

ПРОГРАММНОЕ ОБЕСПЕЧЕНИЕ

**“ПОИСКОВАЯ ПРОТЕОМНАЯ
МАШИНА IdentiPROT”**

v. 1.0.1

Ограниченная гарантия

ВНИМАНИЕ! ВНИМАТЕЛЬНО ПРОЧИТАЙТЕ ЭТУ ОГРАНИЧЕННУЮ ГАРАНТИЮ, ЧТОБЫ ПОНЯТЬ СВОИ ПРАВА И ОБЯЗАННОСТИ.

Программы предназначены и предоставляются в качестве программного обеспечения общего назначения, а не для какой-либо конкретной цели пользователя. Вы соглашаетесь с тем, что любое программное обеспечение может содержать ошибки, поэтому настоятельно рекомендуем регулярно создавать резервные копии своих файлов. Разработчики не несут ответственности за потерю, повреждение, издержки и затраты, понесенные вами или третьим лицом в результате неправильного использования Программы, включая без ограничений обязательства за торговые расходы, простой оборудования, потери, понесенные вами или третьим лицом в результате отсутствия, неисправности, вируса, ошибки или нарушения работы Программы.

Сохранение права собственности на Программы.

Программы защищены законами и соглашениями об авторском праве и других правах на интеллектуальную собственность. Разработчикам принадлежат все авторские права на Программы.

Прочие права и ограничения.

- Изменение Программы допускается, при условии заключения соглашения на условиях партнерства - Partner Collaborative License. По вопросам заключения соглашения обращаться okeresearch@gmail.com.
- **Прокат.** Не разрешается давать программу в прокат или во временное использование.

СОДЕРЖАНИЕ

I. Руководство администратора сервера хранения и обработки пользовательских данных с использованием протеомной поисковой машины IdentiPROT	7
ВВЕДЕНИЕ	9
ИСПОЛЬЗУЕМЫЕ ОПРЕДЕЛЕНИЯ, ОБОЗНАЧЕНИЯ И СОКРАЩЕНИЯ	9
1 ОБЩИЕ СВЕДЕНИЯ	10
1.1 Основные принципы работы сервера	10
1.1.1 Основные понятия	10
1.1.2 Архитектура серверного программного обеспечения	10
1.1.3 Квалификационные требования к администратору сервера	11
2 СИСТЕМНЫЕ ТРЕБОВАНИЯ СЕРВЕРА ППМ IdentiPROT	12
2.1 Минимальная конфигурация сервера	12
2.2 Рекомендуемая конфигурация сервера	12
2.3 Минимальная конфигурация клиентского места	12
2.4 Рекомендуемая конфигурация клиентского места	12
2.5 Требования к операционной системе	13
3 УСТАНОВКА И УДАЛЕНИЕ ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ ППМ IdentiPROT	13
3.1 Программное обеспечение сервера ППМ IdentiPROT	13
3.2 Установка программного обеспечения ППМ IdentiPROT	13
3.3 Удаление программного обеспечения ППМ IdentiPROT	14
4 УПРАВЛЕНИЕ КОМПОНЕНТАМИ СЕРВЕРА	15
4.1 Конфигурация серверного ПО ППМ IdentiPROT	15
4.2 Настройка доступа к интерфейсу ППМ IdentiPROT	16
4.3 Диагностика серверного модуля хранения пользовательских данных	16
4.4 Создание пользовательских учётных записей	17
Лист регистрации измерений	18
II. Руководство пользователя сервера хранения и обработки пользовательских данных с использованием протеомной поисковой машины IdentiPROT	19
АННОТАЦИЯ	21
ВВЕДЕНИЕ	21
Область применения	21
Краткое описание возможностей	21
Уровень подготовки пользователя	22

Перечень необходимой эксплуатационной документации	22
Принятые соглашения	22
Описание вызова команд	22
Иллюстрации	22
Язык веб-интерфейса ППМ IdentiPROT	22
Назначение и условия применения	23
Назначение	23
Условия применения	23
1 ПОДГОТОВКА К РАБОТЕ	23
1.1 Состав и содержание дистрибутивного носителя данных	23
1.2 Порядок загрузки данных и программ	23
1.3 Порядок проверки работоспособности ПО	23
2 ОПИСАНИЕ ОПЕРАЦИЙ	24
2.1 Доступ к веб-интерфейсу ППМ IdentiPROT	24
2.2 Вход в систему	24
2.3 Выход из системы	25
2.4 Загрузка данных	25
2.4.1 Форматы и имена файлов	26
2.5 Запуск поиска	26
2.6 Выбор файлов	27
2.7 Выбор базы данных	28
2.8 Выбор параметров поиска	28
2.8.1 Выбор уровня фильтрации результатов	28
2.8.2 Сохранение использованных параметров	29
2.8.3 Использование ранее сохранённых параметров	30
2.8.4 Задание модификаций	30
2.8.5 Создание новых модификаций	30
3 ПРОСМОТР СПИСКА ЗАПУЩЕННЫХ ПОИСКОВ	31
4 ПРОСМОТР РЕЗУЛЬТАТОВ ПОИСКА	31
4.1 Визуализация результатов поиска	32
4.2 Просмотр списков идентификаций	34
4.2.1 Просмотр списка спектральных идентификаций для данного пептида	34
4.2.2 Просмотр списка спектральных или пептидных идентификаций для данного белка	35
4.2.3 Просмотр информации о белке в базе данных Uniprot	35
4.3 Загрузка результатов на компьютер пользователя	35
5 АВАРИЙНЫЕ СИТУАЦИИ	36
5.1 Отсутствие доступа к веб-интерфейсу ППМ IdentiPROT	36

5.2 Сообщения об ошибках (коды 5xx) при работе с веб-интерфейсом ППМ	36
5.3 Рекомендации по освоению	36
6 ПОДДЕРЖКА И КОНТАКТНАЯ ИНФОРМАЦИЯ	36

**Руководство администратора сервера хранения и обработки
пользовательских данных с использованием
протеомной поисковой машины IdentiPROT**

ВВЕДЕНИЕ

Данный документ содержит инструкцию по управлению компонентами сервера хранения и обработки пользовательских данных с использованием протеомной поисковой машины IdentiPROT, реализующей Интернет-доступный сервис для хранения и обработки данных хромато-масс-спектрометрических измерений.

