

Aparat MiSeq™Dx

DO STOSOWANIA W DIAGNOSTYCE IN VITRO
WYŁĄCZNIE NA EKSPORT

Przeznaczenie

Aparat MiSeqDx jest przeznaczony do ukierunkowanego sekwencjonowania bibliotek DNA z ludzkiego genomowego DNA wyodrębnianego z obwodowej krwi pełnej lub z tkanki utrwalonej w formalinie i zatopionej w parafinie (FFPE, formalin-fixed, paraffin-embedded) z wykorzystaniem oznaczeń diagnostycznych *in vitro* (IVD) przeznaczonych do wykonywania na tym aparacie. Aparat MiSeqDx nie jest przeznaczony do sekwencjonowania genomu pełnego ani sekwencjonowania *de novo*. Aparat MiSeqDx jest przeznaczony do stosowania z zarejestrowanymi i uwzględnionymi w wykazach, dozwolonymi lub zatwierdzonymi do stosowania w IVD odczynnikami i oprogramowaniem analitycznym.

Zasady dotyczące procedury

Aparat Illumina MiSeqDx jest przeznaczony do celowanego resekwencjonowania ludzkiego DNA z użyciem materiałów eksploatacyjnych do sekwencjonowania firmy Illumina i bibliotek ludzkiego genomowego DNA wyodrębnionego z obwodowej krwi pełnej lub tkanek FFPE przy użyciu odczynników wymienionych w wykazach, dozwolonych lub zatwierdzonych do stosowania w diagnostyce *in vitro*. Biblioteki przygotowuje się poprzez amplifikację sekwencji docelowych oraz dodanie indeksów próbek i sekwencji przechwytywania. Biblioteki próbek są przechwytywane w komorze przepływowej i sekwencjonowane w aparacie przy użyciu technologii sekwencjonowania przez syntezę SBS (ang. Sequencing By Synthesis). W technologii SBS stosowana jest metoda terminatora odwracalnego, służąca do wykrywania nukleotydów w miarę ich wiązania w rosnących łańcuchach DNA. Oprogramowanie do analizy w czasie rzeczywistym (RTA) wykonuje analizę obrazów i rozpoznawanie nukleotydów oraz w każdym cyklu sekwencjonowania przypisuje wynik jakościowy do każdego nukleotydu. Po zakończeniu analizy podstawowej w aparacie MiSeqDx odbywa się przetwarzanie rozpoznanych nukleotydów w analizie wtórnej.

Proces przetwarzania obejmuje demultipleksowanie, tworzenie plików FASTQ, dopasowywanie, rozpoznawanie wariantów oraz tworzenie plików rozpoznawania wariantów (VCF, Variant Call Format), które zawierają informacje o wariantach obecnych na określonych pozycjach w genomie referencyjnym. Wtórna analiza w aparacie MiSeqDx wykonywana jest przy użyciu różnych modułów, w zależności od procedury.

Konfiguracja z obsługą dwóch systemów operacyjnych

Konfiguracja z obsługą dwóch systemów operacyjnych obejmuje osprzęt, oprogramowanie i procedury instalacyjne umożliwiające aparatowi MiSeqDx wykonywanie zarówno przebiegów sekwencjonowania diagnostycznego *in vitro* (IVD), jak i sekwencjonowania wyłącznie do celów badawczych (RUO, ang. Research Use Only). Konfiguracja z obsługą dwóch systemów operacyjnych umożliwia użytkownikowi przełączanie aparatu między trybem diagnostycznym a trybem badawczym. Znaczniki do identyfikacji radiowej (RFID, radio-frequency identification) na materiałach eksploatacyjnych wykorzystywanych w sekwencjonowaniu zapobiegają użyciu odczynników do sekwencjonowania RUO w przebiegach sekwencjonowania diagnostycznego.

Ograniczenia dotyczące procedury

- Do celów diagnostyki *in vitro*.
- Wyniki przedstawione w ulotce uzyskano przy użyciu paneli reprezentatywnych oznaczeń z wykorzystaniem obwodowej krwi pełnej lub linii komórkowych w przypadku oznaczania wariantów linii zarodkowych oraz tkanki FFPE lub linii komórkowych FFPE w przypadku oznaczania wariantów somatycznych przy użyciu opisanych odczynników i modułów oprogramowania. Moduł analizy wariantów linii zarodkowej oraz moduł analizy wariantów somatycznych opracowano do oceny skuteczności oznaczeń reprezentatywnych. Cechy wydajności podano wyłącznie w celach informacyjnych. Przedstawione badania walidacyjne służą jedynie jako przykładowa ilustracja ogólnych możliwości aparatu i nie stanowią o możliwości użycia lub odpowiedności aparatu w odniesieniu do jakichkolwiek szczególnych zastosowań. Wszystkie testy diagnostyczne opracowane przy użyciu aparatu wymagają pełnej walidacji we wszystkich aspektach działania.
- Produkt ma następujące ograniczenia:
 - Dane wyjściowe sekwencjonowania ≥ 5 Gb przy długości odczytu 2×150 bp.
 - Odczyty przechodzące przez filtr ≥ 15 milionów przy długości odczytu 2×150 bp.
 - Nukleotydy o jakości $Q30 \geq 80\%$ przy długości odczytu 2×150 pb.
80% lub więcej nukleotydów ma wynik jakościowy w skali Phred powyżej 30, co oznacza rozpoznawanie nukleotydów z dokładnością powyżej 99,9%.
- Aparat MiSeqDx został zwalidowany jedynie do celów sekwencjonowania bibliotek ludzkiego DNA wyodrębnionego z obwodowej krwi pełnej lub tkanek FFPE. W oznaczeniach diagnostycznych *in vitro* z użyciem aparatu nie należy stosować bibliotek wygenerowanych z innych rodzajów próbek. Nie zbadano przydatności urządzenia w sekwencjonowaniu kwasów nukleinowych bakterii lub wirusów w preparatach klinicznych.
- Aparat MiSeqDx jest przeznaczony do zastosowań diagnostycznych *in vitro* z zarejestrowanymi i uwzględnionymi w wykazach, dozwolonymi lub zatwierdzonymi do stosowania w IVD odczynnikami lub oznaczeniami. Ograniczenia w zakresie odczynników i cechy wydajności opisane

w niniejszej ulotce do opakowania są oparte na oznaczeniach reprezentatywnych i modułach oprogramowania. Informacje na temat przeznaczenia oznaczeń IVD, wykrywanych przez nie wariantów i rodzajów próbek można znaleźć w ulotce do opakowania właściwej dla danego oznaczenia.

- Oprogramowanie oznaczenia nie dopasowuje polimorfizmów typu indel (insercje, delecje i ich kombinacje) dłuższych niż 25 bp. W rezultacie oprogramowanie do oznaczeń nie wykrywa polimorfizmów typu indel o długości większej niż 25 bp.
- System został zwalidowany do wykrywania wariantów pojedynczych nukleotydów (SNV) i delecji o długości do 25 bp oraz insercji o długości 24 bp w przypadku stosowania z oprogramowaniem do wykrywania wariantów linii zarodkowej lub wariantów somatycznych. W przypadku rozpoznawania wariantów somatycznych przy częstości wariantów wynoszącej 0,05 wykryto delecje o długości 25 bp oraz insercje o długości 18 bp.
- Oprogramowanie do oznaczeń może nie dopasowywać odczytów ampliconów z ekstremalną zawartością wariantów, co prowadzi do zaraportowania regionu jako typu dzikiego. Taka ekstremalna zawartość obejmuje:
 - Odczyty zawierające ponad trzy polimorfizmy typu indel.
 - Odczyty o długości co najmniej 30 bp i zawartości SNV powyżej 4% całkowitej długości docelowego ampliconu (z wykluczeniem regionów sondy).
 - Odczyty o długości poniżej 30 bp i zawartości SNV ponad 10% całkowitej długości ampliconu (z uwzględnieniem regionów sondy).
- Duże warianty, w tym warianty wielonukleotydowe (MNV) i duże polimorfizmy typu indel, mogą zostać zgłoszone jako odrębne mniejsze warianty w wyjściowym pliku VCF.
- Warianty delecji mogą być filtrowane lub pominięte w przypadku rozciągnięcia na dwa sąsiednie amplicony, jeśli długość delecji jest większa lub równa nałożeniu pomiędzy sąsiednimi ampliconami.
- System nie jest w stanie wykryć polimorfizmów typu indel, jeśli występują bezpośrednio obok startera i nie ma nakładającego się ampliconu. W przypadku regionów z nakładającymi się ampliconami oznaczenie nie pozwala wykryć delecji, kiedy region nałożenia jest mniejszy niż rozmiar wykrywanej delecji. Na przykład, jeśli region nałożenia pomiędzy dwoma sąsiednimi ampliconami obejmuje dwa nukleotydy, oznaczenie nie pozwala wykryć żadnych delecji obejmujących oba te nukleotydy. Delecja pojedynczego nukleotydu na dowolnym z tych nukleotydów może zostać wykryta.
- Podobnie jak w przypadku procedur przygotowania bibliotek opartych na hybrydyzacji odpowiednie polimorfizmy, mutacje, insercje lub delecje w regionach wiązań oligonukleotydów mogą mieć wpływ na badane allele. W rezultacie wpływ jest też wywierany na rozpoznania nukleotydów podczas sekwencjonowania. Na przykład:
 - Wariant w fazie z wariantem w regionie startera może nie ulec amplifikacji, dając wynik fałszywie ujemny.
 - Warianty w regionie startera mogą uniemożliwić amplifikację allelu referencyjnego, prowadząc do nieprawidłowego wykrycia wariantu homozygotycznego.

- Warianty polimorfizmów typu indel w regionie startera mogą doprowadzić do fałszywie dodatniego rozpoznania pod koniec odczytu w sąsiedztwie startera.
- Polimorfizmy typu indel mogą zostać odfiltrowane ze względu na obciążenie systematyczne nici, jeśli występują w pobliżu końca odczytu i są poddawane procesowi soft-clippingu podczas dopasowania.
- Nie walidowano pod kątem małych polimorfizmów typu MNV.
- Nie walidowano pod kątem wariantów liczby kopii lub wariantów strukturalnych, takich jak fuzje czy translokacje.
- Ograniczenia swoiste dla linii zarodkowej:
 - Aparat MiSeqDx wykorzystujący moduł do wykrywania wariantów linii zarodkowej jest przeznaczony do dostarczania wyników jakościowych na potrzeby rozpoznawania wariantów linii zarodkowej (czyli homozygotycznych, heterozygotycznych, typu dzikiego).
 - W przypadku stosowania z modułem do wykrywania wariantów linii zarodkowej, minimalne pokrycie na amplikon niezbędne do dokładnego rozpoznania wariantu wynosi 150x. Liczba próbek i łączna liczba celowanych nukleotydów wpływają na pokrycie. Zawartość GC i inna zawartość genomowa mogą wpływać na pokrycie.
 - Zmienność liczby kopii może wpływać na to, czy wariant zostanie rozpoznany jako homo- czy heterozygotyczny.
 - Warianty w pewnym powtarzalnym kontekście są wyfiltrowywane w plikach VCF. Filtr powtórzeń RMxN jest stosowany do filtrowania wariantów, jeśli całość lub część sekwencji wariantu występuje powtarzalnie w genomie referencyjnym sąsiadującym z pozycją wariantu. W przypadku rozpoznawania wariantów linii zarodkowej do wyfiltrowania wariantu potrzeba co najmniej 9 powtórzeń w regionie referencyjnym, przy czym uwzględniane są tylko powtórzenia o długości do 5 bp (R5x9).
- Ograniczenia swoiste dla wariantów somatycznych:
 - System MiSeqDx wykorzystujący moduł oznaczania wariantów somatycznych jest przeznaczony do dostarczania wyników jakościowych na potrzeby rozpoznawania wariantów somatycznych (np. obecność wariantu somatycznego o częstości wariantu większej lub równej 0,026 przy granicy wykrywalności wynoszącej 0,05).
 - W przypadku stosowania z modułem do wykrywania wariantów somatycznych minimalne pokrycie na amplikon niezbędne do dokładnego rozpoznania wariantu wynosi 450x na pulę oligonukleotydów. Liczba próbek i łączna liczba celowanych nukleotydów wpływają na pokrycie. Zawartość GC i inna zawartość genomowa mogą wpływać na pokrycie.
 - Warianty w pewnym powtarzalnym kontekście są wyfiltrowywane w plikach VCF. Filtr powtórzeń RMxN jest stosowany do filtrowania wariantów, jeśli całość lub część sekwencji wariantu występuje powtarzalnie w genomie referencyjnym sąsiadującym z pozycją wariantu. W przypadku rozpoznawania wariantów somatycznych potrzeba co najmniej 6 powtórzeń w regionie referencyjnym do wyfiltrowania wariantu, przy czym uwzględniane są tylko powtórzenia o długości do 3 bp (R3x6).

- Moduł do wykrywania wariantów somatycznych nie jest w stanie odróżnić wariantów linii zarodkowej od wariantów somatycznych. Moduł jest przeznaczony do wykrywania wariantów o różnych częstościach wariantu, jednak częstości wariantu nie można użyć do odróżnienia wariantów somatycznych od wariantów linii zarodkowej.
- Obecność prawidłowej tkanki w próbce wpływa na wykrywanie wariantów. Zgłoszona granica wykrywalności jest oparta na częstości wariantu względem całkowitego DNA wyekstrahowanego z tkanki nowotworowej i prawidłowej.

Elementy produktu

Aparat Illumina MiSeqDx składa się z następujących elementów:

Aparat MiSeqDx (nr kat.: DX-410-1001)

Do działania aparatu MiSeqDx i analizy danych wymagane są następujące składniki oprogramowania:

Aplikacja	Funkcja	Opis
Oprogramowanie MiSeq Operating Software (MOS)	Sterowanie działaniem aparatu	Aplikacja MOS zarządza działaniem aparatu podczas sekwencjonowania oraz generuje obrazy używane w oprogramowaniu do analizy w czasie rzeczywistym (RTA, Real-Time Analysis). Więcej informacji zawiera dokument <i>MiSeqDx Instrument Reference Guide for MOS v4 (Instrukcja obsługi aparatu MiSeqDx z oprogramowaniem MOS w wer. 4) (nr dokumentu: 200010452)</i> .
Real-Time Analysis (RTA)	Wykonywanie analizy podstawowej	Aplikacja RTA przetwarza obrazy wygenerowane przez oprogramowanie MOS dla poszczególnych płytek w cyklu sekwencjonowania na pliki rozpoznania nukleotydów, stanowiących dane wejściowe modułów analitycznych oprogramowania Local Run Manager. Aplikacja RTA nie udostępnia interfejsu użytkownika.
Lokalny menedżer przebiegu	Interfejs wyboru modułu	Lokalny menedżer przebiegu stanowi zintegrowane z aparatem rozwiązanie do zarządzania użytkownikami, wykonywania wtórnych analiz i monitorowania statusu. Więcej informacji zawiera dokument <i>Local Run Manager v3 Software Reference Guide for MiSeqDx (Instrukcja obsługi lokalnego menedżera przebiegu w wer. 3 na aparacie MiSeqDx) (nr dokumentu: 200003931)</i> .

Przechowywanie i sposób postępowania

Element	Specyfikacja
Temperatura	Transport i przechowywanie: od -10°C do 40°C (od 14°F do 104°F) Warunki robocze: od 19°C do 25°C (od 66°F do 77°F)
Wilgotność	Transport i przechowywanie: wilgotność bez kondensacji Warunki pracy: wilgotność względna 30–75% (bez kondensacji)

Sprzęt i materiały wymagane, ale niedostarczane

Materiały eksploatacyjne do sekwencjonowania

Zestaw odczynników MiSeqDx Reagent Kit v3 (nr kat.: 20037124)

Zestaw odczynników MiSeqDx Reagent Kit v3 Micro (nr kat.: 20063860)

Materiały eksploatacyjne dostarczane przez użytkownika

Przed rozpoczęciem przebiegu należy się upewnić, że materiały eksploatacyjne dostarczane przez użytkownika są dostępne.

Materiał eksploatacyjny	Cel
Chusteczki nasączone 70% alkoholem izopropylowym lub 70% etanolem	Czyszczenie platformy komory przepływowej i szkła komory
Chusteczki laboratoryjne, niestrzępiące się	Czyszczenie platformy komory przepływowej
Papier do czyszczenia soczewek, 4 x 6 cali	Czyszczenie komory przepływowej
Tween 20	Mycie aparatu

Materiał eksploatacyjny	Cel
Plastikowe szczypce z kwadratową końcówką (opcjonalnie)	Wymywanie komory przepływowej z pojemnika transportowego
Woda, jakość laboratoryjna	Mycie aparatu

Wytyczne dotyczące wody laboratoryjnej

Do przeprowadzania procedur w aparacie zawsze należy używać wody laboratoryjnej lub wody dejonizowanej. Nie wolno używać wody wodociągowej.

Należy używać wyłącznie następujących rodzajów wody lub jej odpowiedników:

- Woda dejonizowana
- Illumina PW1
- Woda o rezystywności 18 megaomów (MΩ)
- Woda Milli-Q
- Woda Super-Q
- Woda do biologii molekularnej

Ostrzeżenia i środki ostrożności



OSTRZEŻENIE

Ten zestaw odczynników zawiera potencjalnie niebezpieczne substancje chemiczne. **Wdychanie, połknięcie, kontakt ze skórą i oczami mogą powodować uszczerbek na zdrowiu. Należy nosić wyposażenie ochronne, w tym ochronę oczu, rękawiczki oraz fartuch laboratoryjny odpowiednie do ryzyka narażenia. Zużyte odczynniki należy traktować jako odpady chemiczne i utylizować je zgodnie z odpowiednimi przepisami regionalnymi, krajowymi i lokalnymi.** Dodatkowe informacje dotyczące ochrony środowiska, zdrowia i bezpieczeństwa zawiera karta charakterystyki (SDS) dostępna na stronie support.illumina.com/sds.html.

- Z wszystkimi próbkami krwi należy postępować tak, jak gdyby znana była ich zakaźność w zakresie ludzkiego wirusa niedoboru odporności (HIV), ludzkiego wirusa zapalenia wątroby typu B (HBV) oraz innych patogenów przenoszonych z krwią (uniwersalne środki ostrożności).
- Nieprzestrzeganie przedstawionych procedur może powodować błędy w wynikach lub znaczne obniżenie jakości próbek.

- Należy przestrzegać środków ostrożności dotyczących rutynowych badań laboratoryjnych. Nie należy pipetować ustami. Nie należy spożywać żywności i napojów ani palić tytoniu na wyznaczonych stanowiskach roboczych. Podczas posługiwania się próbkami i odczynnikami zestawu należy nosić jednorazowe rękawice i fartuchy laboratoryjne. Po zakończeniu posługiwania się próbkami i odczynnikami zestawu należy dokładnie umyć ręce.
- Przestrzeganie zasad wykonywania prac oraz higieny prac laboratoryjnych jest wymagane, aby chronić produkty reakcji PCR przed zanieczyszczeniem przez odczynniki, przyrządy lub próbki genomowego DNA. Zanieczyszczenie produktów reakcji PCR może powodować niedokładne i niepewne wyniki.
- Aby zapobiegać zanieczyszczeniu produktów, należy upewnić się, że obszary procesu przed amplifikacją i po niej są wyposażone w specjalne przyrządy i materiały eksploatacyjne (np. pipety, końcówki do pipetowania, bloki grzejne, mieszadła wirowe i wirówki).
- W odpowiednich przypadkach parowanie indeksów i próbek powinno przebiegać w ścisłej zgodności z układem wydrukowanej płytki. Moduł lokalnego menedżera przebiegu automatycznie podstawia startery indeksowe związane z nazwami próbek, jeśli zostały wprowadzone do modułu. Przed rozpoczęciem przebiegu sekwencjonowania użytkownikowi zaleca się zweryfikowanie powiązań starterów indeksowych z próbkami. Niezgodności między rozkładem próbek a ich przypisaniem na płytce mogą powodować dodatkowe rozpoznanie próbki i nieprawidłowe dane w raporcie wyników.
- Zdecydowanie zaleca się instalację dostarczonego przez użytkownika oprogramowania chroniącego komputer przed wirusami. W celu uzyskania instrukcji instalacji należy zapoznać się z podręcznikiem użytkownika.
- Aparatu MiSeqDx nie należy użytkować ze zdjętymi panelami. Obsługa aparatu ze zdjętymi panelami stwarza potencjalne ryzyko ekspozycji użytkownika na działanie napięcia sieciowego oraz napięć prądu stałego.
- Nie należy dotykać stolika w przedziale komory przepływowej. Znajdująca się w tym przedziale grzałka działa w zakresie temperatur od 22°C do 95°C i może spowodować oparzenia.
- Aparat waży około 57 kg (126 lb) i w przypadku upuszczenia lub nieprawidłowego obchodzenia się z nim może powodować poważne obrażenia.
- Wszelkie poważne incydenty związane z tym produktem należy niezwłocznie zgłaszać do firmy Illumina i właściwego organu państwa członkowskiego, w których przebywa użytkownik i/lub pacjent.

Instrukcja użytkowania

Poniższe instrukcje użytkowania aparatu MiSeqDx wymagają zastosowania odczynników dostarczonych w zestawie odczynników MiSeqDx Reagent Kit v3.

Tworzenie przebiegu za pomocą lokalnego menedżera przebiegu

Szczegółowe instrukcje na temat tworzenia przebiegu można znaleźć w dokumencie *Local Run Manager v3 Software Reference Guide for MiSeqDx (Instrukcja obsługi lokalnego menedżera przebiegu w wer. 3 na aparacie MiSeqDx)* (nr dokumentu 200003931) i instrukcji dotyczącej używanego do analizy modułu lokalnego menedżera przebiegu.

Przygotowywanie kasety odczynników

W poniższych instrukcjach opisano, jak rozmrozić odczynniki za pomocą kąpieli wodnej o temperaturze pokojowej.

1. Wyjąć kasetę odczynników z chłodni o temperaturze od -25°C do -15°C .
2. Umieścić kasetę odczynników w kąpieli z wodą dejonizowaną o temperaturze pokojowej, zanurzając dolną część kasety do nadrukowanej na niej linii poziomu wody. Należy uważać, aby nie przekroczyć linii maksymalnego poziomu wody.

Rysunek 1 Maksymalny poziom wody



3. Pozostawić kasetę odczynników do rozmrożenia w kąpieli wodnej o temperaturze pokojowej na około 60–90 minut lub do momentu całkowitego rozmrożenia.
4. Wyjąć kasetę z kąpieli wodnej i delikatnie postukać nią o blat w celu usunięcia wody z dolnej części kasety. Osuszyć dolną część kasety. Upewnić się, że wierzch kasety odczynników nie jest zachlapany wodą.

Sprawdzanie stanu kasety odczynników

1. Odwrócić dziesięciokrotnie kasetę odczynników w celu wymieszania rozmrożonych odczynników, a następnie sprawdzić, czy wszystkie pozycje zostały rozmrożone.
2. Sprawdzić odczynniki w pozycjach 1, 2 i 4, aby upewnić się, że zostały dokładnie wymieszane i nie zawierają osadów.

UWAGA W celu zapewnienia prawidłowego sekwencjonowania ważne jest, aby odczynniki w kasecie były dokładnie rozmrożone i wymieszane.

3. Delikatnie postukać kasetą o blat w celu zmniejszenia liczby pęcherzyków powietrza w odczynnikach.

UWAGA W dolnej części każdego zbiorniczka znajdują się rurki dozownika aparatu MiSeqDx służące do aspiracji odczynników, ważne jest więc, aby w zbiorniczkach nie było pęcherzyków powietrza.

4. Do czasu konfiguracji przebiegu umieścić kasetę odczynników w lodzie lub wstawić do lodówki o temperaturze od 2°C do 8°C (na maksymalnie sześć godzin). W celu uzyskania najlepszych rezultatów należy umieścić próbkę w aparacie i skonfigurować przebieg.

Przygotowanie próbek do sekwencjonowania

Wskazówki dotyczące przygotowywania bibliotek próbek do sekwencjonowania, w tym przygotowywania rozcieńczeń i pul bibliotek, zawarto w ulotce do opakowania zestawu do przygotowywania bibliotek w sekcji „Instrukcja użytkownika”.

Rozcieńczenie bibliotek próbek zależy od złożoności pul oligonukleotydów. W aparacie MiSeqDx wymagana jest optymalizacja gęstości klastrów; optymalna gęstość może się różnić w zależności od konkretnego zestawu do przygotowania bibliotek.

