

УДК 51-76; 519.256; 575.162
DOI: 10.15827/0236-235X.031.1.204-211

Дата подачи статьи: 07.08.17
2018. Т. 31. № 1. С. 204–211

РАЗРАБОТКА ПРОГРАММНОГО КОМПЛЕКСА ДЛЯ УЧЕТА ЛАБОРАТОРНЫХ ЖИВОТНЫХ

Е.Ю. Фурта¹, аспирант, lena-furta@rambler.ru

И.М. Шабалина¹, к.т.н., доцент, irinashabalina09@gmail.com

¹ Петрозаводский государственный университет,
просп. Ленина, 33, г. Петрозаводск, 185910, Россия

Исследования в области молекулярной генетики связаны с накоплением и хранением значительного объема данных, получаемых в ходе проведения экспериментов и предназначенных для дальнейшего анализа и обоснования результатов. В настоящее время развиваются глобальные проекты и ресурсы для хранения и обработки результатов секвенирования генома, предоставляющие данные и результаты исследований широкому кругу пользователей. Тем не менее, прежде чем публиковать информацию, используя глобальные ресурсы, каждое отдельное исследование требует подбора форматов хранения данных и ПО для их обработки, учитывающих уникальные особенности проводимых экспериментов.

При работе с живыми организмами на уровне лаборатории требуется организовать учет животных с соблюдением норм и стандартов, действующих в этой области. Реализуемая для этих целей программная система должна соответствовать задачам проводимых исследований, давать возможность всестороннего анализа данных и формирования отчетов.

В статье описывается программная система учета лабораторных животных, реализованная для лаборатории молекулярной генетики врожденного иммунитета Петрозаводского государственного университета. Помимо хранения данных о животных, в системе реализованы методы хранения и визуализации родословной каждого животного, что позволяет отследить изменения его генотипа в нескольких поколениях.

Предлагаемая система дает возможность проведения анализа выживаемости животных и статистической обработки данных генотипирования. В ней реализованы справочники, журнал учета лабораторных животных, отчеты. Система имеет централизованное хранилище данных и удобный пользовательский веб-интерфейс. В настоящее время программная система внедрена в эксплуатацию, БД содержит данные более чем о 5 000 животных.

Ключевые слова: генетические исследования, наследственность, учет лабораторных животных, хранение генетических данных.

Разработка автоматизированной системы для учета лабораторных животных проводилась в рамках НИР, выполняемых лабораторией молекулярной генетики врожденного иммунитета Петрозаводского государственного университета. Одним из направлений работ является выявление участков генома, обуславливающих высокую чувствительность человека к септическому шоку, с использованием мышинной модели.

Биологическая проблематика

Различные виды млекопитающих обладают разной чувствительностью к септическому шоку. Предварительные исследования показали, что септический шок можно строго моделировать (имитировать) не путем заражения бактериями, а непосредственной инъекцией препарата, искусственно полученного и очищенного TNF (tumor necrosis factor). В исследовании задействованы две линии мышей, полностью противоположные по восприимчивости к TNF: близкие к человеку по чувствительности к TNF (линия B6) и абсолютно устойчивые (линия Msm). Это позволяет использовать метод генетического анализа для поиска участков генома, обуславливающих чувствительность организма к TNF, то есть путем скрещивания особей

противоположного фенотипа и анализа получаемого потомства.

Фенотип определяется двумя признаками: резистентный (мышь выжила после инъекции TNF) и чувствительный (мышь погибла). Генетическая информация представлена тремя типами: А – генотип резистентный, В – генотип чувствительный, Н – гетерозиготный. Единицей информации о генотипе является маркер – участок хромосомы (локус), на котором определяется генотип. Анализ потомства проводится путем сопоставления наблюдаемого фенотипа и устанавливаемого для каждого потомка генотипа.

Задачи и проблемы, пути решения

На первый взгляд, для определения взаимосвязи между генотипом и фенотипом достаточно решить задачу поиска маркеров, обладающих сильной связанностью с фенотипом при помощи количественной меры связи, основанной на применении статистического критерия (например хи-квадрат Пирсона). Действительно, в предположении, что одна группа генов, находящаяся на одной хромосоме в одном локусе, полностью определяет резистентность или чувствительность объекта, задача решается достаточно просто. Однако в ходе пред-

варительной обработки данных обнаружилось, что такая однозначность не подтверждается имеющимися выборочными данными, природа резистентности и чувствительности сложнее и не может быть объяснена генетической информацией, содержащейся на одном локусе [1]. Более того, исследование генетических данных потомков без учета генетики родителей не позволяет точно определить важность наследования отдельных участков генома при сохранении фенотипа.

Стандартное представление генотипических и фенотипических данных в виде таблиц не позволяет выявить индивидуальные характеристики предков и потомков конкретного объекта исследования, можно только анализировать поколение лабораторных животных в целом. Создание структуры данных с реализацией связей «предок–потомок» дает возможность исследования признаков не только в горизонтальной структуре (в рамках одного поколения), но и с учетом вертикальных связей между поколениями.

