможность предоставления для широчайшего круга пользователей огромного количества информации, объем которой продолжает расти с головокружительной скоростью. Средства для просмотра информации получили название браузеры (более точно – Интернет браузеры).

Первоначально, функция браузера в Интернете сводилась к правильному отображению так называемых ГИПЕРТЕКСТОВ, то

---

## БРАУЗЕРЫ В ИНТЕРНЕТЕ

Вечный спор: какой автомобиль лучше (для мужчин) или какой производитель косметики выпускает более натуральную косметику (для женской аудитории) в последнее время часто замещается вопросом какой браузер для просмотра Интернета лучше. Причем для начинающих пользователей вопрос не стоит – Интернет Эксплорер от Майкрософт. Почему? А просто потому, что он уже установлен на каждом компьютере вместе с операционной системой Windows. Лучший ли он? Любой профессионал скажет: «Нет!» И единственное, для чего его надо использовать, – это для скачивания других браузеров. Например, FireFox (Лисы в русском сленге) или Chrome (Хром – от Google).

А Вы думаете, они сильно отличаются друг от друга? Есть всего несколько платформ, на базе которых построены все браузеры (а их, по крайней мере, сотни). Это как разница между Ауди и Фольксвагеном, у которых ручки дверей расположены по-разному, в то же время все остальное может быть одинаково.

Совершенно аналогично и с браузерами. Функция автомобиля – перевозить. Функция браузера – отображать информацию. А то, что они различны, – так на то и рынок. Все мы разные и спорить о вкусах – терять время...

есть текстов, в которых некоторые слова представляли собой ссылки на более подробное их определение. Постепенно, шаг за шагом, функции браузеров пополнялись новыми возможностями для представления разнообразной информации во всех ее проявлениях: картинок, схем, видео и аудио информации, а также средствами обратной связи.

Другой очевидной областью применения идеи браузеров является генетика, а именно анализ генома – многомиллиардной последовательности нуклеотидов, которая повторяется в каждой клеточке каждого живого существа на Земле. Например, в организме человека примерно  $10^{12}$  клеток, и в каждой из них хранятся почти абсолютно идентичные последовательности, по сути, хранятся в закодированном виде весь организм. Длина этой последовательности у человека составляет примерно  $3,2 \cdot 10^9$  «буковок» из четырехбуквенного алфавита {A, C, G, T}. Вроде бы немного – чуть больше тысячи томов объемом с «Война и Мир» каждый в напечатанном виде... Но если посмотреть на количество информации, то даже 300 буковок достаточно для кодирования числа вариантов, сопоставимых с количеством атомов во вселенной<sup>1</sup>... И уж заведомо меньше числа «гуголь» – так шутники-ученые назвали число со ста нулями –  $10^{100}$ , от которого и произошло имя, которым все мы пользуемся ежедневно (сравните googol и Google).

К тому же наши знания о составе генома все еще незначительны, фрагментарны и расширяются буквально каждый день (тут под словом НАШИ, имеются в виду знания всего человечества – всех ученых в мире, работающих в этой области).

Но результат сиквенирования – это еще не значит сборка всего генома. Это просто ряд фрагментов, на которые почти случайным образом полный геном был разбит перед началом эксперимента специальными ферментами. Без указания места каждого фрагмента и связки с другими фрагментами. Расположение фрагментов на геноме

можно представить, сопоставляя с образцом, который у нас имеется (в настоящее время это 19-я версия сборки генома – HG19 – результат кропотливой работы по сшиванию фрагментов путем поиска возможных перекрытий между ними – по сути сборка паззла огромных размеров).

Итак, у нас есть последовательность. Самое время посмотреть на нее подробнее. С точки зрения фрагментов, но не тех фрагментов, из которых она собиралась, а тех, которые реально функционируют.