Ładowanie bibliotek próbek do kasy

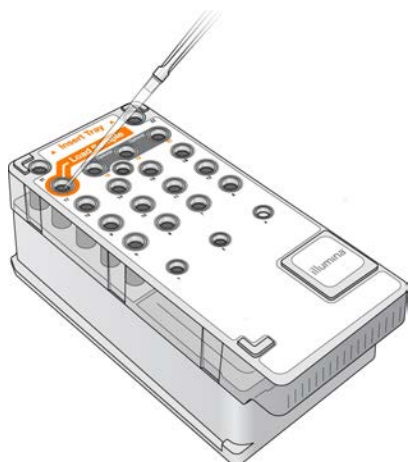
Próbki można załadować do kasy odczynników, gdy jest ona gotowa do użycia po całkowitym rozmrożeniu.

1. Za pomocą osobnej, czystej i pustej 1 ml końcówki do pipety przedziurawić folię zabezpieczającą zbiorniczek kasy odczynników oznaczony **Load Samples** (Ładowanie próbek).

UWAGA Nie nakłuwać innych miejsc na kasecie odczynników. Inne miejsca na kasecie odczynników są przekłuwane automatycznie podczas przebiegu.

2. Wkroplić pipetą 600 µl przygotowanych bibliotek próbek z biblioteki ampliconów w rozcieńczeniu (DAL, diluted amplicon library) do zbiorniczka z oznaczeniem **Load Samples** (Ładowanie próbek). Nie dotykać foliowego zamknięcia.
3. Po załadowaniu próbki sprawdzić, czy w zbiorniczku nie ma pęcherzyków powietrza. Jeśli występują pęcherzyki, delikatnie stuknąć kasetą o blat w celu uwolnienia pęcherzyków.

Rysunek 2 Ładowanie bibliotek



4. Przejść bezpośrednio do konfiguracji przebiegu za pomocą interfejsu oprogramowania MiSeq Operating Software (MOS).

Konfiguracja przebiegu

Pełną instrukcję konfiguracji przebiegu można znaleźć w dokumencie *MiSeqDx Instrument Reference Guide for MOS v4 (Instrukcja obsługi aparatu MiSeqDx z oprogramowaniem MOS w wer. 4)* (nr dokumentu: 200010452).

1. Zalogować się do aparatu MiSeqDx przy użyciu lokalnego hasła do oprogramowania Local Run Manager.
2. Na ekranie głównym oprogramowania MOS wybrać **Sequence** (Sekwencjonowanie).
3. Wybrać przebieg z listy, a następnie nacisnąć **Next** (Dalej).
Zostaną otwarte ekrany konfiguracji przebiegu w następującej kolejności: Load Flow Cell (Ładowanie komory przepływowej), Load Reagents (Ładowanie odczynników), Review (Przegląd) i Pre-Run Check (Przegląd wyników wstępnego testu kontrolnego).
4. Po wyświetleniu ekranu Load Flow Cell (Ładowanie komory przepływowej) wyczyścić, a następnie załadować komorę przepływową.
5. Zamknąć zatrzask i drzwiczki przedziału komory przepływowej.
Przed rozpoczęciem przebiegu zarówno zatrzask, jak i drzwiczki przedziału powinny być zamknięte. Gdy komora przepływowa zostanie załadowana, oprogramowanie odczytuje i rejestruje znacznik RFID. W prawym dolnym rogu ekranu zostaje wyświetlone potwierdzenie pomyślnego odczytu znacznika RFID.
6. Aby załadować butelkę z roztworem MiSeqDx SBS (PR2), postępować zgodnie z monitami oprogramowania. Upewnić się, że butelka na zlewki jest pusta i załadować kasetę odczynników. Po załadowaniu butelki z roztworem MiSeqDx SBS (PR2) i kasety odczynników oprogramowanie odczytuje i rejestruje znacznik RFID. W prawym dolnym rogu ekranu zostaje wyświetlone potwierdzenie pomyślnego odczytu znacznika RFID.

7. Po rozpoczęciu przebiegu otwiera się ekran Sequencing (Sekwencjonowanie). Ten ekran w sposób wizualny przedstawia trwający przebieg, w tym poziomy intensywności i wyniki jakościowe.

Wyniki

Oprogramowanie do analizy w czasie rzeczywistym (RTA) jest zintegrowanym oprogramowaniem, które wykonuje analizę obrazów i rozpoznawanie nukleotydów, a także przypisuje wyniki jakościowe do poszczególnych nukleotydów w każdym cyklu sekwencjonowania. Po zakończeniu analizy podstawowej wybrany w aparacie MiSeqDx moduł oprogramowania opisany w sekcji [Tworzenie przebiegu za pomocą lokalnego menedżera przebiegu na stronie 9](#) rozpoczyna analizę wtórną. Inne procedury opisano w dokumentacji konkretnego oznaczenia.

Procedury kontroli jakości

Oprogramowanie MiSeqDx ocenia każdy przebieg, próbkę i rozpoznanie nukleotydu w oparciu o miary kontroli jakości. Gdy jest to wymagane, ocenie pod kątem oczekiwanych wyników należy również poddać dodatnie i ujemne próby kontrolne uwzględnione w przygotowaniu biblioteki.

Charakterystyka działania

Wszystkie badania wykonywano na aparacie MiSeqDx.

W badaniach linii zarodkowych do przygotowania bibliotek wykorzystywano odczynniki z testu MiSeqDx Cystic Fibrosis 139-Variant Assay do oznaczania 139 wariantów genu mukowiscydozy albo z zestawu do sekwencjonowania niestandardowych ampliconów TruSeq Custom Amplicon Kit Dx. W obu zestawach wykorzystywane są te same odczynniki do przygotowywania bibliotek, z jedynie jedną różnicą na poziomie proceduralnym dotyczącą liczby cykli łańcuchowej reakcji polimerazy (PCR) (odpowiednio 25 i 28). Jak wykazano w badaniu poziomów wejściowych DNA z wykorzystaniem zestawu TruSeq Custom Amplicon Kit Dx, dzięki dodatkowym cyklom reakcji PCR w przypadku zestawu do sekwencjonowania niestandardowych ampliconów TruSeq Custom Amplicon Kit Dx (50 ng) dopuszczalny jest niższy poziom wejściowy DNA niż w teście MiSeqDx Cystic Fibrosis 139-Variant Assay (250 ng) do oznaczania 139 wariantów genu mukowiscydozy. Biblioteki sporządzone z wykorzystaniem testu MiSeqDx Cystic Fibrosis 139-Variant Assay do oznaczania 139 wariantów genu mukowiscydozy sekwencjonowano odczynnikami wchodzącymi w skład zestawu. Biblioteki sporządzone z wykorzystaniem zestawu do sekwencjonowania niestandardowych ampliconów TruSeq Custom Amplicon Kit Dx sekwencjonowano odczynnikami z zestawu MiSeqDx Reagent Kit v3.

Odczynniki sekwencjonujące użyte w tym drugim przypadku charakteryzują się wyższym poziomem wyjściowym w porównaniu z odczynnikami stosowanymi w teście MiSeqDx Cystic Fibrosis 139-Variant Assay do oznaczania 139 wariantów genu mukowiscydozy.

Testy obejmują zakresy przepustowości pod względem próbek obsługiwane przy użyciu zestawu MiSeqDx Reagent Kit v3 Micro. Aparat MiSeqDx może obsłużyć 1–96 próbek/przebieg, w zależności od oznaczenia. Zestaw odczynników MiSeqDx Reagent Kit v3 Micro został opracowany na potrzeby obsługi przy niższej przepustowości pod względem próbek w wybranych oznaczeniach.

W badaniach wariantów somatycznych wykorzystywano zestaw do sekwencjonowania niestandardowych amplikonów TruSeq Custom Amplicon Kit Dx z zestawem odczynników MiSeqDx Reagent Kit v3.

Po procedurach właściwych dla linii zarodkowych lub wariantów somatycznych opisanych dla zestawu do sekwencjonowania niestandardowych amplikonów TruSeq Custom Amplicon Kit Dx następowała analiza z wykorzystaniem odpowiednio modułu wariantów linii zarodkowych lub modułu wariantów somatycznych z dwoma wyjątkami. W badaniach z wykorzystaniem jednego genu (procedura linii zarodkowej, test MiSeqDx Cystic Fibrosis 139-Variant Assay do oznaczania 139 wariantów genu mukowiscydozy) lub dwóch genów (procedura somatyczna) wykorzystywano procedury i moduły analityczne właściwe dla oznaczenia.

UWAGA Zawartość genomową amplikonów opisano względem sekwencjonowanej nici genomu. W przypadku amplikonów zaprojektowanych dla nici antysensownej referencyjna zawartość genomowa jest odwrotnie komplementarna (np. regiony PolyA na amplikonach nici antysensownej odpowiadają regionom PolyT w genomie referencyjnym).

Definicje obliczeń stosowanych w określaniu parametrów użytkowych

- Procentową zgodność wyników dodatnich (PPA, ang. Positive Percent Agreement) oblicza się jako odsetek prawidłowo zgłoszonych w oznaczeniu loci sklasyfikowanych jako warianty w oparciu o metodę referencyjną.
 - $(\text{liczba loci wariantowych prawidłowo zgłoszonych w oznaczeniu}) / (\text{całkowita liczba loci wariantowych})$
Loci wariantowe zgłoszone w oznaczeniu i zgodne z metodą referencyjną stanowią wyniki prawdziwie dodatnie (TP, ang. True Positive). Loci wariantowe zgłoszone w oznaczeniu jako rozpoznania sekwencji referencyjnej lub jako inne rozpoznania wariantowe stanowią wyniki fałszywie ujemne (FN, ang. False Negative).
- Procentową zgodność wyników ujemnych (NPA, ang. Negative Percent Agreement) oblicza się jako odsetek loci sklasyfikowanych metodą referencyjną jako typ dziki i prawidłowo zgłoszonych w oznaczeniu.

- (liczba loci typu dzikiego prawidłowo zgłoszonych w oznaczeniu) / (całkowita liczba loci typu dzikiego)
Loci typu dzikiego zgłoszone przez oznaczenie i zgodne z metodą referencyjną stanowią wyniki prawdziwie ujemne (TN, ang. True Negative). Loci typu dzikiego zgłoszone w oznaczeniu jako warianty stanowią wyniki fałszywie dodatnie (FP, ang. False Positive).
- Ogólną zgodność procentową (OPA, ang. Overall Percent Agreement) oblicza się jako odsetek loci prawidłowo sklasyfikowanych w oznaczeniu w stosunku do metody referencyjnej.
 - ((liczba loci wariantowych prawidłowo zgłoszonych w oznaczeniu) + (liczba loci typu dzikiego prawidłowo zgłoszonych w oznaczeniu)) / ((całkowita liczba loci wariantowych) + (całkowita liczba loci typu dzikiego))
- W przypadku rozpoznawania wariantów obliczenia parametrów PPA, NPA i OPA nie obejmują braków rozpoznań (loci wariantowych lub referencyjnych niespełniających warunków jednego lub większej liczby filtrów jakościowych). W dwóch badaniach w mierze „Odsetek prawidłowych rozpoznań” uwzględniono brak rozpoznań, co zostało zaznaczone przy stosownych tabelach.
- Odsetek rozpoznań oblicza się jako całkowitą liczbę loci, które przeszły przez filtry, podzieloną przez całkowitą liczbę pozycji zsekwencjonowanych dla chromosomów 1–22. Chromosomy X i Y nie są uwzględniane. Metryka ta nie uwzględnia zgodności rozpoznań z metodą referencyjną.

Informacje na temat cech wydajności zależnych od czynników przedanalizacyjnych (np. sposobów ekstrakcji lub wprowadzania DNA) można znaleźć w ulotce dołączonej do opakowania odpowiedniej metody przygotowywania bibliotek.

Indeksowanie próbek

Dodawane podczas przygotowywania biblioteki startery indeksujące próbki służą do przypisywania unikalnych sekwencji do poszczególnych próbek DNA, co umożliwia pulowanie wielu próbek w pojedynczym przebiegu sekwencjonowania. Indeksowanie próbek przebadano w procedurach dotyczących linii zarodkowych i somatycznych.

Przebadano łącznie 96 indeksów próbek przy użyciu oznaczenia reprezentatywnego w celu rozpoznania szeregu genów wchodzących w skład wszystkich 23 ludzkich chromosomów i zawierających po 12 588 nukleotydów na łańcuch w celu weryfikacji stałej zdolności do rozpoznawania genotypu w określonej próbce z zastosowaniem różnych kombinacji starterów indeksujących. Chromosom Y nie zawiera regionów o wysokiej ufności i nie został poddany badaniu. Osiem unikalnych próbek zostało przebadanych przy użyciu 12 różnych kombinacji starterów indeksujących. Wyniki dla próbek uzyskane przy użyciu modułu analizy linii zarodkowych porównano do bazy Platinum Genomes w wersji 2016-01. Parametr PPA (dla SNV i polimorfizmów typu indel) przekroczył 97% (faktycznie dodatnie rozpoznania wynosiły co najmniej 70 dla SNV, 38 dla insercji, 36 dla delecji), zaś parametr NPA wynosił 100% (co najmniej 23 440 pozycji referencyjnych na kombinację indeksów) dla każdej z 96 kombinacji indeksów. Niezależnie od powyższego przeprowadzono badanie pojedynczego indeksu w celu stwierdzenia, że chemiczny mechanizm sekwencjonowania zestawu odczynników MiSeqDx Reagent Kit v3 jest w stanie obsłużyć mniej niż osiem próbek (mechanizm chemiczny wcześniejszego

zestawu uniwersalnego MiSeqDx Universal Kit 1.0 był ograniczony do oznaczania co najmniej ośmiu próbek). Dla pojedynczego indeksu wyznaczono wartości PPA wynoszące 98,9% (180/182) dla SNV, 100% (38/38) dla insercji i 100% (46/46) dla delecji. NPA wyniósł 100% (23 856/23 856).

W celu zmierzenia dokładności indeksowej dla wariantów somatycznych występujących z częstościami z zakresu 0,05–0,10 przy użyciu modułu do oznaczeń wariantów somatycznych przetestowano dwanaście powtórzeń (24 biblioteki) próbki (do rozpoznania w liniach somatycznych wykorzystuje się kombinacje dwóch indeksów w każdym powtórzeniu). PPA wyniósł 100% dla SNV (64/64), insercji (11/11) i delecji (19/19). NPA wyniósł 100% (co najmniej 11 590 pozycji referencyjnych na każdą kombinację indeksów) dla wszystkich kombinacji indeksów.

Przenoszenie próbek

Procedura oznaczeń przy użyciu aparatu MiSeqDx obejmuje przygotowanie biblioteki i sekwencjonowanie wielu próbek i kontroli, które są wszystkie przetwarzane jednocześnie. Przeprowadzono badanie przeniesienia próbki mające na celu stwierdzenie, czy wyniki fałszywie dodatnie, powstałe na skutek przeniesienia zanieczyszczenia z sąsiedniego dołka na etapie przygotowania biblioteki lub przeniesienia zanieczyszczenia między kolejnymi przebiegami, mają wpływ na wyniki oznaczeń. Do tego celu użyto wariantów somatycznych, jako że ich wykrywanie możliwe jest przy niższych częstościach występowania w allelach niż ma to miejsce w przypadku wariantów linii zarodkowych.

Zestaw próbek obejmował cztery próbki genomowego DNA z linii komórkowych; każda próbka zawierała różne mutacje panelowe dwugenowego oznaczenia reprezentatywnego. Próbki były dobrane w taki sposób, by mutacja w danej pozycji w jednej z próbek posiadała punkt odniesienia właściwy dla sekwencji dzikiego typu w pozostałych próbkach.

Przeniesienie między dołkami określa się jako tryb błędu potencjalnie generowanego w ręcznych etapach przetwarzania (pipetowanie, przemieszanie próbek itp.). Ocena efektu przeniesienia między próbkami obejmowała przeprowadzenie dwóch przebiegów:

- Szachownicowy układ z wysokim poziomem wejściowym genomowego DNA (gDNA) zawierającego mutację w genie 1 naprzemiennie z niskim poziomem wejściowym gDNA zawierającego mutację w genie 2.
- Szachownicowy układ z wysokim poziomem wejściowym gDNA zawierającego mutację w genie 2 naprzemiennie z niskim poziomem wejściowym gDNA zawierającego mutację w genie 1.

W każdym z przebiegów dokonano oceny 12 powtórzeń pod kątem wyników fałszywie dodatnich (np. mutacja genu 1 zgłoszona dla dołka oznakowanego jako próbka z mutacją genu 2 lub odwrotnie).

Przeniesienie między przebiegami określa się jako tryb błędu potencjalnie generowanego przez pozostałość z poprzedniego przebiegu sekwencjonowania. W celu stwierdzenia ewentualnego przeniesienia między przebiegami przygotowano dwie płytki, z których każda zawierała 11 powtórzeń pojedynczej, unikalnej próbki gDNA o wysokim poziomie wejściowym oraz próbę ślepą; obie serie sekwencjonowano kolejno na jednym aparacie MiSeqDx, po czym dokonano oceny wyników fałszywie dodatnich. W pierwszym przebiegu sekwencjonowano 11 powtórzeń próbki z mutacją genu 2 i jedną

próbę ślełą. W drugim przebiegu sekwencjonowano 11 powtórzeń próbki z mutacją genu 1 i jedną próbę ślełą. W pierwszej kolejności wykonano sekwencjonowanie biblioteki próbki z mutacją genu 2, po którym wykonano sekwencjonowanie biblioteki próbki z mutacją genu 1, a następnie ponowne sekwencjonowanie biblioteki próbki z mutacją genu 2. Zaobserwowanie mutacji genu 2 w przebiegu obejmującym wyłącznie mutacje genu 1 lub odwrotnie wskazywałoby na zajście efektu przeniesienia. Zgłoszono zero wyników fałszywie dodatnich (0/24, 0%) będących wynikiem przeniesienia *między dołkami*. Wykryto wszystkie oczekiwane mutacje. Zgłoszono zero wyników fałszywie dodatnich (0/24, 0%) będących wynikiem przeniesienia *między przebiegami*. Wykryto wszystkie oczekiwane mutacje. Zgłoszono zero wyników fałszywie dodatnich (0/48, 0%) będących wynikiem przeniesienia *ogółem* (przeniesienia między dołkami i przeniesienia między przebiegami łącznie).

Cechy wydajności – linia zarodkowa

W opisywanych tu badaniach do analizy danych z sekwencjonowania wykorzystywano moduł do oznaczania wariantów zarodkowych, z wyjątkiem badań wykorzystujących panel jednogenowy, w których wykorzystywano moduł swoisty dla oznaczenia.

Dokładność

W celu oceny dokładności aparatu MiSeqDx z zestawem odczynników MiSeqDx Reagent Kit v3 i DNA o wysokiej jakości przeprowadzono opisane poniżej badanie. W tym badaniu użyto oznaczenia reprezentatywnego w celu rozpoznania różnych genów obejmujących 12 588 nukleotydów znajdujących się w 23 różnych chromosomach przy użyciu 150 amplikonów. Chromosom Y nie zawiera regionów o wysokiej ufnosci i nie został poddany badaniu. W badaniu użyto 12 unikalnych próbek pochodzących od dwojga rodziców i 10 dzieci, które to próbki wcześniej były często sekwencjonowane różnymi metodami przez wiele laboratoriów. Dostępnych jest pięć próbek pochodzących od osób płci żeńskiej oraz siedem próbek od osób płci męskiej. Każdą z próbek badano w dwóch powtórzeniach. Dokładność została określona dla wariantów SNV, insercji i delecji przez porównanie danych z badania z dobrze scharakteryzowaną referencyjną bazą danych. Sekwencję w referencyjnej bazie danych (Platinum Genomes wersja 2016-01) uzyskano w wyniku połączenia wielu metod sekwencjonowania, publicznie dostępnych danych oraz informacji o dziedziczeniu. O ile nie podano inaczej, regiony o wysokiej zgodności zostały zdefiniowane w oparciu o ww. metodę referencyjną. Łącznie próbki poddano przebiegowi osiem razy. Tabele mające ilustrować dokładność rozpoznań opierają się na danych z pierwszego przebiegu.

Tabela 1 zawiera dane z badania prezentowane w postaci procentowych zgodności wyników dodatnich i ujemnych dla poszczególnych próbek, przy czym na potrzeby obliczeń zgodności PPA wyniki wariantów są porównywane z dobrze scharakteryzowanymi, złożonymi danymi referencyjnymi. Trzy typy wariantów (SNV, insercje i delecje) zostały ze sobą połączone. Ponieważ metoda referencyjna dostarcza tylko wyniki dotyczące wariantów pojedynczych nukleotydów oraz insercji/delecji, wyniki nukleotydów niewariantowych na potrzeby obliczeń zgodności NPA porównywane są z referencyjną sekwencją 19 genomu ludzkiego.

Tabela 1 Zgodność wyników rozpoznawania nukleotydów przez aparat MiSeqDx w zależności od próbki

Próbka	Średni odsetek rozpoznań	Całkowita liczba wariantów	Całkowita liczba wariantów TP	Całkowita liczba wariantów FN	Łączna liczba nierozpoznanych nukleotydów	Całkowita liczba rozpoznań TN	PPA	NPA	OPA
NA12877	>99,9	152	152	0	4	24024	100	100	100
NA12878	>99,9	270	266	0	4	23856	100	100	100
NA12879	>99,9	192	190	1	1	24054	99,5	100	>99,9
NA12880	>99,9	222	220	0	6	24052	100	100	100
NA12881	>99,9	250	247	1	2	23862	99,6	100	>99,9
NA12882	>99,9	200	196	2	2	23962	99,0	100	>99,9
NA12883	>99,9	226	224	0	6	23870	100	100	100
NA12884	>99,9	228	226	1	1	23942	99,6	100	>99,9
NA12885	>99,9	244	240	2	2	23942	99,2	100	>99,9
NA12886	>99,9	230	228	1	1	23888	99,6	100	>99,9
NA12888	>99,9	216	216	0	4	24002	100	100	100
NA12893	>99,9	236	234	0	2	23810	100	100	100

Oznaczenie reprezentatywne składało się ze 150 amplikonów zaprojektowanych tak, by pokryć szeroki zakres zawartości materiału genetycznego. Zawartość GC w amplikonach mieściła się w zakresie 26–87%. Amplikony charakteryzowały się również pewnym zakresem powtórzeń pojedynczych nukleotydów (np. PolyA, PolyT), dinukleotydów i trinukleotydów. Uzyskane dane skompilowano w oparciu o poszczególne amplikony (tabela 2) w celu oznaczenia wpływu zawartości materiału genetycznego na odsetek prawidłowych rozpoznań. Odsetek prawidłowych rozpoznań obejmował rozpoznania wariantów i rozpoznania referencyjne i był mniejszy niż 100% w przypadku wystąpienia nieprawidłowych rozpoznań lub braku rozpoznań. Warianty nie są rozpoznawane w przypadku niespełnienia warunków jednego lub większej liczby filtrów dla rozpoznania wariantu (np. niewystarczającego pokrycia). Siedem spośród ośmiu wariantów FN w tabeli 2 obejmowało insercję 1 bp w amplikonie 111, zawierającym również homopolimer PolyA i mającym zawartość GC na poziomie 0,29. Ósmy wariant FN (nieprawidłowe rozpoznanie) był spowodowany rozpoznaniem jako wariantu homozygotycznego oczekiwanej heterozygotycznej SNV w amplikonie 125 o zawartości GC na poziomie 0,68. Częstość

wariantu SNV wyniosła 0,71, co jest wartością powyżej progu 0,70 przyjętego dla zaklasyfikowania wariantu jako homozygotycznego. Amplikonem o najniższym odsetku prawidłowych rozpoznań (98,2%) był amplikon 17, dla którego stwierdzono 40 przypadków braku rozpoznania, a który zawierał powtórzenia AT oraz miał zawartość GC na poziomie 27%.

