

## Tờ hướng dẫn sử dụng

DÙNG CHO CHẨN ĐOÁN TRONG ỐNG NGHIỆM. CHỈ DÙNG CHO MỤC ĐÍCH XUẤT KHẨU.

### Danh mục số 20005715

## Mục đích sử dụng

NextSeq 550Dx instrument được thiết kế để giải trình tự các thư viện ADN khi được sử dụng với các xét nghiệm chẩn đoán *trong ống nghiệm*. NextSeq 550Dx instrument phải được sử dụng với thuốc thử chẩn đoán *trong ống nghiệm* và phần mềm phân tích cụ thể đã được đăng ký, chứng nhận hoặc phê duyệt.

## Nguyên tắc của quy trình

Illumina NextSeq 550Dx instrument được thiết kế để giải trình tự các thư viện ADN thông qua các xét nghiệm chẩn đoán *trong ống nghiệm* và dành cho các nhân viên phòng thí nghiệm lâm sàng đủ trình độ và được đào tạo về cách sử dụng các quy trình chẩn đoán trong ống nghiệm được thực hiện trong phòng thí nghiệm lâm sàng. Đối với đầu vào, NextSeq 550Dx sử dụng thư viện được tạo từ ADN trong đó chỉ thị mẫu và trình tự thu thập được thêm vào mục tiêu đã khuếch đại. Thư viện mẫu được thu thập trên tế bào dòng chảy và giải trình tự trên thiết bị nhờ công nghệ hóa học giải trình tự theo nguyên lý tổng hợp (SBS). Công nghệ hóa học SBS sử dụng phương pháp vùng kết thúc thuận nghịch để phát hiện các base nucleotide đơn lẻ đã gắn nhãn huỳnh quang khi các base này được kết hợp vào các sợi ADN đang phát triển. Phần mềm Real-Time Analysis (RTA) thực hiện phân tích hình ảnh và phát hiện base đồng thời chỉ định điểm chất lượng cho từng base cho mỗi chu kỳ giải trình tự. Khi quá trình phân tích chính kết thúc, quá trình phân tích phụ có thể được thực hiện trên thiết bị để xử lý các phát hiện base. NextSeq 550Dx sử dụng các mô-đun phân tích phụ khác nhau tùy theo quy trình công việc. Đối với Germline Variant Module hay Somatic Variant Module, quá trình xử lý bao gồm tách đoạn, tạo tệp FASTQ, so khớp, phát hiện biến thể và tạo tệp định dạng phát hiện biến thể (VCF và gVCF). Các tệp VCF và gVCF chứa thông tin về các biến thể được tìm thấy ở các vị trí cụ thể trong hệ gen tham chiếu.

## Cấu hình khởi động kép

NextSeq 550Dx bao gồm cấu hình khởi động kép để người dùng có thể sử dụng thiết bị ở chế độ chẩn đoán (Dx) hoặc chỉ dùng cho mục đích nghiên cứu (RUO). Các xét nghiệm giải trình tự chẩn đoán *trong ống nghiệm*, bao gồm Germline Variant Module và Somatic Variant Module, được thực hiện ở chế độ chẩn đoán. Chỉ có thể sử dụng thuốc thử giải trình tự chẩn đoán trong ống nghiệm IVD trong chế độ chẩn đoán. Các đặc điểm hiệu suất và giới hạn của quy trình đối với NextSeq 550Dx instrument đã được thiết lập bằng Germline Variant Module và Somatic Variant Module ở chế độ chẩn đoán.

# Giới hạn của quy trình

1. Dùng cho chẩn đoán *trong ống nghiệm*.
2. Germline Variant Module và Somatic Variant Module, khi được sử dụng với NextSeq 550Dx High Output Reagent Kit v2 (300 chu kỳ) hoặc NextSeq 550Dx High Output Reagent Kit v2.5 (300 chu kỳ), có khả năng mang lại:
  - Đầu ra giải trình tự  $\geq 90$  gigabase (Gb)
  - Độ dài đoạn đọc (trong lần chạy kết đôi)  $2 \times 150$  cặp base (bp)
  - Các base bằng hoặc cao hơn Q30  $\geq 75\%$  ở độ dài đoạn đọc  $2 \times 150$  bp  
Từ 75% base trở lên có điểm chất lượng theo thang đo Phred  $\geq 30$ , cho biết độ chính xác khi phát hiện base cao hơn 99,9%
3. Phần mềm xét nghiệm không so khớp các đoạn đọc có indel (vùng chèn, vùng xóa hoặc kết hợp) trong đó độ dài thành phần  $> 25$  bp. Do đó, phần mềm xét nghiệm không phát hiện được các indel có độ dài  $> 25$  bp.
4. Phần mềm xét nghiệm có thể không so khớp đoạn đọc amplicon có thành phần biến thể cực hạn, dẫn đến việc vùng đó được báo cáo là thể đại. Các thành phần cực hạn này bao gồm:
  - Đoạn đọc chứa nhiều hơn ba indel
  - Đoạn đọc có độ dài tối thiểu 30 bp với hàm lượng biến thể nucleotide đơn (SNV)  $> 4\%$  tổng độ dài mục tiêu amplicon (không bao gồm các vùng thăm dò)
  - Đoạn đọc có độ dài  $< 30$  bp với hàm lượng SNV  $> 10\%$  tổng độ dài amplicon (bao gồm các vùng thăm dò)
5. Biến thể lớn, bao gồm các biến thể đa nucleotide (MNV) và các indel lớn, có thể được báo cáo là các biến thể nhỏ hơn riêng biệt trong tệp VCF đầu ra.
6. Các biến thể xóa có thể bị lọc ra hoặc bỏ sót khi mở rộng hai amplicon dạng ô nếu độ dài vùng xóa lớn hơn hoặc bằng vùng chồng lấn giữa các amplicon dạng ô.
7. Hệ thống không thể phát hiện indel nếu các biến thể này xảy ra ngay liền kề mỗi và không có amplicon chồng lấn. Đối với các vùng có amplicon chồng lấn, xét nghiệm không thể phát hiện vùng xóa khi vùng chồng lấn nhỏ hơn kích thước vùng xóa cần phát hiện. Ví dụ: nếu vùng chồng lấn giữa hai amplicon liền kề là hai base, xét nghiệm sẽ không thể phát hiện vùng xóa nào bao gồm cả hai base đó. Có thể phát hiện một vùng xóa base duy nhất tại một trong hai base đó.
8. Như với bất kỳ quy trình công việc chuẩn bị thư viện dựa trên quá trình tạo thể lai nào, các đa hình cơ bản, đột biến, vùng chèn hoặc vùng xóa trong các vùng gắn kết oligonucleotide có thể ảnh hưởng đến các alen được thăm dò và các kết quả phát hiện trong khi giải trình tự. Ví dụ:
  - Một biến thể cùng pha với một biến thể trong vùng mồi có thể không được khuếch đại, dẫn đến trường hợp âm tính giả.
  - Các biến thể trong vùng mồi có thể ngăn cản sự khuếch đại của alen tham chiếu, dẫn đến phát hiện biến thể đồng hợp tử không chính xác.
  - Các biến thể indel trong vùng mồi có thể dẫn đến phát hiện dương tính giả ở cuối đoạn đọc liền kề với mồi.

9. Xét nghiệm có thể lọc các indel do sai lệch sợi nếu các biến thể này xảy ra gần cuối một đoạn đọc và bị cắt mềm trong quá trình so khớp.
10. MNV nhỏ chưa được kiểm nhận và chỉ được báo cáo trong Somatic Variant Module.
11. Các vùng xóa được báo cáo trong VCF tại tọa độ của base nằm trước theo định dạng VCF. Do đó, hãy xem xét các biến thể liền kề trước khi báo cáo một phát hiện base riêng lẻ là tham chiếu đồng hợp tử.
12. Các giới hạn dành riêng cho dòng mầm:
  - NextSeq 550Dx instrument, sử dụng Germline Variant Module của Local Run Manager cho NextSeq 550Dx, được thiết kế để cung cấp kết quả định tính cho phát hiện biến thể dòng mầm (ví dụ: đồng hợp tử, dị hợp tử, thể đại).
  - Khi được sử dụng với Germline Variant Module, phạm vi tối thiểu cần thiết trên mỗi amplicon để phát hiện biến thể chính xác là 150x. Do đó, cần có 150 phân đoạn ADN hỗ trợ, tương đương với 300 đoạn đọc kết đôi chồng lấn. Số lượng mẫu và tổng số base nhằm mục tiêu sẽ ảnh hưởng đến phạm vi. Hàm lượng GC và hàm lượng hệ gen khác có thể ảnh hưởng đến phạm vi.
  - Sự thay đổi số lượng bản sao có thể ảnh hưởng đến việc xác định một biến thể là đồng hợp tử hay dị hợp tử.
  - Các biến thể trong trường hợp lặp lại nhất định được lọc ra khỏi các tệp VCF. Bộ lọc lặp lại RMxN được sử dụng để lọc biến thể nếu tất cả hoặc một phần trình tự biến thể xuất hiện lặp đi lặp lại trong hệ gen tham chiếu liền kề với vị trí biến thể. Đối với phát hiện biến thể dòng mầm, cần ít nhất 9 trình tự lặp lại trong tham chiếu để lọc biến thể. Chỉ xét những trình tự lặp lại có độ dài tối đa 5 bp (R5x9).
  - Một indel và một SNV ở lô-cut đơn có thể chỉ cho một biến thể trong báo cáo.
13. Các giới hạn dành riêng cho sinh dưỡng.
  - NextSeq 550Dx instrument, sử dụng Somatic Variant Module của Local Run Manager cho NextSeq 550Dx, được thiết kế để cung cấp kết quả định tính cho phát hiện biến thể sinh dưỡng (ví dụ: sự hiện diện của một biến thể sinh dưỡng có tần số biến thể lớn hơn hoặc bằng 0,026 với giới hạn phát hiện là 0,05).
  - Khi được sử dụng với Somatic Variant Module, phạm vi tối thiểu cần thiết trên mỗi amplicon để phát hiện biến thể chính xác là 450x trên mỗi nhóm gộp oligonucleotide. Do đó, cần có 450 phân đoạn ADN hỗ trợ trên mỗi nhóm gộp oligonucleotide, tương đương với 900 đoạn đọc kết đôi chồng lấn. Số lượng mẫu và tổng số base nhằm mục tiêu sẽ ảnh hưởng đến phạm vi. Hàm lượng GC và hàm lượng hệ gen khác có thể ảnh hưởng đến phạm vi.
  - Đối với phát hiện biến thể sinh dưỡng, cần ít nhất sáu trình tự lặp lại trong tham chiếu để lọc biến thể và chỉ xét những trình tự lặp lại có độ dài tối đa 3 bp (R3x6).
  - Somatic Variant Module thể sinh dưỡng không thể phân biệt giữa biến thể dòng mầm và sinh dưỡng. Mô-đun này được thiết kế để phát hiện các biến thể trên một loạt các tần số biến thể, nhưng không thể sử dụng tần số biến thể để phân biệt các biến thể sinh dưỡng với các biến thể dòng mầm.
  - Mô bình thường trong mẫu xét nghiệm ảnh hưởng đến việc phát hiện các biến thể. Giới hạn phát hiện được báo cáo dựa trên tần số biến thể liên quan đến tổng số ADN tách chiết từ cả mô khối u và mô bình thường.

# Thành phần sản phẩm

Illumina NextSeq 550Dx bao gồm những điều sau đây:

1. NextSeq 550Dx instrument (Danh mục số 20005715)
2. Các thành phần phần mềm cho NextSeq 550Dx instrument, bao gồm:

Ứng dụng phần mềm	Chức năng	Mô tả
NextSeq 550Dx Operating Software (NOS)	Điều khiển hoạt động của thiết bị	Ứng dụng phần mềm NOS quản lý hoạt động của thiết bị trong quá trình giải trình tự và tạo hình ảnh cho phần mềm Real-Time Analysis (RTA) sử dụng.
Real-time Analysis Software (RTA)	Thực hiện quá trình phân tích chính	Ứng dụng phần mềm RTA chuyển đổi các hình ảnh do NOS tạo cho mỗi ô trong mỗi chu kỳ chạy giải trình tự thành các tệp phát hiện base, là đầu vào cho các mô-đun phân tích của Local Run Manager. Ứng dụng phần mềm RTA không có giao diện người dùng.
Local Run Manager	Giao diện cho phép lựa chọn mô-đun	Phần mềm Local Run Manager là một giải pháp tích hợp trên thiết bị để quản lý người dùng, chọn mô-đun phân tích phù hợp và giám sát trạng thái.
Somatic Variant Module	Thực hiện quá trình phân tích phụ	Phần mềm mô-đun phân tích này của Local Run Manager xử lý phát hiện base thông qua quá trình phân tích phụ. Quá trình xử lý bao gồm tách đoạn, tạo tệp FASTQ, so khớp, phát hiện biến thể và báo cáo. Trình phát hiện biến thể (Pisces) tạo tệp VCF chứa thông tin về các biến thể phát hiện được ở các vị trí cụ thể trong hệ gen tham chiếu và bao gồm tần số biến thể đo được.
Germline Variant Module	Thực hiện quá trình phân tích phụ	Phần mềm mô-đun phân tích này của Local Run Manager xử lý phát hiện base thông qua quá trình phân tích phụ. Quá trình xử lý bao gồm tách đoạn, tạo tệp FASTQ, so khớp, phát hiện biến thể và báo cáo. Trình phát hiện biến thể (Pisces) tạo tệp VCF chứa thông tin về các biến thể phát hiện được ở các vị trí cụ thể trong hệ gen tham chiếu và xác định mỗi biến thể là dị hợp tử hay đồng hợp tử.

3. **Không bắt buộc** Illumina DRAGEN Server for NextSeq 550Dx (Danh mục số 20086130), bao gồm thành phần phần mềm sau:

Ứng dụng phần mềm	Chức năng	Mô tả
Illumina Run Manager	Giao diện để lựa chọn mô-đun ứng dụng	Phần mềm Illumina Run Manager được cài đặt trên máy chủ DRAGEN không cần thiết bị, tùy chọn. Illumina Run Manager cho phép quản lý người dùng, lựa chọn mô-đun phân tích và theo dõi trạng thái chạy và phân tích trình tự.

Illumina DRAGEN Server for NextSeq 550Dx chỉ có sẵn ở một số quốc gia. Liên hệ với đại diện của Illumina để biết tình trạng sẵn có sản phẩm tại khu vực.

## Điều kiện vận hành

Để biết thêm thông tin về các điều kiện vận hành, hãy xem phần Các lưu ý về môi trường trong *Hướng dẫn chuẩn bị khu vực làm việc của NextSeq 550Dx Instrument (tài liệu số # 1000000009869)*.

Yếu tố	Thông số kỹ thuật
Nhiệt độ	Duy trì nhiệt độ phòng thí nghiệm trong khoảng từ 19°C đến 25°C (22°C ±3°C). Đây là nhiệt độ hoạt động của thiết bị. Trong thời gian chạy, không để nhiệt độ môi trường xung quanh thay đổi quá ±2°C.
Độ ẩm	Duy trì độ ẩm tương đối không ngưng tụ trong khoảng 20-80%.

Để biết hướng dẫn, cảnh báo và thông tin mới nhất về bảo mật của các hệ thống Illumina, hãy tham khảo [Bảo mật và Kết nối mạng](#).

## Thiết bị và vật liệu

### Các thiết bị và vật liệu cần tự chuẩn bị, bán riêng

NextSeq 550Dx High Output Reagent Kit v2.5 (75 chu kỳ), Danh mục số 20028870

NextSeq 550Dx High Output Reagent Kit v2.5 (300 chu kỳ), Danh mục số 20028871

## Các thiết bị và vật liệu cần tự chuẩn bị chứ không được cung cấp

### Vật tư tiêu hao do người dùng cung cấp cho các lần chạy giải trình tự

Vật tư tiêu hao	Nhà cung cấp	Mục đích
Khăn lau tấm cồng, isopropyl 70% hoặc Ethanol, 70%	VWR, danh mục số 95041-714 (hoặc tương đương) Nhà cung cấp vật tư phòng thí nghiệm thông thường	Làm sạch tế bào dòng chảy và mục đích thông dụng
Giấy lụa dùng cho phòng thí nghiệm, ít bụi	VWR, danh mục số 21905-026 (hoặc tương đương)	Làm sạch tế bào dòng chảy và mục đích thông dụng

### Vật tư tiêu hao do người dùng cung cấp để bảo trì thiết bị

Vật tư tiêu hao	Nhà cung cấp	Mục đích
NaOCl, 5% (natri hypoclorit)	Sigma-Aldrich, danh mục số 239305 (hoặc loại tương đương ở cấp độ dùng trong phòng thí nghiệm)	Rửa thiết bị bằng quy trình rửa thủ công sau khi chạy; pha loãng đến mức 0,12%
Tween 20	Sigma-Aldrich, danh mục số P7949	Rửa thiết bị bằng tùy chọn rửa thủ công; pha loãng đến mức 0,05%
Nước, cấp độ dùng trong phòng thí nghiệm	Nhà cung cấp vật tư phòng thí nghiệm thông thường	Rửa thiết bị (quy trình rửa thủ công)
Bộ lọc không khí	Illumina, danh mục số 20063988	Làm sạch không khí mà thiết bị hút vào để làm mát

## Hướng dẫn về nước ở cấp độ dùng trong phòng thí nghiệm

Luôn sử dụng nước ở cấp độ dùng trong phòng thí nghiệm hoặc nước khử ion để thực hiện các quy trình của thiết bị. Tuyệt đối không dùng nước máy. Chỉ sử dụng các cấp độ nước hoặc các loại tương đương sau:

- Nước khử ion
- Illumina PW1
- Nước 18 Megohms (MΩ)
- Nước Milli-Q
- Nước Super-Q
- Nước cấp độ sinh học phân tử

## Cảnh báo và biện pháp phòng ngừa



### THẬN TRỌNG

Luật liên bang chỉ cho phép thiết bị này được bán bởi hoặc theo yêu cầu của bác sĩ hoặc bác sĩ hành nghề khác được cấp phép theo luật của Tiểu bang nơi bác sĩ đó hành nghề, sử dụng hoặc yêu cầu sử dụng thiết bị.

