




PARallele™ HLA software


РУКОВОДСТВО ПОЛЬЗОВАТЕЛЯ

Санкт-Петербург 2020

	ООО «ПАРСЕК ЛАБ»	PARAllele™ HLA software	
	РУКОВОДСТВО ПОЛЬЗОВАТЕЛЯ	Версия: 0.1	Стр. 2 из 16

Оглавление

ОГЛАВЛЕНИЕ	2
ВВЕДЕНИЕ.....	3
1. РЕКОМЕНДОВАННЫЕ ТЕХНИЧЕСКИЕ ХАРАКТЕРИСТИКИ.....	3
2. РОЛЕВАЯ СИСТЕМА	3
3. РАБОЧИЙ ПРОЦЕСС	5
4. ЖИЗНЕННЫЙ ЦИКЛ ОБРАЗЦА	6
5. АЛГОРИТМ HLA-ТИПИРОВАНИЯ	9
6. МЕТРИКИ КАЧЕСТВА ДАННЫХ.....	9
6.1. МЕТРИКИ КАЧЕСТВА ЗАПУСКА	9
6.2. МЕТРИКИ КАЧЕСТВА ОБРАЗЦА	9
6.3. МЕТРИКИ КАЧЕСТВА ЛОКУСА.....	10
7. ИНТЕРПРЕТАЦИЯ ГРАФИКОВ	11
8. КОНТРОЛЬ КАЧЕСТВА	12
9. РЕЗУЛЬТАТЫ ГЕНОТИПИРОВАНИЯ	12
10. ЭКСПОРТ РЕЗУЛЬТАТОВ	13
11. ИСПОЛЬЗОВАНИЕ КОНТРОЛЬНОГО ОБРАЗЦА	14
12. ДОСТУП К ПО.....	15
13. ТЕХНИЧЕСКАЯ ПОДДЕРЖКА.....	15
ГЛОССАРИЙ	15

	ООО «ПАРСЕК ЛАБ»	PARAllele™ HLA software	
	РУКОВОДСТВО ПОЛЬЗОВАТЕЛЯ	Версия: 0.1	Стр. 3 из 16

Введение

Программное обеспечение «PARAllele™ HLA software» (далее *ПО*) предназначено для анализа данных высокопроизводительного секвенирования, полученных с помощью набора реагентов для типирования генов HLA в высоком разрешении методом высокопроизводительного секвенирования (производства ООО «ПАРСЕК ЛАБ»), далее *набор реагентов*.

ПО обеспечивает следующие возможности:

- Импорт “сырых” прочтений и метаданных секвенирования
- Анализ загруженных прочтений различными алгоритмами с целью определения HLA-генотипа с разрешением 2 поля (2 fields, 4 digit) по локусам: HLA-A, HLA-B, HLA-C, HLA-DQB1 и HLA-DRB1
- Визуализацию и оценку качества результатов типирования
- Экспорт результатов типирования
- Встроенный канал взаимодействия со службой поддержки
- Разделение пользователей по ролям

Анализ и хранение данных осуществляются в “облачной” вычислительной инфраструктуре.


Потоковые исследования по HLA-генотипированию требуют высокой процессивности, надежности и воспроизводимости результатов, *ПО* назначает генотипы в полностью автоматическом режиме, ручное назначение генотипов пользователем *ПО* не предусмотрено. В случае невозможности автоматического назначения HLA-генотипа хотя бы по одному из локусов, образец автоматически направляется в службу поддержки.

1. Рекомендованные технические характеристики

- Скорость соединения не менее 50 Мбит/с
- Браузер Google Chrome версии 80 и выше
- Разрешение экрана – 1920x1080 пикселей

2. Ролевая система

Рабочий процесс предполагает разделение ролей пользователей на специалиста, выполняющего постановку в лаборатории – роль *Лаборант*, и руководителя,

	ООО «ПАРСЕК ЛАБ»	PARAllele™ HLA software	
	РУКОВОДСТВО ПОЛЬЗОВАТЕЛЯ	Версия: 0.1	Стр. 4 из 16