Tabela 2 Dokładność aparatu MiSeqDx na poziomie amplikonu

Amplikon	Chromosom	Początek amplikonu	Koniec amplikonu	Rozmiar analizowanego fragmentu	Zasady w regionach wysokiej ufności	Materiał genetyczny amplikonu	Zawartość GC	Łączna liczba prawidłowo rozpoznanych nukleotydów	Łączna liczba nieprawidłowo rozpoznanych nukleotydów	Łączna liczba nierozpoznanych nukleotydów	Odsetek prawidłowo rozpoznanych nukleotydów
1	1	36450499	36450591	93	93	Indel	0,22	2232	0	0	100
2	1	109465122	109465200	79	79	PolyA (5), PolyC (5), indel	0,38	1896	0	0	100
3	1	218353867	218353957	91	91	Indel	0,4	2184	0	0	100
4	1	223906657	223906748	92	92	Indel	0,49	2208	0	0	100
5	1	228526602	228526682	81	81	PolyG (5)	0,69	1944	0	0	100
6	1	236372039	236372108	70	70	PolyT (10), indel	0,39	1680	0	0	100
7	1	247812041	247812128	88	88	PolyA (5), CT (3), TAA(3), indel	0,27	2112	0	0	100
8	2	55862774	55862863	90	90	Indel	0,28	2160	0	0	100
9	2	87003930	87004009	80	80	Indel	0,38	1920	0	0	100
10	2	177016721	177016805	85	81	nd.	0,65	1944	0	0	100
11	2	186625727	186625801	75	75	PolyA (8)	0,35	1800	0	0	100
12	2	190323504	190323591	88	88	PolyT (5)	0,42	2112	0	0	100
13	2	200796740	200796826	87	87	PolyT (5), indel	0,31	2088	0	0	100
14	2	212245049	212245139	91	91	PolyT (5), PolyA (6), indel	0,3	2184	0	0	100
15	2	228147052	228147144	93	93	nd.	0,43	2232	0	0	100
16	2	235016350	235016422	73	73	PolyT (5), indel	0,42	1752	0	0	100

Amplikon	Chromosom	Początek amplikonu	Koniec amplikonu	Rozmiar analizowanego fragmentu	Zasady w regionach wysokiej ufności	Materiał genetyczny amplikonu	Zawartość GC	Łączna liczba prawidłowo rozpoznanych nukleotydów	Łączna liczba nieprawidłowo rozpoznanych nukleotydów	Łączna liczba nierozpoznanych nukleotydów	Odsetek prawidłowo rozpoznanych nukleotydów
17	3	4466229	4466321	93	93	AT(3), indel	0,27	2192	0	40	98,2
18	3	46620561	46620643	83	83	nd.	0,43	1992	0	0	100
19	3	49851331	49851400	70	70	CT(3), indel	0,49	1680	0	0	100
20	3	189713161	189713248	88	88	PolyA (5), PolyT (5), PolyA (9), TG (3)	0,41	2112	0	0	100
21	3	190106030	190106104	75	74	Indel	0,57	1774	0	2	99,9
22	4	2233667	2233744	78	78	PolyA (6)	0,26	1872	0	0	100
23	4	7780541	7780637	97	97	PolyG (6), PolyT (5), PolyA (5)	0,42	2328	0	0	100
24	4	15688604	15688681	78	78	nd.	0,29	1872	0	0	100
25	4	56236521	56236586	66	62	PolyA (5), indel	0,36	1488	0	0	100
26	4	102839244	102839314	71	69	PolyA (5)	0,46	1656	0	0	100
27	4	164446743	164446804	62	62	PolyA (7), indel	0,27	1488	0	0	100
28	5	1882081	1882158	78	75	nd.	0,78	1800	0	0	100
29	5	14769061	14769144	84	84	GT(3), CCA (3)	0,62	2016	0	0	100
30	5	41069808	41069871	64	64	nd.	0,39	1536	0	0	100
31	5	74077114	74077196	83	83	PolyA (6), indel	0,3	1992	0	0	100
32	5	147475343	147475409	67	67	PolyT (5)	0,37	1608	0	0	100
33	5	149323731	149323821	91	91	CT(4), AG(3)	0,55	2184	0	0	100
34	5	155662213	155662287	75	75	Indel	0,43	1800	0	0	100
35	6	6318713	6318814	10	10	PolyG (6)	0,68	2448	0	0	100
36	6	24949983	24950074	92	92	Indel	0,63	2208	0	0	100
37	6	31084900	31084999	100	94	GCT(5), indel	0,61	2244	0	12	99,5

Amplikon	Chromosom	Początek amplikonu	Koniec amplikonu	Rozmiar analizowanego fragmentu	Zasady w regionach wysokiej ufności	Materiał genetyczny amplikonu	Zawartość GC	Łączna liczba prawidłowo rozpoznanych nukleotydów	Łączna liczba nieprawidłowo rozpoznanych nukleotydów	Łączna liczba nierozpoznanych nukleotydów	Odsetek prawidłowo rozpoznanych nukleotydów
38	6	32147987	32148084	98	98	PolyT (5), TCT(3), CTT (3)	0,55	2352	0	0	100
39	6	32986864	32986958	95	95	Indel	0,53	2280	0	0	100
40	6	33408498	33408583	86	86	PolyC (6)	0,7	2064	0	0	100
41	6	41647401	41647495	95	94	PolyG (5), indel	0,61	2256	0	0	100
42	6	112435865	112435955	91	91	PolyA (5)	0,44	2184	0	0	100
43	7	22202076	22202148	73	73	nd.	0,44	1752	0	0	100
44	7	66276100	66276187	88	88	Indel	0,35	2112	0	0	100
45	7	77365735	77365821	87	87	PolyA (7), AG (4)	0,26	2088	0	0	100
46	7	110939946	110940030	85	85	Indel	0,38	2040	0	0	100
47	7	128533468	128533557	90	90	PolyG (5), indel	0,62	2160	0	0	100
48	7	149503875	149503965	91	91	PolyG (6), PolyC (6), indel	0,71	2184	0	0	100
49	7	154404519	154404599	81	66	nd.	0,31	1584	0	0	100
50	7	156476507	156476599	93	93	Indel	0,35	2232	0	0	100
51	8	1817312	1817394	83	83	nd.	0,42	1992	0	0	100
52	8	24811020	24811109	90	89	PolyG (7), CTC(4), indel	0,61	2113	0	23	98,9
53	8	76518625	76518691	67	67	Indel	0,3	1608	0	0	100
54	9	103054909	103055006	98	98	PolyG (6)	0,67	2352	0	0	100
55	9	105586150	105586214	65	65	Indel	0,32	1560	0	0	100
56	9	107620823	107620918	96	96	nd.	0,49	2304	0	0	100
57	9	123769149	123769231	83	83	AT(3)	0,37	1992	0	0	100
58	9	138995345	138995441	97	97	PolyC (6), indel	0,68	2328	0	0	100

Amplikon	Chromosom	Początek amplitonu	Koniec amplitonu	Rozmiar analizowanego fragmentu	Zasady w regionach wysokiej ufności	Materiał genetyczny amplitonu	Zawartość GC	Łączna liczba prawidłowo rozpoznanych nukleotydów	Łączna liczba nieprawidłowo rozpoznanych nukleotydów	Łączna liczba nierozpoznanych nukleotydów	Odsetek prawidłowo rozpoznanych nukleotydów
59	10	5987120	5987198	79	78	PolyG (5), indel	0,47	1872	0	0	100
60	10	11784629	11784726	98	91	GC(3)	0,87	2184	0	0	100
61	10	27317777	27317855	79	79	PolyT (5)	0,3	1896	0	0	100
62	10	33018351	33018440	90	90	PolyA (5), PolyT (5)	0,2	2160	0	0	100
63	10	45084159	45084253	95	95	Indel	0,35	2280	0	0	100
64	10	55892599	55892687	89	88	AC(11), indel	0,42	2102	0	10	99,5
65	10	101611250	101611329	80	80	nd.	0,49	1920	0	0	100
66	10	118351373	118351453	81	81	nd.	0,51	1944	0	0	100
67	11	8159816	8159912	97	96	nd.	0,45	2304	0	0	100
68	11	30177648	30177717	70	70	Indel	0,46	1680	0	0	100
69	11	47470345	47470444	100	100	nd.	0,65	2400	0	0	100
70	11	59837679	59837740	62	62	Indel	0,37	1488	0	0	100
71	11	64418856	64418957	102	102	nd.	0,59	2448	0	0	100
72	11	93529612	93529684	73	73	PolyA (5)	0,4	1752	0	0	100
73	11	101347052	101347136	85	85	nd.	0,42	2040	0	0	100
74	11	102477336	102477426	91	91	PolyG (6)	0,55	2184	0	0	100
75	11	118406285	118406369	85	85	Indel	0,53	2040	0	0	100
76	11	120357801	120357885	85	85	PolyA (5), CA (3), indel	0,34	2040	0	0	100
77	11	125769313	125769397	85	85	GA(3)	0,52	2040	0	0	100
78	12	2834770	2834853	84	84	PolyC (5), indel	0,52	2016	0	0	100
79	12	26811004	26811096	93	93	PolyA (7), AC (4)	0,33	2232	0	0	100
80	12	30881766	30881846	81	81	nd.	0,49	1944	0	0	100
81	12	88474105	88474175	71	71	PolyA (6)	0,35	1704	0	0	100
82	12	120966872	120966966	95	95	PolyG (5)	0,68	2280	0	0	100

Amplikon	Chromosom	Początek amplikonu	Koniec amplikonu	Rozmiar analizowanego fragmentu	Zasady w regionach wysokiej ufności	Materiał genetyczny amplikonu	Zawartość GC	Łączna liczba prawidłowo rozpoznanych nukleotydów	Łączna liczba nieprawidłowo rozpoznanych nukleotydów	Łączna liczba nierozpoznanych nukleotydów	Odsetek prawidłowo rozpoznanych nukleotydów
83	13	24167504	24167576	73	73	nd.	0,52	1752	0	0	100
84	13	25816961	25817049	89	88	PolyA (5), PolyT (7), PolyA (7), indel	0,22	2112	0	0	100
85	13	44880112	44880200	89	89	Indel	0,49	2136	0	0	100
86	13	77665218	77665294	77	77	Indel	0,39	1848	0	0	100
87	14	31619327	31619393	67	67	GA(3), TA(3)	0,39	1608	0	0	100
88	14	39517884	39517966	83	83	nd.	0,25	1992	0	0	100
89	14	46958962	46959034	73	72	PolyT (5), indel	0,19	1727	0	1	99,9
90	14	58050030	58050110	81	81	Indel	0,38	1944	0	0	100
91	14	82390559	82390649	91	91	Indel	0,35	2184	0	0	100
92	14	92549544	92549609	66	66	PolyA (5)	0,41	1584	0	0	100
93	14	102808496	102808589	94	94	Indel	0,62	2256	0	0	100
94	15	43170751	43170848	98	96	PolyC (5)	0,45	2304	0	0	100
95	15	63446149	63446216	68	68	Indel	0,25	1632	0	0	100
96	15	77879807	77879901	95	93	PolyG (5), indel	0,68	2232	0	0	100
97	15	81625334	81625428	95	95	PolyT (6)	0,43	2280	0	0	100
98	15	85438263	85438334	72	71	Indel	0,65	1704	0	0	100
99	15	89817413	89817503	91	91	nd.	0,36	2184	0	0	100
100	15	89864274	89864343	70	70	Indel	0,56	1680	0	0	100
101	16	1894910	1894972	63	63	nd.	0,27	1512	0	0	100
102	16	28997904	28997998	95	95	PolyC (5)	0,67	2280	0	0	100
103	16	53682908	53682994	87	87	TA(3)	0,41	2088	0	0	100
104	16	57954406	57954509	104	104	PolyC (5)	0,67	2496	0	0	100
105	16	85706375	85706465	91	91	PolyT (5), indel	0,37	2184	0	0	100

Amplikon	Chromosom	Początek amplikonu	Koniec amplikonu	Rozmiar analizowanego fragmentu	Zasady w regionach wysokiej ufności	Materiał genetyczny amplikonu	Zawartość GC	Łączna liczba prawidłowo rozpoznanych nukleotydów	Łączna liczba nieprawidłowo rozpoznanych nukleotydów	Łączna liczba nierozpoznanych nukleotydów	Odsetek prawidłowo rozpoznanych nukleotydów
106	17	3563920	3564008	89	89	GC(3)	0,64	2136	0	0	100
107	17	3594191	3594277	87	87	PolyC (5), indel	0,67	2088	0	0	100
108	17	3970090	3970180	91	91	Indel	0,46	2184	0	0	100
109	17	16084945	16085037	93	93	Indel	0,26	2232	0	0	100
110	17	33998759	33998849	91	89	PolyT (5)	0,54	2136	0	0	100
111	17	39589691	39589774	84	82	PolyA (13), indel (x2)	0,29	1944	7	17	98,8
112	17	41244394	41244484	91	91	PolyA (5)	0,34	2184	0	0	100
113	17	45438866	45438957	92	92	PolyA (7), AT (3), AT(4), AT (4), indel	0,26	2208	0	0	100
114	17	61502432	61502510	79	79	Indel	0,41	1887	0	9	99,5
115	17	64023582	64023667	86	86	PolyT (7)	0,22	2064	0	0	100
116	17	72308237	72308320	84	84	GAG(3)	0,62	2016	0	0	100
117	18	2616456	2616522	67	67	GA(3)	0,31	1608	0	0	100
118	18	6980478	6980568	91	91	nd.	0,37	2184	0	0	100
119	18	9888026	9888094	69	69	PolyA (6), TG (3)	0,43	1656	0	0	100
120	18	38836999	38837073	75	75	PolyA (5), indel	0,37	1800	0	0	100
121	18	47405382	47405462	81	81	CTC(3), indel	0,47	1944	0	0	100
122	18	54815665	54815749	85	85	CT(3), indel	0,45	2040	0	0	100
123	18	59773996	59774060	65	65	nd.	0,48	1560	0	0	100
124	19	625143	625241	99	99	nd.	0,59	2376	0	0	100
125	19	18121418	18121491	74	74	nd.	0,68	1775	1	0	99,9
126	19	18186574	18186643	70	70	nd.	0,64	1680	0	0	100
127	20	746056	746149	94	94	nd.	0,61	2256	0	0	100
128	20	10633195	10633276	82	82	AC(3)	0,59	1968	0	0	100

Amplikon	Chromosom	Początek amplikonu	Koniec amplikonu	Rozmiar analizowanego fragmentu	Zasady w regionach wysokiej ufności	Materiał genetyczny amplikonu	Zawartość GC	Łączna liczba prawidłowo rozpoznanych nukleotydów	Łączna liczba nieprawidłowo rozpoznanych nukleotydów	Łączna liczba nierozpoznanych nukleotydów	Odsetek prawidłowo rozpoznanych nukleotydów
129	20	17705633	17705708	76	76	CT(3)	0,58	1824	0	0	100
130	20	21766821	21766890	70	70	GT(3), TG(4), indel	0,46	1680	0	0	100
131	20	25278421	25278521	101	101	Indel	0,63	2424	0	0	100
132	20	50897302	50897368	67	67	Indel	0,36	1608	0	0	100
133	20	62331904	62331994	91	88	PolyG (6)	0,73	2112	0	0	100
134	20	62690860	62690946	87	87	Indel	0,57	2088	0	0	100
135	21	30300823	30300888	66	66	Indel	0,35	1584	0	0	100
136	21	33694176	33694273	98	98	PolyT (6), CA (3)	0,54	2352	0	0	100
137	21	36710706	36710792	87	87	GT(3), indel	0,39	2088	0	0	100
138	21	46644924	46644992	69	69	PolyA (6), AG (3), indel	0,32	1656	0	0	100
139	21	46705575	46705664	90	90	PolyT (5), PolyA (6)	0,5	2160	0	0	100
140	22	25750774	25750873	100	100	Indel	0,63	2400	0	0	100
141	22	32439233	32439329	97	97	nd.	0,68	2328	0	0	100
142	22	37409844	37409940	97	97	Indel	0,46	2328	0	0	100
143	22	37637596	37637694	99	99	nd.	0,6	2376	0	0	100
144	22	47081347	47081438	92	92	Indel	0,66	2208	0	0	100
145	X	15870424	15870492	69	69	PolyT (5)	0,26	1656	0	0	100
146	X	135288543	135288611	69	69	PolyC (5)	0,62	1656	0	0	100
147	X	135290777	135290847	71	71	nd.	0,52	1704	0	0	100
148	Y	2655397	2655461	65	0	nd.	0,55	0	0	0	nd.
149	Y	2655519	2655609	91	0	nd.	0,48	0	0	0	nd.
150	Y	2655609	2655679	71	0	PolyA (5)	0,37	0	0	0	nd.

Warianty, dla których stwierdzono brak rozpoznań, wymieniono w [tabeli 3](#). Filtry, których zastosowanie skutkowało brakiem rozpoznań, zostały wymienione w tabeli. Insercją w amplikonie 111 odfiltrowano w dziewięciu na 16 wystąpień, zaś w pozostałych siedmiu przypadkach wariant został rozpoznany jako referencyjny (wynik FN).

Tabela 3 Podsumowanie wariantów, dla których stwierdzono brak rozpoznania

Nr amplikonu	Chr:Pozycja	Wariant	Zawartość odpowiedniego amplikonu	Filtr	Warianty pominięte	Warianty oczekiwane	Rozpoznania fałszywie ujemne
64	10:55892600	TAC>T	AC(11), 42% GC	R5x9 ¹	10	10	0
111	17:39589692	C>CA	PolyA (13), 29% GC	R5x9	9	16	7

¹ R5x9: filtr powtórzeń. Wariant jest odfiltrowywany, jeśli całość lub część sekwencji wariantu występuje powtarzalnie w genomie referencyjnym sąsiednio do pozycji wariantu. W oznaczeniu referencyjnym wymaganych jest co najmniej dziewięć powtórzeń, przy uwzględnianiu jedynie powtórzeń o długości nie większej niż 5 bp.

Wyniki sekwencjonowania próbki NA12878 porównano z genotypem o wysokim poziomie ufności, ustalonym dla próbki NA12878 przez ośrodek NIST (National Institutes of Standards and Technology) (v.2.19). Na sumaryczną liczbę 150 amplikonów 92 w pełni zawierało się w regionach genomu o wysokim poziomie ufności, 41 częściowo nakładało się na te regiony, zaś 17 nie nakładało się na sekwencję NIST. W rezultacie dla każdego powtórzenia uzyskano 10 000 współrzędnych do porównań. Rozpoznane nukleotydy niewariantowe zostały porównane z referencyjną sekwencją 19 genomu ludzkiego. Wyniki oznaczeń dokładności przedstawiono w tabeli 4.

Tabela 4 Zgodność wyników rozpoznań nukleotydów przy użyciu aparatu MiSeqDx dla próbki NA12878 z bazy danych NIST

Próbka	Liczba amplikonów	Średni odsetek rozpoznań	Całkowita liczba wariantów TP	Całkowita liczba wariantów FN	Całkowita liczba rozpoznań TN	Całkowita liczba rozpoznań FP	PPA	NPA	OPA
NA12878	133	99,98	208	0	19380	0	100	100	100

Próbki poddano dalszej analizie pod kątem rozpoznawania małych insercji i delecji (polimorfizmów typu indel) (tabela 5). W niektórych przypadkach polimorfizmy typu indel były wspólne dla dwóch lub większej liczby próbek, co znalazło odzwierciedlenie w kolumnie „Całkowita liczba powtórzeń próbki z polimorfizmem typu indel”. Wyniki dla obu powtórzeń 12 ważnych próbek podano w tabeli 5. Obecnych było w sumie 71 polimorfizmów typu indel o wielkości od 1 do 24 bp w przypadku insercji i od 1 do 25 bp w przypadku delecji. Wykryto 69 polimorfizmów typu indel z dodatnią zgodnością procentową wynoszącą 100%. Jedna delecja (amplikon 64; delecja 2 bp (chr10 55892600 TAC>T)) nie dała prawidłowych rozpoznań z uwagi na fakt, że żaden z wariantów nie został rozpoznany z powodu

zastosowania filtra R5x9. Z tego powodu nie można było obliczyć wartości parametru PPA, którego definicja wyklucza brak rozpoznań. Inny polimorfizm typu indel w postaci insercji 1 bp (chr17 39589692 C>CA w amplikonie 111) również nie dał prawidłowych rozpoznań z uwagi na brak rozpoznania dziewięciu wariantów z powodu zastosowania filtra R5x9 i fałszywie ujemny status siedmiu rozpoznań.

Tabela 5 Podsumowanie wykrywania polimorfizmów typu indel przy użyciu aparatu MiSeqDx

Amplikon	Chromosom	Pozycja	Rozmiar analizowanego fragmentu	Rodzaj i długość polimorfizmu typu indel w amplikonie	Indel	Całkowita liczba powtórzeń próbki z polimorfizmem typu indel	Liczba nierozpoznanych nukleotydów	Całkowita liczba nieprawidłowych rozpoznań polimorfizmów typu indel	Całkowita liczba prawidłowych rozpoznań polimorfizmów typu indel	PPA
1	1	36450544	93	Delecja 25 bp	GAAAATTTAATGAAACACATTGTCCT>G	2	0	0	2	100
2	1	109465165	79	Delecja 3 bp	ACTT>A	12	0	0	12	100
3	1	218353908	91	Insercja 23 bp	T>TTTTAATAGCAAAAAGAGGCTAGA	24	0	0	24	100
4	1	223906701	92	Delecja 17 bp	GACAGACTGTGAGGAAGA>G	10	0	0	10	100
6	1	236372081	70	Insercja 5 bp	C>CTTAAG	10	0	0	10	100
7	1	247812083	88	Insercja 3 bp	C>CATG	10	0	0	10	100
8	2	55862804	90	Insercja 7 bp	T>TTTGGTAA	14	0	0	14	100
9	2	87003972	80	Delecja 6 bp	TTATCTC>T	6	0	0	6	100
13	2	200796749	87	Insercja 5 bp	T>TTAAAA	24	0	0	24	100
14	2	212245090	91	Insercja 12 bp	C>CTGAAAATAGGAT	14	0	0	14	100
16	2	235016388	73	Insercja 2 bp	A>ATG	12	0	0	12	100
17	3	4466274	93	Delecja 23 bp	TAACCTAAAATTACAAAATAACCC>T	2	0	0	2	100
19	3	49851375	70	Insercja 9 bp	C>CCTGGCTCCT	4	0	0	4	100
21	3	190106071	75	Delecja 1 bp	AG>A	20	0	0	20	100
25	4	56236567	66	Delecja 8 bp	TAACCGAAA>T	12	0	0	12	100
27	4	164446785	62	Insercja 11 bp	T>TTATGGTATTGA	12	0	0	12	100
31	5	74077155	83	Delecja 4 bp	TAGTA>T	10	0	0	10	100
34	5	155662255	75	Insercja 8 bp	G>GCCTACTGA	20	0	0	20	100
36	6	24950035	92	Delecja 21 bp	CCCTGGGTGCTATAGCCCACCA>C	10	0	0	10	100
37	6	31084942	100	Delecja 3 bp	GCTT>G	14	0	0	14	100
39	6	32986905	95	Delecja 25 bp	CTTTCACCTTCCCCTCATGCAAAG>C	12	0	0	12	100
41	6	41647442	95	Delecja 23 bp	GGCATGAGGCTTGGTGACATGGCA>G	8	0	0	8	100
44	7	66276142	88	Insercja 1 bp	C>CT	16	0	0	16	100

Amplikon	Chromosom	Pozycja	Rozmiar analizowanego fragmentu	Rodzaj i długość polimorfizmu typu indel w amplikonie	Indel	Całkowita liczba powtórzeń próbki z polimorfizmem typu indel	Liczba nierozpoznanych nukleotydów	Całkowita liczba nieprawidłowych rozpoznań polimorfizmów typu indel	Całkowita liczba prawidłowych rozpoznań polimorfizmów typu indel	PPA
46	7	110939983	85	Delecja 4 bp	CAAGT>C	12	0	0	12	100
47	7	128533514	90	Insercja 1 bp	T>TC	24	0	0	24	100
48	7	149503916	91	Delecja 4 bp	GGATA>G	8	0	0	8	100
50	7	156476548	93	Delecja 11 bp	GAATCTGCACTT>G	12	0	0	12	100
52	8	24811064	90	Delecja 1 bp	AG>A	24	0	0	24	100
53	8	76518677	67	Insercja 4 bp	T>TACTG	14	0	0	14	100
55	9	105586193	65	Insercja 4 bp	C>CAATT	2	0	0	2	100
58	9	138995370	97	Delecja 21 bp	TCTGGGGGGCAGCCCCTGAGGG>T	14	0	0	14	100
59	10	5987158	79	Delecja 3 bp	TAAC>T	10	0	0	10	100
63	10	45084202	95	Delecja 16 bp	AGCGTCTATAACCAAT>A	12	0	0	12	100
64	10	55892600	89	Delecja 2 bp	TAC>T	10	10	0	0	nd.
68	11	30177690	70	Insercja 2 bp	C>CTG	10	0	0	10	100
70	11	59837721	62	Insercja 8 bp	T>TTATGAAAA	12	0	0	12	100
75	11	118406328	85	Delecja 8 bp	CAGTGTGGA>C	10	0	0	10	100
76	11	120357842	85	Delecja 2 bp	CTT>C	10	0	0	10	100
78	12	2834814	84	Insercja 21 bp	T>TTCTCAGTACGGTGAACCCCAG	24	0	0	24	100
84	13	25817002	89	Insercja 19 bp	C>CAAAATATAAAAAGCTCCCT	24	0	0	24	100
85	13	44880152	89	Insercja 4 bp	C>CCTGT	12	0	0	12	100
86	13	77665265	77	Delecja 20 bp	ATCTATTTTCTAATAGACGGC>A	14	0	0	14	100
89	14	46958967	73	Delecja 22 bp	TTTAAATTTGAATGTGATAAAA>T	24	0	0	24	100
90	14	58050081	81	Insercja 4 bp	C>CTGAT	20	0	0	20	100
91	14	82390602	91	Delecja 16 bp	CTTGCTCTATAAACCGT>C	10	0	0	10	100
93	14	102808554	94	Delecja 5 bp	CGTGGA>C	10	0	0	10	100
95	15	63446199	68	Delecja 6 bp	CAAAATT>C	12	0	0	12	100
96	15	77879862	95	Delecja 25 bp	GCCCCTGAGCCAGCCTCCCCTCTTA>G	14	0	0	14	100
98	15	85438311	72	Insercja 3 bp	C>CTTG	8	0	0	8	100
100	15	89864316	70	Insercja 4 bp	G>GCTAC	8	0	0	8	100

Amplikon	Chromosom	Pozycja	Rozmiar analizowanego fragmentu	Rodzaj i długość polimorfizmu typu indel w amplikonie	Indel	Całkowita liczba powtórzeń próbki z polimorfizmem typu indel	Liczba nierozpoznanych nukleotydów	Całkowita liczba nieprawidłowych rozpoznań polimorfizmów typu indel	Całkowita liczba prawidłowych rozpoznań polimorfizmów typu indel	PPA
105	16	85706416	91	Delecja 7 bp	ATTATTTC>A	16	0	0	16	100
107	17	3594276	87	Delecja 1 bp	TG>T	2	0	0	2	100
108	17	3970133	91	Insercja 18 bp	A>ATCCTATTCTACTCTGAAT	10	0	0	10	100
109	17	16084985	93	Insercja 4 bp	A>AACAC	10	0	0	10	100
111	17	39589692	84	Insercja 1 bp	C>CA	16	9	7	0	0
112	17	39589739	84	Insercja 24 bp	T>TTCTGAAGGTCAAGTCTATCCCTGA	24	0	0	24	100
113	17	45438886	92	Delecja 4 bp	CAGTG>C	12	0	0	12	100
114	17	61502459	79	Delecja 12 bp	TTTGTATCTGCTG>T	20	0	0	20	100
120	18	38837054	75	Insercja 22 bp	T>TGTATCTTAGCAAAGTTTCTCA	24	0	0	24	100
121	18	47405425	81	Insercja 3 bp	T>TGAG	20	0	0	20	100
122	18	54815706	85	Delecja 2 bp	ACT>A	20	0	0	20	100
130	20	21766863	70	Delecja 15 bp	TACTTGAGAAGTGG>T	4	0	0	4	100
131	20	25278464	101	Insercja 5 bp	A>AGTGGG	20	0	0	20	100
132	20	50897361	67	Insercja 11 bp	G>GGAATGTCAGCC	24	0	0	24	100
134	20	62690925	87	Delecja 16 bp	TCCTGGCTGGCCTGTGG>T	10	0	0	10	100
135	21	30300873	66	Insercja 11 bp	G>GATAAACTTTA	10	0	0	10	100
137	21	36710749	87	Delecja 21 bp	ACTCAAGATAACTCATGTTATC>A	16	0	0	16	100
138	21	46644985	69	Delecja 5 bp	GTTGTT>G	8	0	0	8	100
140	22	25750814	100	Insercja 6 bp	C>CAGGGCA	20	0	0	20	100
142	22	37409885	97	Insercja 5 bp	C>CTGTTT	2	0	0	2	100
144	22	47081407	92	Delecja 10 bp	GGGCACAGGCA>G	12	0	0	12	100

Odtwarzalność

Przeprowadzono dwa badania mające na celu ocenę odtwarzalności aparatu MiSeqDx w przypadku linii komórkowych (badanie 1 i 2) lub krwi o obniżonej zawartości leukocytów z dodatkiem linii komórkowych (badanie 2). W badaniu 1 wykorzystywano wiele aparatów. W badaniu 2 wykorzystywano wiele stanowisk.

