

NextSeq™ 550Dx Instrument

ВИКОРИСТОВУВАТИ ДЛЯ ДІАГНОСТИКИ *IN VITRO*
ЛИШЕ ДЛЯ ЕКСПОРТУ

№ за каталогом 20005715

Передбачене використання

Прилад NextSeq 550Dx призначено для секвенування бібліотек ДНК під час діагностичних аналізів *in vitro*. Прилад NextSeq 550Dx слід використовувати з певними зареєстрованими, сертифікованими або схваленими реагентами для діагностики *in vitro* й аналітичним програмним забезпеченням.

Принципи виконання процедури

Прилад Illumina NextSeq 550Dx призначено для секвенування бібліотек ДНК під час виконання аналізів для діагностики *in vitro*. До використання приладу допускається кваліфікований персонал клінічної лабораторії, належним чином навчений виконанню процедур, що належать до діагностики *in vitro*, в умовах клінічної лабораторії. Для введення в прилад NextSeq 550Dx використовуються бібліотеки, отримані з ДНК, де індекси зразків і послідовності для захоплення додаються до ампліфікованих цільових фрагментів. Бібліотеки зразків фіксуються на проточній кюветі та секвенуються на приладі за допомогою хімічної реакції секвенування за допомогою синтезу (SBS). У хімічній реакції SBS використовується метод оборотних термінаторів для виявлення флуоресцентно-мічених однонуклеотидних основ, тому що їх включено в нитки ДНК, які зростають. Програмне забезпечення Real-Time Analysis (RTA) виконує аналіз зображень і розпізнавання основ, а потім присвоює оцінку якості до кожної основи для кожного циклу секвенування. Після завершення первинного аналізу можна виконати вторинний аналіз для обробки розпізнавання основ. У приладі NextSeq 550Dx використовуються різні модулі для вторинного аналізу залежно від робочого процесу. Для модулів Germline Variant Module або Somatic Variant Module обробка складається з демультіплексування, генерування файлів FASTQ, вирівнювання, розпізнавання варіантів і генерування файлів у форматі розпізнавання варіанта (VCF або gVCF). Файли VCF і gVCF містять інформацію про варіанти, знайдені в конкретних позиціях референсного геному.

Завантаження з вибором операційної системи

Прилад NextSeq 550Dx має опцію завантаження з вибором операційної системи, щоб прилад можна було використовувати в режимі діагностики (Dx) або в режимі лише для досліджень (research use only, RUO). Діагностичні аналізи секвенування *in vitro*, включно з модулями Germline Variant Module і Somatic Variant Module, запускаються в діагностичному режимі. У діагностичному режимі можна застосовувати лише реагенти для секвенування IVD. Технічні характеристики й обмеження процедури для приладу NextSeq 550Dx встановлено з використанням модулів Germline Variant Module і Somatic Variant Module у діагностичному режимі.

Обмеження процедури

- 1 Використовувати для діагностики *in vitro*.
- 2 У разі використання набору реагентів NextSeq 550Dx High Output Reagent Kit v2 (300 cycles) і набору реагентів NextSeq 550Dx High Output Reagent Kit v2.5 (300 cycles) модулі Germline Variant Module та Somatic Variant Module дають змогу досягти таких показників:
 - ▶ вихідні дані секвенування ≥ 90 мільярдів пар основ (млрд п. о.);

- ▶ довжина зчитування (у парному кінцевому прогоні) 2 ×150 пар основ (п. о.);
 - ▶ основи з показником Q не менше 30 (Q30) ≥ 75 % за довжини зчитування 2 ×150 п. о.; щонайменше 75 % основ, які мають оцінки якості за шкалою Phred ≥ 30, що вказує на точність розпізнавання основ більше 99,9 %.
- 3 Зчитування з інделами (інсерції, делеції або їхні комбінації) довжиною > 25 п. о. програмним забезпеченням аналізу не вирівнюються. Отже, програмне забезпечення аналізу не виявляє індели довжиною > 25 п. о.
- 4 Програмне забезпечення аналізу може не вирівнювати зчитування ампліконів з край високим вмістом варіантів, через що відповідна ділянка буде визначена як дикий тип. Такий край високий вміст передбачає наведене нижче.
- ▶ Зчитування, що містять понад три індели.
 - ▶ Зчитування довжиною щонайменше 30 п. о. з вмістом однуклеотидних варіантів (single nucleotide variants, SNV) > 4 % загальної цільової довжини ампліконів (за винятком ділянок зондів).
 - ▶ Зчитування довжиною < 30 п. о. із вмістом SNV > 10 % загальної довжини ампліконів (включно з ділянками зондів).
- 5 Великі варіанти, зокрема багатонуклеотидні варіанти (multi-nucleotide variants, MNV) і великі індели, можуть звітуватися як окремі менші варіанти у вихідному файлі VCF.
- 6 Варіанти делецій можна фільтрувати або пропускати під час стягування двох сегментованих ампліконів, якщо довжина делеції дорівнює перекриттю між сегментованими ампліконами або більша.
- 7 Система не здатна виявляти індели, коли вони виникають прямо поряд із праймером, а амплікона з перекриттям немає. Для ділянок з ампліконами, що перекриваються, аналіз не здатен виявити делеції, коли така ділянка перекриття менша за розмір делеції, яку потрібно виявити. Наприклад, якщо ділянка перекриття між двома суміжними ампліконами становить дві основи, аналіз не може виявити жодні делеції, зокрема ці обидві основи. Одноосновну делецію на будь-якій із цих основ виявити можна.
- 8 Як і у випадку з будь-яким робочим процесом підготовки бібліотеки на основі гібридизації, первинні поліморфізми, мутації, інсерції або делеції в областях зв'язування олігонуклеотидів можуть впливати на алелі, що досліджуються, і розпізнавання під час секвенування. Приклад наведено далі.
- ▶ Варіант у фазі, де варіант перебуває в ділянці праймера, може не піддаватися ампліфікації, що призводить до хибнонегативних результатів.
 - ▶ Варіанти в ділянці праймера можуть заважати ампліфікації референсного алеля, що призводить до неправильного розпізнавання гомозиготного варіанта.
 - ▶ Варіанти інделів у ділянці праймера можуть призводити до хибнопозитивного розпізнавання в кінці зчитування біля праймера.
- 9 Індели можна фільтрувати завдяки викривленню нитки, якщо вони виникають поруч із кінцем одного зчитування й піддаються м'якому усіченню під час вирівнювання.
- 10 Малі MNV валідацію не проходили, і про них повідомляється лише в модулі Somatic Variant Module.
- 11 Про делеції повідомляється у файлі VCF у координаті попередньої основи відповідно до формату VCF. Тому перш ніж повідомляти про окреме розпізнавання основи як про гомозиготний референсний варіант, розгляньте суміжні варіанти.
- 12 Обмеження, що стосуються зародкової лінії.
- ▶ Прилад NextSeq 550Dx, у якому використовується модуль Germline Variant Module програмного забезпечення Local Run Manager для NextSeq 550Dx, призначено для отримання якісних результатів розпізнавання варіантів зародкової лінії (наприклад, гомозиготних, гетерозиготних, дикого типу).
 - ▶ Коли використовується модуль Germline Variant Module, мінімальне покриття на амплікон, потрібне для точного розпізнавання варіантів, становить 150x. Зрештою потрібно 150 допоміжних фрагментів ДНК, що дорівнює 300 парним зчитуванням, які перекриваються. На покриття впливають кількість зразків і загальна кількість цільових основ. На покриття може впливати вміст GC й інший геномний вміст.
 - ▶ Варіація числа копій може впливати на те, яким буде визначено варіант: гомозиготним чи гетерозиготним.

- ▶ У файлах VCF варіанти в певному повторному контексті відфільтровуються. Для фільтрування варіантів використовується фільтр повторів RМхN, якщо вся послідовність варіанта або її частина повторно наявна в референсному геномі, суміжному з позицією варіанта. Для розпізнавання варіантів зародкової лінії в референсному геномі потрібно принаймні дев'ять повторів, щоб варіант було відфільтровано. До уваги беруться лише повтори довжиною до 5 п. о. (R5x9).
 - ▶ Наявність інделу й SNV в одному локусі може призвести до виявлення лише одного варіанта.
- 13 Обмеження, що стосуються соматичних варіантів.
- ▶ Прилад NextSeq 550Dx, у якому використовується модуль Somatic Variant Module програмного забезпечення Local Run Manager для NextSeq 550Dx, призначено для отримання якісних результатів розпізнавання соматичних варіантів (тобто наявність соматичного варіанта з варіантною частотою не менше 0,026 із межею виявлення 0,05).
 - ▶ Коли використовується модуль Somatic Variant Module, мінімальне покриття на амплікон, потрібне для точного розпізнавання варіантів, становить 450x на пул олігонуклеотидів. Зрештою потрібно 450 допоміжних фрагментів ДНК на пул олігонуклеотидів, що дорівнює 900 парним зчитуванням, які перекриваються. На покриття впливають кількість зразків і загальна кількість цільових основ. На покриття може впливати вміст GC й інший геномний вміст.
 - ▶ Для розпізнавання соматичних варіантів у референсному геномі потрібно принаймні шість повторів, щоб варіант було відфільтровано. До уваги беруться лише повтори завдовжки до 3 п. о. (R3x6).
 - ▶ Somatic Variant Module не здатен відрізнити варіанти зародкової лінії від соматичних варіантів. Модуль призначений для виявлення варіантів у межах діапазону частот варіантів, але частота варіанта не може використовуватися для диференціації соматичних варіантів від варіантів зародкової лінії.
 - ▶ Вміст нормальної тканини в зразку впливає на виявлення варіантів. Зареєстрована межа виявлення базується на частоті варіанта щодо загальної ДНК, екстрагованої як із пухлини, так і з нормальної тканини.

Компоненти виробу

- 1 Прилад NextSeq 550Dx (№ за каталогом 20005715).
- 2 Компоненти програмного забезпечення для приладу NextSeq 550Dx, зокрема представлені в таблиці далі.

Програма	Функція	Опис
Системне програмне забезпечення NextSeq 550Dx (NOS)	Контроль роботи приладу	Прикладне програмне забезпечення NOS керує роботою приладу під час секвенування й генерує зображення для використання в програмному забезпеченні Real-Time Analysis (RTA).
Програмне забезпечення Real-time Analysis (RTA)	Первинний аналіз	Програмне забезпечення RTA перетворює зображення, згенероване NOS для кожного сегмента на цикл прогону секвенування, на файли розпізнавання основ, які є вхідними для модулів аналізу програмного забезпечення Local Run Manager. Програмне забезпечення RTA не має інтерфейсу користувача.
Програмне забезпечення Local Run Manager	Інтерфейс для вибору модуля	Програмне забезпечення Local Run Manager — це інтегроване в прилад рішення для керування користувачами, вибору відповідного модуля аналізу й моніторингу стану.
Модуль Somatic Variant Module	Вторинний аналіз	Цей модуль аналізу програмного забезпечення Local Run Manager обробляє розпізнавання основ за допомогою вторинного аналізу. Обробка включає демультіплексування, генерування файлів FASTQ, вирівнювання, розпізнавання варіантів і звітність. Засіб розпізнавання варіантів (Pisces) генерує файли VCF, які містять інформацію про варіанти, знайдені в певних позиціях у референсному геномі, і включає отриману під час вимірювань частоту варіанта.

Програма	Функція	Опис
Модуль Germline Variant Module	Вторинний аналіз	Цей модуль аналізу програмного забезпечення Local Run Manager обробляє розпізнавання основ за допомогою вторинного аналізу. Обробка включає демультіплексування, генерування файлів FASTQ, вирівнювання, розпізнавання варіантів і звітність. Засіб розпізнавання варіантів (Piscis) генерує файли VCF, які містять інформацію про варіанти, знайдені в певних позиціях у референсному геномі, й ідентифікує кожний варіант як гетерозиготний або гомозиготний.

Умови експлуатації

Показник	Вимога
Температура	Підтримуйте температуру в лабораторії в межах від 19 °C до 25 °C (22 °C ± 3 °C). У цьому температурному діапазоні має працювати прилад. Під час виконання циклу аналізу кімнатна температура навколишнього середовища не повинна змінюватися більше ніж на ±2 °C.
Вологість	Підтримуйте відносну вологість у стані відсутності конденсації в межах 20–80 %.

Обладнання й матеріали

Потрібні обладнання й матеріали, продаються окремо

Набір реагентів NextSeq 550Dx High Output Reagent Kit v2.5 (75 cycles), № за каталогом 20028870

Набір реагентів NextSeq 550Dx High Output Reagent Kit v2.5 (300 cycles), № за каталогом 20028871

Потрібні обладнання й матеріали, які не надано

Витратні матеріали, що замовляє користувач, для прогонів секвенування

Витратний матеріал	Постачальник	Мета
Серветки, просочені ізопропіловим спиртом, 70 %, чи етанолом, 70 %	VWR, № 95041-714 за каталогом (або еквівалентний). Постачальник загальнолабораторного обладнання	Очищення проточної кювети й загальне призначення
Низьковорсова лабораторна серветка	VWR, № 21905-026 за каталогом (або еквівалентний)	Очищення проточної кювети й загальне призначення

Витратні матеріали, що замовляє користувач, для обслуговування приладу

Витратний матеріал	Постачальник	Мета
NaOCl, 5 % (гіпохлорит натрію)	Sigma-Aldrich, № 239305 за каталогом (або еквівалент за хімічною чистотою)	Промивання приладу з використанням ручного промивання після прогону; розведено до концентрації 0,12 %
Tween 20	Sigma-Aldrich, № P7949 за каталогом	Промивання приладу з використанням параметрів ручного промивання; розведено до концентрації 0,05 %
Вода, хімічно чиста	Постачальник загальнолабораторного обладнання	Промивання приладу (промивання вручну)
Повітряний фільтр	Illumina, № 20022240 за каталогом	Очищення повітря, що надходить до приладу для охолодження

Рекомендації щодо хімічно чистої води

Для виконання процедур приладу завжди використовуйте хімічно чисту або деіонізовану воду. Використання водопровідної води заборонено. Використовуйте тільки наведені нижче класи води або їхні еквіваленти.

- ▶ Деіонізована вода.
- ▶ Illumina PW1.
- ▶ Вода 18 мегаом (МОм).
- ▶ Вода Milli-Q.
- ▶ Вода Super-Q.
- ▶ Вода для молекулярної біології.

Попередження й запобіжні заходи

ЗАСТЕРЕЖЕННЯ. Федеральний закон дозволяє продаж цього виробу лише за призначенням або на замовлення лікарів чи інших спеціалістів, які займаються медичною практикою й мають ліцензію від штату, у якому вони здійснюють свою практику, для використання або замовлення на використання цього виробу.

- 1 **Деякі компоненти реагентів, наданих компанією Illumina для використання з приладом NextSeq 550Dx, містять потенційно небезпечні хімічні речовини. Вдихання, проковтування, потрапляння на шкіру та в очі може завдати шкоди здоров'ю. Надягайте захисне приладдя, зокрема захист очей, рукавички та лабораторний одяг, з урахуванням ризику впливу. Поводьтеся з використаними реагентами як із хімічними відходами й утилізуйте їх відповідно до застосованих регіональних, державних і місцевих законів і нормативних правил. Щоб отримати додаткову інформацію про захист навколишнього середовища, аспекти здоров'я й техніки безпеки, див. паспорти безпеки продукції (safety data sheet, SDS) на сайті support.illumina.com/sds.html.**
- 2 негайно повідомляйте про всі серйозні інциденти, пов'язані з цим виробом, у компанію Illumina й компетентні установи держав, у яких перебувають користувач і пацієнт.
- 3 Зі зразками крові слід поводитися так, ніби вони заражені вірусом імунодефіциту людини (ВІЛ), вірусом гепатиту В людини (HBV) або іншими збудниками, що передаються через кров (універсальні заходи застереження).
- 4 Недотримання викладених процедур може призвести до неправильних результатів або значного погіршення якості зразка.
- 5 Використовуйте стандартні заходи застереження, прийняті в лабораторії. Не використовуйте піпетування ротом. Не вживайте їжу або напої й не паліть у робочих зонах. Під час роботи зі зразками та реагентами з набору надягайте одноразові рукавички й лабораторний одяг. Після роботи зі зразками та реагентами з набору ретельно мийте руки.
- 6 Належні правила поводження та гігієна в лабораторії є обов'язковими для попередження забруднення реагентів, приладів та інструментів, а також геномних зразків ДНК продуктами ПЛР. Забруднення продуктами ПЛР може призвести до неправильних і ненадійних результатів.
- 7 Щоб попередити забруднення, виділіть для зон преампліфікації та постапліфікації спеціальне обладнання й витратні матеріали (як-от піпетки, наконечники піпеток, термостати, вихрові мішалки й центрифуги).
- 8 Парування індексів зі зразками має точно відповідати роздрукованому макету планшета. Під час уведення назви зразка в програмне забезпечення Local Run Manager він автоматично заповнює праймери індексу, пов'язані з назвою зразка. Перед початком прогону секвенування рекомендується перевіряти зв'язок праймерів індексу зі зразками. Невідповідність між зразком і результатами макета планшета завадить виявленню позитивного зразка та призведе до отримання хибного результату.
- 9 Наполегливо рекомендовано встановити антивірусне програмне забезпечення, що замовляє користувач, для захисту комп'ютера від вірусів. Інструкції з установами можна знайти в посібнику користувача.