В проводимых исследованиях на данный момент участвовало более 10 поколений животных, общее количество животных достигло нескольких тысяч. Очевидно, что ручной способ учета, оприходования и списания животных требует значительных трудозатрат. Кроме того, учет лабораторных животных регламентируется соответствующими нормами и стандартами. При приобретении животных между медицинским учреждением и питомником заключается контракт (договор), порядок заключения которого должен соответствовать нормам Федерального закона № 44-ФЗ. Списание лабораторных животных оформляется актом на списание подопытных животных (форма 25-МЗ) либо материальных запасов (форма 0504230) [2]. Автоматизация учета животных упрощает соблюдение требований норм и регламентов за счет унификации процедур учета и формирования отчетов, а также сокращает трудозатраты персонала лаборатории за счет улучшения качества учета, упрощения поиска и отбора животных для участия в эксперименте, упрощения процедуры расчета затрат на содержание животных.

Для количественных исследований генетической и фенотипической информации требуется применение базовых статистических процедур (корреляционный анализ, частотный анализ, построение гистограмм распределения и пр.). Реализация основных статистических функций в качестве модулей системы учета позволит проводить оперативные вычисления для сгруппированных по разным критериям данных. Для более детальных исследований возможны выгрузка данных во внешние системы и дальнейшее использование пакетов статистической обработки данных (R, Statistica, MathLab и пр.).

Таким образом, при проведении исследований с применением генетического анализа актуальным

является создание автоматизированной системы для учета лабораторных животных, обладающей следующими функциями:

- сквозной учет всех лабораторных животных, участвующих в эксперименте;
 - учет наследственности: если животное родилось в лаборатории, то точно известны его предки и в дальнейшем потомки;
 - возможность базовой статистической обработки сгруппированных данных (по поколениям, по полу, по типам секвенирования и пр.), применение к данным корреляционного или частотного анализа;
 - удобный интерфейс для внесения информации;
 - система отчетов;
 - выгрузка во внешние системы (R, Statistica)
- для более глубокого статистического исследования данных.

Современное состояние исследований

Многие ученые из таких областей, как биология, информатика и биоинформатика, приложили немало усилий, чтобы найти способ хранения генетических данных и управления ими единым образом.

Так, был разработан целый класс программных систем для поддержки работы лабораторий. Эти системы были названы LIMS (Lab Information Management Systems, лабораторная информационная менеджмент-система). Следует отметить, что у каждой лаборатории свои требования к хранению и управлению данными. Соответственно, LIMS, разработанные для химической лаборатории, будут отличаться от LIMS, необходимых для молекулярно-генетической лаборатории. Авторы интересуют системы, предназначенные для молекулярно-генетических лабораторий. Большинство из них ориентированы на хранение обработанных данных и управление ими, они включают в себя данные микрочипов [3–5] и данные из [6–8]. Wendl и др. разработали информационную систему для отслеживания данных секвенирования [9]. В 2006 году группа исследователей разработала открытую информационную систему для генотипирования AGL-LIMS [10]. Появились некоторые приложения, которые одновременно стали уделять внимание управлению генотипическими данными одиночных нуклеотидных полиморфизмов (SNP) и данными фенотипа [11, 12]. Кроме того, Weissensteiner и др. расширили свои системы разработанной в 2009 году [13], чтобы сделать возможными импорт и хранение данных о митохондриальных ДНК и данных STR (коротких tandemных повторов) [13]. Общая цель этих информационных систем – предоставить пользователям лабораторий средства для хранения данных в одном месте и при необходимости извлечения их для дальнейшего

анализа. Тем не менее, пользователи часто стремятся получить исходные данные из конкретной платформы [11] или импортировать только окончательные данные, игнорируя исходные [13, 14]. Некоторые системы [12–14] не предоставляют способы документирования лабораторных данных. Также существует LIMS под названием MolabLS [15], которая имеет возможность хранения данных для разных видов организмов, привязана к конкретным особям и может формировать различные отчеты по данным секвенирования.

В данном исследовании используется иная, нежели в указанных информационных системах, формат хранения данных. К тому же система предоставляет первичную статистическую обработку, чего не удалось найти в других системах для хранения генетических данных.

Так как доступные информационные системы предназначены для определенных лабораторий, установить и использовать их в других лабораториях, как правило, сложно.

Большинство систем ориентируются на анализ результатов работы секвенаторов, то есть на анализ первичной генетической информации без привязки к объектам исследования.

Учетные системы обычно реализованы для конкретных, достаточно узких предметных областей, поэтому коммерческих продуктов мало и, как правило, они не русскоязычные и не обладают хорошими адаптационными возможностями к задачам, отличающимся от спецификации конкретной программы.

Описание системы

Разработанная система учета лабораторных животных позволяет хранить информацию о всех лабораторных животных, участвовавших в эксперименте, производить базовую статистическую обработку, создавать отчеты по хранимым данным.

В системе реализован удобный пользовательский интерфейс, благодаря которому можно добавлять всю информацию о новом лабораторном животном, производить корреляционный или частотный анализ по сгруппированным данным, а также формировать различные отчеты.