У нас имеются более-менее полные знания о некоторых фрагментах последовательности генома. В первую очередь, конечно, это гены – то есть те участки генома, которые отвечают за синтез белков. Функционирование генов во многом определяется расположенными рядом (или не очень рядом) с ними регуляторными элементами. Кроме генов, важную роль (не ясную в полной мере) играют так называемые повторы – повторяющиеся участки, распределенные по геному. Зачастую мы можем определить, где они расположены в геноме (пусть читателя не смущает термин «зачастую» – дело в том, что большинство последовательностей настолько вариативны, что непросто сказать о принадлежности данной последовательности тому или иному типу).

Кроме самих элементов, очень важную роль играет взаимное расположение фрагментов внутри общей последовательности генома.

Посмотрим на общую картину. Из  $3,2 \cdot 10^9$  буковок мы можем что-то сказать о примерно 2–4 % последовательности – это гены. Около 50 % процентов генома составляют повторы. Огромное количество программ, использующих современные средства искусственного интеллекта, созданы для анализа, распознавания и расположения отдельных элементов генома. И тем не менее, пожалуй, самым надежным и универсальным средством по-прежнему является человек и его зрительное восприятие.

Именно потому, что необходимо пред-

<sup>1</sup> Размер романа «Война и мир» – 3220 Кб. Количество атомов во вселенной оценивается как  $10^{80} = 2^{240} = 4^{239}$ .

ставлять большие объемы плохо структурированной информации, были созданы специальные программные средства, которые получили название геномных браузеров. Почему во множественном числе? Почему нельзя создать одно универсальное средство? Видимо, это дело будущего, а в настоящее время, по крайней мере, более 30 программных средств носят название геномных браузеров. Некоторые из них ориентированы на работу с геномом только конкретных организмов (например, бактерий), некоторые не привязаны к организмам, но зато позволяют увидеть ограниченное количество геномных элементов. Некоторые браузеры ориентированы на автономную работу на компьютере исследователя, некоторые – на облачные вычисления, когда сама програм-

ма выполняется на удаленном сервере, а исследователь просто загружает данные и анализирует результаты на своем компьютере.

Что же привлекает особое внимание генетиков в настоящее время и почему необходим пристальный взгляд на весь геном?

Важными элементами генома являются SNP (single nucleotide polymorphism) – это единичные изменения в геноме, которые отличают одного индивидуума от другого. Единичные – значит, изменения касаются только одной буквы из 3 миллиардов (ну, может быть, нескольких рядом стоящих). Всего таких SNP в геноме человека известно около 1,5 миллионов. Что они означают? Конечно, это зависит от того, где они расположены и какие изменения они представляют. Когда такие SNP расположены внутри ге-

---

## ИСТОРИЯ ЧТЕНИЯ ГЕНОМА

Представление о том, что геном человека был расшифрован в 2000 году, не совсем соответствует действительности. Действительно, в 2000 году громко и четко, скорее рекламно, было объявлено о расшифровке генома человека. Какого человека? Одного. Его имя неизвестно. Реально все мы, люди, отличаемся друг от друга очень незначительно с точки зрения генома как нуклеотидной последовательности. Не более чем на 0,4 %. Все остальное идентично (в каком-то смысле).

Значит ли вышесказанное что:

- Теперь мы знаем о геноме человека все?
- В каждой клеточке нашего тела у каждого из нас содержатся одинаковые последовательности нуклеотидов?

Увы. Ответы на оба вопроса отрицательны. Мы знаем о геноме далеко не все. Сама последовательность (кстати, она известна еще не вся, некоторые фрагменты остались нерасшифрованными – только не думайте, что именно эти нерасшифрованные участки тают что-то самое важное) – это текст, составленный из четырех букв (ACGT) без точек, запятых, пробелов. Более того, слова в тексте могут пересекаться, разрываться, продолжаться позже и т. д. Да и «словарь» слов – элементов, которые мы знаем, далеко не полон.

Идентичность геномов в каждой клеточке тоже не совершенна. В дополнение к неизбежным мутациям («опечаткам»), а это 1 на миллиард, возможны перестройки при клеточном делении, а также постоянно возникающие новые мутации под действием внешних (радиация, особые химические вещества) и внутриклеточных условий.

