

# Aparat NextSeq™ 550Dx Instrument

DO STOSOWANIA W DIAGNOSTYCE IN VITRO  
TYLKO NA EKSPORT

Nr kat. 20005715

## Przeznaczenie

Aparat NextSeq 550Dx jest przeznaczony do sekwencjonowania bibliotek DNA w oznaczeniach diagnostycznych *in vitro*. Aparat NextSeq 550Dx jest przeznaczony do użytku z konkretnymi zarejestrowanymi, certyfikowanymi lub zatwierdzonymi do stosowania *in vitro* odczynnikami diagnostycznymi i oprogramowaniem analitycznym.

## Zasady dotyczące procedury

Aparat Illumina NextSeq 550Dx służy do sekwencjonowania bibliotek DNA w oznaczeniach diagnostycznych *in vitro*. Jest on przeznaczony do stosowania przez wykwalifikowanych pracowników laboratoriów klinicznych, którzy zostali przeszkoleni w zakresie procedur diagnostycznych *in vitro* przeprowadzanych w laboratoriach klinicznych. Funkcje danych wejściowych w przypadku aparatu NextSeq 550Dx Instrument pełnią biblioteki generowane z DNA, w których do amplifikowanych sekwencji docelowych dodawane są indeksy próbek i sekwencje przechwytywania. Biblioteki próbek są przechwytywane w komorze przepływowej i sekwencjonowane w aparacie przy użyciu technologii sekwencjonowania przez syntezę (SBS, Sequencing By Synthesis). W technologii sekwencjonowania przez syntezę stosowana jest metoda terminatora odwracalnego służąca do wykrywania znakowanych fluorescencyjnie pojedynczych nukleotydów w miarę ich wbudowywania do rosnących łańcuchów 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 można przeprowadzić przetwarzanie rozpoznań nukleotydów w analizie wtórnej. Wtórna analiza w aparacie NextSeq 550Dx Instrument wykonywana jest przy użyciu różnych modułów analitycznych, w zależności od procedury. Proces przetwarzania w modułach do wykrywania wariantów linii zarodkowej i wariantów somatycznych obejmuje demultipleksowanie, tworzenie plików FASTQ, dopasowywanie, rozpoznawanie wariantów oraz tworzenie plików rozpoznawania wariantów (VCF i gVCF). Pliki VCF i gVCF zawierają informacje o wariantach znalezionych w określonych pozycjach w genomie referencyjnym.

## Konfiguracja z obsługą dwóch systemów operacyjnych

Aparat NextSeq 550Dx Instrument jest wyposażony w opcję konfiguracji z obsługą dwóch systemów operacyjnych umożliwiającą użytkowanie aparatu w trybie diagnostycznym (Dx, diagnostic) i trybie wyłącznie badawczym (RUO, research use only). Diagnostyczne oznaczenia sekwencjonowania w warunkach *in vitro*, w tym moduły do wykrywania wariantów linii zarodkowej i wariantów somatycznych, wykonuje się w trybie diagnostycznym. W trybie diagnostycznym można używać tylko odczynników do sekwencjonowania przeznaczonych do diagnostyki *in vitro*. Charakterystykę działania i ograniczenia procedur wykonywanych z użyciem aparatu NextSeq 550Dx Instrument wyznaczono przy użyciu modułów do wykrywania wariantów linii zarodkowej i wariantów somatycznych w trybie diagnostycznym.

## Ograniczenia dotyczące procedury

- 1 Do stosowania w diagnostyce *in vitro*.
- 2 Stosowane z zestawem odczynników na 300 cykli NextSeq 550Dx High Output Reagent Kit v2 (300 cycles) lub NextSeq 550Dx High Output Reagent Kit v2.5 (300 cycles) moduły do wykrywania wariantów linii zarodkowej i wariantów somatycznych są w stanie zapewnienia:

- ▶ Dane wyjściowe sekwencjonowania  $\geq 90$  giganukleotydów (Gb).
  - ▶ Długość odczytu (w przebiegu w trybie sparowanych końców)  $2 \times 150$  par nukleotydów (bp).
  - ▶ Nukleotydy o jakości równej lub większej  $Q30 \geq 75\%$  przy długości odczytu  $2 \times 150$  bp. 75% lub więcej nukleotydów ma wynik jakościowy w skali Phred  $\geq 30$ , co oznacza rozpoznawanie nukleotydów z dokładnością powyżej 99,9%.
- 3 Oprogramowanie do oznaczeń nie dopasowuje odczytów z polimorfizmami typu indel (insercjami, delecjami bądź ich kombinacjami), w przypadku których długość zawartości wynosi  $>25$  bp. W rezultacie oprogramowanie do oznaczeń nie wykrywa polimorfizmów typu indel o długości  $>25$  bp.
  - 4 Oprogramowanie do oznaczeń może nie dopasowywać odczytów ampliconów z ekstremalną zawartością wariantów, co prowadzi do zgłoszenia regionu jako typu dzikiego. Taka ekstremalna zawartość obejmuje:
    - ▶ Odczyty zawierające ponad trzy polimorfizmy typu indel.
    - ▶ Odczyty o długości co najmniej 30 bp z wariantem pojedynczego nukleotydu (SNV, Single Nucleotide Variant)  $>4\%$  całkowitej długości docelowego ampliconu (z wykluczeniem regionów sondy).
    - ▶ Odczyty o długości  $<30$  bp o zawartości SNV  $>10\%$  całkowitej długości ampliconu (z uwzględnieniem regionów sondy).
  - 5 Duże warianty, w tym warianty wielonukleotydowe (MNV, Multi-Nucleotide Variant) i duże polimorfizmy typu indel, mogą zostać zgłoszone jako odrębne mniejsze warianty w wyjściowym pliku VCF.
  - 6 Warianty delecji mogą być filtrowane lub pominięte w przypadku rozciągnięcia na dwa sąsiednie amplicony, jeśli długość delecji jest większa lub równa długości regionu nałożenia pomiędzy sąsiednimi ampliconami.
  - 7 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 długość regionu nałożenia jest mniejsza niż długość 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.
  - 8 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 oraz 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 polimorficzne typu indel w regionie startera mogą doprowadzić do fałszywie dodatniego rozpoznania pod koniec odczytu w sąsiedztwie startera.
  - 9 Polimorfizmy typu indel można filtrować na podstawie obciążenia systematycznego nici, jeśli występują w pobliżu końca odczytu i są poddawane procesowi soft-clippingu podczas dopasowania.
  - 10 Małe warianty MNV mogą nie zostać zwalidowane i są zgłaszane wyłącznie w module do wykrywania wariantów somatycznych.
  - 11 Delecje są zgłaszane w pliku VCF we współrzędnej poprzedniego nukleotydu według formatu VCF. W związku z tym przed zgłoszeniem, że rozpoznanie danego nukleotydu jest odniesieniem homozygotycznym, należy uwzględnić sąsiednie warianty.
  - 12 Ograniczenia swoiste dla wariantów linii zarodkowej:
    - ▶ Aparat NextSeq 550Dx Instrument, wykorzystujący moduł do wykrywania wariantów linii zarodkowej lokalnego menedżera przebiegu, jest przeznaczony do dostarczania wyników jakościowych w przypadku rozpoznawania wariantów linii zarodkowej (np. homozygotycznych, heterozygotycznych, typu dzikiego).
    - ▶ W przypadku stosowania z modułem do wykrywania wariantów linii zarodkowej minimalne pokrycie na amplicon niezbędne do dokładnego rozpoznania wariantu wynosi 150x. W konsekwencji wymaganych jest 150 potwierdzających fragmentów DNA, co odpowiada 300 nakładającym się odczytom w trybie sparowanych końców. 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 potrzeba co najmniej dziewięciu powtórzeń w regionie referencyjnym do wyfiltrowania wariantu. Uwzględniane są tylko powtórzenia o długości do 5 bp (R5x9).
  - ▶ Polimorfizm typu indel oraz wariant SNV w pojedynczym locus może skutkować zgłoszeniem tylko jednego wariantu.
- 13 Ograniczenia swoiste dla wariantów somatycznych:
- ▶ Aparat NextSeq 550Dx Instrument, wykorzystujący moduł do wykrywania wariantów somatycznych lokalnego menedżera przebiegu, jest przeznaczony do dostarczania wyników jakościowych w rozpoznawaniu 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 amplicon niezbędne do dokładnego rozpoznania wariantu wynosi 450x na pulę oligonukleotydów. W konsekwencji wymaganych jest 450 potwierdzających fragmentów DNA na pulę oligonukleotydów, co odpowiada 900 nakładającym się odczytom w trybie sparowanych końców. Liczba próbek i łączna liczba celowanych nukleotydów wpływają na pokrycie. Zawartość GC i inna zawartość genomowa mogą wpływać na pokrycie.
  - ▶ W przypadku rozpoznawania wariantów somatycznych potrzeba co najmniej sześciu 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 występowania, jednak częstości występowania 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 występowania wariantu względem całkowitego DNA wyekstrahowanego z tkanki nowotworowej i prawidłowej.

## Elementy produktu

- 1 Aparat NextSeq 550Dx Instrument (nr kat.: 20005715)
- 2 Elementy oprogramowania aparatu NextSeq 550Dx Instrument, w tym:

Aplikacja	Funkcja	Opis
Oprogramowanie NextSeq 550Dx Operating Software (NOS)	Sterowanie działaniem aparatu	Aplikacja NOS zarządza działaniem aparatu podczas sekwencjonowania oraz generuje obrazy używane w oprogramowaniu do analizy w czasie rzeczywistym (RTA, Real-Time Analysis).
Oprogramowanie RTA	Wykonywanie analizy podstawowej	Aplikacja RTA przetwarza obrazy wygenerowane przez oprogramowanie NOS dla poszczególnych płytek w cyklu sekwencjonowania na pliki rozpoznania nukleotydów, stanowiących dane wejściowe modułów analitycznych oprogramowania lokalnego menedżera przebiegu. Aplikacja RTA nie udostępnia interfejsu użytkownika.
Lokalny menedżer przebiegu	Interfejs wyboru modułu	Oprogramowanie lokalnego menedżera przebiegu stanowi zintegrowane z aparatem rozwiązanie do zarządzania użytkownikami, wybierania odpowiedniego modułu analitycznego i monitorowania statusu.

Aplikacja	Funkcja	Opis
Moduł do wykrywania wariantów somatycznych	Wykonywanie analizy wtórnej	Ten moduł analityczny oprogramowania lokalnego menedżera przebiegu przetwarza rozpoznania nukleotydów w analizie wtórnej. Proces przetwarzania obejmuje demultipleksowanie, generowanie plików FASTQ, dopasowywanie, rozpoznawanie wariantów i ich raportowanie. Algorytm rozpoznawania wariantów (Pisces) generuje pliki VCF zawierające informacje o wariantach zidentyfikowanych w określonych pozycjach genomu referencyjnego oraz zmierzoną częstość występowania wariantów.
Moduł do wykrywania wariantów linii zarodkowej	Wykonywanie analizy wtórnej	Ten moduł analityczny oprogramowania lokalnego menedżera przebiegu przetwarza rozpoznania nukleotydów w analizie wtórnej. Proces przetwarzania obejmuje demultipleksowanie, generowanie plików FASTQ, dopasowywanie, rozpoznawanie wariantów i ich raportowanie. Algorytm rozpoznawania wariantów (Pisces) generuje pliki VCF zawierające informacje o wariantach zidentyfikowanych w określonych pozycjach genomu referencyjnego oraz identyfikuje każdy wariant jako homozygotyczny lub heterozygotyczny.

## Warunki pracy

Element	Specyfikacja
Temperatura	Utrzymywać temperaturę w laboratorium w zakresie od 19°C do 25°C (22°C ±3°C). Jest to temperatura robocza aparatu. Podczas przebiegu nie dopuszczać do zmian temperatury otoczenia większych niż ±2°C.
Wilgotność	Utrzymywać wilgotność względną bez kondensacji w zakresie 20–80%.

## Sprzęt i materiały

### Wymagane wyposażenie i materiały, sprzedawane osobno

Zestaw odczynników na 75 cykli NextSeq 550Dx High Output Reagent Kit v2.5 (75 cycles), nr kat. 20028870

Zestaw odczynników na 300 cykli NextSeq 550Dx High Output Reagent Kit v2.5 (300 cycles), nr kat. 20028871

### Sprzęt i materiały wymagane, ale niedostarczane

Materiały eksploatacyjne dostarczane przez użytkownika w celu sekwencjonowania

Materiał eksploatacyjny	Dostawca	Cel
Chusteczki nasączone 70% alkoholem izopropylowym lub 70% etanolem	VWR, nr kat. 95041-714 (lub równoważne) Ogólny dostawca laboratoryjny	Czyszczenie komory przepływowej i przeznaczenie ogólne
Chusteczki laboratoryjne, niestrzępiące się	VWR, nr kat. 21905-026 (lub równoważne)	Czyszczenie komory przepływowej i przeznaczenie ogólne

Materiały eksploatacyjne dostarczane przez użytkownika w celu konserwacji aparatu

Materiał eksploatacyjny	Dostawca	Cel
NaOCl, 5% (podchloryn sodu)	Sigma-Aldrich, nr kat. 239305 (lub odpowiednik o jakości laboratoryjnej)	Plukanie aparatu za pomocą ręcznego płukania po przebiegu; rozcieńczony do 0,12%
Tween 20	Sigma-Aldrich, nr kat. P7949	Plukanie aparatu za pomocą opcji płukania ręcznego; rozcieńczony do 0,05%
Woda, jakość laboratoryjna	Ogólny dostawca laboratoryjny	Plukanie aparatu (płukanie ręczne)
Filtr powietrza	Illumina, nr kat. 20022240	Oczyszczanie powietrza pobieranego przez aparat do chłodzenia

## Wytyczne dotyczące wody laboratoryjnej

Do przeprowadzania procedur w aparacie zawsze należy używać wody laboratoryjnej lub wody dejonizowanej. Nie wolno używać wody z kranu. 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 zastosowań w biologii molekularnej

## Ostrzeżenia i środki ostrożności

PRZESTROGA Prawo federalne dopuszcza sprzedaż tego wyrobu do użytku lub zlecenia użytku wyłącznie przez lekarza lub na jego zlecenie bądź innego specjalistę posiadającego prawo do wykonywania zawodu zgodnie z prawem danego stanu, w którym prowadzi praktykę.

- 1 **Niektóre składniki odczynników dostarczanych przez firmę Illumina do stosowania z aparatem NextSeq 550Dx Instrument zawierają potencjalnie niebezpieczne substancje chemiczne. Wdychanie, połknięcie, kontakt ze skórą i oczami mogą powodować uszczerbek na zdrowiu. Należy nosić środki ochrony osobistej, 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, Safety Data Sheet) dostępna na stronie [support.illumina.com/sds.html](http://support.illumina.com/sds.html).
- 2 Wszelkie poważne incydenty związane z tym produktem należy niezwłocznie zgłaszać do firmy Illumina i właściwych organów państw członkowskich, w których użytkownik ma siedzibę i/lub pacjent mieszka.
- 3 Wszystkie próbki krwi należy traktować tak, jakby były zakażne pod kątem ludzkiego wirusa niedoboru odporności (HIV, Human Immunodeficiency Virus), wirusa zapalenia wątroby typu B (WZW B) oraz innych patogenów przenoszonych przez krew, tj. zachowywać uniwersalne środki ostrożności.
- 4 Nieprzestrzeganie przedstawionych procedur może powodować błędy w wynikach lub znaczne obniżenie jakości próbek.
- 5 Należy przestrzegać standardowych środków ostrożności obowiązujących w laboratorium. Nie pipetować ustami. Nie jeść, nie pić ani nie palić tytoniu w wyznaczonych obszarach roboczych. Podczas posługiwania się próbkami i odczynnikami z zestawów należy nosić jednorazowe rękawiczki i fartuchy laboratoryjne. Po zakończeniu posługiwania się próbkami i odczynnikami zestawu należy dokładnie umyć ręce.
- 6 Aby chronić produkty reakcji PCR przed zanieczyszczeniem przez odczynniki, przyrządy lub próbki genomowego DNA, należy przestrzegać zasad dobrej praktyki laboratoryjnej oraz zasad higieny. Zanieczyszczenie produktów reakcji PCR może powodować niedokładne i niepewne wyniki.
- 7 Aby zapobiegać zanieczyszczeniu produktów, należy upewnić się, że obszary procesu przed amplifikacją i po niej są wyposażone w odrębne przyrządy i materiały eksploatacyjne (np. pipety, końcówki do pipetowania, bloki grzejne, mieszadła wirowe i wirówki).
- 8 Parowanie indeksów z próbkami powinno przebiegać w ścisłej zgodności z wydrukowanym układem 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 dodatnie rozpoznanie próbki i nieprawidłowe dane w raporcie wyników.
- 9 Zdecydowanie zaleca się instalację dostarczonego przez użytkownika oprogramowania chroniącego komputer przed wirusami. W celu uzyskania instrukcji instalacji należy zapoznać się z podręcznikiem użytkownika.

- 10 Aparatu NextSeq 550Dx Instrument nie należy obsługiwać ze zdjętymi panelami. Obsługa aparatu z jakimkolwiek zdjętym panelem stwarza potencjalne ryzyko ekspozycji użytkownika na działanie napięcia sieciowego oraz napięcie prądu stałego.
- 11 Nie dotykać platformy w przedziale komory przepływowej. Znajdująca się w tym przedziale grzałka działa w zakresie temperatur od 22°C do 95°C i może spowodować oparzenia.
- 12 Aparat ma masę około 84 kg (185 funtów) i w przypadku upuszczenia lub nieprawidłowego obchodzenia się z nim może spowodować poważne obrażenia ciała.

## Instrukcja użytkownika

Poniższe instrukcje dotyczą korzystania z modułów do wykrywania wariantów linii zarodkowej i wariantów somatycznych w trybie diagnostycznym aparatu NextSeq 550Dx Instrument z wykorzystaniem zestawu odczynników na 300 cykli NextSeq 550Dx High Output Reagent Kit v2 (300 cycles) lub zestawu odczynników na 300 cykli NextSeq 550Dx High Output Reagent Kit v2.5 (300 cycles).

## Wprowadzanie informacji o przebiegu

Szczegółowe informacje zawiera NextSeq 550Dx Instrument Reference Guide (Instrukcja obsługi aparatu NextSeq 550Dx Instrument), nr dokumentu: 1000000009513, oraz instrukcja odpowiedniego modułu oprogramowania lokalnego menedżera przebiegu.