1. **Một số thành phần của thuốc thử do Illumina cung cấp để sử dụng với NextSeq 550Dx instrument có chứa các hóa chất có thể gây nguy hiểm. Có thể xảy ra thương tích cá nhân nếu hít phải, nuốt phải, tiếp xúc với da và mắt. Mang thiết bị bảo hộ, bao gồm bảo vệ mắt, găng tay và áo choàng phòng thí nghiệm tương ứng với các nguy cơ phơi nhiễm. Xử lý thuốc thử đã sử dụng như chất thải hóa học và thải bỏ theo các luật và quy định hiện hành của địa phương, quốc gia và khu vực. Để biết thêm thông tin về môi trường, sức khỏe và an toàn, hãy xem Bảng dữ liệu an toàn (SDS) tại địa chỉ [support.illumina.com/sds.html](https://support.illumina.com/sds.html).**
2. Xử lý tất cả các mẫu máu như thể các mẫu đó được biết là có khả năng lây nhiễm vi-rút suy giảm miễn dịch ở người (HIV), vi-rút viêm gan B ở người (HBV) và các tác nhân mầm bệnh lây qua đường máu khác (các biện pháp phòng ngừa chung).
3. Việc không tuân thủ các quy trình đã nêu có thể dẫn đến kết quả sai hoặc chất lượng mẫu giảm sút đáng kể.
4. Sử dụng các biện pháp phòng ngừa thường quy của phòng thí nghiệm. Không dùng pipet bằng miệng. Không ăn uống hoặc hút thuốc trong khu vực làm việc được chỉ định. Đeo găng tay và mặc áo khoác phòng thí nghiệm dùng một lần khi xử lý các mẫu xét nghiệm và thuốc thử trong bộ kit. Rửa tay kỹ càng sau khi xử lý các mẫu xét nghiệm và thuốc thử trong bộ kit.
5. Yêu cầu áp dụng phương pháp thực hành trong phòng thí nghiệm thích hợp và vệ sinh trong phòng thí nghiệm sạch sẽ để ngăn chặn các sản phẩm phản ứng chuỗi polymerase (PCR) làm nhiễm bẩn thuốc thử, thiết bị đo và mẫu ADN hệ gen. Tình trạng nhiễm bẩn PCR có thể khiến cho kết quả không chính xác và không đáng tin cậy.
6. Để tránh tình trạng nhiễm bẩn, hãy đảm bảo rằng khu vực tiền khuếch đại và hậu khuếch đại có thiết bị và vật tư tiêu hao chuyên dụng (ví dụ: pipet, đầu pipet, khối ủ nhiệt, máy trộn mẫu và máy ly tâm).

7. Việc ghép cặp mẫu với chỉ thị phải khớp chính xác với cách bố trí khay đã in. Local Run Manager tự động điền các môi chỉ thị liên kết với tên mẫu khi tên mẫu được nhập vào mô-đun. Người dùng nên xác minh mỗi chỉ thị liên kết với mẫu trước khi bắt đầu chạy giải trình tự. Mẫu không khớp với cách bố trí khay sẽ dẫn đến việc mất khả năng nhận dạng mẫu dương tính và báo cáo kết quả không chính xác.
8. Đặc biệt, người dùng nên tự cài đặt phần mềm chống vi-rút để bảo vệ máy tính khỏi vi-rút. Tham khảo hướng dẫn sử dụng để biết hướng dẫn cài đặt.
9. Không vận hành NextSeq 550Dx khi đã tháo bất kỳ tấm vỏ nào. Việc vận hành thiết bị khi đã tháo bất kỳ tấm vỏ nào đều dẫn đến nguy cơ tiếp xúc với điện áp đường dây và điện áp DC.
10. Không chạm vào bộ tế bào dòng chảy trong khoang chứa tế bào dòng chảy. Máy sưởi trong khoang chứa này vận hành trong khoảng từ 22°C đến 95°C và có thể gây bỏng.
11. Thiết bị nặng khoảng 185 lb và có thể gây thương tích nghiêm trọng nếu bị rơi hoặc thao tác sai cách.
12. Illumina nhận ra nhu cầu liên tục phải quản lý các mối đe dọa an ninh mạng, để biết hướng dẫn bảo mật mới nhất, cảnh báo và thông tin cho các hệ thống Illumina, hãy tham khảo mục [An ninh và Mạng](#).

## Hướng dẫn sử dụng

Các hướng dẫn sau đây để sử dụng NextSeq 550Dx instrument sẽ cần các thuốc thử được cung cấp trong NextSeq 550Dx High Output Reagent Kit v2.5 (300 chu kỳ) hoặc NextSeq 550Dx High Output Reagent Kit v2.5 (75 chu kỳ).

## Tạo lần chạy

Tạo một lần chạy trình tự bằng cách sử dụng Local Run Manager hoặc Illumina Run Manager. Hướng dẫn sử dụng Local Run Manager được trình bày dưới đây và trong NextSeq 550Dx Instrument Reference Guide (document # 100000009513). Để biết hướng dẫn về cách tạo một lần chạy bằng Illumina Run Manager, tham khảo Illumina Run Manager for NextSeq 550Dx Software Guide (document # 200025239).

Để biết hướng dẫn chọn giữa Local Run Manager hoặc Illumina Run Manager, tham khảo Illumina Run Manager for NextSeq 550Dx Software Guide (document # 200025239). Để biết hướng dẫn chi tiết về các ứng dụng cụ thể, hãy tham khảo hướng dẫn mô-đun hoặc ứng dụng để biết xét nghiệm cụ thể.

Các hướng dẫn sau đây liên quan đến việc sử dụng các mô-đun Germline và Somatic Variant Local Run Manager.

## Thiết lập thông số

1. Đăng nhập vào Local Run Manager.
2. Chọn **Create Run** (Tạo lần chạy) và chọn **Somatic Variant** (Biến thể sinh dưỡng) hoặc **Germline Variant** (Biến thể dòng mầm).
3. Nhập tên lượt chạy giúp xác định lượt chạy từ bước giải trình tự đến bước phân tích.  
Sử dụng các ký tự chữ và số, dấu cách, dấu gạch dưới hoặc dấu gạch ngang.
4. [Không bắt buộc] Nhập thông tin mô tả lần chạy để giúp xác định lần chạy.



Sử dụng các ký tự chữ và số, dấu cách, dấu gạch dưới hoặc dấu gạch ngang.

5. Chọn số mẫu và bộ chỉ thị từ danh sách thả xuống.  
Hãy cân nhắc thông tin sau đây khi bạn lựa chọn.
  - Danh sách thả xuống chứa số mẫu cùng với một bộ chỉ thị. Ví dụ: 24-Set 1 (24-Bộ 1) thể hiện rằng 24 mẫu sẽ được xét nghiệm, với các chỉ thị từ bộ chỉ thị 1.
  - Số bộ chỉ thị đề cập đến các bộ cặp chỉ thị i5 và i7 khác nhau. Set 1 (Bộ 1) và Set 2 (Bộ 2) đều cung cấp sự đa dạng về chỉ thị. Hai bộ chỉ thị được cung cấp để giúp ngăn tình trạng thiếu hụt nếu chỉ có một bộ duy nhất.
  - Chọn số mẫu gần nhất với số mẫu bạn sẽ xét nghiệm. Nếu danh sách không có số mẫu chính xác bạn cần, hãy chọn số mẫu gần nhất, nhưng ít hơn số mẫu bạn sẽ xét nghiệm. Ví dụ: nếu bạn muốn xét nghiệm 18 mẫu, hãy chọn 16 mẫu.
  - Tổ hợp các giếng mẫu và chỉ thị được đề xuất đáp ứng các yêu cầu về đa dạng chỉ thị được đánh dấu bằng màu xanh lục.

## Nhập tệp phiếu kê khai cho lần chạy

1. Đảm bảo rằng các phiếu kê khai mà bạn muốn nhập có sẵn ở một vị trí mạng có thể truy cập hoặc trên ổ USB.
2. Chọn **Import Manifests** (Nhập phiếu kê khai).
3. Chuyển đến tệp phiếu kê khai và chọn các phiếu kê khai mà bạn muốn thêm.

**LƯU Ý** Để chuẩn bị sẵn sàng tệp phiếu kê khai cho tất cả các lần chạy bằng mô-đun phân tích Germline Variant hoặc Somatic Variant, hãy thêm phiếu kê khai bằng tính năng Module Settings (Cài đặt mô-đun). Tính năng này yêu cầu bạn phải có quyền cấp người dùng quản trị viên. Để biết thêm thông tin, tham khảo *NextSeq 550Dx Instrument Reference Guide (document # 1000000009513)*.

## Chỉ định mẫu cho lần chạy

Chỉ định mẫu cho lượt chạy bằng một trong các tùy chọn và các hướng dẫn sau đó.


**Nhập mẫu theo cách thủ công** — Sử dụng bảng trống trên màn hình Create Run (Tạo lần chạy).

**Nhập mẫu** — Chuyển đến tệp bên ngoài ở định dạng giá trị được phân tách bằng dấu phẩy (\*.csv). Trên màn hình Create Run (Tạo lần chạy) sẽ có sẵn một mẫu cho bạn tải xuống.


## Nhập mẫu theo cách thủ công

1. Nhập một tên mẫu duy nhất (*Mô-đun phân tích Somatic Variant*) hoặc ID mẫu (*Mô-đun phân tích Germline Variant*).  
Sử dụng các ký tự chữ và số, dấu gạch ngang hoặc dấu gạch dưới.
2. [Không bắt buộc] Đối với các mẫu chứng dương hoặc âm, nhấp chuột phải và chọn loại mẫu chứng.

Mẫu chứng trong một giếng mẫu sẽ tự động điền vào giếng tương ứng trong nhóm gộp khác có cùng mẫu chứng.

- [Không bắt buộc] Nhập thông tin mô tả mẫu vào trường Sample Description (Mô tả mẫu). Sử dụng các ký tự chữ và số, dấu gạch ngang hoặc dấu gạch dưới.
- Chọn một adapter Chỉ thị 1 từ danh sách thả xuống Index 1 (i7)[Chỉ thị 1 (i7)]. Khi bạn sử dụng các giếng mẫu được đề xuất, phần mềm sẽ tự động điền các adapter chỉ thị i7 và i5 đáp ứng các yêu cầu về đa dạng chỉ thị. Nếu danh sách không có số mẫu chính xác mà bạn sẽ xét nghiệm, hãy nhớ chọn adapter chỉ thị cho các giếng trội thêm.
- Chọn một adapter Chỉ thị 2 từ danh sách thả xuống Index 2 (i5) [Chỉ thị 2 (i5)].
- Chọn tệp phiếu kê khai từ danh sách thả xuống Manifest (Phiếu kê khai). Các mẫu trong Nhóm gộp A yêu cầu phiếu kê khai khác với các mẫu trong Nhóm gộp B.
- Chọn một tùy chọn để xem, in hoặc lưu cách bố trí khay làm tham chiếu khi chuẩn bị thư viện:
  - Chọn biểu tượng  **Print** (In) để hiển thị cách bố trí khay. Chọn **Print** (In) để in cách bố trí khay.
  - Chọn **Export** (Xuất) để xuất thông tin mẫu ra tệp bên ngoài.
- Chọn **Save Run** (Lưu lần chạy).

## Nhập mẫu

- Chọn **Import Samples** (Nhập mẫu) rồi duyệt đến vị trí của tệp thông tin mẫu. Bạn có thể nhập hai loại tệp.
  - Chọn **Template** (Mẫu) trên màn hình Create Run (Tạo lần chạy) để tạo cách bố trí khay mới. Tệp mẫu chứa các tiêu đề cột chính xác để nhập thông tin. Nhập thông tin mẫu vào mỗi cột cho các mẫu trong lần chạy. Xóa thông tin ví dụ trong các ô không sử dụng, sau đó lưu tệp.
  - Sử dụng tệp thông tin mẫu đã được xuất từ mô-đun Germline Variant hoặc Somatic Variant bằng tính năng Export (Xuất).
- Chọn biểu tượng  **Print** (In) để hiển thị cách bố trí khay.
- Chọn **Print** (In) để in cách bố trí khay làm tham chiếu khi chuẩn bị thư viện.
- Chọn **Save Run** (Lưu lần chạy).

## Chuẩn bị hộp thuốc thử

Bạn cần cẩn thận làm theo các hướng dẫn của hộp thuốc thử để giải trình tự thành công.

- Lấy hộp thuốc thử ra khỏi nơi bảo quản có nhiệt độ từ -25°C đến -15°C.
- Chọn một trong các phương pháp sau đây để rã đông thuốc thử. Không nhấn chìm hộp. Sau khi rã đông hộp, hãy lau khô hộp trước khi chuyển sang bước tiếp theo.

Nhiệt độ	Thời gian cần để rã đông	Giới hạn ổn định
Bồn nước 15°C đến 30°C	60 phút	Không quá 6 giờ
2°C đến 8°C	7 giờ	Không quá 5 ngày

**LƯU Ý** Nếu rã đông nhiều hộp trong cùng một bồn nước, hãy tăng thêm thời gian rã đông.

3. Úp ngửa hộp năm lần để trộn các thuốc thử.
4. Kiểm tra đáy hộp để đảm bảo rằng thuốc thử đã được rã đông và không có kết tủa. Xác nhận rằng các vị trí số 29, 30, 31 và 32 đã được rã đông, vì đây là các vị trí lớn nhất và mất nhiều thời gian rã đông nhất.
5. Gõ nhẹ lên bàn thí nghiệm để giảm bọt khí.  
Để có kết quả tốt nhất, hãy chuyển ngay sang quá trình nạp mẫu và thiết lập lần chạy.

## Chuẩn bị tế bào dòng chảy

1. Lấy một hộp tế bào dòng chảy mới ra khỏi nơi bảo quản có nhiệt độ từ 2°C đến 8°C.
2. Lấy túi nhôm ra khỏi hộp và để ở nhiệt độ phòng trong 30 phút.

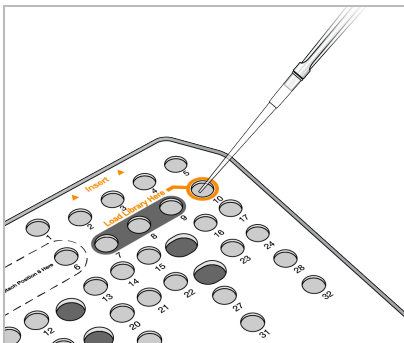
## Chuẩn bị thư viện để giải trình tự

Biến tính và pha loãng thư viện đến mức thể tích nạp 1,3 ml. Trên thực tế, nồng độ nạp có thể thay đổi tùy vào phương pháp chuẩn bị và định lượng thư viện. Việc pha loãng thư viện mẫu phụ thuộc vào độ phức tạp của các nhóm gộp oligonucleotide. Nếu bạn muốn biết hướng dẫn về cách chuẩn bị thư viện mẫu để giải trình tự, bao gồm cả cách pha loãng và gộp nhóm thư viện, hãy xem phần Hướng dẫn sử dụng dành cho bộ kit chuẩn bị thư viện tương ứng. Bạn phải tối ưu hóa mật độ cụm trên NextSeq 550Dx.

## Nạp thư viện vào hộp thuốc thử

1. Dùng giấy lụa ít bụi lau sạch màng nhôm dán kín ngăn chứa số 10 có nhãn **Load Library Here** (Nạp thư viện tại đây).
2. Dùng đầu pipet 1 ml sạch đâm thủng màng nhôm.
3. Nạp 1,3 ml thư viện đã chuẩn bị vào ngăn chứa số 10 có nhãn **Load Library Here** (Nạp thư viện tại đây).  
Tránh chạm vào màng nhôm khi phân phối các thư viện.

Hình 1 Nạp thư viện



## Thiết lập lần chạy giải trình tự

Xem NextSeq 550Dx Instrument Reference Guide (document # 1000000009513) để biết hướng dẫn thiết lập lần chạy hoàn chỉnh.

1. Đăng nhập vào NextSeq 550Dx bằng mật khẩu phần mềm Local Run Manager hoặc Illumina Run Manager.
2. Từ màn hình Home (Chính) của phần mềm NOS, chọn **Sequence** (Trình tự).
3. Chọn lần chạy từ danh sách, sau đó chọn **Next** (Tiếp).

Một loạt màn hình thiết lập chạy mở theo thứ tự sau: Nạp tế bào dòng chảy, Nạp hộp dung dịch đệm, Nạp hộp thuốc thử và Kiểm tra trước khi chạy.

**LƯU Ý** Chỉ có thể truy cập các lần chạy bằng cách sử dụng cùng Run Manager (Trình quản lý lần chạy) được sử dụng khi lập kế hoạch lần chạy. Để biết hướng dẫn về cách đặt Run Manager Software (Phần mềm Quản lý lần chạy), tham khảo Illumina Run Manager for NextSeq 550Dx Software Guide (document # 200025239).