Процесс расшифровки генома продолжается. Каждый год публикуются тысячи и тысячи научных работ на эту тему. Огромное количество ученых во всем мире и не только в развитых странах заняты в этом. Немало открытий совершается в области генетики и смежных областях каждый год.

Каждый день сиквенируются (термин сиквенирование, имеющий англоязычные корни, успешно прижился в словаре генетиков и означает расшифровку ДНК – «превращение» молекулы ДНК в последовательность букв из алфавита {ACGT}) новые геномы для новых организмов. И не только для новых видов организмов, но и для новых особей того же вида. Многие виды организмов сиквенируются повторно для определения отличий одного индивидуума от другого. Поток генетической информации огромен. При этом некоторое количество новостей и публикаций основывается на непроверенных материалах или требует уточнения. Недаром фраза о «Британских ученых» стала расхожей для обозначения сенсационных, но бесполезных или просто невероятных открытий, которые требуют более тщательного анализа, в том числе с использованием геномных браузеров...

нов, это может означать одно из трех: ген будет полностью выключен из функционирования; ген может функционировать как-то иначе; а может и совсем не изменять функцию гена.

Почему знание незначительных фрагментов так важно? Все мы немного разные. И если одни подвержены каким-то заболеваниям, а другие нет – то их различие, вероятно, определяется такими SNP.

А как же узнают о таких различиях, если известен только один геном одного человека?

Да, геном полностью собран только в одном экземпляре – сборка генома из фрагментов секвенирования – очень трудоемкий процесс. И неоднозначный. Но для тысяч людей известны (отсеквенированы) небольшие фрагменты целого генома, и, когда накладывают эти фрагменты на весь геном-образец, видно, какие отличия имеет тот или иной человек. Эти отличия и есть SNP (замена одного нуклеотида, вставка или удаление небольшого фрагмента или комбинация таких отличий).

Далее, сравнивая такие отличия и сопоставляя клинические данные о людях, чьи фрагменты ДНК сравниваются, можно сделать выводы о том, что возможно такие изменения явились причиной возникновения тех или иных заболеваний/особенностей.

Теперь самое время открыть читателям некую тайну, которая, по сути, тайной не является. «Наука принадлежит народу». В том числе большинство генетических данных как научные достижения открыты для всех (конечно, это не касается персональных данных, являющихся объектом врачебной тайны).

Геном человека? А, пожалуйста. В настоящее время все генетики работают с 19 версией собранного полного генома человека – просто наберите HG19 или синоним GRCh37 на сервере [www.ncbi.nlm.nih.gov](http://www.ncbi.nlm.nih.gov) или просто в Google.

Геном туберкулеза? Запросто. Тысячи последовательностей – вариантов, связанных с бактерией туберкулеза, доступны в Интернете, и один из них принят за образец (H37Rv).

---

## ГЕНОМНЫЕ ЭЛЕМЕНТЫ

Наиболее часто встречающимися элементами в геноме, как уже говорилось, являются так называемые повторы (repeats). Это небольшие (от 2–10 до 100 пар оснований) и большие (от 500 до 1000 пар) фрагменты, которые встречаются в составе генома не один и не два раза.

Заинтересованные читатели могут узнать более подробно про геномные элементы по этой ссылке <http://humbio.ru/humbio/molevol/000199bc.htm>. Здесь приведем только выжимки из полного материала.

Геном человека известен на 94%. Повторы занимают, по крайней мере, 50 % генома.

Несеквенированная пока часть генома относится в основном к центромерным областям хромосом.

Функции некоторых повторов уже более-менее известны: например LTR – служат для обозначения начала кодировки ретранспозонов – мобильных элементов генома. Пользуясь обыденной терминологией, такие повторы можно сравнить с пешеходной зеброй на дороге, которая в большинстве случаев (но не всегда) предшествует перекрестку.

Функции многих других повторов абсолютно неясны, но, несомненно, они имеют какую-то функцию (может быть, это какие-то инструменты – «гаечные ключики», разбросанные по геному с такой целью, чтобы ключ всегда оказывался «под рукой»?)