Данная система реализована с помощью платформы .NET.

Функции программной системы учета лабораторных животных:

- ведение БД паспортов животных: добавление и корректировка данных о животных, включая данные генетического исследования;
- построение генеалогического дерева с указанным количеством уровней;
- формирование списка животных с возможностью фильтрации по различным параметрам;
- расчет количественных оценок степени зависимости фенотипов группы животных и геноти-

пов: статистики хи-квадрат, ранговой корреляции.

В паспорте лабораторного животного содержатся следующие блоки информации: номер животного; номер клетки, в которой содержится животное; номер тетради, в которой содержатся данные о животном; биологический вид; пол; дата рождения; дата поступления в лабораторию; ссылки на паспорта обоих родителей; дата подготовки эксперимента; дата начала эксперимента; дата инъекции; ссылка на вещество и концентрацию; дата летального исхода после инъекции; дата выведения животного из эксперимента; причина вывода из эксперимента; дополнительная информация (рис. 1).

Печатная форма должна содержать паспорт животного (рис. 2) с генеалогическим деревом с заданным числом уровней, сводный реестр животных, сводный отчет по линиям с информацией о количестве выживших и умерших после инъекции.

Программная система содержит следующие справочники:

- справочник линий животных (наименование линии, признак «гибрид» или «не гибрид», ссылки на «родительские линии»);
- справочник веса животных в зависимости от возраста и пола;
- справочник вводимых веществ (наименование вещества и единицы измерения концентрации вещества);
- справочник маркеров для хранения наименований маркеров, расстояния в сантиморганах, номера хромосомы для маркера;
- справочник наборов маркеров и справочник маркеров, принадлежащих наборам (номера списков и номера маркеров, относящихся к спискам).

Отчеты в программной системе учета лабораторных животных: акт на оприходование приплода животных и отчеты по выживаемости.

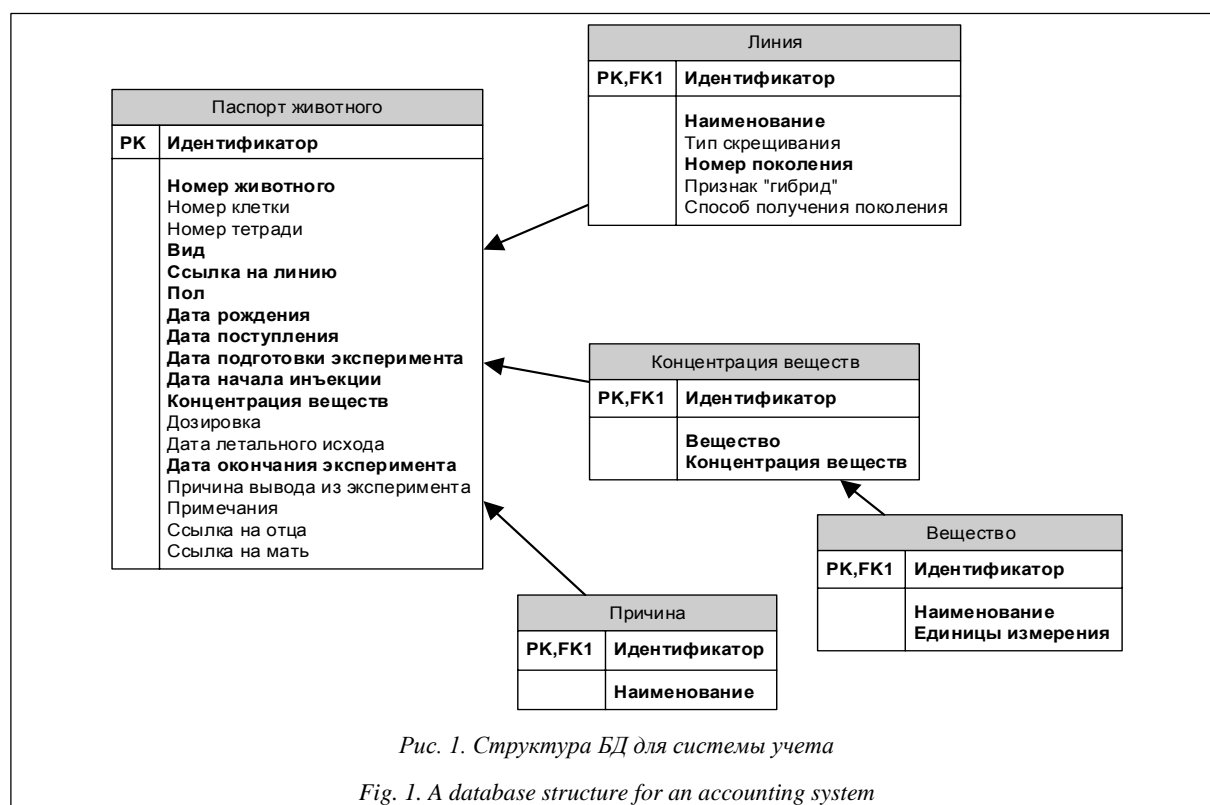
Интерфейс

В системе реализован удобный пользовательский интерфейс. На главном экране программы находятся списки паспортов всех животных. При выборе одного из них открывается паспорт животного, который можно редактировать, также можно в удобном формате отобразить всех предков или потомков особи. В системе имеется календарь с отображением таких событий, как начало эксперимента, подготовка эксперимента, рождение новых особей.