- 10 Забороняється працювати на приладі NextSeq 550Dx, коли з нього знято будь-яку з панелей. Експлуатація приладу, з якого знято хоча б одну з панелей, створює потенційну небезпеку від дії напруги електричної мережі й напруги постійного струму.
- 11 Не торкайтеся платформи проточної кювети у відсіку проточної кювети. Нагрівач у цьому відсіку працює за температури від 22 °C до 95 °C, що може призвести до опіків.
- 12 Цей прилад важить приблизно 84 кг (185 фунтів) і внаслідок падіння або неправильного поводження може призвести до тяжкого травмування.

Інструкції з використання

Наведені далі інструкції з використання призначено для роботи з модулями Germline Variant Module і Somatic Variant Module у діагностичному режимі на приладі NextSeq 550Dx із використанням набору реагентів NextSeq 550Dx High Output Reagent Kit v2 (300 cycles) або набору реагентів NextSeq 550Dx High Output Reagent Kit v2.5 (300 cycles).

Уведення інформації про прогін

Докладні інструкції див. у документі NextSeq 550Dx Instrument Reference Guide (Довідковий посібник до приладу NextSeq 550Dx) (документ № 1000000009513) і відповідному посібнику для модуля програмного забезпечення Local Run Manager.

Налаштування параметрів

- 1 Увійдіть до Local Run Manager.
- 2 Виберіть **Create Run** (Створити прогін), а потім виберіть **Somatic Variant** (Соматичні варіанти) або **Germline Variant** (Варіанти зародкової лінії).
- 3 Уведіть ім'я прогону, яке ідентифікує прогін від секвенування до аналізу. Використовуйте літери й цифри, пробіл, підкреслювання або дефіс.
- 4 **(Додатково.)** Уведіть опис прогону, який допоможе його ідентифікувати. Використовуйте літери й цифри, пробіл, підкреслювання або дефіс.
- 5 Виберіть кількість зразків та набори індексів у розкритому списку. Під час вибору враховуйте наведену далі інформацію.
 - ▶ Розкритий список містить номери зразків зі встановленими індексами. Наприклад, 24-Set 1 показує 24 зразки для тестування, з індексами з набору індексів 1.
 - ▶ Номер наборів індексів належать до різних наборів пар індексів i5 і i7. Обидва набори, i Set 1, і Set 2, забезпечують різноманітність індексів. Для запобігання виснаженню одного набору пропонуються два набори індексів.
 - ▶ Виберіть кількість зразків, яка є найближчою до кількості зразків, що ви тестуєте. Якщо точної кількості зразків немає в списку, виберіть найближчу, але меншу за кількість, яку ви тестуєте. Наприклад, якщо ви хочете протестувати 18 зразків, виберіть 16 зразків.
 - ▶ Зеленим кольором виділено запропоновані лунки й комбінації індексів зразків, які відповідають вимогам щодо різноманітності індексів.

Імпорт файлів маніфесту для прогону

- 1 Переконайтеся, що маніфести, які вам потрібно імпортувати, доступні в мережі або на USB-накопичувачі.
- 2 Натисніть **Import Manifests** (Імпорт маніфестів).
- 3 Перейдіть до файлу маніфесту й виберіть маніфести, які хочете додати.


ПРИМІТКА. Щоб зробити файли маніфесту доступними для всіх прогонів за допомогою модуля аналізу Germline Variant Module або Somatic Variant Module, додайте маніфести, використовуючи функцію Module Settings (Налаштування модуля). Ця функція вимагає дозволів на рівні адміністратора. Додаткову інформацію див. у документі *NextSeq 550Dx Instrument Reference Guide (Довідковий посібник до приладу NextSeq 550Dx)* (документ № 1000000009513).

Зазначення зразків для прогону


Укажіть зразки для прогону, використовуючи один із варіантів і наведені далі інструкції.

- ▶ **Enter samples manually** (Увести зразки вручну) — скористайтеся порожньою таблицею на екрані Create Run (Створити прогін).
- ▶ **Import samples** (Імпорт зразків) — перейдіть до зовнішнього файлу у форматі розділених комами значень (*.csv). Шаблон можна завантажити на екрані Create Run (Створити прогін).

Уведення зразків вручну.

- 1 Уведіть унікальну назву зразка (для **модуля аналізу Somatic Variant Module**) або ідентифікаційний номер зразка (для **модуля аналізу Germline Variant Module**).
Використовуйте літери та цифри, дефіс або підкреслювання.
- 2 **[Додатково]** Для зразків позитивного й негативного контролю клацніть правою клавшею миші й виберіть тип контролю.
Контроль у лунці одного зразка автоматично додається у відповідну лунку в іншому пулі з таким самим контролем.
- 3 **[Додатково]** Уведіть опис зразка в поле Sample Description (Опис зразка).
Використовуйте літери та цифри, дефіс або підкреслювання.
- 4 Виберіть адаптер індексу 1 у розкритому списку Index 1 (Індекс 1) (i7).
У разі використання запропонованих лунок зразка програмне забезпечення автоматично заповнює індексні адаптери i7 і i5, які відповідають вимогам щодо різноманітності індексів. Якщо точної кількості зразків для тестування немає в списку, обов'язково виберіть індексні адаптери для додаткових лунок.
- 5 Виберіть адаптер індексу 2 у розкритому списку Index 2 (Індекс 2) (i5).
- 6 Виберіть файл маніфесту з розкритого списку Manifest (Маніфест).
Зразкам у Pool A (Пул А) потрібен інший маніфест, ніж зразкам у Pool B (Пул В).
- 7 Оберіть варіант для перегляду, друку або збереження макета планшета як довідкового матеріалу для підготовки бібліотек.
 - Натисніть значок  **Print** (Друк), щоб відобразити макет планшета. Виберіть **Print** (Друк), щоб роздрукувати макет планшета.
 - Виберіть **Export** (Експорт), щоб експортувати інформацію про зразки до зовнішнього файлу.
- 8 Виберіть **Save Run** (Зберегти прогін).

Імпорт зразків

- 1 Виберіть **Import Samples** (Імпорт зразків) і перейдіть до розташування інформаційного файлу зразка. Ви можете імпортувати два типи файлів.
 - Виберіть **Template** (Шаблон) на екрані Create Run (Створити прогін), щоб створити новий макет планшета. Файл шаблону містить правильні заголовки стовпців для імпорту. Уведіть інформацію про зразок у кожний стовпець для зразків у прогоні. Видаліть приклад даних із невикористаних клітинок і збережіть файл.
 - Використайте файл з інформацією про зразок, який було експортовано з модуля Germline Variant Module або Somatic Variant Module за допомогою функції Export (Експорт).
- 2 Натисніть значок  **Print** (Друк), щоб відобразити макет планшета.
- 3 Виберіть **Print** (Друк), щоб роздрукувати макет планшета як довідковий матеріал під час підготовки бібліотек.
- 4 Виберіть **Save Run** (Зберегти прогін).

Підготовка картриджа з реагентами

Ретельно дотримуйтеся вказівок щодо картриджа з реагентами, щоб успішно виконати секвенування.

- 1 Вийміть картридж із реагентами з місця зберігання, у якому підтримується температура від -25 до -15 °C.
- 2 Виберіть один із наведених нижче методів розмороження реагентів. Не занурюйте картридж. Після розморожування картриджа висушіть його, перш ніж перейти до наступного етапу.

Температура	Тривалість розморожування	Межа стабільності
Водяна баня від 15 °C до 30 °C	60 хвилин	Не більше ніж 6 годин
Від 2 до 8 °C	7 годин	Не більше 5 днів

ПРИМІТКА. У разі розморожування декількох картриджів на одній водяній бані тривалість розморожування збільшується.

- 3 Переверніть картридж п'ять разів для перемішування реагентів.
- 4 Перевірте нижню частину картриджа, щоб переконатися, що реагенти відтанули й не містять осадів. Переконайтеся, що положення 29, 30, 31 і 32 відтанули, тому що вони найбільші й розморожуються найдовше.
- 5 Обережно постукайте по столу, щоб зменшити кількість бульбашок повітря.
Для отримання найкращих результатів перейдіть безпосередньо до завантаження зразка та налаштування прогону.

Підготовка проточної кювети

- 1 Витягніть коробку з новою проточною кюветою з місця для зберігання за температури від 2 °C до 8 °C.
- 2 Вийміть з коробки пакування з фольги й відкладіть для витримки за кімнатної температури протягом 30 хвилин.

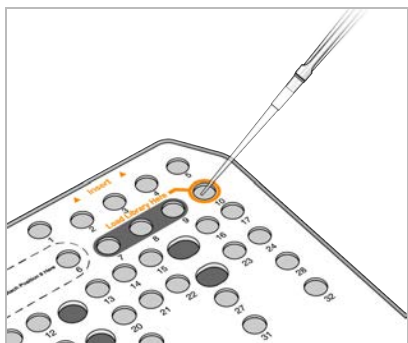
Підготовка бібліотек для секвенування

Денатуруйте й розведіть свої бібліотеки до об'єму завантаження 1,3 мл. На практиці концентрація завантаження може залежати від підготовки бібліотеки й методів кількісної оцінки. Розведення бібліотек зразків залежить від складності олігонуклеотидних пулів. Указівки щодо підготовки бібліотек зразків для секвенування, зокрема щодо розведення й об'єднання бібліотек, див. в розділі «Інструкції з використання» застосовного набору для підготовки бібліотек. Щільність кластерів у NextSeq 550Dx потрібно оптимізувати.

Завантаження бібліотек у картридж із реагентами

- 1 Очистьте фольгове ущільнення, що закриває резервуар № 10 із написом **Load Library Here** (Завантажити бібліотеку тут), з використанням низьковорсової серветки.
- 2 Проткніть ущільнення чистим наконечником піпетки об'ємом 1 мл.
- 3 Завантажте 1,3 мл підготовлених бібліотек у резервуар № 10, позначений написом **Load Library Here** (Завантажити бібліотеку тут). Уникайте дотику до фольгового ущільнення під час розподілу бібліотек.

Рисунок 1 Завантаження бібліотек



Налаштування прогону секвенування

- 1 Увійдіть до системи приладу NextSeq 550Dx, скориставшись своїм паролем до програмного забезпечення Local Run Manager.
- 2 На екрані Home (Головний екран) програмного забезпечення NOS виберіть **Sequence** (Послідовність).
- 3 Виберіть прогін зі списку й натисніть **Next** (Далі).
Декілька екранів налаштування прогону відкриються в такому порядку: Load Flow Cell (Завантаження проточної кювети), Load Buffer Cartridge (Завантаження картриджу з буферами), Load Reagent Cartridge (Завантаження картриджу з реагентами) і Pre-Run check (Перевірка перед прогоном).
- 4 Коли з'явиться екран Load Flow Cell (Завантаження проточної кювети), очистьте проточну кювету, а потім завантажте її.
 - ▶ Вийміть проточну кювету з фольгової упаковки.
 - ▶ Відкрийте прозорий пластиковий двостулковий контейнер і дістаньте проточну кювету.
 - ▶ Протріть скляну поверхню проточної кювети просоченою спиртом безворсовою серветкою. Досуша витріть скло низьковорсовою лабораторною серветкою.
 - ▶ Переконайтеся, що скляна поверхня проточної кювети чиста. За потреби повторіть етап очищення.
 - ▶ Зніміть використану проточну кювету від попереднього прогону.
 - ▶ Зіставте проточну кювету з напрямними штифтами й помістіть проточну кювету на платформу.
- 5 Виберіть **Load** (Завантажити).
Дверця автоматично закриваються, на екрані з'являється ідентифікатор проточної кювети, і перевіряються датчики.
- 6 Дотримуючись вказівок програмного забезпечення, випорожніть використаний контейнер із реагентом, завантажте картридж із буферами NextSeq 550Dx, а також картридж із реагентом NextSeq 550Dx.
Після завантаження картриджів із буферами й реагентом для NextSeq 550D програма зчитує та записує RFID. ID картриджів із буферами й реагентом з'являється на екрані, а також буде відзначено датчики.
- 7 Після завершення автоматизованої перевірки перед прогоном виберіть **Start** (Запуск). (Не потрібно, якщо налаштовано автоматичний запуск.)
- 8 Після початку прогону відкриється екран Sequencing (Секвенування). На цьому екрані ви побачите візуальне представлення виконання прогону, включно з даними про інтенсивність і показники оцінки якості (Q-показники).

Результати

Програмне забезпечення Real-Time Analysis (RTA) — це інтегроване програмне забезпечення, яке виконує аналіз зображень і розпізнавання основ, а потім присвоює оцінку якості до кожної основи для кожного циклу

секвенування. Після завершення первинного аналізу вибраний модуль програмного забезпечення Local Run Manager на приладі NextSeq 550Dx автоматично почне вторинний аналіз. Описані в цьому документі процеси вторинного аналізу стосуються модулів Germline Variant Module або Somatic Variant Module.

Демультіплексування

Під час демультіплексування порівнюється кожна послідовність зчитування індексу з послідовностями індексів, заданими для кожного прогону. На цьому етапі якісні значення не враховуються.

Зчитування індексів ідентифікуються під час описаних далі етапів.

- ▶ Зразки нумеруються від 1 залежно від порядку, у якому їх перелічено для прогону.
- ▶ Номер зразка 0 резервується для кластерів, які не було присвоєно зразку.
- ▶ Кластери присвоюються зразку в разі точної відповідності послідовності індексів або коли є одна розбіжність на зчитування індексу.

Генерування файлів FASTQ

Після демультіплексування програмне забезпечення генерує проміжні файли аналізу у форматі FASTQ, який є текстовим форматом для представлення послідовностей. Файли FASTQ містять зчитування для кожного зразка та пов'язані оцінки якості. Кластери, які не пройшли крізь фільтри, виключено.

Кожний файл FASTQ містить зчитування лише для одного зразка, а ім'я цього зразка включається в ім'я файлу FASTQ. У модулях Germline Variant Module і Somatic Variant Module генерується вісім файлів FASTQ на зразок на пул олігонуклеотидів: чотири зі зчитування 1 і чотири зі зчитування 2. Отже, створюється загалом 8 і 16 файлів FASTQ на зразок для варіантів зародкової лінії та соматичних варіантів відповідно. Файли FASTQ містять первинні вхідні дані для вирівнювання.

Вирівнювання

Під час етапу вирівнювання попарний алгоритм Сміта — Уотермана вирівнює кластери з одного зразка відносно послідовностей ампліконів, заданих у файлі маніфесту.

Попарний алгоритм Сміта — Уотермана виконує напівглобальне вирівнювання послідовностей для визначення схожих областей між двома послідовностями. Замість порівняння всієї послідовності алгоритм Сміта — Уотермана порівнює сегменти всіх можливих довжин.

Кожне зчитування парних кінців оцінюється з точки зору його вирівнювання щодо відповідних послідовностей зондів для цього зчитування.

- ▶ Зчитування 1 оцінюється відносно зворотно комплементарної ділянки низхідного локус-специфічного олігонуклеотиду (Downstream Locus-Specific Oligos, DLSO).
- ▶ Зчитування 2 оцінюється відносно висхідного локус-специфічного олігонуклеотиду (Upstream Locus-Specific Oligos, ULSO).
- ▶ Якщо початок зчитування збігається з послідовністю зонда не більше ніж із однією розбіжністю, повна довжина зчитування вирівнюється щодо цільового амплікона для цієї послідовності.
- ▶ Якщо початок зчитування збігається з послідовністю зонда не більше ніж із трьома відмінностями (розбіжності або зсуви через переважання інделів), повна довжина зчитування вирівнюється щодо цільового амплікона для цієї послідовності.
- ▶ Індели для DLSO й ULSO не спостерігалися з огляду на хімічні властивості аналізу.

Вирівнювання фільтруються з результатів вирівнювання на підставі частоти розбіжності для досліджуваної області або повного амплікона, залежно від довжини амплікона. Відфільтровані вирівнювання записуються у файли вирівнювання як невирівняні та не використовуються в розпізнаванні варіантів.

Розпізнавання варіантів

Програму розпізнавання варіантів Pisces розроблено для здійснення розпізнавань SNV й інделів із бібліотек, підготовлених для приладу.

Звіти й додаткові вихідні файли

Модулі аналізу варіантів створюють звіти у форматі PDF і текстових файлів із поділом табуляцією (*.txt), у яких відображаються показники на кшталт довжини секвенування й підрахунку варіантів. Крім того, модулі генерують вихідні файли, наприклад файли VCF або файли розпізнавання геномних варіантів (gVCF) для програм розпізнавання варіантів.

Процедури контролю якості

Програмне забезпечення приладу NextSeq 550Dx оцінює кожний прогін, кожний зразок і кожне розпізнавання основ за показниками контролю якості. Позитивний і негативний контролю також рекомендуються для підготовки бібліотеки та потребують оцінки. Оцінка контролів описана далі.