В случае существенных модификаций в программном обеспечении сервера и его конфигурации, предприятие-изготовитель предоставляет обновленную редакцию данной документации.

ИСПОЛЬЗУЕМЫЕ ОПРЕДЕЛЕНИЯ, ОБОЗНАЧЕНИЯ И СОКРАЩЕНИЯ

IdentiPROT – прикладное программное обеспечение для обработки данных панорамированного протеомного анализа

ППМ – протеомная поисковая машина

ПО – программное обеспечение

БД – база данных

1 ОБЩИЕ СВЕДЕНИЯ

1.1. Основные принципы работы сервера

1.1.1 Основные понятия

В Руководстве используются следующие основные понятия:

Сервер ППМ IdentiPROT – компьютер под управлением операционной системы Windows или Linux с установленным программным обеспечением ППМ IdentiPROT.

Клиентское место – рабочее место пользователя ППМ IdentiPROT, которое можно запустить на любом компьютере, имеющем связь с сервером ППМ IdentiPROT через локальную сеть или сеть Интернет, или непосредственно на сервере. Установленная операционная система на компьютере - Windows XP (SP2 или SP3), Windows 7, Windows 8, Windows 10 (как 32, так и 64-битные), а также Mac OS или Linux. Количество клиентских мест на сервере неограниченно.

Логин – уникальное имя клиентского места на сервере. Как правило, адрес электронной почты пользователя, предоставленный администратору сервера.

Пароль – уникальный код доступа пользователя к клиентскому месту на сервере с идентификатором Логин. Пароль предоставляется пользователю администратором сервера.

Серверный модуль хранения пользовательских данных, представленных как в инструментальных, так и стандартизированных форматах, включая результаты обработки экспериментальных данных ППМ IdentiPROT.

Серверный модуль обработки пользовательских данных – область размещения ППМ IdentiPROT, доступная для обращения из клиентского места.

1.1.2 Архитектура серверного программного обеспечения

В состав серверного ПО IdentiPROT входит:

- ядро ППМ IdentiPROT, реализующее два программных модуля: модуль обработки данных и модуль вывода и анализа результатов обработки;
- графический пользовательский интерфейс для настройки параметров работы ППМ и ввода данных;

Ядро ППМ реализовано на языке программирования Python 2.7 и содержит точку входа, к которой происходит обращение из пользовательского интерфейса. Ядро ППМ включает в себя программные модули обработки данных масс-спектрометрических измерений, представленных в стандартизированных форматах mzML и MGF, а также ПО для анализа результатов обработки данных.

Пользовательский интерфейс реализован с помощью языка программирования Python 2.7 и веб-фреймворка Django 1.8 и служит для настройки параметров работы ППМ и ввода данных, а также вывода результатов обработки данных.

В состав пользовательского интерфейса входят два приложения:

- приложение для загрузки и хранения данных (datasets);
- приложение для анализа протеомных данных (identiprot_server).

Каждое из вышеупомянутых приложений организовано согласно архитектуре “Модель - Представление - Контроллер” и требованиям платформы Django. Соответственно, они содержат модули моделей, модули представлений и шаблоны веб-страниц.

Серверное ПО IdentiPROT предназначено для обращения пользователя к серверу при реализации Интернет-доступного сервиса для хранения и обработки данных хромато-масс-спектрометрических измерений. Для реализации этой функциональности серверное ПО работает по протоколу WSGI связи с веб-сервером общего назначения, таким как Apache. Веб-сервер должен быть настроен на делегирование HTTP-запросов по заданному доменному имени через WSGI-приложение indetiprot_server. Приложение identiprot_server содержит файл wsgi.py, путь к которому должен быть указан в конфигурационном файле веб-сервера.

1.1.3 Квалификационные требования к администратору сервера

Для работы с сервером Администратору следует ознакомиться с настоящим руководством. Для работы необходимы базовые навыки администрирования соответствующей ОС (Windows или Linux). Требуется умение производить установку совместимых версий библиотек языка Python 2 и другого ПО на серверной ОС, а также навыки установки, запуска и наладки, администрирования и конфигурации веб-сервера. Администратору требуются базовые знания о принципах работы компьютерных сетей.

Администратору сервера рекомендуется ознакомиться с руководством пользователя на ПО ППМ IdentiPROT.

2 СИСТЕМНЫЕ ТРЕБОВАНИЯ СЕРВЕРА ППМ IdentiPROT

2.1 Минимальная конфигурация сервера

- ОС Linux;
- Оперативная память: от 2 Гб;
- Тактовая частота процессора: от 1.5 Гц;
- Свободное место на диске: 1 Гб на системном диске (разделе) для установки ППМ IdentiPROT, библиотек функций и вспомогательных программ и 1 Тб на серверном модуле хранения пользовательских данных, определяемое типичным размером экспериментальных данных панорамного хромато-масс-спектрометрического анализа 12 биологических проб ткани человека (или фракций одной пробы) с тремя техническими повторами при стандартных условиях хроматографического разделения (4-часовой градиент ВЭЖХ на один технический повтор);
- Подключение к сети Интернет.

2.2 Рекомендуемая конфигурация сервера

- ОС Linux 64 bit;
- Процессор: Intel Core i7 или аналогичный, с тактовой частотой от 2 Гц;
- Оперативная память: от 32 Гб;
- Свободное место на диске: 1 Гб на системном диске (разделе) для установки ППМ IdentiPROT, библиотек функций и вспомогательных программ и от 4 Тб на серверном модуле хранения пользовательских данных;
- Подключение к сети пропускной способностью от 100 Мбит/с.

2.3 Минимальная конфигурация клиентского места

Поддержка работы современного веб-браузера (Internet Explorer, Mozilla Firefox, Google Chrome).

2.4 Рекомендуемая конфигурация клиентского места

- Процессор: от 1 ГГц;
- Оперативная память: от 2 Гб;
- Свободное место на жестком диске: от 100 Гб.

Возможна установка клиентского места и сервера на одном ПК, при этом следует придерживаться рекомендуемой конфигурации для сервера, описанной в п. 4.2.

Для комфортной работы пользователя ППМ IdentiPROT с данными протеомного анализа рекомендуется обеспечить скорость работы сети между клиентским местом и сервером не менее 100 Мбит/с.

2.5 Требования к операционной системе

Установка ПО ППМ IdentiPROT производится на компьютеры под управлением операционной системы Linux 64 bit.