Анализ данных

1. Построение таблицы сопряженности признаков, расчет статистики ХИ-квадрат, ранговой корреляции.

В таблице сопряженности строки – это особи (номера паспортов), а столбцы – маркеры, которые



берутся из справочника маркеров. На их пересечении указывается значение маркера для данной особи (А, В, Н).

Расчет статистики по маркерам представляет собой значения ХИ-квадрат и ранговой корреляции для каждой хромосомы по всем выбранным особям.

2. Составление отчетов по выживаемости.

Отчет «Анализ выживаемости» содержит информацию о количестве выживших и умерших потомков (в первом поколении) от указанных родителей. Из данного отчета можно узнать, какая комбинация генов родителей дала наиболее устойчивое потомство.

Отчет «Анализ выживаемости потомства» содержит информацию о количестве выживших или умерших потомков по отцу или по матери в каждом поколении. Из него можно узнать, у какой особи наиболее устойчивое потомство, сравнить с анализом выживаемости по обоим родителям и выяснить, достаточно ли генов одного родителя или необходима комбинация.

Отчет «Выживаемость линий» (рис. 3) содержит информацию о количестве самцов и самок в каждой линии и выживаемости особей в каждой линии. Отчет представляет собой таблицу, в которой наглядно видно количество выживших (желтый цвет) и умерших (зеленый цвет) как отдельно по самцам или самкам, так и в целом.

Отчет «Выживаемость линий по родителям» содержит информацию о линии родителей, линии детей, количестве детей и выживаемости родителей и

детей. Отчет позволяет проследить выживаемость как детей, так и родителей.

Отчет «Распределение маркеров по поколениям» отображает график количества значений по каждому маркеру в конкретном поколении (рис. 4). Он позволяет оценить количество каждого значения маркеров в поколении и сделать предположение о влиянии этих значений на устойчивость особей. Эти предположения можно подтвердить при помощи значения ХИ-квадратов и ранговой корреляции, подсчитанных в программной системе.

3. Возможности выгрузки данных для обработки в прочих системах.

Данные из системы выгружаются в таблицу Excel, которую можно использовать во многих пакетах обработки статистических данных, например, в Statistica, R, MatLab и др.

Результаты, выводы

Разработанная программная система учета лабораторных животных внесла существенный вклад в развитие информационной инфраструктуры лабораторий молекулярной генетики врожденного иммунитета Петрозаводского государственного университета. На данный момент система введена в промышленную эксплуатацию, в ее БД содержится информация более чем о 5 000 лабораторных животных, из них около 10 % генотипированы, известны значения генетических маркеров. Применение системы повышает эффективность работы как лаборантов, отвечающих за учет и содержание жи-

| ПАСПОРТ лабораторного животного | | | | | | |
|------------------------------------|-----------------------------|---------------------|--|-------|------|------|
| 1 | № клетки | | 45 | | | |
| 2 | Вид | | mus musculus | | | |
| 3 | Линия | | гиб. N3F1 | | | |
| 4 | Пол | | ♂ самец | | | |
| 5 | № чипа | | | | | |
| 6 | Возраст (месяц/дни) | квартал | I | II | III | IV |
| | | 2013 г. | | | 1/42 | 1/53 |
| | | 2014 г. | | | | |
| 7 | Вес, г (средний) | квартал | I | II | III | IV |
| | | 2013 г. | | | 15,1 | 17,9 |
| | | 2014 г. | | | | |
| 8 | Дата поступления | | | | | |
| 9 | Дата рождения | | 20.08.2013 | | | |
| 10 | Родители | | ♂ гиб. N2F1 №1521 × ♀ B6 №1578 | | | |
| 11 | Дата начала эксперимента | 1. подготовка | 20.08.2013 | | | |
| | | 2. начало | 10.09.2013 | | | |
| | | 3. инъекции: | | | | |
| | | - вещество | TNF | GalN | NaCl | |
| | | - концентрация | | 20 мг | 0,9% | |
| | - вводимая доза | 0,5 см ³ | | | | |
| 12 | Дата окончания эксперимента | | | | | |
| 13 | Особые отметки | | 2р ♂ B6G2 №496 × ♀ гиб. F1 №131 3р ♂ B6G1 №203 контроль × ♀ B6G1 №244 контроль 4р ♂ B6 №1 × ♀ B6 №86 | | | |
| 14 | Примечание | | | | | |

Рис. 2. Пример интерфейса пользователя (печатная форма паспорта лабораторного животного)

Fig. 2. The example of the user interface (printing form of a laboratory animal passport)

вотных, так и научных сотрудников, проводящих исследовательские работы.