- **Негативний контроль (контроль без додавання матриці) або інший негативний контроль** має генерувати прогнозований результат. Якщо негативний контроль генерує результат, відмінний від прогнозованого, можливо, має місце помилка у відстеженні зразка, неправильне записування праймерів індексациї або забруднення.
- **Позитивний контрольний зразок** має генерувати прогнозований результат. Якщо позитивний контроль генерує результат, відмінний від прогнозованого, можливо, має місце помилка у відстеженні зразка або неправильне записування праймерів індексациї.

Технічні характеристики

Технічні характеристики для приладу NextSeq 550Dx було встановлено з використанням модулів Germline Variant Module і Somatic Variant Module зі спеціальним набором ампліконів TruSeq Custom Amplicon Kit Dx і набором реагентів NextSeq 550Dx High Output Reagent Kit v2 (300 cycles), а потім підтверджено за допомогою набору реагентів NextSeq 550Dx High Output Reagent Kit v2.5 (300 cycles). Дослідження включали індексацию зразків, перенесення зразків, уведення ДНК, аналітичну чутливість (межа холостої проби / межа виявлення), точність, прецизійність, порівняння методів і відтворюваність.

Аналітичні дослідження з використанням набору реагентів NextSeq 550Dx High Output Reagent Kit v2.5 (300 cycles) було призначено для оцінювання заявлених технічних характеристик, установлених попередньо за допомогою набору реагентів NextSeq 550Dx High Output Reagent Kit v2 (300 cycles). Результати демонструють, що набори реагентів (версії 2 і 2.5) мають зіставні технічні характеристики в разі використання спеціального набору ампліконів TruSeq Custom Amplicon Kit Dx. Технічні характеристики, пов'язані з преаналітичними факторами на кшталт методів екстракції або інтерферувальних речовин, див. в *інструкції з використання спеціального набору ампліконів TruSeq Custom Amplicon Kit Dx*.

Визначення обчислень, використаних у технічних характеристиках

- 1 Показник збігу позитивних результатів (Positive Percent Agreement, PPA) обчислюється як кількісне відношення локусів, класифікованих як варіанти за еталонним методом, які правильно реєструються в процесі аналізу.
 - ▶ (Кількість локусів варіантів, правильно зареєстрованих у процесі аналізу) / (загальна кількість локусів варіантів).
 - Локуси варіантів, зареєстровані в процесі аналізу, що відповідають еталонному методу, є істиннопозитивними (ІП). Локуси варіантів, зареєстровані як референсні розпізнавання або як інші розпізнавання варіантів у процесі аналізу, є хибнонегативними (ХН).
- 2 Показник збігу негативних результатів (Negative Percent Agreement, NPA) обчислюється як кількісне відношення локусів, класифікованих як дикий тип за еталонним методом, які правильно реєструються в процесі аналізу.
 - ▶ (Кількість локусів дикого типу, правильно зареєстрованих у процесі аналізу) / (загальна кількість локусів дикого типу)
 - Локуси дикого типу, зареєстровані в процесі аналізу, що відповідають еталонному методу, є істинно негативними (ІН). Локуси дикого типу, зареєстровані в процесі аналізу як варіанти, є хибнопозитивними (ХП).

- 3 Відсоток загальної узгодженості (Overall Percent Agreement, OPA) обчислюється як кількісне відношення локусів, які правильно реєструються в процесі аналізу відносно еталонного методу.
 - ▶ ((Кількість локусів варіантів, відображених аналізом) + (кількість локусів дикого типу, відображених аналізом)) / ((загальна кількість локусів варіантів) + (загальна кількість локусів дикого типу)).
- 4 Обчислення параметрів PPA, NPA і OPA не включає нерозпізнавання (варіантні або референсні локуси, які не відповідають одному чи декільком фільтрам якості).
- 5 Частота аутосомних розпізнавань обчислюється як загальна кількість локусів, які пройшли фільтр, поділена на загальну кількість позицій, секвенованих для хромосом 1–22. Хромосоми X та Y виключаються. Цей показник не враховує узгодженість розпізнавань з еталонним методом.

Технічні характеристики набору реагентів NextSeq 550Dx High Output Reagent Kit v2 (300 Cycles)

Індексація зразків

Праймери індексу зразка, додані під час підготовки бібліотеки, присвоюють унікальну послідовність кожному зразку ДНК. Ці унікальні послідовності дають змогу об'єднувати декілька зразків у пул в одному прогоні секвенування. Індексація зразків використовується в робочому процесі як для генів зародкової лінії, так і для соматичних генів. Мета цього дослідження — установити мінімальну (8) і максимальну (96) кількість зразків, які можна обробити в одному прогоні секвенування за допомогою приладу NextSeq 550Dx. Вісім унікальних зразків Platinum Genome було протестовано з 12 різними комбінаціями праймерів індексації на зразок. Результати зразків із чотирьох прогонів секвенування за допомогою модуля Germline Variant Module було порівняно з результатами зразків Platinum Genomes версії 2016-1.0.

Для першого набору прогонів 96 унікально індексованих бібліотек зразків було протестовано за допомогою репрезентативного аналізу, призначеного для пошуку низки генів, які покривають 12 588 основ на нитку в усіх 23 хромосомах людини, щоб перевірити здатність аналізу стабільно виконувати розпізнавання під час генотипування заданого зразка в різних комбінаціях праймерів індексації. Для другого набору прогонів вісім унікально індексованих бібліотек зразків було секвеновано у двох прогонах секвенування, щоб верифікувати мінімальну кількість підтримуваних індексів.

Для 96-індексних прогонів показник PPA для SNV варіювався від 98,7 % до 100 %, PPA для інсерцій і делецій становив 100 %, а NPA становив 100 % для кожної з 96 комбінацій індексів. У 8-індексних прогонах значення PPA становили 100 % (SNV, інсерції та делеції), а NPA — 100 % для кожної з восьми комбінацій індексів.

Перенесення зразків

Прилад NextSeq 550Dx дає змогу виконувати секвенування декількох зразків, а також контроль в одному прогоні секвенування. Дослідження було проведено для оцінювання перенесення зразків у межах прогону секвенування (у межах прогону) і між прогонами секвенування (між прогонами). Два зразки Platinum Genome (один із них отримано від чоловіка, інший — від жінки) було протестовано з використанням репрезентативного аналізу, призначеного для пошуку різних генів, що покривають 12 588 основ (150 ампліконів) у 23 різних хромосомах, включно з обома статевими хромосомами. Бібліотеки секвенувалися на приладі NextSeq 550Dx за допомогою модуля Germline Variant Module. Перенесення чоловічих зразків у жіночі було виявлено внаслідок наявності читувань амплікона Y-хромосоми в жіночих зразках.

Перенесення в межах прогону може відбуватися під час генерування кластерів, циклів індексації, розпізнавання основ і демультіплексування зразків. Щоб перевірити перенесення зразка в межах прогону секвенування, бібліотеку, яка складалась із 46 повторностей для кожного зразка — чоловічого й жіночого, — а також чотирьох контрольних зразків без додавання матриці, секвенували один раз на приладі NextSeq 550Dx. Перенесення зразка в межах прогону було оцінено під час порівняння покриття амплікона Y-хромосоми кожної жіночої повторності з покриттям амплікона всіх чоловічих повторностей у пулі. Медіана отриманого показника перенесення в межах прогону становила 0,084 %.

Для тестування перенесення зразка між прогонами два пули бібліотек було підготовано й секвеновано послідовно на одному приладі NextSeq 550Dx. Перший пул містив 46 повторностей жіночого зразка, а також два контролю без додавання матриці. Другий пул містив 46 повторностей чоловічого зразка, а також два контролю без додавання матриці. В обох пулах використано однаковий набір адаптерів індексу. Жіночий пул секвенували першим, після чого виконували секвенування чоловічого пулу. Потім жіночий пул секвенували ще раз. Перенесення зразка між прогонами оцінювалося під час порівняння покриття амплікона Y-хромосоми між відповідними повторностями повторного прогону жіночого пулу й прогону чоловічого пулу. Медіана отриманого показника перенесення між прогонами становила 0,0076 %.

Уведення ДНК

Кров (зародкова лінія)

Для приладу NextSeq 550Dx було встановлено діапазон уведення ДНК зі зразків крові для підготовки бібліотеки за допомогою спеціального набору ампліконів TruSeq Custom Amplicon Kit Dx із використанням робочого процесу модуля Germline Variant Module. Цей діапазон було оцінено за допомогою дослідження серійного розведення з використанням 13 зразків Platinum Genome; дослідження включало репрезентативний аналіз, призначений для пошуку різних генів, які покривають 12 588 основ для 23 різних хромосом. Бібліотеку було секвеновано на двох приладах NextSeq 550Dx із використанням однієї серії набору реагентів NextSeq 550Dx High Output Reagent Kit v2 (300 cycles).

П'ять зразків було перевірено у двох повторностях на п'яти рівнях уведення ДНК від 250 нг до 12 нг (250 нг, 100 нг, 50 нг, 25 нг і 12 нг). Вісім зразків було перевірено в одній повторності на кожному з п'яти рівнів уведення ДНК. Для визначення точності зразки генотипів було порівняно зі зразками Platinum Genomes версії 2016-1.0. Результати було визначено для кожного рівня введення. Значення PPA для кожного типу варіантів (SNV, інсерції та делеції) представлено в Таблиця 1. Значення NPA представлено в Таблиця 2. Усі рівні введення мають подібну точність. Рекомендоване введення ДНК для спеціального набору ампліконів TruSeq Custom Amplicon Kit Dx становить 50 нг за значень нижньої межі на рівні 25 нг і верхньої межі на рівні 100 нг для відповідності технічним характеристикам.

Таблиця 1 Результати PPA для кожного введення ДНК за типом варіантів

Уведення ДНК (нг)	Тип варіанта	Прогнозовані варіанти	ІП	ХН	Варіанти без розпізнавань	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	Інсерція	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	Делеція	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

Таблиця 2 NPA для кожного введення ДНК

Уведення ДНК (нг)	ІН	ХП	Референсні нерозпізнавання	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 (Соматичні варіанти)

Для приладу NextSeq 550Dx було встановлено діапазон введення ДНК із зафіксованої формаліном і залитої парафіном (formalin-fixed paraffin-embedded, FFPE) тканини для підготовки бібліотеки за допомогою спеціального набору ампліконів TruSeq Custom Amplicon Kit Dx із використанням робочого процесу модуля Somatic Variant Module. Діапазон введення ДНК було оцінено за допомогою дослідження серійного розведення з використанням трьох зразків Platinum Genome. Дослідження включало репрезентативний аналіз, призначений для пошуку різних генів, які покривають 12 588 основ для 23 різних хромосом. Клітинні лінії Platinum Genome GM12878 і GM12877 фіксувались у формаліні та заливалися парафіном, після чого було проведено екстракцію ДНК. Зразки лінії GM12878 було розведено зразками лінії GM12877 таким чином, щоб частоти варіантних алелів (variant allele frequencies, VAF) для 81 варіанта (55 SNV, 10 інсерцій і 16 делецій) становили приблизно 0,025, 0,05 або 0,10. Крім того, кожний зразок містив 91 варіант із вищими значеннями частоти, які сягали 1,0 для VAF. Зразки було оброблено у двох повторностях на п'яти рівнях введення ДНК із середньою різницею циклів кількісного аналізу (delta quantitative cycle, dCq), яка становила 2,1, 3,6, 4,6, 6,0 і 7,8 відповідно за результатами вимірювання за допомогою спеціального набору ампліконів TruSeq Custom Amplicon Kit Dx — FFPE QC. Кожну бібліотеку було секвеновано на двох приладах NextSeq 550Dx із використанням двох серій набору реагентів NextSeq 550Dx High Output Reagent Kit v2 (300 cycles). Для визначення точності отримані показники розпізнавання варіантів зразків було порівняно з Platinum Genomes версії 2016-1.0. Значення PPA для кожного типу варіанта (SNV, інсерції та делеції) представлено в [Таблиця 3](#). Значення NPA представлено в [Таблиця 4](#). Рекомендоване введення ДНК для варіантів із VAF 0,05 або вище характеризується показником dCq ≤ 4 , водночас значення 4,6 відповідає нижній межі для дотримання технічних характеристик.

Таблиця 3 Результати PPA для кожного введення ДНК за типом варіантів

Середня dCq	Тип варіанта	Прогнозовані варіанти	Прогнозовані нерозпізнавання	Цільова VAF розведення					
				0,025		0,05		0,10	
				Варіанти без розпізнавань	PPA (%)	Варіанти без розпізнавань	PPA (%)	Варіанти без розпізнавань	PPA (%)
2,1	SNV	808	Не застосовно.	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	Інсерція	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	Делеція	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

Таблиця 4 NPA для кожного введення ДНК

Середня dCq	Прогнозований дикий тип	Цільова VAF розведення					
		0,025		0,05		0,10	
		Референсні нерозпізнання	NPA (%)	Референсні нерозпізнання	NPA (%)	Референсні нерозпізнання	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

Аналітична чутливість (межа холостої проби (МХП) і межа виявлення (МВ))

Це дослідження було проведено для оцінювання межі холостої проби (МХП) і межі виявлення (МВ) для модуля Somatic Variant Module на приладі NextSeq 550Dx. У дослідженні було застосовано репрезентативний аналіз для пошуку низки генів, які покривають 12 588 основ для 23 різних хромосом. Клітинні лінії Platinum Genome GM12878 і GM12877 фіксувались у формаліні та заливалися парафіном, після чого було проведено екстракцію ДНК. Матеріал лінії GM12878 було розведено матеріалом лінії GM12877 таким чином, щоб частоти 74 варіантів (53 SNV, 7 інсерцій і 14 делецій) становили приблизно $0,05 \pm 0,02$. Матеріал лінії GM12877 і розведений матеріал лінії GM12878 (GM12878-D) було протестовано протягом шести днів запуску поспіль із використанням одного приладу та з чергуванням двох серій набору реагентів NextSeq 550Dx High Output Reagent Kit v2 (300 cycles), загалом упродовж шести прогонів секвенування. У результаті цього тесту було отримано 60 повторностей для кожного варіанта лінії GM12878-D і 72 повторності для кожної відповідної координати дикого типу лінії GM12877 для кожної серії реагентів. Межу холостої проби (МХП) і МВ було обчислено за допомогою класичного підходу, описаного в стандарті CLSI EP17-A2, з використанням непараметричних методів оцінки. МХП і МВ було обчислено для SNV, інсерцій і делецій окремо за допомогою об'єднання частот варіантів для заданого типу варіанта. Помилку першого роду було встановлено на рівні 0,01, помилку другого роду — на рівні 0,05.

Для МХП пул частот варіантів було відсортовано від найнижчої до найвищої частоти, також для кожної серії реагентів для кожного типу варіанта було обчислено 99-те значення рангу (Таблиця 5). У модулі Somatic Variant Module для визначення межі якісного виявлення варіантів використовується граничне значення (межа вимірювання МХП) із VAF, яка становить 0,026. Обчислені МХП для цього граничного значення дають помилку першого роду не більше 0,01.

Таблиця 5 Межа холостої проби

Тип варіанта	Загалом спостережень	МХП, серія реагентів 1 (%)	МХП, серія реагентів 2 (%)
SNV	3816	0,77	0,77
Інсерція	504	0,56	0,56
Делеція	1008	1,20	1,20

Для МВ було обчислено відсоток частоти окремої мутації для кожної серії реагентів для кожного типу варіанта, який був нижче граничного значення 0,026 (Таблиця 6). Значення відсотків були менше помилки другого роду в 5 % (0,05), тому медіану суми частот варіантів було обчислено як МВ (Таблиця 6). МВ для кожного типу варіанта було прийнято як найбільше з двох значень, обчислених для двох серій реагентів — 4,97 % для SNV, 5,12 % для інсерцій і 5,26 % для делецій.

Таблиця 6 Межа виявлення

Серія реагентів	Тип варіанта	Загалом спостережень	К-сть вимірювань VAF < 2,6 %	% вимірювань VAF < 2,6 %	Межа виявлення (%)
1	SNV	3180	53	1,7	4,94
	Інсерція	420	6	1,4	5,08
	Делеція	840	7	0,8	5,22
2	SNV	3180	51	1,6	4,97
	Інсерція	420	5	1,2	5,12
	Делеція	840	7	0,80	5,26

Точність

Зародкова лінія

Це дослідження було проведено для оцінювання точності розпізнавання варіантів за допомогою модуля Germline Variant Module на приладі NextSeq 550Dx із використанням набору реагентів NextSeq 550Dx High Output Reagent Kit v2 (300 cycles). Тринадцять унікальних зразків Platinum Genome було протестовано за допомогою репрезентативного аналізу, призначеного для пошуку низки генів, які покривають 12 588 основ (150 ампліконів) для 23 різних хромосом. Загалом було виконано дев'ять прогонів із використанням трьох приладів секвенування, трьох серій реагентів і трьох операторів протягом п'яти днів запуску. Точність було встановлено для SNV, інсерцій й делецій за допомогою порівняння отриманих результатів із показниками, установленними за допомогою точно охарактеризованого складеного еталонного методу — Platinum Genomes версії 2016-1.0. Довірчі геномні області було визначено на підставі цього еталонного методу, якщо не вказано інше.