4. Khi màn hình Load Flow Cell (Nạp tế bào dòng chảy) xuất hiện, hãy dọn sạch và sau đó nạp tế bào dòng chảy.
  - Lấy tế bào dòng chảy ra khỏi túi nhôm.
  - Mở hộp nhựa trong suốt dạng vỏ sò và lấy tế bào dòng chảy ra.
  - Làm sạch bề mặt thủy tinh của tế bào dòng chảy bằng khăn lau mịn tẩm cồn. Dùng giấy lụa ít xơ dùng cho phòng thí nghiệm để lau khô phần thủy tinh.
  - Đảm bảo bề mặt thủy tinh của tế bào dòng chảy sạch sẽ. Nếu cần, hãy lặp lại bước làm sạch.
  - Loại bỏ tế bào dòng chảy đã sử dụng từ lần chạy trước.
  - Căn chỉnh tế bào dòng chảy bằng các ghim căn chỉnh và đặt tế bào dòng chảy lên bệ.
5. Chọn **Load** (Nạp).

Cửa sẽ tự động đóng lại, ID tế bào dòng chảy xuất hiện trên màn hình và các cảm biến được kiểm tra.
6. Làm theo lời nhắc của phần mềm để dốc hết hộp đựng thuốc thử đã sử dụng, nạp hộp dung dịch đệm NextSeq 550Dx và nạp hộp thuốc thử NextSeq 550Dx.

Khi bạn đã nạp hộp dung dịch đệm và thuốc thử NextSeq 550Dx, phần mềm sẽ đọc và ghi lại RFID. ID hộp dung dịch đệm và thuốc thử sẽ xuất hiện trên màn hình và các cảm biến được kiểm tra.
7. Khi bước kiểm tra tự động trước lần chạy hoàn tất, hãy chọn **Start** (Bắt đầu). (Không bắt buộc nếu được cấu hình để tự động bắt đầu.)
8. Màn hình Sequencing (Giải trình tự) sẽ mở ra khi lần chạy bắt đầu. Màn hình này cung cấp phần trình bày trực quan về lần chạy đang diễn ra, bao gồm cường độ và điểm chất lượng (điểm Q).

## Kết quả

Real-Time Analysis (RTA) là phần mềm tích hợp thực hiện phân tích hình ảnh và phát hiện base đồng thời chỉ định điểm chất lượng cho từng base cho mỗi chu kỳ giải trình tự. Khi quá trình phân tích chính kết thúc, mô-đun ứng dụng đã chọn sẽ tự động bắt đầu quá trình phân tích phụ. Các quy trình phân tích phụ được mô tả ở đây là dành cho Local Run Manager Mô-đun Germline và Somatic Variant NextSeq 550Dx instrument.

## Tách đoạn

Bước tách đoạn so sánh từng trình tự. Đoạn đọc chỉ thị với các trình tự chỉ thị được chỉ định cho lần chạy. Bước này không xem xét giá trị chất lượng nào.

Đoạn đọc chỉ thị được xác định bằng các bước sau:

- Các mẫu được đánh số bắt đầu từ 1 dựa trên thứ tự mẫu được liệt kê cho lần chạy.
- Mẫu số 0 được dành riêng cho các cụm không được chỉ định cho một mẫu.
- Các cụm được chỉ định cho một mẫu khi trình tự chỉ thị khớp chính xác hoặc khi có tối đa một lần không khớp trên mỗi Đoạn đọc chỉ thị.

## Tạo tệp FASTQ

Sau khi tách đoạn, phần mềm tạo các tệp phân tích trung gian ở định dạng FASTQ, đây là định dạng văn bản được sử dụng để trình bày các trình tự. Tệp FASTQ chứa các đoạn đọc cho từng mẫu và điểm chất lượng liên quan. Các cụm không đi qua bộ lọc đều bị loại trừ.

Mỗi tệp FASTQ chứa các đoạn đọc cho riêng một mẫu và tên của mẫu đó sẽ có trong tên tệp FASTQ. Trong các Mô-đun Germline và Somatic Variant, tám tập tin FASTQ được tạo ra trên mỗi mẫu trên mỗi oligo pool, bốn tập tin từ Đoạn đọc 1 và bốn tập tin từ Đoạn đọc 2. Kết quả đầu ra này lần lượt là 8 và 16 tập tin FASTQ trên mỗi mẫu cho Germline và Somatic. Tệp FASTQ là đầu vào chính để so khớp.

## So khớp

Trong bước so khớp, thuật toán Smith-Waterman theo dải băng sẽ so khớp các cụm từ mỗi mẫu với trình tự amplicon được chỉ định trong tệp phiếu kê khai.

Thuật toán Smith-Waterman theo dải băng thực hiện so khớp trình tự bán toàn cầu để xác định các vùng tương tự giữa hai trình tự. Thay vì so sánh tổng trình tự, thuật toán Smith-Waterman so sánh các đoạn thuộc tất cả các độ dài có thể.

Mỗi đoạn đọc kết đôi được đánh giá về mức độ so khớp với các trình tự đầu dò liên quan cho đoạn đọc đó.

- Đoạn đọc 1 được đánh giá dựa trên phần bổ sung ngược của Oligo dành riêng cho lô-cut xuôi hướng (DLSO).
- Đoạn đọc 2 được đánh giá dựa trên Oligo dành riêng cho lô-cut ngược hướng (ULSO).

- Nếu phần đầu của đoạn đọc khớp với trình tự đầu dò có không quá một lần không khớp thì toàn bộ độ dài của đoạn đọc được so khớp với mục tiêu amplicon cho trình tự đó.
- Nếu phần đầu của đoạn đọc khớp với trình tự đầu dò có không quá ba khác biệt (số lần không khớp hoặc thay đổi do indel dẫn đầu) thì toàn bộ độ dài của đoạn đọc được so khớp với mục tiêu amplicon cho trình tự đó.
- Indel trong DLSO và ULSO không được quan sát, xét công nghệ hóa học của xét nghiệm.

Các chuỗi so khớp được lọc từ kết quả so khớp dựa trên tỷ lệ không khớp trên vùng quan tâm hoặc toàn bộ amplicon, tùy thuộc vào độ dài amplicon. Các chuỗi so khớp đã lọc được ghi trong các tệp so khớp là không so khớp và không được sử dụng trong việc phát hiện biến thể.

## Phát hiện biến thể

Trình phát hiện biến thể Pisces được thiết kế để thực hiện các phát hiện biến thể SNV và indel từ thư viện được chuẩn bị cho thiết bị.

## Báo cáo và tệp đầu ra bổ sung

Các mô-đun phân tích biến thể tạo báo cáo dạng PDF và dạng định giới bằng tab (\*.txt). Các báo cáo này trình bày số liệu như độ sâu giải trình tự và số lượng biến thể. Các mô-đun cũng tạo các tệp đầu ra như tệp Định dạng phát hiện biến thể (VCF) và Định dạng phát hiện biến thể hệ gen (gVCF) cho các ứng dụng phát hiện biến thể.

## Quy trình đối chứng chất lượng

Phần mềm NextSeq 550Dx đánh giá từng lần chạy, mẫu và phát hiện base dựa trên các số liệu đối chứng chất lượng. Mẫu chứng dương và âm cũng được đề xuất trong quá trình chuẩn bị thư viện và cần được đánh giá. Đánh giá các mẫu chứng như sau:

- **Mẫu chứng âm (Chứng không có khuôn) hoặc mẫu chứng âm khác** — Phải tạo ra kết quả dự kiến. Nếu mẫu chứng âm tạo ra kết quả khác với dự kiến, thì có thể đã xảy ra lỗi trong việc theo dõi mẫu, ghi sai các môi lập chỉ thị hoặc tình trạng nhiễm bẩn.
- **Mẫu chứng dương** — Phải tạo ra kết quả dự kiến. Nếu mẫu chứng dương tạo ra kết quả khác với dự kiến, thì có thể đã xảy ra lỗi trong việc theo dõi mẫu hoặc ghi sai các môi lập chỉ thị.

## Đặc điểm hiệu suất

Các đặc điểm hiệu suất của NextSeq 550Dx instrument đã được thiết lập bằng Germline Variant Module và Somatic Variant Module cùng với TruSeq Custom Amplicon Kit Dx và NextSeq 550Dx High Output Reagent Kit v2 (300 chu kỳ) và được xác nhận bằng NextSeq 550Dx High Output Reagent Kit v2.5 (300 chu kỳ). Các nghiên cứu bao gồm Lập chỉ thị mẫu, Tồn đọng mẫu, Đầu vào ADN, Độ nhạy phân tích (Giới hạn mẫu trắng/Giới hạn phát hiện), Độ chính xác, Độ chụm, So sánh phương pháp và Khả năng tái lập.

Các nghiên cứu phân tích sử dụng NextSeq 550Dx High Output Reagent Kit v2.5 (300 chu kỳ) được thiết kế để đánh giá các tuyên bố về hiệu suất đã được thiết lập trước đó với NextSeq 550Dx High Output Reagent Kit v2 (300 chu kỳ). Kết quả chứng minh rằng các bộ kit thuốc thử (v2 và v2.5) có hiệu suất tương đương khi sử dụng TruSeq Custom Amplicon Kit Dx. Xem *tờ hướng dẫn sử dụng TruSeq Custom Amplicon Kit Dx* để biết các đặc điểm hiệu suất liên quan đến các yếu tố trước khi phân tích, chẳng hạn như phương pháp tách chiết hoặc các chất can thiệp.

## Định nghĩa các phép tính được sử dụng trong đặc điểm hiệu suất

- Tỷ lệ phần trăm độ tương hợp dương (PPA) được tính bằng tỷ lệ các lô-cut được phân loại là biến thể theo một phương pháp tham chiếu mà xét nghiệm báo cáo chính xác.
  - $(\text{số lô-cut biến thể được báo cáo chính xác thông qua xét nghiệm}) / (\text{tổng số lô-cut biến thể})$   
 Các lô-cut biến thể được báo cáo thông qua xét nghiệm và phù hợp với phương pháp tham chiếu là dương tính thật (TP). Các lô-cut biến thể được báo cáo là phát hiện tham chiếu hoặc phát hiện biến thể khác nhau thông qua xét nghiệm là âm tính giả (FP).
- Tỷ lệ phần trăm độ tương hợp âm (NPA) được tính bằng tỷ lệ các lô-cut được phân loại là thể đại theo một phương pháp tham chiếu mà xét nghiệm báo cáo chính xác.
  - $(\text{số lô-cut thể đại được báo cáo chính xác thông qua xét nghiệm}) / (\text{tổng số lô-cut thể đại})$   
 Các lô-cut thể đại được báo cáo thông qua xét nghiệm và phù hợp với phương pháp tham chiếu là âm tính thật (TN). Các lô-cut thể đại được báo cáo là biến thể thông qua xét nghiệm là dương tính giả (FP).
- Tỷ lệ phần trăm độ tương hợp tổng thể (OPA) được tính bằng tỷ lệ các lô-cut được báo cáo chính xác thông qua xét nghiệm so với một phương pháp tham chiếu.
  - $((\text{số lô-cut biến thể được báo cáo chính xác thông qua xét nghiệm}) + (\text{số lô-cut thể đại được báo cáo chính xác thông qua xét nghiệm})) / ((\text{tổng số lô-cut biến thể}) + (\text{tổng số lô-cut thể đại}))$
- Các phép tính PPA, NPA và OPA không bao gồm trường hợp không phát hiện (các lô-cut biến thể hoặc tham chiếu không đáp ứng một hoặc nhiều bộ lọc chất lượng).
- Tỷ lệ phát hiện nhiễm sắc thể thường được tính bằng cách lấy tổng số lô-cut đi qua bộ lọc chia cho tổng số vị trí giải trình tự đối với nhiễm sắc thể 1 – 22; loại trừ nhiễm sắc thể X và Y. Số liệu này không xét đến độ tương hợp của kết quả phát hiện với phương pháp tham chiếu.

## Hiệu suất của NextSeq 550Dx High Output Reagent Kit v2 (300 chu kỳ)

### Lập chỉ thị mẫu

Các mẫu chỉ thị mẫu, được thêm vào trong quá trình chuẩn bị thư viện, chỉ định một trình tự duy nhất cho mỗi ADN mẫu. Các trình tự duy nhất này cho phép nhiều mẫu được gộp lại với nhau trong một lần chạy giải trình tự duy nhất. Việc lập chỉ thị mẫu được sử dụng cho cả quy trình công việc dòng mầm và sinh dưỡng. Mục đích của

nghiên cứu này là thiết lập số lượng mẫu tối thiểu (8) và tối đa (96) có thể được xử lý trong một lần chạy giải trình tự duy nhất trên NextSeq 550Dx instrument. Tám mẫu Platinum Genome duy nhất đã được xét nghiệm với 12 tổ hợp mỗi lập chỉ thị khác nhau trên mỗi mẫu. Kết quả mẫu từ bốn lần chạy giải trình tự bằng Germline Variant Module được so sánh với Platinum Genome phiên bản 2016-1.0.

Đối với loạt các lần chạy đầu tiên, 96 thư viện mẫu được lập chỉ thị duy nhất đã được xét nghiệm thông qua một xét nghiệm đại diện được thiết kế để truy vấn nhiều loại gen bao gồm 12.588 base mỗi sợi trên tất cả 23 nhiễm sắc thể ở người. Mục đích là xác minh khả năng của xét nghiệm trong việc thực hiện nhất quán phát hiện định kiểu gen cho một mẫu nhất định trên các tổ hợp mỗi lập chỉ thị khác nhau. Đối với loạt các lần chạy thứ hai, tám thư viện mẫu được lập chỉ thị duy nhất đã được giải trình tự trong hai lần chạy giải trình tự để xác minh số lượng chỉ thị tối thiểu được hỗ trợ.

Đối với các lần chạy 96 chỉ thị, PPA đối với SNV dao động từ 98,7% đến 100%, PPA đối với vùng chèn và vùng xóa là 100% và NPA là 100% cho mỗi tổ hợp 96 chỉ thị. Các lần chạy 8 chỉ thị có giá trị PPA là 100% (SNV, vùng chèn và vùng xóa) và NPA là 100% cho mỗi tổ hợp 8 chỉ thị.

## Tồn đọng mẫu

NextSeq 550Dx instrument cho phép giải trình tự nhiều mẫu cùng với các mẫu chứng trong một lần chạy giải trình tự duy nhất. Một nghiên cứu đã được thực hiện để đánh giá mức độ tồn đọng mẫu trong một lần chạy giải trình tự (trong một lần chạy) và giữa các lần chạy giải trình tự (giữa các lần chạy). Hai mẫu Platinum Genome, một nam và một nữ, đã được xét nghiệm thông qua một xét nghiệm đại diện được thiết kế để truy vấn nhiều gen khác nhau bao gồm 12.588 base (150 amplicon) trên 23 nhiễm sắc thể khác nhau, bao gồm cả hai nhiễm sắc thể giới tính. Các thư viện đã được giải trình tự trên NextSeq 550Dx instrument bằng Germline Variant Module. Sự tồn đọng từ các mẫu nam sang các mẫu nữ đã được quan sát nhờ sự hiện diện của đoạn đọc amplicon nhiễm sắc thể Y trong các mẫu nữ.

Sự tồn đọng trong một lần chạy có thể xảy ra trong quá trình tạo cụm, phát hiện base chu kỳ chỉ thị và tách đoạn mẫu. Để kiểm tra sự tồn đọng mẫu trong một lần chạy giải trình tự, một nhóm gộp thư viện bao gồm 46 mẫu lập cho mỗi mẫu nam và nữ cộng với bốn mẫu chứng không có khuôn đã được giải trình tự một lần trên NextSeq 550Dx instrument. Sự tồn đọng mẫu trong một lần chạy được đánh giá bằng cách so sánh phạm vi amplicon nhiễm sắc thể Y ở mỗi mẫu lập nữ với phạm vi amplicon nhiễm sắc thể Y trung bình ở tất cả các mẫu lập nam trong nhóm gộp. Tỷ lệ tồn đọng mẫu trung bình trong một lần chạy quan sát được là 0,084%.

Để kiểm tra sự tồn đọng mẫu giữa các lần chạy, hai nhóm gộp thư viện đã được chuẩn bị và giải trình tự liên tiếp trên một NextSeq 550Dx instrument. Nhóm gộp đầu tiên bao gồm 46 mẫu lập của mẫu nữ cộng với hai mẫu chứng không có khuôn. Nhóm gộp thứ hai bao gồm 46 mẫu lập của mẫu nam cộng với hai mẫu chứng không có khuôn. Cả hai nhóm gộp đều sử dụng cùng một bộ adapter chỉ thị. Nhóm gộp nữ được giải trình tự trước, rồi đến một lần chạy giải trình tự với nhóm gộp nam, sau đó là một lần chạy giải trình tự lặp lại khác đối với nhóm gộp nữ. Sự tồn đọng mẫu giữa các lần chạy được đánh giá bằng cách so sánh phạm vi amplicon nhiễm sắc thể Y giữa các mẫu lập tương ứng của lần chạy nhóm gộp nữ lặp lại và lần chạy nhóm gộp nam. Tỷ lệ tồn đọng mẫu trung bình giữa các lần chạy quan sát được là 0,0076%.