### Ustawianie parametrów

- 1 Zalogować się do lokalnego menedżera przebiegu.
- 2 Wybrać opcję **Create Run** (Utwórz przebieg), a następnie opcję **Somatic Variant** (Wariant somatyczny) lub **Germline Variant** (Wariant linii zarodkowej).
- 3 Wprowadzić nazwę przebiegu, która identyfikuje przebieg od sekwencjonowania po analizę. Można używać znaków alfanumerycznych, spacji, znaków podkreślenia lub łączników.
- 4 **[Opcjonalnie]** Wprowadzić opis przebiegu, aby ułatwić jego identyfikację. Można używać znaków alfanumerycznych, spacji, znaków podkreślenia lub łączników.
- 5 Wybrać liczbę próbek i zbiór indeksów z listy rozwijanej. Dokonując wyboru, należy uwzględnić poniższe informacje.
  - ▶ Lista rozwijana zawiera liczby próbek ze zbiorem indeksów. Na przykład „24-Set 1” oznacza 24 próbki do przetestowania z indeksami ze zbioru 1 indeksów.
  - ▶ Numery zbiorów indeksów odnoszą się do różnych zestawów par indeksów i5 oraz i7. Oba zbiory Set 1 (Zbiór 1) i Set 2 (Zbiór 2) zapewniają zróżnicowanie indeksów. Proponowane są dwa zbiory indeksów, aby zapobiec wyczerpaniu pojedynczego zbioru.
  - ▶ Należy wybrać liczbę próbek najbliższą liczbie próbek do testowania. Jeśli dokładnej liczby próbek nie ma na liście, należy wybrać najbliższą wartość, jednak mniejszą od liczby testowanych próbek. Na przykład: jeśli testowanych ma być 18 próbek, należy wybrać liczbę 16.
  - ▶ Sugerowane kombinacje dołków próbek i indeksów spełniające wymagania dotyczące zróżnicowania indeksów zostały zaznaczone na zielono.

### Importowanie plików wykazów dla przebiegu

- 1 Należy się upewnić, że wykazy, które mają zostać zaimportowane, znajdują się w dostępnej lokalizacji sieciowej lub na dysku USB.
- 2 Wybrać opcję **Import Manifests** (Importuj wykazy).
- 3 Przejść do pliku wykazu i wybrać wykazy do dodania.


UWAGA Aby udostępnić pliki wykazów dla wszystkich przebiegów przy użyciu modułu badania wariantów linii zarodkowej lub wariantów somatycznych, należy dodać wykazy za pomocą funkcji Module Settings (Ustawienia modułu). Do jej użycia wymagane są uprawnienia administratora. Więcej informacji na ten temat zawiera *Instrukcja obsługi aparatu NextSeq 550Dx* (nr dokumentu: 100000009513)

### Określanie próbek do przebiegu


Próbki do przebiegu określa się za pomocą jednej z poniższych opcji, zgodnie z podanymi wskazówkami.

- ▶ **Enter samples manually** (Ręczne wprowadzanie próbek) – należy użyć pustej tabeli na ekranie Create Run (Tworzenie przebiegu).
- ▶ **Import samples** (Importowanie próbek) – należy przejść do pliku zewnętrznego w formacie wartości rozdzielonych przecinkami (\*.csv). Szablon jest dostępny do pobrania na ekranie Create Run (Tworzenie przebiegu).

### Ręczne wprowadzanie próbek

- 1 Wpisać niepowtarzalną nazwę próbki (**moduł analizy wariantów somatycznych**) lub identyfikator próbki (**moduł analizy wariantów linii zarodkowej**).  
Można używać znaków alfanumerycznych, łączników lub znaków podkreślenia.
- 2 **[Opcjonalnie]** W przypadku dodatnich lub ujemnych próbek kontrolnych kliknąć prawym przyciskiem myszy i wybrać typ próbki kontrolnej.  
Określenie kontroli w jednym dołku próbki powoduje automatyczne wpisanie tej samej kontroli do odpowiadającego mu dołka w drugiej puli.
- 3 **[Opcjonalnie]** Wprowadzić opis próbki w polu Sample Description (Opis próbki).  
Można używać znaków alfanumerycznych, łączników lub znaków podkreślenia.
- 4 Wybrać adapter Indeksu 1 z listy rozwijanej Index 1 (i7) (Indeks 1, i7).  
Jeśli używane są sugerowane dołki próbek, oprogramowanie automatycznie wypełnia adaptory indeksów i7 oraz i5, które spełniają wymagania dotyczące zróżnicowania indeksów. Jeśli na liście nie ma dokładnej liczby testowanych próbek, należy wybrać adaptory indeksu dla dodatkowych dołków.
- 5 Wybrać adapter Indeksu 2 z listy rozwijanej Index 2 (i5) (Indeks 2, i5).
- 6 Wybrać plik wykazu z listy rozwijanej Manifest (Wykaz).  
Próbki w puli A wymagają innego wykazu niż próbki w puli B.
- 7 Wybrać odpowiednią opcję, aby wyświetlić, wydrukować lub zapisać układ płytki jako odniesienie do przygotowywania bibliotek.
  - Wybrać ikonę  **Drukuj**, aby wyświetlić układ płytki. Wybrać opcję **Print** (Drukuj), aby wydrukować układ płytki.
  - Wybrać opcję **Export** (Eksportuj), aby wyeksportować informacje o próbce do pliku zewnętrznego.
- 8 Wybrać opcję **Save Run** (Zapisz przebieg).

### Importowanie próbek

- 1 Wybrać opcję **Import Samples** (Importuj próbki) i przejść do lokalizacji pliku z informacjami o próbce. Można zaimportować dwa rodzaje plików.
  - Wybrać opcję **Template** (Szablon) na ekranie Create Run (Utwórz przebieg), aby przygotować nowy układ płytki. Plik szablonu zawiera właściwe nagłówki kolumn do zaimportowania. Dla próbek w przebiegu w każdej kolumnie wprowadzić informacje o próbce. Usunąć przykładowe informacje w nieużywanych komórkach, a następnie zapisać plik.
  - Użyć pliku z informacjami o próbce, który został wyeksportowany z modułu do wykrywania wariantów linii zarodkowej lub wariantów somatycznych przy użyciu funkcji eksportowania.
- 2 Wybrać ikonę  **Drukuj**, aby wyświetlić układ płytki.
- 3 Wybrać opcję **Print** (Drukuj), aby wydrukować układ płytki jako odniesienie do przygotowania bibliotek.
- 4 Wybrać opcję **Save Run** (Zapisz przebieg).

## Przygotowywanie kasety odczynników

W celu pomyślnego wykonania sekwencjonowania należy dokładnie przestrzegać wskazówek dotyczących kasety odczynników.

- 1 Wyjąć kasetę odczynników z chłodni o temperaturze od  $-25^{\circ}\text{C}$  do  $-15^{\circ}\text{C}$ .
- 2 W celu rozmrożenia odczynników wybrać jedną z podanych poniżej metod. Nie zanurzać kasety. Po rozmrożeniu kasety osuszyć ją przed przejściem do następnego etapu.

Temperatura	Czas rozmrażania	Czas stabilności
Kąpiel wodna o temperaturze od $15^{\circ}\text{C}$ do $30^{\circ}\text{C}$	60 minut	Nie przekraczać 6 godzin
Od $2^{\circ}\text{C}$ do $8^{\circ}\text{C}$	7 godzin	Nie przekraczać 5 dni

**UWAGA** Jeśli w tej samej kąpeli wodnej rozmrażanych jest więcej kaset, należy wydłużyć czas rozmrażania.

- 3 Wymieszać odczynniki, odwracając kasetę pięciokrotnie.
- 4 Sprawdzić dno kasety, aby upewnić się, że odczynniki zostały rozmrożone i nie zawierają osadów. Sprawdzić, czy pozycje 29, 30, 31 i 32 zostały rozmrożone, ponieważ są one największe, a ich rozmrażanie trwa najdłużej.
- 5 Delikatnie stuknąć o blat w celu zmniejszenia liczby pęcherzyków powietrza.  
W celu uzyskania najlepszych rezultatów należy umieścić próbkę w aparacie i skonfigurować przebieg.

## Przygotowywanie komory przepływowej

- 1 Wyjąć opakowanie z nową komorą przepływową z miejsca przechowywania w temperaturze od  $2^{\circ}\text{C}$  do  $8^{\circ}\text{C}$ .
- 2 Zdjąć folię zabezpieczającą z opakowania i zostawić zawartość na 30 minut w temperaturze pokojowej.

## Przygotowanie bibliotek do sekwencjonowania

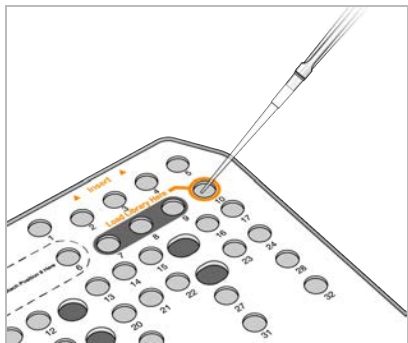
Zdenaturować i rozcieńczyć biblioteki do objętości ładowania wynoszącej 1,3 ml. W praktyce stężenie ładowania może się różnić w zależności od metod przygotowania i oznaczenia ilościowego biblioteki. Rozcieńczenie bibliotek próbek zależy od złożoności pul oligonukleotydów. Wskazówki dotyczące przygotowywania bibliotek próbek do sekwencjonowania, w tym przygotowywania rozcieńczeń i pul bibliotek, zawiera część instrukcji użytkownika dotycząca odpowiedniego zestawu do przygotowania biblioteki. Wymagana jest optymalizacja gęstości klastra aparatu NextSeq 550Dx Instrument.

## Ładowanie bibliotek do kasety odczynników

- 1 Oczyszczyć folię uszczelniającą zbiorniczek nr 10 z oznaczeniem **Load Library Here** (Tutaj załaduj bibliotekę) niestrzępiącą się ściereczką.
- 2 Przedziurawić folię końcówką czystej pipety 1 ml.
- 3 Załadować 1,3 ml przygotowanych bibliotek do zbiorniczka nr 10 z oznaczeniem **Load Library Here** (Tutaj załaduj bibliotekę). Należy uważać, aby podczas pipetowania bibliotek nie dotknąć folii uszczelniającej.



Rysunek 1 Ładowanie bibliotek



## Konfiguracja sekwencjonowania

- 1 Zalogować się do aparatu NextSeq 550Dx przy użyciu swojego lokalnego hasła do oprogramowania lokalnego menedżera przebiegu.
- 2 Na ekranie głównym oprogramowania NOS wybrać **Sequence** (Sekwencjonowanie).
- 3 Wybrać z listy przebieg, a następnie nacisnąć **Next** (Dalej).  
Zostaną otwarte ekrany konfiguracji przebiegu w następującej kolejności: Load Flow Cell (Ładowanie komory przepływowej), Load Buffer Cartridge (Ładowanie kasety z buforem), Load Reagent Cartridge (Ładowanie kasety odczynników) 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.
  - ▶ Wyjąć komorę przepływową z opakowania foliowego.
  - ▶ Otworzyć przezroczystą, plastikową kasetkę i wyjąć komorę przepływową.
  - ▶ Oczyszczyć szklaną powierzchnię komory przepływowej niestrzępiącą się ściereczką nasączoną alkoholem. Osuszyć szklaną powierzchnię niestrzępiącą się chusteczką laboratoryjną.
  - ▶ Upewnić się, że szklana powierzchnia komory przepływowej jest czysta. W razie potrzeby powtórzyć etap czyszczenia.
  - ▶ Wyjąć zużytą komorę przepływową z poprzedniego przebiegu.
  - ▶ Ustawić komorę przepływową nad bolcami regulacyjnymi i umieścić ją na platformie.
- 5 Wybrać **Load** (Załaduj).  
Drzwiczki zamkną się automatycznie, na ekranie pojawi się identyfikator komory przepływowej, a czujniki zostaną sprawdzone.
- 6 W celu opróżnienia zużytego pojemnika na odczynniki, załadowania kasety z buforem NextSeq 550Dx i załadowania kasety odczynników NextSeq 550Dx należy postępować zgodnie z monitami wyświetlanymi przez oprogramowanie. Po załadowaniu kasety z buforem i kasety odczynników NextSeq 550Dx oprogramowanie odczytuje i rejestruje znaczniki RFID. Identyfikatory kasety z buforem i kasety odczynników wyświetlane są na ekranie, po czym następuje sprawdzenie czujników.
- 7 Po zakończeniu automatycznego wstępnego testu kontrolnego systemu wybrać **Start** (Uruchom). (Nie jest wymagane, jeśli ustawiono uruchamianie automatyczne).
- 8 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 moduł oprogramowania lokalnego

menedżera przebiegów w aparacie NextSeq 550Dx Instrument uruchamia automatycznie analizę wtórną. Opisane tu procesy analizy wtórnej dotyczą modułów do wykrywania wariantów linii zarodkowej i wariantów somatycznych.

## Demultipleksowanie

Demultipleksowanie polega na porównywaniu każdej sekwencji odczytu indeksów z sekwencjami indeksów określonymi dla przebiegu. Na tym etapie nie są brane pod uwagę żadne wartości dotyczące jakości.

Odczyty indeksów identyfikuje się poprzez następujące czynności:

- ▶ Próbkę są ponumerowane, począwszy od numeru 1, na podstawie kolejności, w jakiej je wymieniono dla przebiegu.
- ▶ Próbkę numer 0 jest zarezerwowana dla klastrów, które nie zostały przypisane do próbki.
- ▶ Klastry są przypisywane do próbki, gdy sekwencja indeksu jest dokładnie zgodna lub gdy na odczyt indeksu przypada maksymalnie jedna niezgodność.

## Generowanie pliku FASTQ

Po zakończeniu demultipleksowania oprogramowanie generuje pliki analizy pośredniej w formacie FASTQ, który przedstawia sekwencje w postaci tekstowej. Pliki FASTQ zawierają odczyty dla każdej próbki i powiązane z nimi wyniki jakościowe. Klastry, które nie przeszły przez filtr, są wykluczane.

Każdy plik FASTQ zawiera odczyty dla tylko jednej próbki. Nazwa tej próbki wchodzi w skład nazwy pliku FASTQ. W modułach do wykrywania wariantów linii zarodkowej i wariantów somatycznych generowanych jest osiem plików FASTQ na próbkę dla każdej puli oligonukleotydów: cztery z odczytu 1 i cztery z odczytu 2, co daje łącznie 8 i 16 plików FASTQ na próbkę odpowiednio dla linii zarodkowej i somatycznej. Pliki FASTQ to podstawowe dane wejściowe do dopasowywania.

## Dopasowywanie

Podczas etapu dopasowywania klastry z poszczególnych próbek są dopasowywane za pomocą pasmowego algorytmu Smitha-Watermana do sekwencji amplikonów określonych w pliku wykazu.

Pasmowy algorytm Smitha-Watermana wykonuje półglobalne dopasowania sekwencji w celu określenia podobnych regionów między dwiema sekwencjami. Zamiast porównywać całą sekwencję, algorytm Smitha-Watermana porównuje segmenty o wszystkich możliwych długościach.

Każdy odczyt w trybie sparowanych końców jest oceniany pod względem dopasowania do sekwencji sondy odpowiednich dla tego odczytu.

- ▶ Odczyt 1 jest oceniany względem odwrotnej zgodności oligonukleotydów dalszych zależnych od umiejscowienia (DLSO, Downstream Locus-Specific Oligos).
- ▶ Odczyt 2 jest oceniany względem zgodności oligonukleotydów wcześniejszych zależnych od umiejscowienia (ULSO, Upstream Locus-Specific Oligos).
- ▶ Jeśli początek sekwencji odczytu jest niezgodny maksymalnie w jednym miejscu z sekwencją sondy, pełna długość odczytu jest następnie dopasowywana do sekwencji docelowej amplikonu dla tej sekwencji sondy.
- ▶ Jeśli przy dopasowaniu początku odczytu do sekwencji sondy występują maksymalnie trzy różnice (niedopasowania lub przesunięcia z powodu wiodących polimorfizmów typu indel), pełna długość odczytu jest dopasowywana do amplikonu docelowego dla tej sekwencji.
- ▶ Ze względu na mechanizm działania oznaczenia, polimorfizmy typu indel w oligonukleotydach DLSO i ULSO nie są obserwowane.

Dopasowania są odfiltrowywane z wyników na podstawie liczby niedopasowań w regionie docelowym lub całym amplikonie, w zależności od długości amplikonu. Odfiltrowane dopasowania są zapisywane w plikach dopasowania jako niedopasowane i nie są używane do rozpoznawania wariantów.

## Rozpoznawanie wariantów

Algorytm rozpoznawania wariantów Pisces służy do rozpoznawania wariantowych polimorfizmów typu SNV i indel z zastosowaniem bibliotek przygotowanych do użycia w aparacie.

## Raporty i dodatkowe pliki wyjściowe

Moduły do wykrywania wariantów generują raporty w plikach PDF i rozdzielonych znakami tabulacji plikach tekstowych (\*.txt), w których znajdują się takie metryki, jak głębokość sekwencjonowania czy liczba wariantów. Moduły mogą również generować pliki wyjściowe, takie jak pliki w formatach VCF i gVCF (plik rozpoznawania wariantów genomowych) wykorzystywane przy rozpoznawaniu wariantów.

## Procedury kontroli jakości

Oprogramowanie NextSeq 550Dx ocenia każdy przebieg, próbkę i rozpoznanie nukleotydu w oparciu o metryki kontroli jakości. Ocenie należy również poddać dodatnie i ujemne kontrole, które zaleca się uwzględnić przy przygotowaniu biblioteki. Kontrole należy ocenić w następujący sposób:

- **Kontrola ujemna (kontrola bez wzorca) lub inna kontrola ujemna** — musi generować oczekiwany wynik. Jeśli kontrola ujemna wygeneruje wynik inny od oczekiwanego, oznacza to, że mógł wystąpić błąd w śledzeniu próbki lub rejestracja starterów indeksujących była nieprawidłowa, lub też, że doszło do zanieczyszczenia próbki.
- **Dodatnia próbka kontrolna** — musi generować oczekiwany wynik. Jeśli kontrola dodatnia wygeneruje wynik inny od oczekiwanego, oznacza to, że mógł wystąpić błąd w śledzeniu próbki lub rejestracja starterów indeksujących była nieprawidłowa.

## Charakterystyka działania

Charakterystykę działania oznaczeń z użyciem aparatu NextSeq 550Dx Instrument ustalono z wykorzystaniem modułów do wykrywania wariantów linii zarodkowej i wariantów somatycznych oraz zestawu TruSeq Custom Amplicon Kit Dx i zestawu odczynników na 300 cykli NextSeq 550Dx High Output Reagent Kit v2 (300 cycles), a następnie potwierdzono z użyciem zestawu odczynników na 300 cykli NextSeq 550Dx High Output Reagent Kit v2.5 (300 cycles). Wykonane badania obejmowały indeksowanie próbek, przenoszenie próbek, poziom wejściowy DNA, czułość analityczną (LoB/LoD), dokładność, precyzję, porównanie metod i odtwarzalność.

Badania analityczne z wykorzystaniem zestawu odczynników na 300 cykli NextSeq 550Dx High Output Reagent Kit v2.5 (300 cycles) zostały zaprojektowane tak, aby ocenić deklarowane parametry wydajności ustalone uprzednio przy użyciu zestawu odczynników na 300 cykli NextSeq 550Dx High Output Reagent Kit v2 (300 cycles). Wyniki uzyskane przy użyciu zestawu TruSeq Custom Amplicon Kit Dx wykazały, że zestawy odczynników (v2 i v2.5) charakteryzują się podobnymi parametrami. Informacje na temat cech charakterystyki działania zależnych od czynników przedanalizacyjnych, takich jak sposób ekstrakcji czy substancje zakłócające, można znaleźć w *ulotce dołączonej do opakowania zestawu TruSeq Custom Amplicon Kit Dx*.