Таблиця 7 Зведені дані узгодженості для варіантів зародкової лінії

Критерії	Загальні спостереження ¹	Результат за спостереженням ²	Результат за прогоном ³
PPA для SNV	819	98,7	> 99,9
PPA для інсерцій	819	95,0	98,9
PPA для делецій	819	100	100
NPA	819	100	100
OPA	819	> 99,9	> 99,9

¹Обчислено як кількість зразків на прогін (91) × кількість прогонів (9) = 819.

²Найменше отримане значення за повторністю зразка в усіх 9 прогонах.

³Найменше значення, коли дані з кожного прогону аналізуються сукупно.

У Таблиця 8 наведено дані дослідження з відсотками збігу позитивних і негативних результатів залежно від зразка, де результати для варіантів порівняно з даними Platinum Genomes версії 2016-1.0 для обчислення PPA. Поєднано три типи варіантів (SNV, інсерції та делеції). Еталонний метод дає результати лише для одонуклеотидних варіантів та інсерцій/делецій, тому для обчислення NPA результати розпізнавання основ, які не належать до варіантів, порівнюються зі збіркою 19 референсної послідовності геному людини (hg19).

Таблиця 8 Узгодженість варіантів зародкової лінії за зразками

Зразок	Середня частота розпізнавань	Прогнозовані варіанти ¹	ІП	ХН	Варіанти без розпізнавань	ІН	ХП	РРА	НРА	ОРА
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

¹ Загальна кількість варіантів у всіх повторностях зразків для 9 прогонів.

У **Таблиця 9** містяться дані, представлені залежно від зразка, де результати варіантів порівняно з показниками, установленими за допомогою точно охарактеризованого складеного еталонного методу. Виявлення оцінюється для кожного типу варіанта — SNV, інсерції й делеції окремо. Референсні позиції виключено.

Таблиця 9 Узгодженість варіантів зародкової лінії для зразків за типом варіанта

>Зразок	SNV			Інсерції			Делеції		
	>Прогнозовано	>ІП	>ХН	>Прогнозовано	>ІП	>ХН	Прогнозовано	ІП	ХН
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

Зразки було додатково проаналізовано для розпізнавання незначних інсерцій і делецій (інделів). Загальні зведені дані наведено в **Таблиця 10**. Загалом було виявлено 71 індел розміром 1–24 п. о. для інсерцій і 1–25 п. о. для делецій.

Таблиця 10 Зведені дані виявлення інделів для зародкової лінії

Тип варіанта	Прогнозовані варіанти	ІП	ХН	Варіанти без розпізнавань	PPA
Інсерція	18522	18018	27	477	99,9
Делеція	17388	17073	0	315	100

Репрезентативний аналіз складався зі 150 ампліконів, які покривали різноманітний геномний матеріал. Уміст GC в ампліконах варіювався в межах 0,19–0,87. Також амплікони мали низку одонуклеотидних (наприклад, полі-А, полі-Т), динуклеотидних і тринуклеотидних повторів. Дані було скомпільовано залежно від амплікона (Таблиця 11) для визначення впливу геномного вмісту на відсоток правильних розпізнавань. Відсоток правильних розпізнавань складається з варіантних і референсних розпізнавань, він менше 100 %, якщо наявні неправильні розпізнавання або нерозпізнавання.

Таблиця 11 Точність на рівні ампліконів для зародкової лінії

Амплікон	Хромо-сома	Початок амплікона	Кінець амплікона	Розмір проаналізованого фрагмента	Основи в довірчих зонах	Геномний уміст амплікона	Вміст GC	Правильні розпізнавання	Неправильні розпізнавання	Нерозпізнавання	Відсоток правильних розпізнавань
1	1	36450499	36450591	93	93	Індел	0,22	76167	0	0	100
2	1	109465122	109465200	79	79	Полі-А (5), полі-С (5), індел	0,38	64701	0	0	100
3	1	218353867	218353957	91	91	Індел	0,4	74529	0	0	100
4	1	223906657	223906748	92	92	Індел	0,49	75348	0	0	100
5	1	228526602	228526682	81	81	Полі-С (5)	0,69	66339	0	0	100
6	1	236372039	236372108	70	70	Полі-Т (10), індел	0,39	57330	0	0	100
7	1	247812041	247812128	88	88	Полі-А (5), СТ (3), ТАА (3), індел	0,27	72072	0	0	100
8	2	55862774	55862863	90	90	Індел	0,28	73710	0	0	100
9	2	87003930	87004009	80	80	Індел	0,38	65520	0	0	100
10	2	177016721	177016805	85	81	Н/Д	0,65	66339	0	0	100
11	2	186625727	186625801	75	75	Полі-А (8)	0,35	61425	0	0	100
12	2	190323504	190323591	88	88	Полі-Т (5)	0,42	72072	0	0	100
13	2	200796740	200796826	87	87	Полі-Т (5), індел	0,31	71253	0	0	100
14	2	212245049	212245139	91	91	Полі-Т (5), полі-А (6), індел	0,3	74529	0	0	100
15	2	228147052	228147144	93	93	Індел	0,43	76167	0	0	100
16	2	235016350	235016422	73	73	Полі-Т (5), індел	0,42	59787	0	0	100
17	3	4466229	4466321	93	93	АТ (3), індел	0,27	74823	0	1344	98,2
18	3	46620561	46620643	83	83	Н/Д	0,43	67977	0	0	100
19	3	49851331	49851400	70	70	СТ (3), індел	0,49	57330	0	0	100

Амплікон	Хромо-сома	Початок амплікона	Кінець амплікона	Розмір проаналізованого фрагмента	Основи в довірчих зонах	Геномний уміст амплікона	Вміст GC	Правильні розпізнавання	Неправильні розпізнавання	Нерозпізнавання	Відсоток правильних розпізнавань
20	3	189713161	189713248	88	88	Полі-А (5), полі-Т (5), полі-А (9), TG (3)	0,41	72072	0	0	100
21	3	190106030	190106104	75	74	Індел	0,57	60543	0	63	99,9
22	4	2233667	2233744	78	78	Полі-А (6)	0,26	63882	0	0	100
23	4	7780541	7780637	97	97	Полі-G (6), полі-Т (5), полі-А (5)	0,42	79443	0	0	100
24	4	15688604	15688681	78	78	Н/Д	0,29	63882	0	0	100
25	4	56236521	56236586	66	62	Полі-А (5), індел	0,36	50778	0	0	100
26	4	102839244	102839314	71	69	Полі-А (5)	0,46	56511	0	0	100
27	4	164446743	164446804	62	62	Полі-А (7), індел	0,27	50778	0	0	100
28	5	1882081	1882158	78	75	Н/Д	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	Н/Д	0,39	52416	0	0	100
31	5	74077114	74077196	83	83	Полі-А (6), індел	0,3	67977	0	0	100
32	5	147475343	147475409	67	67	Полі-Т (5)	0,37	54873	0	0	100
33	5	149323731	149323821	91	91	СТ (4), AG (3)	0,55	74529	0	0	100
34	5	155662213	155662287	75	75	Індел	0,43	61425	0	0	100
35	6	6318713	6318814	102	102	Полі-G (6)	0,68	83538	0	0	100
36	6	24949983	24950074	92	92	Індел	0,63	75348	0	0	100
37	6	31084900	31084999	100	94	GCT (5), індел	0,61	76608	0	378	99,5
38	6	32147987	32148084	98	98	Полі-Т (5), TCT (3), CTT (3)	0,55	80262	0	0	100
39	6	32986864	32986958	95	95	Індел	0,53	77805	0	0	100
40	6	33408498	33408583	86	86	Полі-С (6)	0,7	70434	0	0	100
41	6	41647401	41647495	95	94	Полі-G (5), індел	0,61	76986	0	0	100
42	6	112435865	112435955	91	91	Полі-А (5)	0,44	74529	0	0	100

Амплікон	Хромо-сома	Початок амплікона	Кінець амплікона	Розмір проаналізованого фрагмента	Основи в довірчих зонах	Геномний уміст амплікона	Вміст GC	Правильні розпізнавання	Неправильні розпізнавання	Нерозпізнавання	Відсоток правильних розпізнавань
43	7	22202076	22202148	73	73	Н/Д	0,44	59787	0	0	100
44	7	66276100	66276187	88	88	Індел	0,35	72072	0	0	100
45	7	77365735	77365821	87	87	Полі-А (7), АГ (4)	0,26	71253	0	0	100
46	7	110939946	110940030	85	85	Індел	0,38	69615	0	0	100
47	7	128533468	128533557	90	90	Полі-Г (5), індел	0,62	73710	0	0	100
48	7	149503875	149503965	91	91	Полі-Г (6), полі-С (6), індел	0,71	74529	0	0	100
49	7	154404519	154404599	81	66	Н/Д	0,31	54054	0	0	100
50	7	156476507	156476599	93	93	Індел	0,35	76167	0	0	100
51	8	1817312	1817394	83	83	Н/Д	0,42	67977	0	0	100
52	8	24811020	24811109	90	89	Полі-Г (7), СТС (4), індел	0,61	72171	0	720	99,0
53	8	76518625	76518691	67	67	Індел	0,3	54873	0	0	100
54	9	103054909	103055006	98	98	Полі-Г (6)	0,67	80262	0	0	100
55	9	105586150	105586214	65	65	Індел	0,32	53235	0	0	100
56	9	107620823	107620918	96	96	Н/Д	0,49	78624	0	0	100
57	9	123769149	123769231	83	83	АТ (3)	0,37	67977	0	0	100
58	9	138995345	138995441	97	97	Полі-С (6), індел	0,68	79443	0	0	100
59	10	5987120	5987198	79	78	Полі-Г (5), індел	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	Полі-Т (5)	0,3	64701	0	0	100
62	10	33018351	33018440	90	90	Полі-А (5), полі-Т (5)	0,2	73710	0	0	100
63	10	45084159	45084253	95	95	Індел	0,35	77805	0	0	100
64	10	55892599	55892687	89	88	АС (11), індел	0,42	71747	0	325	99,5
65	10	101611250	101611329	80	80	Н/Д	0,49	65520	0	0	100
66	10	118351373	118351453	81	81	Н/Д	0,51	66339	0	0	100

Амплікон	Хромо- сома	Початок амплікона	Кінець амплікона	Розмір проаналізо- ваного фрагмента	Основи в довірчих зонах	Геномний уміст амплікона	Вміст GC	Правильні розпізна- вання	Неправильні розпізна- вання	Нерозпізна- вання	Відсоток правильних розпізнавань
67	11	8159816	8159912	97	96	Н/Д	0,45	78624	0	0	100
68	11	30177648	30177717	70	70	Індел	0,46	57330	0	0	100
69	11	47470345	47470444	100	100	Н/Д	0,65	81900	0	0	100
70	11	59837679	59837740	62	62	Індел	0,37	50778	0	0	100
71	11	64418856	64418957	102	102	Н/Д	0,59	83538	0	0	100
72	11	93529612	93529684	73	73	Полі-А (5)	0,4	59787	0	0	100
73	11	101347052	101347136	85	85	Н/Д	0,42	69615	0	0	100
74	11	102477336	102477426	91	91	Полі-Г (6)	0,55	74529	0	0	100
75	11	118406285	118406369	85	85	Індел	0,53	69615	0	0	100
76	11	120357801	120357885	85	85	Полі-А (5), СА (3), індел	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	Полі-С (5), індел	0,52	68796	0	0	100
79	12	26811004	26811096	93	93	Полі-А (7), АС (4)	0,33	76167	0	0	100
80	12	30881766	30881846	81	81	Н/Д	0,49	66339	0	0	100
81	12	88474105	88474175	71	71	Полі-А (6)	0,35	58149	0	0	100
82	12	120966872	120966966	95	95	Полі-Г (5)	0,68	77805	0	0	100
83	13	24167504	24167576	73	73	Н/Д	0,52	59787	0	0	100
84	13	25816961	25817049	89	88	Полі-А (5), полі-Т (7), полі-А (7), індел	0,22	72072	0	0	100
85	13	44880112	44880200	89	89	Індел	0,49	72891	0	0	100
86	13	77665218	77665294	77	77	Індел	0,39	63063	0	0	100
87	14	31619327	31619393	67	67	GA (3), ТА (3)	0,39	54873	0	0	100
88	14	39517884	39517966	83	83	Н/Д	0,25	67977	0	0	100
89	14	46958962	46959034	73	72	Полі-Т (5), індел	0,19	58642	0	326	99,4
90	14	58050030	58050110	81	81	Індел	0,38	66339	0	0	100

Амплікон	Хромо-сома	Початок амплікона	Кінець амплікона	Розмір проаналізованого фрагмента	Основи в довірчих зонах	Геномний уміст амплікона	Вміст GC	Правильні розпізнавання	Неправильні розпізнавання	Нерозпізнавання	Відсоток правильних розпізнавань
91	14	82390559	82390649	91	91	Індел	0,35	74529	0	0	100
92	14	92549544	92549609	66	66	Полі-А (5)	0,41	54054	0	0	100
93	14	102808496	102808589	94	94	Індел	0,62	76986	0	0	100
94	15	43170751	43170848	98	96	Полі-С (5)	0,45	78624	0	0	100
95	15	63446149	63446216	68	68	Індел	0,25	55692	0	0	100
96	15	77879807	77879901	95	93	Полі-Г (5), індел	0,68	76167	0	0	100
97	15	81625334	81625428	95	95	Полі-Т (6)	0,43	77805	0	0	100
98	15	85438263	85438334	72	71	Індел	0,65	58149	0	0	100
99	15	89817413	89817503	91	91	Н/Д	0,36	74529	0	0	100
100	15	89864274	89864343	70	70	Індел	0,56	57330	0	0	100
101	16	1894910	1894972	63	63	Н/Д	0,27	51597	0	0	100
102	16	28997904	28997998	95	95	Полі-С (5)	0,67	77805	0	0	100
103	16	53682908	53682994	87	87	ТА (3)	0,41	71253	0	0	100
104	16	57954406	57954509	104	104	Полі-С (5)	0,67	85176	0	0	100
105	16	85706375	85706465	91	91	Полі-Т (5), індел	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	Полі-С (5), індел	0,67	71247	0	6	100
108	17	3970090	3970180	91	91	Індел	0,46	74529	0	0	100
109	17	16084945	16085037	93	93	Індел	0,26	76167	0	0	100
110	17	33998759	33998849	91	89	Полі-Т (5)	0,54	72891	0	0	100
111	17	39589691	39589774	84	82	Полі-А (13), індел (x2)	0,29	66343	27	788	98,8
112	17	41244394	41244484	91	91	Полі-А (5)	0,34	74529	0	0	100
113	17	45438866	45438957	92	92	Полі-А (7), АТ (3), АТ (4), АТ (4), індел	0,26	75348	0	0	100
114	17	61502432	61502510	79	79	Індел	0,41	64413	0	288	99,6

Амплікон	Хромо-сома	Початок амплікона	Кінець амплікона	Розмір проаналізованого фрагмента	Основи в довірчих зонах	Геномний уміст амплікона	Вміст GC	Правильні розпізнавання	Неправильні розпізнавання	Нерозпізнавання	Відсоток правильних розпізнавань
115	17	64023582	64023667	86	86	Полі-Т (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	Н/Д	0,37	74529	0	0	100
119	18	9888026	9888094	69	69	Полі-А (6), TG (3)	0,43	56511	0	0	100
120	18	38836999	38837073	75	75	Полі-А (5), індел	0,37	61425	0	0	100
121	18	47405382	47405462	81	81	СТС (3), індел	0,47	66339	0	0	100
122	18	54815665	54815749	85	85	СТ (3), індел	0,45	69615	0	0	100
123	18	59773996	59774060	65	65	Н/Д	0,48	53235	0	0	100
124	19	625143	625241	99	99	Н/Д	0,59	81081	0	0	100
125	19	18121418	18121491	74	74	Н/Д	0,68	60605	1	0	100
126	19	18186574	18186643	70	70	Н/Д	0,64	57330	0	0	100
127	20	746056	746149	94	94	Н/Д	0,61	76986	0	0	100
128	20	10633195	10633276	82	82	АС (3)	0,59	67158	0	0	100
129	20	17705633	17705708	76	76	СТ (3)	0,58	62244	0	0	100
130	20	21766821	21766890	70	70	GT (3), TG (4), індел	0,46	57330	0	0	100
131	20	25278421	25278521	101	101	Індел	0,63	82719	0	0	100
132	20	50897302	50897368	67	67	Індел	0,36	54873	0	0	100
133	20	62331904	62331994	91	88	Полі-G (6)	0,73	72072	0	0	100
134	20	62690860	62690946	87	87	Індел	0,57	71253	0	0	100
135	21	30300823	30300888	66	66	Індел	0,35	54054	0	0	100
136	21	33694176	33694273	98	98	Полі-Т (6), СА (3)	0,54	80262	0	0	100
137	21	36710706	36710792	87	87	GT (3), індел	0,39	71253	0	0	100
138	21	46644924	46644992	69	69	Полі-А (6), AG (3), індел	0,32	56439	0	72	99,9

Амплікон	Хромо- сома	Початок амплікона	Кінець амплікона	Розмір проаналізо- ваного фрагмента	Основи в довірчих зонах	Геномний уміст амплікона	Вміст GC	Правильні розпізна- вання	Неправильні розпізна- вання	Нерозпізна- вання	Відсоток правильних розпізнавань
139	21	46705575	46705664	90	90	Полі-Т (5), полі-А (6)	0,5	73710	0	0	100
140	22	25750774	25750873	100	100	Індел	0,63	81900	0	0	100
141	22	32439233	32439329	97	97	Н/Д	0,68	79443	0	0	100
142	22	37409844	37409940	97	97	Індел	0,46	79443	0	0	100
143	22	37637596	37637694	99	99	Н/Д	0,6	81081	0	0	100
144	22	47081347	47081438	92	92	Індел	0,66	75348	0	0	100
145	X	15870424	15870492	69	69	Полі-Т (5)	0,26	56511	0	0	100
146	X	135288543	135288611	69	69	Полі-С (5)	0,62	56511	0	0	100
147	X	135290777	135290847	71	71	Н/Д	0,52	58149	0	0	100
148	Y	2655397	2655461	65	0	Н/Д	0,55	0	0	0	Н/Д
149	Y	2655519	2655609	91	0	Н/Д	0,48	0	0	0	Н/Д
150	Y	2655609	2655679	71	0	Полі-А (5)	0,37	0	0	0	Н/Д

Результати секвенування для зразка лінії NA12878 було порівняно з генотипом із високим ступенем вірогідності для NA12878, затвердженим Національними інститутами стандартів і технологій (National Institutes of Standards and Technology, NIST) (версія 2.19). 92 зі 150 ампліконів було виявлено повністю в межах геномних зон із високим ступенем вірогідності, 41 амплікон мав часткове перекриття, і 17 ампліконів не мали перекриття в послідовності, затвердженій NIST. Цей результат призвів до появи 10 000 координат на повторність для порівняння. Результати розпізнавання основ, які не належали до варіантів, було порівняно зі збіркою референсної послідовності геному людини hg19. Результати дослідження точності наведено в [Таблиця 12](#).