Возможны комбинации сервера и клиентского места на удаленном компьютере под управлением разных ОС, включая ОС Mac OS для клиентского места.

Независимо от типа используемой операционной системы, необходима установка на нее последних обновлений, выпущенных производителем для данной ОС.

3 УСТАНОВКА И УДАЛЕНИЕ ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ ППМ IdentiPROT

3.1 Программное обеспечение сервера ППМ IdentiPROT

Программное обеспечение сервера ППМ IdentiPROT построено на основе клиент-серверной архитектуры и состоит из двух частей: ядра ППМ IndetiPROT и веб-приложений. Веб-приложения `identiprot_server` и `datasets` предоставляют доступ всем программным компонентам ППМ IdentiPROT к соответствующим экспериментальным данным, загруженным на сервер из клиентского места, а также к результатам обработки данных. Веб-приложение `identiprot_server` является точкой входа, доступ к которой обеспечивается настройкой WSGI-совместимого веб-сервера.

Для работы на сервере необходимо наличие зарегистрированного клиентского места. Работа с ППМ IdentiPROT на сервере инициализируется из клиентского места после ввода комбинации логина и пароля.

Программным обеспечением клиентского места является Интернет-обозреватель, используемый на удаленном компьютере, имеющем связь с сервером по сети. Возможна установка клиентского места непосредственно на сервер ППМ IdentiPROT.

3.2 Установка программного обеспечения ППМ IdentiPROT

Серверная часть ПО ППМ IndetiPROT может работать под управлением различных операционных систем и требует для работы только установленной реализации языка Python 2.7 и совместимых с ней версий следующих библиотек:

- `setuptools`
- `pip`
- `numpy`

- scipy
- matplotlib
- django
- mechanize
- lxml
- jinja2
- pyteomics.biolccc
- pyteomics.pepxmltk
- django-multiupload

Для повышения производительности ядра ППМ IdentiPROT также рекомендуется установить следующее ПО:

- cython
- pyteomics.cythonize

Процедура установки ПО зависит от используемой серверной ОС. В случае использования ОС Linux рекомендуется использовать системные репозитории. Пакеты, отсутствующие в системных репозиториях, могут быть установлены с помощью утилиты pip.

В случае использования ОС Windows рекомендуется установить последнюю версию Python 2.7 с сайта <http://python.org>, установочные файлы остальных компонентов или их исходный код могут быть загружены с соответствующих веб-сайтов, либо установка может быть произведена с помощью утилиты pip.

По усмотрению администратора установка может быть произведена внутри виртуального окружения (virtualenv). В этом случае требуется установка ПО Python 2.7 и virtualenv, а остальные необходимые библиотеки могут быть установлены внутри виртуального окружения, что уже не требует прав администратора. Однако, права администратора в любом случае могут потребоваться для установки и корректной настройки WSGI-совместимого веб-сервера.

3.3 Удаление программного обеспечения ППМ IdentiPROT

Программные файлы серверного ПО ППМ IdentiPROT при установке не копируются в системные директории серверной ОС, поэтому специальных действий по удалению ПО не требуется. Для прекращения работы сервера ППМ IdentiPROT достаточно удалить соответствующую запись в конфигурационном файле используемого веб-сервера. Вспомогательное ПО может быть удалено администратором с помощью системного менеджера пакетов или утилиты pip, в зависимости от способа установки.

4 УПРАВЛЕНИЕ КОМПОНЕНТАМИ СЕРВЕРА

4.1 Конфигурация серверного ПО ППМ IdentiPROT

Конфигурация серверного ПО ППМ IdentiPROT для корректной работы с веб-сервером и серверным оборудованием производится путём редактирования файла конфигурации `settings.py` в корневой директории приложения `identiprot_server`. Стандартные переменные, используемые для конфигурации приложений Django, описаны в документации на сайте <https://docs.djangoproject.com/en/1.8/ref/settings>. Дополнительные переменные, которые должны быть определены в конфигурационном файле, связанных с доставкой сообщений, отправленных пользователями через форму поддержки:

- `EMAIL_USE_TLS` (True или False)
- `EMAIL_HOST` (str)
- `EMAIL_PORT` (int)
- `EMAIL_HOST_USER` (str)
- `EMAIL_HOST_PASSWORD` (str)
- `EMAIL_SEND_TO` (list)

Эти переменные содержат данные, необходимые для аутентификации на сервере `EMAIL_HOST` под именем `EMAIL_HOST_USER` и паролем `EMAIL_HOST_PASSWORD`, с которого пользовательские сообщения будут отправляться на адреса из списка `EMAIL_SEND_TO`. Данные адреса должны принадлежать администраторам сервера IdentiPROT. Примеры всех значений указаны в поставляемой версии файла `settings.py`.

Ещё одна переменная, `NUMBER_OF_PARALLEL_RUNS`, принимает целочисленные значения и означает количество одновременно обрабатываемых поисков. В случае, если пользователями запускается большее количество поисков, они ставятся в очередь и выполняются последовательно. Значение `NUMBER_OF_PARALLEL_RUNS` выбирается администратором исходя из количества доступной оперативной памяти и типичного размера файлов с хромато-масс-спектрометрическими данными, обрабатываемых пользователями. В случае невозможности выделения оперативной памяти в достаточном количестве в работе сервера произойдёт ошибка, и один или несколько поисков не будут завершены. Статус незавершённых поисков может быть обновлён не ранее перезапуска сервера.

Перед первым запуском серверного ПО необходимо также выполнить следующие действия:

1. Перейти в поддиректорию `identiprot_server` директории с исходным кодом ПО;

2. Выполнить команды:

```
python2 manage.py makemigrations
```

```
python2 manage.py migrate,
```

где `python2` — имя исполняемого файла интерпретатора Python 2.7.