Повышение качества и эффективности учета животных достигается за счет

- учета всех животных, находящихся в лаборатории, ведения подробных паспортов животных (электронная и печатная формы);
- четкой привязки к поколениям, ведения дополнительной информации для поиска животных в помещении вивария;
- упрощения формирования отчетности по НИР, проводимых на базе лаборатории;
- наличия нескольких вариантов поиска информации – по поколениям, по родителям/детям, по календарю (в случае привязки к датам проведения инъекций);
- верификации введенных данных;
- использования печатных форм, рекомендованных к использованию нормами и правилами учета лабораторных животных.

Программная система предоставляет следующие возможности для планирования экспериментальной работы и отражения результатов экспериментов:

- формирование списка животных для генотипирования с учетом наследственности;

- оперативный контроль и фиксация результатов эксперимента, заключающегося в проведении инъекций с последующей фиксацией фенотипа;

- ведение наглядной БД генетических данных по маркерам;

- наглядное представление рождаемости, смертности, дат эксперимента и летальных исходов на календаре;

- создание структуры данных с реализацией связей «предок–потомок» для исследования признаков не только в горизонтальной структуре (в рамках одного поколения), но и с учетом вертикальных связей между поколениями;

- выявление особей, наиболее подходящих для проведения эксперимента.

Упрощение получения основной аналитической информации о ходе эксперимента:

- количество животных, выживших и умерших в результате эксперимента;

- определение связанности генотипических маркеров и фенотипов в разных поколениях и линиях животных на основе статистических и графических методов.

Кроме того, в системе имеется возможность загрузки данных о генотипе и фенотипе во внешние системы в виде таблицы с данными для тех

| Выживаемость линий | | | | | | | | |
|--------------------|------------|--------|-------|------------------|-----|-------------------|-----|-------------------|
| Линия | Количество | | | Выживаемость | | | | |
| | самцов | самцов | всего | самцов | | самок | | общая |
| F2 | 458 | 469 | 927 | | | 291 -> 92s + 199d | | 291 -> 92s + 199d |
| | | | | | | 32% | 68% | 32% 68% |
| intercross N2F2 | 122 | 128 | 250 | 80 -> 14s + 66d | | 80 -> 9s + 71d | | 160 -> 23s + 137d |
| | | | | 18% | 82% | 11% | 89% | 14% 86% |
| intercross N3F2 | 218 | 210 | 428 | 1 -> 1s | | 159 -> 30s + 129d | | 160 -> 31s + 129d |
| | | | | 100% | | 19% | 81% | 19% 81% |
| intercross N4F2 | 52 | 46 | 98 | 21 -> 15s + 6d | | 20 -> 13s + 7d | | 41 -> 28s + 13d |
| | | | | 71% | 29% | 65% | 35% | 68% 32% |
| intercross N5F2 | 86 | 86 | 172 | 83 -> 44s + 39d | | 83 -> 35s + 48d | | 166 -> 79s + 87d |
| | | | | 53% | 47% | 42% | 58% | 48% 52% |
| backcross N3F1 | 138 | 123 | 261 | 107 -> 16s + 91d | | 94 -> 29s + 65d | | 201 -> 45s + 156d |
| | | | | 15% | 85% | 31% | 69% | 22% 78% |
| backcross N4F1 | 167 | 156 | 323 | | | 120 -> 15s + 105d | | 120 -> 15s + 105d |
| | | | | | | 13% | 87% | 13% 87% |
| backcross N5F1 | 56 | 76 | 132 | 39 -> 6s + 33d | | 56 -> 13s + 43d | | 95 -> 19s + 76d |
| | | | | 15% | 85% | 23% | 77% | 20% 80% |

Рис. 3. Пример интерфейса пользователя (выживаемость линий)

Fig. 3. The example of the user interface (lines survival)

Поколение intercross N3F2

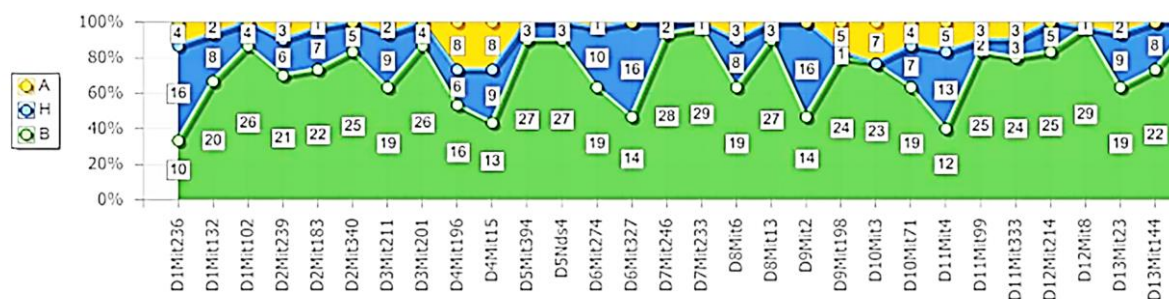


Рис. 4. Пример интерфейса пользователя (отчет о распределении маркеров среди умерших особей)

Fig. 4. The example of the user interface (the report on markers distribution for dead animals)

случаев, когда требуется более глубокий анализ данных.