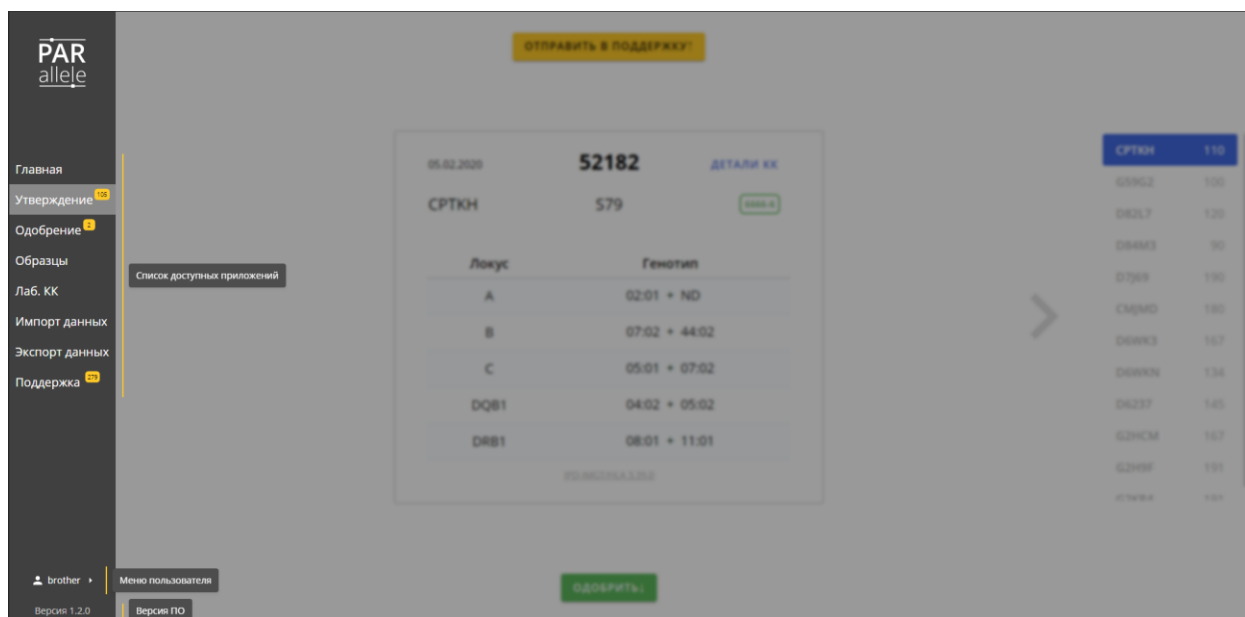
осуществляющего утверждение генотипов образцов – роль *Супервизор*. В ПО реализовано соответствующее разграничение прав доступа.

В соответствии с п. ЕЗ «Computer assisted analysis» руководства «Standards for Histocompatibility & Immunogenetics Testing», версия 8 (EFI, 2020) руководитель должен просмотреть, утвердить и подписать HLA-генотип, определенный автоматически с помощью ПО. Свое согласие с генотипом, установленным автоматически с помощью ПО, *Супервизор* выражает путем нажатия на кнопку «Утвердить».

На данный момент предусмотрены следующие роли:


- *Лаборант* - выполняет загрузку данных в приложение и оценивает качество запусков
- *Супервизор* - осуществляет утверждение генотипов образцов и экспорт результатов

Работа начинается с выбора приложения для работы. Список доступных приложений находится в навигационной панели.



Доступность приложений по пользовательским ролям:

Приложение	Описание	Лаборант	Супервизор
Лаб. КК	Оценка лабораторных метрик качества запусков	✓	✓
Импорт данных	Загрузка результатов секвенирования в ПО	✓	✓

	ООО «ПАРСЕК ЛАБ»	PARAllele™ HLA software	
	РУКОВОДСТВО ПОЛЬЗОВАТЕЛЯ	Версия: 0.1	Стр. 5 из 16

Приложение	Описание	Лаборант	Супервизор
Образцы	Просмотр статусов и результатов генотипирования всех образцов, загруженных в аккаунт, экспорт результатов для выбранных образцов	✓	✓
Главная	Схема анализа образца	х	✓
Утверждение	Утверждение результатов автоматического генотипирования образцов, которые прошли все проверки контроля качества	х	✓
Одобрение	Одобрение результатов ручного генотипирования образцов	х	✓
Экспорт данных	Экспорт результатов по запускам	х	✓

Пользователи работают в пределах своей организации и имеют доступ к данным, импортированным другими пользователями в пределах своей организации.

3. Рабочий процесс

Рабочий процесс состоит из согласованной работы пользователей обеих ролей. Первая часть, включающая загрузку данных осуществляется *Лаборантом*. Последующий анализ результатов проводится *Супервизором*.