Badanie 1

Odtwarzalność oznaczeń przy użyciu aparatu MiSeqDx wyznaczano przy użyciu dwóch aparatów, dwóch operatorów i dwóch zestawów odczynników w całkowitej liczbie ośmiu przebiegów. Oznaczenie reprezentatywne, próbki i metoda referencyjna były takie same, jak opisane dla badania dokładności.

Wyniki przedstawiono w podziale na poszczególne amplikony dla każdego z aparatów (tabela 6) tak, by wykazać odtwarzalność rozpoznań na różnych aparatach. Odsetek prawidłowych rozpoznań wyznaczano po uwzględnieniu zarówno nieprawidłowych rozpoznań, jak i braku rozpoznań (niespełnienia jednego lub większej liczby filtrów rozpoznawania wariantów). Oba aparaty charakteryzowały się podobną liczbą braków rozpoznań i nieprawidłowych rozpoznań dla poszczególnych amplikonów.

Tabela 6 Badanie – wyniki odtwarzalności badań między aparatami MiSeqDx (na poziomie amplikonu)

Amplikon	Chromosom	Początek amplikonu	Koniec amplikonu	Rozmiar analizowanego fragmentu	Zasady w regionach wysokiej ufności	Materiał genetyczny amplikonu	Zawartość GC	MiSeqDx 1		MiSeqDx 2			
								Łączna liczba prawidłowo rozpoznanych nukleotydów	Łączna liczba nieprawidłowo rozpoznanych nukleotydów	Łączna liczba nierozpoznanych nukleotydów	Łączna liczba prawidłowo rozpoznanych nukleotydów	Łączna liczba nieprawidłowo rozpoznanych nukleotydów	Łączna liczba nierozpoznanych nukleotydów
1	1	36450499	36450591	93	93	Indel	0,22	8928	0	0	8928	0	0
2	1	109465122	109465200	79	79	PolyA (5), PolyC (5), indel	0,38	7584	0	0	7584	0	0
3	1	218353867	218353957	91	91	Indel	0,4	8736	0	0	8736	0	0
4	1	223906657	223906748	92	92	Indel	0,49	8832	0	0	8832	0	0
5	1	228526602	228526682	81	81	PolyG (5)	0,69	7776	0	0	7776	0	0
6	1	236372039	236372108	70	70	PolyT (10), indel	0,39	6720	0	0	6720	0	0
7	1	247812041	247812128	88	88	PolyA (5), CT(3), TAA (3), indel	0,27	8448	0	0	8448	0	0
8	2	55862774	55862863	90	90	Indel	0,28	8640	0	0	8640	0	0
9	2	87003930	87004009	80	80	Indel	0,38	7680	0	0	7680	0	0
10	2	177016721	177016805	85	81	nd.	0,65	7775	1	0	7775	1	0
11	2	186625727	186625801	75	75	PolyA (8)	0,35	7200	0	0	7200	0	0

Ampli- kon	Chromo- som	Początek amplikonu	Koniec amplikonu	Rozmiar analizo- wanego fragmentu	Zasady w regionach wysokiej ufności	Materiał genetyczny amplikonu	Zawartość GC	MiSeqDx 1			MiSeqDx 2		
								Łączna liczba prawidłowo rozpoznanych nukleotydów	Łączna liczba nieprawidłowo rozpoznanych nukleotydów	Łączna liczba nieroz- poznanych nukleotydów	Łączna liczba prawidłowo rozpo- znanych nukleotydów	Łączna liczba niepra- widłowo rozpo- znanych nukleotydów	Łączna liczba nieroz- poznanych nukleotydów
12	2	190323504	190323591	88	88	PolyT (5)	0,42	8448	0	0	8448	0	0
13	2	200796740	200796826	87	87	PolyT (5), indel	0,31	8352	0	0	8352	0	0
14	2	212245049	212245139	91	91	PolyT (5), PolyA (6), indel	0,3	8736	0	0	8736	0	0
15	2	228147052	228147144	93	93	nd.	0,43	8928	0	0	8928	0	0
16	2	235016350	235016422	73	73	PolyT (5), indel	0,42	7008	0	0	7008	0	0
17	3	44662229	4466321	93	93	AT(3), indel	0,27	8761	0	167	8760	0	168
18	3	46620561	46620643	83	83	nd.	0,43	7968	0	0	7968	0	0
19	3	49851331	49851400	70	70	CT(3), indel	0,49	6720	0	0	6720	0	0
20	3	189713161	189713248	88	88	PolyA (5), PolyT (5), PolyA (9), TG(3)	0,41	8448	0	0	8448	0	0
21	3	190106030	190106104	75	74	Indel	0,57	7096	0	8	7096	0	8
22	4	2233667	2233744	78	78	PolyA (6)	0,26	7488	0	0	7488	0	0
23	4	7780541	7780637	97	97	PolyG (6), PolyT (5), PolyA (5)	0,42	9312	0	0	9312	0	0
24	4	15688604	15688681	78	78	nd.	0,29	7488	0	0	7488	0	0
25	4	56236521	56236586	66	62	PolyA (5), indel	0,36	5952	0	0	5952	0	0
26	4	102839244	102839314	71	69	PolyA (5)	0,46	6624	0	0	6624	0	0
27	4	164446743	164446804	62	62	PolyA (7), indel	0,27	5952	0	0	5952	0	0
28	5	1882081	1882158	78	75	nd.	0,78	7200	0	0	7200	0	0

Ampli- kon	Chromo- som	Początek amplikonu	Koniec amplikonu	Rozmiar analizo- wanego fragmentu	Zasady w regionach wysokiej ufności	Materiał genetyczny amplikonu	Zawartość GC	MiSeqDx 1		MiSeqDx 2			
								Łączna liczba prawidłowo rozpoznanych nukleotydów	Łączna liczba nieprawidłowo rozpoznanych nukleotydów	Łączna liczba nieroz- poznanych nukleotydów	Łączna liczba prawidłowo rozpo- znanych nukleotydów	Łączna liczba niepra- widłowo rozpo- znanych nukleotydów	Łączna liczba nieroz- poznanych nukleotydów
29	5	14769061	14769144	84	84	GT(3), CCA (3)	0,62	8064	0	0	8064	0	0
30	5	41069808	41069871	64	64	nd.	0,39	6144	0	0	6144	0	0
31	5	74077114	74077196	83	83	PolyA (6), indel	0,3	7968	0	0	7968	0	0
32	5	147475343	147475409	67	67	PolyT (5)	0,37	6432	0	0	6432	0	0
33	5	149323731	149323821	91	91	CT(4), AG (3)	0,55	8736	0	0	8736	0	0
34	5	155662213	155662287	75	75	Indel	0,43	7200	0	0	7200	0	0
35	6	6318713	6318814	102	102	PolyG (6)	0,68	9792	0	0	9792	0	0
36	6	24949983	24950074	92	92	Indel	0,63	8832	0	0	8832	0	0
37	6	31084900	31084999	100	94	GCT(5), indel	0,61	8979	0	45	8979	0	45
38	6	32147987	32148084	98	98	PolyT (5), TCT(3), CTT(3)	0,55	9408	0	0	9408	0	0
39	6	32986864	32986958	95	95	Indel	0,53	9120	0	0	9120	0	0
40	6	33408498	33408583	86	86	PolyC (6)	0,7	8256	0	0	8256	0	0
41	6	41647401	41647495	95	94	PolyG (5), indel	0,61	9024	0	0	9024	0	0
42	6	112435865	112435955	91	91	PolyA (5)	0,44	8736	0	0	8736	0	0
43	7	22202076	22202148	73	73	nd.	0,44	7008	0	0	7008	0	0
44	7	66276100	66276187	88	88	Indel	0,35	8448	0	0	8448	0	0
45	7	77365735	77365821	87	87	PolyA (7), AG(4)	0,26	8352	0	0	8352	0	0
46	7	110939946	110940030	85	85	Indel	0,38	8160	0	0	8160	0	0

Ampli- kon	Chromo- som	Początek amplikonu	Koniec amplikonu	Rozmiar analizo- wanego fragmentu	Zasady w regionach wysokiej ufności	Materiał genetyczny amplikonu	Zawartość GC	MiSeqDx 1		MiSeqDx 2			
								Łączna liczba prawidłowo rozpoznanych nukleotydów	Łączna liczba nieprawidłowo rozpoznanych nukleotydów	Łączna liczba nieroz- poznanych nukleotydów	Łączna liczba prawidłowo rozpo- znanych nukleotydów	Łączna liczba niepra- widłowo rozpo- znanych nukleotydów	Łączna liczba nieroz- poznanych nukleotydów
47	7	128533468	128533557	90	90	PolyG (5), indel	0,62	8550	0	90	8550	0	90
48	7	149503875	149503965	91	91	PolyG (6), PolyC (6), indel	0,71	8736	0	0	8736	0	0
49	7	154404519	154404599	81	66	nd.	0,31	6336	0	0	6336	0	0
50	7	156476507	156476599	93	93	Indel	0,35	8928	0	0	8928	0	0
51	8	1817312	1817394	83	83	nd.	0,42	7968	0	0	7968	0	0
52	8	24811020	24811109	90	89	PolyG (7), CTC(4), indel	0,61	8452	0	92	8449	0	95
53	8	76518625	76518691	67	67	Indel	0,3	6432	0	0	6432	0	0
54	9	103054909	103055006	98	98	PolyG (6)	0,67	9408	0	0	9408	0	0
55	9	105586150	105586214	65	65	Indel	0,32	6240	0	0	6240	0	0
56	9	107620823	107620918	96	96	nd.	0,49	9216	0	0	9216	0	0
57	9	123769149	123769231	83	83	AT(3)	0,37	7968	0	0	7968	0	0
58	9	138995345	138995441	97	97	PolyC (6), indel	0,68	9312	0	0	9312	0	0
59	10	5987120	5987198	79	78	PolyG (5), indel	0,47	7488	0	0	7488	0	0
60	10	11784629	11784726	98	91	GC(3)	0,87	8644	1	91	8644	1	91
61	10	27317777	27317855	79	79	PolyT (5)	0,3	7584	0	0	7584	0	0
62	10	33018351	33018440	90	90	PolyA (5), PolyT (5)	0,2	8640	0	0	8640	0	0
63	10	45084159	45084253	95	95	Indel	0,35	9120	0	0	9120	0	0
64	10	55892599	55892687	89	88	AC(11), indel	0,42	8408	0	40	8407	0	41
65	10	101611250	101611329	80	80	nd.	0,49	7680	0	0	7680	0	0

Ampli- kon	Chromo- som	Początek amplikonu	Koniec amplikonu	Rozmiar analizo- wanego fragmentu	Zasady w regionach wysokiej ufności	Materiał genetyczny amplikonu	Zawartość GC	MiSeqDx 1		MiSeqDx 2			
								Łączna liczba prawidłowo rozpoznanych nukleotydów	Łączna liczba nieprawidłowo rozpoznanych nukleotydów	Łączna liczba nieroz- poznanych nukleotydów	Łączna liczba prawidłowo rozpo- znanych nukleotydów	Łączna liczba niepra- widłowo rozpo- znanych nukleotydów	Łączna liczba nieroz- poznanych nukleotydów
66	10	118351373	118351453	81	81	nd.	0,51	7776	0	0	7776	0	0
67	11	8159816	8159912	97	96	nd.	0,45	9216	0	0	9216	0	0
68	11	30177648	30177717	70	70	Indel	0,46	6720	0	0	6720	0	0
69	11	47470345	47470444	100	100	nd.	0,65	9600	0	0	9600	0	0
70	11	59837679	59837740	62	62	Indel	0,37	5952	0	0	5952	0	0
71	11	64418856	64418957	102	102	nd.	0,59	9792	0	0	9792	0	0
72	11	93529612	93529684	73	73	PolyA (5)	0,4	7008	0	0	7008	0	0
73	11	101347052	101347136	85	85	nd.	0,42	8160	0	0	8160	0	0
74	11	102477336	102477426	91	91	PolyG (6)	0,55	8736	0	0	8736	0	0
75	11	118406285	118406369	85	85	Indel	0,53	8160	0	0	8160	0	0
76	11	120357801	120357885	85	85	PolyA (5), CA(3), indel	0,34	8160	0	0	8160	0	0
77	11	125769313	125769397	85	85	GA(3)	0,52	8160	0	0	8160	0	0
78	12	2834770	2834853	84	84	PolyC (5), indel	0,52	8064	0	0	8064	0	0
79	12	26811004	26811096	93	93	PolyA (7), AC(4)	0,33	8928	0	0	8928	0	0
80	12	30881766	30881846	81	81	nd.	0,49	7776	0	0	7776	0	0
81	12	88474105	88474175	71	71	PolyA (6)	0,35	6816	0	0	6816	0	0
82	12	120966872	120966966	95	95	PolyG (5)	0,68	9117	3	0	9119	1	0
83	13	24167504	24167576	73	73	nd.	0,52	7008	0	0	7008	0	0
84	13	25816961	25817049	89	88	PolyA (5), PolyT (7), PolyA (7), indel	0,22	8448	0	0	8448	0	0
85	13	44880112	44880200	89	89	Indel	0,49	8544	0	0	8544	0	0

Ampli- kon	Chromo- som	Początek amplikonu	Koniec amplikonu	Rozmiar analizo- wanego fragmentu	Zasady w regionach wysokiej ufności	Materiał genetyczny amplikonu	Zawartość GC	MiSeqDx 1		MiSeqDx 2			
								Łączna liczba prawidłowo rozpoznanych nukleotydów	Łączna liczba nieprawidłowo rozpoznanych nukleotydów	Łączna liczba nieroz- poznanych nukleotydów	Łączna liczba prawidłowo rozpo- znanych nukleotydów	Łączna liczba niepra- widłowo rozpo- znanych nukleotydów	Łączna liczba nieroz- poznanych nukleotydów
86	13	77665218	77665294	77	77	Indel	0,39	7392	0	0	7392	0	0
87	14	31619327	31619393	67	67	GA(3), TA(3)	0,39	6432	0	0	6432	0	0
88	14	39517884	39517966	83	83	nd.	0,25	7968	0	0	7968	0	0
89	14	46958962	46959034	73	72	PolyT (5), indel	0,19	6830	0	82	6835	0	77
90	14	58050030	58050110	81	81	Indel	0,38	7776	0	0	7776	0	0
91	14	82390559	82390649	91	91	Indel	0,35	8736	0	0	8736	0	0
92	14	92549544	92549609	66	66	PolyA (5)	0,41	6336	0	0	6336	0	0
93	14	102808496	102808589	94	94	Indel	0,62	9024	0	0	9024	0	0
94	15	43170751	43170848	98	96	PolyC (5)	0,45	9216	0	0	9216	0	0
95	15	63446149	63446216	68	68	Indel	0,25	6528	0	0	6528	0	0
96	15	77879807	77879901	95	93	PolyG (5), indel	0,68	8928	0	0	8926	2	0
97	15	81625334	81625428	95	95	PolyT (6)	0,43	9120	0	0	9120	0	0
98	15	85438263	85438334	72	71	Indel	0,65	6816	0	0	6816	0	0
99	15	89817413	89817503	91	91	nd.	0,36	8736	0	0	8736	0	0
100	15	89864274	89864343	70	70	Indel	0,56	6720	0	0	6720	0	0
101	16	1894910	1894972	63	63	nd.	0,27	6048	0	0	6048	0	0
102	16	28997904	28997998	95	95	PolyC (5)	0,67	9120	0	0	9120	0	0
103	16	53682908	53682994	87	87	TA(3)	0,41	8352	0	0	8352	0	0
104	16	57954406	57954509	104	104	PolyC (5)	0,67	9984	0	0	9984	0	0
105	16	85706375	85706465	91	91	PolyT (5), indel	0,37	8736	0	0	8736	0	0
106	17	3563920	3564008	89	89	GC(3)	0,64	8544	0	0	8544	0	0
107	17	3594191	3594277	87	87	PolyC (5), indel	0,67	8347	0	5	8347	0	5

Ampli- kon	Chromo- som	Początek amplikonu	Koniec amplikonu	Rozmiar analizo- wanego fragmentu	Zasady w regionach wysokiej ufności	Materiał genetyczny amplikonu	Zawartość GC	MiSeqDx 1			MiSeqDx 2		
								Łączna liczba prawidłowo rozpoznanych nukleotydów	Łączna liczba nieprawidłowo rozpoznanych nukleotydów	Łączna liczba nieroz- poznanych nukleotydów	Łączna liczba prawidłowo rozpo- znanych nukleotydów	Łączna liczba niepra- widłowo rozpo- znanych nukleotydów	Łączna liczba nieroz- poznanych nukleotydów
108	17	3970090	3970180	91	91	Indel	0,46	8736	0	0	8736	0	0
109	17	16084945	16085037	93	93	Indel	0,26	8928	0	0	8928	0	0
110	17	33998759	33998849	91	89	PolyT (5)	0,54	8544	0	0	8544	0	0
111	17	39589691	39589774	84	82	PolyA (13), indel (x2)	0,29	7776	7	89	7777	12	83
112	17	41244394	41244484	91	91	PolyA (5)	0,34	8736	0	0	8736	0	0
113	17	45438866	45438957	92	92	PolyA (7), AT(3), AT (4), AT(4), indel	0,26	8832	0	0	8832	0	0
114	17	61502432	61502510	79	79	Indel	0,41	7546	0	38	7547	0	37
115	17	64023582	64023667	86	86	PolyT (7)	0,22	8256	0	0	8256	0	0
116	17	72308237	72308320	84	84	GAG(3)	0,62	8064	0	0	8064	0	0
117	18	2616456	2616522	67	67	GA(3)	0,31	6432	0	0	6432	0	0
118	18	6980478	6980568	91	91	nd.	0,37	8736	0	0	8736	0	0
119	18	9888026	9888094	69	69	PolyA (6), TG(3)	0,43	6624	0	0	6624	0	0
120	18	38836999	38837073	75	75	PolyA (5), indel	0,37	7200	0	0	7200	0	0
121	18	47405382	47405462	81	81	CTC(3), indel	0,47	7776	0	0	7776	0	0
122	18	54815665	54815749	85	85	CT(3), indel	0,45	8160	0	0	8160	0	0
123	18	59773996	59774060	65	65	nd.	0,48	6240	0	0	6240	0	0
124	19	625143	625241	99	99	nd.	0,59	9504	0	0	9504	0	0
125	19	18121418	18121491	74	74	nd.	0,68	7102	2	0	7104	0	0
126	19	18186574	18186643	70	70	nd.	0,64	6718	2	0	6718	2	0
127	20	746056	746149	94	94	nd.	0,61	9024	0	0	9024	0	0

Ampli- kon	Chromo- som	Początek amplikonu	Koniec amplikonu	Rozmiar analizo- wanego fragmentu	Zasady w regionach wysokiej ufności	Materiał genetyczny amplikonu	Zawartość GC	MiSeqDx 1		MiSeqDx 2			
								Łączna liczba prawidłowo rozpoznanych nukleotydów	Łączna liczba nieprawidłowo rozpoznanych nukleotydów	Łączna liczba nieroz- poznanych nukleotydów	Łączna liczba prawidłowo rozpo- znanych nukleotydów	Łączna liczba niepra- widłowo rozpo- znanych nukleotydów	Łączna liczba nieroz- poznanych nukleotydów
128	20	10633195	10633276	82	82	AC(3)	0,59	7872	0	0	7872	0	0
129	20	17705633	17705708	76	76	CT(3)	0,58	7296	0	0	7296	0	0
130	20	21766821	21766890	70	70	GT(3), TG (4), indel	0,46	6720	0	0	6720	0	0
131	20	25278421	25278521	101	101	Indel	0,63	9696	0	0	9696	0	0
132	20	50897302	50897368	67	67	Indel	0,36	6432	0	0	6432	0	0
133	20	62331904	62331994	91	88	PolyG (6)	0,73	8360	0	88	8360	0	88
134	20	62690860	62690946	87	87	Indel	0,57	8352	0	0	8352	0	0
135	21	30300823	30300888	66	66	Indel	0,35	6336	0	0	6336	0	0
136	21	33694176	33694273	98	98	PolyT (6), CA(3)	0,54	9408	0	0	9408	0	0
137	21	36710706	36710792	87	87	GT(3), indel	0,39	8352	0	0	8352	0	0
138	21	46644924	46644992	69	69	PolyA (6), AG(3), indel	0,32	6603	0	21	6601	0	23
139	21	46705575	46705664	90	90	PolyT (5), PolyA (6)	0,5	8640	0	0	8640	0	0
140	22	25750774	25750873	100	100	Indel	0,63	9600	0	0	9600	0	0
141	22	32439233	32439329	97	97	nd.	0,68	9312	0	0	9312	0	0
142	22	37409844	37409940	97	97	Indel	0,46	9312	0	0	9312	0	0
143	22	37637596	37637694	99	99	nd.	0,6	9504	0	0	9504	0	0
144	22	47081347	47081438	92	92	Indel	0,66	8832	0	0	8832	0	0
145	X	15870424	15870492	69	69	PolyT (5)	0,26	6624	0	0	6624	0	0
146	X	135288543	135288611	69	69	PolyC (5)	0,62	6624	0	0	6624	0	0
147	X	135290777	135290847	71	71	nd.	0,52	6816	0	0	6816	0	0
148	Y	2655397	2655461	65	0	nd.	0,55	0	0	0	0	0	0
149	Y	2655519	2655609	91	0	nd.	0,48	0	0	0	0	0	0
150	Y	2655609	2655679	71	0	PolyA (5)	0,37	0	0	0	0	0	0

Wyniki badania odtwarzalności analizowano dla poszczególnych operatorów w oparciu o częstość wariantów (tabela 7). Analiza wykazała spójność częstości wariantów wykrywanych przez operatorów. Przedstawiono wartości średniej częstości wariantu +/- 1 odchylenie standardowe.