Заключение

Авторы разработали программную систему, предназначенную для применения в виварии Петрозаводского государственного университета для ведения БД лабораторных животных, участвующих в экспериментах, проводимых лабораторией молекулярной генетики врожденного иммунитета.

Данная система необходима, так как исследование генотипа невозможно без учета наследственности. Система обеспечивает учет не только всех ла-

бораторных животных, участвовавших в эксперименте, но и их предков и потомков, возможность базовой статистической обработки генетических данных животного, наличие различных справочников и отчетов, удобный пользовательский интерфейс.

Развитие системы продолжается. Рассматривается возможность более детального исследования данных сразу в системе, без выгрузки во внешние программы.

На момент подготовки публикации в системе записаны данные более чем о 5 000 особей, прослежены 50 различных линий и 12 поколений животных.

Литература

1. Volkova T., Furta E., Dmitrieva O., Shabalina I. Pattern building methods in genetic data processing. *Jour. on Selected Topics in Nano Electronics and Computing*, 2014, vol. 2, no. 1, pp. 2–6. URL: http://jstnec.petrsu.ru/journal/article_en.php?id=3041&seq (дата обращения: 6.08.2017).
2. Учет лабораторных животных в медицинских учреждениях. URL: <http://economic-club.ru/st/uchet-laboratornykh-zhivotnykh-v-medicinskikh-uchrezhdeniyakh> (дата обращения: 6.08.2017).
3. Kokocinski F., Wrobel G., Hahn M., Lichter P. QuickLIMS: facilitating the data management for DNA-microarray fabrication. *Bioinformatics*, 2003, vol. 19, no. 2, pp. 283–284. DOI: 10.1093/bioinformatics/19.2.283.
4. Swertz M.A., de Brock E.O., van Hijum S.A.F.T., de Jong A., Buist G., Baerends R.J.S., Kok J., Kuipers O.P., Jansen R.C. Molecular genetics information system (MOLGENIS): alternatives in developing local experimental genomics databases. *Bioinformatics*, 2004, vol. 20, no. 13, pp. 2075–2083. DOI: 10.1093/bioinformatics/bth206.
5. Monnier S., Cox D.G., Albion T., Canzian F. T.I.M.S: Tag-Man information management system, tools to organize data flow in a genotyping laboratory. *BMC Bioinformatics*, 2005, vol. 6, no. 246. DOI: 10.1186/1471-2105-6-246. URL: <https://bmcbioinformatics.biomedcentral.com/articles/10.1186/1471-2105-6-246> (дата обращения: 17.07.2017).
6. Goh C.S., Lan N., Echols N., Douglas S.M., Milburn D., Bertone P., Xiao R., Chung Ma L., Zheng D., Wunderlich Z., Aceton T., Montelione G.T., Gerstein M. SPINE 2: a system for collaborative structural proteomics within a federated database framework. *Nucleic Acids Res*, 2003, vol. 31, no. 11, pp. 2833–2838. DOI: 10.1093/nar/gkg397.
7. Morisawa H., Hirota M., Toda T. Development of an open source laboratory information management system for 2-D gel electrophoresis-based proteomics workflow. *BMC Bioinformatics*, 2006, vol. 7, no. 430. DOI: 10.1186/1471-2105-7-430. URL: <https://bmcbioinformatics.biomedcentral.com/articles/10.1186/1471-2105-7-430> (дата обращения: 17.07.2017).
8. Droit A., Hunter J., Rouleau M., Ethier C., Picard-Cloutier A., Bourgeois D., Poirier G. PARPs database: A LIMS systems for protein-protein interaction data mining or laboratory information management system. *BMC Bioinformatics*, 2007, vol. 8, no. 483. DOI: 10.1186/1471-2105-8-483. URL: <https://bmcbioinformatics.biomedcentral.com/articles/10.1186/1471-2105-8-483> (дата обращения: 17.07.2017).
9. Wendl M., Smith S., Pohl C., Dooling D., Chinwalla A., Crouse K., Hepler T., Leong S., Carmichael L., Nhan M., Oberkfell B., Mardis E., Hillier L., Wilson R. Design and implementation of a generalized laboratory data model. *BMC Bioinformatics*, 2007, vol. 8, no. 362. DOI: 10.1186/1471-2105-8-362. URL: <https://bmcbioinformatics.biomedcentral.com/articles/10.1186/1471-2105-8-362> (дата обращения: 17.07.2017).
10. Jayashree B., Reddy P.T., Leeladevi Y., Crouch J.H., Mahalakshmi V., Buhariwalla H.K., Eshwar K.E., Mace E., Folkertsma R., Senthilvel S., Varshney R.K., Seetha K., Rajalakshmi R., Prasanth V.P., Chandra S., Swarupa L., Srikalyani P., Hoisington D.A. Laboratory information management software for genotyping workflows: applications in high throughput crop genotyping. *BMC Bioinformatics*, 2006, vol. 7, no. 383. DOI: 10.1186/1471-2105-7-383. URL: <https://bmcbioinformatics.biomedcentral.com/articles/10.1186/1471-2105-7-383> (дата обращения: 17.07.2017).
11. Orro A., Guffanti G., Salvi E., Macciardi F., Milanese L. SNPLims: a data management system for genome wide association studies. *BMC Bioinformatics*, 2008, vol. 9, no. 2. URL: <https://bmcbioinformatics.biomedcentral.com/articles/10.1186/1471-2105-9-S2-S13> (дата обращения: 17.07.2017).
12. Schönherr S., Weissensteiner H., Coassin S., Specht G., Kronenberg F., Brandstätter A. eCOMPAGT – efficient Combination and Management of Phenotypes and Genotypes for Genetic Epidemiology. *BMC Bioinformatics*, 2009, vol. 10, no. 139. URL: <https://bmcbioinformatics.biomedcentral.com/articles/10.1186/1471-2105-10-139> (дата обращения: 17.07.2017).
13. Weissensteiner H., Schönherr S., Specht G., Kronenberg F., Brandstätter A. eCOMPAGT integrates mtDNA: import, validation and export of mitochondrial DNA profiles for population genetics, tumour dynamics and genotype-phenotype association studies. *BMC Bioinformatics*, 2010, vol. 11, no. 122. URL: <https://bmcbioinformatics.biomedcentral.com/articles/10.1186/1471-2105-11-122> (дата обращения: 17.07.2017).
14. Dunca S., Sirkanungo R., Miller L., Phillips G. J. DraGNET: Software for storing, managing and analyzing annotated draft genome sequence data. *BMC Bioinformatics*, 2010, vol. 11, no. 100. URL: <https://bmcbioinformatics.biomedcentral.com/articles/10.1186/1471-2105-11-100> (дата обращения: 17.07.2017).
15. Truong C.V., Groeneveld L.F., Morgenstern B., Groeneveld E. MolabIS – An integrated information system for storing and managing molecular genetics data. *BMC Bioinformatics*, 2011, vol. 12, no. 425. DOI: 10.1186/1471-2105-12-425. URL: <https://bmcbioinformatics.biomedcentral.com/articles/10.1186/1471-2105-12-425> (дата обращения: 17.07.2017).