## Definicje obliczeń stosowanych w określaniu charakterystyki działania

- 1 Procentową zgodność wyników dodatnich (PPA, 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, True Positive). Loci wariantowe zgłoszone w oznaczeniu jako rozpoznania sekwencji referencyjnej lub jako inne rozpoznania wariantowe stanowią wyniki fałszywie ujemne (FN, False Negative).
- 2 Procentową zgodność wyników ujemnych (NPA, Negative Percent Agreement) oblicza się jako odsetek loci sklasyfikowanych metodą referencyjną jako typ dziki i prawidłowo zgłoszonych w oznaczeniu.
  - ▶  $(\text{liczba loci typu dzikiego prawidłowo zgłoszonych w oznaczeniu}) / (\text{całkowita liczba loci typu dzikiego})$   
Loci typu dzikiego zgłoszone przez oznaczenie i zgodne z metodą referencyjną stanowią wyniki prawdziwie ujemne (TN, True Negative). Loci typu dzikiego zgłoszone w oznaczeniu jako warianty stanowią wyniki fałszywie dodatnie (FP, False Positive).
- 3 Ogólną zgodność procentową (OPA, Overall Percent Agreement) oblicza się jako odsetek loci prawidłowo sklasyfikowanych w oznaczeniu w stosunku do metody referencyjnej.

- ▶  $((\text{liczba loci wariantowych prawidłowo zgłoszonych w oznaczeniu}) + (\text{liczba loci typu dzikiego prawidłowo zgłoszonych w oznaczeniu})) / ((\text{całkowita liczba loci wariantowych}) + (\text{całkowita liczba loci typu dzikiego}))$
- 4 Obliczenia PPA, NPA i OPA nie uwzględniają braków rozpoznań (loci wariantowych lub referencyjnych niespełniających warunków jednego lub większej liczby filtrów jakościowych).
- 5 Odsetek rozpoznań autosomalnych 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.

## Wydajność zestawu odczynników na 300 cykli NextSeq 550Dx High Output Reagent Kit v2 (300 Cycles)

### Indeksowanie próbek

Startery indeksujące próbki dodawane na etapie przygotowywania biblioteki przypisują do każdego DNA próbki unikalną sekwencję. Takie unikalne sekwencje umożliwiają pulowanie wielu próbek w tym samym przebiegu sekwencjonowania. Indeksowanie próbek stosuje się zarówno w procedurach dotyczących linii zarodkowych, jak i somatycznych. Celem badania było wyznaczenie minimalnej (8) i maksymalnej (96) liczby próbek, które mogą być przetwarzane w pojedynczym przebiegu sekwencjonowania w aparacie NextSeq 550Dx Instrument. Ośiem unikalnych próbek Platinum Genome przebadano przy użyciu 12 różnych kombinacji starterów indeksujących na próbkę. Wyniki dla próbek uzyskane z czterech przebiegów sekwencjonowania przy użyciu modułu do wykrywania wariantów linii zarodkowej porównano z bazą danych Platinum Genomes w wersji 2016-1.0.

W pierwszej serii przebiegów przebadano łącznie 96 unikalnie zindeksowanych bibliotek próbek przy użyciu oznaczenia reprezentatywnego w celu rozpoznania szeregu genów wchodzących w skład wszystkich 23 ludzkich chromosomów i zawierających po 12 588 nukleotydów na nić oraz weryfikacji zdolności oznaczenia do rozpoznawania genotypu w określonej próbce z zastosowaniem różnych kombinacji starterów indeksujących. W drugiej serii przebiegów ośiem unikalnie zindeksowanych bibliotek próbek poddano sekwencjonowaniu w dwóch przebiegach w celu zweryfikowania minimalnej liczby obsługiwanych indeksów.

Dla przebiegów z 96 indeksami wartość PPA dla SNV wynosiła od 98,7% do 100%, wartość PPA dla insercji i delecji wynosiła 100%, natomiast wartość NPA wynosiła 100% dla każdej kombinacji 96 indeksów. Dla przebiegów z 8 indeksami wartości PPA dla SNV, insercji i delecji wynosiły 100%, zaś wartość NPA wynosiła 100% dla każdej kombinacji ośmiu indeksów.

### Przeniesienie próbek

Aparat NextSeq 550Dx Instrument umożliwia sekwencjonowanie wielu próbek w pojedynczym przebiegu łącznie z kontrolami. Przeprowadzono badanie mające na celu ocenę stopnia przenoszenia próbek w pojedynczym przebiegu sekwencjonowania oraz między przebiegami sekwencjonowania. Dwie próbki wzorca Platinum Genome, jedną męską i jedną żeńską, zbadano przy użyciu oznaczenia reprezentatywnego mającego na celu rozpoznanie różnych genów obejmujących 12 588 nukleotydów (150 amplikonów) znajdujących się w 23 różnych chromosomach, w tym w obu chromosomach płciowych. Biblioteki sekwencjonowano przy użyciu aparatu NextSeq 550Dx Instrument i modułu do wykrywania wariantów linii zarodkowej. Przeniesienie próbek męskich do próbek żeńskich stwierdzono na podstawie obecności odczytów właściwych dla amplikonów chromosomu Y w próbkach żeńskich.

Przeniesienie w ramach pojedynczego przebiegu może zajść w trakcie generowania klastrów, rozpoznawania nukleotydów podczas cyklu odczytu indeksów i demultipleksowania próbek. Na potrzeby badania przeniesień próbek w pojedynczym przebiegu sekwencjonowania wykonano jedno sekwencjonowanie puli biblioteki składającej się z 46 powtórzeń próbek męskich i żeńskich oraz czterech kontroli bez wzorca przy użyciu aparatu NextSeq 550Dx Instrument. Przeniesienie próbek w ramach przebiegu oceniano, porównując pokrycie amplikonów chromosomu Y w każdym powtórzeniu próbki żeńskiej w porównaniu ze średnim pokryciem amplikonów chromosomu Y we wszystkich powtórzeniach próbek męskich. Obserwowana mediana przeniesień w danym przebiegu wyniosła 0,084%.

Na potrzeby badania przeniesienia między przebiegami przygotowano dwie pulę bibliotek, które kolejno poddano sekwencjonowaniu na jednym aparacie NextSeq 550Dx Instrument. Pierwsza pula składała się z 46 powtórzeń próbek żeńskich i dwóch kontroli bez wzorca. Druga pula składała się z 46 powtórzeń próbek męskich i dwóch kontroli bez wzorca. W obu pulach stosowano ten sam zestaw adapterów indeksów. Pulę żeńską sekwencjonowano w pierwszej

kolejności, następnie sekwencjonowano pulę męską, a w dalszej kolejności wykonywano powtórzenie sekwencjonowania puli żeńskiej. Przenoszenie próbek między przebiegami oceniano, porównując pokrycie amplikonów chromosomu Y między odpowiednimi powtórzeniami powtórnego przebiegu puli żeńskiej i przebiegu puli męskiej. Obserwowana mediana przeniesień między przebiegami wyniosła 0,0076%.

## Poziom wejściowy DNA

### Krew (linia zarodkowa)

Dla aparatu NextSeq 550Dx Instrument ustalono zakres poziomów wejściowych DNA z krwi w bibliotekach przygotowywanych z użyciem zestawu TruSeq Custom Amplicon Kit Dx z wykorzystaniem procedury modułu do wykrywania wariantów linii zarodkowej. Oceny ww. zakresu dokonano, wykonując badanie serii rozcieńczeń 13 próbek Platinum Genome w reprezentatywnym oznaczeniu rozpoznającym różne geny obejmujące 12 588 nukleotydów na 23 różnych chromosomach. Bibliotekę sekwencjonowano na dwóch aparatach NextSeq 550Dx Instrument przy użyciu jednej serii zestawu odczytników na 300 cykli NextSeq 550Dx High Output Reagent Kit v2 (300 cycles).

Pięć próbek poddano badaniu w dwóch powtórzeniach przy pięciu poziomach wejściowych DNA z zakresu od 250 ng do 12 ng (250 ng, 100 ng, 50 ng, 25 ng i 12 ng). Osiem próbek poddano badaniu w jednym powtórzeniu dla każdego z pięciu poziomów wejściowych DNA. W celu wyznaczenia dokładności genotypy próbek porównano do genotypów z bazy danych Platinum Genomes w wersji 2016-1.0. Wyniki pozyskano dla każdego poziomu wejściowego. Wartości PPA dla każdego rodzaju wariantu (SNV, insercje i delecje) przedstawiono w [Tabela 1](#); wartości NPA zawiera [Tabela 2](#). Wszystkie poziomy wejściowe charakteryzowały się podobną dokładnością. Zalecany poziom wejściowy DNA dla zestawu TruSeq Custom Amplicon Kit Dx wynosi 50 ng, natomiast 25 ng i 100 ng stanowią odpowiednio dolną i górną granicę wartości, dla których zachowana jest zgodność z charakterystyką działania.

Tabela 1 Wyniki PPA dla każdego poziomu wejściowego DNA w zależności od rodzaju wariantu

Poziom wejściowy DNA (ng)	Typ wariantu	Warianty oczekiwane	TP	FN	Brak rozpoznania wariantu	PPA (%)
12	SNV	2412	2381	31	0	98,7
25			2404	8	0	99,7
50			2403	9	0	99,6
100			2412	0	0	100
250			2412	0	0	100
12	Insercja	808	784	3	21	99,6
25			781	5	22	99,4
50			786	2	20	99,8
100			786	0	22	100
250			786	0	22	100
12	Delecja	758	732	12	14	98,4
25			737	7	14	99,1
50			742	2	14	99,7
100			744	0	14	100
250			744	0	14	100

Tabela 2 NPA dla każdego poziomu wejściowego DNA

Poziom wejściowy DNA (ng)	TN	FP	Brak rozpoznania wzorca	NPA (%)
12	430940	4	26	>99,9
25	430936	0	34	100

50	430936	2	32	>99,9
100	430942	0	28	100
250	430942	0	28	100

**FFPE (somaticzne)**

Ustalono zakres poziomów wejściowych DNA z utwalonych formaliną i zatopionych w parafinie skrawków (FFPE) w bibliotekach przygotowywanych z użyciem zestawu TruSeq Custom Amplicon Kit Dx dla aparatu NextSeq 550Dx Instrument z wykorzystaniem modułu do wykrywania wariantów somatycznych. Oceny ww. zakresu dokonano, wykonując badanie serii rozcieńczeń trzech próbek Platinum Genome w reprezentatywnym oznaczeniu rozpoznającym różne geny obejmujące 12 588 nukleotydów na 23 różnych chromosomach. Linie komórkowe Platinum Genome GM12878 i GM12877 utwalono w formalinie i zatopiono w parafinie, po czym dokonano ekstrakcji DNA. Linie komórkową GM12878 rozcieńczono linią GM12877 w taki sposób, aby częstości występowania allelu wariantu (VAF) dla 81 wariantów (55 SNV, 10 insercji i 16 delecji) były bliskie odpowiednio 0,025; 0,05 lub 0,10. Ponadto w każdej próbce znajdowało się 91 wariantów o wyższej częstości występowania aż do VAF o wartości 1,0. Próbkę przetwarzano w dwóch powtórzeniach przy pięciu poziomach wejściowych DNA o średniej różnicy cykli ilościowych (dCq) zmierzonej przy użyciu zestawu TruSeq Custom Amplicon Dx – FFPE QC Kit i wynoszącej 2,1; 3,6; 4,6; 6,0 i 7,8. Każdą bibliotekę sekwencjonowano na dwóch aparatach NextSeq 550Dx Instrument przy użyciu dwóch serii zestawu odczynników na 300 cykli NextSeq 550Dx High Output Reagent Kit v2 (300 cycles). W celu wyznaczenia dokładności rozpoznania wariantów w próbkach porównano do wyników wzorcowych uzyskanych z użyciem bazy danych Platinum Genomes w wersji 2016-1.0. Wartości PPA dla każdego rodzaju wariantu (SNV, insercje i delecje) zawiera [Tabela 3](#); wartości NPA przedstawia [Tabela 4](#). Zalecany poziom wejściowy DNA dla wariantów o VAF = 0,05 lub powyżej wynosi dCq ≤4; wartość 4,6 stanowi dolną granicę zachowania zgodności z charakterystyka działania.

Tabela 3 Wyniki PPA dla każdego poziomu wejściowego DNA w zależności od rodzaju wariantu

Średnia dCq	Typ wariantu	Warianty oczekiwane	Oczekiwany brak rozpoznania	VAF w docelowym rozcieńczeniu					
				0,025		0,05		0,10	
				Brak rozpoznania wariantu	PPA (%)	Brak rozpoznania wariantu	PPA (%)	Brak rozpoznania wariantu	PPA (%)
2,1	SNV	808	Nie dotyczy.	196	100	0	100	0	100
3,6				250	99,3	4	100	0	100
4,6				251	94,6	51	99,2	5	100
6,0				257	65,3	213	91,4	100	100
7,8				254	69,3	185	90,7	100	100
2,1	Insercja	264	8	66	96,5	8	100	8	100
3,6				62	97,0	8	100	8	100
4,6				48	96,3	21	100	8	100
6,0				40	80,4	47	98,2	24	95,8
7,8				57	87,0	56	96,2	31	100
2,1	Delecja	304	16	58	100	16	100	16	100
3,6				80	100	16	100	16	100
4,6				65	95,4	28	100	16	100
6,0				78	74,8	105	94,0	36	100
7,8				76	75,0	79	95,1	57	98,8

Tabela 4 NPA dla każdego poziomu wejściowego DNA

Średnia dCq	Oczekiwany typ dziki	VAF w docelowym rozcieńczeniu					
		0,025		0,05		0,10	
		Brak rozpoznań wzorca	NPA (%)	Brak rozpoznań wzorca	NPA (%)	Brak rozpoznań wzorca	NPA (%)
2,1	93688	344	100	260	100	324	100
3,6		400	100	332	100	380	100
4,6		1308	100	1336	100	784	100
6,0		3900	>99,9	3296	>99,9	2996	100
7,8		3020	>99,9	2880	>99,9	2448	>99,9

### Czułość analityczna (granica próby ślepej [LoB] i granica wykrywalności [LoD])

Badanie to przeprowadzono w celu oceny granicy próby ślepej (LoB) i granicy wykrywalności (LoD) modułu do wykrywania wariantów somatycznych aparatu NextSeq 550Dx Instrument. W tym badaniu użyto oznaczenia reprezentatywnego w celu rozpoznania różnych genów obejmujących 12 588 nukleotydów znajdujących się na 23 różnych chromosomach. Linie komórkowe Platinum Genome GM12878 i GM12877 utrwalono w formalinie i zatopiono w parafinie, po czym dokonano ekstrakcji DNA. Linie komórkową GM12878 rozcieńczono linią GM12877 w taki sposób, aby częstość występowania wariantu dla 74 wariantów (53 SNV, 7 insercji i 14 delecji) była bliska  $0,05 \pm 0,02$ . Linie komórkową GM12877 i rozcieńczoną linię komórkową GM12878 (GM12878-D) badano w sześciu kolejnych dniach pracy na jednym aparacie przy naprzemiennym stosowaniu dwóch serii zestawu odczynników na 300 cykli NextSeq 550Dx High Output Reagent Kit v2 (300 cycles); w sumie wykonano sześć przebiegów sekwencjonowania. W rezultacie przy użyciu każdej serii odczynników uzyskano 60 powtórzeń dla każdego wariantu linii komórkowej GM12878-D i 72 powtórzenia dla każdej odpowiadającej im współrzędnej typu dzikiego w linii komórkowej GM12877. Wartości granicy próby ślepej (LoB) i granicy wykrywalności (LoD) obliczono przy użyciu klasycznego podejścia przedstawionego w dokumencie EP17-A2 Instytutu Norm Kliniknych i Laboratoryjnych (CLSI, Clinical & Laboratory Standards Institute) z wykorzystaniem opcji nieparametrycznej. Wartości LoB i LoD obliczono oddzielnie dla SNV, insercji i delecji, sumując częstości występowania wariantów danego rodzaju. Prawdopodobieństwo błędu pierwszego rodzaju zdefiniowano jako 0,01, zaś prawdopodobieństwo błędu drugiego rodzaju zdefiniowano jako 0,05.

Dla LoB zsumowane częstości występowania wariantów posortowano od najniższej do najwyższej, wyznaczając 99. pozycję w rankingu dla każdej serii odczynnika i dla każdego rodzaju wariantu (Tabela 5). W module do wykrywania wariantów somatycznych do jakościowego wykrywania wariantów stosuje się punkt odcięcia (efektywną wartość LoB) odpowiadający VAF na poziomie 0,026. W oparciu o obliczone wartości LoB potwierdzono, że ww. punkt odcięcia prowadzi do błędu pierwszego rodzaju z prawdopodobieństwem nie większym niż 0,01.

Tabela 5 Granica próby ślepej

Typ wariantu	Obserwacje łącznie	LoB dla odczynnika z serii 1 (%)	LoB dla odczynnika z serii 2 (%)
SNV	3816	0,77	0,77
Insercja	504	0,56	0,56
Delecja	1008	1,20	1,20

W celu wyznaczenia granicy wykrywalności (LoD) dla każdej serii odczynników i dla każdego rodzaju wariantu obliczono odsetek częstości poszczególnych mutacji plasujący się poniżej punktu odcięcia wynoszącego 0,026 (Tabela 6). Ponieważ te odsetki były niższe niż prawdopodobieństwo błędu drugiego rodzaju wynoszące 5% (0,05), jako granicę wykrywalności obliczono medianę połączonych częstości występowania wariantów (Tabela 6). Za wartość LoD dla każdego rodzaju wariantu przyjęto większą spośród dwóch wartości obliczonych dla dwóch serii odczynników: 4,97% dla SNV; 5,12% dla insercji i 5,26% dla delecji.

Tabela 6 Granica wykrywalności

Seria odczynnika	Typ wariantu	Obserwacje łącznie	Liczba pomiarów VAF <2,6%	Odsetek pomiarów VAF <2,6%	Granica wykrywalności (%)
1	SNV	3180	53	1,7	4,94
	Insercja	420	6	1,4	5,08
	Delecja	840	7	0,8	5,22
2	SNV	3180	51	1,6	4,97
	Insercja	420	5	1,2	5,12
	Delecja	840	7	0,80	5,26

## Dokładność

### Linia zarodkowa

W celu dokonania oceny dokładności rozpoznawania wariantów przez moduł do wykrywania wariantów linii zarodkowej aparatu NextSeq 550Dx Instrument przeprowadzono następujące badanie przy użyciu zestawu odczynników na 300 cykli NextSeq 550Dx High Output Reagent Kit v2 (300 cycles): 13 unikalnych próbek Platinum Genome poddano badaniu w reprezentatywnym oznaczeniu rozpoznającym różne geny obejmujące 12 588 nukleotydów (150 amplikonów) na 23 różnych chromosomach. Wykonano w sumie dziewięć przebiegów z użyciem trzech aparatów do sekwencjonowania, trzech serii odczynników z udziałem trzech operatorów w okresie pięciu dni rozpoczęcia oznaczeń. Dokładność rozpoznania SNV, insercji i delecji oznaczono, porównując wyniki do wyników uzyskanych dobrze scharakteryzowaną, złożoną metodą referencyjną — bazą danych Platinum Genomes w wersji 2016-1.0. O ile nie podano inaczej, regiony o wysokim poziomie ufności zostały zdefiniowane w oparciu o ww. metodę referencyjną.