4. Жизненный цикл образца




Жизненный цикл начинается с импорта данных, после чего запускается автоматический анализ данных. Каждый образец анализируется независимо от остальных образцов в запуске. Образец может иметь *терминальный* или *промежуточный* статус.



Два типа переходов между статусами:

Переход	Описание
Отмеченный на схеме пунктирной линией	Переход, не требующий активного участия пользователей (переход осуществляется автоматически, либо техническими специалистами производителя ПО)

	ООО «ПАРСЕК ЛАБ»	PARAllele™ HLA software	
	РУКОВОДСТВО ПОЛЬЗОВАТЕЛЯ	Версия: 0.1	Стр. 7 из 16

Переход	Описание
Отмеченный на схеме сплошной линией	Переход, требующий активных действий от пользователя (<i>Лаборанта и/или Супервизора</i>) или специалиста поддержки

В случае, если произошла техническая ошибка при анализе данных (например, загружены данные, полученные не с набором *реагентов*) образцу будет присвоен статус **ТЕХ. ОШИБКА**.

По результатам корректного завершения автоматического анализа образцу будет присвоен один из статусов: **ОЖИДАЕТ УТВЕРЖДЕНИЯ, В ПОДДЕРЖКЕ** и **ЛАБ. ОШИБКА**.

Статус **ЛАБ. ОШИБКА** присваивается образцам, не прошедшим контроль качества лабораторного этапа. Статус **ЛАБ. ОШИБКА** является терминальным статусом образца, установить генотип в этом случае невозможно, требуется повторное секвенирование образца.

Статус **ОЖИДАЕТ УТВЕРЖДЕНИЯ** присваивается образцам, прошедшим все этапы контроля качества. Эти образцы готовы для утверждения генотипа пользователем в роли *Супервизора*. Если результат типирования не вызывает сомнений, *Супервизор* утверждает генотип, после чего образец переходит в терминальный статус **УТВЕРЖДЕН**. Если генотип вызывает сомнения, *Супервизор* направляет образец в техническую поддержку, после чего образцу автоматически присваивается статус **В ПОДДЕРЖКЕ**.


В случае возникновения вопросов, связанных с анализом конкретного образца или локуса до утверждения генотипа *Супервизором*, рекомендуется отправить образец в поддержку непосредственно через *ПО*.

Статус **В ПОДДЕРЖКЕ** присваивается образцам, прошедшим контроль качества лабораторного этапа, но не прошедшим контроль качества на этапе анализа данных, либо образцам, направленным в поддержку *Супервизором*. После дополнительного анализа специалистами ООО «ПАРСЕК ЛАБ», образцы из данного статуса "переходят" в статус **ОЖИДАЕТ ОДОБРЕНИЯ**. Эти образцы готовы для подтверждения генотипа *Супервизором*. Если генотип образца не вызывает сомнений, *Супервизор* утверждает генотип, после чего образец переходит в терминальный статус **УТВЕРЖДЕН**. Если по результатам дополнительного анализа генотип установить не удалось, образец переходит в терминальный статус **ОШИБКА ТИПИРОВАНИЯ**.

После того, как все образцы в запуске достигают терминального статуса, результаты по запуску могут быть экспортированы.

Возможные статусы образцов приведены в таблице:

Статус образца	Описание	Тип
АНАЛИЗИРУЕТСЯ	Образец в процессе анализа	Промежуточный
В ПОДДЕРЖКЕ	Образец на рассмотрении специалистов поддержки	Промежуточный
ОЖИДАЕТ ОДОБРЕНИЯ	Образец рассмотрен специалистами поддержки и ожидает одобрения результата <i>Супервизором</i>	Промежуточный
ОЖИДАЕТ УТВЕРЖДЕНИЯ	Образец проанализирован и ожидает утверждения генотипа <i>Супервизором</i>	Промежуточный
ЛАБОРАТОРНАЯ ОШИБКА	Образец не прошел контроль качества лабораторных метрик	Терминальный
ОШИБКА ТИПИРОВАНИЯ	Генотип хотя бы по одному из локусов не может быть установлен	Терминальный
ТЕХНИЧЕСКАЯ ОШИБКА	Анализ образца не завершен корректно	Терминальный
УТВЕРЖДЕН	Генотип по всем анализируемым локусам установлен, результаты по образцу готовы к экспорту	Терминальный

	ООО «ПАРСЕК ЛАБ»	PARAllele™ HLA software	
	РУКОВОДСТВО ПОЛЬЗОВАТЕЛЯ	Версия: 0.1	Стр. 9 из 16

5. Алгоритм HLA-типирования

После импорта данных в ПО автоматически запускается алгоритм HLA-типирования **mercury-pipeline**, представляющий собой программный конвейер, принимающий на вход “сырые” прочтения и выдающий результаты HLA-типирования соответствующего разрешения результаты контроля качества. ПО поддерживает работу с различными версиями **mercury-pipeline**.