## Đầu vào ADN

### Máu (Dòng mầm)



Phạm vi đầu vào ADN trong máu để chuẩn bị thư viện TruSeq Custom Amplicon Kit Dx bằng quy trình công việc Germline Variant Module đã được thiết lập cho NextSeq 550Dx instrument. Phạm vi này đã được đánh giá bằng cách thực hiện nghiên cứu pha loãng tuần tự sử dụng 13 mẫu Platinum Genome với một xét nghiệm đại diện được thiết kế để truy vấn nhiều gen bao gồm 12.588 base trên 23 nhiễm sắc thể khác nhau. Thư viện đã được giải trình tự trên hai thiết bị NextSeq 550Dx bằng một lô NextSeq 550Dx High Output Reagent Kit v2 (300 chu kỳ).

Năm mẫu được xét nghiệm lặp lại ở năm mức đầu vào ADN khác nhau, từ 250 ng đến 12 ng (250 ng, 100 ng, 50 ng, 25 ng và 12 ng). Tám mẫu được xét nghiệm như một mẫu lặp duy nhất ở mỗi mức trong số năm mức đầu vào ADN. Để xác định độ chính xác, các kiểu gen mẫu đã được so sánh với Platinum Genome phiên bản 2016-1.0. Kết quả được xác định cho từng mức đầu vào. PPA cho từng loại biến thể (SNV, vùng chèn và vùng xóa) được trình bày trong [Bảng 1](#); NPA được trình bày trong [Bảng 2](#). Tất cả các mức đầu vào đều có độ chính xác tương tự. Đầu vào ADN được đề xuất cho TruSeq Custom Amplicon Kit Dx là 50 ng với 25 ng và 100 ng cung cấp giới hạn dưới và trên để đáp ứng các đặc điểm hiệu suất.

Bảng 1 Kết quả PPA cho mỗi đầu vào ADN theo loại biến thể

Đầu vào ADN (ng)	Loại biến thể	Biến thể theo dự kiến	TP	FN	Biến thể không có phát hiện	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	Vùng chèn	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	Vùng xóa	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

Bảng 2 NPA cho mỗi đầu vào ADN

Đầu vào ADN (ng)	TN	FP	Tham chiếu không có phát hiện	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 (Sinh dưỡng)

Phạm vi đầu vào ADN được cố định bằng formalin và nhúng parafin (FFPE) để chuẩn bị thư viện TruSeq Custom Amplicon Kit Dx bằng quy trình công việc Somatic Variant Module đã được thiết lập cho NextSeq 550Dx instrument. Phạm vi đầu vào ADN đã được đánh giá bằng cách thực hiện nghiên cứu pha loãng tuần tự sử dụng ba mẫu Platinum Genome với một xét nghiệm đại diện được thiết kế để truy vấn nhiều gen bao gồm 12.588 base trên 23 nhiễm sắc thể khác nhau. Các dòng tế bào Platinum Genome GM12878 và GM12877 đã được cố định bằng formalin và nhúng parafin, sau đó ADN được tách chiết. GM12878 được pha loãng với GM12877 sao cho tần số alen biến thể (VAF) của 79 biến thể (55 SNV, 9 vùng chèn và 15 vùng xóa) gần bằng 0,025, 0,05 hoặc 0,10. Ngoài ra, mỗi mẫu có 91 biến thể có tần số biến thể cao hơn, với VAF lên đến 1,0. Các mẫu đã được xử lý lặp lại ở năm mức đầu vào ADN với chu kỳ định lượng delta (dCq) trung bình là 2,1, 3,6, 4,6, 6,0 và 7,8 như được đo bằng TruSeq Custom Amplicon Dx - FFPE QC Kit. Mỗi thư viện đã được giải trình tự trên hai thiết bị NextSeq 550Dx bằng hai lô NextSeq 550Dx High Output Reagent Kit v2 (300 chu kỳ). Để xác định độ chính xác, các kiểu biến thể mẫu đã được so sánh với Platinum Genome phiên bản 2016-1.0. PPA cho từng loại biến thể (SNV, vùng chèn và vùng xóa) được trình bày trong [Bảng 3](#); NPA được trình bày trong [Bảng 4](#). Đầu vào ADN được đề xuất cho các biến thể ở mức VAF 0,05 trở lên là  $dCq \leq 4$  với 4,6 cung cấp giới hạn dưới để đáp ứng các đặc điểm hiệu suất.

Bảng 3 Kết quả PPA cho mỗi đầu vào ADN theo loại biến thể

dCq trung bình	Loại biến thể	Biến thể theo dự kiến	Không có phát hiện theo dự kiến	VAF pha loãng mục tiêu					
				0,025		0,05		0,10	
				Biến thể không có phát hiện	PPA (%)	Biến thể không có phát hiện	PPA (%)	Biến thể không có phát hiện	PPA (%)

2,1	SNV	808	Không áp dụng.	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	Vùng chèn	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	Vùng xóa	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

Bảng 4 NPA cho mỗi đầu vào ADN

dCq trung bình	Thể đại dự kiến	VAF pha loãng mục tiêu					
		0,025		0,05		0,10	
		Tham chiếu không có phát hiện	NPA (%)	Tham chiếu không có phát hiện	NPA (%)	Tham chiếu không có phát hiện	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

## Độ nhạy phân tích (Giới hạn mẫu trắng [LoB] và Giới hạn phát hiện [LoD])

Nghiên cứu này được thực hiện để đánh giá Giới hạn mẫu trắng (LoB) và Giới hạn phát hiện (LoD) cho Somatic Variant Module trên NextSeq 550Dx instrument. Nghiên cứu được tiến hành bằng cách sử dụng một xét nghiệm đại diện được thiết kế để truy vấn nhiều loại gen bao gồm 12.588 base trên 23 nhiễm sắc thể khác nhau. Các dòng tế bào Platinum Genome GM12878 và GM12877 đã được cố định bằng formalin và nhúng parafin, sau đó ADN được tách chiết. GM12878 được pha loãng với GM12877 sao cho tần số biến thể của 74 biến thể (53 SNV,

7 vùng chèn và 14 vùng xóa) là  $0,05 \pm 0,02$ . GM12877 và GM12878 đã pha loãng (GM12878-D) được xét nghiệm trong sáu ngày bắt đầu liên tiếp bằng một thiết bị duy nhất, xen kẽ giữa hai lô NextSeq 550Dx High Output Reagent Kit v2 (300 chu kỳ) cho tổng số sáu lần chạy giải trình tự. Kết quả của xét nghiệm này là 60 mẫu lặp cho mỗi biến thể ở GM12878-D và 72 mẫu lặp cho mỗi tọa độ thể đại tương ứng ở GM12877 đối với mỗi lô thuốc thử. LoB và LoD được tính bằng phương pháp cổ điển nêu trong tài liệu EP17-A2 của CLSI, sử dụng tùy chọn không tham số. LoB và LoD được tính riêng biệt cho SNV, vùng chèn và vùng xóa bằng cách gộp các tần số biến thể cho một loại biến thể nhất định. Sai số Loại I được xác định là 0,01 và sai số Loại II được xác định là 0,05.

Đối với LoB, tần số biến thể đã gộp nhóm được sắp xếp từ thấp nhất đến cao nhất và vị trí xếp hạng thứ 99 cho mỗi lô thuốc thử đối với mỗi loại biến thể đã được tính toán [Bảng 5](#). Somatic Variant Module dùng sử dụng một ngưỡng (LoB hiệu dụng) là VAF 0,026 để xác định khả năng phát hiện biến thể về mặt định tính. LoB được tính toán đã xác minh rằng ngưỡng này dẫn đến sai số Loại I không quá 0,01.

Bảng 5 Giới hạn mẫu trắng

Loại biến thể	Tổng số lần quan sát	LoB lô thuốc thử 1 (%)	LoB lô thuốc thử 2 (%)
SNV	3816	0,77	0,77
Vùng chèn	504	0,56	0,56
Vùng xóa	1008	1,20	1,20

Đối với LoD, tỷ lệ phần trăm tần số đột biến riêng lẻ cho mỗi lô thuốc thử đối với mỗi loại biến thể nằm dưới ngưỡng 0,026 đã được tính toán [Bảng 6](#). Vì tỷ lệ phần trăm này thấp hơn sai số Loại II là 5% (0,05), giá trị trung bình của các tần số biến thể kết hợp được tính toán là LoD ([Bảng 6](#)). LoD cho mỗi loại biến thể được lấy theo giá trị lớn hơn trong hai giá trị được tính toán cho hai lô thuốc thử – 4,97% đối với SNV, 5,12% đối với vùng chèn và 5,26% đối với vùng xóa.

Bảng 6 Giới hạn phát hiện

Lô thuốc thử	Loại biến thể	Tổng số lần quan sát	Số phép đo VAF < 2,6%	% phép đo VAF < 2,6%	Giới hạn phát hiện (%)
1	SNV	3180	53	1,7	4,94
	Vùng chèn	420	6	1,4	5,08
	Vùng xóa	840	7	0,8	5,22
2	SNV	3180	51	1,6	4,97
	Vùng chèn	420	5	1,2	5,12
	Vùng xóa	840	7	0,80	5,26

## Độ chính xác

### Germline

Nghiên cứu sau đây được thực hiện để đánh giá độ chính xác về khả năng phát hiện biến thể của Germline Variant Module trên NextSeq 550Dx instrument bằng NextSeq 550Dx High Output Reagent Kit v2 (300 chu kỳ). 13 mẫu Platinum Genome duy nhất đã được xét nghiệm bằng một xét nghiệm đại diện được thiết kế để truy vấn nhiều loại gen bao gồm 12.588 base (150 amplicon) trên 23 nhiễm sắc thể khác nhau. Tổng cộng có chín lần chạy được thực hiện bằng ba thiết bị giải trình tự, ba lô thuốc thử và giao cho ba người vận hành trong năm ngày bắt đầu. Độ chính xác được xác định đối với các loại đột biến SNV, vùng chèn và vùng xóa bằng cách so sánh kết quả với phương pháp tham chiếu tổng hợp có đặc điểm rõ ràng, Platinum Genome phiên bản 2016-1.0. Các vùng gen tin cậy được xác định dựa trên phương pháp tham chiếu này trừ khi có quy định khác.

Bảng 7 Tóm tắt về độ tương hợp dòng mầm

Tiêu chí	Tổng số lần quan sát <sup>1</sup>	Kết quả theo lần quan sát <sup>2</sup>	Kết quả theo lần chạy <sup>3</sup>
PPA đối với SNV	819	98,7	>99,9
PPA với vùng chèn	819	95,0	98,9
PPA với vùng xóa	819	100	100
NPA	819	100	100
OPA	819	>99,9	>99,9

<sup>1</sup>Được tính bằng số mẫu mỗi lần chạy (91) x số lần chạy (9) = 819.

<sup>2</sup>Giá trị thấp nhất quan sát được theo mẫu lặp của mẫu trên tất cả 9 lần chạy.

<sup>3</sup>Giá trị thấp nhất khi dữ liệu từ mỗi lần chạy được phân tích tổng hợp.

**Bảng 8** bao gồm các dữ liệu nghiên cứu được trình bày với tỷ lệ phần trăm độ tương hợp dương và âm trên cơ sở mỗi mẫu, trong đó các kết quả biến thể được so sánh với Platinum Genome phiên bản 2016-1.0 để tính PPA. Ba loại biến thể (SNV, vùng chèn và vùng xóa) được kết hợp. Vì phương pháp tham chiếu chỉ cung cấp kết quả cho các biến thể nucleotide đơn và vùng chèn/vùng xóa, kết quả base không phải biến thể được so sánh với trình tự tham chiếu hệ gen người bản dựng hg19 để tính NPA.

Bảng 8 Độ tương hợp dòng mầm trên mỗi mẫu

Mẫu	Tỷ lệ phát hiện trung bình	Biến thể dự kiến <sup>1</sup>	TP	FN	Biến thể không có phát hiện	TN	FP	PPA	NPA	OPA
NA12877	>99,9	4788	4788	0	0	756762	0	100	100	100
NA12878	>99,9	8505	8379	1	125	751464	0	>99,9	100	>99,9
NA12879	>99,9	6048	5985	5	58	757701	0	99,9	100	>99,9
NA12880	>99,9	6993	6930	0	63	757638	0	100	100	100
NA12881	>99,9	7875	7811	3	61	751653	0	>99,9	100	>99,9

Mẫu	Tỷ lệ phát hiện trung bình	Biến thể dự kiến <sup>1</sup>	TP	FN	Biến thể không có phát hiện	TN	FP	PPA	NPA	OPA
NA12882	>99,9	6300	6174	3	123	754803	0	>99,9	100	>99,9
NA12883	>99,9	7119	7056	0	63	751905	0	100	100	100
NA12884	>99,9	7182	7119	6	57	754146	0	99,9	100	>99,9
NA12885	>99,9	7686	7560	2	124	754173	0	>99,9	100	>99,9
NA12886	>99,9	7245	7182	7	56	752469	0	99,9	100	>99,9
NA12887	>99,9	7119	7119	0	0	750645	0	100	100	100
NA12888	>99,9	6804	6804	0	0	756065	0	100	100	100
NA12893	>99,9	7434	7371	1	62	750015	0	>99,9	100	>99,9

<sup>1</sup> Tổng số biến thể trong tất cả các mẫu lặp của mẫu trong 9 lần chạy.

**Bảng 9** bao gồm các dữ liệu nghiên cứu được trình bày trên cơ sở mỗi mẫu, trong đó các kết quả biến thể được so sánh với phương pháp tham chiếu tổng hợp có đặc điểm rõ ràng. Khả năng phát hiện được đánh giá cho riêng từng loại biến thể – SNV, vùng chèn và vùng xóa. Loại trừ các vị trí tham chiếu.

**Bảng 9** Độ tương hợp dòng mầm trên mỗi mẫu theo loại biến thể

Mẫu	SNV			Vùng chèn			Vùng xóa		
	Dự kiến	TP	FN	Dự kiến	TP	FN	Dự kiến	TP	FN
NA12877	2331	2331	0	1323	1323	0	1134	1134	0
NA12878	5733	5733	0	1260	1197	1	1512	1449	0
NA12879	3591	3591	0	1323	1260	5	1134	1134	0
NA12880	4221	4221	0	1512	1512	0	1260	1197	0
NA12881	4914	4913	1	1512	1449	2	1449	1449	0
NA12882	3717	3717	0	1386	1323	3	1197	1134	0
NA12883	4284	4284	0	1449	1449	0	1386	1323	0
NA12884	4284	4284	0	1575	1512	6	1323	1323	0
NA12885	4725	4725	0	1575	1512	2	1386	1323	0
NA12886	4347	4347	0	1449	1386	7	1449	1449	0
NA12887	4284	4284	0	1323	1323	0	1512	1512	0
NA12888	4158	4158	0	1449	1449	0	1197	1197	0
NA12893	4599	4599	0	1386	1323	1	1449	1449	0

Các mẫu đã được phân tích thêm để phát hiện vùng chèn và vùng xóa nhỏ (indel). Tóm tắt tổng thể được trình bày trong [Bảng 10](#). Có tổng cộng 71 indel có kích thước từ 1 – 24 bp đối với vùng chèn và 1 – 25 bp đối với vùng xóa.

Bảng 10 Tóm tắt về khả năng phát hiện indel dòng mầm

<b>Biến thể Loại</b>	<b>Biến thể theo dự kiến</b>	<b>TP</b>	<b>FN</b>	<b>Biến thể không có phát hiện</b>	<b>PPA</b>
Vùng chèn	18522	18018	27	477	99,9
Vùng xóa	17388	17073	0	315	100

Xét nghiệm đại diện bao gồm 150 amplicon được thiết kế để bao gồm nhiều hàm lượng hệ gen. Hàm lượng GC của các amplicon dao động từ 0,19 – 0,87. Amplicon cũng có một loạt các trình tự nucleotide đơn (ví dụ: PolyA, PolyT), dinucleotide và trinucleotide lặp lại. Dữ liệu được tổng hợp trên cơ sở mỗi amplicon (**Bảng 11**) để xác định ảnh hưởng của hàm lượng hệ gen đối với tỷ lệ phần trăm phát hiện chính xác. Tỷ lệ phần trăm phát hiện chính xác bao gồm số lần phát hiện biến thể và phát hiện tham chiếu và thấp hơn 100% nếu có số lần phát hiện không chính xác hoặc không có phát hiện.