4.2 Настройка доступа к интерфейсу ППМ IdentiPROT

Управление доступом к интерфейсу ППМ IdentiPROT осуществляется путём настройки используемого WSGI-совместимого веб-сервера. Пример записи в конфигурационном файле веб-сервера Apache с использованием виртуальных хостов:

```
<VirtualHost *:80>
```

```
    WSGIScriptAlias / /path/to/identiprot_server/identiprot_server/wsgi.py
```

```
    WSGIDaemonProcess identiprot python-path=/path/to/identiprot_server:/path/to/identiprot
user=admin group=admin threads=5
```

```
    WSGIProcessGroup identiprot
```

```
    ServerAdmin admin@lab.ru
```

```
    ServerName identiprot.domainname.ru
```

```
    Alias /static/ /path/to/static/
```

```
    <Directory "/path/to/identiprot_server/identiprot_server/">
```

```
        Order deny,allow
```

```
        Require all granted
```

```
    </Directory>
```

```
    <Directory "/path/to/static/">
```

```
        Order deny,allow
```

```
        Require all granted
```

```
    </Directory>
```

```
</VirtualHost>
```

4.3 Диагностика серверного ПО ППМ

Диагностическая информация о работе серверного ПО ППМ IdentiPROT выводится в файл журнала ошибок веб-сервера. Расположение файла настраивается путём редактирования конфигурационного файла веб-сервера.

При корректной работе серверного ПО ППМ IdentiPROT при запросе сконфигурированного URL через веб-браузер должна отображаться веб-страница с формой входа или экраном начала поиска. Если запрос приводит к ошибке 404, то веб-сервер

настроен некорректно. Если происходит ошибка 403, то вероятная причина — некорректно настроенные права доступа процесса веб-сервера к файлам с исходным кодом ППМ. В случае возникновения ошибки 500 следует изучить журнал ошибок веб-сервера, либо изменить значение конфигурационной переменной DEBUG в файле settings.py с False на True и повторить запрос. При этом вместо стандартной страницы ошибки 500 будет отображена подробная информация об ошибке. Если тип ошибки — ImportError, то часть вспомогательного ПО может быть не установлена или установлена некорректно. В противном случае, следует убедиться, что пп. 3.2 и 4.1 настоящего Руководства выполнены полностью и при повторении ошибки обратиться к поставщику ПО ППМ, сообщив полную информацию об ошибке, отображаемую на странице при включённом параметре DEBUG.

4.4 Создание пользовательских учётных записей

Для создания пользовательской учётной записи следует перейти в поддиректорию `identiprot_server` и выполнить следующую команду:

```
python2 manage.py createuser user@email.org userpassword
```

где `python2` — исполняемый файл интерпретатора Python 2.7, `user@email.org` — адрес электронной почты пользователя, `userpassword` — пароль учётной записи.

**Руководство пользователя сервера хранения и обработки
пользовательских данных с использованием протеомной поисковой
машины IdentiPROT**

АННОТАЦИЯ

Назначение протеомной поисковой машины IdentiPROT программного обеспечения: инструмент для обработки данных протеомного анализа методами хромато-масс-спектрометрии и идентификация пептидов и белков.

Назначение документа: представить руководство пользователя (исследователя, оператора, аналитика) для работы с протеомной поисковой машиной IdentiPROT. Документ содержит перечень основных функциональных возможностей ППМ IdentiPROT и подробное описание необходимых действий пользователя для использования каждой из функциональных возможностей.

ВВЕДЕНИЕ

Область применения

Программное обеспечение “Поисковая протеомная машина (ППМ) IdentiPROT” предназначено для обработки данных протеомного анализа биологических проб, получаемых методами хромато-масс-спектрометрии, и идентификации на основе результатов обработки пептидов и белков, присутствовавших в пробах. Доступ пользователя к ППМ IdentiPROT и работа с данными осуществляется через веб-интерфейс (веб-интерфейс ППМ IdentiPROT), в котором реализовано управление всеми программным модулям ППМ, необходимыми для работы с данными. Пользователи могут использовать собственные пространства работы с данными (личные кабинеты), анализировать результаты идентификации белков и пептидов, а также передавать по сети как сами данные, так и результаты обработки, на локальные ПК.

Краткое описание возможностей

Серверное ПО ППМ IdentiPROT обеспечивает следующие возможности:

- авторизация пользователя на сервере;
- загрузка на сервер данных панорамного хромато-масс-спектрометрического протеомного анализа в формате MGF и mzML;
- загрузка на сервер протеомных баз данных в формате FASTA;
- настройка параметров идентификации пептидов и белков;
- запуск процесса идентификации пептидов и белков (поиска);
- просмотр информации о запущенных поисках;
- просмотр результатов завершённого поиска;
- загрузка результатов поиска в различных форматах на ПК пользователя.

Уровень подготовки пользователя

Для работы пользователь должен обладать следующими навыками:

- базовые навыки работы с персональным компьютером;
- работа с любым интернет-обозревателем;
- работа с файлами формата CSV;
- опыт работы со следующими информационными системами или аналогами: UniProt, PubMed

Перечень необходимой эксплуатационной документации

Для работы с ППМ IdentiPROT пользователю следует ознакомиться с настоящим руководством.

Принятые соглашения

Описание вызова команд

Для использования той или иной функции системы, пользователь вызывает соответствующую команду. Во многих случаях одна и та же команда может быть вызвана несколькими способами. Для краткости изложения в тексте настоящего документа там, где это возможно, предлагается использование наиболее простого способа.

Иллюстрации

Приведённые в документе иллюстрации отражают вид веб-страниц и элементов графического интерфейса в интернет-обозревателе Chromium 47.0. При просмотре страниц сайта в другом интернет-обозревателе, вид страниц и элементов графического интерфейса может незначительно отличаться от приведенного в данном документе.

Язык веб-интерфейса ППМ IdentiPROT

Язык веб-интерфейса программного обеспечения ППМ IdentiPROT - английский. В настоящем руководстве названия элементов управления приводятся на английском языке.

Назначение и условия применения

Назначение

Программное обеспечение “Поисковая протеомная машина (ППМ) IdentiPROT” предназначено для автоматизации обработки данных протеомного анализа биологических проб, получаемых методами хромато-масс-спектрометрии с целью идентификации пептидов и белков, присутствовавших в пробах.

Условия применения

Для функционирования клиентской части системы, компьютер пользователя должен быть подключён к сети Интернет. На рабочем месте пользователя должна быть установлен Интернет-обозреватель Internet Explorer, Mozilla Firefox или Google Chrome последней версии. Операционная система ПК пользователя должна позволять использование Интернет-обозревателя.

1 ПОДГОТОВКА К РАБОТЕ

1.1 Состав и содержание дистрибутивного носителя данных

Для функционирования клиентской части системы установка ПО с дистрибутивного носителя данных не требуется.