Информацию о версии алгоритма **mercury-pipeline**, используемой при анализе данных конкретного образца можно увидеть в карточке образца (бейдж Лот).

Обновление версии IPD-IMGT/HLA в ПО осуществляется каждые 6 месяцев.

6. Метрики качества данных

В ПО предусмотрено несколько типов метрик качества: по запуску, по образцу и по локусу. Все метрики качества данных рассчитываются и оцениваются в автоматическом режиме.

6.1. Метрики качества запуска

Среднее качество прочтений в образцах запуска

Среднее качество по запуску определяется по минимальному значению среднего качества образца в данном запуске.

Зависит от:

- качества реагентов для секвенирования
- соблюдения сроков и условий их хранения
- соблюдения протокола, в том числе при выборе параметров секвенирования


Среднее качество прочтений по запуску не является блокирующей метрикой при анализе данных, однако при значении ниже заданного порога рекомендуется повторить постановку.

6.2. Метрики качества образца

Общее количество прочтений

Основная метрика качества секвенирования образца. Общее количество прочтений складывается из:

- целевых прочтений - выровненных на анализируемые локусы HLA

	ООО «ПАРСЕК ЛАБ»	PARAllele™ HLA software	
	РУКОВОДСТВО ПОЛЬЗОВАТЕЛЯ	Версия: 0.1	Стр. 10 из 16

- гомологичных прочтений - выровненных на гомологи анализируемых локусов HLA
- нецелевых прочтений - выровненных на локусы, отличные от HLA
- триммированных по качеству прочтений (отсечение с начала и конца прочтений нуклеотидов с низким качеством)

Зависит от:

- соблюдения протокола на этапах измерения концентрации готовых библиотек, разведения и пулирования образцов.

Превышение рекомендованного значения не способствует повышению качества результата, но при этом негативно влияет на распределение прочтений по другим образцам.

Среднее качество прочтений

Среднее качество прочтений образца рассчитывается как среднее качество по всем прочитанным нуклеотидам. Качество прочтения отдельных нуклеотидов рассчитывается секвенатором и записывается в файл вместе с последовательностью прочтения. Отражает вероятность ошибки. В норме качество несколько снижается к концу прочтения.

Зависит от:

- качества реагентов для секвенирования
- соблюдения сроков и условий их хранения
- соблюдения протокола, в том числе при выборе параметров секвенирования

Размер вставки

Средняя длина секвенируемых фрагментов. Под вставкой понимается длина фрагмента без учета последовательностей адаптеров.

Зависит от:

- точности соблюдения протокола на этапе фрагментации и при очистках магнитными частицами

6.3. Метрики качества локуса

Количество прочтений

Число прочтений, выровненных на локус. Учитываются только те, прочтения в которых оба прочтения из пары выровнялись на данный локус. Недостаточное покрытие по одному из анализируемых локусов может свидетельствовать о выпадении аллелей (drop out, потеря гетерозиготности), т.е. наличии в образце аллелей, которые не могут быть амплифицированы с помощью набора реагентов.

Контаминация

Контаминация исследуемых образцов посторонней геномной ДНК человека или продуктами её амплификации, идентифицируется автоматически как высокая степень неоднозначности полученных результатов генотипирования.

При обнаружении контаминации рекомендуется повторно забрать биоматериал, выделить ДНК и секвенировать образец с помощью *набора реагентов*.

Качество генотипирования

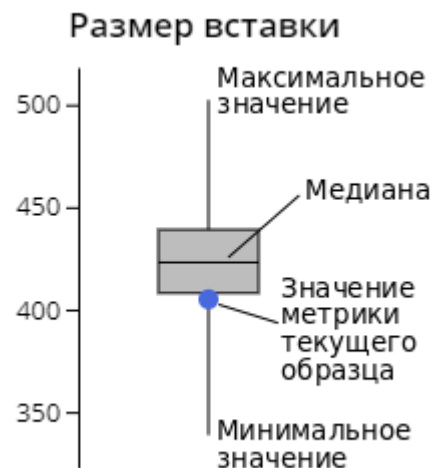
После определения наиболее вероятного генотипа оценивается качество генотипирования путем оценки степени неоднозначности по генотипам и по аллелям.