Таблиця 12 Узгодженість зразків лінії NA12878 із базою даних NIST для зародкової лінії

Зразок	Кількість ампліконів	Середня частота розпізнавань	ІП	ХН	ІН	ХП	РРА	NPA	OPA
NA12878	133	> 99,9	6552	1	610470	0	> 99,9	100	> 99,9

На підставі даних, отриманих у цьому дослідженні матеріалу зародкової лінії протягом дев'яти прогонів, прилад NextSeq 550Dx може секвенувати з отриманням стабільних результатів указані далі послідовності.

- ▶ Зі вмістом GC $\geq 19\%$ (усі основи, що розпізнавались у 819 секвенованих ампліконах зі вмістом GC 19%, розпізнано правильно з частотою нерозпізнавання 0,6%)
- ▶ Зі вмістом GC $\leq 87\%$ (усі основи, що розпізнавались у 819 секвенованих ампліконах зі вмістом GC 87%, розпізнано правильно без нерозпізнавань)
- ▶ Фрагменти полі-А довжиною ≤ 9 (усі основи, що розпізнавались у 819 секвенованих ампліконах із повтором полі-А з дев'яти нуклеотидів, розпізнано правильно без нерозпізнавань)
- ▶ Фрагменти полі-Т довжиною ≤ 10 (усі основи, що розпізнавались у 819 секвенованих ампліконах із повтором полі-Т з десяти нуклеотидів, розпізнано правильно без нерозпізнавань)
- ▶ Фрагменти полі-Г довжиною ≤ 7 (усі основи, що розпізнавались у 819 секвенованих ампліконах із повтором полі-Г із семи нуклеотидів, розпізнано правильно з частотою нерозпізнавання 1,0%)
- ▶ Фрагменти полі-С довжиною ≤ 6 (усі основи, що розпізнавались у 2457 секвенованих ампліконах із повтором полі-С із шести нуклеотидів, розпізнано правильно без нерозпізнавань)
- ▶ Довжини динуклеотидних повторів $\leq 11x$ (усі основи, що розпізнавались у 819 секвенованих ампліконах з 11-кратним (11x) динуклеотидним повтором, розпізнано правильно з частотою нерозпізнавання 0,5%)
- ▶ Довжини тринуклеотидних повторів $\leq 5x$ (усі основи, що розпізнавались у 819 секвенованих ампліконах із 5-кратним (5x) тринуклеотидним повтором, розпізнано правильно з частотою нерозпізнавання 0,5%)
- ▶ Інсерції довжиною ≤ 24 нуклеотидів (66 343 з 66 370 основ, які розпізнавались у 819 секвенованих ампліконах із 24-нуклеотидною інсерцією, розпізнано правильно з частотою нерозпізнавання 1,2%, водночас неправильні розпізнавання не відбувались у регіоні з 24-нуклеотидною інсерцією)
- ▶ Делеції довжиною ≤ 25 нуклеотидів (усі основи, що розпізнавались у 2457 секвенованих ампліконах із 25-нуклеотидною делецією, розпізнано правильно без нерозпізнавань)

Соматичний

Описане далі дослідження проводилося для оцінювання точності розпізнавання варіантів модуля Somatic Variant Module приладу NextSeq 550Dx із використанням набору реагентів NextSeq 550Dx High Output Reagent Kit v2 (300 cycles).

У цьому дослідженні застосовано репрезентативний аналіз для пошуку низки генів, які покривають 12 588 основ (150 ампліконів) для 23 різних хромосом. ДНК для зразків Platinum Genome було екстраговано з блоків, оброблених за методом FFPE, з метою отримання шести унікальних зразків із подальшим оцінюванням у дослідженні.

Зразок ДНК лінії GM12877 було розведено зразком ДНК лінії GM12878 для отримання ліній GM12877-D5 і GM12877-D7 як набору унікальних гетерозиготних варіантів із частотою варіантів приблизно 5% і 7% відповідно. Зразок ДНК лінії GM12878 було розведено зразком ДНК лінії GM12877, щоб отримати лінії GM12878-D5 і GM12878-D7. Кожний зі зразків протестовано в трьох повторностях, окрім розведених зразків, які було протестовано в шести повторностях. Загалом було виконано дев'ять прогонів із використанням трьох приладів секвенування, трьох серій реагентів і трьох операторів протягом п'яти днів запуску.

Точність було встановлено для SNV, інсерцій і делецій за допомогою порівняння отриманих результатів із

показниками, установленими за допомогою точно охарактеризованого складеного еталонного методу — Platinum Genomes версії 2016-1.0. Довірчі геномні області було визначено на основі цього еталонного методу, якщо не зазначено інше.

Таблиця 13 Зведені дані узгодженості для соматичних варіантів

Критерії	Загальні спостереження ¹	Результат за спостереженням ²	Результат за прогоном ³
PPA для SNV	378	98,9	99,9
PPA для інсерцій	378	96,9	99,9
PPA для делецій	378	97,1	99,9
NPA	378	> 99,9	> 99,9
OPA	378	> 99,9	> 99,9

¹Обчислено як кількість зразків на прогін (42) × кількість прогонів (9) = 378.

²Найменше отримане значення за повторністю зразка в усіх 9 прогонах.

³Найменше значення, коли дані з кожного прогону аналізуються сукупно.

У Таблиця 14 містяться дані, представлені з відсотками збігу позитивних і негативних результатів залежно від зразка, де результати варіантів порівняно з показниками, встановленими за допомогою точно охарактеризованого складеного еталонного методу для обчислення PPA. Поєднано три типи варіантів (SNV, інсерції та делеції). Еталонний метод дає результати лише для одонуклеотидних варіантів та інсерцій/делецій, тому для обчислення NPA результати розпізнавання основ, які не належать до варіантів, порівнюються зі збіркою 19 референсної послідовності геному людини (hg19).

Таблиця 14 Узгодженість соматичних варіантів за зразками

Зразок	Середня частота розпізнавань	Прогнозовано	ІП	ХН	Варіанти без розпізнавань	ІН	ХП	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

У Таблиця 15 містяться дані, представлені залежно від зразка, де результати варіантів порівняно з показниками, установленими за допомогою точно охарактеризованого складеного еталонного методу. Виявлення оцінюється для кожного типу варіанта — SNV, інсерції й делеції окремо. Референсні позиції виключено.

Таблиця 15 Узгодженість соматичних варіантів для зразків за типом варіанта

Зразок	SNV			Інсерції			Делеції		
	Прогнозовано	ІП	ХН	Прогнозовано	ІП	ХН	Прогнозовано	ІП	ХН
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

Десять зразків було додатково проаналізовано для розпізнавання незначних інсерцій і делецій (інделів) (Таблиця 16). Загалом було виявлено 71 індел розміром 1–24 п. о. для інсерцій і 1–25 п. о. для делецій.

Таблиця 16 Зведені дані виявлення інделів для соматичного варіанта

Тип варіанта	Прогнозовані варіанти	ІП	ХН	Варіанти без розпізнавань	PPA
Інсерція	10773	10282	9	482	99,2
Делеція	11502	10667	5	830	> 99,9

150 ампліконів було розроблено з метою покриття різноманітного геномного матеріалу. Уміст GC в ампліконах варіювався в межах 0,19–0,87 %. Також амплікони мали низку одонуклеотидних (наприклад, полі-А, полі-Т), динуклеотидних і тринуклеотидних повторів. Дані було скопійовано залежно від амплікона (Таблиця 17) для визначення впливу геномного вмісту на відсоток правильних розпізнавань. Відсоток правильних розпізнавань складається з варіантних і референсних розпізнавань, він менше 100 %, якщо наявні неправильні розпізнавання або нерозпізнавання.

Таблиця 17 Точність на рівні ампліконів для соматичних варіантів

Амплікон	Хромо-сома	Початок амплікона	Кінець амплікона	Розмір проаналізованого фрагмента	Основи в довірчих зонах	Геномний уміст амплікона	Вміст GC	Правильні розпізнавання	Неправильні розпізнавання	Нерозпізнавання	Відсоток правильних розпізнавань
1	1	36450499	36450591	93	93	Індел	0,22	35066	0	88	99,7
2	1	109465122	109465200	79	79	Полі-А (5), полі-С (5), індел	0,38	29827	0	35	99,9
3	1	218353867	218353957	91	91	Індел	0,4	34202	0	283	99,2
4	1	223906657	223906748	92	92	Індел	0,49	34613	0	163	99,5
5	1	228526602	228526682	81	81	Полі-G (5)	0,69	30571	0	47	99,8
6	1	236372039	236372108	70	70	Полі-Т (10), індел	0,39	26452	0	8	100,0
7	1	247812041	247812128	88	88	Полі-А (5), СТ (3), ТАА (3), індел	0,27	33148	0	116	99,7
8	2	55862774	55862863	90	90	Індел	0,28	33928	0	92	99,7
9	2	87003930	87004009	80	80	Індел	0,38	30218	0	22	99,9
10	2	177016721	177016805	85	81	Н/Д	0,65	30616	0	2	> 99,9
11	2	186625727	186625801	75	75	Полі-А (8)	0,35	28017	0	499	98,3
12	2	190323504	190323591	88	88	Полі-Т (5)	0,42	33207	0	57	99,8
13	2	200796740	200796826	87	87	Полі-Т (5), індел	0,31	32524	9	718	97,8
14	2	212245049	212245139	91	91	Полі-Т (5), полі-А (6), індел	0,3	33972	0	456	98,7
15	2	228147052	228147144	93	93	Н/Д	0,43	35051	0	103	99,7

Амплікон	Хромо-сома	Початок амплікона	Кінець амплікона	Розмір проаналізованого фрагмента	Основи в довірчих зонах	Геномний уміст амплікона	Вміст GC	Правильні розпізнавання	Неправильні розпізнавання	Нерозпізнавання	Відсоток правильних розпізнавань
16	2	235016350	235016422	73	73	Полі-Т (5), індел	0,42	27459	0	136	99,5
17	3	4466229	4466321	93	93	АТ (3), індел	0,27	34534	0	620	98,2
18	3	46620561	46620643	83	83	Н/Д	0,43	31339	0	44	99,9
19	3	49851331	49851400	70	70	СТ (3), індел	0,49	26373	0	87	99,7
20	3	189713161	189713248	88	88	Полі-А (5), полі-Т (5), полі-А (9), TG (3)	0,41	32829	0	857	97,5
21	3	190106030	190106104	75	74	Індел	0,57	27925	0	47	99,8
22	4	2233667	2233744	78	78	Полі-А (6)	0,26	29327	4	162	99,4
23	4	7780541	7780637	97	97	Полі-Г (6), полі-Т (5), полі-А (5)	0,42	36585	0	117	99,7
24	4	15688604	15688681	78	78	Н/Д	0,29	29427	0	57	99,8
25	4	56236521	56236586	66	62	Полі-А (5), індел	0,36	23356	5	75	99,7
26	4	102839244	102839314	71	69	Полі-А (5)	0,46	25942	0	140	99,5
27	4	164446743	164446804	62	62	Полі-А (7), індел	0,27	22944	0	560	97,6
28	5	1882081	1882158	78	75	Н/Д	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	Н/Д	0,39	24120	0	72	99,7
31	5	74077114	74077196	83	83	Полі-А (6), індел	0,3	31297	0	77	99,8
32	5	147475343	147475409	67	67	Полі-Т (5)	0,37	25277	0	55	99,8
33	5	149323731	149323821	91	91	СТ (4), AG (3)	0,55	34308	0	90	99,7
34	5	155662213	155662287	75	75	Індел	0,43	28266	0	163	99,4

Амплікон	Хромо-сома	Початок амплікона	Кінець амплікона	Розмір проаналізованого фрагмента	Основи в довірчих зонах	Геномний уміст амплікона	Вміст GC	Правильні розпізнавання	Неправильні розпізнавання	Нерозпізнавання	Відсоток правильних розпізнавань
35	6	6318713	6318814	102	102	Полі-G (6)	0,68	38489	0	67	99,8
36	6	24949983	24950074	92	92	Індел	0,63	34730	0	46	99,9
37	6	31084900	31084999	100	94	ГСТ (5), індел	0,61	35057	0	483	98,6
38	6	32147987	32148084	98	98	Полі-Т (5), ТСТ (3), СТТ (3)	0,55	36647	0	406	98,9
39	6	32986864	32986958	95	95	Індел	0,53	35681	0	238	99,3
40	6	33408498	33408583	86	86	Полі-С (6)	0,7	32438	0	70	99,8
41	6	41647401	41647495	95	94	Полі-G (5), індел	0,61	35441	0	91	99,7
42	6	112435865	112435955	91	91	Полі-А (5)	0,44	34354	0	44	99,9
43	7	22202076	22202148	73	73	Н/Д	0,44	27575	0	28	99,9
44	7	66276100	66276187	88	88	Індел	0,35	33060	0	213	99,4
45	7	77365735	77365821	87	87	Полі-А (7), АГ (4)	0,26	32423	0	489	98,5
46	7	110939946	110940030	85	85	Індел	0,38	32074	0	56	99,8
47	7	128533468	128533557	90	90	Полі-G (5), індел	0,62	33791	0	281	99,2
48	7	149503875	149503965	91	91	Полі-G (6), полі-С (6), індел	0,71	34316	0	82	99,8
49	7	154404519	154404599	81	66	Н/Д	0,31	24901	0	47	99,8
50	7	156476507	156476599	93	93	Індел	0,35	35067	0	87	99,8
51	8	1817312	1817394	83	83	Н/Д	0,42	31365	0	9	> 99,9
52	8	24811020	24811109	90	89	Полі-G (7), СТС (4), індел	0,61	32781	0	890	97,4
53	8	76518625	76518691	67	67	Індел	0,3	25228	0	146	99,4

Амплікон	Хромо-сома	Початок амплікона	Кінець амплікона	Розмір проаналізованого фрагмента	Основи в довірчих зонах	Геномний уміст амплікона	Вміст GC	Правильні розпізнавання	Неправильні розпізнавання	Нерозпізнавання	Відсоток правильних розпізнавань
54	9	103054909	103055006	98	98	Полі-G (6)	0,67	36968	0	76	99,8
55	9	105586150	105586214	65	65	Індел	0,32	24472	0	100	99,6
56	9	107620823	107620918	96	96	Н/Д	0,49	36203	0	85	99,8
57	9	123769149	123769231	83	83	АТ (3)	0,37	31329	0	45	99,9
58	9	138995345	138995441	97	97	Полі-С (6), індел	0,68	36472	0	201	99,5
59	10	5987120	5987198	79	78	Полі-G (5), індел	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	Полі-Т (5)	0,3	29843	0	19	99,9
62	10	33018351	33018440	90	90	Полі-А (5), полі-Т (5)	0,2	33968	0	68	99,8
63	10	45084159	45084253	95	95	Індел	0,35	35829	0	81	99,8
64	10	55892599	55892687	89	88	АС (11), індел	0,42	32098	88	2048	93,8
65	10	101611250	101611329	80	80	Н/Д	0,49	30217	0	28	99,9
66	10	118351373	118351453	81	81	Н/Д	0,51	30531	0	96	99,7
67	11	8159816	8159912	97	96	Н/Д	0,45	36105	0	192	99,5
68	11	30177648	30177717	70	70	Індел	0,46	26318	0	153	99,4
69	11	47470345	47470444	100	100	Н/Д	0,65	37785	0	24	99,9
70	11	59837679	59837740	62	62	Індел	0,37	23368	0	68	99,7
71	11	64418856	64418957	102	102	Н/Д	0,59	38546	0	10	> 99,9
72	11	93529612	93529684	73	73	Полі-А (5)	0,4	27516	0	78	99,7
73	11	101347052	101347136	85	85	Н/Д	0,42	32083	0	48	99,9
74	11	102477336	102477426	91	91	Полі-G (6)	0,55	34047	0	369	98,9
75	11	118406285	118406369	85	85	Індел	0,53	32065	0	74	99,8
76	11	120357801	120357885	85	85	Полі-А (5), СА (3), індел	0,34	32083	0	47	99,9