1.2 Порядок загрузки данных и программ

Для использования клиентской части системы требуется:

1. Убедиться, что на клиентском ПК установлена последняя версия Интернет-обозревателя Google Chrome, Mozilla Firefox или Internet Explorer.
2. Загрузить на клиентский ПК данные результатов хромато-масс-спектрометрического анализа в стандартизованных форматах mzML и MGF. Если данные содержатся на внешнем носителе, подключить его к ПК.

1.3 Порядок проверки работоспособности ПО

Для проверки работоспособности ППМ IdentiPROT необходимо:

1. Открыть Интернет-обозреватель;
2. Ввести в адресную строку доменное имя или адрес сервера ППМ;
3. Нажать клавишу Enter.

После этого должна открыться веб-страница интерфейса ППМ. Если этого не происходит, следует обратиться к администратору сервера.

2 ОПИСАНИЕ ОПЕРАЦИЙ

2.1 Доступ к веб-интерфейсу ППМ IdentiPROT

Для доступа к веб-интерфейсу ППМ IdentiPROT пользователю необходимы следующие сведения:

1. Доменное имя или IP-адрес, по которому доступен веб-интерфейс;
2. Имя и пароль учётной записи пользователя, зарегистрированной на сервере.

Эти сведения необходимо получить у администратора сервера.

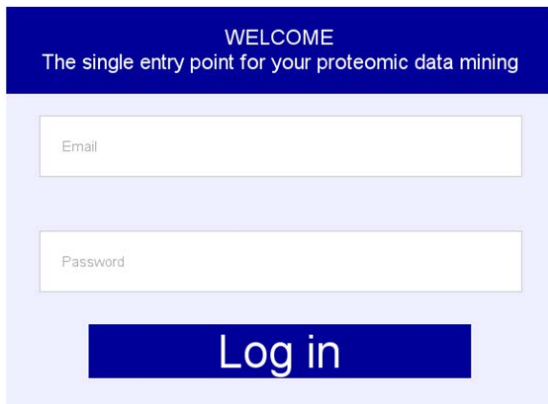
После этого для доступа к веб-интерфейсу необходимо:

1. Открыть Интернет-обозреватель;
2. Ввести в адресную строку доменное имя или адрес сервера ППМ;
3. Нажать клавишу Enter.

После этого должна открыться веб-страница интерфейса ППМ. Если этого не происходит, следует обратиться к администратору сервера.

2.2 Вход в систему

При первом входе по адресу интерфейса ППМ отображается страница, содержащая форму входа. Вход в систему производится путём ввода логина и пароля в поля формы входа и нажатия на кнопку “Log in” (рис.1)



The image shows a login form for the IdentiPROT web interface. At the top, there is a dark blue banner with the text "WELCOME" and "The single entry point for your proteomic data mining". Below this banner are two white input fields: "Email" and "Password". At the bottom of the form is a dark blue button with the text "Log in".

Рис. 1. Форма входа в веб-интерфейс ППМ IdentiPROT

2.3 Выход из системы

Если вход в систему произведён, в правом верхнем углу любого экрана системы будет отображаться имя учётной записи, под которой произведён вход. Рядом с именем отображается кнопка “Log out”, нажатие на которую приведёт к выходу из учётной записи (рис. 2).

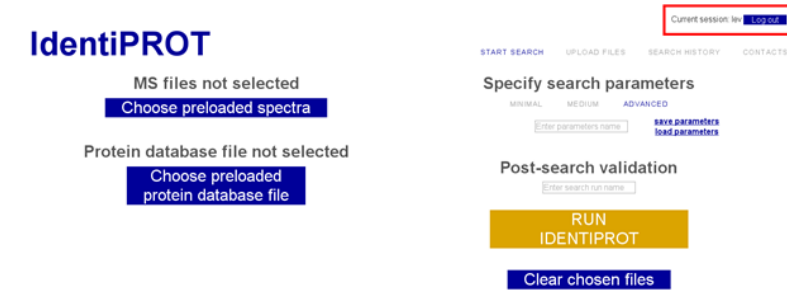


Рис. 2. Имя текущей учётной записи и кнопка выхода.

2.4 Загрузка данных

Для загрузки данных на сервер ППМ следует перейти к экрану “Upload files” по ссылке в верхней панели любого экрана системы. Для выбора файлов, которые необходимо загрузить на сервер, используется кнопка “Choose files”. При нажатии на неё открывается диалоговое окно, в котором нужно выбрать один или несколько файлов и нажать на кнопку “Open” или “Открыть”. После этого необходимо нажать кнопку “Upload files” для загрузки выбранных файлов (рис. 3). После нажатия на кнопку “Upload files” начнётся процесс загрузки. Необходимо дождаться окончания процесса загрузки файлов, прежде чем покинуть страницу. По окончании загрузки на экране появится уведомление об успешной загрузке (рис. 4).

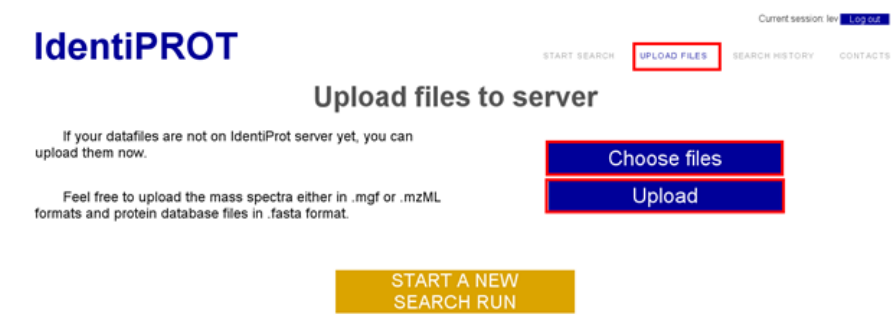


Рис. 3. Экран загрузки файлов.