Неоднозначные генотипы и неоднозначные аллели, входящие в список известных неразрешимых неоднозначностей, **не являются некорректным результатом** и подлежат выдаче наряду с однозначными генотипами и аллелями.

7. Интерпретация графиков

Ящик с усами (боксплот)

На примере графика размера вставки



	ООО «ПАРСЕК ЛАБ»	PARAllele™ HLA software	
	РУКОВОДСТВО ПОЛЬЗОВАТЕЛЯ	Версия: 0.1	Стр. 12 из 16

8. Контроль качества

Контроль качества состоит в последовательной оценке метрик качества образца и классификации на их основании образца в один из трех возможных статусов: **ОЖИДАЕТ УТВЕРЖДЕНИЯ, В ПОДДЕРЖКЕ и ЛАБ. ОШИБКА**. Пороговые значения определяются версией алгоритма анализа. По каждой метрике существуют зоны с оптимальными и недостаточными значениями, некоторые метрики также могут иметь субоптимальную зону.

В случае, если в запуске присутствует хотя бы один образец среднее качество прочтений которого ниже порогового, весь запуск признается не прошедшим контроль качества.

9. Результаты генотипирования

Метаданные по образцу, результаты HLA-генотипирования и детали контроля качества (КК) представлены на карточке образца, доступной через приложения: «Утверждение», «Одобрение» и «Образцы».

Метаданные по образцу включают следующую информацию:


- имя образца
- дату секвенирования
- номер запуска
- номер образца в запуске
- лот
- релиз базы данных IPD-IMGT/HLA

Результаты HLA-генотипирования представлены в таблице, каждая строка которой содержит название локуса и генотип в формате: **аллель 1 + аллель 2**, где аллель 1 и аллель 2 – пара гомологичных аллелей.

Неоднозначные генотипы записываются друг под другом в одной строке соответствующего локуса. Неоднозначные аллели разделены символом «/».

В соответствии с требованием п. F1.2.3 «Reporting Homozygosity and Heterozygosity» руководства «Standards for Histocompatibility & Immunogenetics Testing», версия 8 (EFI, 2020), гомозиготный генотип записывается как **аллель 1 + ND**, где ND (not determined) – аллель не определена.

Редкие аллели (аллели с частотой по образцам в аккаунте организации менее 5%) подчеркнуты желтой чертой, при наведении на аллель отображается соответствующий бейдж.

	ООО «ПАРСЕК ЛАБ»	PARAllele™ HLA software	
	РУКОВОДСТВО ПОЛЬЗОВАТЕЛЯ	Версия: 0.1	Стр. 13 из 16

В приложении «Одобрение» результаты генотипирования содержат «Старый генотип» (определенный автоматически в рамках ЛО) и «Новый генотип» (генотип, установленный специалистами технической поддержки производителя), в приложении «Утверждение» содержится только генотип, определенный автоматически.

Детали контроля качества (Детали КК) содержат следующую информацию:

- общее количество прочтений по образцу
- распределение прочтений по типам (триммированные по качеству, нецелевые, гомологичные и целевые)
- размер вставки по образцу (п.н.) и диаграмму размаха по запуску
- среднее качество по образцу (phred) и диаграмму размаха по запуску
- количество прочтений по локусам и соответствующие диаграммы размаха по запуску

Также в разделе **Детали контроля качества** представлены результаты контроля качества по образцу, а именно общие проблемы (субоптимальные значения покрытия по образцу) и локус-специфичные проблемы (субоптимальные значения покрытия по локусу, неопределенный генотип по локусу).

10. Экспорт результатов

Экспорт результатов доступен через приложения «Образцы» и «Экспорт данных». В приложении «Образцы» осуществляется экспорт результатов по образцам, выбранным с помощью фильтров. В приложении «Экспорт данных» осуществляется экспорт результатов по всем образцам запуска, после утверждения и одобрения всех образцов прошедших контроль качества (финализация работы с запуском).

Экспорт осуществляется в текстовый файл формата CSV. Экспортируется следующая информация по образцу:

- статус образца (Resolution)
- имя образца (Sample name)
- лот реагентов (Reagent kit lot)
- номер запуска (Run Id)
- номер образца в запуске (Sample id in run)
- имя пользователя, загрузившего запуск (Uploaded by)
- дата анализа данных (Analysis date)
- версия алгоритма анализа данных (Pipeline)
- релиз IPD-IMGT/HLA (IMGT)

- имя *Супервизора* (Approved by)
- дата утверждения/одобрения (Approved at)
- генотип по локусу HLA-A
- генотип по локусу HLA-B
- генотип по локусу HLA-C
- генотип по локусу HLA-DQB1
- генотип по локусу HLA-DRB1

Генотипы по локусам указаны в формате **GL string**, без префиксов «HLA-A», «HLA-B», «HLA-C», «HLA-DQB1», «HLA-DRB1».