Амплікон	Хромо-сома	Початок амплікона	Кінець амплікона	Розмір проаналізованого фрагмента	Основи в довірчих зонах	Геномний уміст амплікона	Вміст GC	Правильні розпізнавання	Неправильні розпізнавання	Нерозпізнавання	Відсоток правильних розпізнавань
77	11	125769313	125769397	85	85	GA (3)	0,52	32103	0	27	99,9
78	12	2834770	2834853	84	84	Полі-С (5), індел	0,52	31645	16	525	98,3
79	12	26811004	26811096	93	93	Полі-А (7), АС (4)	0,33	34824	0	330	99,1
80	12	30881766	30881846	81	81	Н/Д	0,49	30497	0	121	99,6
81	12	88474105	88474175	71	71	Полі-А (6)	0,35	26773	0	65	99,8
82	12	120966872	120966966	95	95	Полі-G (5)	0,68	35830	9	72	99,8
83	13	24167504	24167576	73	73	Н/Д	0,52	27498	0	114	99,6
84	13	25816961	25817049	89	88	Полі-А (5), полі-Т (7), полі-А (7), індел	0,22	32824	0	566	98,3
85	13	44880112	44880200	89	89	Індел	0,49	33574	0	77	99,8
86	13	77665218	77665294	77	77	Індел	0,39	29075	0	31	99,9
87	14	31619327	31619393	67	67	GA (3), ТА (3)	0,39	25313	0	13	99,9
88	14	39517884	39517966	83	83	Н/Д	0,25	31360	0	22	99,9
89	14	46958962	46959034	73	72	Полі-Т (5), індел	0,19	26499	0	717	97,4
90	14	58050030	58050110	81	81	Індел	0,38	30494	0	133	99,6
91	14	82390559	82390649	91	91	Індел	0,35	34313	0	86	99,7
92	14	92549544	92549609	66	66	Полі-А (5)	0,41	24555	0	1527	94,1
93	14	102808496	102808589	94	94	Індел	0,62	35472	0	69	99,8
94	15	43170751	43170848	98	96	Полі-С (5)	0,45	36264	0	24	99,9
95	15	63446149	63446216	68	68	Індел	0,25	25667	0	37	99,9
96	15	77879807	77879901	95	93	Полі-G (5), індел	0,68	34745	0	432	98,8
97	15	81625334	81625428	95	95	Полі-Т (6)	0,43	35870	0	40	99,9

Амплікон	Хромо-сома	Початок амплікона	Кінець амплікона	Розмір проаналізованого фрагмента	Основи в довірчих зонах	Геномний уміст амплікона	Вміст GC	Правильні розпізнавання	Неправильні розпізнавання	Нерозпізнавання	Відсоток правильних розпізнавань
98	15	85438263	85438334	72	71	Індел	0,65	26762	0	76	99,7
99	15	89817413	89817503	91	91	Н/Д	0,36	34286	0	112	99,7
100	15	89864274	89864343	70	70	Індел	0,56	26449	0	11	> 99,9
101	16	1894910	1894972	63	63	Н/Д	0,27	23809	0	5	> 99,9
102	16	28997904	28997998	95	95	Полі-С (5)	0,67	35860	0	50	99,9
103	16	53682908	53682994	87	87	ТА (3)	0,41	32835	0	60	99,8
104	16	57954406	57954509	104	104	Полі-С (5)	0,67	39177	0	144	99,6
105	16	85706375	85706465	91	91	Полі-Т (5), індел	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	Полі-С (5), індел	0,67	32752	0	134	99,6
108	17	3970090	3970180	91	91	Індел	0,46	34343	0	82	99,8
109	17	16084945	16085037	93	93	Індел	0,26	35077	0	78	99,8
110	17	33998759	33998849	91	89	Полі-Т (5)	0,54	33553	0	89	99,7
111	17	39589691	39589774	84	82	Полі-А (13), індел (x2)	0,29	30554	53	2296	92,9
112	17	41244394	41244484	91	91	Полі-А (5)	0,34	34360	0	38	99,9
113	17	45438866	45438957	92	92	Полі-А (7), АТ (3), АТ (4), АТ (4), індел	0,26	34367	0	418	98,8
114	17	61502432	61502510	79	79	Індел	0,41	29751	0	119	99,6
115	17	64023582	64023667	86	86	Полі-Т (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	Н/Д	0,37	34386	0	12	> 99,9
119	18	9888026	9888094	69	69	Полі-А (6), TG (3)	0,43	25692	0	399	98,5

Амплікон	Хромо-сома	Початок амплікона	Кінець амплікона	Розмір проаналізованого фрагмента	Основи в довірчих зонах	Геномний уміст амплікона	Вміст GC	Правильні розпізнавання	Неправильні розпізнавання	Нерозпізнавання	Відсоток правильних розпізнавань
120	18	38836999	38837073	75	75	Полі-А (5), індел	0,37	27923	0	893	96,9
121	18	47405382	47405462	81	81	СТС (3), індел	0,47	30598	0	20	99,9
122	18	54815665	54815749	85	85	СТ (3), індел	0,45	31969	0	161	99,5
123	18	59773996	59774060	65	65	Н/Д	0,48	24531	0	48	99,8
124	19	625143	625241	99	99	Н/Д	0,59	37298	0	124	99,7
125	19	18121418	18121491	74	74	Н/Д	0,68	27881	0	109	99,6
126	19	18186574	18186643	70	70	Н/Д	0,64	26442	0	26	99,9
127	20	746056	746149	94	94	Н/Д	0,61	35501	0	31	99,9
128	20	10633195	10633276	82	82	АС (3)	0,59	30951	0	72	99,8
129	20	17705633	17705708	76	76	СТ (3)	0,58	28686	0	42	99,9
130	20	21766821	21766890	70	70	GT (3), TG (4), індел	0,46	26372	0	88	99,7
131	20	25278421	25278521	101	101	Індел	0,63	38159	0	20	99,9
132	20	50897302	50897368	67	67	Індел	0,36	25188	0	544	97,9
133	20	62331904	62331994	91	88	Полі-G (6)	0,73	32969	0	309	99,1
134	20	62690860	62690946	87	87	Індел	0,57	32818	0	77	99,8
135	21	30300823	30300888	66	66	Індел	0,35	24758	9	181	99,2
136	21	33694176	33694273	98	98	Полі-Т (6), СА (3)	0,54	36902	0	160	99,6
137	21	36710706	36710792	87	87	GT (3), індел	0,39	32841	0	48	99,9
138	21	46644924	46644992	69	69	Полі-А (6), AG (3), індел	0,32	25939	0	280	98,9
139	21	46705575	46705664	90	90	Полі-Т (5), полі-А (6)	0,5	33942	0	78	99,8
140	22	25750774	25750873	100	100	Індел	0,63	37733	0	86	99,8
141	22	32439233	32439329	97	97	Н/Д	0,68	36617	0	49	99,9

Амплікон	Хромо-сома	Початок амплікона	Кінець амплікона	Розмір проаналізованого фрагмента	Основи в довірчих зонах	Геномний уміст амплікона	Вміст GC	Правильні розпізнавання	Неправильні розпізнавання	Нерозпізнавання	Відсоток правильних розпізнавань
142	22	37409844	37409940	97	97	Індел	0,46	36525	0	162	99,6
143	22	37637596	37637694	99	99	Н/Д	0,6	37398	0	24	99,9
144	22	47081347	47081438	92	92	Індел	0,66	34754	0	22	99,9
145	X	15870424	15870492	69	69	Полі-Т (5)	0,26	26046	0	36	99,9
146	X	135288543	135288611	69	69	Полі-С (5)	0,62	26019	0	63	99,8
147	X	135290777	135290847	71	71	Н/Д	0,52	26780	0	58	99,8
148	Y	2655397	2655461	65	0	Н/Д	0,55	0	0	0	Н/Д
149	Y	2655519	2655609	91	0	Н/Д	0,48	0	0	0	Н/Д
150	Y	2655609	2655679	71	0	Полі-А (5)	0,37	0	0	0	Н/Д

Результати секвенування для зразків лінії GM12878 було порівняно з генотипом із високим ступенем вірогідності для лінії NA12878, затвердженим Національними інститутами стандартів і технологій (National Institutes of Standards and Technology, NIST) (версія 2.19). 92 зі 150 ампліконів було виявлено повністю в межах геномних областей із високим ступенем вірогідності, 41 амплікон мав часткове перекриття, і 17 ампліконів не мали перекриття в послідовності, затвердженій NIST. Цей результат призвів до появи 10 000 координат на повторність для порівняння. Результати розпізнавання основ, які не належали до варіантів, було порівняно зі збіркою референсної послідовності геному людини hg19. Результати дослідження точності наведено в [Таблиця 18](#).

Таблиця 18 Узгодженість зразків лінії GM12878 із базою даних NIST для соматичних варіантів

Зразок	Кількість ампліконів	Середня частота розпізнавань	ІП	ХН	ІН	ХП	РРА	NPA	OPA
GM12878	133	98,8	2808	0	258488	0	100	100	100

На підставі даних, отриманих у цьому дослідженні соматичного матеріалу протягом дев'яти прогонів, прилад NextSeq 550Dx може секвенувати з отриманням стабільних результатів наведені далі послідовності.

- ▶ Зі вмістом GC $\geq 19\%$ (усі основи, що розпізнавались у 378 секвенованих ампліконах зі вмістом GC 19%, розпізнано правильно з частотою нерозпізнавання 2,6%)
- ▶ Зі вмістом GC $\leq 87\%$ (усі основи, що розпізнавались у 378 секвенованих ампліконах зі вмістом GC 87%, розпізнано правильно з частотою нерозпізнавання 0,6%)
- ▶ Фрагменти полі-А довжиною ≤ 9 (усі основи, що розпізнавались у 378 секвенованих ампліконах із повтором полі-А з дев'яти нуклеотидів, розпізнано правильно з частотою нерозпізнавання 2,5%)
- ▶ Фрагменти полі-Т довжиною ≤ 10 (усі основи, що розпізнавались у 378 секвенованих ампліконах із повтором полі-Т з десяти нуклеотидів, розпізнано правильно з частотою нерозпізнавання менше 0,1%)
- ▶ Фрагменти полі-Г довжиною ≤ 6 (усі основи, що розпізнавались у 2268 секвенованих ампліконах із повтором полі-Г із шести нуклеотидів, розпізнано правильно з частотою нерозпізнавання 0,5%)
- ▶ Фрагменти полі-С довжиною ≤ 6 (усі основи, що розпізнавались у 756 секвенованих ампліконах із повтором полі-С із шести нуклеотидів, розпізнано правильно з частотою нерозпізнавання 0,4%)
- ▶ Довжини динуклеотидних повторів $\leq 4x$ (усі основи, що розпізнавались у 1890 секвенованих ампліконах із 4-кратним (4x) динуклеотидним повтором, розпізнано правильно з частотою нерозпізнавання 0,9%)
- ▶ Довжини тринуклеотидних повторів $\leq 5x$ (усі основи, що розпізнавались у 378 секвенованих ампліконах із 5-кратним (5x) тринуклеотидним повтором, розпізнано правильно з частотою нерозпізнавання 1,4%)
- ▶ Інсерції довжиною ≤ 23 (усі основи, що розпізнавались у 378 секвенованих ампліконах із 23-нуклеотидною інсерцією, розпізнано правильно з частотою нерозпізнавання 0,8%)
- ▶ Делеції довжиною ≤ 25 (усі основи, що розпізнавались у 1134 секвенованих ампліконах із 25-нуклеотидною делецією, розпізнано правильно з частотою нерозпізнавання 0,7%)

Прецизійність

Прецизійність приладу NextSeq 550Dx було встановлено після тестування 13 унікальних зразків Platinum Genome за допомогою трьох приладів, трьох серій реагентів і трьох операторів для виконання дев'яти прогонів секвенування протягом п'яти днів запуску. Репрезентативний аналіз, зразки й еталонний метод були аналогічними описаним у дослідженні точності для варіантів зародкової лінії. Дані щодо прецизійності було встановлено за допомогою аналізу компонентів дисперсії, у якому VAF використано як залежну змінну (відгук) і в якому стандартні відхилення обчислюються на рівні компонента для приладу, серії реагентів, оператора та дня запуску ([Таблиця 19](#)). Загальна кількість спостережень, використаних у аналізі для кожного компонента варіабельності приладу, оператора або серії реагентів, становила 699, 176 і 235 для SNV, інсерцій і делецій відповідно.

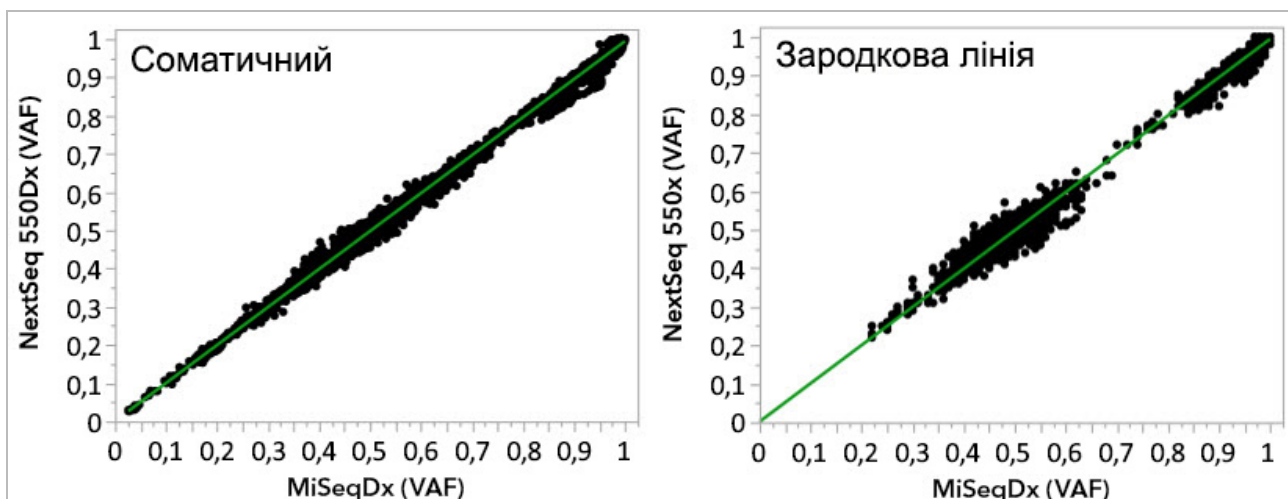
Таблиця 19 Результати дослідження прецизійності для приладу NextSeq 550Dx (стандартне відхилення)

Компонент	Тип варіанта	СВ компонента		Загальне СВ	
		Макс.	Медіана	Макс.	Медіана
Серія	SNV	0,0076	0,0002	0,0833	0,0154
	Інсерція	0,0104	0,0000	0,0410	0,0157
	Делеція	0,0046	0,0005	0,0560	0,0187
Прилад	SNV	0,0114	0,0003	0,0840	0,0153
	Інсерція	0,0138	0,0009	0,0407	0,0161
	Делеція	0,0079	0,0008	0,0549	0,0187
Оператор	SNV	0,0226	0,0008	0,0841	0,0155
	Інсерція	0,0344	0,0010	0,0417	0,0164
	Делеція	0,0083	0,0013	0,0547	0,0187
День	SNV	0,0277	0,0012	0,0825	0,0160
	Інсерція	0,0235	0,0012	0,0409	0,0169
	Делеція	0,0271	0,0014	0,0548	0,0188

Порівняння методів (залежно від платформи секвенування)

Зразки, отримані з цільної крові, і зразки FFPE було оцінено на приладах NextSeq 550Dx і MiSeqDx із використанням робочих процесів для варіантів зародкової лінії та соматичних варіантів за допомогою спеціального набору ампліконів TruSeq Custom Amplicon Kit Dx. Узгодженість частоти варіантів для зразків, отриманих із крові, і зразків FFPE було оцінено за допомогою декількох репрезентативних аналізів. На [Рисунок 2](#) показано графік кореляції VAF між двома приладами для одного репрезентативного аналізу, а в [Таблиця 20](#) підсумовано цю кореляцію за панеллю аналізу. На підставі сильної кореляції між приладами MiSeqDx і NextSeq 550Dx було встановлено, що технічні характеристики, пов'язані з преаналітичними факторами (наприклад, методами екстракції або інтерферувальними речовинами), є застосовними до обох приладів. Додаткову інформацію див. в інструкції з використання спеціального набору ампліконів TruSeq Custom Amplicon Kit Dx.