**Bảng 11** Độ chính xác cấp độ amplicon của dòng mầm

Amplicon	Nhiệm sắc thể	Phần bắt đầu của amplicon	Phần kết thúc của amplicon	Kích thước phân đoạn được phân tích	Số base trong vùng tin cậy	Hàm lượng hệ gen của amplicon	Hàm lượng GC	Số lần phát hiện chính xác	Số lần phát hiện không chính xác	Số lần không có phát hiện	% số lần phát hiện chính xác
1	1	36450499	36450591	93	93	Indel	0,22	76167	0	0	100
2	1	109465122	109465200	79	79	Poly A (5), Poly C (5), indel	0,38	64701	0	0	100
3	1	218353867	218353957	91	91	Indel	0,4	74529	0	0	100
4	1	223906657	223906748	92	92	Indel	0,49	75348	0	0	100
5	1	228526602	228526682	81	81	Poly G (5)	0,69	66339	0	0	100
6	1	236372039	236372108	70	70	Poly T (10), indel	0,39	57330	0	0	100
7	1	247812041	247812128	88	88	Poly A (5), CT(3), TAA(3), indel	0,27	72072	0	0	100
8	2	55862774	55862863	90	90	Indel	0,28	73710	0	0	100
9	2	87003930	87004009	80	80	Indel	0,38	65520	0	0	100
10	2	177016721	177016805	85	81	Không áp dụng	0,65	66339	0	0	100
11	2	186625727	186625801	75	75	Poly A (8)	0,35	61425	0	0	100
12	2	190323504	190323591	88	88	Poly T (5)	0,42	72072	0	0	100
13	2	200796740	200796826	87	87	Poly T (5), indel	0,31	71253	0	0	100
14	2	212245049	212245139	91	91	Poly T (5), Poly A (6), indel	0,3	74529	0	0	100
15	2	228147052	228147144	93	93	Indel	0,43	76167	0	0	100
16	2	235016350	235016422	73	73	Poly T (5), indel	0,42	59787	0	0	100
17	3	4466229	4466321	93	93	AT(3), indel	0,27	74823	0	1344	98,2
18	3	46620561	46620643	83	83	Không áp dụng	0,43	67977	0	0	100
19	3	49851331	49851400	70	70	CT(3), indel	0,49	57330	0	0	100



Amplicon	Nhiễm sắc thể	Phần bắt đầu của amplicon	Phần kết thúc của amplicon	Kích thước phân đoạn được phân tích	Số base trong vùng tin cậy	Hàm lượng hệ gen của amplicon	Hàm lượng GC	Số lần phát hiện chính xác	Số lần phát hiện không chính xác	Số lần không có phát hiện	% số lần phát hiện chính xác
20	3	189713161	189713248	88	88	Poly A (5), Poly T (5), Poly A (9), TG (3)	0,41	72072	0	0	100
21	3	190106030	190106104	75	74	Indel	0,57	60543	0	63	99,9
22	4	2233667	2233744	78	78	Poly A (6)	0,26	63882	0	0	100
23	4	7780541	7780637	97	97	Poly G (6), Poly T (5), Poly A (5)	0,42	79443	0	0	100
24	4	15688604	15688681	78	78	Không áp dụng	0,29	63882	0	0	100
25	4	56236521	56236586	66	62	Poly A (5), indel	0,36	50778	0	0	100
26	4	102839244	102839314	71	69	Poly A (5)	0,46	56511	0	0	100
27	4	164446743	164446804	62	62	Poly A (7), indel	0,27	50778	0	0	100
28	5	1882081	1882158	78	75	Không áp dụng	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	Không áp dụng	0,39	52416	0	0	100
31	5	74077114	74077196	83	83	Poly A (6), indel	0,3	67977	0	0	100
32	5	147475343	147475409	67	67	Poly T (5)	0,37	54873	0	0	100
33	5	149323731	149323821	91	91	CT(4), AG(3)	0,55	74529	0	0	100
34	5	155662213	155662287	75	75	Indel	0,43	61425	0	0	100
35	6	6318713	6318814	102	102	Poly G (6)	0,68	83538	0	0	100
36	6	24949983	24950074	92	92	Indel	0,63	75348	0	0	100
37	6	31084900	31084999	100	94	GCT(5), indel	0,61	76608	0	378	99,5
38	6	32147987	32148084	98	98	Poly T (5), TCT(3), CTT(3)	0,55	80262	0	0	100
39	6	32986864	32986958	95	95	Indel	0,53	77805	0	0	100
40	6	33408498	33408583	86	86	Poly C (6)	0,7	70434	0	0	100
41	6	41647401	41647495	95	94	Poly G (5), indel	0,61	76986	0	0	100
42	6	112435865	112435955	91	91	Poly A (5)	0,44	74529	0	0	100
43	7	22202076	22202148	73	73	Không áp dụng	0,44	59787	0	0	100
44	7	66276100	66276187	88	88	Indel	0,35	72072	0	0	100
45	7	77365735	77365821	87	87	Poly A (7), AG (4)	0,26	71253	0	0	100

Amplicon	Nhiễm sắc thể	Phần bắt đầu của amplicon	Phần kết thúc của amplicon	Kích thước phân đoạn được phân tích	Số base trong vùng tin cậy	Hàm lượng hệ gen của amplicon	Hàm lượng GC	Số lần phát hiện chính xác	Số lần phát hiện không chính xác	Số lần không có phát hiện	% số lần phát hiện chính xác
46	7	110939946	110940030	85	85	Indel	0,38	69615	0	0	100
47	7	128533468	128533557	90	90	Poly G (5), indel	0,62	73710	0	0	100
48	7	149503875	149503965	91	91	Poly G (6), Poly C (6), indel	0,71	74529	0	0	100
49	7	154404519	154404599	81	66	Không áp dụng	0,31	54054	0	0	100
50	7	156476507	156476599	93	93	Indel	0,35	76167	0	0	100
51	8	1817312	1817394	83	83	Không áp dụng	0,42	67977	0	0	100
52	8	24811020	24811109	90	89	Poly G (7), CTC(4), indel	0,61	72171	0	720	99,0
53	8	76518625	76518691	67	67	Indel	0,3	54873	0	0	100
54	9	103054909	103055006	98	98	Poly G (6)	0,67	80262	0	0	100
55	9	105586150	105586214	65	65	Indel	0,32	53235	0	0	100
56	9	107620823	107620918	96	96	Không áp dụng	0,49	78624	0	0	100
57	9	123769149	123769231	83	83	AT(3)	0,37	67977	0	0	100
58	9	138995345	138995441	97	97	Poly C (6), indel	0,68	79443	0	0	100
59	10	5987120	5987198	79	78	Poly G (5), indel	0,47	63882	0	0	100
60	10	11784629	11784726	98	91	GC(3)	0,87	74529	0	0	100
61	10	27317777	27317855	79	79	Poly T (5)	0,3	64701	0	0	100
62	10	33018351	33018440	90	90	Poly A (5), Poly T (5)	0,2	73710	0	0	100
63	10	45084159	45084253	95	95	Indel	0,35	77805	0	0	100
64	10	55892599	55892687	89	88	AC(11), indel	0,42	71747	0	325	99,5
65	10	101611250	101611329	80	80	Không áp dụng	0,49	65520	0	0	100
66	10	118351373	118351453	81	81	Không áp dụng	0,51	66339	0	0	100
67	11	8159816	8159912	97	96	Không áp dụng	0,45	78624	0	0	100
68	11	30177648	30177717	70	70	Indel	0,46	57330	0	0	100
69	11	47470345	47470444	100	100	Không áp dụng	0,65	81900	0	0	100
70	11	59837679	59837740	62	62	Indel	0,37	50778	0	0	100
71	11	64418856	64418957	102	102	Không áp dụng	0,59	83538	0	0	100

Amplicon	Nhiễm sắc thể	Phần bắt đầu của amplicon	Phần kết thúc của amplicon	Kích thước phân đoạn được phân tích	Số base trong vùng tin cậy	Hàm lượng hệ gen của amplicon	Hàm lượng GC	Số lần phát hiện chính xác	Số lần phát hiện không chính xác	Số lần không có phát hiện	% số lần phát hiện chính xác
72	11	93529612	93529684	73	73	Poly A (5)	0,4	59787	0	0	100
73	11	101347052	101347136	85	85	Không áp dụng	0,42	69615	0	0	100
74	11	102477336	102477426	91	91	Poly G (6)	0,55	74529	0	0	100
75	11	118406285	118406369	85	85	Indel	0,53	69615	0	0	100
76	11	120357801	120357885	85	85	Poly A (5), CA (3), indel	0,34	69615	0	0	100
77	11	125769313	125769397	85	85	GA(3)	0,52	69615	0	0	100
78	12	2834770	2834853	84	84	Poly C (5), indel	0,52	68796	0	0	100
79	12	26811004	26811096	93	93	Poly A (7), AC (4)	0,33	76167	0	0	100
80	12	30881766	30881846	81	81	Không áp dụng	0,49	66339	0	0	100
81	12	88474105	88474175	71	71	Poly A (6)	0,35	58149	0	0	100
82	12	120966872	120966966	95	95	Poly G (5)	0,68	77805	0	0	100
83	13	24167504	24167576	73	73	Không áp dụng	0,52	59787	0	0	100
84	13	25816961	25817049	89	88	Poly A (5), Poly T (7), Poly A (7), indel	0,22	72072	0	0	100
85	13	44880112	44880200	89	89	Indel	0,49	72891	0	0	100
86	13	77665218	77665294	77	77	Indel	0,39	63063	0	0	100
87	14	31619327	31619393	67	67	GA(3),TA(3)	0,39	54873	0	0	100
88	14	39517884	39517966	83	83	Không áp dụng	0,25	67977	0	0	100
89	14	46958962	46959034	73	72	Poly T (5), indel	0,19	58642	0	326	99,4
90	14	58050030	58050110	81	81	Indel	0,38	66339	0	0	100
91	14	82390559	82390649	91	91	Indel	0,35	74529	0	0	100
92	14	92549544	92549609	66	66	Poly A (5)	0,41	54054	0	0	100
93	14	102808496	102808589	94	94	Indel	0,62	76986	0	0	100
94	15	43170751	43170848	98	96	Poly C (5)	0,45	78624	0	0	100
95	15	63446149	63446216	68	68	Indel	0,25	55692	0	0	100
96	15	77879807	77879901	95	93	Poly G (5), indel	0,68	76167	0	0	100
97	15	81625334	81625428	95	95	Poly T (6)	0,43	77805	0	0	100
98	15	85438263	85438334	72	71	Indel	0,65	58149	0	0	100

Amplicon	Nhiễm sắc thể	Phần bắt đầu của amplicon	Phần kết thúc của amplicon	Kích thước phân đoạn được phân tích	Số base trong vùng tin cậy	Hàm lượng hệ gen của amplicon	Hàm lượng GC	Số lần phát hiện chính xác	Số lần phát hiện không chính xác	Số lần không có phát hiện	% số lần phát hiện chính xác
99	15	89817413	89817503	91	91	Không áp dụng	0,36	74529	0	0	100
100	15	89864274	89864343	70	70	Indel	0,56	57330	0	0	100
101	16	1894910	1894972	63	63	Không áp dụng	0,27	51597	0	0	100
102	16	28997904	28997998	95	95	Poly C (5)	0,67	77805	0	0	100
103	16	53682908	53682994	87	87	TA(3)	0,41	71253	0	0	100
104	16	57954406	57954509	104	104	Poly C (5)	0,67	85176	0	0	100
105	16	85706375	85706465	91	91	Poly T (5), indel	0,37	74529	0	0	100
106	17	3563920	3564008	89	89	GC(3)	0,64	72891	0	0	100
107	17	3594191	3594277	87	87	Poly C (5), indel	0,67	71247	0	6	100
108	17	3970090	3970180	91	91	Indel	0,46	74529	0	0	100
109	17	16084945	16085037	93	93	Indel	0,26	76167	0	0	100
110	17	33998759	33998849	91	89	Poly T (5)	0,54	72891	0	0	100
111	17	39589691	39589774	84	82	Poly A (13), indel (x2)	0,29	66343	27	788	98,8
112	17	41244394	41244484	91	91	Poly A (5)	0,34	74529	0	0	100
113	17	45438866	45438957	92	92	Poly A (7), AT(3), AT(4), AT(4), indel	0,26	75348	0	0	100
114	17	61502432	61502510	79	79	Indel	0,41	64413	0	288	99,6
115	17	64023582	64023667	86	86	Poly T (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	Không áp dụng	0,37	74529	0	0	100
119	18	9888026	9888094	69	69	Poly A (6), TG(3)	0,43	56511	0	0	100
120	18	38836999	38837073	75	75	Poly A (5), indel	0,37	61425	0	0	100
121	18	47405382	47405462	81	81	CTC(3), indel	0,47	66339	0	0	100
122	18	54815665	54815749	85	85	CT(3), indel	0,45	69615	0	0	100
123	18	59773996	59774060	65	65	Không áp dụng	0,48	53235	0	0	100
124	19	625143	625241	99	99	Không áp dụng	0,59	81081	0	0	100
125	19	18121418	18121491	74	74	Không áp dụng	0,68	60605	1	0	100

Amplicon	Nhiễm sắc thể	Phần bắt đầu của amplicon	Phần kết thúc của amplicon	Kích thước phân đoạn được phân tích	Số base trong vùng tin cậy	Hàm lượng hệ gen của amplicon	Hàm lượng GC	Số lần phát hiện chính xác	Số lần phát hiện không chính xác	Số lần không có phát hiện	% số lần phát hiện chính xác
126	19	18186574	18186643	70	70	Không áp dụng	0,64	57330	0	0	100
127	20	746056	746149	94	94	Không áp dụng	0,61	76986	0	0	100
128	20	10633195	10633276	82	82	AC(3)	0,59	67158	0	0	100
129	20	17705633	17705708	76	76	CT(3)	0,58	62244	0	0	100
130	20	21766821	21766890	70	70	GT(3),TG(4), indel	0,46	57330	0	0	100
131	20	25278421	25278521	101	101	Indel	0,63	82719	0	0	100
132	20	50897302	50897368	67	67	Indel	0,36	54873	0	0	100
133	20	62331904	62331994	91	88	Poly G (6)	0,73	72072	0	0	100
134	20	62690860	62690946	87	87	Indel	0,57	71253	0	0	100
135	21	30300823	30300888	66	66	Indel	0,35	54054	0	0	100
136	21	33694176	33694273	98	98	Poly T (6), CA(3)	0,54	80262	0	0	100
137	21	36710706	36710792	87	87	GT(3), indel	0,39	71253	0	0	100
138	21	46644924	46644992	69	69	Poly A (6), AG(3), indel	0,32	56439	0	72	99,9
139	21	46705575	46705664	90	90	Poly T (5), Poly A (6)	0,5	73710	0	0	100
140	22	25750774	25750873	100	100	Indel	0,63	81900	0	0	100
141	22	32439233	32439329	97	97	Không áp dụng	0,68	79443	0	0	100
142	22	37409844	37409940	97	97	Indel	0,46	79443	0	0	100
143	22	37637596	37637694	99	99	Không áp dụng	0,6	81081	0	0	100
144	22	47081347	47081438	92	92	Indel	0,66	75348	0	0	100
145	X	15870424	15870492	69	69	Poly T (5)	0,26	56511	0	0	100
146	X	135288543	135288611	69	69	Poly C (5)	0,62	56511	0	0	100
147	X	135290777	135290847	71	71	Không áp dụng	0,52	58149	0	0	100
148	Y	2655397	2655461	65	0	Không áp dụng	0,55	0	0	0	Không áp dụng
149	Y	2655519	2655609	91	0	Không áp dụng	0,48	0	0	0	Không áp dụng
150	Y	2655609	2655679	71	0	Poly A (5)	0,37	0	0	0	Không áp dụng

Kết quả giải trình tự cho mẫu NA12878 được so sánh với một kiểu gen có độ tin cậy cao đối với NA12878, do Viện Tiêu chuẩn và Công nghệ Quốc gia (NIST) thiết lập (v.2.19). Trong số 150 amplicon, 92 amplicon nằm trọn trong vùng hệ gen có độ tin cậy cao, 41 amplicon có sự chồng lấn một phần và 17 amplicon không có sự chồng lấn trong trình tự NIST. Kết quả này dẫn đến 10.000 tọa độ trên mỗi mẫu lặp để so sánh. Các phát hiện base không phải biến thể được so sánh với trình tự tham chiếu hệ gen người (bản dựng hg19). Kết quả về độ chính xác được thể hiện trong [Bảng 12](#).

Bảng 12 Độ tương hợp dòng mầm về mẫu NA12878 với cơ sở dữ liệu NIST

Mẫu	Số amplicon	Tỷ lệ phát hiện trung bình	TP	FN	TN	FP	PPA	NPA	OPA
NA12878	133	>99,9	6552	1	610470	0	>99,9	100	>99,9

Dựa trên dữ liệu từ nghiên cứu Dòng mầm gồm chín lần chạy này, NextSeq 550Dx instrument có thể giải trình tự một cách nhất quán:

- Hàm lượng GC  $\geq 19\%$  (tất cả các base được phát hiện ở 819 amplicon được giải trình tự có hàm lượng GC 19% được phát hiện chính xác với tỷ lệ không có phát hiện là 0,6%)
- Hàm lượng GC  $\leq 87\%$  (tất cả các base được phát hiện ở 819 amplicon được giải trình tự có hàm lượng GC 87% được phát hiện chính xác với 0 lần không có phát hiện)
- Độ dài PolyA  $\leq 9$  (tất cả các base được phát hiện ở 819 amplicon được giải trình tự có đoạn lặp lại PolyA gồm chín nucleotide được phát hiện chính xác với 0 lần không có phát hiện)
- Độ dài PolyT  $\leq 10$  (tất cả các base được phát hiện ở 819 amplicon được giải trình tự có đoạn lặp lại PolyT gồm mười nucleotide được phát hiện chính xác với 0 lần không có phát hiện)
- Độ dài PolyG  $\leq 7$  (tất cả các base được phát hiện ở 819 amplicon được giải trình tự có đoạn lặp lại PolyG gồm bảy nucleotide được phát hiện chính xác với tỷ lệ không có phát hiện là 1,0%)
- Độ dài PolyC  $\leq 6$  (tất cả các base được phát hiện ở 2457 amplicon được giải trình tự có đoạn lặp lại PolyC gồm sáu nucleotide được phát hiện chính xác với 0 lần không có phát hiện)
- Độ dài đoạn lặp lại dinucleotide  $\leq 11x$  (tất cả các base được phát hiện ở 819 amplicon được giải trình tự có đoạn lặp lại dinucleotide 11x được phát hiện chính xác với tỷ lệ không có phát hiện là 0,5%)
- Độ dài đoạn lặp lại trinucleotide  $\leq 5x$  (tất cả các base được phát hiện ở 819 amplicon được giải trình tự có đoạn lặp lại trinucleotide 5x được phát hiện chính xác với tỷ lệ không có phát hiện là 0,5%)
- Độ dài vùng chèn  $\leq 24$  (66343/66370 base được phát hiện ở 819 amplicon được giải trình tự có vùng chèn 24 nucleotide được phát hiện chính xác với tỷ lệ không có phát hiện là 1,2%; không xảy ra lần phát hiện không chính xác nào trong vùng có vùng chèn 24 nucleotide)
- Độ dài vùng xóa  $\leq 25$  (tất cả các base được phát hiện ở 2457 amplicon được giải trình tự có vùng xóa 25 nucleotide được phát hiện chính xác với 0 lần không có phát hiện)

## Sinh dưỡng

Nghiên cứu được mô tả ở đây đã được sử dụng để đánh giá độ chính xác về khả năng phát hiện biến thể của Somatic Variant Module trên NextSeq 550Dx instrument bằng NextSeq 550Dx High Output Reagent Kit v2 (300 chu kỳ).