Software & Systems

DOI: 10.15827/0236-235X.031.1.204-211

Received 07.08.17

2018, vol. 31, no. 1, pp. 204–211

DEVELOPMENT OF A SOFTWARE PACKAGE FOR THE REGISTRATION OF LABORATORY ANIMALS

*E.Yu. Furta*¹, Postgraduate Student, lana-furta@rambler.ru*I.M. Shabalina*¹, Ph.D. (Engineering), Associate Professor, irinashabalina09@gmail.com¹ *Petrozavodsk State University, Lenina Ave. 33, Petrozavodsk, 185910, Russian Federation*

Abstract. Research in molecular genetics are associated with accumulation and storage of a large amount of data obtained during experiments. The collected data are used for further analysis and validation of the study results. Currently global projects and resources for storage and processing genome sequencing results are developing (such as "Human Genome", "1000 Genomes" and others). Such projects provide data and research to a wide range of users. However, every study needs its own storage format and software for processing these data. When people work with living organisms at the laboratory, they need to organize animals registration, considering the norms and standards applicable in this field. The software system for this purpose should be consistent with the research objectives and provide the opportunity to conduct a comprehensive data analysis and reporting. This article describes a software system for recording laboratory animals, which is suitable for research carried out

by the Laboratory of Molecular Genetics of the innate immunity of Petrozavodsk State University. In addition to the storing data on animals, the system includes the methods of storage and visualization of the animals' pedigree that allows tracking changes in animal's genotype for several generations. The proposed system allows analyzing animal survival and making statistical processing of data genotyping. This program provides a record of all laboratory animals involved in the experiment, the animal ancestors and descendant's record. It also provides the animal genetic data basic statistical processing possibility and contains various reference books and reports, a user-friendly interface. Currently, the software system is put into operation. The database contains data on more than 5,000 animals.

Keywords: genetic research, heredity, accounting laboratory animals, genetic data storage.