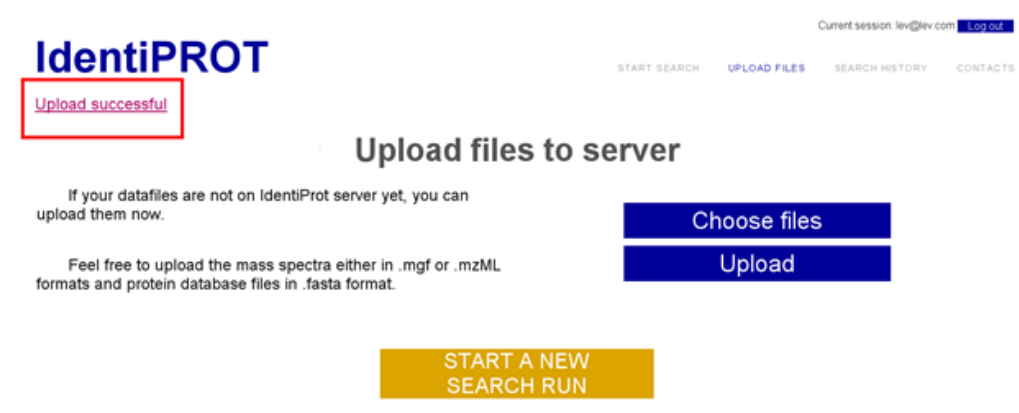


Рис. 4. Уведомление об успешном окончании загрузки файлов на сервер.

2.4.1 Форматы и имена файлов

ППМ IdentiPROT поддерживает форматы хромато-масс-спектрометрических данных MGF и mzML и формат баз данных белковых последовательностей FASTA. При загрузке файлов будут загружены только файлы с соответствующими расширениями (.mgf, .mzml, .fasta).

2.5 Запуск поиска

Для запуска нового поиска необходимо перейти на экран запуска с помощью ссылки "Start Search" в верхней панели ссылок (рис. 5).

После этого требуется:

- выбрать входные файлы для поиска (MGF или mzML);
- выбрать белковую базу данных для поиска (FASTA);
- установить параметры поиска;
- ввести идентификационное имя поиска;
- нажать кнопку "Start Search".

IdentiPROT

MS files not selected
Choose preloaded spectra

Protein database file not selected
Choose preloaded protein database file

Current session: lev@lev.com [Log out](#)

START SEARCH [UPLOAD FILES](#) [SEARCH HISTORY](#) [CONTACTS](#)

Specify search parameters

MINIMAL MEDIUM ADVANCED

Enter parameters name [save parameters](#)
[load parameters](#)

send email notification

use auto optimization

enzyme: [add custom cleavage rule](#) trypsin

number of missed cleavages: 2

precursor accuracy unit: ppm

precursor accuracy left: 10

precursor accuracy right: 10

product accuracy, Da: 0.1

FDR: 1.0

FDR type: psm

minimum charge: 2

maximum charge: 3

generate decoy db: no

decoy method: reverse

decoy prefix: DECOY_

dynamic range: 100

peptide minimum length: 5

peptide maximum length: 50

peptide minimum mass: 300

peptide maximum mass: 10000

fragments in spectra, min: 4

fragments in spectra, max: 50

product minimum m/z: 150

maximum fragment charge: 0

matched fragments, min: 1

use scoring function: hyperscore

score threshold: 0

Рис. 5. Экран запуска поиска.

2.6 Выбор файлов

Для выбора файлов необходимо нажать кнопку “Choose preloaded spectra”. При этом откроется страница со списком ранее загруженных пользователем файлов в форматах MGF и mzML. Пользователь может выбрать один или несколько файлов из списка. После этого необходимо нажать кнопку “Submit” (рис. 6). После нажатия на кнопку “Submit” пользователь будет возвращён на страницу запуска, где будет отображаться информация о выбранных файлах.

IdentiPROT

Current session: lev@lev.com [Log out](#)

[START SEARCH](#) [UPLOAD FILES](#) [SEARCH HISTORY](#) [CONTACTS](#)

Select files:

submit [cancel](#) [delete](#) [upload new files](#)

file1.mgf

file2.mgf

file3.mgf

Рис. 6. Выбор файлов для поиска

2.7 Выбор базы данных

Для выбора базы данных необходимо нажать кнопку “Choose preloaded protein database file”. При этом откроется страница со списком ранее загруженных пользователем файлов в формате FASTA. Пользователь может выбрать один файл из списка. После этого необходимо нажать кнопку “Submit” (рис. 7). После нажатия на кнопку “Submit” пользователь будет возвращён на страницу запуска, где будет отображаться информация о выбранной базе данных.

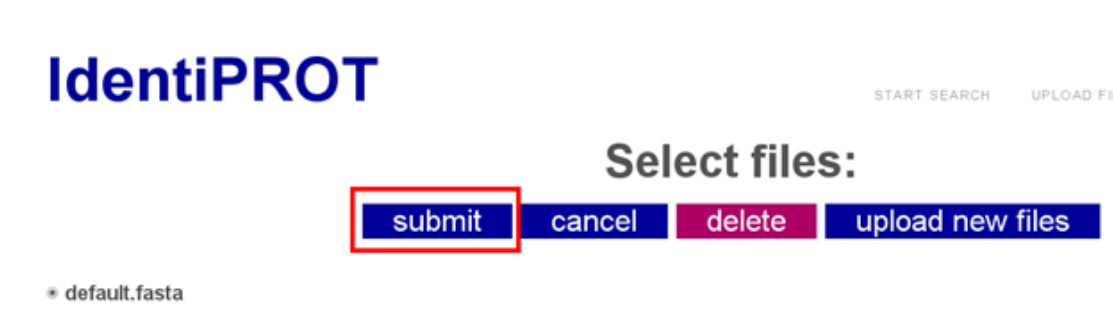


Рис. 7. Выбор базы данных для поиска.

2.8 Выбор параметров поиска

Настройка параметров поиска производится в правой части окна запуска. Веб-интерфейс ППМ предусматривает три режима настройки: минимальный, средний и продвинутый. Переключение между ними осуществляется путём нажатия на надписи “Minimal”, “Medium” и “Advanced” в правой части экрана запуска поиска. После этого следует выставить требуемые значения перечисленных на странице параметров.

2.8.1 Выбор уровня фильтрации результатов

ППМ IdentiPROT реализует фильтрацию результатов поиска методом использования объединённой целевой и декойной базы данных. Уровень достоверности для фильтрации задаётся параметром FDR (в процентах). Ещё одним параметром поиска является FDR type, принимающий значения PSM, peptide и protein. В зависимости от значения этого параметра фильтрация идентификаций будет происходить на спектральном, пептидном или белковом уровне, соответственно.

2.8.3 Использование ранее сохранённых параметров

Чтобы загрузить ранее сохранённые параметры, следует нажать на ссылку “load parameters”. При этом откроется страница со списком ранее сохранённых параметров. После выбора нужного набора и нажатия кнопки “Submit” пользователь будет возвращён к странице запуска поиска. При необходимости загруженные параметры можно отредактировать и/или сохранить под новым именем.