Tabela 7 Wyniki oznaczeń przy użyciu aparatu MiSeqDx dokonywanych przez poszczególnych operatorów

Zakres częstości wariantu	Liczba unikalnych wariantów	Całkowita liczba wariantów analizowanych przez operatora 1	Całkowita liczba wariantów analizowanych przez operatora 2	Średnia (SD) częstość wariantu raportowana przez operatora 1	Średnia (SD) częstość wariantu raportowana przez operatora 2
Homozygotyczne (0,70–1,00)	2424	2424	2422	0,94 +/- 0,07	0,96 +/- 0,05
Heterozygotyczne (0,20–0,70)	8240	8132	8128	0,48 +/- 0,04	0,49 +/- 0,04

Wyniki badania odtwarzalności dla poszczególnych próbek zawierają dane ze wszystkich ośmiu przebiegów (tabela 8). Wykrywalność została oceniona oddzielnie dla każdego rodzaju wariantu – SNV, insercji i delecji. Pozycje referencyjne zostały wykluczone. Analiza wykazała, że uzyskane wyniki dotyczące wariantów były powtarzalne we wszystkich próbkach.

Tabela 8 Zgodność wyników rozpoznawania nukleotydów przez aparat MiSeqDx w zależności od próbki

Próbka	Warianty SNV					Insercje			Delecje			
	Łączna liczba	Łączna liczba TP	Łączna liczba F	Łączna liczba F	Łączna liczba F	Łączna liczba TP	Łączna liczba FP	Łączna liczba FN	Łączna liczba	Łączna liczba TP	Łączna liczba FP	Łączna liczba FN
NA12877	592	592	0	0	336	336	0	0	288	288	0	0
NA12878	1456	1456	0	0	320	304	0	0	384	368	0	0
NA12879	912	912	0	0	336	320	0	2	288	288	0	0
NA12880	1072	1071	0	0	384	384	0	0	320	304	0	0
NA12881	1248	1247	0	1	384	368	0	0	368	368	0	0
NA12882	944	943	0	1	352	336	0	4	304	288	0	0
NA12883	1088	1087	0	1	368	368	0	0	352	335	0	0
NA12884	1088	1088	0	0	400	384	0	5	336	336	0	0
NA12885	1200	1189	0	7	400	382	0	4	352	336	0	0

Próbka	Łączna liczba	Warianty SNV				Insercje			Delecje			
		Łączna liczba TP	Łączna liczba F	Łączna liczba F	Łączna liczba F	Łączna liczba TP	Łączna liczba FP	Łączna liczba FN	Łączna liczba TP	Łączna liczba FP	Łączna liczba FN	
NA12886	1104	1102	0	2	368	352	0	3	368	368	0	0
NA12888	1056	1054	0	2	368	368	0	0	304	304	0	0
NA12893	1168	1168	0	0	352	336	0	1	368	368	0	0

Dane uzyskane w oparciu o osiem przebiegów w przedstawionym badaniu odtwarzalności potwierdzają twierdzenie, że aparat MiSeqDx jest w stanie w sposób spójny sekwencjonować:

- Zawartość GC $\geq 19\%$ (wszystkie rozpoznane nukleotydy w 192 ze 192 sekwencjonowanych ampliconów z 19% zawartością GC rozpoznane prawidłowo z odsetkiem nierozpoznanych nukleotydów na poziomie 1,1%).
- Zawartość GC $\leq 78\%$ (wszystkie rozpoznane nukleotydy w 192 ze 192 sekwencjonowanych ampliconów z 78% zawartością GC rozpoznane prawidłowo przy braku nierozpoznanych nukleotydów).
- Łańcuchy PolyA o długości ≤ 8 (powtórzenie PolyA złożone z 8 nukleotydów rozpoznane prawidłowo w 192 ze 192 sekwencjonowanych ampliconów zawierających PolyA = 8).
- Łańcuchy PolyT o długości ≤ 10 (powtórzenie PolyT złożone z 10 nukleotydów rozpoznane prawidłowo w 192 ze 192 sekwencjonowanych ampliconów zawierających PolyT = 10).
- Łańcuchy PolyG o długości ≤ 7 (powtórzenie PolyG złożone z 7 nukleotydów rozpoznane prawidłowo w 192 ze 192 sekwencjonowanych ampliconów zawierających PolyG = 7).
- Łańcuchy PolyC o długości ≤ 6 (powtórzenie PolyC złożone z 6 nukleotydów rozpoznane prawidłowo w 576 z 576 sekwencjonowanych ampliconów zawierających PolyC = 6).
- Powtórzenia dinukleotydowe o długości $\leq 11x$ (wszystkie zidentyfikowane nukleotydy z 11-krotnym powtórzeniem dinukleotydu w 192 ze 192 sekwencjonowanych ampliconów zostały rozpoznane prawidłowo przy odsetku braku rozpoznań na poziomie 0,5%).
- Powtórzenia trinukleotydowe o długości $\leq 5x$ (wszystkie zidentyfikowane nukleotydy z 5-krotnym powtórzeniem trinukleotydu w 192 ze 192 sekwencjonowanych ampliconów zostały rozpoznane prawidłowo przy odsetku braku rozpoznań na poziomie 0,5%).
- Insercje 24 lub mniej nukleotydów i delecje 25 lub mniej nukleotydów:
 - insercje 24 lub mniej nukleotydów zidentyfikowane prawidłowo w 192 ze 192 próbek,
 - delecje 25 nukleotydów zidentyfikowane poprawnie w 223 próbkach przy nieprawidłowym rozpoznaniu 1 na 224 próbki.

Badanie 2

Przeprowadzono badanie odtwarzalności wyników między stanowiskami przy użyciu reprezentatywnego oznaczenia – testu Illumina MiSeqDx Cystic Fibrosis 139-Variant Assay do oznaczania 139 wariantów genu mukowiscydozy, obejmujące zmiany genu *CFTR* o znaczeniu klinicznym, przeanalizowane za pomocą oprogramowania MiSeq Reporter z wykorzystaniem procedury ukierunkowanego sekwencjonowania DNA przy użyciu aparatu MiSeqDx. Zaślepienie badanie przeprowadzono na 3 stanowiskach badawczych; na każdym z nich było 2 operatorów. Każdy z operatorów na każdym stanowisku przebadał dwa dobrze scharakteryzowane panele po 46 próbek, co dało łącznie 810 rozpoznań na stanowisko. Panele zawierały mieszaninę genomowego DNA z linii komórkowych ze znanymi wariantami w genie *CFTR*, a także krew o obniżonej zawartości leukocytów zmieszaną z liniami komórkowymi ze znanymi wariantami w genie *CFTR*. Próbkę krwi dostarczono w celu przeprowadzenia czynności ekstrakcji do przygotowania gDNA będącego materiałem wejściowym w procedurze oznaczenia. Odsetek prawidłowych próbek, określany jako liczba próbek spełniających wymagania parametrów kontroli jakości w pierwszej próbce, wyniósł 99,88%. Wszystkie wyniki dotyczą pierwszego badania.

Tabela 9 Podsumowanie wyników badania odtwarzalności przeprowadzonego przy użyciu reprezentatywnym testu MiSeqDx Cystic Fibrosis 139-Variant Assay do oznaczania 139 wariantów genu mukowiscydozy

Panel	Nr próbki	Genotyp próbki	Warianty	Łączna liczba rozpoznanych nukleotydów na stanowisko	Rozpoznania ze zgodnością dodatnią (warianty)			Rozpoznania ze zgodnością ujemną (typu dzikiego)			Liczba nieprawidłowo rozpoznanych nukleotydów	Liczba nierozpoznanych nukleotydów	Zgodność dodatnia (%)	Zgodność ujemna (%)	Zgodność ogólna (%)
					Stano-wisko 1	Stano-wisko 2	Stano-wisko 3	Stano-wisko 1	Stano-wisko 2	Stano-wisko 3					
A	1	S549N (HET)		810	6	6	6	804	804	804	0	0	100	100	100
A	2	1812-1G>A (HET)		810	6	6	6	804	804	804	0	0	100	100	100
A	3	Q493X/F508del (HET)		810	12	12	12	798	798	798	0	0	100	100	100
A	4 ¹	F508del/2184delA (HET)		810	12	12	12	797	798	798	0	1 ¹	100	100	100
A	5 ²	Y122X/R1158X (HET)		810	12	10	12	798	665	798	0	135 ²	94,44	94,44	94,44
A	6	F508del/2183AA>G (HET)		810	12	12	12	798	798	798	0	0	100	100	100
A	7	R75X (HET)		810	6	6	6	804	804	804	0	0	100	100	100
A	8	I507del/F508del (HET)		810	12	12	12	798	798	798	0	0	100	100	100
A	9 ³	F508del/W1282X (HET)		810	12	11	12	798	797	798	2 ³	0	97,22	99,96	99,92

Panel	Nr próbki	Genotyp próbki	Warianty	Łączna liczba rozpoznanych nukleotydów na stanowisko	Rozpoznanie ze zgodnością dodatnią (warianty)			Rozpoznanie ze zgodnością ujemną (typu dzikiego)			Liczba nieprawidłowo rozpoznanych nukleotydów	Liczba nierozpoznanych nukleotydów	Zgodność dodatnia (%)	Zgodność ujemna (%)	Zgodność ogólna (%)
					Stano-wisko 1	Stano-wisko 2	Stano-wisko 3	Stano-wisko 1	Stano-wisko 2	Stano-wisko 3					
A	10 ³	F508del/3272-26A>G (HET)		810	12	11	12	798	797	798	2 ³	0	97,22	99,96	99,92
A	11	F508del/3849+10kbC>T (HET)		810	12	12	12	798	798	798	0	0	100	100	100
A	12	621+1G>T/3120+1G>A (HET)		810	12	12	12	798	798	798	0	0	100	100	100
A	13	E60X/F508del (HET)		810	12	12	12	798	798	798	0	0	100	100	100
A	14	M1101K (HET)		810	6	6	6	804	804	804	0	0	100	100	100
A	15	M1101K (HOM)		810	6	6	6	804	804	804	0	0	100	100	100
A	16	F508del (HOM)	I506V, I507V, F508C – brak	828	6	6	6	822	822	822	0	0	100	100	100
A	17	F508del/3659delC (HET)		810	12	12	12	798	798	798	0	0	100	100	100
A	18	R117H/F508del (HET)	(TG)10 (T)9/ (TG)12 (T)5	816	18	18	18	798	798	798	0	0	100	100	100
A	19	621+1G>T/711+1G>T (HET)		810	12	12	12	798	798	798	0	0	100	100	100
A	20	G85E/621+1G>T (HET)		810	12	12	12	798	798	798	0	0	100	100	100
A	21	A455E/F508del (HET)		810	12	12	12	798	798	798	0	0	100	100	100
A	22	F508del/R560T (HET)		810	12	12	12	798	798	798	0	0	100	100	100
A	23	F508del/Y1092X (C>A) (HET)		810	12	12	12	798	798	798	0	0	100	100	100
A	24	N1303K (HET)		810	6	6	6	804	804	804	0	0	100	100	100

Panel	Nr próbki	Genotyp próbki	Warianty	Łączna liczba rozpoznanych nukleotydów na stanowisko	Rozpoznania ze zgodnością dodatnią (warianty)			Rozpoznania ze zgodnością ujemną (typu dzikiego)			Liczba nieprawidłowo rozpoznanych nukleotydów	Liczba nierozpoznanych nukleotydów	Zgodność dodatnia (%)	Zgodność ujemna (%)	Zgodność ogólna (%)
					Stano-wisko 1	Stano-wisko 2	Stano-wisko 3	Stano-wisko 1	Stano-wisko 2	Stano-wisko 3					
A	25	G542X (HOM)		810	6	6	6	804	804	804	0	0	100	100	100
A	26	G542X (HET)		810	6	6	6	804	804	804	0	0	100	100	100
A	27	G551D/R553X (HET)		810	12	12	12	798	798	798	0	0	100	100	100
A	28	3849+10kbC>T (HOM)		810	6	6	6	804	804	804	0	0	100	100	100
A	29	WT		810	0	0	0	810	810	810	0	0	nd.	100	100
A	30	F508del (HET)		810	6	6	6	804	804	804	0	0	100	100	100
A	31	1717-1G>A (HET)		810	6	6	6	804	804	804	0	0	100	100	100
A	32	R1162X (HET)		810	6	6	6	804	804	804	0	0	100	100	100
A	33	R347P/G551D (HET)		810	12	12	12	798	798	798	0	0	100	100	100
A	34	R334W (HET)		810	6	6	6	804	804	804	0	0	100	100	100
A	35	WT		810	0	0	0	810	810	810	0	0	nd.	100	100
A	36	G85E (HET)		810	6	6	6	804	804	804	0	0	100	100	100
A	37	I336K (HET)		810	6	6	6	804	804	804	0	0	100	100	100
A	38	WT		810	0	0	0	810	810	810	0	0	nd.	100	100
A	39	F508del/3849+10kbC>T (HET)		810	12	12	12	798	798	798	0	0	100	100	100
A	40	621+1G>T/3120+1G>A (HET)		810	12	12	12	798	798	798	0	0	100	100	100
A	41	F508del/3659delC (HET)		810	12	12	12	798	798	798	0	0	100	100	100
A	42	R117H/F508del (HET)	(TG)10 (T)9/ (TG)12 (T)5	816	18	18	18	798	798	798	0	0	100	100	100

Panel	Nr próbki	Genotyp próbki	Warianty	Łączna liczba rozpoznanych nukleotydów na stanowisko	Rozpoznania ze zgodnością dodatnią (warianty)			Rozpoznania ze zgodnością ujemną (typu dzikiego)			Liczba nieprawidłowo rozpoznanych nukleotydów	Liczba nierozpoznanych nukleotydów	Zgodność dodatnia (%)	Zgodność ujemna (%)	Zgodność ogólna (%)
					Stano-wisko 1	Stano-wisko 2	Stano-wisko 3	Stano-wisko 1	Stano-wisko 2	Stano-wisko 3					
A	43	G85E/621+1G>T (HET)		810	12	12	12	798	798	798	0	0	100	100	100
A	44	A455E/F508del (HET)		810	12	12	12	798	798	798	0	0	100	100	100
A	45	N1303K (HET)		810	6	6	6	804	804	804	0	0	100	100	100
A	46	G551D/R553X (HET)		810	12	12	12	798	798	798	0	0	100	100	100
B	47	2789+5G>A (HOM)		810	6	6	6	804	804	804	0	0	100	100	100
B	48	CFTR dele2, 3/F508del (HET)		810	12	12	12	798	798	798	0	0	100	100	100
B	49	F508del/1898+1G>A (HET)		810	12	12	12	798	798	798	0	0	100	100	100
B	50	WT		810	0	0	0	810	810	810	0	0	nd.	100	100
B	51	F508del/2143delT (HET)		810	12	12	12	798	798	798	0	0	100	100	100
B	52	3876delA (HET)		810	6	6	6	804	804	804	0	0	100	100	100
B	53	3905insT (HET)		810	6	6	6	804	804	804	0	0	100	100	100
B	54	394delTT (HET)		810	6	6	6	804	804	804	0	0	100	100	100
B	55	F508del (HET)		810	6	6	6	804	804	804	0	0	100	100	100
B	56	WT		810	0	0	0	810	810	810	0	0	nd.	100	100
B	57	WT		810	0	0	0	810	810	810	0	0	nd.	100	100
B	58	F508del (HET)		810	6	6	6	804	804	804	0	0	100	100	100
B	59	WT		810	0	0	0	810	810	810	0	0	nd.	100	100
B	60	L206W (HET)		810	6	6	6	804	804	804	0	0	100	100	100
B	61	WT		810	0	0	0	810	810	810	0	0	nd.	100	100

Panel	Nr próbki	Genotyp próbki	Warianty	Łączna liczba rozpoznanych nukleotydów na stanowisko	Rozpoznania ze zgodnością dodatnią (warianty)			Rozpoznania ze zgodnością ujemną (typu dzikiego)			Liczba nieprawidłowo rozpoznanych nukleotydów	Liczba nierozpoznanych nukleotydów	Zgodność dodatnia (%)	Zgodność ujemna (%)	Zgodność ogólna (%)
					Stano-wisko 1	Stano-wisko 2	Stano-wisko 3	Stano-wisko 1	Stano-wisko 2	Stano-wisko 3					
B	62	G330X (HET)		810	6	6	6	804	804	804	0	0	100	100	100
B	63	WT		810	0	0	0	810	810	810	0	0	nd.	100	100
B	64	R347H (HET)		810	6	6	6	804	804	804	0	0	100	100	100
B	65	1078delT (HET)		810	6	6	6	804	804	804	0	0	100	100	100
B	66	G178R/F508del (HET)		810	12	12	12	798	798	798	0	0	100	100	100
B	67	S549R (c.1647T>G) (HET)		810	6	6	6	804	804	804	0	0	100	100	100
B	68	S549N (HET)		810	6	6	6	804	804	804	0	0	100	100	100
B	69	W846X (HET)		810	6	6	6	804	804	804	0	0	100	100	100
B	70	WT		810	0	0	0	810	810	810	0	0	nd.	100	100
B	71	E92X/F508del (HET)		810	12	12	12	798	798	798	0	0	100	100	100
B	72 ⁴	621+1G>T/1154insTC (HET)		810	12	12	12	798	798	797	0	1 ⁴	100	99,96	99,96
B	73	G542X (HET)		810	6	6	6	804	804	804	0	0	100	100	100
B	74	F508del (HET)		810	6	6	6	804	804	804	0	0	100	100	100
B	75 ²	F508del (HET)		810	6	5	6	804	670	804	0	135 ²	94,44	94,44	94,44
B	76	F508del (HET)		810	6	6	6	804	804	804	0	0	100	100	100
B	77	621+1G>T/A455E (HET)		810	12	12	12	798	798	798	0	0	100	100	100
B	78	1812-1 G>A (HET)		810	6	6	6	804	804	804	0	0	100	100	100
B	79	WT		810	0	0	0	810	810	810	0	0	nd.	100	100
B	80	F508del/R553X (HET)		810	12	12	12	798	798	798	0	0	100	100	100

Panel	Nr próbki	Genotyp próbki	Warianty	Łączna liczba rozpoznanych nukleotydów na stanowisko	Rozpoznania ze zgodnością dodatnią (warianty)			Rozpoznania ze zgodnością ujemną (typu dzikiego)			Liczba nieprawidłowo rozpoznanych nukleotydów	Liczba nierozpoznanych nukleotydów	Zgodność dodatnia (%)	Zgodność ujemna (%)	Zgodność ogólna (%)
					Stano-wisko 1	Stano-wisko 2	Stano-wisko 3	Stano-wisko 1	Stano-wisko 2	Stano-wisko 3					
B	81	F508del/G551D (HET)		810	12	12	12	798	798	798	0	0	100	100	100
B	82	R347P/F508del (HET)		810	12	12	12	798	798	798	0	0	100	100	100
B	83	R117H/F508del (HET)	(TG)10 (T)9/ (TG)12 (T)5	816	18	18	18	798	798	798	0	0	100	100	100
B	84	I507del (HET)		810	6	6	6	804	804	804	0	0	100	100	100
B	85	2789+5G>A (HOM)		810	6	6	6	804	804	804	0	0	100	100	100
B	86 ⁴	CFTR dele2, 3/F508del (HET)		810	12	12	12	798	797	798	0	1 ⁴	100	99,96	99,96
B	87	F508del/1898+1G>A (HET)		810	12	12	12	798	798	798	0	0	100	100	100
B	88	WT		810	0	0	0	810	810	810	0	0	nd.	100	100
B	89	F508del/2143delT (HET)		810	12	12	12	798	798	798	0	0	100	100	100
B	90	3905insT (HET)		810	6	6	6	804	804	804	0	0	100	100	100
B	91	394delTT (HET)		810	6	6	6	804	804	804	0	0	100	100	100
B	92	F508del (HET)		810	6	6	6	804	804	804	0	0	100	100	100
Łącznie				74556	2209			221182			4	273	99,77	99,88	99,88

¹ Lokacja typu dzikiego dotycząca wariantu N1303K w jednym powtórzeniu dała wynik nierozpoznanych nukleotydów ze względu na niewystarczające pokrycie.

² Jedno powtórzenie próbek 5 i 75 dało stopień rozpoznań 0%. Dalsze badania wykazały, że próbki mogły nie zostać umieszczone na płytce przed przygotowaniem biblioteki, ponieważ objętości próbek pozostałe w probówkach wskazywały, że nie usunięto żadnej objętości.

³ Badanie wykazało, że próbki 9 i 10 zostały prawdopodobnie zamienione przez operatora przed przygotowaniem biblioteki.

⁴ Lokacja typu dzikiego dotycząca wariantu M1V w jednym powtórzeniu każdej z dwóch próbek dała wynik No Call (Brak rozpoznania) ze względu na niewystarczające pokrycie.

Cechy wydajności – linia somatyczna

W opisywanych tu badaniach do analizy danych z sekwencjonowania wykorzystywano moduł do oznaczania wariantów somatycznych, z wyjątkiem badań wykorzystujących panel dwugenowy, w których wykorzystywano moduł swoisty dla oznaczenia.

Dokładność

Przeprowadzono trzy badania mające na celu ocenę dokładności wykonywanych w aparacie MiSeqDx oznaczeń DNA wyodrębnionego z próbek FFPE.

Badanie 1

W tym badaniu użyto oznaczenia reprezentatywnego w celu rozpoznania różnych genów obejmujących 12 588 nukleotydów znajdujących się w 23 różnych chromosomach przy użyciu 150 amplikonów. Chromosom Y nie zawiera regionów o wysokiej ufności i nie został poddany badaniu. W badaniu użyto pięciu unikalnych próbek pochodzących od jednej rodziny – dwojga rodziców i trojga dzieci – które wcześniej były często sekwencjonowane różnymi metodami przez wiele laboratoriów. Dostępne są trzy próbki pochodzące od osób płci żeńskiej oraz dwie próbki pochodzące od osób płci męskiej. Przed ekstrakcją DNA na potrzeby badania wszystkie próbki utrwalono w formalinie i zatopiono w parafinie. Próbkę GM12877 rozcieńczono na poziomie DNA próbką GM12878, tworząc próbkę GM12877-D zawierającą zestaw wariantów o częstościach zbliżonych do 5% i 10%. Każdą próbkę badano w dwóch powtórzeniach, z wyjątkiem próbki GM12877-D, którą badano w pięciu powtórzeniach. Dokładność została określona dla wariantów SNV, insercji i delecji przez porównanie danych z badania z dobrze scharakteryzowaną referencyjną bazą danych. Sekwencję w referencyjnej bazie danych (Platinum Genomes wersja 2016-01) uzyskano w wyniku połączenia wielu metod sekwencjonowania, publicznie dostępnych danych oraz informacji o dziedziczeniu. O ile nie podano inaczej, regiony o wysokiej zgodności zostały zdefiniowane w oparciu o ww. metodę referencyjną. Łącznie próbki poddano przebiegowi osiem razy. Tabele przedstawione w celu wykazania dokładności bazują na danych z pierwszego przebiegu.

Tabela 10 zawiera dane z badania prezentowane w postaci procentowych zgodności wyników dodatnich i ujemnych dla poszczególnych próbek, przy czym na potrzeby obliczeń zgodności PPA wyniki wariantów są porównywane z dobrze scharakteryzowanymi, złożonymi danymi referencyjnymi. Trzy typy wariantów (SNV, insercje i delecje) zostały ze sobą połączone. Ponieważ metoda referencyjna dostarcza tylko wyniki dotyczące wariantów pojedynczych nukleotydów oraz insercji/delecji, wyniki nukleotydów niewariantowych na potrzeby obliczeń zgodności NPA porównywane są z referencyjną sekwencją 19 genomu ludzkiego.

Tabela 10 Zgodność wyników rozpoznawania nukleotydów przy użyciu aparatu MiSeqDx z danymi referencyjnymi dotyczącymi 6 dobrze scharakteryzowanych próbek

Próbka	Średni odsetek rozpoznań	Całkowita liczba wariantów	Całkowita liczba wariantów TP	Całkowita liczba wariantów FN	Całkowita liczba rozpoznań TN	PPA	NPA	OPA
GM12877	98,7	152	147	0	23719	100	100	100
GM12878	98,4	270	260	0	23482	100	100	100
GM12879	98,7	192	186	0	23744	100	100	100
GM12885	99,1	244	236	0	23713	100	100	100
GM12886	98,7	230	226	0	23652	100	100	100
GM12877-D ¹		675	650	0		100	100	100
GM12877-D ²	98,4	155	155	0	57608	100	100	100

¹ Warianty z częstością większą niż 20%.