Nghiên cứu này sử dụng một xét nghiệm đại diện được thiết kế để truy vấn nhiều loại gen bao gồm 12.588 base (150 amplicon) trên 23 nhiễm sắc thể khác nhau. ADN Platinum Genome được tách chiết từ các khối đã được xử lý FFPE để tạo ra sáu mẫu duy nhất phục vụ mục đích đánh giá trong nghiên cứu.

ADN mẫu GM12877 được pha loãng với ADN mẫu GM12878 để tạo ra GM12877-D5 và GM12877-D7 dưới dạng một tập hợp các biến thể dị hợp tử duy nhất với tần số biến thể gần bằng 5% và 7%. Tương tự, ADN mẫu GM12878 được pha loãng với ADN mẫu GM12877 để tạo ra GM12878-D5 và GM12878-D7. Mỗi mẫu được xét nghiệm ba lần ngoại trừ các mẫu đã pha loãng được xét nghiệm thành sáu mẫu lặp. Tổng cộng có chín lần chạy được thực hiện bằng ba thiết bị giải trình tự, ba lô thuốc thử và giao cho ba người vận hành trong năm ngày bắt đầu. Độ chính xác được xác định đối với các loại đột biến SNV, vùng chèn và vùng xóa bằng cách so sánh kết quả với phương pháp tham chiếu tổng hợp có đặc điểm rõ ràng, Platinum Genome phiên bản 2016- 1.0. Các vùng gen tin cậy được xác định dựa trên phương pháp tham chiếu này trừ khi có quy định khác.

Bảng 13 Tóm tắt về độ tương hợp sinh dưỡng

Tiêu chí	Tổng số lần quan sát <sup>1</sup>	Kết quả theo lần quan sát <sup>2</sup>	Kết quả theo lần chạy <sup>3</sup>
PPA đối với SNV	378	98,9	99,9
PPA với vùng chèn	378	96,9	99,9
PPA với vùng xóa	378	97,1	99,9
NPA	378	>99,9	>99,9
OPA	378	>99,9	>99,9

<sup>1</sup>Được tính bằng số mẫu mỗi lần chạy (42) x số lần chạy (9) = 378.

<sup>2</sup>Giá trị thấp nhất quan sát được theo mẫu lặp của mẫu trên tất cả 9 lần chạy.

<sup>3</sup>Giá trị thấp nhất khi dữ liệu từ mỗi lần chạy được phân tích tổng hợp.

**Bảng 14** bao gồm các dữ liệu nghiên cứu được trình bày với tỷ lệ phần trăm độ tương hợp dương và âm trên cơ sở mỗi mẫu, trong đó các kết quả biến thể được so sánh với phương pháp tham chiếu tổng hợp có đặc điểm rõ ràng để tính PPA. Ba loại biến thể (SNV, vùng chèn và vùng xóa) được kết hợp. Vì phương pháp tham chiếu chỉ cung cấp kết quả cho các biến thể nucleotide đơn và vùng chèn/vùng xóa, kết quả base không phải biến thể được so sánh với trình tự tham chiếu hệ gen người bản dựng hg19 để tính NPA.

Bảng 14 Độ tương hợp sinh dưỡng trên mỗi mẫu

Mẫu	Tỷ lệ phát hiện trung bình	Dự kiến	TP	FN	Biến thể không có phát hiện	TN	FP	PPA	NPA	OPA
GM12877	98,7	2052	2025	0	27	318682	15	100	>99,9	>99,9
GM12878	98,8	3645	3564	0	81	317645	0	100	100	100
GM12879	99,8	2592	2538	0	54	323614	2	100	>99,9	>99,9
GM12884	99,8	3078	3024	0	54	322038	5	100	>99,9	>99,9
GM12885	99,8	3294	3213	0	81	322121	0	100	100	100
GM12888	99,8	2916	2889	0	27	323048	2	100	>99,9	>99,9
GM12877-D5	99,8	9288	8930	0	358	630621	0	100	100	100
GM12877-D7	99,7	9288	9032	0	256	629719	0	100	100	100
GM12878-D5	99,5	9288	8699	42	547	628582	0	99,5	100	>99,9
GM12878-D7	99,7	9288	9108	0	180	629803	0	100	100	100

**Bảng 15** bao gồm các dữ liệu nghiên cứu được trình bày trên cơ sở mỗi mẫu, trong đó các kết quả biến thể được so sánh với phương pháp tham chiếu tổng hợp có đặc điểm rõ ràng. Khả năng phát hiện được đánh giá cho riêng từng loại biến thể – SNV, vùng chèn và vùng xóa. Loại trừ các vị trí tham chiếu.

Bảng 15 Độ tương hợp sinh dưỡng trên mỗi mẫu theo loại biến thể

Mẫu	SNV			Vùng chèn			Vùng xóa		
	Dự kiến	TP	FN	Dự kiến	TP	FN	Dự kiến	TP	FN
GM12877	999	999	0	567	567	0	486	459	0
GM12878	2457	2457	0	540	513	0	648	594	0
GM12879	1539	1539	0	567	540	0	486	459	0
GM12884	1836	1836	0	675	648	0	567	540	0
GM12885	2025	2025	0	675	648	0	594	540	0
GM12888	1782	1782	0	621	621	0	513	486	0



Mẫu	Dự kiến	SNV			Vùng chèn			Vùng xóa		
		TP	FN		Dự kiến	TP	FN	Dự kiến	TP	FN
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

Mười mẫu đã được phân tích thêm để phát hiện vùng chèn và vùng xóa nhỏ (indel) (Bảng 16). Có tổng cộng 71 indel có kích thước từ 1 – 24 bp đối với vùng chèn và 1 – 25 bp đối với vùng xóa.

Bảng 16 Tóm tắt về khả năng phát hiện indel sinh dưỡng

Loại biến thể	Biến thể theo dự kiến	TP	FN	Biến thể không có phát hiện	PPA
Vùng chèn	10773	10282	9	482	99,2
Vùng xóa	11502	10667	5	830	>99,9

150 amplicon được thiết kế để bao gồm nhiều hàm lượng hệ gen. Hàm lượng GC của các amplicon dao động từ 0,19 – 0,87%. Amplicon cũng có một loạt các trình tự nucleotide đơn (ví dụ: PolyA, PolyT), dinucleotide và trinucleotide lặp lại. Dữ liệu được tổng hợp trên cơ sở mỗi amplicon (Bảng 17) để xác định ảnh hưởng của hàm lượng hệ gen đối với tỷ lệ phần trăm phát hiện chính xác. Tỷ lệ phần trăm phát hiện chính xác bao gồm số lần phát hiện biến thể và phát hiện tham chiếu và thấp hơn 100% nếu có số lần phát hiện không chính xác hoặc không có phát hiện.

Bảng 17 Độ chính xác cấp độ amplicon của sinh dưỡng

Amplicon	Nhiễm sắc thể	Phần bắt đầu của amplicon	Phần kết thúc của amplicon	Kích thước phân đoạn được phân tích	Số base trong vùng tin cậy	Hàm lượng hệ gen của amplicon	Hàm lượng GC	Số lần phát hiện chính xác	Số lần phát hiện không chính xác	Số lần không có phát hiện	% số lần phát hiện chính xác
1	1	36450499	36450591	93	93	Indel	0,22	35066	0	88	99,7
2	1	109465122	109465200	79	79	Poly A (5), Poly C (5), indel	0,38	29827	0	35	99,9
3	1	218353867	218353957	91	91	Indel	0,4	34202	0	283	99,2
4	1	223906657	223906748	92	92	Indel	0,49	34613	0	163	99,5
5	1	228526602	228526682	81	81	Poly G (5)	0,69	30571	0	47	99,8
6	1	236372039	236372108	70	70	Poly T (10), indel	0,39	26452	0	8	100,0
7	1	247812041	247812128	88	88	Poly A (5), CT(3), TAA(3), indel	0,27	33148	0	116	99,7
8	2	55862774	55862863	90	90	Indel	0,28	33928	0	92	99,7
9	2	87003930	87004009	80	80	Indel	0,38	30218	0	22	99,9
10	2	177016721	177016805	85	81	Không áp dụng	0,65	30616	0	2	>99,9
11	2	186625727	186625801	75	75	Poly A (8)	0,35	28017	0	499	98,3
12	2	190323504	190323591	88	88	Poly T (5)	0,42	33207	0	57	99,8
13	2	200796740	200796826	87	87	Poly T (5), indel	0,31	32524	9	718	97,8
14	2	212245049	212245139	91	91	Poly T (5), Poly A (6), indel	0,3	33972	0	456	98,7
15	2	228147052	228147144	93	93	Không áp dụng	0,43	35051	0	103	99,7
16	2	235016350	235016422	73	73	Poly T (5), indel	0,42	27459	0	136	99,5
17	3	4466229	4466321	93	93	AT(3), indel	0,27	34534	0	620	98,2
18	3	46620561	46620643	83	83	Không áp dụng	0,43	31339	0	44	99,9
19	3	49851331	49851400	70	70	CT(3), indel	0,49	26373	0	87	99,7

Amplicon	Nhiễm sắc thể	Phần bắt đầu của amplicon	Phần kết thúc của amplicon	Kích thước phân đoạn được phân tích	Số base trong vùng tin cậy	Hàm lượng hệ gen của amplicon	Hàm lượng GC	Số lần phát hiện chính xác	Số lần phát hiện không chính xác	Số lần không có phát hiện	% số lần phát hiện chính xác
20	3	189713161	189713248	88	88	Poly A (5), Poly T (5), Poly A (9), TG (3)	0,41	32829	0	857	97,5
21	3	190106030	190106104	75	74	Indel	0,57	27925	0	47	99,8
22	4	2233667	2233744	78	78	Poly A (6)	0,26	29327	4	162	99,4
23	4	7780541	7780637	97	97	Poly G (6), Poly T (5), Poly A (5)	0,42	36585	0	117	99,7
24	4	15688604	15688681	78	78	Không áp dụng	0,29	29427	0	57	99,8
25	4	56236521	56236586	66	62	Poly A (5), indel	0,36	23356	5	75	99,7
26	4	102839244	102839314	71	69	Poly A (5)	0,46	25942	0	140	99,5
27	4	164446743	164446804	62	62	Poly A (7), indel	0,27	22944	0	560	97,6
28	5	1882081	1882158	78	75	Không áp dụng	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	Không áp dụng	0,39	24120	0	72	99,7
31	5	74077114	74077196	83	83	Poly A (6), indel	0,3	31297	0	77	99,8
32	5	147475343	147475409	67	67	Poly T (5)	0,37	25277	0	55	99,8
33	5	149323731	149323821	91	91	CT(4), AG(3)	0,55	34308	0	90	99,7
34	5	155662213	155662287	75	75	Indel	0,43	28266	0	163	99,4
35	6	6318713	6318814	102	102	Poly G (6)	0,68	38489	0	67	99,8
36	6	24949983	24950074	92	92	Indel	0,63	34730	0	46	99,9
37	6	31084900	31084999	100	94	GCT(5), indel	0,61	35057	0	483	98,6
38	6	32147987	32148084	98	98	Poly T (5), TCT(3), CTT(3)	0,55	36647	0	406	98,9
39	6	32986864	32986958	95	95	Indel	0,53	35681	0	238	99,3
40	6	33408498	33408583	86	86	Poly C (6)	0,7	32438	0	70	99,8
41	6	41647401	41647495	95	94	Poly G (5), indel	0,61	35441	0	91	99,7
42	6	112435865	112435955	91	91	Poly A (5)	0,44	34354	0	44	99,9
43	7	22202076	22202148	73	73	Không áp dụng	0,44	27575	0	28	99,9
44	7	66276100	66276187	88	88	Indel	0,35	33060	0	213	99,4
45	7	77365735	77365821	87	87	Poly A (7), AG(4)	0,26	32423	0	489	98,5

Amplicon	Nhiễm sắc thể	Phần bắt đầu của amplicon	Phần kết thúc của amplicon	Kích thước phân đoạn được phân tích	Số base trong vùng tin cậy	Hàm lượng hệ gen của amplicon	Hàm lượng GC	Số lần phát hiện chính xác	Số lần phát hiện không chính xác	Số lần không có phát hiện	% số lần phát hiện chính xác
46	7	110939946	110940030	85	85	Indel	0,38	32074	0	56	99,8
47	7	128533468	128533557	90	90	Poly G (5), indel	0,62	33791	0	281	99,2
48	7	149503875	149503965	91	91	Poly G (6), Poly C (6), indel	0,71	34316	0	82	99,8
49	7	154404519	154404599	81	66	Không áp dụng	0,31	24901	0	47	99,8
50	7	156476507	156476599	93	93	Indel	0,35	35067	0	87	99,8
51	8	1817312	1817394	83	83	Không áp dụng	0,42	31365	0	9	>99,9
52	8	24811020	24811109	90	89	Poly G (7), CTC(4), indel	0,61	32781	0	890	97,4
53	8	76518625	76518691	67	67	Indel	0,3	25228	0	146	99,4
54	9	103054909	103055006	98	98	Poly G (6)	0,67	36968	0	76	99,8
55	9	105586150	105586214	65	65	Indel	0,32	24472	0	100	99,6
56	9	107620823	107620918	96	96	Không áp dụng	0,49	36203	0	85	99,8
57	9	123769149	123769231	83	83	AT(3)	0,37	31329	0	45	99,9
58	9	138995345	138995441	97	97	Poly C (6), indel	0,68	36472	0	201	99,5
59	10	5987120	5987198	79	78	Poly G (5), indel	0,47	29473	0	11	>99,9
60	10	11784629	11784726	98	91	GC(3)	0,87	34188	0	213	99,4
61	10	27317777	27317855	79	79	Poly T (5)	0,3	29843	0	19	99,9
62	10	33018351	33018440	90	90	Poly A (5), Poly T (5)	0,2	33968	0	68	99,8
63	10	45084159	45084253	95	95	Indel	0,35	35829	0	81	99,8
64	10	55892599	55892687	89	88	AC(11), indel	0,42	32098	88	2048	93,8
65	10	101611250	101611329	80	80	Không áp dụng	0,49	30217	0	28	99,9
66	10	118351373	118351453	81	81	Không áp dụng	0,51	30531	0	96	99,7
67	11	8159816	8159912	97	96	Không áp dụng	0,45	36105	0	192	99,5
68	11	30177648	30177717	70	70	Indel	0,46	26318	0	153	99,4
69	11	47470345	47470444	100	100	Không áp dụng	0,65	37785	0	24	99,9
70	11	59837679	59837740	62	62	Indel	0,37	23368	0	68	99,7
71	11	64418856	64418957	102	102	Không áp dụng	0,59	38546	0	10	>99,9
72	11	93529612	93529684	73	73	Poly A (5)	0,4	27516	0	78	99,7

Amplicon	Nhiễm sắc thể	Phần bắt đầu của amplicon	Phần kết thúc của amplicon	Kích thước phân đoạn được phân tích	Số base trong vùng tin cậy	Hàm lượng hệ gen của amplicon	Hàm lượng GC	Số lần phát hiện chính xác	Số lần phát hiện không chính xác	Số lần không có phát hiện	% số lần phát hiện chính xác
73	11	101347052	101347136	85	85	Không áp dụng	0,42	32083	0	48	99,9
74	11	102477336	102477426	91	91	Poly G (6)	0,55	34047	0	369	98,9
75	11	118406285	118406369	85	85	Indel	0,53	32065	0	74	99,8
76	11	120357801	120357885	85	85	Poly A (5), CA(3), indel	0,34	32083	0	47	99,9
77	11	125769313	125769397	85	85	GA(3)	0,52	32103	0	27	99,9
78	12	2834770	2834853	84	84	Poly C (5), indel	0,52	31645	16	525	98,3
79	12	26811004	26811096	93	93	Poly A (7), AC(4)	0,33	34824	0	330	99,1
80	12	30881766	30881846	81	81	Không áp dụng	0,49	30497	0	121	99,6
81	12	88474105	88474175	71	71	Poly A (6)	0,35	26773	0	65	99,8
82	12	120966872	120966966	95	95	Poly G (5)	0,68	35830	9	72	99,8
83	13	24167504	24167576	73	73	Không áp dụng	0,52	27498	0	114	99,6
84	13	25816961	25817049	89	88	Poly A (5), Poly T (7), Poly A (7), indel	0,22	32824	0	566	98,3
85	13	44880112	44880200	89	89	Indel	0,49	33574	0	77	99,8
86	13	77665218	77665294	77	77	Indel	0,39	29075	0	31	99,9
87	14	31619327	31619393	67	67	GA(3),TA(3)	0,39	25313	0	13	99,9
88	14	39517884	39517966	83	83	Không áp dụng	0,25	31360	0	22	99,9
89	14	46958962	46959034	73	72	Poly T (5), indel	0,19	26499	0	717	97,4
90	14	58050030	58050110	81	81	Indel	0,38	30494	0	133	99,6
91	14	82390559	82390649	91	91	Indel	0,35	34313	0	86	99,7
92	14	92549544	92549609	66	66	Poly A (5)	0,41	24555	0	1527	94,1
93	14	102808496	102808589	94	94	Indel	0,62	35472	0	69	99,8
94	15	43170751	43170848	98	96	Poly C (5)	0,45	36264	0	24	99,9
95	15	63446149	63446216	68	68	Indel	0,25	25667	0	37	99,9
96	15	77879807	77879901	95	93	Poly G (5), indel	0,68	34745	0	432	98,8
97	15	81625334	81625428	95	95	Poly T (6)	0,43	35870	0	40	99,9
98	15	85438263	85438334	72	71	Indel	0,65	26762	0	76	99,7
99	15	89817413	89817503	91	91	Không áp dụng	0,36	34286	0	112	99,7