Конечно, другим важнейшим и наиболее изученным элементом генома являются гены – фрагменты ДНК, которые непосредственно кодируют белки. Всего в организме человека насчитывается около 30 000 генов. Но это не значит, что и белков у человека столько же. Наиболее достоверная база данных белков (UNIPROT) содержит около полумиллиона белков – дело в том, что один и тот же ген может быть источником многих белков, благодаря недавно открытому механизму альтернативного сплайсинга.

Каждый элемент в геноме выполняет свою функцию. Причем, возможно, не одну. При этом распределение элементов по геному не просто важно – зачастую имеет решающее значение.



Подробный список всех белков, которые уже открыты и синтезируются в живых организмах? Еще проще...

Геном HIV – вируса иммунодефицита человека также представлен в виде сотен тысяч вариантов.

Среди множества баз данных, хранящих генетическую информацию, самым авторитетным банком является NCBI (National Center for Biotechnology Information – адрес в Интернете [www.ncbi.nlm.nih.gov](http://www.ncbi.nlm.nih.gov)). Доступ к этой базе данных и полную информацию о ней можно получить любому человеку, подключенному к Интернету. Единственное ограничение – знание английского языка, который *de facto* стал языком науки.

Чтобы не быть голословным, давайте вместе шаг за шагом проследим, как же работает браузер и что мы можем увидеть. Ради упрощения вместо генома человека будем рассматривать геном простого существа, значительно меньшего размера – сути это не меняет.

Но, чтобы не отвлекаться, приведем краткое описание того, что мы можем сделать прямо сейчас.

Для оперативной поддержки этой статьи создана страница в Интернете, на которой подробно разбираются примеры работы с геномным браузером. На сайте <http://genomebrowser.dobzhanskycenter.org> можно найти подробные инструкции и примеры. Поскольку работы по развитию технологии продолжаются и браузер совершенствуется, возможно что рекомендации на странице отличаются от рекомендаций в данной статье. Кроме того, страница служит существенным дополнением к тексту и будет совершенствоваться.

*GBrowse* – это программа, написанная для исполнения на веб сервере на языке Perl и его расширении BioPerl. Для работы с ней Вам не понадобится знание программирования.

Существует четыре способа работы с программой:

1. Использовать ее на сервере разработчиков. Это не самый гибкий способ, но самый простой для начинающих пользователей браузера.

2. Вторым по сложности, но реально тоже совсем не сложным является установка на Ваш компьютер всего программного обеспечения, которое эмулирует как веб сервер, так и браузер. В этом случае Вы не будете зависеть от подключения к сети и сможете работать автономно. Процедура установки всего программного обеспечения достаточно проста и состоит из двух частей: установка виртуальной Линукс-машины на Ваш компьютер, а затем установка самого браузера. Реально это достаточно несложно и потребует примерно часа работы. На странице поддержки есть описание процесса.

3. Установка браузера на ваш собственный веб сервер. Это достаточно непростая процедура и потребует немалых знаний по администрированию сервера, баз данных и их «тонкой» настройке.

4. Наконец, четвертый способ, включающий в себя предыдущий, позволяет встроить браузер в Ваш собственный интерфейс и адаптировать его под решение Ваших задач.

Итак, самый простой путь работы с браузером – это использование его на сервере разработчиков

<http://webgbrowse.cgb.indiana.edu>.

---

---

## GBROWSE

Что нам позволяет делать GBrowse?

GBrowse – это облачные вычисления, то есть программа, выполняемая где-то на удаленном компьютере – сервере. А Вы, используя Ваш обычный браузер (Chrome, Opera, FireFox и т. д.) на Вашем компьютере, подключенном Интернету, можете пересылать туда данные, запускать вычисления, скачивать или просматривать результаты, смотреть расположение элементов генома на экране Вашего компьютера с различным разрешением.

Рис. 1

Конечно, минимальные знания английского языка не помешают, но подробные рекомендации дают возможность даже не владеющим языком начать работу с браузером.