Tabela 7 Podsumowanie zgodności dla linii zarodkowej

Kryteria	Obserwacje łącznie <sup>1</sup>	Wynik dla obserwacji <sup>2</sup>	Wynik dla przebiegu <sup>3</sup>
PPA dla SNV	819	98,7	>99,9
PPA dla insercji	819	95,0	98,9
PPA dla delecji	819	100	100
NPA	819	100	100
OPA	819	>99,9	>99,9

<sup>1</sup> Obliczone jako liczba próbek na przebieg (91) × liczba przebiegów (9) = 819.

<sup>2</sup> Najniższa obserwowana wartość dla powtórzenia próbki we wszystkich dziewięciu przebiegach.

<sup>3</sup> Najniższa wartość w połączonej analizie danych z każdego przebiegu.

**Tabela 8** 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 bazą danych Platinum Genomes w wersji 2016-1.0. 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 hg19 genomu ludzkiego.



Tabela 8 Zgodność wariantów linii zarodkowej w zależności od próbki

Próbka	Średni odsetek rozpoznania	Warianty oczekiwane <sup>1</sup>	TP	FN	Brak rozpoznania wariantu	TN	FP	PPA	NPA	OPA
NA12877	>99,9	4788	4788	0	0	756762	0	100	100	100
NA12878	>99,9	8505	8379	1	125	751464	0	>99,9	100	>99,9
NA12879	>99,9	6048	5985	5	58	757701	0	99,9	100	>99,9
NA12880	>99,9	6993	6930	0	63	757638	0	100	100	100
NA12881	>99,9	7875	7811	3	61	751653	0	>99,9	100	>99,9
NA12882	>99,9	6300	6174	3	123	754803	0	>99,9	100	>99,9
NA12883	>99,9	7119	7056	0	63	751905	0	100	100	100
NA12884	>99,9	7182	7119	6	57	754146	0	99,9	100	>99,9
NA12885	>99,9	7686	7560	2	124	754173	0	>99,9	100	>99,9
NA12886	>99,9	7245	7182	7	56	752469	0	99,9	100	>99,9
NA12887	>99,9	7119	7119	0	0	750645	0	100	100	100
NA12888	>99,9	6804	6804	0	0	756065	0	100	100	100
NA12893	>99,9	7434	7371	1	62	750015	0	>99,9	100	>99,9

<sup>1</sup> Całkowita liczba wariantów we wszystkich powtórzeniach próbki w 9 przebiegach.

Tabela 9 zawiera dane z badania prezentowane dla poszczególnych próbek, przy czym wyniki wariantów zostały porównane z dobrze scharakteryzowaną, złożoną metodą referencyjną. Wykrywalność została oceniona oddzielnie dla każdego rodzaju wariantu – SNV, insercji i delecji. Pozycje referencyjne zostały wykluczone.

Tabela 9 Zgodność wariantów linii zarodkowej w zależności od typu wariantu w próbce

>Próbka	Warianty SNV			Insercje			Delecje		
	>Oczekiwane	>TP	>FN	>Oczekiwane	>TP	>FN	Oczekiwane	TP	FN
NA12877	2331	2331	0	1323	1323	0	1134	1134	0
NA12878	5733	5733	0	1260	1197	1	1512	1449	0
NA12879	3591	3591	0	1323	1260	5	1134	1134	0
NA12880	4221	4221	0	1512	1512	0	1260	1197	0
NA12881	4914	4913	1	1512	1449	2	1449	1449	0
NA12882	3717	3717	0	1386	1323	3	1197	1134	0
NA12883	4284	4284	0	1449	1449	0	1386	1323	0
NA12884	4284	4284	0	1575	1512	6	1323	1323	0
NA12885	4725	4725	0	1575	1512	2	1386	1323	0
NA12886	4347	4347	0	1449	1386	7	1449	1449	0
NA12887	4284	4284	0	1323	1323	0	1512	1512	0
NA12888	4158	4158	0	1449	1449	0	1197	1197	0
NA12893	4599	4599	0	1386	1323	1	1449	1449	0

Próbki poddano dalszej analizie pod kątem rozpoznawania małych insercji i delecji (polimorfizmów typu indel). Ogólne podsumowanie przedstawia Tabela 10. 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.

Tabela 10 Podsumowanie wykrywania polimorfizmów typu indel w module do wykrywania wariantów linii zarodkowej

Typ wariantu	Warianty oczekiwane	TP	FN	Brak rozpoznania wariantu	PPA
Insercja	18522	18018	27	477	99,9
Delecja	17388	17073	0	315	100

Oznaczenie reprezentatywne składało się ze 150 amplikonów zaprojektowanych tak, aby pokryć zróżnicowaną zawartość genomową. Zawartość GC w amplikonach mieściła się w zakresie 0,19–0,87. Amplikony charakteryzowały się również pewnym zakresem powtórzeń pojedynczych nukleotydów (np. PolyA, PolyT), dinukleotydów i trinukleotydów. Dane skompilowano w oparciu o poszczególne amplikony (Tabela 11) w celu oznaczenia wpływu zawartości genomowej na odsetek prawidłowych rozpoznań. Odsetek prawidłowych rozpoznań obejmuje rozpoznania wariantu i wzorca i jest mniejszy od 100% w przypadku nieprawidłowych rozpoznań lub braku rozpoznań.

Tabela 11 Dokładność oznaczeń wariantów linii zarodkowej na poziomie amplikonów

Amplikon	Chromosom	Początek amplikonu	Koniec amplikonu	Rozmiar analizowanego fragmentu	Nukleotydy w regionach o wysokim poziomie ufności	Materiał genetyczny amplikonu	Zawartość GC	Prawidłowe rozpoznania	Nieprawidłowe rozpoznania	Brak rozpoznań	Odsetek prawidłowo rozpoznanych nukleotydów
1	1	36450499	36450591	93	93	Indel	0,22	76167	0	0	100
2	1	109465122	109465200	79	79	PolyA (5), PolyC (5), indel	0,38	64701	0	0	100
3	1	218353867	218353957	91	91	Indel	0,4	74529	0	0	100
4	1	223906657	223906748	92	92	Indel	0,49	75348	0	0	100
5	1	228526602	228526682	81	81	PolyG (5)	0,69	66339	0	0	100
6	1	236372039	236372108	70	70	PolyT (10), indel	0,39	57330	0	0	100
7	1	247812041	247812128	88	88	PolyA (5), CT(3), TAA (3), indel	0,27	72072	0	0	100
8	2	55862774	55862863	90	90	Indel	0,28	73710	0	0	100
9	2	87003930	87004009	80	80	Indel	0,38	65520	0	0	100
10	2	177016721	177016805	85	81	Nd.	0,65	66339	0	0	100
11	2	186625727	186625801	75	75	PolyA (8)	0,35	61425	0	0	100
12	2	190323504	190323591	88	88	PolyT (5)	0,42	72072	0	0	100
13	2	200796740	200796826	87	87	PolyT (5), indel	0,31	71253	0	0	100
14	2	212245049	212245139	91	91	PolyT (5), PolyA (6), indel	0,3	74529	0	0	100

Amplikon	Chromosom	Początek amplikonu	Koniec amplikonu	Rozmiar analizowanego fragmentu	Nukleotydy w regionach o wysokim poziomie ufności	Materiał genetyczny amplikonu	Zawartość GC	Prawidłowe rozpoznania	Nieprawidłowe rozpoznania	Brak rozpoznań	Odsetek prawidłowo rozpoznanych nukleotydów
15	2	228147052	228147144	93	93	Indel	0,43	76167	0	0	100
16	2	235016350	235016422	73	73	PolyT (5), indel	0,42	59787	0	0	100
17	3	4466229	4466321	93	93	AT(3), indel	0,27	74823	0	1344	98,2
18	3	46620561	46620643	83	83	Nd.	0,43	67977	0	0	100
19	3	49851331	49851400	70	70	CT(3), indel	0,49	57330	0	0	100
20	3	189713161	189713248	88	88	PolyA (5), PolyT (5), PolyA (9), TG(3)	0,41	72072	0	0	100
21	3	190106030	190106104	75	74	Indel	0,57	60543	0	63	99,9
22	4	2233667	2233744	78	78	PolyA (6)	0,26	63882	0	0	100
23	4	7780541	7780637	97	97	PolyG (6), PolyT (5), PolyA (5)	0,42	79443	0	0	100
24	4	15688604	15688681	78	78	Nd.	0,29	63882	0	0	100
25	4	56236521	56236586	66	62	PolyA (5), indel	0,36	50778	0	0	100
26	4	102839244	102839314	71	69	PolyA (5)	0,46	56511	0	0	100
27	4	164446743	164446804	62	62	PolyA (7), indel	0,27	50778	0	0	100
28	5	1882081	1882158	78	75	Nd.	0,78	61425	0	0	100
29	5	14769061	14769144	84	84	GT(3), CCA (3)	0,62	68796	0	0	100
30	5	41069808	41069871	64	64	Nd.	0,39	52416	0	0	100
31	5	74077114	74077196	83	83	PolyA (6), indel	0,3	67977	0	0	100
32	5	147475343	147475409	67	67	PolyT (5)	0,37	54873	0	0	100

Amplikon	Chromosom	Początek amplikonu	Koniec amplikonu	Rozmiar analizowanego fragmentu	Nukleotydy w regionach o wysokim poziomie ufności	Materiał genetyczny amplikonu	Zawartość GC	Prawidłowe rozpoznania	Nieprawidłowe rozpoznania	Brak rozpoznań	Odsetek prawidłowo rozpoznanych nukleotydów
33	5	149323731	149323821	91	91	CT(4), AG(3)	0,55	74529	0	0	100
34	5	155662213	155662287	75	75	Indel	0,43	61425	0	0	100
35	6	6318713	6318814	102	102	PolyG(6)	0,68	83538	0	0	100
36	6	24949983	24950074	92	92	Indel	0,63	75348	0	0	100
37	6	31084900	31084999	100	94	GCT(5), indel	0,61	76608	0	378	99,5
38	6	32147987	32148084	98	98	PolyT(5), TCT(3), CTT(3)	0,55	80262	0	0	100
39	6	32986864	32986958	95	95	Indel	0,53	77805	0	0	100
40	6	33408498	33408583	86	86	PolyC(6)	0,7	70434	0	0	100
41	6	41647401	41647495	95	94	PolyG(5), indel	0,61	76986	0	0	100
42	6	112435865	112435955	91	91	PolyA(5)	0,44	74529	0	0	100
43	7	22202076	22202148	73	73	Nd.	0,44	59787	0	0	100
44	7	66276100	66276187	88	88	Indel	0,35	72072	0	0	100
45	7	77365735	77365821	87	87	PolyA(7), AG(4)	0,26	71253	0	0	100
46	7	110939946	110940030	85	85	Indel	0,38	69615	0	0	100
47	7	128533468	128533557	90	90	PolyG(5), indel	0,62	73710	0	0	100
48	7	149503875	149503965	91	91	PolyG(6), PolyC(6), indel	0,71	74529	0	0	100
49	7	154404519	154404599	81	66	Nd.	0,31	54054	0	0	100
50	7	156476507	156476599	93	93	Indel	0,35	76167	0	0	100
51	8	1817312	1817394	83	83	Nd.	0,42	67977	0	0	100

Amplikon	Chromosom	Początek amplikonu	Koniec amplikonu	Rozmiar analizowanego fragmentu	Nukleotydy w regionach o wysokim poziomie ufności	Materiał genetyczny amplikonu	Zawartość GC	Prawidłowe rozpoznania	Nieprawidłowe rozpoznania	Brak rozpoznań	Odsetek prawidłowo rozpoznanych nukleotydów
52	8	24811020	24811109	90	89	PolyG (7), CTC(4), indel	0,61	72171	0	720	99,0
53	8	76518625	76518691	67	67	Indel	0,3	54873	0	0	100
54	9	103054909	103055006	98	98	PolyG (6)	0,67	80262	0	0	100
55	9	105586150	105586214	65	65	Indel	0,32	53235	0	0	100
56	9	107620823	107620918	96	96	Nd.	0,49	78624	0	0	100
57	9	123769149	123769231	83	83	AT(3)	0,37	67977	0	0	100
58	9	138995345	138995441	97	97	PolyC (6), indel	0,68	79443	0	0	100
59	10	5987120	5987198	79	78	PolyG (5), indel	0,47	63882	0	0	100
60	10	11784629	11784726	98	91	GC(3)	0,87	74529	0	0	100
61	10	27317777	27317855	79	79	PolyT (5)	0,3	64701	0	0	100
62	10	33018351	33018440	90	90	PolyA (5), PolyT (5)	0,2	73710	0	0	100
63	10	45084159	45084253	95	95	Indel	0,35	77805	0	0	100
64	10	55892599	55892687	89	88	AC(11), indel	0,42	71747	0	325	99,5
65	10	101611250	101611329	80	80	Nd.	0,49	65520	0	0	100
66	10	118351373	118351453	81	81	Nd.	0,51	66339	0	0	100
67	11	8159816	8159912	97	96	Nd.	0,45	78624	0	0	100
68	11	30177648	30177717	70	70	Indel	0,46	57330	0	0	100
69	11	47470345	47470444	100	100	Nd.	0,65	81900	0	0	100
70	11	59837679	59837740	62	62	Indel	0,37	50778	0	0	100
71	11	64418856	64418957	102	102	Nd.	0,59	83538	0	0	100
72	11	93529612	93529684	73	73	PolyA (5)	0,4	59787	0	0	100

Amplikon	Chromosom	Początek amplikonu	Koniec amplikonu	Rozmiar analizowanego fragmentu	Nukleotydy w regionach o wysokim poziomie ufności	Materiał genetyczny amplikonu	Zawartość GC	Prawidłowe rozpoznania	Nieprawidłowe rozpoznania	Brak rozpoznań	Odsetek prawidłowo rozpoznanych nukleotydów
73	11	101347052	101347136	85	85	Nd.	0,42	69615	0	0	100
74	11	102477336	102477426	91	91	PolyG (6)	0,55	74529	0	0	100
75	11	118406285	118406369	85	85	Indel	0,53	69615	0	0	100
76	11	120357801	120357885	85	85	PolyA (5), CA(3), indel	0,34	69615	0	0	100
77	11	125769313	125769397	85	85	GA(3)	0,52	69615	0	0	100
78	12	2834770	2834853	84	84	PolyC (5), indel	0,52	68796	0	0	100
79	12	26811004	26811096	93	93	PolyA (7), AC(4)	0,33	76167	0	0	100
80	12	30881766	30881846	81	81	Nd.	0,49	66339	0	0	100
81	12	88474105	88474175	71	71	PolyA (6)	0,35	58149	0	0	100
82	12	120966872	120966966	95	95	PolyG (5)	0,68	77805	0	0	100
83	13	24167504	24167576	73	73	Nd.	0,52	59787	0	0	100
84	13	25816961	25817049	89	88	PolyA (5), PolyT (7), PolyA (7), indel	0,22	72072	0	0	100
85	13	44880112	44880200	89	89	Indel	0,49	72891	0	0	100
86	13	77665218	77665294	77	77	Indel	0,39	63063	0	0	100
87	14	31619327	31619393	67	67	GA(3), TA (3)	0,39	54873	0	0	100
88	14	39517884	39517966	83	83	Nd.	0,25	67977	0	0	100
89	14	46958962	46959034	73	72	PolyT (5), indel	0,19	58642	0	326	99,4
90	14	58050030	58050110	81	81	Indel	0,38	66339	0	0	100
91	14	82390559	82390649	91	91	Indel	0,35	74529	0	0	100
92	14	92549544	92549609	66	66	PolyA (5)	0,41	54054	0	0	100

Amplikon	Chromosom	Początek amplikonu	Koniec amplikonu	Rozmiar analizowanego fragmentu	Nukleotydy w regionach o wysokim poziomie ufności	Materiał genetyczny amplikonu	Zawartość GC	Prawidłowe rozpoznania	Nieprawidłowe rozpoznania	Brak rozpoznań	Odsetek prawidłowo rozpoznanych nukleotydów
93	14	102808496	102808589	94	94	Indel	0,62	76986	0	0	100
94	15	43170751	43170848	98	96	PolyC (5)	0,45	78624	0	0	100
95	15	63446149	63446216	68	68	Indel	0,25	55692	0	0	100
96	15	77879807	77879901	95	93	PolyG (5), indel	0,68	76167	0	0	100
97	15	81625334	81625428	95	95	PolyT (6)	0,43	77805	0	0	100
98	15	85438263	85438334	72	71	Indel	0,65	58149	0	0	100
99	15	89817413	89817503	91	91	Nd.	0,36	74529	0	0	100
100	15	89864274	89864343	70	70	Indel	0,56	57330	0	0	100
101	16	1894910	1894972	63	63	Nd.	0,27	51597	0	0	100
102	16	28997904	28997998	95	95	PolyC (5)	0,67	77805	0	0	100
103	16	53682908	53682994	87	87	TA(3)	0,41	71253	0	0	100
104	16	57954406	57954509	104	104	PolyC (5)	0,67	85176	0	0	100
105	16	85706375	85706465	91	91	PolyT (5), indel	0,37	74529	0	0	100
106	17	3563920	3564008	89	89	GC(3)	0,64	72891	0	0	100
107	17	3594191	3594277	87	87	PolyC (5), indel	0,67	71247	0	6	100
108	17	3970090	3970180	91	91	Indel	0,46	74529	0	0	100
109	17	16084945	16085037	93	93	Indel	0,26	76167	0	0	100
110	17	33998759	33998849	91	89	PolyT (5)	0,54	72891	0	0	100
111	17	39589691	39589774	84	82	PolyA (13), indel (x2)	0,29	66343	27	788	98,8
112	17	41244394	41244484	91	91	PolyA (5)	0,34	74529	0	0	100
113	17	45438866	45438957	92	92	PolyA (7), AT(3), AT (4), AT(4), indel	0,26	75348	0	0	100



Amplikon	Chromosom	Początek amplikonu	Koniec amplikonu	Rozmiar analizowanego fragmentu	Nukleotydy w regionach o wysokim poziomie ufności	Materiał genetyczny amplikonu	Zawartość GC	Prawidłowe rozpoznania	Nieprawidłowe rozpoznania	Brak rozpoznań	Odsetek prawidłowo rozpoznanych nukleotydów
114	17	61502432	61502510	79	79	Indel	0,41	64413	0	288	99,6
115	17	64023582	64023667	86	86	PolyT (7)	0,22	70434	0	0	100
116	17	72308237	72308320	84	84	GAG(3)	0,62	68796	0	0	100
117	18	2616456	2616522	67	67	GA(3)	0,31	54873	0	0	100
118	18	6980478	6980568	91	91	Nd.	0,37	74529	0	0	100
119	18	9888026	9888094	69	69	PolyA (6), TG(3)	0,43	56511	0	0	100
120	18	38836999	38837073	75	75	PolyA (5), indel	0,37	61425	0	0	100
121	18	47405382	47405462	81	81	CTC(3), indel	0,47	66339	0	0	100
122	18	54815665	54815749	85	85	CT(3), indel	0,45	69615	0	0	100
123	18	59773996	59774060	65	65	Nd.	0,48	53235	0	0	100
124	19	625143	625241	99	99	Nd.	0,59	81081	0	0	100
125	19	18121418	18121491	74	74	Nd.	0,68	60605	1	0	100
126	19	18186574	18186643	70	70	Nd.	0,64	57330	0	0	100
127	20	746056	746149	94	94	Nd.	0,61	76986	0	0	100
128	20	10633195	10633276	82	82	AC(3)	0,59	67158	0	0	100
129	20	17705633	17705708	76	76	CT(3)	0,58	62244	0	0	100
130	20	21766821	21766890	70	70	GT(3), TG (4), indel	0,46	57330	0	0	100
131	20	25278421	25278521	101	101	Indel	0,63	82719	0	0	100
132	20	50897302	50897368	67	67	Indel	0,36	54873	0	0	100
133	20	62331904	62331994	91	88	PolyG (6)	0,73	72072	0	0	100
134	20	62690860	62690946	87	87	Indel	0,57	71253	0	0	100
135	21	30300823	30300888	66	66	Indel	0,35	54054	0	0	100

Amplikon	Chromosom	Początek amplikonu	Koniec amplikonu	Rozmiar analizowanego fragmentu	Nukleotydy w regionach o wysokim poziomie ufności	Materiał genetyczny amplikonu	Zawartość GC	Prawidłowe rozpoznania	Nieprawidłowe rozpoznania	Brak rozpoznań	Odsetek prawidłowo rozpoznanych nukleotydów
136	21	33694176	33694273	98	98	PolyT (6), CA(3)	0,54	80262	0	0	100
137	21	36710706	36710792	87	87	GT(3), indel	0,39	71253	0	0	100
138	21	46644924	46644992	69	69	PolyA (6), AG(3), indel	0,32	56439	0	72	99,9
139	21	46705575	46705664	90	90	PolyT (5), PolyA (6)	0,5	73710	0	0	100
140	22	25750774	25750873	100	100	Indel	0,63	81900	0	0	100
141	22	32439233	32439329	97	97	Nd.	0,68	79443	0	0	100
142	22	37409844	37409940	97	97	Indel	0,46	79443	0	0	100
143	22	37637596	37637694	99	99	Nd.	0,6	81081	0	0	100
144	22	47081347	47081438	92	92	Indel	0,66	75348	0	0	100
145	X	15870424	15870492	69	69	PolyT (5)	0,26	56511	0	0	100
146	X	135288543	135288611	69	69	PolyC (5)	0,62	56511	0	0	100
147	X	135290777	135290847	71	71	Nd.	0,52	58149	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.