² Warianty z częstością mniejszą niż 20%.

Całkowita liczba 150 amplikonów była zaprojektowana tak, by pokryć szeroki zakres zawartości materiału genetycznego. Zawartość GC w amplikonach mieściła się w zakresie 26–87%. Amplikony charakteryzowały się również pewnym zakresem powtórzeń pojedynczych nukleotydów (np. PolyA, PolyT), dinukleotydów i trinukleotydów. W oznaczeniu wykorzystano 6 unikalnych próbek. Uzyskane dane skompilowano w oparciu o poszczególne amplikony ([tabela 11](#)) w celu oznaczenia wpływu zawartości materiału genetycznego na odsetek prawidłowych rozpoznań. Odsetek prawidłowych rozpoznań obejmuje rozpoznania wariantów i rozpoznania referencyjne i jest mniejszy niż 100% w przypadku wystąpienia nieprawidłowych rozpoznań lub braku rozpoznań. Warianty nie są rozpoznawane w przypadku niespełnienia warunków jednego lub większej liczby filtrów dla rozpoznania wariantu (np. niewystarczającego pokrycia). Nie stwierdzono nieprawidłowych rozpoznań. Liczba przypadków braku rozpoznania znacznie różniła się między poszczególnymi amplikonami. Najistotniejszymi predyktorami braku rozpoznania były zawartość GC oraz niektóre interakcje z zawartością GC. Na 2580 przypadków braku rozpoznania 2040 (79%) z nich wynikało z niespełnienia wymogów dotyczących pokrycia. Największą liczbą braków rozpoznań charakteryzowały się amplikony o zawartości GC większej niż 78%. Reprezentatywny amplikon o zawartości GC na poziomie 78% dał w sumie 675 braków rozpoznań. Reprezentatywny amplikon o zawartości GC na poziomie 87% dał w sumie 1365 braków rozpoznań. Stopień pokrycia można zwiększyć, zmniejszając liczbę próbek wprowadzanych do komory przepływowej, co umożliwi wykrywanie amplikonów o wysokiej zawartości GC.

Tabela 11 Dane dotyczące dokładności na poziomie ampliconów

Amplicon	Chromosom	Początek ampliconu	Koniec ampliconu	Rozmiar analizowanego fragmentu	Zasady w regionach wysokiej ufności	Materiał genetyczny ampliconu	Zawartość GC	Łączna liczba prawidłowo rozpoznanych nukleotydów	Łączna liczba nieprawidłowo rozpoznanych nukleotydów	Łączna liczba nierozpoznanych nukleotydów	Odsetek prawidłowo rozpoznanych nukleotydów
1	1	36450499	36450591	93	93	Indel	0,22	1395	0	0	100
2	1	109465122	109465200	79	79	PolyA (5), PolyC (5), indel	0,38	1185	0	0	100
3	1	218353867	218353957	91	91	Indel	0,4	1364	0	1	99,9
4	1	223906657	223906748	92	92	Indel	0,49	1380	0	0	100
5	1	228526602	228526682	81	81	PolyG (5)	0,69	1215	0	0	100
6	1	236372039	236372108	70	70	PolyT (10), indel	0,39	1050	0	0	100
7	1	247812041	247812128	88	88	PolyA (5), CT (3), TAA(3), indel	0,27	1320	0	0	100
8	2	55862774	55862863	90	90	Indel	0,28	1350	0	0	100
9	2	87003930	87004009	80	80	Indel	0,38	1200	0	0	100
10	2	177016721	177016805	85	81	nd.	0,65	1215	0	0	100
11	2	186625727	186625801	75	75	PolyA (8)	0,35	1117	0	10	99,1
12	2	190323504	190323591	88	88	PolyT (5)	0,42	1320	0	0	100
13	2	200796740	200796826	87	87	PolyT (5), indel	0,31	1302	0	8	99,4
14	2	212245049	212245139	91	91	PolyT (5), PolyA (6), indel	0,3	1365	0	0	100
15	2	228147052	228147144	93	93	nd.	0,43	1395	0	0	100
16	2	235016350	235016422	73	73	PolyT (5), indel	0,42	1095	0	0	100
17	3	4466229	4466321	93	93	AT(3), indel	0,27	1349	0	46	96,7
18	3	46620561	46620643	83	83	nd.	0,43	1245	0	0	100
19	3	49851331	49851400	70	70	CT(3), indel	0,49	1050	0	0	100

Amplikon	Chromosom	Początek amplikonu	Koniec amplikonu	Rozmiar analizowanego fragmentu	Zasady w regionach wysokiej ufności	Materiał genetyczny amplikonu	Zawartość GC	Łączna liczba prawidłowo rozpoznanych nukleotydów	Łączna liczba nieprawidłowo rozpoznanych nukleotydów	Łączna liczba nierozpoznanych nukleotydów	Odsetek prawidłowo rozpoznanych nukleotydów
20	3	189713161	189713248	88	88	PolyA (5), PolyT (5), PolyA (9), TG (3)	0,41	1305	0	30	97,8
21	3	190106030	190106104	75	74	Indel	0,57	1108	0	2	99,8
22	4	2233667	2233744	78	78	PolyA (6)	0,26	1170	0	0	100
23	4	7780541	7780637	97	97	PolyG (6), PolyT (5), PolyA (5)	0,42	1455	0	0	100
24	4	15688604	15688681	78	78	nd.	0,29	1169	0	1	99,9
25	4	56236521	56236586	66	62	PolyA (5), indel	0,36	930	0	0	100
26	4	102839244	102839314	71	69	PolyA (5)	0,46	1035	0	0	100
27	4	164446743	164446804	62	62	PolyA (7), indel	0,27	920	0	10	98,9
28	5	1882081	1882158	78	75	nd.	0,78	450	0	675	40,0
29	5	14769061	14769144	84	84	GT(3), CCA (3)	0,62	1260	0	0	100
30	5	41069808	41069871	64	64	nd.	0,39	960	0	0	100
31	5	74077114	74077196	83	83	PolyA (6), indel	0,3	1245	0	0	100
32	5	147475343	147475409	67	67	PolyT (5)	0,37	1005	0	0	100
33	5	149323731	149323821	91	91	CT(4), AG(3)	0,55	1365	0	0	100
34	5	155662213	155662287	75	75	Indel	0,43	1125	0	0	100
35	6	6318713	6318814	102	102	PolyG (6)	0,68	1530	0	0	100
36	6	24949983	24950074	92	92	Indel	0,63	1380	0	0	100
37	6	31084900	31084999	100	94	GCT(5), indel	0,61	1383	0	27	98,1
38	6	32147987	32148084	98	98	PolyT (5), TCT(3), CTT (3)	0,55	1455	0	15	99,0
39	6	32986864	32986958	95	95	Indel	0,53	1425	0	0	100

Amplikon	Chromosom	Początek amplitonu	Koniec amplitonu	Rozmiar analizowanego fragmentu	Zasady w regionach wysokiej ufności	Materiał genetyczny amplitonu	Zawartość GC	Łączna liczba prawidłowo rozpoznanych nukleotydów	Łączna liczba nieprawidłowo rozpoznanych nukleotydów	Łączna liczba nierozpoznanych nukleotydów	Odsetek prawidłowo rozpoznanych nukleotydów
40	6	33408498	33408583	86	86	PolyC (6)	0,7	1290	0	0	100
41	6	41647401	41647495	95	94	PolyG (5), indel	0,61	1410	0	0	100
42	6	112435865	112435955	91	91	PolyA (5)	0,44	1365	0	0	100
43	7	22202076	22202148	73	73	nd.	0,44	1095	0	0	100
44	7	66276100	66276187	88	88	Indel	0,35	1320	0	0	100
45	7	77365735	77365821	87	87	PolyA (7), AG (4)	0,26	1299	0	6	99,5
46	7	110939946	110940030	85	85	Indel	0,38	1275	0	0	100
47	7	128533468	128533557	90	90	PolyG (5), indel	0,62	1350	0	0	100
48	7	149503875	149503965	91	91	PolyG (6), PolyC (6), indel	0,71	1365	0	0	100
49	7	154404519	154404599	81	66	nd.	0,31	990	0	0	100
50	7	156476507	156476599	93	93	Indel	0,35	1395	0	0	100
51	8	1817312	1817394	83	83	nd.	0,42	1245	0	0	100
52	8	24811020	24811109	90	89	PolyG (7), CTC(4), indel	0,61	1305	0	30	97,8
53	8	76518625	76518691	67	67	Indel	0,3	1005	0	0	100
54	9	103054909	103055006	98	98	PolyG (6)	0,67	1470	0	0	100
55	9	105586150	105586214	65	65	Indel	0,32	973	0	2	99,8
56	9	107620823	107620918	96	96	nd.	0,49	1440	0	0	100
57	9	123769149	123769231	83	83	AT(3)	0,37	1242	0	3	99,8
58	9	138995345	138995441	97	97	PolyC (6), indel	0,68	1455	0	0	100
59	10	5987120	5987198	79	78	PolyG (5), indel	0,47	1170	0	0	100
60	10	11784629	11784726	98	91	GC(3)	0,87	0	0	1365	0
61	10	27317777	27317855	79	79	PolyT (5)	0,3	1185	0	0	100

Amplikon	Chromosom	Początek amplikonu	Koniec amplikonu	Rozmiar analizowanego fragmentu	Zasady w regionach wysokiej ufności	Materiał genetyczny amplikonu	Zawartość GC	Łączna liczba prawidłowo rozpoznanych nukleotydów	Łączna liczba nieprawidłowo rozpoznanych nukleotydów	Łączna liczba nierozpoznanych nukleotydów	Odsetek prawidłowo rozpoznanych nukleotydów
62	10	33018351	33018440	90	90	PolyA (5), PolyT (5)	0,2	1350	0	0	100
63	10	45084159	45084253	95	95	Indel	0,35	1425	0	0	100
64	10	55892599	55892687	89	88	AC(11), indel	0,42	1290	0	69	94,9
65	10	101611250	101611329	80	80	nd.	0,49	1200	0	0	100
66	10	118351373	118351453	81	81	nd.	0,51	1215	0	0	100
67	11	8159816	8159912	97	96	nd.	0,45	1440	0	0	100
68	11	30177648	30177717	70	70	Indel	0,46	1050	0	0	100
69	11	47470345	47470444	100	100	nd.	0,65	1500	0	0	100
70	11	59837679	59837740	62	62	Indel	0,37	930	0	0	100
71	11	64418856	64418957	102	102	nd.	0,59	1530	0	0	100
72	11	93529612	93529684	73	73	PolyA (5)	0,4	1095	0	0	100
73	11	101347052	101347136	85	85	nd.	0,42	1275	0	0	100
74	11	102477336	102477426	91	91	PolyG (6)	0,55	1365	0	0	100
75	11	118406285	118406369	85	85	Indel	0,53	1275	0	0	100
76	11	120357801	120357885	85	85	PolyA (5), CA (3), indel	0,34	1275	0	0	100
77	11	125769313	125769397	85	85	GA(3)	0,52	1275	0	0	100
78	12	2834770	2834853	84	84	PolyC (5), indel	0,52	1260	0	14	98,9
79	12	26811004	26811096	93	93	PolyA (7), AC (4)	0,33	1395	0	0	100
80	12	30881766	30881846	81	81	nd.	0,49	1215	0	0	100
81	12	88474105	88474175	71	71	PolyA (6)	0,35	1065	0	0	100
82	12	120966872	120966966	95	95	PolyG (5)	0,68	1425	0	0	100
83	12	24167504	24167576	73	73	nd.	0,52	1095	0	0	100
84	13	25816961	25817049	89	88	PolyA (5), PolyT (7), PolyA (7), indel	0,22	1305	0	15	98,9

Amplikon	Chromosom	Początek amplikonu	Koniec amplikonu	Rozmiar analizowanego fragmentu	Zasady w regionach wysokiej ufności	Materiał genetyczny amplikonu	Zawartość GC	Łączna liczba prawidłowo rozpoznanych nukleotydów	Łączna liczba nieprawidłowo rozpoznanych nukleotydów	Łączna liczba nierozpoznanych nukleotydów	Odsetek prawidłowo rozpoznanych nukleotydów
85	13	44880112	44880200	89	89	Indel	0,49	1335	0	0	100
86	13	77665218	77665294	77	77	Indel	0,39	1155	0	0	100
87	14	31619327	31619393	67	67	GA(3), TA(3)	0,39	1005	0	0	100
88	14	39517884	39517966	83	83	nd.	0,25	1245	0	0	100
89	14	46958962	46959034	73	72	PolyT (5), indel	0,19	1038	0	42	96,1
90	14	58050030	58050110	81	81	Indel	0,38	1215	0	0	100
91	14	82390559	82390649	91	91	Indel	0,35	1365	0	0	100
92	14	92549544	92549609	66	66	PolyA (5)	0,41	975	0	60	94,2
93	14	102808496	102808589	94	94	Indel	0,62	1410	0	0	100
94	15	43170751	43170848	98	96	PolyC (5)	0,45	1440	0	0	100
95	15	63446149	63446216	68	68	Indel	0,25	1020	0	0	100
96	15	77879807	77879901	95	93	PolyG (5), indel	0,68	1395	0	0	100
97	15	81625334	81625428	95	95	PolyT (6)	0,43	1425	0	0	100
98	15	85438263	85438334	72	71	Indel	0,65	1065	0	0	100
99	15	89817413	89817503	91	91	nd.	0,36	1365	0	0	100
100	15	89864274	89864343	70	70	Indel	0,56	1050	0	0	100
101	16	1894910	1894972	63	63	nd.	0,27	945	0	0	100
102	16	28997904	28997998	95	95	PolyC (5)	0,67	1425	0	0	100
103	16	3682908	53682994	87	87	TA(3)	0,41	1305	0	0	100
104	16	57954406	57954509	104	104	PolyC (5)	0,67	1560	0	0	100
105	16	85706375	85706465	91	91	PolyT (5), indel	0,37	1362	0	3	99,8
106	17	3563920	3564008	89	89	GC(3)	0,64	1335	0	0	100
107	17	3594191	3594277	87	87	PolyC (5), indel	0,67	1303	0	2	99,8
108	17	3970090	3970180	91	91	Indel	0,46	1365	0	0	100
109	17	16084945	16085037	93	93	Indel	0,26	1395	0	0	100

Amplikon	Chromosom	Początek amplikonu	Koniec amplikonu	Rozmiar analizowanego fragmentu	Zasady w regionach wysokiej ufności	Materiał genetyczny amplikonu	Zawartość GC	Łączna liczba prawidłowo rozpoznanych nukleotydów	Łączna liczba nieprawidłowo rozpoznanych nukleotydów	Łączna liczba nierozpoznanych nukleotydów	Odsetek prawidłowo rozpoznanych nukleotydów
110	17	33998759	33998849	91	89	PolyT (5)	0,54	1335	0	0	100
111	17	39589691	39589774	84	82	PolyA (13), indel (x2)	0,29	1215	0	78	94,0
112	17	41244394	41244484	91	91	PolyA (5)	0,34	1365	0	0	100
113	17	45438866	45438957	92	92	PolyA (7), AT (3), AT(4), AT (4), indel	0,26	1365	0	15	98,9
114	17	61502432	61502510	79	79	Indel	0,41	1175	0	10	99,2
115	17	64023582	64023667	86	86	PolyT (7)	0,22	1289	0	1	99,9
116	17	72308237	72308320	84	84	GAG(3)	0,62	1260	0	0	100
117	18	2616456	2616522	67	67	GA(3)	0,31	1005	0	0	100
118	18	6980478	6980568	91	91	nd.	0,37	1365	0	0	100
119	18	9888026	9888094	69	69	PolyA (6), TG (3)	0,43	1035	0	0	100
120	18	38836999	38837073	75	75	PolyA (5), indel	0,37	1121	0	19	98,3
121	18	47405382	47405462	81	81	CTC(3), indel	0,47	1215	0	0	100
122	18	54815665	54815749	85	85	CT(3), indel	0,45	1275	0	0	100
123	18	59773996	59774060	65	65	nd.	0,48	975	0	0	100
124	19	625143	625241	99	99	nd.	0,59	1478	0	7	99,5
125	19	18121418	18121491	74	74	nd.	0,68	1110	0	0	100
126	19	18186574	18186643	70	70	nd.	0,64	1050	0	0	100
127	20	746056	746149	94	94	nd.	0,61	1410	0	0	100
128	20	10633195	10633276	82	82	AC(3)	0,59	1230	0	0	100
129	20	17705633	17705708	76	76	CT(3)	0,58	1140	0	0	100
130	20	21766821	21766890	70	70	GT(3), TG(4), indel	0,46	1050	0	0	100
131	20	25278421	25278521	101	101	Indel	0,63	1515	0	0	100
132	20	50897302	50897368	67	67	Indel	0,36	1005	0	6	99,4

Amplikon	Chromosom	Początek amplikonu	Koniec amplikonu	Rozmiar analizowanego fragmentu	Zasady w regionach wysokiej ufności	Materiał genetyczny amplikonu	Zawartość GC	Łączna liczba prawidłowo rozpoznanych nukleotydów	Łączna liczba nieprawidłowo rozpoznanych nukleotydów	Łączna liczba nierozpoznanych nukleotydów	Odsetek prawidłowo rozpoznanych nukleotydów
133	20	62331904	62331994	91	88	PolyG (6)	0,73	1320	0	0	100
134	20	62690860	62690946	87	87	Indel	0,57	1305	0	0	100
135	21	30300823	30300888	66	66	Indel	0,35	990	0	0	100
136	21	33694176	33694273	98	98	PolyT (6), CA (3)	0,54	1470	0	0	100
137	21	36710706	36710792	87	87	GT(3), indel	0,39	1305	0	0	100
138	21	46644924	46644992	69	69	PolyA (6), AG (3), indel	0,32	1029	0	7	99,3
139	21	46705575	46705664	90	90	PolyT (5), PolyA (6)	0,5	1350	0	0	100
140	22	25750774	25750873	100	100	Indel	0,63	1500	0	1	99,9
141	22	32439233	32439329	97	97	nd.	0,68	1455	0	0	100
142	22	37409844	37409940	97	97	Indel	0,46	1455	0	0	100
143	22	37637596	37637694	99	99	nd.	0,6	1485	0	0	100
144	22	47081347	47081438	92	92	Indel	0,66	1380	0	0	100
145	X	15870424	15870492	69	69	PolyT (5)	0,26	1035	0	0	100
146	X	135288543	135288611	69	69	PolyC (5)	0,62	1035	0	0	100
147	X	135290777	135290847	71	71	nd.	0,52	1065	0	0	100
148	Y	2655397	2655461	65	0	nd.	0,55	0	0	0	nd.
149	Y	2655519	2655609	91	0	nd.	0,48	0	0	0	nd.
150	Y	2655609	2655679	71	0	PolyA (5)	0,37	0	0	0	nd.

Warianty, dla których stwierdzono brak rozpoznań, wymieniono w tabeli 12. Filtry, których zastosowanie skutkowało brakiem rozpoznań, zostały wymienione w tabeli.

Tabela 12 Podsumowanie wariantów, dla których stwierdzono brak rozpoznania

Nr amplikonu	Chr:Pozycja	Wariant	Zawartość odpowiedniego amplikonu	Filtr	Warianty pominięte	Warianty oczekiwane
28	5:1882129	T>G	78% GC	LowDP ¹	8	13
52	8:24811064	AG>A	PolyG (7), CTC(4), 61% GC	R3x6 ²	15	15

Nr amplikonu	Chr:Pozycja	Wariant	Zawartość odpowiedniego amplikonu	Filtr	Warianty pominięte	Warianty oczekiwane
60	10:11784633	C>T	PolyGC (3), 87% GC	LowDP	13	13
64	10:55892600	TAC>T	AC(11), 42% GC	R3x6	9	9
111	17:39589692	C>CA	PolyA (13), 29% GC	R3x6	13	13

¹ LowDP: niskie pokrycie. Wariant jest odfiltrowywany, jeśli głębokość pokrycia w co najmniej jednej z pul w danej pozycji wynosi poniżej 900.

² R3x6: filtr powtórzeń. Wariant jest odfiltrowywany, jeśli całość lub część sekwencji wariantu występuje powtarzalnie w genomie referencyjnym sąsiednio do pozycji wariantu. W oznaczeniu referencyjnym wymaganych jest co najmniej sześć powtórzeń, przy uwzględnieniu jedynie powtórzeń o długości nie większej niż 3 bp.

Wyniki sekwencjonowania próbki porównano z genotypem o wysokim poziomie ufności, ustalonym dla próbki NA12878 przez ośrodek NIST (National Institutes of Standards and Technology) (v.2.19). Na summaryczną liczbę 150 amplikonów 92 w pełni zawierało się w regionach genomu o wysokim poziomie ufności, 41 częściowo nakładało się na te regiony, zaś 17 nie nakładało się na sekwencję NIST. W rezultacie dla każdego powtórzenia uzyskano 10 000 współrzędnych do porównań. Rozpoznane nukleotydy niewariantowe zostały porównane z referencyjną sekwencją 19 genomu ludzkiego. Wyniki oznaczeń dokładności przedstawiono w tabeli 13.

Tabela 13 Zgodność wyników rozpoznań nukleotydów przy użyciu aparatu MiSeqDx dla próbki GM12878 z bazy wzorców NIST

Próbka	Liczba amplikonów	Średni odsetek rozpoznań	Całkowita liczba rozpoznań wariantów TP	Całkowita liczba rozpoznań wariantów FN	Całkowita liczba rozpoznań TN	Całkowita liczba rozpoznań FP	PPA	NPA	OPA
GM12878	150	98,43	206	0	19231	0	100	100	100

Pięć nierozcieńczonych próbek poddano dalszej analizie pod kątem rozpoznawania małych insercji i delecji (polimorfizmów typu indel) (tabela 14). W niektórych przypadkach polimorfizmy typu indel były wspólne dla dwóch lub większej liczby próbek, co znalazło odzwierciedlenie w kolumnie „Całkowita liczba powtórzeń próbki z polimorfizmem typu indel”. Wyniki dla obu powtórzeń pięciu próbek podano w tabeli 14. Obecnych było w sumie 71 polimorfizmów typu indel o wielkości od 1 do 24 bp w przypadku insercji i od 1 do 25 bp w przypadku delecji. Wykryto 68 polimorfizmów typu indel z dodatnią zgodnością procentową wynoszącą 1. Trzy insercje i delecje nie dały prawidłowych rozpoznań z uwagi na fakt, że żaden z wariantów nie został rozpoznany z powodu zastosowania filtra R3x6. Z tego powodu nie można było obliczyć wartości parametru PPA, którego definicja wyklucza brak rozpoznań. Trzy warianty obejmowały delecję 1 bp (chr8 24811064 AG>A), delecję 2 bp (chr10 55892600 TAC>T) oraz insercję 1 bp (chr17 39589692 C>CA).