Итак. Внешний вид страницы изображен на рис. 1.

На этой странице предлагается выбрать и загрузить на сервер файл с данными. Для простоты предлагается использовать имеющийся, приготовленный заранее текстовый файл, содержащий геном и его описание для небольшого организма. Например VOLVOX. Что это за организм? Это просто зеленая тина, покрывающая стоячие водоемы – люди считают, что это лягушачья икра – на самом деле, – это разновидность сине-зеленых водорослей.

Геном организма можно загрузить, кликнув на ссылку 1 (для простоты геном обрезан до размера 50 тысяч пар нуклеотидов). В дальнейшем вы можете пользоваться настоящими необрезанными геномами, но при этом размеры файлов и время работы будут существенно другими.

Кликнув на ссылку 1 текстовый файл будет загружен в Ваш компьютер либо в папку (загруженные – Download), либо туда, куда вы укажете.

Теперь вы можете посмотреть этот файл. (после загрузки его можно открыть для просмотра простым текстовым редактором, на-

пример Notepad, если вы пользуетесь компьютером с Windows.)

Посмотрели? Интересно? Можно ли что-то увидеть? Сомнительно, без подробного машинного анализа. Файл имеет популярный среди генетиков формат gff3, подробное описание которого можно найти на странице поддержки.

Небольшие комментарии по поводу увиденного содержимого файла. Файл состоит из трех частей: Комментариев, отмеченных знаком # в начале строки (при этом строка может быть очень длинной и завершаться символом перевода каретки – «Enter»). Затем следуют строки – описание геномных элементов. Каждая строка этого раздела файла содержит 9 полей, разделенных табуляцией, в том числе имя и расположение в геноме каждого описанного элемента (номер первого и последнего нуклеотида элемента в геноме).

Завершающая часть файла – это сам геном – множество букв в алфавите acgt (пусть не смущает то, что в данном случае используются те же буквы ACGT, но в другом регистре), – обозначающих последовательность генома как она есть (для удобства просмотра последовательность разделена на строки, в реальности никакого деления нет, и это просто одна длинная строка длиной в 50 000 букв.

Для дальнейшего анализа Вам потребуется файл, который описывает, КАК Вы будете просматривать этот геном, – так называемый конфигурационный файл. Стандартный файл для этого генома можно загрузить по ссылке 2. Файл также является текстовым, доступен для просмотра текстовым редактором Notepad, и в дальнейшем вы можете корректировать этот файл, чтобы настроить представление так, как Вам удобнее.

Две следующих операции – это загрузка этих файлов на сервер (тут стандартная процедура):

- кликнув на ссылку 3, выберите Ваш файл с геномом, который Вы только что загрузили по ссылке 1;

- кликнув на ссылку 4, выберите Ваш только что загруженный конфигурационный файл.

В этом месте (ссылка 5) Вы можете указать ваш электронный адрес – авторы браузера гарантируют, что единственное использование этого адреса – это информирование Вас о том, что файл загружен и его обработка завершена (для больших файлов это может потребовать какого-то времени).

Все. Ссылка 6 инициирует загрузку файлов и построение графического представления генома.

Спустя минуты, Ваш экран будет выглядеть примерно как на рис. 2.

При этом Вам предлагают интерактивно отредактировать вид представления (конфигурационный файл, который вы всегда сможете отредактировать перед очередной загрузкой в текстовом редакторе). Для начала можно ничего не трогать, а просто нажать на кнопку в самом низу странички, на которой написано Display in Gbrowse 2.0.

Всё. Следующий экран – фрагмент которого приведен на рис. 3 – это графическое представление генома.

Небольшие комментарии к картинке на экране. Здесь Вы можете

- изменить размер отображаемого фрагмента генома – просто заменив «1..10000» на диапазон, который вы хотите увидеть;

- варьировать размер просматриваемого окна, а также расположение на геноме просматриваемого «окна» в геноме;

- получить наглядное представление расположения элементов в геноме (в данном случае просто «гипотетических» элементов);

- увидеть элементы-совпадения, белок-кодирующие гены.