11. Использование контрольного образца

В состав *набора реагентов* входит контрольный образец с известным HLA-генотипом, установленным референсным методом. Рекомендуется включать контрольный образец в каждую постановку.

В случае если контрольный образец указан при импорте данных, программа автоматически проанализирует результат, полученный по контрольному образцу, и сравнит его с референсным генотипом. В случае, если контрольный образец не проходит контроль качества или по генотипу выявлено неполное совпадение, программа автоматически маркирует запуск как непрошедший контроль качества (в этом случае рекомендуется повторить постановку целиком).

Статус прохождения контроля качества контрольным образцом (отображаемый на бейдже Лот) может быть одним из:

Статус	Описание
НЕИЗВЕСТНЫЙ	Контрольный образец не был поставлен в запуске (бейдж Лот серый)
ПРОЙДЕН	Контрольный образец прошел контроль качества и генотипы полностью совпали с референсными значениями (бейдж Лот зеленый)
НЕ ПРОЙДЕН	Контрольный образец не прошел контроль качества или генотипы не полностью совпали с референсными значениями (бейдж Лот красный)

	ООО «ПАРСЕК ЛАБ»	PARAllele™ HLA software	
	РУКОВОДСТВО ПОЛЬЗОВАТЕЛЯ	Версия: 0.1	Стр. 15 из 16

12. Доступ к ПО


Предоставляется исключительно для покупателей набора реагентов для типирования генов HLA в высоком разрешении методом высокопроизводительного секвенирования (производства ООО «ПАРСЕК ЛАБ»).

13. Техническая поддержка

В случае возникновения общих вопросов и вопросов, связанных с работой ПО следует обратиться в службу технической поддержки производителя через сайт: <https://support.parseq.pro/>.

Глоссарий

- **HLA** (Human Leukocyte Antigens): человеческие лейкоцитарные антигены, представляющие собой поверхностные рецепторы иммунных клеток, кодируемые группой тесно сцепленных генов на 6-й хромосоме. Антигены HLA играют важнейшую роль в регуляции иммунного ответа на чужеродные антигены, определяют уникальность тканевого типа человека и ответственны за отторжение чужеродных тканей при трансплантации.
- **IPD-IMGT/HLA**: специализированная база данных последовательностей HLA номенклатурного комитета по HLA-системе Всемирной Организации Здравоохранения, представленная на сайте <http://www.ebi.ac.uk/imgt/hla>.
- **Аллель**: вариант последовательности ДНК в локусе.
- **Аллельное разрешение HLA**: это результат HLA-типирования, соответствующий одной аллели, согласно текущей версии WHO HLA Nomenclature Report. Каждая аллель в зависимости от разрешения может иметь 2-, 4-, 6- или 8-значный идентификатор, определяемый ее нуклеотидной последовательностью. Первое поле описывает тип, соответствующий серологическому антигену. Второе поле используются для обозначения специфического подтипа белка. Аллели, номера которых отличаются по второму полю, должны отличаться по одной или нескольким нуклеотидным заменам, изменяющим аминокислотную последовательность кодируемого белка. Аллели, которые отличаются только синонимичными нуклеотидными заменами в кодирующей последовательности, описываются с использованием третьего поля. Аллели, которые отличаются только вариантами последовательностей в интронах или нетранслируемых областях, описываются с использованием четвертого поля.
- **Генотип**: комбинация аллелей гена.

	ООО «ПАРСЕК ЛАБ»	PARAllele™ HLA software	
	РУКОВОДСТВО ПОЛЬЗОВАТЕЛЯ	Версия: 0.1	Стр. 16 из 16

- **Аналитическая чувствительность:** доля совпадающих аллелей между референсными генотипами и генотипами, установленными с помощью *набора реагентов*.
- **Аналитическая специфичность:** доля неоднозначных аллелей от общего числа проанализированных аллелей с помощью *набора реагентов*.
- **Контаминированный образец:** аналит, представляющий собой смесь ДНК неродственных индивидуумов.
- **Локус:** местоположение искомого региона генома на генетической карте хромосомы.