Tabela 14 Podsumowanie wykrywania polimorfizmów typu indel przy użyciu aparatu MiSeqDx

Amplikon	Chromosom	Pozycja	Rozmiar analizowanego fragmentu	Rodzaj i długość polimorfizmu typu indel w amplikonie	Indel	Całkowita liczba powtórzeń próbek z polimorfizmem typu indel	Liczba nierozpoznanych nukleotydów	Całkowita liczba nieprawidłowych rozpoznań polimorfizmów typu indel	Całkowita liczba prawidłowych rozpoznań polimorfizmów typu indel	PPA
1	1	36450544	93	Delecja 25 bp	GAAAATTTAATGAAACACATTGCCT>G	7	0	0	7	100
2	1	109465165	79	Delecja 3 bp	ACTT>A	9	0	0	9	100
3	1	218353908	91	Insercja 23 bp	T>TTTTAATAGCAAAAAGAGGCTAGA	15	0	0	15	100
4	1	223906701	92	Delecja 17 bp	GACAGACTGTGAGGAAGA>G	11	0	0	11	100
6	1	236372081	70	Insercja 5 bp	C>CTTAAG	9	0	0	9	100
7	1	247812083	88	Insercja 3 bp	C>CATG	9	0	0	9	100
8	2	55862804	90	Insercja 7 bp	T>TTTGGTAA	13	0	0	13	100
9	2	87003972	80	Delecja 6 bp	TTATCTC>T	11	0	0	11	100
13	2	200796749	87	Insercja 5 bp	T>TTAAA	15	0	0	15	100
14	2	212245090	91	Insercja 12 bp	C>CTGAAAATAGGAT	11	0	0	11	100
16	2	235016388	73	Insercja 2 bp	A>ATG	9	0	0	9	100
17	3	4466274	93	Delecja 23 bp	TAACTTAAAATTACAAAATAACCC>T	13	0	0	13	100
19	3	49851375	70	Insercja 9 bp	C>CCTGGCTCCT	7	0	0	7	100
21	3	190106071	75	Delecja 1 bp	AG>A	13	0	0	13	100
25	4	56236567	66	Delecja 8 bp	TAACCGAAA>T	9	0	0	9	100
27	4	164446785	62	Insercja 11 bp	T>TTATGGTATTGA	9	0	0	9	100
31	5	74077155	83	Delecja 4 bp	TAGTA>T	7	0	0	7	100
34	5	155662255	75	Insercja 8 bp	G>GCCTACTGA	13	0	0	13	100
36	6	24950035	92	Delecja 21 bp	CCCTGGGTGCTATAGCCCACCA>C	11	0	0	11	100
37	6	31084942	100	Delecja 3 bp	GCTT>G	15	0	0	15	100
39	6	32986905	95	Delecja 25 bp	CTTTCACCTTCCCGTCTCATGCAAAG>C	7	0	0	7	100
41	6	41647442	95	Delecja 23 bp	GGCATGAGGCTTGGTGACATGGCA>G	11	0	0	11	100
44	7	66276142	88	Insercja 1 bp	C>CT	13	0	0	13	100
46	7	110939983	85	Delecja 4 bp	CAAGT>C	13	0	0	13	100
47	7	128533514	90	Insercja 1 bp	T>TC	15	0	0	15	100
48	7	149503916	91	Delecja 4 bp	GGATA>G	7	0	0	7	100

Amplikon	Chromosom	Pozycja	Rozmiar analizowanego fragmentu	Rodzaj i długość polimorfizmu typu indel w amplikonie	Indel	Całkowita liczba powtórzeń próbki z polimorfizmem typu indel	Liczba nierozpoznanych nukleotydów	Całkowita liczba nieprawidłowych rozpoznań polimorfizmów typu indel	Całkowita liczba prawidłowych rozpoznań polimorfizmów typu indel	PPA
50	7	156476548	93	Delecja 11 bp	GAATCTGCACTT>G	13	0	0	13	100
52	8	24811064	90	Delecja 1 bp	AG>A	15	15	0	0	nd.
53	8	76518677	67	Insercja 4 bp	T>TACTG	9	0	0	9	100
55	9	105586193	65	Insercja 4 bp	C>CAATT	13	0	0	13	100
58	9	138995370	97	Delecja 21 bp	TCTGGGGGGCAGCCCCTGAGGG>T	9	0	0	9	100
59	10	5987158	79	Delecja 3 bp	TAAC>T	11	0	0	11	100
63	10	45084202	95	Delecja 16 bp	AGCGTCTATAACCAAAT>A	11	0	0	11	100
64	10	55892600	89	Delecja 2 bp	TAC>T	9	9	0	0	100
68	11	30177690	70	Insercja 2 bp	C>CTG	7	0	0	7	100
70	11	59837721	62	Insercja 8 bp	T>TTATGAAAA	11	0	0	11	100
75	11	118406328	85	Delecja 8 bp	CAGTGTGGA>C	9	0	0	9	100
76	11	120357842	85	Delecja 2 bp	CTT>C	11	0	0	11	100
78	12	2834814	84	Insercja 21 bp	T>TTCTCAGTACGGTGAACCCAG	15	0	0	15	100
84	13	25817002	89	Insercja 19 bp	C>CAAAATATAAAAAGCTCCCT	15	0	0	15	100
85	13	44880152	89	Insercja 4 bp	C>CCTGT	11	0	0	11	100
86	13	77665265	77	Delecja 20 bp	ATCTATTTTCTAATAGACGGC>A	9	0	0	9	100
89	14	46958967	73	Delecja 22 bp	TTTAAATTTGAATGTGATAAAA>T	15	0	0	15	100
90	14	58050081	81	Insercja 4 bp	C>CTGAT	13	0	0	13	100
91	14	82390602	91	Delecja 16 bp	CTTGCTCTATAAACCGT>C	11	0	0	11	100
93	14	102808554	94	Delecja 5 bp	CGTGGA>C	9	0	0	9	100
95	15	63446199	68	Delecja 6 bp	CAAAATT>C	11	0	0	11	100
96	15	77879862	95	Delecja 25 bp	GCCCCTGAGCCAGCCTCCCCTCTTA>G	9	0	0	9	100
98	15	85438311	72	Insercja 3 bp	C>CTTG	9	0	0	9	100
100	15	89864316	70	Insercja 4 bp	G>GCTAC	9	0	0	9	100
105	16	85706416	91	Delecja 7 bp	ATTATTTTC>A	11	0	0	11	100
107	17	3594276	87	Delecja 1 bp	TG>T	13	0	0	13	100
108	17	3970133	91	Insercja 18 bp	A>ATCCTATTCTACTCTGAAT	11	0	0	11	100

Amplikon	Chromosom	Pozycja	Rozmiar analizowanego fragmentu	Rodzaj i długość polimorfizmu typu indel w amplikonie	Indel	Całkowita liczba powtórzeń próbki z polimorfizmem typu indel	Liczba nierozpoznanych nukleotydów	Całkowita liczba nieprawidłowych rozpoznań polimorfizmów typu indel	Całkowita liczba prawidłowych rozpoznań polimorfizmów typu indel	PPA
109	17	16084985	93	Insercja 4 bp	A>AACAC	7	0	0	7	100
111	17	39589692	84	Insercja 1 bp	C>CA	13	13	0	0	100
112	17	39589739	84	Insercja 24 bp	T>TTCTGAAGGTCAAGTCTATCCCTGA	15	0	0	15	100
113	17	45438886	92	Delecja 4 bp	CAGTG>C	7	0	0	7	100
114	17	61502459	79	Delecja 12 bp	TTTGTATCTGCTG>T	13	0	0	13	100
120	18	38837054	75	Insercja 22 bp	T>TGTATCTTAGCAAAGTTTCTCA	15	0	0	15	100
121	18	47405425	81	Insercja 3 bp	T>TGAG	11	0	0	11	100
122	18	54815706	85	Delecja 2 bp	ACT>A	13	0	0	13	100
130	20	21766863	70	Delecja 15 bp	TACTTGAGAAGTGG>T	9	0	0	9	100
131	20	25278464	101	Insercja 5 bp	A>AGTGGG	13	0	0	13	100
132	20	50897361	67	Insercja 11 bp	G>GGAATGTCAGCC	15	0	0	15	100
134	20	62690925	87	Delecja 16 bp	TCCTGGCTGGCCTGTGG>T	9	0	0	9	100
135	21	30300873	66	Insercja 11 bp	G>GATAAACTTTA	9	0	0	9	100
137	21	36710749	87	Delecja 21 bp	ACTCAAGATAACTCATGTTATC>A	9	0	0	9	100
138	21	46644985	69	Delecja 5 bp	GTTGTT>G	13	0	0	13	100
140	22	25750814	100	Insercja 6 bp	C>CAGGGCA	13	0	0	13	100
142	22	37409885	97	Insercja 5 bp	C>CTGTTT	13	0	0	13	100
144	22	47081407	92	Delecja 10 bp	GGGCACAGGCA>G	7	0	0	7	100

Badanie 2

W badaniu wykorzystano przechowywane w banku, utrwalone w formalinie i zatopione w parafinie (FFPE) próbki tkanki raka jelita grubego oraz reprezentatywne oznaczenie dwugenowe, które porównano z metodą referencyjną, tj. dwukierunkowym sekwencjonowaniem metodą Sangera (metoda Sangera). Wyniki otrzymane metodą Sangera i reprezentatywnego oznaczenia dwugenowego potwierdzono u 441 z łącznej liczby 1183 uczestników. W ocenie na poziomie poszczególnych uczestników (tabela 15) u 230 spośród 441 uczestników uzyskano dodatni wynik metodą Sangera (mutacja wykryta metodą Sangera). U 227 uczestników z tej liczby uzyskano dodatni wynik oznaczenia reprezentatywnego. U pozostałych 211 spośród 441 uczestników uzyskano ujemny wynik

metodą Sangera. U 206 uczestników z tej liczby uzyskano ujemny wynik oznaczenia reprezentatywnego. Oznacza to, że uzyskano procentową zgodność wyników dodatnich (PPA) na poziomie 98,7% i procentową zgodność wyników ujemnych (NPA) na poziomie 97,6% (tabela 15).

Tabela 15 Procentowa zgodność wyników dodatnich i ujemnych na poziomie poszczególnych uczestników

Oznaczenie reprezentatywne	Metoda Sangera		Łącznie
	Dodatnie	Ujemne	
Dodatnie	227 ¹	5	232
Ujemne	3 ²	206	209
Łącznie	230	211	441

Podsumowanie wyników		
Statystyki zgodności	Estymator punktowy	Dokładny 95% CI
PPA	227/230 = 98,7%	[96,2%, 99,7%]
NPA	206/211 = 97,6%	[94,6%, 99,2%]

¹ Stwierdzono 224 dokładne dopasowania dla wyników na poziomie wszystkich mutacji u poszczególnych uczestników. U dwóch uczestników w oznaczeniu MiSeqDx wykryto mutację wykrytą metodą Sangera i jedną dodatkową mutację. U jednego uczestnika w oznaczeniu MiSeqDx i oznaczeniu metodą Sangera wykryto różne mutacje.

² U jednego uczestnika wykryto dwie mutacje metodą Sangera. U dwóch uczestników wykryto jedną mutację metodą Sangera.

Badanie 3

W badaniu oceniano biblioteki DNA przygotowane z preparatów FFPE o różnych rodzajach tkanek. Łącznie wykorzystano 109 preparatów FFPE z ośmiu różnych tkanek (jelito grube, jajnik, trzustka, nadnercze, pęcherz, wątroba, tarczyca i piers) z co najmniej 11 preparatami FFPE dla każdego rodzaju tkanki. W tkance nadnercza obecne były przerzuty z nowotworów przelyku, płuca i jelita grubego. Pozostałe tkanki zawierały guzy pierwotne. W tym badaniu użyto oznaczenia reprezentatywnego w celu rozpoznania 26 genów obejmujących 21 577 nukleotydów znajdujących się w 17 różnych chromosomach. Metodą Sangera zsekwencjonowano sześć różnych genów (*KRAS*, *NRAS*, *TP53*, *PIK3CA*, *EGFR* i *BRAF*), przy czym każdy z guzów zawierał od 1 do 3 genów zsekwencjonowanych metodą Sangera w oparciu o spodziewane występowanie właściwych dla niego mutacji somatycznych. Za pomocą sekwencjonowania metodą Sangera zidentyfikowano 39 mutacji somatycznych typu SNV w 33 spośród 109 preparatów FFPE. Aparat MiSeqDx zidentyfikował 36 mutacji somatycznych typu SNV w 32 spośród 109 preparatów FFPE, przy jednym wyniku fałszywie ujemnym i braku rozpoznania dwóch pozycji wariantowych. Parametr PPA wyniósł 97,3%. Aparat MiSeqDx zidentyfikował 78 975 referencyjnych nukleotydów w 109 preparatach FFPE, przy 29 wynikach fałszywie dodatnich w stosunku do rezultatów sekwencjonowania metodą Sangera i 2437 brakach rozpoznań. Parametr NPA wyniósł 99,9%. Dla obu metod stwierdzono zgodność w zakresie delekcji dwóch nukleotydów. [Tabela 16](#) stanowi podsumowanie wyników w zależności od rodzaju tkanki.

Tabela 16 Procentowa zgodność wyników dodatnich i ujemnych w zależności od rodzaju tkanki

Rodzaj tkanki	Liczba próbek	Całkowita liczba wariantów	Całkowita liczba wariantów TP	Całkowita liczba wariantów FN	Całkowita liczba rozpoznań TN	Całkowita liczba rozpoznań FP	Łączna liczba nierozpoznanych nukleotydów	PPA	NPA
Nadnercze	16	6	4	1	11823	2	607	80	>99,9
Pęcherz	12	4	4	0	7070	3	273	100	>99,9
Piers	16	3	3	0	13439	7	479	100	99,9
Jelito grube	11	6	5	0	8720	2	133	100	>99,9
Wątroba	13	3	3	0	7984	1	59	100	>99,9
Jajnik	13	7	7	0	10581	1	724	100	>99,99
Trzustka	17	7	7	0	11929	12	489	100	99,9
Tarczyca	11	3	3	0	7429	1	652	100	>99,9
Łącznie	109	39	36	1	78974	29	3416	97,3	>99,9

Odtwarzalność

Przeprowadzono dwa badania mające na celu ocenę odtwarzalności wykonywanych w aparacie MiSeqDx oznaczeń DNA wyodrębnionego z próbek FFPE. W badaniu 1 wykorzystywano wiele aparatów. W badaniu 2 wykorzystywano wiele stanowisk.

Badanie 1

Odtwarzalność oznaczeń przy użyciu aparatu MiSeqDx wyznaczano przy użyciu dwóch aparatów i dwóch wyszkolonych operatorów w całkowitej liczbie ośmiu przebiegów. Oznaczenie reprezentatywne, kontekst genomowy ampliconów, próbki i metoda referencyjna były takie same, jak opisane powyżej dla badania dokładności 1. Wyniki przedstawiono w podziale na poszczególne amplicony dla każdego z aparatów (tabela 17) tak, by wykazać odtwarzalność rozpoznań na różnych aparatach. Odsetek prawidłowych rozpoznań wyznaczano po uwzględnieniu zarówno nieprawidłowych rozpoznań, jak i braku rozpoznań (niespełnienia jednego lub większej liczby filtrów rozpoznawania wariantów). Oba aparaty charakteryzowały się podobną liczbą braków rozpoznań dla poszczególnych ampliconów. W przypadku aparatu MiSeqDx 1 stwierdzono pojedyncze nieprawidłowe rozpoznanie w regionie o wysokiej ufności określone w oparciu o wzorzec Platinum Genomes. Nieprawidłowe rozpoznanie było fałszywie dodatnim rozpoznanie wariantu insercyjnego w ampliconie 64 próbującym chromosom 10 w pozycjach od 55892599 do 55892687. Amplicon zawierał 11 kolejnych powtórzeń jednego dinukleotydu.

Tabela 17 Badanie – wyniki odtwarzalności badań między aparatami MiSeqDx (na poziomie ampliconu)

Amplicon	Chromosom	Rozmiar analizowanego fragmentu	Zasady w regionach wysokiej ufności	Materiał genetyczny ampliconu	Zawartość GC	MiSeqDx 1			MiSeqDx 2		
						Łączna liczba prawidłowo rozpoznanych nukleotydów	Łączna liczba nieprawidłowo rozpoznanych nukleotydów	Łączna liczba nierozpoznanych nukleotydów	Łączna liczba prawidłowo rozpoznanych nukleotydów	Łączna liczba nieprawidłowo rozpoznanych nukleotydów	Łączna liczba nierozpoznanych nukleotydów
1	1	93	93	Indel	0,22	5580	0	0	5580	0	0
2	1	79	79	PolyA (5), PolyC (5), indel	0,38	4740	0	0	4740	0	0
3	1	91	91	Indel	0,4	5448	0	12	5453	0	8
4	1	92	92	Indel	0,49	5518	0	2	5518	0	2
5	1	81	81	PolyG (5)	0,69	4858	0	2	4860	0	0

Amplikon	Chromosom	Rozmiar analizowanego fragmentu	Zasady w regionach wysokiej ufności	Materiał genetyczny amplikonu	Zawartość GC	MiSeqDx 1			MiSeqDx 2		
						Łączna liczba prawidłowo rozpoznanych nukleotydów	Łączna liczba nieprawidłowo rozpoznanych nukleotydów	Łączna liczba nierozpoznanych nukleotydów	Łączna liczba prawidłowo rozpoznanych nukleotydów	Łączna liczba nieprawidłowo rozpoznanych nukleotydów	Łączna liczba nierozpoznanych nukleotydów
6	1	70	70	PolyT (10), indel	0,39	4200	0	0	4200	0	0
7	1	88	88	PolyA (5), CT(3), TAA (3), indel	0,27	5279	0	1	5279	0	1
8	2	90	90	Indel	0,28	5400	0	0	5400	0	0
9	2	80	80	Indel	0,38	4800	0	0	4800	0	0
10	2	85	81	nd.	0,65	4859	0	1	4859	0	1
11	2	75	75	PolyA (8)	0,35	4468	0	40	4468	0	40
12	2	88	88	PolyT (5)	0,42	5280	0	0	5280	0	0
13	2	87	87	PolyT (5), indel	0,31	5211	0	43	5214	0	40
14	2	91	91	PolyT (5), PolyA (6), indel	0,3	5453	0	7	5449	0	11
15	2	93	93	nd.	0,43	5579	0	1	5579	0	1
16	2	73	73	PolyT (5), indel	0,42	4378	0	2	4379	0	1
17	3	93	93	AT(3), indel	0,27	5396	0	184	5396	0	184
18	3	83	83	nd.	0,43	4980	0	0	4980	0	0
19	3	70	70	CT(3), indel	0,49	4193	0	7	4194	0	6
20	3	88	88	PolyA (5), PolyT (5), PolyA (9), TG(3)	0,41	5220	0	120	5220	0	120
21	3	75	74	Indel	0,57	4432	0	8	4432	0	8
22	4	78	78	PolyA (6)	0,26	4676	0	4	4676	0	4

Amplikon	Chromosom	Rozmiar analizowanego fragmentu	Zasady w regionach wysokiej ufności	Materiał genetyczny amplitonu	Zawartość GC	MiSeqDx 1			MiSeqDx 2		
						Łączna liczba prawidłowo rozpoznanych nukleotydów	Łączna liczba nieprawidłowo rozpoznanych nukleotydów	Łączna liczba nierozpoznanych nukleotydów	Łączna liczba prawidłowo rozpoznanych nukleotydów	Łączna liczba nieprawidłowo rozpoznanych nukleotydów	Łączna liczba nierozpoznanych nukleotydów
23	4	97	97	PolyG (6), PolyT (5), PolyA (5)	0,42	5820	0	0	5820	0	0
24	4	78	78	nd.	0,29	4679	0	1	4677	0	3
25	4	66	62	PolyA (5), indel	0,36	3720	0	0	3720	0	0
26	4	71	69	PolyA (5)	0,46	4140	0	0	4140	0	0
27	4	62	62	PolyA (7), indel	0,27	3676	0	45	3671	0	51
28	5	78	75	nd.	0,78	3368	0	1132	3485	0	1015
29	5	84	84	GT(3), CCA (3)	0,62	5040	0	0	5040	0	0
30	5	64	64	nd.	0,39	3840	0	0	3840	0	0
31	5	83	83	PolyA (6), indel	0,3	4979	0	1	4980	0	0
32	5	67	67	PolyT (5)	0,37	4020	0	0	4020	0	0
33	5	91	91	CT(4), AG (3)	0,55	5460	0	0	5460	0	0
34	5	75	75	Indel	0,43	4498	0	6	4500	0	1
35	6	102	102	PolyG (6)	0,68	6120	0	0	6120	0	0
36	6	92	92	Indel	0,63	5520	0	0	5520	0	0
37	6	100	94	GCT(5), indel	0,61	5532	0	108	5532	0	108
38	6	98	98	PolyT (5), TCT(3), CTT(3)	0,55	5820	0	60	5820	0	60
39	6	95	95	Indel	0,53	5697	0	3	5698	0	2

Amplikon	Chromosom	Rozmiar analizowanego fragmentu	Zasady w regionach wysokiej ufności	Materiał genetyczny amplikonu	Zawartość GC	MiSeqDx 1			MiSeqDx 2		
						Łączna liczba prawidłowo rozpoznanych nukleotydów	Łączna liczba nieprawidłowo rozpoznanych nukleotydów	Łączna liczba nierozpoznanych nukleotydów	Łączna liczba prawidłowo rozpoznanych nukleotydów	Łączna liczba nieprawidłowo rozpoznanych nukleotydów	Łączna liczba nierozpoznanych nukleotydów
40	6	86	86	PolyC (6)	0,7	5159	0	1	5160	0	0
41	6	95	94	PolyG (5), indel	0,61	5638	0	2	5638	0	2
42	6	91	91	PolyA (5)	0,44	5460	0	0	5460	0	0
43	7	73	73	nd.	0,44	4380	0	0	4380	0	0
44	7	88	88	Indel	0,35	5279	0	1	5276	0	4
45	7	87	87	PolyA (7), AG(4)	0,26	5184	0	36	5181	0	39
46	7	85	85	Indel	0,38	5100	0	0	5100	0	0
47	7	90	90	PolyG (5), indel	0,62	5398	0	2	5399	0	1
48	7	91	91	PolyG (6), PolyC (6), indel	0,71	5460	0	0	5459	0	1
49	7	81	66	nd.	0,31	3960	0	0	3960	0	0
50	7	93	93	Indel	0,35	5580	0	0	5579	0	1
51	8	83	83	nd.	0,42	4980	0	0	4980	0	0
52	8	90	89	PolyG (7), CTC(4), indel	0,61	5219	0	121	5220	0	120
53	8	67	67	Indel	0,3	4020	0	0	4020	0	0
54	9	98	98	PolyG (6)	0,67	5879	0	1	5880	0	0
55	9	65	65	Indel	0,32	3894	0	6	3895	0	5
56	9	96	96	nd.	0,49	5760	0	0	5760	0	0
57	9	83	83	AT(3)	0,37	4973	0	7	4978	0	2
58	9	97	97	PolyC (6), indel	0,68	5817	0	3	5818	0	2

Amplikon	Chromosom	Rozmiar analizowanego fragmentu	Zasady w regionach wysokiej ufności	Materiał genetyczny amplikonu	Zawartość GC	MiSeqDx 1			MiSeqDx 2		
						Łączna liczba prawidłowo rozpoznanych nukleotydów	Łączna liczba nieprawidłowo rozpoznanych nukleotydów	Łączna liczba nierozpoznanych nukleotydów	Łączna liczba prawidłowo rozpoznanych nukleotydów	Łączna liczba nieprawidłowo rozpoznanych nukleotydów	Łączna liczba nierozpoznanych nukleotydów
59	10	79	78	PolyG (5), indel	0,47	4679	0	1	4680	0	0
60	10	98	91	GC(3)	0,87	450	0	5010	632	0	4828
61	10	79	79	PolyT (5)	0,3	4740	0	0	4740	0	0
62	10	90	90	PolyA (5), PolyT (5)	0,2	5400	0	0	5400	0	0
63	10	95	95	Indel	0,35	5699	0	1	5699	0	1
64	10	89	88	AC(11), indel	0,42	5157	0	276	5153	2	273
65	10	80	80	nd.	0,49	4800	0	0	4800	0	0
66	10	81	81	nd.	0,51	4860	0	0	4860	0	0
67	11	97	96	nd.	0,45	5760	0	0	5760	0	0
68	11	70	70	Indel	0,46	4199	0	2	4200	0	1
69	11	100	100	nd.	0,65	5999	0	1	5998	0	2
70	11	62	62	Indel	0,37	3720	0	0	3720	0	0
71	11	102	102	nd.	0,59	6120	0	0	6118	0	2
72	11	73	73	PolyA (5)	0,4	4380	0	0	4380	0	0
73	11	85	85	nd.	0,42	5100	0	0	5100	0	0
74	11	91	91	PolyG (6)	0,55	5437	0	23	5441	0	19
75	11	85	85	Indel	0,53	5100	0	0	5100	0	0
76	11	85	85	PolyA (5), CA(3), indel	0,34	5100	0	0	5100	0	0
77	11	85	85	GA(3)	0,52	5100	0	0	5100	0	0
78	12	84	84	PolyC (5), indel	0,52	5040	0	60	5038	0	63

Amplikon	Chromosom	Rozmiar analizowanego fragmentu	Zasady w regionach wysokiej ufności	Materiał genetyczny amplikonu	Zawartość GC	MiSeqDx 1			MiSeqDx 2		
						Łączna liczba prawidłowo rozpoznanych nukleotydów	Łączna liczba nieprawidłowo rozpoznanych nukleotydów	Łączna liczba nierozpoznanych nukleotydów	Łączna liczba prawidłowo rozpoznanych nukleotydów	Łączna liczba nieprawidłowo rozpoznanych nukleotydów	Łączna liczba nierozpoznanych nukleotydów
79	12	93	93	PolyA (7), AC(4)	0,33	5577	0	3	5573	0	7
80	12	81	81	nd.	0,49	4860	0	0	4860	0	0
81	12	71	71	PolyA (6)	0,35	4260	0	0	4260	0	0
82	2	95	95	PolyG (5)	0,68	5605	0	95	5605	0	95
83	13	73	73	nd.	0,52	4380	0	0	4379	0	1
84	13	89	88	PolyA (5), PolyT (7), PolyA (7), indel	0,22	5220	0	60	5220	0	60
85	13	89	89	Indel	0,49	5340	0	0	5340	0	0
86	13	77	77	Indel	0,39	4620	0	0	4620	0	0
87	14	67	67	GA(3), TA (3)	0,39	4020	0	0	4020	0	0
88	14	83	83	nd.	0,25	4980	0	0	4980	0	0
89	14	73	72	PolyT (5), indel	0,19	4173	0	147	4173	0	147
90	14	81	81	Indel	0,38	4860	0	2	4860	0	0
91	14	91	91	Indel	0,35	5459	0	1	5460	0	0
92	14	66	66	PolyA (5)	0,41	3900	0	240	3900	0	240
93	14	94	94	Indel	0,62	5637	0	3	5637	0	3
94	15	98	96	PolyC (5)	0,45	5760	0	0	5760	0	0
95	15	68	68	Indel	0,25	4079	0	1	4078	0	2
96	15	95	93	PolyG (5), indel	0,68	5475	0	105	5487	0	93
97	15	95	95	PolyT (6)	0,43	5699	0	1	5700	0	0
98	15	72	71	Indel	0,65	4260	0	0	4260	0	0