Экран действительно интерактивный, и ваши действия на нем мышкой немедленно будут иметь последствия. В частности, Вы можете выбрать район генома, требующий более пристального внимания с помощью желтых стрелок или изменить размер рассматриваемого фрагмента. В конфигурацион-

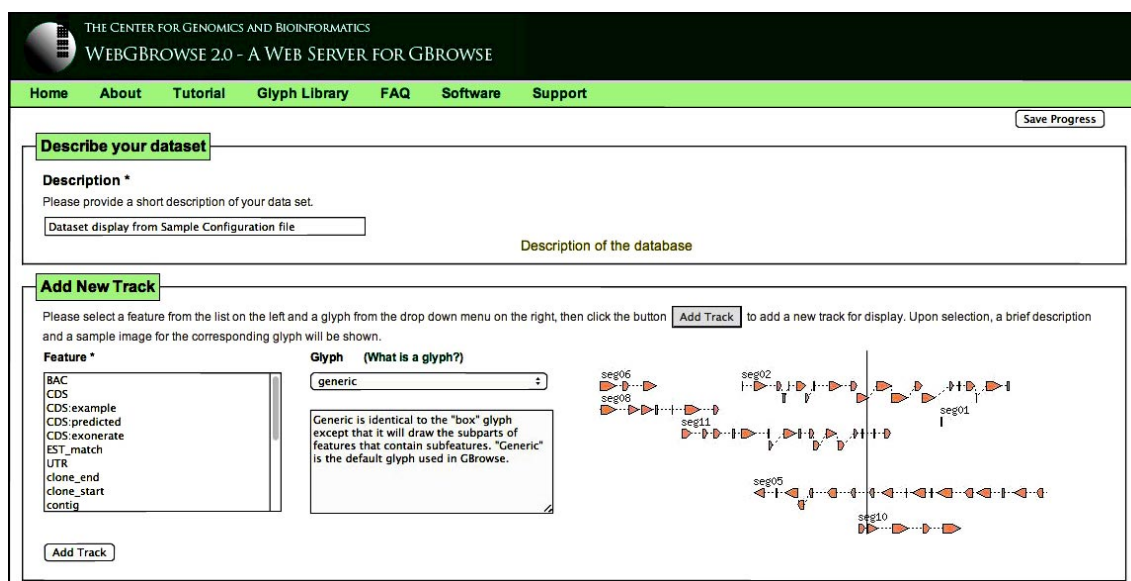


Рис. 2



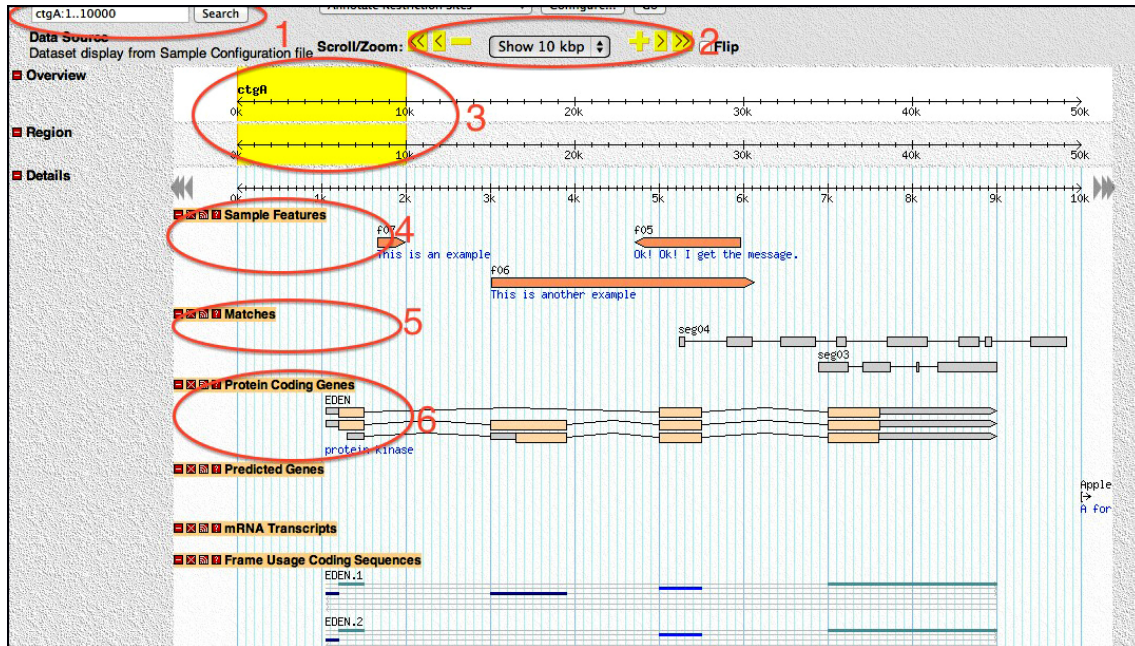


Рис. 3

ном файле можно изменить символы, обозначающие различные геномные элементы, сделать те или иные элементы видимыми или невидимыми. Посмотреть подробно (или загрузить на Ваш компьютер) рассматриваемый фрагмент в виде последовательности нуклеотидов. Или, в конце концов, полностью заменить весь геном и посмотреть другой.

Как это сделать и как описать те или иные геномные элементы, доступно описано на поддерживающей странице. Там же можно задать вопросы.

И самое главное – не бойтесь экспериментировать с программой. Отображаемые треки (линии расположения геномных элементов), как указано в 4, 5 и 6, могут быть удалены из рассмотрения, переставлены местами или заменены. Новые треки могут быть дополнены и подгружены в браузер в виде нового gff3 файла. Новые примеры как геномов, так и конфигурационных файлов можно загрузить с поддерживающей странички.

Надо сказать, что Интернет-версия этого браузера, которую мы указали для ознакомления, предоставляя гибкий интерфейс

и легкость в освоении, не предоставляет всех возможностей, которыми пользуются исследователи, просто в силу упрощения интерфейса. Но это никак не умаляет его достоинств с точки зрения наглядного представления структуры генома.