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, natomiast 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 hg19 genomu ludzkiego. Wyniki oznaczeń dokładności zawiera [Tabela 12](#).

Tabela 12 Zgodność wariantów linii zarodkowych próbki NA12878 z bazą danych NIST

Próbka	Liczba amplikonów	Średni odsetek rozpoznań	TP	FN	TN	FP	PPA	NPA	OPA
NA12878	133	>99,9	6552	1	610470	0	>99,9	100	>99,9

W oparciu o dane dostarczone z dziewięciu przebiegów analizy wariantów linii zarodkowej, można stwierdzić, że aparat NextSeq 550Dx Instrument jest w stanie powtarzalnie sekwencjonować:

- ▶ Zawartość GC  $\geq 19\%$  (wszystkie rozpoznane nukleotydy w 819 sekwencjonowanych amplikonach z 19% zawartością GC zostały rozpoznane prawidłowo przy odsetku braku rozpoznań na poziomie 0,6%).
- ▶ Zawartość GC  $\leq 87\%$  (wszystkie rozpoznane nukleotydy w 819 sekwencjonowanych amplikonach z 87% zawartością GC zostały rozpoznane prawidłowo przy braku nierozpoznanych nukleotydów).
- ▶ Łańcuchy PolyA o długości  $\leq 9$  (wszystkie rozpoznane nukleotydy w 819 sekwencjonowanych amplikonach zawierających powtórzenie PolyA złożone z dziewięciu nukleotydów zostały rozpoznane prawidłowo przy braku nierozpoznanych powtórzeń).
- ▶ Łańcuchy PolyT o długości  $\leq 10$  (wszystkie rozpoznane nukleotydy w 819 sekwencjonowanych amplikonach zawierających powtórzenie PolyT złożone z dziesięciu nukleotydów zostały rozpoznane prawidłowo przy braku nierozpoznanych powtórzeń).
- ▶ Łańcuchy PolyG o długości  $\leq 7$  (wszystkie rozpoznane nukleotydy w 819 sekwencjonowanych amplikonach zawierających powtórzenie PolyG złożone z siedmiu nukleotydów zostały rozpoznane prawidłowo przy odsetku braku rozpoznań na poziomie 1,0%).
- ▶ Łańcuchy PolyC o długości  $\leq 6$  (wszystkie rozpoznane nukleotydy w 2457 sekwencjonowanych amplikonach zawierających powtórzenie PolyC złożone z sześciu nukleotydów zostały rozpoznane prawidłowo przy braku nierozpoznanych powtórzeń).
- ▶ Powtórzenia dinukleotydowe o długości  $\leq 11 \times$  (wszystkie rozpoznane nukleotydy w 819 sekwencjonowanych amplikonach z 11-krotnym powtórzeniem dinukleotydu zostały rozpoznane prawidłowo przy odsetku braku rozpoznań na poziomie 0,5%).
- ▶ Powtórzenia trinukleotydowe o długości  $\leq 5 \times$  (wszystkie rozpoznane nukleotydy w 819 sekwencjonowanych amplikonach z 5-krotnym powtórzeniem trinukleotydu zostały rozpoznane prawidłowo przy odsetku braku rozpoznań na poziomie 0,5%).
- ▶ Insercje o długości  $\leq 24$  (66343 spośród 66370 rozpoznanych nukleotydów w 819 sekwencjonowanych amplikonach zawierających insercję 24-nukleotydową zostało zidentyfikowanych poprawnie przy odsetku braku rozpoznań na poziomie 1,2%; w regionach zawierających insercję 24-nukleotydową nie stwierdzano nieprawidłowych rozpoznań).
- ▶ Delecje o długości  $\leq 25$  (wszystkie rozpoznane nukleotydy w 2457 sekwencjonowanych amplikonach zawierających delecję 25-nukleotydową zostały zidentyfikowane poprawnie przy braku nierozpoznanych powtórzeń).

### Warianty somatyczne

Opisywane tu badanie miało na celu dokonanie oceny dokładności rozpoznawania wariantów przez moduł do wykrywania wariantów somatycznych aparatu NextSeq 550Dx Instrument przy użyciu zestawu odczynników na 300 cykli NextSeq 550Dx High Output Reagent Kit v2 (300 cycles).

W tym badaniu użyto oznaczenia reprezentatywnego w celu rozpoznania różnych genów obejmujących 12 588 nukleotydów (150 amplikonów) znajdujących się w 23 różnych chromosomach. Materiał DNA Platinum Genome wyekstrahowano z poddanych obróbce bloków FFPE, tworząc sześć unikalnych próbek ocenianych w badaniu.

DNA z próbki linii komórkowej GM12877 rozcieńczono DNA z próbki linii komórkowej GM12878, uzyskując próbki GM12877-D5 i GM12877-D7 będące zestawami unikalnych wariantów heterozygotycznych o częstościach występowania wariantów zbliżonych do 5% i 7%. W podobny sposób DNA z próbki linii komórkowej GM12878 rozcieńczono DNA z próbki linii komórkowej GM12877, uzyskując próbki GM12878-D5 i GM12878-D7. Każdą z próbek oznaczano w trzech powtórzeniach, z wyjątkiem próbek rozcieńczonych, które oznaczano w sześciu powtórzeniach. Wykonano w sumie dziewięć przebiegów z użyciem trzech aparatów do sekwencjonowania, trzech serii odczynników z udziałem trzech operatorów w okresie pięciu dni rozpoczynania oznaczeń. Dokładność została określona dla wariantów SNV, insercji i delecji przez porównanie wyników badania z dobrze scharakteryzowaną, złożoną metodą referencyjną — bazą danych Platinum Genomes w wersji 2016-1.0. O ile nie podano inaczej, regiony o wysokim poziomie ufności zostały zdefiniowane w oparciu o ww. metodę referencyjną.

Tabela 13 Podsumowanie zgodności wariantów somatycznych

Kryteria	Obserwacje łącznie <sup>1</sup>	Wynik dla obserwacji <sup>2</sup>	Wynik dla przebiegu <sup>3</sup>
PPA dla SNV	378	98,9	99,9
PPA dla insercji	378	96,9	99,9
PPA dla delecji	378	97,1	99,9
NPA	378	>99,9	>99,9
OPA	378	>99,9	>99,9

<sup>1</sup> Obliczone jako liczba próbek na przebieg (42) × liczba przebiegów (9) = 378.

<sup>2</sup> Najniższa obserwowana wartość dla powtórzenia próbki we wszystkich dziewięciu przebiegach.

<sup>3</sup> Najniższa wartość w połączonej analizie danych z każdego przebiegu.

Tabela 14 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 hg19 genomu ludzkiego.

Tabela 14 Zgodność wariantów somatycznych w zależności od próbki

Próbka	Średni odsetek rozpoznań	Oczekiwane	TP	FN	Brak rozpoznania wariantu	TN	FP	PPA	NPA	OPA
GM12877	98,7	2052	2025	0	27	318682	15	100	>99,9	>99,9
GM12878	98,8	3645	3564	0	81	317645	0	100	100	100
GM12879	99,8	2592	2538	0	54	323614	2	100	>99,9	>99,9
GM12884	99,8	3078	3024	0	54	322038	5	100	>99,9	>99,9
GM12885	99,8	3294	3213	0	81	322121	0	100	100	100
GM12888	99,8	2916	2889	0	27	323048	2	100	>99,9	>99,9
GM12877-D5	99,8	9288	8930	0	358	630621	0	100	100	100
GM12877-D7	99,7	9288	9032	0	256	629719	0	100	100	100
GM12878-D5	99,5	9288	8699	42	547	628582	0	99,5	100	>99,9
GM12878-D7	99,7	9288	9108	0	180	629803	0	100	100	100

Tabela 15 zawiera dane z badania prezentowane dla poszczególnych próbek, przy czym wyniki wariantów zostały porównane z dobrze scharakteryzowaną, złożoną metodą referencyjną. Wykrywalność została oceniona oddzielnie dla każdego rodzaju wariantu – SNV, insercji i delecji. Pozycje referencyjne zostały wykluczone.

Tabela 15 Zgodność wariantów somatycznych w zależności od rodzaju wariantu w próbce

Próbka	Warianty SNV			Insercje			Delecje		
	Oczekiwane	TP	FN	Oczekiwane	TP	FN	Oczekiwane	TP	FN
GM12877	999	999	0	567	567	0	486	459	0
GM12878	2457	2457	0	540	513	0	648	594	0
GM12879	1539	1539	0	567	540	0	486	459	0
GM12884	1836	1836	0	675	648	0	567	540	0
GM12885	2025	2025	0	675	648	0	594	540	0
GM12888	1782	1782	0	621	621	0	513	486	0
GM12877-D5	5454	5392	0	1782	1647	0	2052	1891	0
GM12877-D7	5454	5406	0	1782	1728	0	2052	1898	0
GM12878-D5	5454	5192	28	1782	1651	9	2052	1856	5
GM12878-D7	5454	5445	0	1782	1719	0	2052	1944	0

Dziesięć próbek poddano dalszej analizie pod kątem rozpoznawania małych insercji i delecji (polimorfizmów typu indel) (Tabela 16). 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.

Tabela 16 Podsumowanie wykrywania polimorfizmów typu indel wariantów somatycznych

Typ wariantu	Warianty oczekiwane	TP	FN	Brak rozpoznania wariantu	PPA
Insercja	10773	10282	9	482	99,2
Delecja	11502	10667	5	830	>99,9

Liczba 150 amplikonów była zaprojektowana tak, aby pokryć zróżnicowaną zawartość genomową. Zawartość GC w amplikonach mieściła się w zakresie 0,19–0,87%. Amplikony charakteryzowały się również pewnym zakresem powtórzeń pojedynczych nukleotydów (np. PolyA, PolyT), dinukleotydów i trinukleotydów. Dane skompilowano w oparciu o poszczególne amplikony (Tabela 17) w celu oznaczenia wpływu zawartości genomowej na odsetek prawidłowych rozpoznań. Odsetek prawidłowych rozpoznań obejmuje rozpoznania wariantu i wzorca i jest mniejszy od 100% w przypadku nieprawidłowych rozpoznań lub braku rozpoznań.

Tabela 17 Dokładność oznaczeń wariantów somatycznych na poziomie amplikonów

Amp-likon	Chromo-som	Początek amplikonu	Koniec amplikonu	Rozmiar analizowanego fragmentu	Nukleotydy w regionach o wysokim poziomie ufności	Materiał genetyczny amplikonu	Zawartość GC	Prawidłowe rozpoznania	Nieprawidłowe rozpoznania	Brak rozpoznań	Odsetek prawidłowo rozpoznanych nukleotydów
1	1	36450499	36450591	93	93	Indel	0,22	35066	0	88	99,7
2	1	109465122	109465200	79	79	PolyA (5), PolyC (5), indel	0,38	29827	0	35	99,9
3	1	218353867	218353957	91	91	Indel	0,4	34202	0	283	99,2
4	1	223906657	223906748	92	92	Indel	0,49	34613	0	163	99,5
5	1	228526602	228526682	81	81	PolyG (5)	0,69	30571	0	47	99,8
6	1	236372039	236372108	70	70	PolyT (10), indel	0,39	26452	0	8	100,0
7	1	247812041	247812128	88	88	PolyA (5), CT(3), TAA (3), indel	0,27	33148	0	116	99,7
8	2	55862774	55862863	90	90	Indel	0,28	33928	0	92	99,7
9	2	87003930	87004009	80	80	Indel	0,38	30218	0	22	99,9
10	2	177016721	177016805	85	81	Nd.	0,65	30616	0	2	>99,9
11	2	186625727	186625801	75	75	PolyA (8)	0,35	28017	0	499	98,3
12	2	190323504	190323591	88	88	PolyT (5)	0,42	33207	0	57	99,8
13	2	200796740	200796826	87	87	PolyT (5), indel	0,31	32524	9	718	97,8
14	2	212245049	212245139	91	91	PolyT (5), PolyA (6), indel	0,3	33972	0	456	98,7

Amp-likon	Chromosom	Początek amplikonu	Koniec amplikonu	Rozmiar analizowanego fragmentu	Nukleotydy w regionach o wysokim poziomie ufności	Materiał genetyczny amplikonu	Zawartość GC	Prawidłowe rozpoznania	Nieprawidłowe rozpoznania	Brak rozpoznań	Odsetek prawidłowo rozpoznanych nukleotydów
15	2	228147052	228147144	93	93	Nd.	0,43	35051	0	103	99,7
16	2	235016350	235016422	73	73	PolyT (5), indel	0,42	27459	0	136	99,5
17	3	4466229	4466321	93	93	AT(3), indel	0,27	34534	0	620	98,2
18	3	46620561	46620643	83	83	Nd.	0,43	31339	0	44	99,9
19	3	49851331	49851400	70	70	CT(3), indel	0,49	26373	0	87	99,7
20	3	189713161	189713248	88	88	PolyA (5), PolyT (5), PolyA (9), TG(3)	0,41	32829	0	857	97,5
21	3	190106030	190106104	75	74	Indel	0,57	27925	0	47	99,8
22	4	2233667	2233744	78	78	PolyA (6)	0,26	29327	4	162	99,4
23	4	7780541	7780637	97	97	PolyG (6), PolyT (5), PolyA (5)	0,42	36585	0	117	99,7
24	4	15688604	15688681	78	78	Nd.	0,29	29427	0	57	99,8
25	4	56236521	56236586	66	62	PolyA (5), indel	0,36	23356	5	75	99,7
26	4	102839244	102839314	71	69	PolyA (5)	0,46	25942	0	140	99,5
27	4	164446743	164446804	62	62	PolyA (7), indel	0,27	22944	0	560	97,6
28	5	1882081	1882158	78	75	Nd.	0,78	28299	0	53	99,8
29	5	14769061	14769144	84	84	GT(3), CCA (3)	0,62	31658	0	94	99,7
30	5	41069808	41069871	64	64	Nd.	0,39	24120	0	72	99,7
31	5	74077114	74077196	83	83	PolyA (6), indel	0,3	31297	0	77	99,8
32	5	147475343	147475409	67	67	PolyT (5)	0,37	25277	0	55	99,8

Amp-likon	Chromosom	Początek amplikonu	Koniec amplikonu	Rozmiar analizowanego fragmentu	Nukleotydy w regionach o wysokim poziomie ufności	Materiał genetyczny amplikonu	Zawartość GC	Prawidłowe rozpoznania	Nieprawidłowe rozpoznania	Brak rozpoznań	Odsetek prawidłowo rozpoznanych nukleotydów
33	5	149323731	149323821	91	91	CT(4), AG (3)	0,55	34308	0	90	99,7
34	5	155662213	155662287	75	75	Indel	0,43	28266	0	163	99,4
35	6	6318713	6318814	102	102	PolyG (6)	0,68	38489	0	67	99,8
36	6	24949983	24950074	92	92	Indel	0,63	34730	0	46	99,9
37	6	31084900	31084999	100	94	GCT(5), indel	0,61	35057	0	483	98,6
38	6	32147987	32148084	98	98	PolyT (5), TCT(3), CTT(3)	0,55	36647	0	406	98,9
39	6	32986864	32986958	95	95	Indel	0,53	35681	0	238	99,3
40	6	33408498	33408583	86	86	PolyC (6)	0,7	32438	0	70	99,8
41	6	41647401	41647495	95	94	PolyG (5), indel	0,61	35441	0	91	99,7
42	6	112435865	112435955	91	91	PolyA (5)	0,44	34354	0	44	99,9
43	7	22202076	22202148	73	73	Nd.	0,44	27575	0	28	99,9
44	7	66276100	66276187	88	88	Indel	0,35	33060	0	213	99,4
45	7	77365735	77365821	87	87	PolyA (7), AG(4)	0,26	32423	0	489	98,5
46	7	110939946	110940030	85	85	Indel	0,38	32074	0	56	99,8
47	7	128533468	128533557	90	90	PolyG (5), indel	0,62	33791	0	281	99,2
48	7	149503875	149503965	91	91	PolyG (6), PolyC (6), indel	0,71	34316	0	82	99,8
49	7	154404519	154404599	81	66	Nd.	0,31	24901	0	47	99,8
50	7	156476507	156476599	93	93	Indel	0,35	35067	0	87	99,8
51	8	1817312	1817394	83	83	Nd.	0,42	31365	0	9	>99,9