Amplicon	Nhiễm sắc thể	Phần bắt đầu của amplicon	Phần kết thúc của amplicon	Kích thước phân đoạn được phân tích	Số base trong vùng tin cậy	Hàm lượng hệ gen của amplicon	Hàm lượng GC	Số lần phát hiện chính xác	Số lần phát hiện không chính xác	Số lần không có phát hiện	% số lần phát hiện chính xác
100	15	89864274	89864343	70	70	Indel	0,56	26449	0	11	>99,9
101	16	1894910	1894972	63	63	Không áp dụng	0,27	23809	0	5	>99,9
102	16	28997904	28997998	95	95	Poly C (5)	0,67	35860	0	50	99,9
103	16	53682908	53682994	87	87	TA(3)	0,41	32835	0	60	99,8
104	16	57954406	57954509	104	104	Poly C (5)	0,67	39177	0	144	99,6
105	16	85706375	85706465	91	91	Poly T (5), indel	0,37	34075	0	323	99,1
106	17	3563920	3564008	89	89	GC(3)	0,64	33632	0	11	>99,9
107	17	3594191	3594277	87	87	Poly C (5), indel	0,67	32752	0	134	99,6
108	17	3970090	3970180	91	91	Indel	0,46	34343	0	82	99,8
109	17	16084945	16085037	93	93	Indel	0,26	35077	0	78	99,8
110	17	33998759	33998849	91	89	Poly T (5)	0,54	33553	0	89	99,7
111	17	39589691	39589774	84	82	Poly A (13), indel (x2)	0,29	30554	53	2296	92,9
112	17	41244394	41244484	91	91	Poly A (5)	0,34	34360	0	38	99,9
113	17	45438866	45438957	92	92	Poly A (7), AT(3), AT(4), AT(4), indel	0,26	34367	0	418	98,8
114	17	61502432	61502510	79	79	Indel	0,41	29751	0	119	99,6
115	17	64023582	64023667	86	86	Poly T (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	Không áp dụng	0,37	34386	0	12	>99,9
119	18	9888026	9888094	69	69	Poly A (6), TG(3)	0,43	25692	0	399	98,5
120	18	38836999	38837073	75	75	Poly A (5), indel	0,37	27923	0	893	96,9
121	18	47405382	47405462	81	81	CTC(3), indel	0,47	30598	0	20	99,9
122	18	54815665	54815749	85	85	CT(3), indel	0,45	31969	0	161	99,5
123	18	59773996	59774060	65	65	Không áp dụng	0,48	24531	0	48	99,8
124	19	625143	625241	99	99	Không áp dụng	0,59	37298	0	124	99,7
125	19	18121418	18121491	74	74	Không áp dụng	0,68	27881	0	109	99,6
126	19	18186574	18186643	70	70	Không áp dụng	0,64	26442	0	26	99,9

Amplicon	Nhiễm sắc thể	Phần bắt đầu của amplicon	Phần kết thúc của amplicon	Kích thước phân đoạn được phân tích	Số base trong vùng tin cậy	Hàm lượng hệ gen của amplicon	Hàm lượng GC	Số lần phát hiện chính xác	Số lần phát hiện không chính xác	Số lần không có phát hiện	% số lần phát hiện chính xác
127	20	746056	746149	94	94	Không áp dụng	0,61	35501	0	31	99,9
128	20	10633195	10633276	82	82	AC(3)	0,59	30951	0	72	99,8
129	20	17705633	17705708	76	76	CT(3)	0,58	28686	0	42	99,9
130	20	21766821	21766890	70	70	GT(3),TG(4), indel	0,46	26372	0	88	99,7
131	20	25278421	25278521	101	101	Indel	0,63	38159	0	20	99,9
132	20	50897302	50897368	67	67	Indel	0,36	25188	0	544	97,9
133	20	62331904	62331994	91	88	Poly G (6)	0,73	32969	0	309	99,1
134	20	62690860	62690946	87	87	Indel	0,57	32818	0	77	99,8
135	21	30300823	30300888	66	66	Indel	0,35	24758	9	181	99,2
136	21	33694176	33694273	98	98	Poly T (6), CA(3)	0,54	36902	0	160	99,6
137	21	36710706	36710792	87	87	GT(3), indel	0,39	32841	0	48	99,9
138	21	46644924	46644992	69	69	Poly A (6), AG(3), indel	0,32	25939	0	280	98,9
139	21	46705575	46705664	90	90	Poly T (5), Poly A (6)	0,5	33942	0	78	99,8
140	22	25750774	25750873	100	100	Indel	0,63	37733	0	86	99,8
141	22	32439233	32439329	97	97	Không áp dụng	0,68	36617	0	49	99,9
142	22	37409844	37409940	97	97	Indel	0,46	36525	0	162	99,6
143	22	37637596	37637694	99	99	Không áp dụng	0,6	37398	0	24	99,9
144	22	47081347	47081438	92	92	Indel	0,66	34754	0	22	99,9
145	X	15870424	15870492	69	69	Poly T (5)	0,26	26046	0	36	99,9
146	X	135288543	135288611	69	69	Poly C (5)	0,62	26019	0	63	99,8
147	X	135290777	135290847	71	71	Không áp dụng	0,52	26780	0	58	99,8
148	Y	2655397	2655461	65	0	Không áp dụng	0,55	0	0	0	Không áp dụng
149	Y	2655519	2655609	91	0	Không áp dụng	0,48	0	0	0	Không áp dụng
150	Y	2655609	2655679	71	0	Poly A (5)	0,37	0	0	0	Không áp dụng

Kết quả giải trình tự cho mẫu GM12878 đã được so sánh với kiểu gen có độ tin cậy cao đối với NA12878, do Viện Tiêu chuẩn và Công nghệ Quốc gia (NIST) thiết lập (v.2.19). Trong số 150 amplicon, 92 amplicon nằm trọn trong vùng hệ gen có độ tin cậy cao, 41 amplicon có sự chồng lấn một phần và 17 amplicon không có sự chồng lấn trong trình tự NIST. Kết quả này dẫn đến 10.000 tọa độ trên mỗi mẫu lặp để so sánh. Các phát hiện base không phải biến thể được so sánh với trình tự tham chiếu hệ gen người (bản dựng hg19). Kết quả về độ chính xác được thể hiện trong [Bảng 18](#).

**Bảng 18** Độ tương hợp sinh dưỡng về mẫu GM12878 với cơ sở dữ liệu NIST

Mẫu	Số amplicon	Tỷ lệ phát hiện trung bình	TP	FN	TN	FP	PPA	NPA	OPA
GM12878	133	98,8	2808	0	258488	0	100	100	100

Dựa trên dữ liệu từ nghiên cứu Sinh dưỡng gồm chín lần chạy này, NextSeq 550Dx instrument có thể giải trình tự một cách nhất quán:

- Hàm lượng GC  $\geq 19\%$  (tất cả các base được phát hiện ở 378 amplicon được giải trình tự có hàm lượng GC 19% được phát hiện chính xác với tỷ lệ không có phát hiện là 2,6%)
- Hàm lượng GC  $\leq 87\%$  (tất cả các base được phát hiện ở 378 amplicon được giải trình tự có hàm lượng GC 87% được phát hiện chính xác với tỷ lệ không có phát hiện là 0,6%)
- Độ dài PolyA  $\leq 9$  (tất cả các base được phát hiện ở 378 amplicon được giải trình tự có đoạn lặp lại PolyA gồm chín nucleotide được phát hiện chính xác với tỷ lệ không có phát hiện là 2,5%)
- Độ dài PolyT  $\leq 10$  (tất cả các base được phát hiện ở 378 amplicon được giải trình tự có đoạn lặp lại PolyT gồm mười nucleotide được phát hiện chính xác với tỷ lệ không có phát hiện dưới 0,1%)
- Độ dài PolyG  $\leq 6$  (tất cả các base được phát hiện ở 2268 amplicon được giải trình tự có đoạn lặp lại PolyG gồm sáu nucleotide được phát hiện chính xác với tỷ lệ không có phát hiện là 0,5%)
- Độ dài PolyC  $\leq 6$  (tất cả các base được phát hiện ở 756 amplicon được giải trình tự có đoạn lặp lại PolyC gồm sáu nucleotide được phát hiện chính xác với tỷ lệ không có phát hiện là 0,4%)
- Độ dài đoạn lặp lại dinucleotide  $\leq 4x$  (tất cả các base được phát hiện ở 1890 amplicon được giải trình tự có đoạn lặp lại dinucleotide 4x được phát hiện chính xác với tỷ lệ không có phát hiện là 0,9%)
- Độ dài đoạn lặp lại trinucleotide  $\leq 5x$  (tất cả các base được phát hiện ở 378 amplicon được giải trình tự có đoạn lặp lại trinucleotide 5x được phát hiện chính xác với tỷ lệ không có phát hiện là 1,4%)
- Độ dài vùng chèn  $\leq 23$  (tất cả các base được phát hiện ở 378 amplicon được giải trình tự có vùng chèn 23 nucleotide được phát hiện chính xác với tỷ lệ không có phát hiện là 0,8%)
- Độ dài vùng xóa  $\leq 25$  (tất cả các base được phát hiện ở 1134 amplicon được giải trình tự có vùng xóa 25 nucleotide được phát hiện chính xác với tỷ lệ không có phát hiện là 0,7%)



## Độ chụm

Độ chụm của NextSeq 550Dx instrument được xác định bằng cách xét nghiệm 13 mẫu Platinum Genome duy nhất, sử dụng ba thiết bị, ba lô thuốc thử và giao cho ba người vận hành để tạo ra chín lần chạy giải trình tự trong năm ngày bắt đầu. Xét nghiệm đại diện, mẫu và phương pháp tham chiếu giống như thông tin mô tả cho nghiên cứu về độ chính xác dòng mầm. Các yếu tố đóng góp cho độ chụm được xác định bằng cách phân tích thành phần phương sai, sử dụng VAF làm biến phản hồi và tính toán độ lệch chuẩn ở cấp độ thành phần cho thiết bị, lô thuốc thử, người vận hành và ngày bắt đầu (Bảng 19). Tổng số lần quan sát được sử dụng trong quá trình phân tích về độ biến thiên của từng thành phần gồm thiết bị, người vận hành hoặc lô thuốc thử lần lượt là 699, 176 và 235 đối với SNV, vùng chèn và vùng xóa.

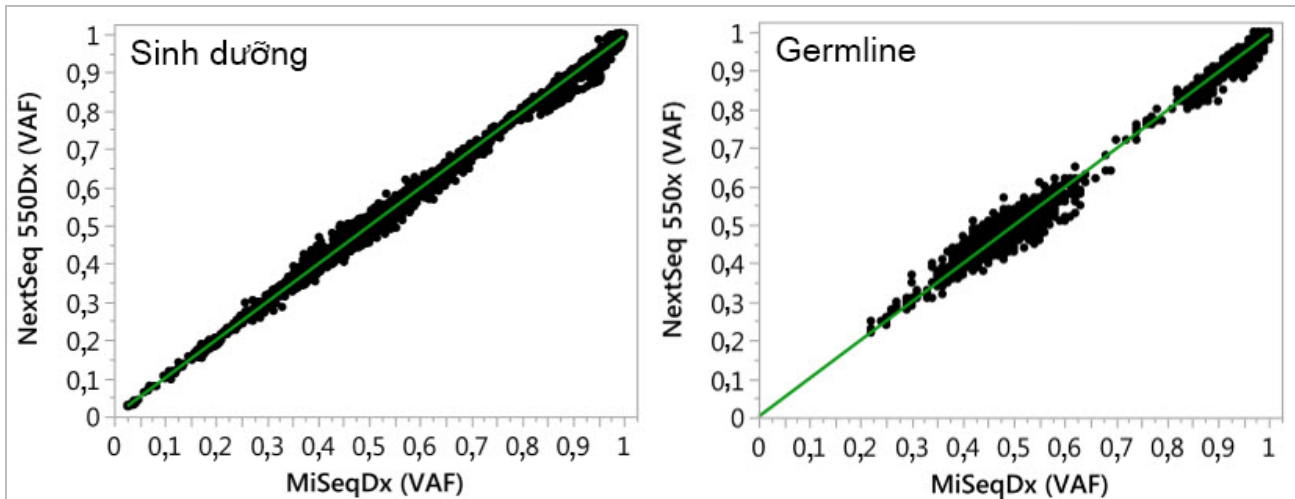
Bảng 19 Kết quả về độ chụm cho NextSeq 550Dx instrument (Độ lệch chuẩn/SD)

Thành phần	Loại biến thể	SD thành phần		Tổng SD	
		Tối đa	Trung bình	Tối đa	Trung bình
Lô	SNV	0,0076	0,0002	0,0833	0,0154
	Vùng chèn	0,0104	0,0000	0,0410	0,0157
	Vùng xóa	0,0046	0,0005	0,0560	0,0187
Thiết bị	SNV	0,0114	0,0003	0,0840	0,0153
	Vùng chèn	0,0138	0,0009	0,0407	0,0161
	Vùng xóa	0,0079	0,0008	0,0549	0,0187
Người vận hành	SNV	0,0226	0,0008	0,0841	0,0155
	Vùng chèn	0,0344	0,0010	0,0417	0,0164
	Vùng xóa	0,0083	0,0013	0,0547	0,0187
Ngày	SNV	0,0277	0,0012	0,0825	0,0160
	Vùng chèn	0,0235	0,0012	0,0409	0,0169
	Vùng xóa	0,0271	0,0014	0,0548	0,0188

## So sánh phương pháp (Nền tảng giải trình tự)

Các mẫu máu toàn phần và FFPE đã được đánh giá trên NextSeq 550Dx instrument và thiết bị MiSeqDx, sử dụng quy trình công việc dòng mầm và sinh dưỡng của TruSeq Custom Amplicon Kit Dx. Độ tương hợp về tần số biến thể đối với các mẫu máu và FFPE được đánh giá bằng nhiều xét nghiệm đại diện. Hình 2 minh họa mối tương quan về VAF giữa hai thiết bị cho một xét nghiệm đại diện và Bảng 20 tóm tắt mối tương quan này theo bảng xét nghiệm. Dựa trên mối tương quan chặt chẽ giữa thiết bị MiSeqDx và NextSeq 550Dx instrument, các đặc điểm hiệu suất liên quan đến các yếu tố trước khi phân tích (ví dụ: phương pháp tách chiết hoặc chất can thiệp) được xác định là có thể áp dụng cho cả hai thiết bị. Xem tờ hướng dẫn sử dụng TruSeq Custom Amplicon Kit Dx để biết thêm chi tiết.

Hình 2 Mối tương quan về VAF của thiết bị MiSeqDx với thiết bị NextSeq 550Dx cho mẫu FFPE (trái) và mẫu máu (phải) khi sử dụng xét nghiệm 1



Bảng 20 Kết quả so sánh phương pháp bằng cách sử dụng các mẫu máu và FFPE duy nhất

Nguồn gADN	Xét nghiệm (Bảng Oligo)	Số mẫu lập sinh học (Mẫu)	Số mẫu lập kỹ thuật (mỗi Mẫu)	Số lần quan sát (Số biến thể)	Hệ số góc	Hệ số chặn	Mối tương quan ( $R^2$ )
Máu	Xét nghiệm 1	45	2	8369 <sup>1</sup>	0,992	0,002	0,995 <sup>2</sup>
Máu	Xét nghiệm 2	45	2	5457	0,995	0,005	0,981
FFPE	Xét nghiệm 1	46	2	8319	0,993	0,000	0,997 <sup>2</sup>
FFPE	Xét nghiệm 3	40	1	280	0,969	0,015	0,978

<sup>1</sup>Hai điểm dữ liệu đã được loại bỏ dựa trên giới hạn đã nêu đối với mô-đun Germline Variant.

<sup>2</sup>Hệ số xác định cho biểu đồ VAF như được minh họa trong Hình 2.

## Khả năng tái lập

Khả năng tái lập của NextSeq 550Dx instrument được đánh giá bằng cách sử dụng các mẫu Platinum Genome với một xét nghiệm đại diện được thiết kế để truy vấn nhiều loại gen bao gồm 12.588 base trên 23 nhiễm sắc thể khác nhau bằng 150 amplicon. Xét nghiệm dòng mầm bao gồm bảy mẫu lập của 13 mẫu; xét nghiệm sinh dưỡng bao gồm sáu mẫu lập của bảy mẫu ở các mức VAF khác nhau. TruSeq Custom Amplicon Kit Dx được sử dụng để chuẩn bị các mẫu.