Рисунок 2 Кореляція значень, отриманих на приладах MiSeqDx і NextSeq 550Dx, для зразків FFPE (ліворуч) і зразків, отриманих із крові (праворуч), з використанням аналізу 1



Таблиця 20 Результати порівняння методів із використанням унікальних зразків, отриманих із крові, і зразків FFPE

Джерело геномної ДНК	Аналіз (панель олігонуклеотидів)	Біологічні повторності (зразки)	Технічні повторності (за зразком)	Спостереження (к-сть варіантів)	Нахил	Відрізок, який відсікається	Коефіцієнт кореляції (R ²)
Кров	Аналіз 1	45	2	8369 ¹	0,992	0,002	0,995 ²
Кров	Аналіз 2	45	2	5457	0,995	0,005	0,981
FFPE	Аналіз 1	46	2	8319	0,993	0,000	0,997 ²
FFPE	Аналіз 3	40	1	280	0,969	0,015	0,978

¹Дві точки даних видалено на підставі вказаного обмеження для модуля Germline Variant Module.

²Коефіцієнт детермінації для графіків VAF, як показано на рисунку 2.

Відтворюваність

Відтворюваність приладу NextSeq 550Dx було оцінено за допомогою зразків Platinum Genome із репрезентативним аналізом для пошуку низки генів, які покривають 12 588 основ для 23 різних хромосом із використанням 150 ампліконів. Тестування зародкової лінії складалося з семи повторностей 13 зразків, а тестування соматичних варіантів — з шести повторностей семи зразків із різними рівнями VAF. Зразки було підготовлено з використанням спеціального набору ампліконів TruSeq Custom Amplicon Kit Dx.

Тестування було проведено в трьох зовнішніх центрах із використанням однієї серії набору реагентів NextSeq 550Dx High Output Reagent Kit v2 (300 cycles) У кожному центрі використовувався один прилад NextSeq 550Dx. У кожному центрі тестування проводили два оператори. Кожний оператор проводив тестування протягом трьох неспідовних днів запуску для кожного типу зразка в загальній кількості 36 прогонів для трьох центрів. У результаті цього тестування було виконано по 18 прогонів робочого процесу як для генів зародкової лінії, так і для соматичних генів.

Зародкова лінія

Зародкові варіанти з рівнем VAF $\geq 0,2$ позначено як позитивні (варіант). Для прогнозованих позитивних варіантів зародкової лінії дані було оцінено з точки зору частоти нерозпізнання й частоти правильних позитивних розпізнавань у межах типу варіанта (SNV, інсерція, делеція). У [Таблиця 21](#) для кожного типу варіанта підбито підсумок отриманих частот, а також нижньої й верхньої межі рівня вірогідності 95 % (НРВ/ВРВ), обчислених за методом оцінки Вілсона.

Таблиця 21 Дані розпізнання варіанта зародкової лінії для прогнозованих позитивних результатів за типом варіанта

Тип варіанта	Нерозпізнання			Правильне позитивне розпізнання				
	Спостереження	Загалом	Відсоток	Спостереження	Загалом	Відсоток	95 % НРВ	95 % ВРВ
SNV	16	110 376	0,014	110 349	110 360	99,99	99,98	99,99
Інсерції	1026	37 044	2,77	36 018	36 018	100	99,99	100,00
Делеції	648	34 776	1,86	34 128	34 128	100	99,99	100,00

Зародкові варіанти з рівнем VAF $\geq 0,2$ позначено як негативні (дикий тип). Для прогнозованих розташувань із негативними результатами щодо варіантів зародкової лінії дані було оцінено з точки зору частоти нерозпізнання й частоти правильного розпізнання дикого типу. У [Таблиця 22](#) підбито підсумок отриманих частот, а також нижньої й верхньої межі рівня вірогідності 95 % (НРВ/ВРВ), обчислених за методом оцінки Вілсона.

Таблиця 22 Дані розпізнавання варіанта зародкової лінії для прогнозованих негативних результатів

Тип варіанта	Нерозпізнавання			Правильне негативне розпізнавання				
	Спостереження	Загалом	Відсоток	Спостереження	Загалом	Відсоток	95 % НРВ	95 % ВРВ
Дикий тип	4883	19 600 182	0,025	19 595 299	19 595 299	100	100,00	100,00

Варіанти зародкової лінії з рівнем VAF $\geq 0,2$ і $< 0,7$ називаються позитивними гетерозиготними для варіанта, а варіанти з рівнем VAF $\geq 0,7$ — позитивними гомозиготними. Зразки зародкової лінії з гетерозиготними варіантами використано для визначення, чи вплине властива аналізу варіабельність на розпізнавання генотипу. Сх було визначено для обох граничних значень (0,2 для гетерозиготних і 0,7 для гомозиготних генотипів), де x — це частка повторних тестів, які перевищують граничне значення. Для нижнього граничного значення VAF 0,2 Сх становив $> 99,999\%$, вказуючи на те, що $> 99,999\%$ гетерозиготних варіантів будуть зчитуватися як гетерозиготні. Що стосується верхнього граничного значення VAF 0,7, Сх становив $\leq 0,001\%$, вказуючи на те, що $\leq 0,001\%$ гетерозиготних варіантів будуть зчитуватися як гомозиготні. У [Таблиця 23](#) підсумовано результати за типом варіанта.

Варіанти зародкової лінії з рівнем VAF $\geq 0,2$ і $< 0,7$ називаються позитивними гетерозиготними для варіанта, а варіанти з рівнем VAF $\geq 0,7$ — позитивними гомозиготними. Зразки зародкової лінії з гетерозиготними варіантами використано для визначення, чи вплине властива аналізу варіабельність на розпізнавання генотипу. Сх було визначено для обох граничних значень (0,2 для гетерозиготних і 0,7 для гомозиготних генотипів), де x — це частка повторних тестів, які перевищують граничне значення. Що стосується нижнього граничного значення VAF 0,2, Сх становив $\geq 99,999\%$, вказуючи на те, що $\geq 99,999\%$ гетерозиготних варіантів будуть зчитуватися як гетерозиготні. Для верхнього граничного значення VAF 0,7 Сх становив $\leq 0,001\%$, вказуючи на те, що $\leq 0,001\%$ гетерозиготних варіантів будуть зчитуватися як гомозиготні. У [Таблиця 23](#) підсумовано результати за типом варіанта.

Таблиця 23 Значення Сх для гетерозиготних варіантів зародкової лінії

Тип варіанта	Граничне значення VAF 0,2	Граничне значення VAF 0,7
	$\geq 99,999\%$	$\leq 0,001\%$
SNV	94/94	94/94
Інсерції	24/24	24/24
Делеції	35/35	35/35
Загалом	153	153

Соматичний

Соматичні варіанти з рівнями VAF $\geq 0,026$ позначено як позитивні (варіант). Дані з рівнями VAF $\geq 0,01$ і $< 0,026$ вважалися неоднозначними з метою цього аналізу (ані позитивними, ані негативними, позначеними як ті, що мають низьку частоту варіантів). Для оцінки характеристик результату було обчислено трьома способами:

- ▶ Найкращий випадок: будь-який неоднозначний результат вважався правильним позитивним розпізнаванням (узгодженість із прогнозованими результатами).
- ▶ Найгірший випадок: будь-який неоднозначний результат вважався неправильним розпізнаванням (узгодженість із прогнозованими результатами).
- ▶ Випадок виключення: будь-який неоднозначний результат було виключено з аналізу.

У трьох таблицях — [Таблиця 24](#), [Таблиця 25](#) і [Таблиця 26](#) — підбито підсумок результатів розпізнавання для найкращого випадку, найгіршого випадку й випадку виключення відповідно, а також нижньої й верхньої межі рівня вірогідності 95 % (НРВ/ВРВ), обчислених за методом оцінки Вілсона.

Таблиця 24 Дані розпізнавання соматичних варіантів для прогнозованих позитивних результатів за типом варіанта (найкращий випадок)

Тип варіанта	Правильне позитивне розпізнавання				
	Спостереження	Загалом	Відсоток	95 % НРВ	95 % ВРВ
SNV	54 346	54 346	100	99,99	100,00
Інсерції	18 036	18 036	100	99,98	100,00
Делеції	18 381	18 381	100	99,98	100,00

Таблиця 25 Дані розпізнавання соматичних варіантів для прогнозованих позитивних результатів за типом варіанта (найгірший випадок)

Тип варіанта	Правильне позитивне розпізнавання				
	Спостереження	Загалом	Відсоток	95 % НРВ	95 % ВРВ
SNV	54 346	54 346	100	99,99	100,00
Інсерції	18 000	18 036	99,8	99,72	99,86
Делеції	18 381	18 381	100	99,98	100,00

Таблиця 26 Дані розпізнавання соматичних варіантів для прогнозованих позитивних результатів за типом варіанта (випадок виключення)

Тип варіанта	Правильне позитивне розпізнавання				
	Спостереження	Загалом	Відсоток	95 % НРВ	95 % ВРВ
SNV	54 346	54 346	100	99,99	100,00
Інсерції	18 000	18 000	100	99,98	100,00
Делеції	18 381	18 381	100	99,98	100,00

Соматичні варіанти з рівнем VAF < 0,01 позначено як негативні розпізнавання (дикий тип). Для прогнозованих розташувань із негативними результатами щодо соматичних варіантів дані було оцінено з точки зору частоти нерозпізнавання й частоти правильного розпізнавання дикого типу. Правильні розпізнавання дикого типу було визначено через виключення нерозпізнавань і віднімання від загального числа отриманих розпізнавань, які підпадають під визначення неоднозначних (рівні VAF $\geq 0,01$ і < 0,026), а також неправильних розпізнавань, які перевищили граничне значення (рівні VAF $\geq 0,026$). У [Таблиця 27](#) підбито підсумок отриманих, загальних і відсоткових результатів для розташувань із негативними результатами щодо соматичних варіантів для частоти нерозпізнавання й частоти правильного розпізнавання дикого типу, а також нижньої й верхньої межі рівня вірогідності 95 % (НРВ/ВРВ), обчислених за методом оцінки Вілсона.

Таблиця 27 Дані розпізнавання соматичного варіанта для прогнозованих негативних результатів

Тип варіанта	Нерозпізнавання			Правильне розпізнавання						
	Спостереження	Загалом	Відсоток	Неоднозначно	Неправильно	Правильно	Загалом	Відсоток	95 % НРВ	95 % ВРВ
Дикий тип	36 326	8 909 676	0,408	2254	121	8 870 975	8 873 350	99,97	99,972	99,974

Соматичні зразки з різними рівнями VAF для одного варіанта було оцінено для визначення С95 аналізу (у межах кожного типу варіанта). Для оцінювання граничного значення варіабельності аналізу використовували зразки, у яких прогнозовані рівні VAF становили від 0,02 до 0,07. С95 було визначено для кожного варіанта, а найвищий С95 для кожного типу варіанта наведено в [Таблиця 28](#).

Таблиця 28 Зведені дані про C95 для соматичних варіантів

Тип варіанта	N	C95
SNV	74	0,0613
Інсерція	24	0,0573
Делеція	33	0,0575

Технічні характеристики набору реагентів NextSeq 550Dx High Output Reagent Kit v2.5 (300 Cycles)

Огляд

З приладом NextSeq 550Dx використовуються два набори реагентів: набір реагентів NextSeq 550Dx High Output Reagent Kit v2 (300 cycles) і набір реагентів NextSeq 550Dx High Output Reagent Kit v2.5 (300 cycles). Щоб продемонструвати відповідність набору реагентів NextSeq 550Dx High Output Reagent Kit v2.5 (300 cycles) аналітичним вимогам, верифікованим і валідованим для набору реагентів NextSeq 550Dx High Output Reagent Kit v2 (300 cycles), було проведено дослідження з набором реагентів NextSeq 550Dx High Output Reagent Kit v2.5 (300 cycles). За допомогою спеціального набору ампліконів TruSeq Custom Amplicon Kit Dx було підготовлено дві бібліотеки: одна з використанням робочого процесу для варіантів зародкової лінії, інша — з використанням робочого процесу для соматичних варіантів. Бібліотеки, отримані під час кожного робочого процесу, було протестовано з трьома серіями набору реагентів NextSeq 550Dx High Output Reagent Kit v2.5 (300 cycles) на трьох приладах NextSeq 550Dx. Крім того, тестування для кожного робочого процесу охоплювало один прогін із набором реагентів NextSeq 550Dx High Output Reagent Kit v2 (300 cycles).

Аналітична чутливість (межа холостої проби (МХП) і межа виявлення (МВ))

Верифікація за допомогою набору реагентів NextSeq 550Dx High Output Reagent Kit v2 (300 cycles) продемонструвала, що прилад NextSeq 550Dx має змогу виявляти варіанти за VAF 0,05 із помилкою другого роду $\leq 0,05$ і що граничне значення VAF 0,026, яке застосовується в модулі Somatic Variant Module, так звана ефективна МХП, підтримує помилку першого роду на рівні $\leq 0,01$. На підставі цих заявлених значень очікується, що варіант із VAF 0,05 більше або дорівнює граничному значенню з VAF 0,026 у 95 % випадків і що VAF для позиції дикого типу становить менше 0,026 у 99 % випадків. Щоб переконатися, що ці заявлені значення є прийнятними для набору реагентів NextSeq 550Dx High Output Reagent Kit v2.5 (300 cycles), на приладі NextSeq 550Dx було проведено повторні вимірювання зі зразками дикого типу (зразками МХП) і зі зразками, які містять варіанти з VAF 0,05 (зразки МВ), за допомогою набору реагентів NextSeq 550Dx High Output Reagent Kit v2.5 (300 cycles). Потім частку розпізнавань для значень із VAF вище й нижче 0,026 було порівняно із заявленими значеннями для набору реагентів NextSeq 550Dx High Output Reagent Kit v2 (300 cycles).

Тестування включало два зразки МВ, кожен із яких містив унікальний набір варіантів із цільовим значенням VAF 0,05, і відповідні зразки МХП, які містили відповідні фрагменти дикого типу для цільових варіантів. Для підготовки бібліотеки зразки МХП і МВ було оброблено у восьми та семи повторностях відповідно з використанням спеціального набору ампліконів TruSeq Custom Amplicon Kit Dx. Від початку бібліотеки було секвеновано з використанням набору реагентів NextSeq 550Dx High Output Reagent Kit v2 (300 cycles), щоб ідентифікувати варіанти/геномні координати для оцінки МХП/МВ із набором реагентів NextSeq 550Dx High Output Reagent Kit v2.5 (300 cycles). Для аналізу МХП (N = 51) було використано всі варіанти із середнім значенням VAF від 0,045 до 0,055 (варіанти МВ) з набору реагентів NextSeq 550Dx High Output Reagent Kit v2 (300 cycles) Для аналізу МХП було оцінено 51 відповідну геномну координату.

Для оцінки набору реагентів NextSeq 550Dx High Output Reagent Kit v2.5 (300 cycles) бібліотеки було секвеновано в трьох прогонах протягом трьох днів поспіль із використанням одного приладу й однієї серії набору реагентів. Це тестування нараховувало 24 повторності для кожного з 51 варіанта МВ і 21 повторність для кожної відповідної позиції дикого типу. Частку розпізнавань варіантів дикого типу з VAF < 0,026 вказано в [Таблиця 29](#). Частку розпізнавань варіантів МВ з VAF, яка дорівнювала або перевищувала 0,026, вказано в [Таблиця 30](#).

Таблиця 29 Частка розпізнавань з VAF < 0,026 для позицій дикого типу (оцінка відповідності заявленим значенням МХП)

Тип варіанта	Оцінено позицій	Загалом спостережень	К-сть вимірювань VAF $\geq 2,6\%$	Частка < 2,6 %	Частка, що входить до 95%-го довірчого інтервалу
SNV	32	672	0	1	0,994–1
Інсерція	11	231	0	1	0,984–1
Делеція	8	168	0	1	0,978–1

Таблиця 30 Частка розпізнавань із VAF $\geq 0,026$ для варіантів МВ (оцінка відповідності заявленим значенням МВ)

Тип варіанта	Оцінено позицій	Загалом спостережень	К-сть вимірювань VAF < 2,6 %	К-сть вимірювань VAF $\geq 2,6\%$	Частка $\geq 2,6\%$	Частка, що входить до 95%-го довірчого інтервалу
SNV	32	768	1	767	0,999	0,993–1
Інсерція	11	264	3	261	0,989	0,967–0,996
Делеція	8	192	2	190	0,99	0,963–0,997

Точність

Зародкова лінія

Це дослідження було проведено для оцінювання точності розпізнавання варіантів за допомогою модуля Germline Variant Module з використанням набору реагентів NextSeq 550Dx High Output Reagent Kit v2.5 (300 cycles). Дванадцять унікальних зразків Platinum Genome було протестовано за допомогою репрезентативного аналізу. Загалом було виконано 11 прогонів із використанням трьох приладів NextSeq 550Dx і трьох наборів реагентів NextSeq 550Dx High Output Reagent Kit v2.5 (300 cycles).

Точність було встановлено для SNV, інсерцій та делецій методом порівняння отриманих результатів із показниками, установленними за допомогою точно охарактеризованого складеного еталонного методу — Platinum Genomes версії 2016-1.0. Результати дослідження точності з одного прогону секвенування з набором реагентів NextSeq 550Dx High Output Reagent Kit v2 (300 cycles) наведено для довідки. Зведення результатів представлено в [Таблиця 31](#).