Amp-likon	Chromosom	Początek amplikonu	Koniec amplikonu	Rozmiar analizowanego fragmentu	Nukleotydy w regionach o wysokim poziomie ufności	Materiał genetyczny amplikonu	Zawartość GC	Prawidłowe rozpoznania	Nieprawidłowe rozpoznania	Brak rozpoznań	Odsetek prawidłowo rozpoznanych nukleotydów
52	8	24811020	24811109	90	89	PolyG(7), CTC(4), indel	0,61	32781	0	890	97,4
53	8	76518625	76518691	67	67	Indel	0,3	25228	0	146	99,4
54	9	103054909	103055006	98	98	PolyG(6)	0,67	36968	0	76	99,8
55	9	105586150	105586214	65	65	Indel	0,32	24472	0	100	99,6
56	9	107620823	107620918	96	96	Nd.	0,49	36203	0	85	99,8
57	9	123769149	123769231	83	83	AT(3)	0,37	31329	0	45	99,9
58	9	138995345	138995441	97	97	PolyC(6), indel	0,68	36472	0	201	99,5
59	10	5987120	5987198	79	78	PolyG(5), indel	0,47	29473	0	11	>99,9
60	10	11784629	11784726	98	91	GC(3)	0,87	34188	0	213	99,4
61	10	27317777	27317855	79	79	PolyT(5)	0,3	29843	0	19	99,9
62	10	33018351	33018440	90	90	PolyA(5), PolyT(5)	0,2	33968	0	68	99,8
63	10	45084159	45084253	95	95	Indel	0,35	35829	0	81	99,8
64	10	55892599	55892687	89	88	AC(11), indel	0,42	32098	88	2048	93,8
65	10	101611250	101611329	80	80	Nd.	0,49	30217	0	28	99,9
66	10	118351373	118351453	81	81	Nd.	0,51	30531	0	96	99,7
67	11	8159816	8159912	97	96	Nd.	0,45	36105	0	192	99,5
68	11	30177648	30177717	70	70	Indel	0,46	26318	0	153	99,4
69	11	47470345	47470444	100	100	Nd.	0,65	37785	0	24	99,9
70	11	59837679	59837740	62	62	Indel	0,37	23368	0	68	99,7
71	11	64418856	64418957	102	102	Nd.	0,59	38546	0	10	>99,9
72	11	93529612	93529684	73	73	PolyA(5)	0,4	27516	0	78	99,7

Amp-likon	Chromosom	Początek amplikonu	Koniec amplikonu	Rozmiar analizowanego fragmentu	Nukleotydy w regionach o wysokim poziomie ufności	Materiał genetyczny amplikonu	Zawartość GC	Prawidłowe rozpoznania	Nieprawidłowe rozpoznania	Brak rozpoznań	Odsetek prawidłowo rozpoznanych nukleotydów
73	11	101347052	101347136	85	85	Nd.	0,42	32083	0	48	99,9
74	11	102477336	102477426	91	91	PolyG (6)	0,55	34047	0	369	98,9
75	11	118406285	118406369	85	85	Indel	0,53	32065	0	74	99,8
76	11	120357801	120357885	85	85	PolyA (5), CA(3), indel	0,34	32083	0	47	99,9
77	11	125769313	125769397	85	85	GA(3)	0,52	32103	0	27	99,9
78	12	2834770	2834853	84	84	PolyC (5), indel	0,52	31645	16	525	98,3
79	12	26811004	26811096	93	93	PolyA (7), AC(4)	0,33	34824	0	330	99,1
80	12	30881766	30881846	81	81	Nd.	0,49	30497	0	121	99,6
81	12	88474105	88474175	71	71	PolyA (6)	0,35	26773	0	65	99,8
82	12	120966872	120966966	95	95	PolyG (5)	0,68	35830	9	72	99,8
83	13	24167504	24167576	73	73	Nd.	0,52	27498	0	114	99,6
84	13	25816961	25817049	89	88	PolyA (5), PolyT (7), PolyA (7), indel	0,22	32824	0	566	98,3
85	13	44880112	44880200	89	89	Indel	0,49	33574	0	77	99,8
86	13	77665218	77665294	77	77	Indel	0,39	29075	0	31	99,9
87	14	31619327	31619393	67	67	GA(3), TA (3)	0,39	25313	0	13	99,9
88	14	39517884	39517966	83	83	Nd.	0,25	31360	0	22	99,9
89	14	46958962	46959034	73	72	PolyT (5), indel	0,19	26499	0	717	97,4
90	14	58050030	58050110	81	81	Indel	0,38	30494	0	133	99,6
91	14	82390559	82390649	91	91	Indel	0,35	34313	0	86	99,7
92	14	92549544	92549609	66	66	PolyA (5)	0,41	24555	0	1527	94,1

Amp-likon	Chromosom	Początek amplikonu	Koniec amplikonu	Rozmiar analizowanego fragmentu	Nukleotydy w regionach o wysokim poziomie ufności	Materiał genetyczny amplikonu	Zawartość GC	Prawidłowe rozpoznania	Nieprawidłowe rozpoznania	Brak rozpoznań	Odsetek prawidłowo rozpoznanych nukleotydów
93	14	102808496	102808589	94	94	Indel	0,62	35472	0	69	99,8
94	15	43170751	43170848	98	96	PolyC (5)	0,45	36264	0	24	99,9
95	15	63446149	63446216	68	68	Indel	0,25	25667	0	37	99,9
96	15	77879807	77879901	95	93	PolyG (5), indel	0,68	34745	0	432	98,8
97	15	81625334	81625428	95	95	PolyT (6)	0,43	35870	0	40	99,9
98	15	85438263	85438334	72	71	Indel	0,65	26762	0	76	99,7
99	15	89817413	89817503	91	91	Nd.	0,36	34286	0	112	99,7
100	15	89864274	89864343	70	70	Indel	0,56	26449	0	11	>99,9
101	16	1894910	1894972	63	63	Nd.	0,27	23809	0	5	>99,9
102	16	28997904	28997998	95	95	PolyC (5)	0,67	35860	0	50	99,9
103	16	53682908	53682994	87	87	TA(3)	0,41	32835	0	60	99,8
104	16	57954406	57954509	104	104	PolyC (5)	0,67	39177	0	144	99,6
105	16	85706375	85706465	91	91	PolyT (5), indel	0,37	34075	0	323	99,1
106	17	3563920	3564008	89	89	GC(3)	0,64	33632	0	11	>99,9
107	17	3594191	3594277	87	87	PolyC (5), indel	0,67	32752	0	134	99,6
108	17	3970090	3970180	91	91	Indel	0,46	34343	0	82	99,8
109	17	16084945	16085037	93	93	Indel	0,26	35077	0	78	99,8
110	17	33998759	33998849	91	89	PolyT (5)	0,54	33553	0	89	99,7
111	17	39589691	39589774	84	82	PolyA (13), indel (×2)	0,29	30554	53	2296	92,9
112	17	41244394	41244484	91	91	PolyA (5)	0,34	34360	0	38	99,9
113	17	45438866	45438957	92	92	PolyA (7), AT(3), AT (4), AT(4), indel	0,26	34367	0	418	98,8

Amp-likon	Chromosom	Początek amplikonu	Koniec amplikonu	Rozmiar analizowanego fragmentu	Nukleotydy w regionach o wysokim poziomie ufności	Materiał genetyczny amplikonu	Zawartość GC	Prawidłowe rozpoznania	Nieprawidłowe rozpoznania	Brak rozpoznań	Odsetek prawidłowo rozpoznanych nukleotydów
114	17	61502432	61502510	79	79	Indel	0,41	29751	0	119	99,6
115	17	64023582	64023667	86	86	PolyT (7)	0,22	32176	0	340	99,0
116	17	72308237	72308320	84	84	GAG(3)	0,62	31604	7	141	99,5
117	18	2616456	2616522	67	67	GA(3)	0,31	25273	8	45	99,8
118	18	6980478	6980568	91	91	Nd.	0,37	34386	0	12	>99,9
119	18	9888026	9888094	69	69	PolyA (6), TG(3)	0,43	25692	0	399	98,5
120	18	38836999	38837073	75	75	PolyA (5), indel	0,37	27923	0	893	96,9
121	18	47405382	47405462	81	81	CTC(3), indel	0,47	30598	0	20	99,9
122	18	54815665	54815749	85	85	CT(3), indel	0,45	31969	0	161	99,5
123	18	59773996	59774060	65	65	Nd.	0,48	24531	0	48	99,8
124	19	625143	625241	99	99	Nd.	0,59	37298	0	124	99,7
125	19	18121418	18121491	74	74	Nd.	0,68	27881	0	109	99,6
126	19	18186574	18186643	70	70	Nd.	0,64	26442	0	26	99,9
127	20	746056	746149	94	94	Nd.	0,61	35501	0	31	99,9
128	20	10633195	10633276	82	82	AC(3)	0,59	30951	0	72	99,8
129	20	17705633	17705708	76	76	CT(3)	0,58	28686	0	42	99,9
130	20	21766821	21766890	70	70	GT(3), TG (4), indel	0,46	26372	0	88	99,7
131	20	25278421	25278521	101	101	Indel	0,63	38159	0	20	99,9
132	20	50897302	50897368	67	67	Indel	0,36	25188	0	544	97,9
133	20	62331904	62331994	91	88	PolyG (6)	0,73	32969	0	309	99,1
134	20	62690860	62690946	87	87	Indel	0,57	32818	0	77	99,8
135	21	30300823	30300888	66	66	Indel	0,35	24758	9	181	99,2

Amp-likon	Chromosom	Początek amplikonu	Koniec amplikonu	Rozmiar analizowanego fragmentu	Nukleotydy w regionach o wysokim poziomie ufności	Materiał genetyczny amplikonu	Zawartość GC	Prawidłowe rozpoznania	Nieprawidłowe rozpoznania	Brak rozpoznań	Odsetek prawidłowo rozpoznanych nukleotydów
136	21	33694176	33694273	98	98	PolyT (6), CA(3)	0,54	36902	0	160	99,6
137	21	36710706	36710792	87	87	GT(3), indel	0,39	32841	0	48	99,9
138	21	46644924	46644992	69	69	PolyA (6), AG(3), indel	0,32	25939	0	280	98,9
139	21	46705575	46705664	90	90	PolyT (5), PolyA (6)	0,5	33942	0	78	99,8
140	22	25750774	25750873	100	100	Indel	0,63	37733	0	86	99,8
141	22	32439233	32439329	97	97	Nd.	0,68	36617	0	49	99,9
142	22	37409844	37409940	97	97	Indel	0,46	36525	0	162	99,6
143	22	37637596	37637694	99	99	Nd.	0,6	37398	0	24	99,9
144	22	47081347	47081438	92	92	Indel	0,66	34754	0	22	99,9
145	X	15870424	15870492	69	69	PolyT (5)	0,26	26046	0	36	99,9
146	X	135288543	135288611	69	69	PolyC (5)	0,62	26019	0	63	99,8
147	X	135290777	135290847	71	71	Nd.	0,52	26780	0	58	99,8
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.

Wyniki sekwencjonowania próbki linii komórkowej GM12878 porównano z genotypem o wysokim poziomie ufności, ustalonym dla próbki NA12878 przez Narodowy Instytut Standaryzacji i Technologii (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, natomiast 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 hg19 genomu ludzkiego. Wyniki oznaczeń dokładności zawiera [Tabela 18](#).

Tabela 18 Zgodność wariantów somatycznych próbki linii komórkowej GM12878 z bazą danych NIST

Próbka	Liczba amplikonów	Średni odsetek rozpoznań	TP	FN	TN	FP	PPA	NPA	OPA
GM12878	133	98,8	2808	0	258488	0	100	100	100

W oparciu o dane dostarczone z dziewięciu przebiegów analizy wariantów somatycznych, można stwierdzić, że aparat NextSeq 550Dx Instrument jest w stanie powtarzalnie sekwencjonować:

- ▶ Zawartość GC  $\geq 19\%$  (wszystkie rozpoznane nukleotydy w 378 sekwencjonowanych amplikonach z 19% zawartością GC zostały rozpoznane prawidłowo przy odsetku braku rozpoznań na poziomie 2,6%).
- ▶ Zawartość GC  $\leq 87\%$  (wszystkie rozpoznane nukleotydy w 378 sekwencjonowanych amplikonach z 87% zawartością GC rozpoznane prawidłowo przy odsetku braku rozpoznań na poziomie 0,6%).
- ▶ Łańcuchy PolyA o długości  $\leq 9$  (wszystkie rozpoznane nukleotydy w 378 sekwencjonowanych amplikonach zawierających powtórzenie PolyA złożone z dziewięciu nukleotydów zostały rozpoznane prawidłowo przy odsetku braku rozpoznań na poziomie 2,5%).
- ▶ Łańcuchy PolyT o długości  $\leq 10$  (wszystkie rozpoznane nukleotydy w 378 sekwencjonowanych amplikonach zawierających powtórzenie PolyT złożone z dziesięciu nukleotydów zostały rozpoznane prawidłowo przy odsetku braku rozpoznań na poziomie 0,1%).
- ▶ Łańcuchy PolyG o długości  $\leq 6$  (wszystkie rozpoznane nukleotydy w 2268 sekwencjonowanych amplikonach zawierających powtórzenie PolyG złożone z sześciu nukleotydów zostały rozpoznane prawidłowo przy odsetku braku rozpoznań na poziomie 0,5%).
- ▶ Łańcuchy PolyC o długości  $\leq 6$  (wszystkie rozpoznane nukleotydy w 756 sekwencjonowanych amplikonach zawierających powtórzenie PolyC złożone z sześciu nukleotydów zostały rozpoznane prawidłowo przy odsetku braku rozpoznań na poziomie 0,4%).
- ▶ Powtórzenia dinukleotydy o długości  $\leq 4\times$  (wszystkie rozpoznane nukleotydy w 1890 sekwencjonowanych amplikonach z 4-krotnym powtórzeniem dinukleotydu zostały rozpoznane prawidłowo przy odsetku braku rozpoznań na poziomie 0,9%).
- ▶ Powtórzenia trinukleotydy o długości  $\leq 5\times$  (wszystkie zidentyfikowane nukleotydy w 378 sekwencjonowanych amplikonach z 5-krotnym powtórzeniem trinukleotydu zostały rozpoznane prawidłowo przy odsetku braku rozpoznań na poziomie 1,4%).
- ▶ Insercje o długości  $\leq 23$  (wszystkie rozpoznane nukleotydy w 378 sekwencjonowanych amplikonach zawierających insercję 23-nukleotydową zostały rozpoznane prawidłowo przy odsetku braku rozpoznań na poziomie 0,8%).
- ▶ Delecje o długości  $\leq 25$  (wszystkie rozpoznane nukleotydy w 1134 sekwencjonowanych amplikonach zawierających delecję 25-nukleotydową zostały rozpoznane poprawnie przy odsetku braku rozpoznań na poziomie 0,7%).

## Precyzja

Precyzję oznaczeń z wykorzystaniem aparatu NextSeq 550Dx Instrument wyznaczono, badając 13 unikalnych próbek Platinum Genome przy użyciu trzech aparatów, trzech serii odczynników przez trzech operatorów; w sumie dziewięć przebiegów sekwencjonowania inicjowanych w pięciu dniach. Oznaczenie reprezentatywne, próbki i metoda referencyjna były takie same, jak opisane dla badania dokładności oznaczania wariantów linii zarodkowej. Udziały w precyzji wyznaczono w analizie komponentów wariacyjnych z wykorzystaniem VAF jako zmiennej odpowiedzi oraz obliczając odchylenia standardowe na poziomie komponentu dla aparatu, serii odczynników, operatora i dnia rozpoczęcia oznaczenia ([Tabela 19](#)). Całkowita liczba obserwacji wykorzystanych w analizie dla każdego komponentu zmienności, tj. aparatu, operatora i serii odczynników, wyniosła 699, 176 i 235 odpowiednio w przypadku wariantów SNV, insercji i delecji.

Tabela 19 Wyniki precyzji oznaczeń przy użyciu aparatu NextSeq 550Dx Instrument (odchylenie standardowe)

Komponent	Typ wariantu	Komponent SD		Całkowite SD	
		Maks.	Mediana	Maks.	Mediana
Seria	SNV	0,0076	0,0002	0,0833	0,0154
	Insercja	0,0104	0,0000	0,0410	0,0157
	Delecja	0,0046	0,0005	0,0560	0,0187
Aparat	SNV	0,0114	0,0003	0,0840	0,0153
	Insercja	0,0138	0,0009	0,0407	0,0161
	Delecja	0,0079	0,0008	0,0549	0,0187
Operator	SNV	0,0226	0,0008	0,0841	0,0155
	Insercja	0,0344	0,0010	0,0417	0,0164
	Delecja	0,0083	0,0013	0,0547	0,0187
Dzień	SNV	0,0277	0,0012	0,0825	0,0160
	Insercja	0,0235	0,0012	0,0409	0,0169
	Delecja	0,0271	0,0014	0,0548	0,0188

### Porównanie metod (platforma sekwencjonująca)

Próbki krwi pełnej i próbki FFPE poddawano badaniom w aparatach NextSeq 550Dx Instrument i MiSeqDx Instrument przy użyciu zestawu TruSeq Custom Amplicon Kit Dx zgodnie z procedurami analiz wariantów linii zarodkowej i wariantów somatycznych. Ocenę zgodności częstości występowania wariantów w próbkach krwi i FFPE oceniano w wielokrotnych oznaczeniach reprezentatywnych. Rysunek 2 przedstawia wykres korelacji VAF między dwoma aparatami w jednym oznaczeniu reprezentatywnym, a Tabela 20 zawiera podsumowanie tej korelacji dla panelu oznaczeń. W oparciu o silną korelację między aparatami MiSeqDx i NextSeq 550Dx Instrument stwierdzono, że cechy charakterystyki działania zależne od czynników przedanalizacyjnych (np. sposobów ekstrakcji czy substancji zakłócających) dotyczą obu tych urządzeń. Dodatkowe informacje znajdują się w ulotce dołączonej do opakowania zestawu TruSeq Custom Amplicon Kit Dx.