Amplikon	Chromosom	Rozmiar analizowanego fragmentu	Zasady w regionach wysokiej ufności	Materiał genetyczny amplikonu	Zawartość GC	MiSeqDx 1			MiSeqDx 2		
						Łączna liczba prawidłowo rozpoznanych nukleotydów	Łączna liczba nieprawidłowo rozpoznanych nukleotydów	Łączna liczba nierozpoznanych nukleotydów	Łączna liczba prawidłowo rozpoznanych nukleotydów	Łączna liczba nieprawidłowo rozpoznanych nukleotydów	Łączna liczba nierozpoznanych nukleotydów
99	15	91	91	nd.	0,36	5460	0	0	5460	0	0
100	15	70	70	Indel	0,56	4200	0	0	4200	0	0
101	16	63	63	nd.	0,27	3780	0	0	780	0	0
102	16	95	95	PolyC (5)	0,67	5700	0	0	5700	0	0
103	16	87	87	TA(3)	0,41	5220	0	0	5220	0	0
104	16	104	104	PolyC (5)	0,67	6238	0	3	6238	0	3
105	16	91	91	PolyT (5), indel	0,37	5443	0	17	5444	0	16
106	17	89	89	GC(3)	0,64	5251	0	89	5339	0	1
107	17	87	87	PolyC (5), indel	0,67	5212	0	8	5212	0	8
108	17	91	91	Indel	0,46	5459	0	1	5459	0	1
109	17	93	93	Indel	0,26	5580	0	0	5580	0	0
110	17	91	89	PolyT (5)	0,54	5340	0	0	5340	0	0
111	17	84	82	PolyA (13), indel (x2)	0,29	4860	0	308	4860	0	07
112	17	91	91	PolyA (5)	0,34	5459	0	1	5459	0	1
113	17	92	92	PolyA (7), AT(3), AT(4), AT(4), indel	0,26	5460	0	60	5460	0	60
114	17	79	79	Indel	0,41	4699	0	41	4700	0	40
115	17	86	86	PolyT (7)	0,22	5153	0	7	5156	0	4
116	17	84	84	GAG(3)	0,62	5039	0	1	5039	0	1
117	18	67	67	GA(3)	0,31	4020	0	0	4020	0	0
118	18	91	91	nd.	0,37	5460	0	0	5460	0	0

Amplikon	Chromosom	Rozmiar analizowanego fragmentu	Zasady w regionach wysokiej ufności	Materiał genetyczny amplitonu	Zawartość GC	MiSeqDx 1			MiSeqDx 2		
						Łączna liczba prawidłowo rozpoznanych nukleotydów	Łączna liczba nieprawidłowo rozpoznanych nukleotydów	Łączna liczba nierozpoznanych nukleotydów	Łączna liczba prawidłowo rozpoznanych nukleotydów	Łączna liczba nieprawidłowo rozpoznanych nukleotydów	Łączna liczba nierozpoznanych nukleotydów
119	18	69	69	PolyA (6), TG(3)	0,43	4132	0	8	4131	0	9
120	18	75	75	PolyA (5), indel	0,37	4475	0	85	4480	0	79
121	18	81	81	CTC(3), indel	0,47	4860	0	0	4860	0	0
122	18	85	85	CT(3), indel	0,45	5098	0	2	5098	0	2
123	18	65	65	nd.	0,48	3900	0	0	3900	0	0
124	19	99	99	nd.	0,59	5926	0	14	5924	0	16
125	19	74	74	nd.	0,68	4440	0	0	4438	0	2
126	19	70	70	nd.	0,64	4199	0	1	4200	0	0
127	20	94	94	nd.	0,61	5640	0	1	5638	0	3
128	20	82	82	AC(3)	0,59	4920	0	0	4920	0	0
129	20	76	76	CT(3)	0,58	4559	0	1	4558	0	2
130	20	70	70	GT(3), TG(4), indel	0,46	4200	0	0	4200	0	0
131	20	101	101	Indel	0,63	6060	0	0	6060	0	0
132	20	67	67	Indel	0,36	4020	0	31	4020	0	25
133	20	91	88	PolyG (6)	0,73	5277	0	3	5274	0	6
134	20	87	87	Indel	0,57	5218	0	2	5218	0	2
135	21	66	66	Indel	0,35	3959	0	1	3957	0	3
136	21	98	98	PolyT (6), CA(3)	0,54	5880	0	0	5880	0	0
137	21	87	87	GT(3), indel	0,39	5220	0	0	5220	0	0
138	21	69	69	PolyA (6), AG(3), indel	0,32	4119	0	31	4113	0	37

Amplikon	Chromosom	Rozmiar analizowanego fragmentu	Zasady w regionach wysokiej ufności	Materiał genetyczny amplikonu	Zawartość GC	MiSeqDx 1			MiSeqDx 2		
						Łączna liczba prawidłowo rozpoznanych nukleotydów	Łączna liczba nieprawidłowo rozpoznanych nukleotydów	Łączna liczba nierozpoznanych nukleotydów	Łączna liczba prawidłowo rozpoznanych nukleotydów	Łączna liczba nieprawidłowo rozpoznanych nukleotydów	Łączna liczba nierozpoznanych nukleotydów
139	21	90	90	PolyT (5), PolyA (6)	0,5	5399	0	1	5399	0	1
140	22	100	100	Indel	0,63	5998	0	7	5997	0	5
141	22	97	97	nd.	0,68	5819	0	1	5819	0	1
142	22	97	97	Indel	0,46	5818	0	2	5816	0	4
143	22	99	99	nd.	0,6	5940	0	0	5940	0	0
144	22	92	92	Indel	0,66	5519	0	1	5519	0	1
145	X	69	69	PolyT (5)	0,26	4139	0	1	4140	0	0
146	X	69	69	PolyC (5)	0,62	4136	0	4	4137	0	3
147	X	71	71	nd.	0,52	4260	0	0	4260	0	0
148	Y	65	0	nd.	0,55	0	0	0	0	0	0
149	Y	91	0	nd.	0,48	0	0	0	0	0	0
150	Y	71	0	PolyA (5)	0,37	0	0	0	0	0	0

Wyniki badania odtwarzalności analizowano dla poszczególnych operatorów w oparciu o częstość wariantów (tabela 18). Analiza wykazała spójność częstości wariantów wykrywanych przez operatorów. Przedstawiono wartości średniej częstości wariantu +/- 1 odchylenie standardowe.

Tabela 18 Wyniki oznaczeń przy użyciu aparatu MiSeqDx dokonywanych przez poszczególnych operatorów

Zakres częstości wariantu	Liczba unikalnych wariantów	Całkowita liczba wariantów analizowanych przez operatora 1	Całkowita liczba wariantów analizowanych przez operatora 2	Średnia (SD) częstość wariantu raportowana przez operatora 1	Średnia (SD) częstość wariantu raportowana przez operatora 2
Wysoka częstość (~100%)	1112	1072	1072	0,96 +/- 0,05	0,96 +/- 0,05
Średnia częstość (~50%)	3240	3151	3161	0,49 +/- 0,04	0,49 +/- 0,04
Niska częstość (3–7%)	620	618	612	0,05 +/- 0,01	0,05 +/- 0,01

Wyniki badania odtwarzalności dla poszczególnych próbek zawierają dane ze wszystkich ośmiu przebiegów (tabela 19). Wykrywalność ocenia się oddzielnie dla każdego rodzaju wariantu – SNV, insercji i delecji. Pozycje referencyjne wyklucza się. Analiza wykazała, że uzyskane wyniki dotyczące wariantów były powtarzalne we wszystkich próbkach.

Tabela 19 Zgodność wyników rozpoznawania nukleotydów przez aparat MiSeqDx w zależności od próbki

Próbka	Warianty SNV				Insercje				Delecje			
	Łączna liczba	Łączna liczba TP	Łączna liczba FP	Łączna liczba FN	Łączna liczba	Łączna liczba TP	Łączna liczba FP	Łączna liczba FN	Łączna liczba	Łączna liczba TP	Łączna liczba FP	Łączna liczba FN
GM12877	592	574	2	0	336	336	0	0	228	272	0	0
GM12878	1456	1432	0	0	320	304	0	0	384	352	0	0
GM12879	912	896	0	0	336	320	0	0	288	272	0	0
GM12885	1200	1192	0	0	400	384	0	0	352	320	0	0
GM12886	1104	1104	0	0	368	352	0	0	368	352	0	0
GM12877-D1 ¹	3640	3582	0	0	800	760	0	0	960	880	0	0
GM12877-D2 ²	400	398	0	0	520	516	0	0	560	556	0	0

¹ Warianty z częstością większą niż 20%.

² Warianty z częstością mniejszą niż 20%.

Dane uzyskane w oparciu o 8 przebiegów w przedstawionym badaniu odtwarzalności potwierdzają twierdzenie, że aparat MiSeqDx jest w stanie w sposób spójny sekwencjonować:

- Zawartość GC \geq 19% (wszystkie rozpoznane nukleotydy w 120 ze 120 sekwencjonowanych amplikonów z 19% zawartością GC rozpoznane prawidłowo z odsetkiem nierozpoznanych nukleotydów na poziomie 3,4%).
- Zawartość GC \leq 73% (wszystkie rozpoznane nukleotydy w 120 ze 120 sekwencjonowanych amplikonów z 73% zawartością GC rozpoznane prawidłowo z odsetkiem nierozpoznanych nukleotydów na poziomie 0,1%).
- Łańcuchy PolyA o długości \leq 8 (powtórzenie PolyA złożone z 8 nukleotydów rozpoznane prawidłowo w 120 ze 120 sekwencjonowanych amplikonów zawierających PolyA = 8).
- Łańcuchy PolyT o długości \leq 10 (powtórzenie PolyT złożone z 10 nukleotydów rozpoznane prawidłowo w 120 ze 120 sekwencjonowanych amplikonów zawierających PolyT = 10).

- Łańcuchy PolyG o długości ≤ 6 (powtórzenie PolyG złożone z 6 nukleotydów rozpoznane prawidłowo w 720 z 720 sekwencjonowanych ampliconów zawierających PolyG = 6).
- Łańcuchy PolyC o długości ≤ 6 (powtórzenie PolyC złożone z 6 nukleotydów rozpoznane prawidłowo w 359 z 360 sekwencjonowanych ampliconów zawierających PolyC = 6, 1 brak rozpoznania).
- Powtórzenia dinukleotydy o długości $\leq 4x$ (wszystkie zidentyfikowane nukleotydy z 4-krotnym powtórzeniem dinukleotydu w 600 z 600 sekwencjonowanych ampliconów zostały rozpoznane prawidłowo przy odsetku braku rozpoznań na poziomie 0,4%).
- Powtórzenia trinukleotydy o długości $\leq 5x$ (wszystkie zidentyfikowane nukleotydy z 5-krotnym powtórzeniem trinukleotydu w 120 ze 120 sekwencjonowanych ampliconów zostały rozpoznane prawidłowo przy odsetku braku rozpoznań na poziomie 1,9%).
- Insercje 24 lub mniej nukleotydów i delecje 25 lub mniej nukleotydów:
 - insercje 24 nukleotydów zidentyfikowane prawidłowo w 120 ze 120 próbek,
 - delecje 25 nukleotydów zidentyfikowane poprawnie w 182 próbkach przy 2 próbkach na 184 próbki oznaczonych jako brak rozpoznania.

Badanie 2

Wykonano zewnętrzne badanie mające na celu ocenę odtwarzalności reprezentatywnego oznaczenia dwugenowego, opisane jako badanie dokładności 2, które przeprowadzono w trzech zewnętrznych ośrodkach badawczych (dwóch operatorów w każdym ośrodku), używając jednej serii odczytników, w trzech nienastępujących po sobie dniach. Badanie prowadzono, wykorzystując sześć dobrze scharakteryzowanych paneli próbek genomowego DNA z utrwalonych w formalinie i zatopionych w parafinie preparatów klinicznych lub linii komórkowych. Każdy panel składał się z 10 elementów, co w sumie dało 60 elementów we wszystkich panelach.

Te 60 elementów paneli obejmowało duplikaty czterech próbek z unikalnym typem dzikim (dla mutacji panelowych), 12 próbek z unikalną mutacją (o pojedynczej mutacji) przygotowanych w wersjach z wysoką i niską częstością mutacji oraz dwóch próbek z unikalną mutacją (o pojedynczej mutacji) przygotowanych wyłącznie w wersji o niskiej częstości mutacji. Każdy unikalny preparat (próbka) o danej częstości mutacji (w każdej serii oznaczany w duplikatach) mógł dostarczyć 36 wyników (2 powtórzenia \times 2 operatorów \times 3 dni \times 3 ośrodki), pod warunkiem że wszystkie oznaczenia były ważne.

Dokonano oceny oczekiwanego odsetka rozpoznanych nukleotydów (PEC, ang. Percent Expected Call) dla wszystkich wariantów dodatnich i ujemnych, porównując wynik oznaczenia reprezentatywnego z oczekiwanym wynikiem mutacji (oczekiwanym wykryciem lub niewykryciem mutacji) w każdej próbce. Wartość PEC oblicza się jako 100% pomnożone przez liczbę oczekiwanych rozpoznań podzieloną przez liczbę podjętych prób rozpoznania. Dwustronny 95-procentowy przedział ufności oblicza się metodą punktacji Wilsona.

Dla połączonych ośrodków odsetki powodzeń wyniosły $\geq 94,7\%$ dla pierwszego przebiegu próbki, tj. próbek testowanych w przebiegach uznanych za ważne za pierwszym razem. Wartość PEC na poziomie mutacji dla wszystkich zmutowanych próbek wyniosła 99,6% (905/909) (95% CI; 98,9, 99,8). Liczba prób rozpoznania podjętych dla wszystkich 56 mutacji (niezależnie od tego, czy wykryta mutacja była oczekiwana czy też nie) we wszystkich ważnych próbkach wyniosła 58 856 (56×1051). Wśród wymienionych wyżej 58 856 obserwacji na poziomie mutacji stwierdzono tylko sześć przypadków niezgodności między wynikiem obserwowanym a oczekiwanym. Wartość PEC na poziomie mutacji dla wszystkich dodatnich i ujemnych wariantów wszystkich elementów paneli – zarówno z mutacją, jak i typu dzikiego – wyniosła 99,99% (58 850/58 856).

Czułość analityczna (granica próby ślepej (LoB) i granica wykrywalności (LoD))

W tym badaniu zweryfikowano próg odcięcia oznaczenia dla urządzenia MiSeqDx z zastosowaniem reprezentatywnego panelu i wyznaczono granicę wykrywalności (LoD, ang. Limit of Detection). Krótko mówiąc, dobrze scharakteryzowane linie komórkowe Platinum Genome GM12878 i GM12877 utrwalono w formalinie i zatopiono w parafinie, po czym dokonano ekstrakcji DNA. GM12878 rozcieńczono GM12877 w taki sposób, by częstość występowania wariantu dla 70 wariantów (52 SNV, 9 insercji i 9 delecji) była bliska 0,05. Dwie próbki DNA były badane przez dwóch operatorów przy użyciu dwóch aparatów i dwóch serii odczytników; w sumie przy użyciu aparatu MiSeqDx wykonano 10 przebiegów sekwencjonowania. W rezultacie przy użyciu każdej serii odczytników uzyskano 40 powtórzeń dla każdego wariantu GM12878 i 60 powtórzeń dla każdej odpowiadającej im współrzędnej typu dzikiego w GM12877. Wartości granicy próby ślepej (LoB) i granicy wykrywalności (LoD) obliczono przy użyciu klasycznego podejścia przedstawionego w CLSI EP17-A2 z wykorzystaniem opcji nieparametrycznej. Wartości LoB i LoD obliczono oddzielnie dla SNV, insercji i delecji, sumując częstości występowania wariantów dla danego rodzaju. Prawdopodobieństwo błędu pierwszego rodzaju wyznaczono na 0,01, zaś prawdopodobieństwo błędu drugiego rodzaju wyznaczono na 0,05.

Dla LoB zsumowane częstości wariantów posortowano od najniższej do najwyższej, wyznaczając 99. pozycję w rankingu dla każdego rodzaju wariantu (tabela 20). W oprogramowaniu MiSeqDx do jakościowego wykrywania wariantów stosuje się próg odcięcia (efektywną wartość LoB) odpowiadający częstości wariantu na poziomie 0,026. W oparciu o obliczone granice próby ślepej potwierdzono, że ww. próg prowadzi do błędu pierwszego rodzaju z prawdopodobieństwem nie większym niż 0,01.

Tabela 20 Granica próby ślepej

Typ wariantu	Całkowita liczba częstości wariantu	LoB dla odczytnika z serii 1 (%)	LoB dla odczytnika z serii 2 (%)
SNV	3120	0,87	0,75
Insercja	540	0,79	0,60
Delecja	540	0,96	0,84

W celu wyznaczenia granicy wykrywalności (LoD) dla każdej serii odczynników i dla każdego rodzaju wariantu obliczono odsetek częstości poszczególnych mutacji plasujący się poniżej progu odcięcia wynoszącego 0,026 (tabela 21). Ponieważ te odsetki były niższe niż prawdopodobieństwo błędu drugiego rodzaju wynoszące 5% (0,05), jako granicę wykrywalności obliczono medianę połączonych częstości wariantów (tabela 22). Za wartość LoD dla każdego rodzaju wariantu przyjęto większą spośród dwóch wartości obliczonych dla dwóch serii odczynników: 5,45% dla SNV, 4,88% dla insercji i 5,44% dla delecji.

Tabela 21 Granica wykrywalności

Seria odczynnika	Typ wariantu	Całkowita liczba częstości wariantu	Liczba pomiarów częstości wariantów <2,6%	Odsetek pomiarów częstości wariantów <2,6%	Granica wykrywalności (%)
1	SNV	2080	5	0,20	5,45
	Insercja	360	0	0,00	4,86
	Delecja	360	3	0,80	5,44
2	SNV	2080	26	1,30	5,44
	Insercja	360	0	0,00	4,88
	Delecja	360	0	0,00	5,24

Przedstawione poniżej badania ilustrują cechy wydajności aparatu MiSeqDx z innym reprezentatywnym oznaczeniem kierowanym na 56 mutacji w dwóch istotnych klinicznie genach odpowiedzialnych za powstawanie nowotworów (panel mutacji). Panel mutacji został zaprojektowany do swoistego wykrywania 56 mutacji w dwóch klinicznie istotnych genach odpowiedzialnych za powstawanie nowotworu (gen 1 i gen 2). Oznaczenie umożliwia równoczesne stwierdzenie obecności lub braku każdej z 56 mutacji we wszystkich sekwencjonowanych próbkach. Metodą referencyjną dla przedstawionych badań było dwukierunkowe sekwencjonowanie metodą Sanger.

Precyzja oznaczeń między seriami

W celu zbadania cech wydajności aparatu MiSeqDx dla różnych serii produkcyjnych zestawów odczynników (obejmujących odczynniki do kwalifikacji próbek, przygotowywania bibliotek i sekwencjonowania) przeprowadzono badanie precyzji oznaczeń między seriami w oparciu o dwugenowe oznaczenie reprezentatywne przy użyciu panelu pięciu zmieszanych preparatów FFPE spełniających wymagania kwalifikacyjne dla próbek. Każdy preparat FFPE zawierał dwie unikalne mutacje: jedną obecną na niższym (ok. 8%), zaś drugą na wyższym (ok. 14%) poziomie częstości. W trzech nienastępujących po sobie dniach przy użyciu trzech różnych serii zestawów odczynników wykonano po dwanaście (12) obserwacji każdej z pięciu mieszanin preparatów. Całkowita liczba obserwacji dla wszystkich serii odczynników wyniosła 180 we wszystkich mieszaninach preparatów i 360 dla wszystkich poziomów częstości występowania mutacji. W odniesieniu do wszystkich serii i dni oznaczeń oczekiwany wynik oznaczenia mutacji uzyskano w 99,7% (359/360) obserwacji. Jedną mutację o niskiej częstości nieprawidłowo rozpoznano jako typ dziki. W celu oszacowania zmienności układu wykonano analizę komponentów wariacyjnych dla każdej mutacji/poziomu częstości mutacji. Całkowite odchylenie standardowe mieściło się w zakresie od 0,011 do 0,029. Wartość komponentu całkowitego odchylenia standardowego, za który odpowiadała seria odczynnika, mieściła się w zakresie od 0 do 0,015.

Historia wersji

Nr dokumentu	Data	Opis zmiany
Dokument nr 200006218 wer. 01	Maj 2022 r.	Dodano numer katalogowy ulotki dołączonej do opakowania zestawu odczynników MiSeqDx Reagent Kit v3 Micro w części Sprzęt i materiały wymagane, ale niedostarczane. Dodano informacje o testach z użyciem zestawu odczynników MiSeqDx Reagent Kit v3 Micro w części Charakterystyka działania. Usunięto przestrożę dotyczącą Stanów Zjednoczonych z części Ostrzeżenia i środki ostrożności.
Dokument nr 200006218 wer. 00	Listopad 2021 r.	Wydanie pierwsze dotyczące oprogramowania MOS w wer. 4.0 i lokalnego menedżera przebiegu w wer. 3.0.

Patenty i znaki towarowe

Niniejszy dokument oraz jego treść stanowią własność firmy Illumina, Inc. oraz jej podmiotów zależnych („Illumina”) i są przeznaczone wyłącznie do użytku zgodnego z umową przez klienta firmy w związku z użytkowaniem produktów opisanych w niniejszym dokumencie, z wyłączeniem innych celów. Niniejszy dokument oraz jego treść nie będą wykorzystywane ani rozpowszechniane do innych celów i/lub publikowane w inny sposób, ujawniane ani kopiowane bez pisemnej zgody firmy Illumina. Firma Illumina na podstawie niniejszego dokumentu nie przenosi żadnych licencji podlegających przepisom w zakresie patentów, znaków towarowych czy praw autorskich ani prawu powszechnemu lub prawom pokrewnym osób trzecich.

W celu zapewnienia właściwego i bezpiecznego użytkowania produktów opisanych w niniejszym dokumencie podane instrukcje powinny być ściśle przestrzegane przez wykwalifikowany i właściwie przeszkolony personel. Przed rozpoczęciem użytkowania tych produktów należy zapoznać się z całą treścią niniejszego dokumentu.

NIEZAPOZNANIE SIĘ LUB NIEDOKŁADNE PRZESTRZEGANIE WSZYSTKICH INSTRUKCJI PODANYCH W NINIEJSZYM DOKUMENCIE MOŻE SPOWODOWAĆ USZKODZENIE PRODUKTÓW LUB OBRAŻENIA CIAŁA UŻYTKOWNIKÓW LUB INNYCH OSÓB ORAZ USZKODZENIE INNEGO MIENIA, A TAKŻE SPOWODUJE UNIEWAŻNIENIE WSZELKICH GWARANCJI DOTYCZĄCYCH PRODUKTÓW.

FIRMA ILLUMINA NIE PONOSI ODPOWIEDZIALNOŚCI ZA NIEWŁAŚCIWE UŻYTKOWANIE PRODUKTÓW (W TYM ICH CZĘŚCI I OPROGRAMOWANIA) OPISANYCH W NINIEJSZYM DOKUMENCIE.

© 2022 Illumina, Inc. Wszelkie prawa zastrzeżone.

Wszystkie znaki towarowe są własnością firmy Illumina, Inc. lub ich odpowiednich właścicieli. Szczegółowe informacje na temat znaków towarowych można znaleźć na stronie www.illumina.com/company/legal.html.

Informacje kontaktowe



Illumina
5200 Illumina Way
San Diego, California 92122, USA
+1 800 809 ILMN (4566)
+1 858 202 4566 (poza Ameryką Północną)
techsupport@illumina.com
www.illumina.com



Illumina Netherlands B.V.
Steenoven 19
5626 DK Eindhoven,
Holandia

Etykiety produktu

Objaśnienia symboli zamieszczonych na opakowaniu i samym produkcie znajdują się w kluczu symboli użytych w danym zestawie, dostępnym na stronie support.illumina.com.