### ЗАКЛЮЧЕНИЕ

В статье описаны возможности, предоставляемые исследователям-генетикам по визуализации генома и составляющих его элементов, по аналогии с Интернет браузерами названные геномными браузерами. На примере одного из наиболее распространенных браузеров (GBrowse) показано, какие возможности открываются для анализа генома. Электронная поддержка статьи на странице Интернета [genomebrowser.dobzhanskycenter.org](http://genomebrowser.dobzhanskycenter.org) представляет более подробное описание браузера, форматов данных и примеров его использования. Кроме того, указанная страничка предполагается стать неким новостным источником информации о состоянии дел в области визуализации исследований геномов.

## BROWSER FOR A GENOME

### Abstract

In the article describe opportunities give to researchers by a computer tools called Genome browsers. The tools implement visualization of genome sequences and genetic elements inside genome. An observe of approaches for a set of genome browsers includes. A open source tool – Gbrowse is describe as an example. Detail description of an opportunities and short instruction to use includes as well as a link to web page for support the article.

**Keywords:** genome, genome elements, visualization, genome browser.

*Симонов Сергей Александрович,  
кандидат физико-математических  
наук, главный специалист  
лаборатории «Центр геномной  
биоинформатики  
им. Ф.Г. Добржанского» СПбГУ,  
SergueiSimonov@gmail.com*

*Макунин Алексей Игоревич,  
лаборант-исследователь  
лаборатории «Центр геномной  
биоинформатики  
им. Ф.Г. Добржанского» СПбГУ  
alex.makunin@gmail.com*

*Stephen J. O'Brien,  
доктор философии в области  
биологии (PhD in biology)  
заведующий лабораторией «Центр  
геномной биоинформатики  
им. Ф.Г. Добржанского» СПбГУ,  
lgdchief@gmail.com.*



Наши авторы, 2013.  
Our authors, 2013.