Rysunek 2 Korelacja wartości VAF między aparatem MiSeqDx Instrument a aparatem NextSeq 550Dx Instrument w oznaczeniu 1 dla próbek FFPE (po lewej) i próbek krwi (po prawej)

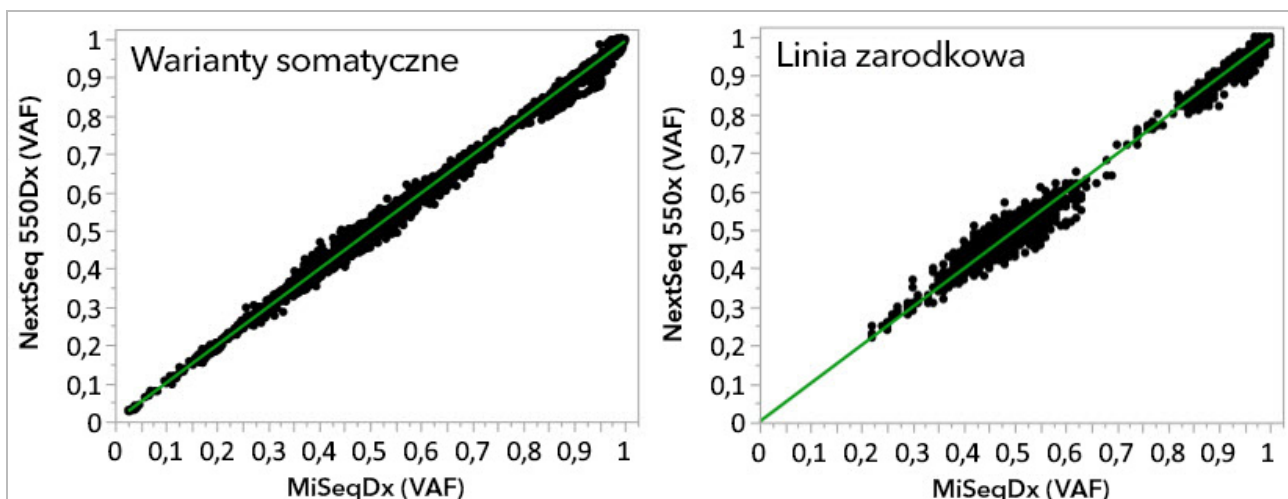


Tabela 20 Wyniki porównania metod z wykorzystaniem unikalnych próbek krwi i FFPE

Źródło gDNA	Oznaczenie (panel oligonukleotydów)	Powtórzenia biologiczne (próbki)	Powtórzenia techniczne (na próbkę)	Obserwacje (liczba wariantów)	Nachylenie	Punkt przecięcia	Korelacja (R <sup>2</sup> )
Krew	Oznaczenie 1	45	2	8369 <sup>1</sup>	0,992	0,002	0,995 <sup>2</sup>
Krew	Oznaczenie 2	45	2	5457	0,995	0,005	0,981
FFPE	Oznaczenie 1	46	2	8319	0,993	0,000	0,997 <sup>2</sup>
FFPE	Oznaczenie 3	40	1	280	0,969	0,015	0,978

<sup>1</sup>Dwa punkty danych usunięto w oparciu o podane ograniczenie dla modułu do wykrywania wariantów linii zarodkowej.

<sup>2</sup>Współczynnik determinacji dla wykresów VAF przedstawionych na rysunku 2.

## Odtwarzalność

Odtwarzalność oznaczeń z użyciem aparatu NextSeq 550Dx Instrument oceniano przy użyciu próbek Platinum Genome w oznaczeniu reprezentatywnym mającym na celu rozpoznanie różnych genów obejmujących 12 588 nukleotydów znajdujących się w 23 różnych chromosomach przy użyciu 150 amplikonów. Badanie wariantów linii zarodkowej obejmowało siedem powtórzeń każdej z 13 próbek; badanie wariantów somatycznych obejmowało sześć powtórzeń każdej z siedmiu próbek przy różnych poziomach VAF. Próbki przygotowano przy użyciu zestawu TruSeq Custom Amplicon Kit Dx.

Badanie przeprowadzono w trzech zewnętrznych ośrodkach z wykorzystaniem jednej serii zestawu odczynników na 300 cykli NextSeq 550Dx High Output Reagent Kit v2 (300 cycles). W każdym ośrodku korzystano z pojedynczego aparatu NextSeq 550Dx Instrument. W każdym ośrodku badania wykonywało dwóch operatorów. Każdy z operatorów wykonywał badania w trzech nienastępujących po sobie dniach rozpoczynania oznaczeń; łącznie wykonano 36 przebiegów w trzech ośrodkach. W badaniu wykonano po 18 przebiegów dla procedur oznaczania wariantów linii zarodkowej i wariantów somatycznych.

## Linia zarodkowa

Warianty linii zarodkowej o VAF na poziomie  $\geq 0,2$  raportowane są jako wyniki (warianty) dodatnie. W przypadku spodziewanych dodatnich wariantów linii zarodkowej dokonywano oceny danych dotyczących odsetka braku rozpoznań i odsetka prawidłowych rozpoznań dodatnich dla każdego rodzaju wariantu (SNV, insercja, delecja). Tabela 21 zawiera podsumowanie obserwowanych wskaźników wraz z obliczonymi metodą punktacji Wilsona górnymi i dolnymi granicami przedziału ufności 95% (LCL/UCL) dla każdego typu wariantu.

Tabela 21 Obserwacje w zakresie rozpoznań wariantów linii zarodkowej dla spodziewanych wyników dodatnich w zależności od typu wariantu

Typ wariantu	Brak rozpoznania			Prawidłowe rozpoznanie dodatnie				
	Obserwacja	Łącznie	Odsetek	Obserwacja	Łącznie	Odsetek	95% LCL	95% UCL
SNV	16	110 376	0,014	110 349	110 360	99,99	99,98	99,99
Insercje	1026	37 044	2,77	36 018	36 018	100	99,99	100,00
Delecje	648	34 776	1,86	34 128	34 128	100	99,99	100,00

Warianty linii zarodkowej o VAF na poziomie  $< 0,2$  są raportowane jako wyniki ujemne (typ dziki). W odniesieniu do spodziewanych ujemnych wyników oznaczeń wariantów linii zarodkowej dane oceniano w oparciu o odsetki braków rozpoznań i prawidłowych rozpoznań typu dzikiego. Tabela 22 zawiera podsumowanie obserwowanych wskaźników wraz z obliczonymi metodą punktacji Wilsona górnymi i dolnymi granicami przedziału ufności 95% (LCL/UCL).



Tabela 22 Obserwacje w zakresie rozpoznania wariantów linii zarodkowej dla spodziewanych wyników ujemnych

Typ wariantu	Brak rozpoznania			Prawidłowe rozpoznanie ujemne				
	Obserwacja	Łącznie	Odsetek	Obserwacja	Łącznie	Odsetek	95% LCL	95% UCL
Typ dziki	4883	19 600 182	0,025	19 595 299	19 595 299	100	100,00	100,00

Warianty linii zarodkowej o VAF na poziomie  $\geq 0,2$  i  $< 0,7$  określane są jako dodatnie heterozygotyczne, a warianty linii zarodkowej o VAF na poziomie  $\geq 0,7$  określane są jako dodatnie homozygotyczne. Do stwierdzenia, czy wewnętrzna zmienność oznaczenia wpływa na rozpoznanie genotypu, użyto próbek linii zarodkowych z wariantami heterozygotycznymi. Dla obu punktów odcięcia (0,2 dla genotypów heterozygotycznych i 0,7 dla genotypów homozygotycznych) wyznaczono wartości Cx, gdzie x jest odsetkiem powtarzanych prób przekraczających punkt odcięcia. Dla dolnego punktu odcięcia przy wartości VAF równej 0,2 wartość Cx wyniosła  $\geq 99,999\%$ , co wskazuje, że  $\geq 99,999\%$  wariantów heterozygotycznych zostałoby rozpoznanych jako heterozygotyczne. Dla górnego punktu odcięcia przy wartości VAF równej 0,7 wartość Cx wyniosła  $\leq 0,001\%$ , co wskazuje, że  $\leq 0,001\%$  wariantów heterozygotycznych zostałoby rozpoznanych jako homozygotyczne. Tabela 23 przedstawia podsumowanie wyników dla poszczególnych typów wariantów.

Warianty linii zarodkowej o VAF na poziomie  $\geq 0,2$  i  $< 0,7$  określane są jako dodatnie heterozygotyczne, a warianty linii zarodkowej o VAF na poziomie  $\geq 0,7$  określane są jako dodatnie homozygotyczne. Do stwierdzenia, czy wewnętrzna zmienność oznaczenia wpływa na rozpoznanie genotypu, użyto próbek linii zarodkowych z wariantami heterozygotycznymi. Dla obu punktów odcięcia (0,2 dla genotypów heterozygotycznych i 0,7 dla genotypów homozygotycznych) wyznaczono wartości Cx, gdzie x jest odsetkiem powtarzanych prób przekraczających punkt odcięcia. W odniesieniu do dolnego punktu odcięcia przy wartości VAF równej 0,2 wartość Cx wyniosła  $\geq 99,999\%$ , wskazując, że  $\geq 99,999\%$  wariantów heterozygotycznych zostałoby rozpoznanych jako heterozygotyczne. Dla górnego punktu odcięcia przy wartości VAF równej 0,7 wartość Cx wyniosła  $\leq 0,001\%$ , co wskazuje, że  $\leq 0,001\%$  wariantów heterozygotycznych zostałoby rozpoznanych jako homozygotyczne. Tabela 23 przedstawia podsumowanie wyników dla poszczególnych typów wariantów.

Tabela 23 Wartości Cx linii zarodkowej dla wariantów heterozygotycznych

Typ wariantu	Punkt odcięcia przy wartości VAF równej 0,2	Punkt odcięcia przy wartości VAF równej 0,7
	$\geq 99,999\%$	$\leq 0,001\%$
SNV	94/94	94/94
Insercje	24/24	24/24
Delekcje	35/35	35/35
Łącznie	153	153

### Warianty somatyczne

Warianty somatyczne o VAF na poziomie  $\geq 0,026$  raportowane są jako wyniki (warianty) dodatnie. Na potrzeby analiz obserwacje o VAF na poziomie  $\geq 0,01$  i  $< 0,026$  uznawano za niejednoznaczne (ani dodatnie, ani ujemne, znakowane jako warianty o niskiej częstości występowania). W ramach oceny wydajności oznaczeń wyniki obliczono na trzy sposoby:

- ▶ Najlepszy scenariusz: każdy wynik niejednoznaczny traktowano jako prawidłowe rozpoznanie dodatnie (zgodność z oczekiwanymi wynikami).
- ▶ Najgorszy scenariusz: każdy wynik niejednoznaczny traktowano jako rozpoznanie nieprawidłowe (niezgodność z oczekiwanymi wynikami).
- ▶ Scenariusz wykluczenia: wyniki niejednoznaczne wykluczano z analizy.

Trzy tabele – Tabela 24, Tabela 25 i Tabela 26 – zawierają podsumowanie wyników rozpoznania odpowiednio dla scenariusza najlepszego, scenariusza najgorszego i scenariusza wykluczenia wraz z obliczonymi metodą punktacji Wilsona górnymi i dolnymi granicami przedziału ufności 95% (LCL/UCL).

Tabela 24 Obserwacje w zakresie rozpoznania wariantów somatycznych dla spodziewanych wyników dodatnich w zależności od rodzaju wariantu (najlepszy scenariusz)

Typ wariantu	Prawidłowe rozpoznanie dodatnie				
	Obserwacja	Łącznie	Odsetek	95% LCL	95% UCL
SNV	54 346	54 346	100	99,99	100,00
Insercje	18 036	18 036	100	99,98	100,00
Delecje	18 381	18 381	100	99,98	100,00

Tabela 25 Obserwacje w zakresie rozpoznania wariantów somatycznych dla spodziewanych wyników dodatnich w zależności od rodzaju wariantu (najgorszy scenariusz)

Typ wariantu	Prawidłowe rozpoznanie dodatnie				
	Obserwacja	Łącznie	Odsetek	95% LCL	95% UCL
SNV	54 346	54 346	100	99,99	100,00
Insercje	18 000	18 036	99,8	99,72	99,86
Delecje	18 381	18 381	100	99,98	100,00

Tabela 26 Obserwacje w zakresie rozpoznania wariantów somatycznych dla spodziewanych wyników dodatnich w zależności od rodzaju wariantu (wykluczenie niejednoznacznego rozpoznania)

Typ wariantu	Prawidłowe rozpoznanie dodatnie				
	Obserwacja	Łącznie	Odsetek	95% LCL	95% UCL
SNV	54 346	54 346	100	99,99	100,00
Insercje	18 000	18 000	100	99,98	100,00
Delecje	18 381	18 381	100	99,98	100,00

Warianty somatyczne o VAF na poziomie  $<0,01$  są raportowane jako wyniki ujemne (typ dziki). W odniesieniu do spodziewanych ujemnych wyników oznaczeń wariantów somatycznych dane oceniano w oparciu o odsetki braków rozpoznania i prawidłowych rozpoznania typu dzikiego. Prawidłowe rozpoznania typu dzikiego stwierdzano w oparciu o wykluczenie braków rozpoznania i odjęcie od sumarycznej liczby obserwowanych rozpoznania mieszczących się w strefie niejednoznaczności ( $VAF \geq 0,01$  i  $<0,026$ ) oraz nieprawidłowych rozpoznania powyżej punktu odcięcia ( $VAF \geq 0,026$ ). Tabela 27 zawiera podsumowanie obserwowanych, łącznych i procentowych wyników ujemnych lokalizacji somatycznych pod kątem braku rozpoznania i prawidłowych rozpoznania typu dzikiego wraz z obliczonymi metodą punktacji Wilsona górnymi i dolnymi granicami przedziału ufności 95% (LCL/UCL).

Tabela 27 Obserwacje w zakresie rozpoznania wariantów somatycznych dla spodziewanych wyników ujemnych

Typ wariantu	Brak rozpoznania			Prawidłowe rozpoznanie						
	Obserwacja	Łącznie	Odsetek	Niejednoznaczne	Nieprawidłowe	Prawidłowe	Łącznie	Odsetek	95% LCL	95% UCL
Typ dziki	36 326	8 909 676	0,408	2254	121	8 870 975	8 873 350	99,97	99,972	99,974

W celu wyznaczenia wartości C95 oznaczenia (w ramach każdego typu wariantu) dokonano oceny próbek somatycznych przy różnych poziomach VAF tego samego wariantu. Do oceny zmienności w pobliżu punktu odcięcia oznaczenia użyto próbek o oczekiwanych poziomach VAF z zakresu od 0,02 do 0,07. Wartości C95 wyznaczono dla każdego wariantu; najwyższe wartości C95 dla każdego typu wariantu przedstawia Tabela 28.

Tabela 28 Podsumowanie wartości C95 dla wariantów somatycznych

Typ wariantu	N	C95
SNV	74	0,0613
Insercja	24	0,0573
Delecja	33	0,0575

## Wydajność zestawu odczynników na 300 cykli NextSeq 550Dx High Output Reagent Kit v2.5 (300 Cycle)

### Przegląd

Aparat NextSeq 550Dx Instrument obsługuje dwa zestawy odczynników: zestaw odczynników na 300 cykli NextSeq 550Dx High Output Reagent Kit v2 (300 cycles) i zestaw odczynników na 300 cykli NextSeq 550Dx High Output Reagent Kit v2.5 (300 cycles). W celu wykazania, że zestaw odczynników na 300 cykli NextSeq 550Dx High Output Reagent Kit v2.5 (300 cycles) spełnia wymagania w zakresie parametrów analitycznych zweryfikowanych i zwalidowanych dla zestawu odczynników na 300 cykli NextSeq 550Dx High Output Reagent Kit v2 (300 cycles), zestaw odczynników na 300 cykli NextSeq 550Dx High Output Reagent Kit v2.5 (300 cycles) poddano serii testów. Przy użyciu zestawu TruSeq Custom Amplicon Kit Dx sporządzono dwie biblioteki — jedną w procedurze dla linii zarodkowej, a drugą w procedurze dla wariantów somatycznych. Biblioteki przygotowane w każdej z procedur poddano testom z wykorzystaniem trzech serii zestawu odczynników na 300 cykli NextSeq 550Dx High Output Reagent Kit v2.5 (300 cycles) na trzech aparatach NextSeq 550Dx Instrument. Dodatkowo testy w ramach każdej procedury obejmowały pojedynczy przebieg z użyciem zestawu odczynników na 300 cykli NextSeq 550Dx High Output Reagent Kit v2 (300 cycles).

### Czułość analityczna (granica próby ślepej [LoB] i granica wykrywalności [LoD])

Weryfikacja z użyciem zestawu odczynników na 300 cykli NextSeq 550Dx High Output Reagent Kit v2 (300 cycles) wykazała, że aparat NextSeq 550Dx Instrument był w stanie wykrywać warianty przy wartości VAF równej 0,05 z błędem typu II wynoszącym  $\leq 0,05$  oraz że punkt odcięcia przy wartości VAF równej 0,026 wykorzystywany w module do wykrywania wariantów somatycznych (efektywna wartość LoB) dopuszcza błąd typu I wynoszący  $\leq 0,01$ . W oparciu o powyższe stwierdzenia oczekuje się, że wariant o VAF 0,05 ma częstość występowania większą lub równą 0,026 VAF w 95% przypadków, zaś pozycja typu dzikiego występuje z częstością mniejszą niż 0,026 VAF w 99% przypadków. W celu upewnienia się, że powyższe wartości deklarowane pozostają prawdziwe dla zestawu odczynników na 300 cykli NextSeq 550Dx High Output Reagent Kit v2.5 (300 cycles), wykonano serię powtórzeń pomiarów próbek typu dzikiego (próbek LoB) oraz próbek zawierających warianty z VAF na poziomie 0,05 (próbek LoD) z użyciem aparatu NextSeq 550Dx Instrument i zestawu odczynników na 300 cykli NextSeq 550Dx High Output Reagent Kit v2.5 (300 cycles). Odsetek rozpoznań powyżej i poniżej punktu odcięcia wynoszącego 0,026 porównano z wartościami deklarowanymi wyznaczonymi przy użyciu zestawu odczynników na 300 cykli NextSeq 550Dx High Output Reagent Kit v2 (300 cycles).

W testach użyto dwóch próbek LoD, każdej o unikalnym zestawie wariantów o docelowej częstości występowania VAF równej 0,05 oraz odpowiednich próbek LoB zawierających typ dziki docelowych wariantów. W ramach przygotowywania bibliotek próbki LoD i LoB przetwarzano odpowiednio w ośmiu i siedmiu powtórzeniach przy użyciu zestawu TruSeq Custom Amplicon Kit Dx. Biblioteki początkowo sekwencjonowano z wykorzystaniem zestawu odczynników na 300 cykli NextSeq 550Dx High Output Reagent Kit v2 (300 cycles) w celu zidentyfikowania wariantów/współrzędnych genomowych na potrzeby oceny LoB/LoD przy użyciu zestawu odczynników na 300 cykli NextSeq 550Dx High Output Reagent Kit v2.5 (300 cycles). W analizie LoD wykorzystano wszystkie warianty o wyznaczonej w oparciu o zestaw odczynników na 300 cykli NextSeq 550Dx High Output Reagent Kit v2 (300 cycles) średniej wartości VAF z zakresu 0,045–0,055 (warianty LoD) (N = 51 wariantów). W analizie LoB oceniano 51 odpowiadających im współrzędnych genomowych.

Na potrzeby oceny zestawu odczynników na 300 cykli NextSeq 550Dx High Output Reagent Kit v2.5 (300 cycles) biblioteki sekwencjonowano w trzech przebiegach w trzech kolejnych dniach przy użyciu tego samego aparatu i tej samej serii zestawu odczynników. Testy obejmowały 24 powtórzenia dla każdego z 51 wariantów LoD i 21 powtórzeń dla każdej z odpowiadających im pozycji typu dzikiego. Odsetek rozpoznań typu dzikiego przy VAF o wartości <0,026 przedstawia [Tabela 29](#). Odsetek rozpoznań wariantów LoD przy VAF o wartości ≥0,026 przedstawia [Tabela 30](#).