Таблиця 31 Зведені дані узгодженості для варіантів зародкової лінії

Критерії	Загалом спостережень (вер. 2.5) ¹	Результат за спостереженням (вер. 2.5) ²	Результат за спостереженням (вер. 2) ³	Результат за прогнозом (вер. 2.5) ⁴	Результат за прогнозом (вер. 2) ⁴
PPA для SNV	1056	98,7	98,7	> 99,9	> 99,9
PPA для інсерцій	1056	100	100	100	98,9
PPA для делецій	1056	95,2	95,2	> 99,9	100
NPA	1056	100	100	100	100
OPA	1056	> 99,9	> 99,9	> 99,9	> 99,9

¹Обчислено як кількість зразків на прогін / кількість прогонів (96 зразків на прогін \times 11 прогонів = 1056 спостережень).

²Найменше отримане значення за повторністю зразка для всіх прогонів (на підставі 11 прогонів для набору реагентів NextSeq 550Dx High Output Reagent Kit v2.5).

³Найменше отримане значення за повторністю зразка в 1 прогоні (загалом 96 спостережень).

⁴Найменше значення, коли дані з кожного прогону аналізуються сукупно.

Соматичний

Описане далі дослідження було проведено для оцінювання точності розпізнавання варіантів за допомогою модуля Somatic Variant Module приладу NextSeq 550Dx із використанням набору реагентів NextSeq 550Dx High Output Reagent Kit v2.5 (300 cycles). Десять зразків FFPE Platinum Genome (два з варіантами, розведеними до отримання значення VAF 0,05) було протестовано за допомогою репрезентативного аналізу. Загалом було виконано 11 прогонів із використанням трьох приладів NextSeq 550Dx і трьох серій набору реагентів NextSeq 550Dx High Output Reagent Kit v2.5 (300 cycles).

Точність було встановлено для SNV, інсерцій та делецій методом порівняння отриманих результатів із показниками, установленими за допомогою точно охарактеризованого складеного еталонного методу — Platinum Genomes версії 2016-1.0. Результати дослідження точності з одного прогону секвенування з набором реагентів NextSeq 550Dx High Output Reagent Kit v2 (300 cycles) наведено для довідки. Зведення результатів представлено в [Таблиця 32](#).

Таблиця 32 Зведені дані узгодженості для соматичних варіантів

Критерії	Загалом спостережень (вер. 2.5) ¹	Результат за спостереженням (вер. 2.5) ²	Результат за спостереженням (вер. 2) ³	Результат за прогнозом (вер. 2.5) ⁴	Результат за прогнозом (вер. 2) ⁴
PPA для SNV	528	100	100	100	100
PPA для інсерцій	528	96,9	96,9	> 99,9	> 99,9
PPA для делецій	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

¹Обчислено як кількість зразків на прогін жількість прогонів (48 зразків на прогін × 11 прогонів = 528 спостережень).

²Найменше отримане значення за повторністю зразка для всіх прогонів (на підставі 11 прогонів для набору реагентів NextSeq 550Dx High Output Reagent Kit v2.5).

³Найменше отримане значення за повторністю зразка в 1 прогоні (загалом 96 спостережень).

⁴Найменше значення, коли дані з кожного прогону аналізуються сукупно.

Прецизійність

Зародкова лінія

Прецизійність набору реагентів NextSeq 550Dx High Output Reagent Kit v2.5 (300 cycles) для модуля Germline Variant Module було оцінено з використанням зразків Platinum Genome і репрезентативного аналізу. Тестування полягало в підготовці однієї бібліотеки за допомогою спеціального набору ампліконів TruSeq Custom Amplicon Kit Dx і включало 12 зразків, кожен із яких оброблено з вісьмома повторностями. Бібліотеки було секвеновано з трьома серіями набору реагентів NextSeq 550Dx High Output Reagent Kit v2.5 (300 cycles) на трьох приладах NextSeq 550Dx загалом за дев'ять прогонів секвенування.

Зразки з гетерозиготними варіантами використано для визначення, чи вплине властива аналізу варіабельність на розпізнавання генотипу (N = 153 унікальні гетерозиготні варіанти). Показник Sx було визначено для обох граничних значень модуля Germline Variant Module (які становлять 0,2 для гетерозиготних і 0,7 для гомозиготних генотипів), де x — це частка повторних тестів, які перевищують граничне значення. Для нижнього граничного значення VAF 0,2 частка варіанта з мінімальним Sx для набору реагентів NextSeq 550Dx Reagent Kit v2.5 (300 cycles) становила > 99,9 %, указуючи на те, що > 99,9 % гетерозиготних варіантів будуть розпізнаватися як гетерозиготні. Для нижнього граничного значення VAF 0,7 частка варіанта з максимальним Sx для набору реагентів NextSeq 550Dx Reagent Kit v2.5 (300 cycles) становила < 1,5 %, указуючи на те, що ≤ 1,5 % гетерозиготних варіантів будуть розпізнаватися як гомозиготні. У [Таблиця 33](#) підсумовано результати за типом варіанта. Значення Sx із одного прогону секвенування з використанням набору реагентів NextSeq 550Dx High Output Reagent Kit v2 (300 cycles) наведено для довідки.

Таблиця 33 Значення Cx для гетерозиготних варіантів зародкової лінії

Тип варіанта	N	Граничне значення VAF 0,2		Граничне значення VAF 0,7	
		Мін. Cx (вер. 2.5) ¹	Мін. Cx (вер. 2) ²	Макс. Cx (вер. 2.5) ¹	Макс. Cx (вер. 2) ²
SNV	94	> 99,9 %	> 99,9 %	1,5 %	1,0 %
Інсерції	24	100 %	100 %	0 %	< 0,1 %
Делеції	35	100 %	> 99,9 %	< 0,1 %	< 0,1 %

¹Значення Cx базуються на розрахункових значеннях загального стандартного відхилення за результатами аналізу компонентів дисперсії.

²Значення Cx базуються на значеннях стандартного відхилення зразків.

Соматичний

Прецизійність набору реагентів NextSeq 550Dx High Output Reagent Kit v2.5 (300 cycles) для модуля Somatic Variant Module було оцінено з використанням зразків FFPE Platinum Genome за допомогою репрезентативного аналізу. Тестування полягало в підготовці однієї бібліотеки за допомогою спеціального набору ампліконів TruSeq Custom Amplicon Kit Dx і включало два зразки з вісьмома повторностями кожний. Бібліотеки було секвеновано з використанням трьох серій набору реагентів NextSeq 550Dx High Output Reagent Kit v2.5 (300 cycles) за допомогою трьох приладів NextSeq 550Dx загалом за дев'ять прогонів секвенування.

Соматичні варіанти з прогнозованими рівнями $VAF \leq 0,10$ (N = 131 унікальний варіант) було використано для оцінки варіабельності приладу приблизно на рівні граничного значення для модуля Somatic Variant Module (соматичні варіанти з рівнем $VAF \geq 0,026$ розпізнаються як позитивні для варіанта). Значення C95 було визначено для кожного із соматичних варіантів. Значення C95 представляють VAF, за яких імовірність перевищення граничного значення для модуля Somatic Variant Module становить 95 %. Найвищі значення C95 за типом варіанта наведено в [Таблиця 34](#). Результати C95 з одного прогону секвенування з використанням набору реагентів NextSeq 550Dx High Output Reagent Kit v2 (300 cycles) наведено для довідки.

Таблиця 34 Зведені дані про C95 для соматичних варіантів

Тип варіанта	К-сть оцінених варіантів	C95 (вер. 2.5) ¹	C95 (вер. 2) ²
SNV	74	0,064	0,063
Інсерції	24	0,062	0,061
Делеції	33	0,060	0,060

¹Значення C95 базуються на розрахункових значеннях загального стандартного відхилення за результатами аналізу компонентів дисперсії.

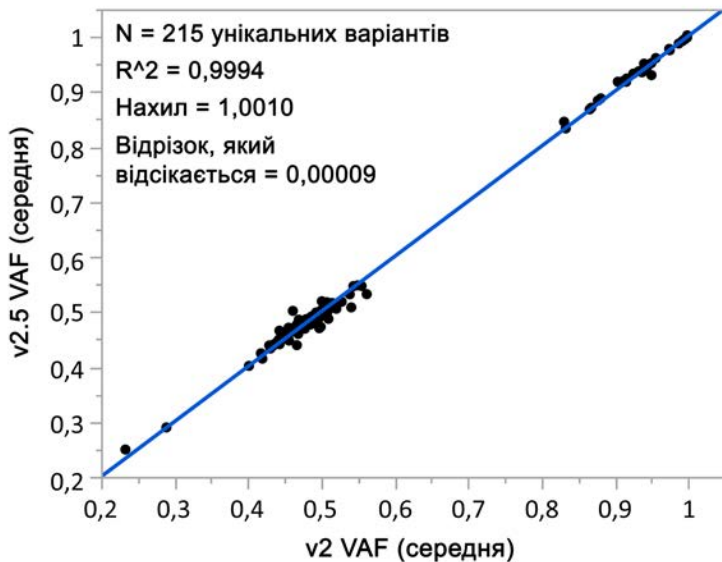
²Значення C95 базуються на значеннях стандартного відхилення зразків.

Порівняння методів (залежно від набору реагентів)

Зародкова лінія

Середні значення VAF із 215 унікальних варіантів було оцінено для набору реагентів NextSeq 550Dx High Output Reagent Kit v2 (300 cycles) і набору реагентів NextSeq 550Dx High Output Reagent Kit v2.5 (300 cycles) з використанням результатів, згенерованих із модуля Germline Variant Module. Середні значення VAF було обчислено з 11 прогонів секвенування (версія 2.5) й одного прогону секвенування (версія 2). Для обчислення середнього значення для кожного варіанта було застосовано не менше восьми повторностей. На [Рисунок 3](#) показано графік кореляції VAF між двома наборами реагентів. На підставі сильної лінійної кореляції VAF і подібності в результатах між наборами реагентів технічні характеристики, початково верифіковані й валідовані за допомогою набору реагентів NextSeq 550Dx High Output Reagent Kit v2 (300 cycles) з модулем Germline Variant Module, визначено як застосовні до набору реагентів NextSeq 550Dx High Output Reagent Kit v2.5 (300 cycles).

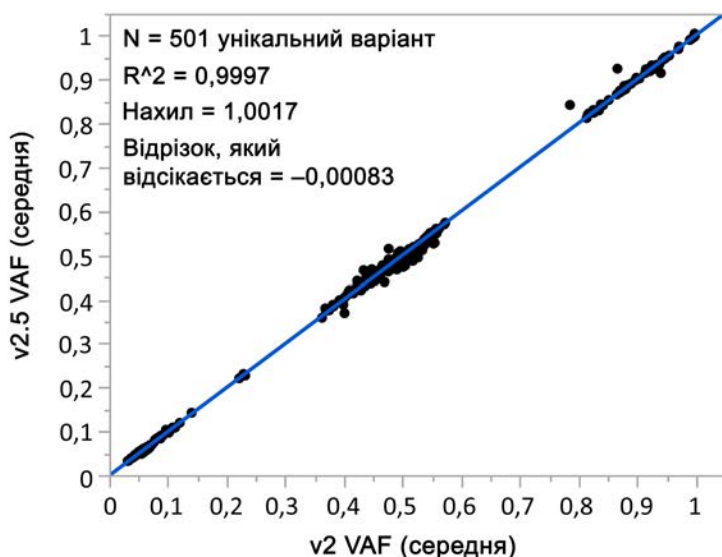
Рисунок 3 Кореляція частоти варіантних алелів (Variant Allele Frequency, VAF) для модуля Germline Variant Module між набором реагентів 550Dx High Output Reagent Kit v2 (300 cycles) і набором реагентів NextSeq 550Dx High Output Reagent Kit v2.5 (300 cycles).



Соматичний

Середні значення VAF із 501 унікального варіанта було оцінено для набору реагентів NextSeq 550Dx High Output Reagent Kit v2 (300 cycles) і набору реагентів NextSeq 550Dx High Output Reagent Kit v2.5 (300 cycles) з використанням результатів, згенерованих із модуля Somatic Variant Module. Середні значення VAF було обчислено з 11 прогонів секвенування (версія 2.5) й одного прогону секвенування (версія 2). Для обчислення середнього значення для кожного унікального варіанта було застосовано не менше трьох повторностей. На [Рисунок 4](#) показано графік кореляції VAF між двома наборами реагентів. На підставі кореляції значень VAF і подібності результатів між наборами реагентів технічні характеристики, верифіковані й валідовані за допомогою набору реагентів NextSeq 550Dx High Output Reagent Kit v2 (300 cycles) з модулем Somatic Variant Module, визначено як застосовні до набору реагентів NextSeq 550Dx High Output Reagent Kit v2.5 (300 cycles).

Рисунок 4 Кореляція частоти варіантних алелів (Variant Allele Frequency, VAF) для модуля Somatic Variant Module між набором реагентів NextSeq 550Dx High Output Reagent Kit v2 (300 cycles) і набором реагентів NextSeq 550Dx High Output Reagent Kit v2.5 (300 cycles).



Історія редакцій

Документ	Дата	Опис змінення
Документ № 1000000030326, версія 06	Травень 2022 р.	Оновлено з метою виправлення інформації, доданої ненавмисно з вихідного програмного забезпечення.
Документ № 1000000030326, версія 05	Листопад 2021 р.	У розділ «Застереження» додано пункт про звітність щодо серйозних інцидентів. У розділ «Принципи виконання процедури» додано пункт, у якому визначено передбачуваних користувачів. Видалено посилання на набір реагентів High Output Reagent Kit v2 (300 cycles). Додано посилання на набір реагентів High Output Reagent Kit v2.5 (75 cycles).
Документ № 1000000030326, версія 04	Серпень 2021 р.	Додано таблицю з історією редакцій. Оновлено адресу уповноваженого представника в Європейському Союзі.

Патенти й товарні знаки

Цей документ і його зміст є власністю компанії Illumina, Inc. і її філій (надалі — Illumina). Він призначений лише для того, щоб користувач використовував вироби тільки за угодою в цілях, описаних у цьому документі. Цей документ і його зміст не слід використовувати або поширювати з будь-якою іншою метою та/або для іншого обговорення, розкриття або відтворення тим або іншим чином без попередньої письмової згоди компанії Illumina. Цим документом компанія Illumina не надає жодного дозволу на свій патент, товарний знак, авторське право або загальноприйняті права, а також на подібні права будь-яких третіх сторін.

Щоб гарантувати правильне та безпечне використання виробів, описаних у цьому документі, кваліфікований і належно навчений персонал повинен суворо та чітко дотримуватись інструкцій, описаних у цьому документі. Перед використанням цих виробів потрібно повністю прочитати й зрозуміти весь зміст цього документа.

НЕПОВНЕ ВИВЧЕННЯ ВСІХ ЗАЗНАЧЕНИХ У ЦЬОМУ ДОКУМЕНТІ ВКАЗІВОК І ЇХНЄ НЕЧІТКЕ ДОТРИМАННЯ МОЖЕ ПРИЗВОДИТИ ДО ПОШКОДЖЕННЯ ЦИХ ВИРОБІВ, ТРАВМУВАННЯ ЛЮДЕЙ, ВКЛЮЧНО З КОРИСТУВАЧАМИ АБО ІНШИМИ ОСОБАМИ, І ПОШКОДЖЕННЯ ІНШОЇ ВЛАСНОСТІ, А ТАКОЖ ПРИЗВЕДЕ ДО ВТРАТИ БУДЬ-ЯКИХ ГАРАНТІЙНИХ ЗОБОВ'ЯЗАНЬ, ЗАСТОСОВНИХ ДО ЦИХ ВИРОБІВ.

КОМПАНІЯ ILLUMINA НЕ НЕСЕ ЖОДНОЇ ВІДПОВІДАЛЬНОСТІ, ЩО ВИНΙΚАЄ ВНАСЛІДОК НЕНАЛЕЖНОГО ВИКОРИСТАННЯ ВИРОБІВ, ОПИСАНИХ У ЦЬОМУ ДОКУМЕНТІ (ВКЛЮЧНО З ЙОГО ЧАСТИНАМИ АБО ПРОГРАМНИМ ЗАБЕЗПЕЧЕННЯМ).

© Illumina, Inc., 2022. Усі права застережено.

Усі товарні знаки — власність компанії Illumina, Inc. або їхніх відповідних власників. Конкретна інформація про товарні знаки зазначена на сторінці www.illumina.com/company/legal.html.

Контактна інформація



Illumina

5200 Illumina Way

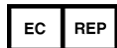
San Diego, California 92122 U.S.A. (США)

+1 800 809.ILMN (4566)

+1 858 202 4566 (за межами Північної Америки)

techsupport@illumina.com

www.illumina.com



Illumina Netherlands B.V.
Steenoven 19
5626 DK Eindhoven
The Netherlands (Нідерланди)

Австралійський спонсор

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

Маркування виробу

Повний список символів, які може бути зображено на упаковці або маркуванні виробу, див. в поясненні символів на вебсайті: support.illumina.com.

Уповноважений представник в Україні (Authorized Representative in Ukraine):



ТОВ «БІОЛАБТЕХ ЛТД»

проспект Героїв Сталінграда буд.42-А, кв.45

м. Київ, 04213, Україна

Тел.: +38 044 492 81 88

Електронна адреса: info@biolabtech.com.ua, ЄДРПОУ

34891619