2.8.4 Задание модификаций

Выбор фиксированных и потенциальных модификаций осуществляется путём нажатия на ссылку “select fixed modifications” либо “select potential modifications” на странице запуска поиска, соответственно. При нажатии на ссылку отображается страница со списком предустановленных либо созданных ранее модификаций. После выбора нужных элементов и списка и нажатия кнопки “Submit” пользователь возвращается к странице запуска поиска.

2.8.5 Создание новых модификаций

Чтобы указать фиксированную или потенциальную модификацию, отсутствующую в списке, следует нажать на кнопку “add custom modification” и в открывшейся форме ввести информацию о модификации (рис. 9). После ввода информации и нажатия кнопки “submit” созданная модификация будет доступна для выбора.

Add custom modification

Modification name:

Modification label:

Mass change or formula:

Amino acids:

Рис. 9. Форма ввода информации о модификации.

3 ПРОСМОТР СПИСКА ЗАПУЩЕННЫХ ПОИСКОВ

Для просмотра списка запущенных поисков используется ссылка “Search history” в верхней части любого экрана. На экране отображается список запущенных поисков в обратном хронологическом порядке с указанием названия, времени запуска и статуса поиска (рис. 10). Список можно отфильтровать по части идентификационного имени для быстрого доступа к информации о нужном запуске.

Page 1 of 1 [Search previous runs by name](#)

Delete selected	Search name	Search date	Status
<input type="checkbox"/>	Best search ever	Dec. 8, 2015, 5:59 p.m.	Task is finished

Рис. 10. Список поисков пользователя.

4 ПРОСМОТР РЕЗУЛЬТАТОВ ПОИСКА

Для просмотра результатов завершённого поиска нужно нажать на название нужного поиска на экране истории поисков (рис. 11).

Page 1 of 1 [Search previous runs by name](#)

Delete selected	Search name	Search date	Status
<input type="checkbox"/>	Best search ever	Dec. 8, 2015, 5:59 p.m.	Task is finished

Рис. 11. Выбор поиска для просмотра результатов.

В случае, если поиск производился по одному файлу, пользователю будет сразу представлена визуализация результатов. Если же поиск производился по нескольким файлам, пользователю будет представлена сводная таблица с информацией о результатах поиска по каждому файлу (рис. 12).

Searches: Best search ever

Search name	MSMS	Unfiltered PSMs	FDR, %	FDR type	PSMs	Peptides	Proteins
file1	0	0	1.0	psm	0	0	0
file2	0	0	1.0	psm	0	0	0
union	0	0	1.0	psm	0	0	0

Download: [csv files](#) [pepxml files](#) [mgf files](#)

BACK TO
SEARCH HISTORY

Рис. 12. Таблица результатов поиска по файлам.

4.1 Визуализация результатов поиска

В случае поиска по нескольким файлам, экран визуализации открывается при нажатии на название файла или слово “union” на экране просмотра результатов.

Данный экран визуализирует статистическую информацию, собранную в процессе обработки и валидации сделанных идентификаций (рис. 13).

Отображаемые диаграммы характеризуют статистические свойства идентификаций, полученных при обработке конкретного входного файла, либо объединённого набора идентификаций, полученных при поиске по всем файлам (при нажатии “union”).

Экран содержит информацию о распределениях спектральных идентификаций (раздел “PSM distributions”), пептидных идентификаций (раздел “Peptide distributions”), результатов безметочного количественного анализа (“Label-free quantitation”) и распределение по признакам, используемым для валидации идентификаций алгоритмом MP-score (“MP-score descriptors”).

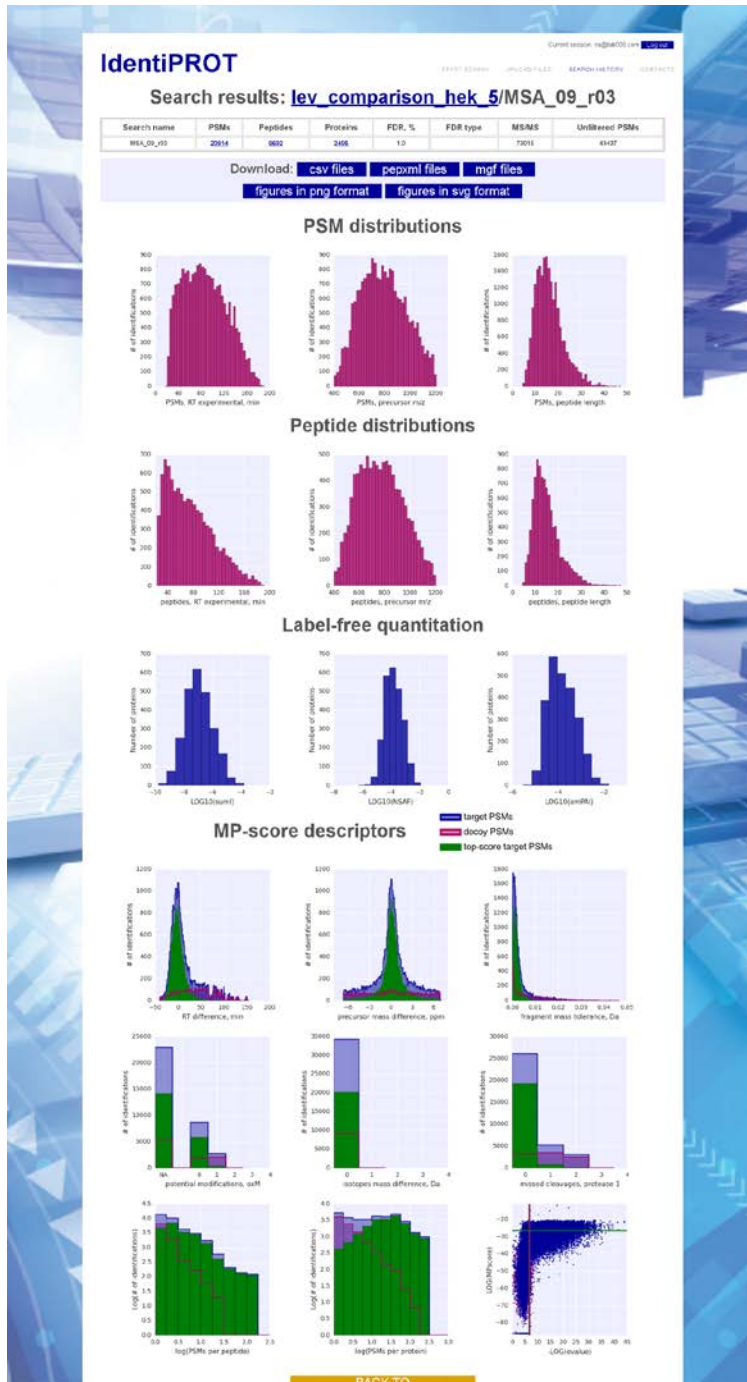


Рис. 13. Экран визуализации результатов поиска.