References

1. Volkova T., Furta E., Dmitrieva O., Shabalina I. Pattern building methods in genetic data processing. *Jour. on Selected Topics in Nano Electronics and Computing*. 2014, vol. 2, no. 1, pp. 2–6. Available at: http://jstnec.petrsu.ru/journal/article_en.php?id=3041&seq (accessed Augus 06, 2017).
2. *Uchet laboratornykh zhivotnykh v meditsinskikh uchrezhdeniyakh* [Registration of Laboratory Animals in Healthcare Settings] Available at: <http://economic-club.ru/st/uchet-laboratornyx-zhivotnyx-v-medicinskix-uchrezhdeniyax> (accessed Augus 06, 2017).
3. Kokocinski F., Wrobel G., Hahn M., Lichter P. QuickLIMS: facilitating the data management for DNA-microarray fabrication. *Bioinformatics*. 2003, vol. 19, no. 2, pp. 283–284.
4. Swertz M.A., de Brock E.O., van Hijum S.A. F. T., de Jong A., Buist G., Baerends R.J.S., Kok J., Kuipers O.P., Jansen R.C. Molecular Genetics Information System (MOLGENIS): alternatives in developing local experimental genomics databases. *Bioinformatics*. 2004, vol. 20, no. 13, pp. 2075–2083.
5. Monnier S., Cox D.G., Albion T., Canzian F. T.I.M.S: TaqMan Information Management System, tools to organize data flow in a genotyping laboratory. *BMC Bioinformatics*. 2005, vol. 6, no. 246. Available at: <https://bmcbioinformatics.biomedcentral.com/articles/10.1186/1471-2105-6-246> (accessed July 17, 2017).
6. Goh C.S., Lan N., Echols N., Douglas S.M., Milburn D., Bertone P., Xiao R., Chung Ma L., Zheng D., Wunderlich Z., Acton T., Montelione G.T., Gerstein M. SPINE 2: a system for collaborative structural proteomics within a federated database framework. *Nucleic Acids Res*. 2003, vol. 31, no. 11, pp. 2833–2838.
7. Morisawa H., Hirota M., Toda T. Development of an open source laboratory information management system for 2-D gel electrophoresis-based proteomics workflow. *BMC Bioinformatics*. 2006, vol. 7, no. 430. Available at: <https://bmcbioinformatics.biomedcentral.com/articles/10.1186/1471-2105-7-430> (accessed July 17, 2017).
8. Droit A., Hunter J., Rouleau M., Ethier C., Picard-Cloutier A., Bourgeois D., Poirier G. PARPs database: A LIMS systems for protein-protein interaction data mining or laboratory information management system. *BMC Bioinformatics*. 2007, vol. 8, no. 483. Available at: <https://bmcbioinformatics.biomedcentral.com/articles/10.1186/1471-2105-8-483> (accessed July 17, 2017).
9. Wendl M., Smith S., Pohl C., Dooling D., Chinwalla A., Crouse K., Hepler T., Leong S., Carmichael L., Nhan M., Oberkfell B., Mardis E., Hillier L., Wilson R. Design and implementation of a generalized laboratory data model. *BMC Bioinformatics*. 2007, vol. 8, no. 362. Available at: <https://bmcbioinformatics.biomedcentral.com/articles/10.1186/1471-2105-8-362> (accessed July 17, 2017).
10. Jayashree B., Reddy P.T., Leeladevi Y., Crouch J.H., Mahalakshmi V., Buhariwalla H.K., Eshwar K.E., Mace E., Folkertsma R., Senthilvel S., Varshney R.K., Seetha K., Rajalakshmi R., Prasanth V. P., Chandra S., Swarupa L., Srikalyani P., Hoisington D.A. Laboratory Information Management Software for genotyping workflows: applications in high throughput crop genotyping. *BMC Bioinformatics*. 2006, vol. 7, no. 383. Available at: <https://bmcbioinformatics.biomedcentral.com/articles/10.1186/1471-2105-7-383> (accessed July 17, 2017).
11. Orro A., Guffanti G., Salvi E., Macchiardi F., Milanese L. SNPLims: a data management system for genome wide association studies. *BMC Bioinformatics*. 2008, vol. 9, no. 2. Available at: <https://bmcbioinformatics.biomedcentral.com/articles/10.1186/1471-2105-9-S2-S13> (accessed July 17, 2017).
12. Schönherr S., Weissensteiner H., Coassin S., Specht G., Kronenberg F., Brandstätter A. eCOMPAGT – efficient Combination and Management of Phenotypes and Genotypes for Genetic Epidemiology. *BMC Bioinformatics*. 2009, vol. 10, no. 139. Available at: <https://bmcbioinformatics.biomedcentral.com/articles/10.1186/1471-2105-10-139> (accessed July 17, 2017).
13. Weissensteiner H., Schönherr S., Specht G., Kronenberg F., Brandstätter A. eCOMPAGT integrates mtDNA: import, validation and export of mitochondrial DNA profiles for population genetics, tumour dynamics and genotype-phenotype association studies. *BMC Bioinformatics*. 2010, vol. 11, no. 122. Available at: <https://bmcbioinformatics.biomedcentral.com/articles/10.1186/1471-2105-11-122> (accessed July 17, 2017).
14. Dunca S., Sirkanungo R., Miller L., Phillips G.J. DraGnET: Software for storing, managing and analyzing annotated draft genome sequence data. *BMC Bioinformatics*. 2010, vol. 11, no. 100. Available at: <https://bmcbioinformatics.biomedcentral.com/articles/10.1186/1471-2105-11-100> (accessed July 17, 2017).
15. Truong C.V., Groeneveld L.F., Morgenstern B., Groeneveld E. MolabIS – An integrated information system for storing and managing molecular genetics data. *BMC Bioinformatics*. 2011, vol. 12, no. 425. Available at: <https://bmcbioinformatics.biomedcentral.com/articles/10.1186/1471-2105-12-425> (accessed July 17, 2017).