Tabela 29 Odsetek rozpoznań przy VAF o wartości < 0,026 dla pozycji typu dzikiego (ocena deklarowanej wartości LoB)

Typ wariantu	Oceniane pozycje	Obserwacje łącznie	Liczba pomiarów VAF ≥2,6%	Odsetek <2,6%	95-procentowy przedział ufności dla odsetka
SNV	32	672	0	1	0,994 – 1
Insercja	11	231	0	1	0,984 – 1
Delecja	8	168	0	1	0,978 – 1

Tabela 30 Odsetek rozpoznań ≥ 0,026 VAF dla wariantów LoD (ocena deklarowanej wartości LoD)

Typ wariantu	Oceniane pozycje	Obserwacje łącznie	Liczba pomiarów VAF <2,6%	Liczba pomiarów VAF ≥2,6%	Odsetek ≥2,6%	95-procentowy przedział ufności dla odsetka
SNV	32	768	1	767	0,999	0,993 – 1
Insercja	11	264	3	261	0,989	0,967 – 0,996
Delecja	8	192	2	190	0,99	0,963 – 0,997

## Dokładność

### Linia zarodkowa

W celu dokonania oceny dokładności rozpoznawania wariantów przez moduł do wykrywania wariantów linii zarodkowej przeprowadzono następujące badanie z wykorzystaniem zestawu odczynników na 300 cykli NextSeq 550Dx High Output Reagent Kit v2.5 (300 cycles). Dwanaście unikalnych próbek Platinum Genome poddano badaniu z wykorzystaniem oznaczenia reprezentatywnego. Wykonano w sumie 11 przebiegów na trzech aparatach NextSeq 550Dx Instrument z wykorzystaniem trzech zestawów odczynników na 300 cykli NextSeq 550Dx High Output Reagent Kit v2.5 (300 cycles).

Dokładność została określona dla wariantów SNV, insercji i delecji przez porównanie danych z badania z dobrze scharakteryzowaną, złożoną metodą referencyjną — bazą danych Platinum Genomes w wersji 2016-1.0. W charakterze odniesienia podano wyniki oznaczenia dokładności pojedynczego przebiegu sekwencjonowania z wykorzystaniem zestawu odczynników na 300 cykli NextSeq 550Dx High Output Reagent Kit v2 (300 cycles). Podsumowanie wyników przedstawia [Tabela 31](#).

Tabela 31 Podsumowanie zgodności dla linii zarodkowej

Kryteria	Obserwacje łącznie (v2.5) <sup>1</sup>	Wynik dla obserwacji (v2.5) <sup>2</sup>	Wynik dla obserwacji (v2) <sup>3</sup>	Wynik dla przebiegu (v2.5) <sup>4</sup>	Wynik dla przebiegu (v2) <sup>4</sup>
PPA dla SNV	1056	98,7	98,7	>99,9	>99,9
PPA dla insercji	1056	100	100	100	98,9
PPA dla delecji	1056	95,2	95,2	>99,9	100
NPA	1056	100	100	100	100
OPA	1056	>99,9	>99,9	>99,9	>99,9

<sup>1</sup> Obliczone jako liczba próbek w przebiegu × liczba przebiegów (96 próbek na przebieg × 11 przebiegów = 1056 obserwacji).

<sup>2</sup> Najniższa obserwowana wartość dla powtórzenia próbki we wszystkich przebiegach (w oparciu o 11 przebiegów z wykorzystaniem zestawu odczynników NextSeq 550Dx High Output Reagent Kit v2.5).

<sup>3</sup> Najniższa obserwowana wartość dla powtórzenia próbki w jednym przebiegu (łącznie 96 obserwacji).

<sup>4</sup> Najniższa wartość w połączonej analizie danych z każdego przebiegu.

### Warianty somatyczne

Opisane poniżej badanie miało na celu dokonanie oceny dokładności rozpoznawania wariantów przez moduł do wykrywania wariantów somatycznych aparatu NextSeq 550Dx Instrument przy użyciu zestawu odczynników na 300 cykli NextSeq 550Dx High Output Reagent Kit v2.5 (300 cycles). Dziesięć próbek Platinum Genome FFPE (dwie z wariantami rozcieńczonym do VAF o wartości 0,05) poddano badaniu przy użyciu oznaczenia reprezentatywnego. Wykonano w sumie 11 przebiegów na trzech aparatach NextSeq 550Dx Instrument z wykorzystaniem trzech serii zestawu odczynników na 300 cykli NextSeq 550Dx High Output Reagent Kit v2.5 (300 cycles).

Dokładność została określona dla wariantów SNV, insercji i delekcji przez porównanie danych z badania z dobrze scharakteryzowaną, złożoną metodą referencyjną — bazą danych Platinum Genomes w wersji 2016-1.0. W charakterze odniesienia podano wyniki oznaczenia dokładności pojedynczego przebiegu sekwencjonowania z wykorzystaniem zestawu odczynników na 300 cykli NextSeq 550Dx High Output Reagent Kit v2 (300 cycles). Podsumowanie wyników przedstawia [Tabela 32](#).

Tabela 32 Podsumowanie zgodności wariantów somatycznych

Kryteria	Obserwacje łącznie (v2.5) <sup>1</sup>	Wynik wg obserwacji (v2.5) <sup>2</sup>	Wynik wg obserwacji (v2) <sup>3</sup>	Wynik wg przebiegu (v2.5) <sup>4</sup>	Wynik wg przebiegu (v2) <sup>4</sup>
PPA dla SNV	528	100	100	100	100
PPA dla insercji	528	96,9	96,9	>99,9	>99,9
PPA dla delekcji	528	100	100	100	100
NPA	528	>99,9	>99,9	>99,9	>99,9
OPA	528	>99,9	>99,9	>99,9	>99,9

<sup>1</sup> Obliczone jako liczba próbek w przebiegu × liczba przebiegów (48 próbek na przebieg × 11 przebiegów = 528 obserwacji).

<sup>2</sup> Najniższa obserwowana wartość dla powtórzenia próbki we wszystkich przebiegach (w oparciu o 11 przebiegów z wykorzystaniem zestawu odczynników NextSeq 550Dx High Output Reagent Kit v2.5).

<sup>3</sup> Najniższa obserwowana wartość dla powtórzenia próbki w jednym przebiegu (łącznie 96 obserwacji).

<sup>4</sup> Najniższa wartość w połączonej analizie danych z każdego przebiegu.

### Precyzja

#### Linia zarodkowa

Precyzję oznaczeń z wykorzystaniem zestawu odczynników na 300 cykli NextSeq 550Dx High Output Reagent Kit v2.5 (300 cycles) przy użyciu modułu do wykrywania wariantów linii zarodkowej oceniano w oparciu o próbki Platinum Genome i oznaczenie reprezentatywne. Badanie polegało na przygotowaniu pojedynczej biblioteki przy użyciu zestawu TruSeq Custom Amplicon Kit Dx, obejmującej 12 próbek, z których każda została poddana przetworzeniu w ośmiu powtórzeniach. Biblioteki sekwencjonowano z wykorzystaniem trzech serii zestawu odczynników na 300 cykli NextSeq 550Dx High Output Reagent Kit v2.5 (300 cycles) na trzech aparatach NextSeq 550Dx Instrument – razem dziewięć przebiegów sekwencjonowania.

Do stwierdzenia, czy wewnętrzna zmienność oznaczenia wpływa na rozpoznanie genotypu użyto próbek z wariantami heterozygotycznymi (N = 153 unikalne warianty heterozygotyczne). Dla obu punktów odcięcia stosowanych w module do wykrywania wariantów linii zarodkowej (0,2 dla genotypów heterozygotycznych i 0,7 dla genotypów homozygotycznych) wyznaczono wartości Cx, gdzie x jest odsetkiem powtarzanych prób przekraczających punkt odcięcia. Dla dolnego punktu odcięcia przy wartości VAF równej 0,2 wariant z minimalną wartością Cx był identyfikowany przy użyciu zestawu odczynników na 300 cykli NextSeq 550Dx Reagent Kit v2.5 (300 cycles)

w >99,9% przypadków, co wskazuje, że >99,9% wariantów heterozygotycznych zostałyby rozpoznanych jako heterozygotyczne. Dla górnego punktu odcięcia przy wartości VAF równej 0,7 wariant z minimalną wartością Cx był identyfikowany przy użyciu zestawu odczynników na 300 cykli NextSeq 550Dx Reagent Kit v2.5 (300 cycles) w <1,5% przypadków, co wskazuje, że  $\leq 1,5\%$  wariantów heterozygotycznych zostałyby rozpoznanych jako homozygotyczne. **Tabela 33** zawiera podsumowanie wyników dla poszczególnych rodzajów wariantów. Dla porównania podano wartości Cx z pojedynczego sekwencjonowania przy użyciu zestawu odczynników na 300 cykli NextSeq 550Dx High Output Reagent Kit v2 (300 cycles).

Tabela 33 Wartości Cx linii zarodkowej dla wariantów heterozygotycznych

Typ wariantu	N	Punkt odcięcia przy wartości VAF równej 0,2		Punkt odcięcia przy wartości VAF równej 0,7	
		Min. Cx (v2.5) <sup>1</sup>	Min. Cx (v2) <sup>2</sup>	Maks. Cx (v2.5) <sup>1</sup>	Maks. Cx (v2) <sup>2</sup>
SNV	94	>99,9%	>99,9%	1,5%	1,0%
Insercje	24	100%	100%	0%	<0,1%
Delecje	35	100%	>99,9%	<0,1%	<0,1%

<sup>1</sup>Wartości Cx oparte na szacunkach całkowitego odchylenia standardowego w analizie komponentów wariacyjnych.

<sup>2</sup>Wartości Cx oparte na odchyleniach standardowych dla próbek.

### Warianty somatyczne

Precyzję oznaczeń z wykorzystaniem zestawu odczynników NextSeq 550Dx High Output Reagent Kit v2.5 (300 cycles) przy użyciu modułu do wykrywania wariantów somatycznych oceniano w oparciu o próbki Platinum Genome FFPE i oznaczenie reprezentatywne. Badanie polegało na sporządzeniu pojedynczej biblioteki przy użyciu zestawu TruSeq Custom Amplicon Kit Dx, obejmującej dwie próbki, z których każda została poddana przetwarzaniu w ośmiu powtórzeniach. Biblioteki sekwencjonowano z wykorzystaniem trzech serii zestawu odczynników na 300 cykli NextSeq 550Dx High Output Reagent Kit v2.5 (300 cycles) na trzech aparatach NextSeq 550Dx Instrument — razem dziewięć przebiegów sekwencjonowania.

Do oceny zmienności oznaczeń aparatu w pobliżu punktu odcięcia VAF dla modułu do wykrywania wariantów somatycznych (dodatkowo rozpoznania wariantów dla wariantów somatycznych o  $VAF \geq 0,026$ ) użyto wariantów somatycznych o oczekiwanych poziomach  $VAF \leq 0,10$  (N = 131 unikalnych wariantów). Dla każdego z wariantów somatycznych wyznaczono wartości C95. Wartości C95 oznaczają poziom VAF, dla którego prawdopodobieństwo bycia większym od punktu odcięcia VAF w module do wykrywania wariantów somatycznych wynosi 95%. Najwyższe wartości C95 dla poszczególnych typów wariantów przedstawia **Tabela 34**. Dla porównania podano wartości C95 z pojedynczego przebiegu sekwencjonowania przy użyciu zestawu odczynników na 300 cykli NextSeq 550Dx High Output Reagent Kit v2 (300 cycles).

Tabela 34 Podsumowanie wartości C95 dla wariantów somatycznych

Typ wariantu	Liczba ocenianych wariantów	C95 (v2.5) <sup>1</sup>	C95 (v2) <sup>2</sup>
SNV	74	0,064	0,063
Insercje	24	0,062	0,061
Delecje	33	0,060	0,060

<sup>1</sup>Wartości C95 oparte na szacunkach całkowitego odchylenia standardowego w analizie komponentów wariacyjnych.

<sup>2</sup>Wartości C95 oparte na odchyleniach standardowych dla próbek.

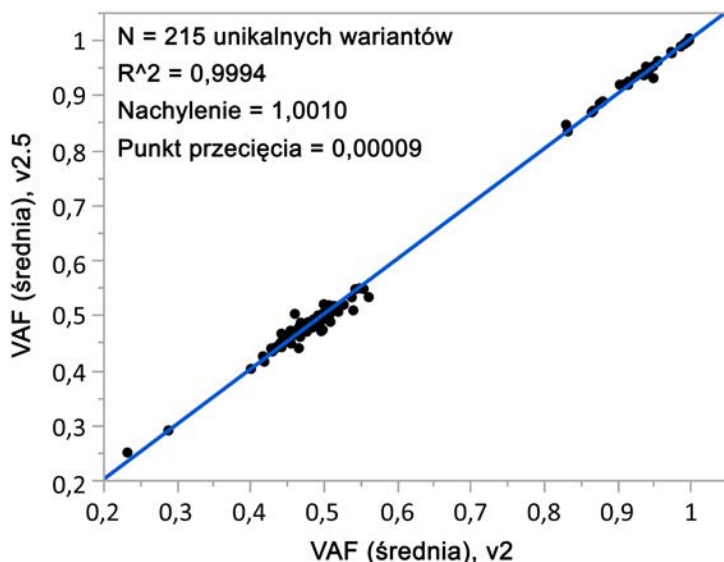
## Porównanie metod (zestaw odczynników)

### Linia zarodkowa

Dokonano oceny średnich wartości VAF dla 215 unikalnych wariantów z zestawu odczynników na 300 cykli NextSeq 550Dx High Output Reagent Kit v2 (300 cycles) i zestawu odczynników na 300 cykli NextSeq 550Dx High Output Reagent Kit v2.5 (300 cycles) w oparciu o wyniki wygenerowane przez moduł do wykrywania wariantów linii

zarodkowej. Średnie wartości VAF obliczono w oparciu o 11 przebiegów sekwencjonowania (v2.5) i jeden przebieg sekwencjonowania (v2). Do obliczenia średniej wartości dla każdego wariantu wykorzystano co najmniej osiem powtórzeń. **Rysunek 3** przedstawia wykres korelacji wartości VAF między oboma zestawami odczynników. W oparciu o silną korelację liniową wartości VAF i podobieństwo wyników uzyskanych przy użyciu obu zestawów odczynników, stwierdza się, że charakterystyka działania oznaczenia początkowo zweryfikowana i zwalidowana dla modułu do wykrywania wariantów linii zarodkowej przy użyciu zestawu odczynników na 300 cykli NextSeq 550Dx High Output Reagent Kit v2 (300 cycles) ma również zastosowanie w przypadku zestawu odczynników na 300 cykli NextSeq 550Dx High Output Reagent Kit v2.5 (300 cycles).

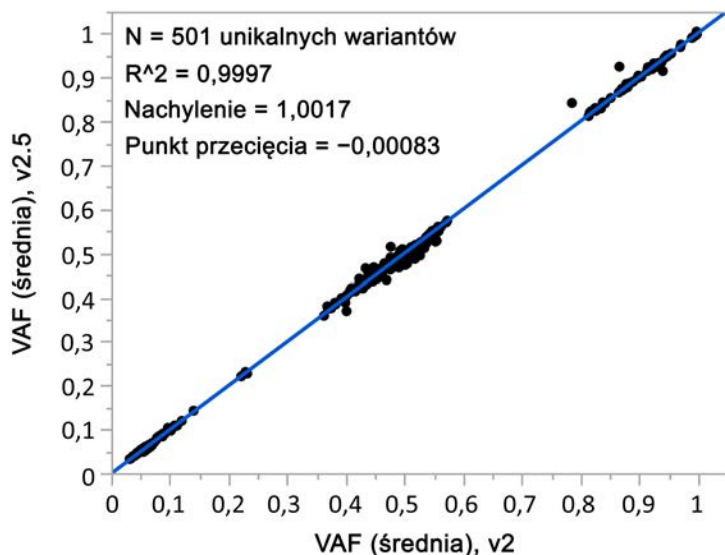
Rysunek 3 Korelacja częstości występowania allelu wariantu (VAF) wyznaczona przy użyciu modułu do wykrywania Kit wariantów linii zarodkowej dla zestawów odczynników na 300 cykli NextSeq 550Dx High Output Reagent Kit v2 (300 cycles) i zestawu odczynników na 300 cykli NextSeq 550Dx High Output Reagent Kit v2.5 Kit (300 cycles).



### Warianty somatyczne

Dokonano oceny średnich wartości VAF dla 501 unikalnych wariantów z zestawu odczynników na 300 cykli NextSeq 550Dx High Output Reagent Kit v2 (300 cycles) i zestawu odczynników na 300 cykli NextSeq 550Dx High Output Reagent Kit v2.5 (300 cycles) w oparciu o wyniki wygenerowane przez moduł do wykrywania wariantów somatycznych. Średnie wartości VAF obliczono w oparciu o 11 przebiegów sekwencjonowania (v2.5) i jeden przebieg sekwencjonowania (v2). Do obliczenia średniej wartości dla każdego unikalnego wariantu wykorzystano co najmniej trzy powtórzenia. **Rysunek 4** przedstawia wykres korelacji wartości VAF między oboma zestawami odczynników. W oparciu o korelację wartości VAF i podobieństwo wyników uzyskanych przy użyciu obu zestawów odczynników, stwierdza się, że charakterystyka działania oznaczenia zweryfikowana i zwalidowana dla modułu do wykrywania wariantów somatycznych przy użyciu zestawu odczynników na 300 cykli NextSeq 550Dx High Output Reagent Kit v2 (300 cycles) ma również zastosowanie w przypadku zestawu odczynników na 300 cykli NextSeq 550Dx High Output Reagent Kit v2.5 (300 cycles).

Rysunek 4 Korelacja częstości występowania allelu wariantu (VAF) wyznaczona przy użyciu modułu do wykrywania wariantów somatycznych dla zestawów odczynników na 300 cykli NextSeq 550Dx High Output Reagent Kit v2 (300 cycles) i zestawu odczynników na 300 cykli NextSeq 550Dx High Output Reagent Kit v2.5 (300 cycles).



## Historia wersji

Dokument	Data	Opis zmiany
Nr dokumentu: 1000000030326, wer. 06	Maj 2022 r.	Aktualizacje dokonane w celu poprawienia treści dodanych nieumyślnie z oprogramowania źródłowego.
Nr dokumentu: 1000000030326, wer. 05	Listopad 2021 r.	Dodano oświadczenie „Ostrzeżenia i środki ostrożności” dotyczące zgłaszania poważnych incydentów. Do części Zasady dotyczące procedury dodano oświadczenie określające docelowego użytkownika. Usunięto odniesienie do zestawu odczynników na 300 cykli High Output Reagent Kit v2 (300 cycles). Dodano odniesienie do zestawu odczynników na 75 cykli High Output Reagent Kit v2.5 (75 cycles).
Nr dokumentu: 1000000030326, wer. 04	Sierpień 2021	Dodano tabelę z historią wersji. Zaktualizowano adres autoryzowanego przedstawiciela w UE.

## 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 NIEWŁĄCZYWE UŻYTKOWANIE PRODUKTÓW (W TYM ICH CZĘŚCI I OPROGRAMOWANIA) OPISANYCH W NINIEJSZYM DOKUMENCIE.

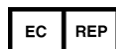
© 2022 Illumina, Inc. Wszelkie prawa zastrzeżone.

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](http://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](mailto:techsupport@illumina.com)  
[www.illumina.com](http://www.illumina.com)



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

### Sponsor australijski

Illumina Australia Pty Ltd  
Nursing Association Building  
Level 3, 535 Elizabeth Street  
Melbourne, VIC 3000  
Australia

## 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](http://support.illumina.com).