4.2 Просмотр списков идентификаций

В верхней части экрана визуализации отображается таблица, содержащая количество спектральных, пептидных и белковых идентификаций (рис. 13). При нажатии на количество идентификаций отображается страница со списком этих идентификаций (рис. 14).

Current session: ira@lab006.com [Log out](#)

IdentiPROT [START SEARCH](#) [UPLOAD FILES](#) [SEARCH HISTORY](#) [CONTACTS](#)

Search results: [lev_comparison_hek_5/MSA_09_r01](#)

Search name	PSMs	Peptides	Proteins	FDR, %	FDR type	MS/MS	Unfiltered PSMs
MSA_09_r01	21118	9825	2524	1.0	-	71783	43815

Select columns for psms:

sequence modified_sequence m/z exp m/z error in ppm missed cleavages e-value MPscore RT exp spectrum proteins
 proteins description Sin

[show selected columns](#)

sequence	m/z exp	missed cleavages	RT exp
LSTDKTVK	446.259999533	1	23.383853
ATCAPQHGAPGPGPADASK	895.41840476	0	23.430968
YHTINGHNAEVR	705.84741343	0	23.433886
LEAVSHTSDMHR	691.828509005	0	23.436807
GAPAAATAPAPTAHK	666.357336425	0	23.439728
QVHPDTGISSK	584.801849215	0	23.462488
IHEGCEEPATHNALAK	888.92072379	0	23.481513
AVTEQGHELSNEER	799.873788415	0	23.484436
FQRPQDPQSAQDK	737.35774571	0	23.487357
HATAEEVEEEER	714.81972886	0	23.493205

Рис. 14. Список спектральных идентификаций.

Перемещаться между списками можно, нажимая на ссылки в суммирующей таблице в верхней части экрана.

Настройка информации, отображаемой в списке, производится путём выбора интересующих пользователя столбцов и нажатия кнопки “show selected columns”.

Сортировка списка по значениям в определённом столбце производится путём нажатия на заголовок столбца.

4.2.1 Просмотр списка спектральных идентификаций для данного пептида

Если на странице со списком пептидных идентификаций нажать на последовательность пептида (столбец “sequence”), то отобразится страница со списком спектральных идентификаций, соответствующих данной пептидной последовательности (рис. 15).

Search results: [lev_comparison_hek_5/MSA_09_r01](#)

Search name	PSMs	Peptides	Proteins	FDR, %	FDR type	MS/MS	Unfiltered PSMs
MSA_09_r01	21118	9825	2524	1.0	-	71783	43815

Select columns for psm:

 sequence modified_sequence m/z exp m/z error in ppm missed cleavages e-value MPscore RT exp spectrum proteins
 proteins description Sin
[show selected columns](#)

sequence	m/z exp	missed cleavages	RT exp
NMGGPYGGGNYGPGSGSGGGYGGRR	1095.45655395	0	37.632367
NMGGPYGGGNYGPGSGSGGGYGGRR	1095.45656108	0	37.941831
NMGGPYGGGNYGPGSGSGGGYGGRR	730.640062133	0	37.944752
NMGGPYGGGNYGPGSGSGGGYGGRR	730.639981893	0	37.649923

Рис. 15. Сверху: выбор пептидной последовательности в списке пептидных идентификаций для просмотра списка спектральных идентификаций. Снизу: список спектральных идентификаций для выбранной последовательности.

4.2.2 Просмотр списка спектральных или пептидных идентификаций для данного белка

При просмотре списка белковых идентификаций в столбцах “PSMs” и “peptides” отображается количество спектральных и пептидных идентификаций для данного белка, соответственно. При нажатии на отображаемое число открывается список спектральных либо пептидных идентификаций.

4.2.3 Просмотр информации о белке в базе данных Uniprot

Если для поиска использовалась база данных на основе Uniprot, то при просмотре списка белковых идентификаций столбец “dbname” будет содержать ссылку на страницу с информацией о данном белке на сайте <http://www.uniprot.org>. Страница содержит протеомную и геномную информацию о структуре, функции и классификации белка.

4.3 Загрузка результатов на компьютер пользователя

На страницах визуализации результатов, а также просмотра списков идентификаций и сводных таблиц о результатах имеются кнопки “download csv files”, “download png files” и др. При нажатии на эти кнопки инициируется загрузка zip-архива с файлами на компьютер пользователя.

5 АВАРИЙНЫЕ СИТУАЦИИ

5.1 Отсутствие доступа к веб-интерфейсу ППМ IdentiPROT

В случае отсутствия доступа к веб-интерфейсу ППМ IdentiPROT необходимо убедиться в исправности сетевого подключения клиентского ПК и работоспособности установленного Интернет-обозревателя. В случае исправности подключения, следует обратиться к администратору сервера ППМ. В случае отсутствия исправного подключения или неработоспособности установленного Интернет-обозревателя, следует обратиться к сетевому администратору или администратору клиентского ПК.

5.2 Сообщения об ошибках (коды 5xx) при работе с веб-интерфейсом ППМ

При возникновении сообщений об ошибках с кодами 5xx следует обратиться к администратору сервера ППМ.

5.3 Рекомендации по освоению

Для освоения функциональности ППМ можно использовать любой файл в формате MGF или mzML. Его следует загрузить на сервер в соответствии с настоящей инструкцией. Простейший пример поиска включает использование стандартных параметров и стандартной базы данных (default.fasta), доступной автоматически.

6 ПОДДЕРЖКА И КОНТАКТНАЯ ИНФОРМАЦИЯ

Разработчик: Общество с ограниченной ответственностью «Куб»

Техническая поддержка: identiprot@gmail.com