Xét nghiệm được thực hiện tại ba địa điểm bên ngoài, sử dụng một lô NextSeq 550Dx High Output Reagent Kit v2 (300 chu kỳ). Mỗi địa điểm sử dụng một NextSeq 550Dx instrument duy nhất. Hai người vận hành thực hiện xét nghiệm tại mỗi địa điểm. Mỗi người vận hành thực hiện xét nghiệm trong ba ngày bắt đầu không liên tiếp đối với mỗi loại mẫu với tổng số 36 lần chạy tại ba địa điểm. Xét nghiệm này dẫn đến 18 lần chạy cho mỗi quy trình công việc dòng mầm và sinh dưỡng.

## Germline

Các biến thể dòng mầm có mức VAF  $\geq 0,2$  được báo cáo là dương tính (biến thể). Đối với các biến thể dòng mầm dương tính dự kiến, dữ liệu được đánh giá về tỷ lệ không có phát hiện và tỷ lệ phát hiện dương tính chính xác ở mỗi loại biến thể (SNV, vùng chèn, vùng xóa). **Bảng 21** tóm tắt các tỷ lệ quan sát được, cùng với mức tin cậy 95% dưới và trên (LCL/UCL) được tính bằng phương pháp Điểm Wilson, cho từng loại biến thể.

Bảng 21 Các lần quan sát phát hiện dòng mầm cho kết quả dương tính dự kiến theo loại biến thể

Loại biến thể	Không có phát hiện			Phát hiện dương tính chính xác				
	Quan sát được	Tổng	Tỷ lệ phần trăm	Quan sát được	Tổng	Tỷ lệ phần trăm	LCL 95%	UCL 95%
SNV	16	110.376	0,014	110.349	110.360	99,99	99,98	99,99
Vùng chèn	1026	37.044	2,77	36.018	36.018	100	99,99	100,00
Vùng xóa	648	34.776	1,86	34.128	34.128	100	99,99	100,00

Các biến thể dòng mầm có mức VAF  $< 0,2$  được báo cáo là âm tính (thể đại). Đối với các vị trí dòng mầm âm tính dự kiến, dữ liệu đã được đánh giá về tỷ lệ không có phát hiện và tỷ lệ phát hiện thể đại chính xác. **Bảng 22** tóm tắt các tỷ lệ quan sát được, cùng với mức tin cậy 95% dưới và trên (LCL/UCL) được tính bằng phương pháp Điểm Wilson.

Bảng 22 Quan sát phát hiện dòng mầm cho kết quả âm tính dự kiến

Loại biến thể	Không có phát hiện			Phát hiện âm tính chính xác				
	Quan sát được	Tổng	Tỷ lệ phần trăm	Quan sát được	Tổng	Tỷ lệ phần trăm	LCL 95%	UCL 95%
Thể đại	4883	19.600.182	0,025	19.595.299	19.595.299	100	100,00	100,00

Các biến thể dòng mầm có mức VAF  $\geq 0,2$  và  $< 0,7$  được phát hiện là dị hợp tử dương tính đối với biến thể và các biến thể có mức VAF  $\geq 0,7$  được phát hiện là đồng hợp tử dương tính đối với biến thể. Các mẫu dòng mầm có biến thể dị hợp tử được sử dụng để xác định xem liệu sự biến thiên vốn có của xét nghiệm có ảnh hưởng đến việc phát hiện kiểu gen hay không. Cx đã được xác định cho cả hai ngưỡng (0,2 đối với kiểu gen dị hợp tử và 0,7 đối với kiểu gen đồng hợp tử), trong đó x là tỷ lệ các xét nghiệm lặp lại vượt quá ngưỡng. Đối với ngưỡng dưới là VAF 0,2, Cx  $\geq 99,999\%$ , cho thấy rằng  $\geq 99,999\%$  biến thể dị hợp tử sẽ được phát hiện là dị hợp tử. Đối với ngưỡng trên là VAF 0,7, Cx  $\leq 0,001\%$ , do đó cho thấy rằng  $\leq 0,001\%$  biến thể dị hợp tử sẽ được phát hiện là đồng hợp tử. **Bảng 23** tóm tắt kết quả theo loại biến thể.

Các biến thể dòng mầm có mức VAF  $\geq 0,2$  và  $< 0,7$  được phát hiện là dị hợp tử dương tính đối với biến thể và các biến thể có mức VAF  $\geq 0,7$  được phát hiện là đồng hợp tử dương tính đối với biến thể. Các mẫu dòng mầm có biến thể dị hợp tử được sử dụng để xác định xem liệu sự biến thiên vốn có của xét nghiệm có ảnh hưởng đến việc phát hiện kiểu gen hay không. Cx đã được xác định cho cả hai ngưỡng (0,2 đối với kiểu gen dị hợp tử và 0,7 đối với kiểu gen đồng hợp tử), trong đó x là tỷ lệ các xét nghiệm lặp lại vượt quá ngưỡng. Đối với ngưỡng dưới là

VAF 0,2, Cx  $\geq 99,999\%$ , do đó cho thấy rằng  $\geq 99,999\%$  biến thể dị hợp tử sẽ được phát hiện là dị hợp tử. Đối với ngưỡng trên là VAF 0,7, Cx  $\leq 0,001\%$ , cho thấy rằng  $\leq 0,001\%$  biến thể dị hợp tử sẽ được phát hiện là đồng hợp tử. **Bảng 23** tóm tắt kết quả theo loại biến thể.

Bảng 23 Giá trị Cx dòng mầm cho các biến thể dị hợp tử

Loại biến thể	Ngưỡng ở VAF 0,2	Ngưỡng ở VAF 0,7
	$\geq C99,999\%$	$\leq C0,001\%$
SNV	94/94	94/94
Vùng chèn	24/24	24/24
Vùng xóa	35/35	35/35
Tổng	153	153

## Sinh dưỡng

Các biến thể sinh dưỡng có mức VAF  $\geq 0,026$  được báo cáo là dương tính (biến thể). Các lần quan sát có mức VAF  $\geq 0,01$  và  $< 0,026$  được coi là không rõ ràng khi xét các mục đích của quá trình phân tích này (không dương tính cũng không âm tính, được gán cờ là tần số biến thể thấp). Để đánh giá hiệu suất, các kết quả được tính theo ba cách:

- Trường hợp tốt nhất: Mọi kết quả không rõ ràng đều được coi là phát hiện dương tính chính xác (tương hợp với kết quả dự kiến)
- Trường hợp xấu nhất: Mọi kết quả không rõ ràng đều được coi là phát hiện không chính xác (không tương hợp với kết quả dự kiến)
- Trường hợp loại trừ: Mọi kết quả không rõ ràng đều bị loại trừ khỏi quá trình phân tích

Ba bảng gồm **Bảng 24**, **Bảng 25** và **Bảng 26** lần lượt tóm tắt kết quả phát hiện cho trường hợp tốt nhất, trường hợp xấu nhất và trường hợp loại trừ, cùng với mức tin cậy 95% dưới và trên (LCL/UCL) được tính bằng phương pháp Điểm Wilson.

Bảng 24 Các lần quan sát phát hiện sinh dưỡng cho kết quả dương tính dự kiến theo loại biến thể (Trường hợp tốt nhất)

Loại biến thể	Phát hiện dương tính chính xác				
	Quan sát được	Tổng	Tỷ lệ phần trăm	LCL 95%	UCL 95%
SNV	54.346	54.346	100	99,99	100,00
Vùng chèn	18.036	18.036	100	99,98	100,00
Vùng xóa	18.381	18.381	100	99,98	100,00

Bảng 25 Các lần quan sát phát hiện sinh dưỡng cho kết quả dương tính dự kiến theo loại biến thể (Trường hợp xấu nhất)

Loại biến thể	Phát hiện dương tính chính xác				
	Quan sát được	Tổng	Tỷ lệ phần trăm	LCL 95%	UCL 95%
SNV	54.346	54.346	100	99,99	100,00
Vùng chèn	18.000	18.036	99,8	99,72	99,86
Vùng xóa	18.381	18.381	100	99,98	100,00

Bảng 26 Các lần quan sát phát hiện sinh dưỡng cho kết quả dương tính dự kiến theo loại biến thể (Loại bỏ phát hiện không rõ ràng)

Loại biến thể	Phát hiện dương tính chính xác				
	Quan sát được	Tổng	Tỷ lệ phần trăm	LCL 95%	UCL 95%
SNV	54.346	54.346	100	99,99	100,00
Vùng chèn	18.000	18.000	100	99,98	100,00
Vùng xóa	18.381	18.381	100	99,98	100,00

Các biến thể sinh dưỡng có mức VAF < 0,01 được báo cáo là phát hiện âm tính (thể đại). Đối với các vị trí sinh dưỡng âm tính dự kiến, dữ liệu đã được đánh giá về tỷ lệ không có phát hiện và tỷ lệ phát hiện thể đại chính xác. Các phát hiện thể đại chính xác được xác định bằng cách loại trừ các trường hợp không có phát hiện và lấy tổng trừ đi các phát hiện quan sát được rơi vào vùng không rõ ràng (mức VAF  $\geq$  0,01 và < 0,026) cũng như các phát hiện không chính xác vượt quá ngưỡng (mức VAF  $\geq$  0,026). [Bảng 27](#) tóm tắt các kết quả quan sát được, tổng số kết quả và kết quả theo tỷ lệ phần trăm đối với các vị trí sinh dưỡng âm tính về tỷ lệ không có phát hiện và tỷ lệ phát hiện thể đại chính xác cùng với mức tin cậy 95% dưới và trên (LCL/UCL) được tính bằng phương pháp Điểm Wilson.

Bảng 27 Các lần quan sát phát hiện sinh dưỡng cho kết quả âm tính dự kiến

Biến thể Loại	Không có phát hiện			Phát hiện chính xác						
	Quan sát được	Tổng	Tỷ lệ phần trăm	Không rõ ràng	Không chính xác	Chính xác	Tổng	Tỷ lệ phần trăm	LCL 95%	UCL 95%
Thể đại	36.326	8.909.676	0,408	2254	121	8.870.975	8.873.350	99,97	99,972	99,974

Các mẫu sinh dưỡng ở các mức VAF khác nhau đối với cùng một biến thể đã được đánh giá để xác định C95 của xét nghiệm (trong mỗi loại biến thể). Để đánh giá độ biến thiên gần ngưỡng xét nghiệm, các mẫu có mức VAF dự kiến trong khoảng từ 0,02 đến 0,07 đã được sử dụng. C95 được xác định cho từng biến thể, với C95 cao nhất cho từng loại biến thể được báo cáo trong [Bảng 28](#).

Bảng 28 Tóm tắt giá trị C95 sinh dưỡng

Loại biến thể	N	C95
SNV	74	0,0613
Vùng chèn	24	0,0573
Vùng xóa	33	0,0575

## Hiệu suất của NextSeq 550Dx High Output Reagent Kit v2.5 (300 chu kỳ)

### Tổng quan

NextSeq 550Dx được hỗ trợ bởi hai bộ kit thuốc thử: NextSeq 550Dx High Output Reagent Kit v2 (300 chu kỳ) và NextSeq 550Dx High Output Reagent Kit v2.5 (300 chu kỳ). Để chứng minh rằng NextSeq 550Dx High Output Reagent Kit v2.5 (300 chu kỳ) có thể đáp ứng các yêu cầu về hiệu suất phân tích đã được xác minh và xác nhận với NextSeq 550Dx High Output Reagent Kit v2 (300 chu kỳ), các nghiên cứu đã được thực hiện với NextSeq 550Dx High Output Reagent Kit v2.5 (300 chu kỳ). Hai quá trình chuẩn bị thư viện bằng TruSeq Custom Amplicon Kit Dx đã được thực hiện, một với quy trình công việc Dòng mầm và một với quy trình công việc Sinh dưỡng. Các thư viện từ mỗi quy trình công việc đã được xét nghiệm với ba lô NextSeq 550Dx High Output Reagent Kit v2.5 (300 chu kỳ), sử dụng ba NextSeq 550Dx instrument. Ngoài ra, việc xét nghiệm cho mỗi quy trình công việc bao gồm một lần chạy với NextSeq 550Dx High Output Reagent Kit v2 (300 chu kỳ).

### Độ nhạy phân tích (Giới hạn mẫu trắng [LoB] và Giới hạn phát hiện [LoD])

Việc xác minh với NextSeq 550Dx High Output Reagent Kit v2 (300 chu kỳ) đã chứng minh rằng NextSeq 550Dx instrument có thể phát hiện các biến thể ở VAF 0,05 với sai số Loại II  $\leq 0,05$  và ngưỡng VAF 0,026 được sử dụng bởi Somatic Variant Module (LoB hiệu dụng) chấp nhận sai số Loại I  $\leq 0,01$ . Dựa trên những tuyên bố này, dự kiến rằng một biến thể ở VAF 0,05 sẽ lớn hơn hoặc bằng ngưỡng VAF 0,026 95% thời gian và vị trí thể đại sẽ nhỏ hơn ngưỡng VAF 0,026 99% thời gian. Để đảm bảo rằng những tuyên bố này phù hợp với NextSeq 550Dx High Output Reagent Kit v2.5 (300 chu kỳ), các phép đo lặp lại đã được tiến hành trên NextSeq 550Dx instrument với các mẫu thể đại (mẫu LoB) và với các mẫu có chứa các biến thể ở VAF 0,05 (Mẫu LoD) bằng NextSeq 550Dx High Output Reagent Kit v2.5 (300 chu kỳ). Tỷ lệ phát hiện nằm trên và dưới ngưỡng 0,026 sau đó được so sánh với các tuyên bố đã được thiết lập với NextSeq 550Dx High Output Reagent Kit v2 (300 chu kỳ).

Quá trình xét nghiệm bao gồm hai mẫu LoD, mỗi mẫu có một tập hợp biến thể duy nhất được nhắm mục tiêu đến VAF 0,05 và các mẫu LoB tương ứng là thể đại cho các biến thể được nhắm mục tiêu. Để chuẩn bị thư viện, các mẫu LoD và LoB lần lượt được xử lý theo dạng tám và bảy mẫu lặp, sử dụng TruSeq Custom Amplicon Kit Dx. Ban đầu, các thư viện được giải trình tự bằng NextSeq 550Dx High Output Reagent Kit v2 (300 chu kỳ) nhằm xác định các biến thể/tọa độ hệ gen để đánh giá LoB/LoD với NextSeq 550Dx High Output Reagent Kit v2.5 (300 chu kỳ). Tất cả các biến thể có VAF trung bình từ 0,045 – 0,055 (Biến thể LoD) dựa trên kết quả từ NextSeq 550Dx High Output Reagent Kit v2 (300 chu kỳ) được sử dụng cho quá trình phân tích LoD (N = 51 biến thể). Đối với quá trình phân tích LoB, 51 tọa độ hệ gen tương ứng đã được đánh giá.

Để đánh giá NextSeq 550Dx High Output Reagent Kit v2.5 (300 chu kỳ), các thư viện được giải trình tự trong ba lần chạy trong ba ngày liên tiếp bằng cùng một thiết bị và lô bộ kit thuốc thử. Xét nghiệm này lên đến 24 mẫu lặp cho mỗi biến thể trong số 51 biến thể LoD và 21 mẫu lặp cho mỗi vị trí trong số các vị trí thể đại tương ứng. Tỷ lệ phát hiện thể đại với VAF < 0,026 được cung cấp trong [Bảng 29](#). Tỷ lệ phát hiện biến thể LoD với VAF lớn hơn hoặc bằng 0,026 được cung cấp trong [Bảng 30](#).

Bảng 29 Tỷ lệ phát hiện < 0,026 cho các vị trí thể đại (Đánh giá tuyên bố về LoB)

Biến thể Loại	Vị trí được đánh giá	Tổng số lần quan sát	Số phép đo VAF $\geq 2,6\%$	Tỷ lệ < 2,6%	Tỷ lệ 95% khoảng tin cậy
SNV	32	672	0	1	0,994 – 1
Vùng chèn	11	231	0	1	0,984 – 1
Vùng xóa	8	168	0	1	0,978 – 1

Bảng 30 Tỷ lệ phát hiện với VAF  $\geq 0,026$  cho biến thể LoD (Đánh giá tuyên bố về LoD)

Biến thể Loại	Vị trí được đánh giá	Tổng số lần quan sát	Số phép đo VAF < 2,6%	Số phép đo VAF $\geq 2,6\%$	Tỷ lệ $\geq 2,6\%$	Tỷ lệ 95% khoảng tin cậy
SNV	32	768	1	767	0,999	0,993 – 1
Vùng chèn	11	264	3	261	0,989	0,967 – 0,996
Vùng xóa	8	192	2	190	0,99	0,963 – 0,997

## Độ chính xác

### Germline

Nghiên cứu sau đây được thực hiện để đánh giá độ chính xác về khả năng phát hiện biến thể với Germline Variant Module bằng NextSeq 550Dx High Output Reagent Kit v2.5 (300 chu kỳ). 12 mẫu Platinum Genome duy nhất đã được xét nghiệm bằng một xét nghiệm đại diện. Tổng cộng 11 lần chạy đã được thực hiện bằng ba thiết bị NextSeq 550Dx và ba NextSeq 550Dx High Output Reagent Kit v2.5 (300 chu kỳ).

Độ chính xác được xác định đối với các loại đột biến SNV, vùng chèn và vùng xóa bằng cách so sánh kết quả với phương pháp tham chiếu tổng hợp có đặc điểm rõ ràng, Platinum Genome phiên bản 2016-1.0. Kết quả về độ chính xác từ một lần chạy giải trình tự đơn lẻ bằng NextSeq 550Dx High Output Reagent Kit v2 (300 chu kỳ) được cung cấp cho mục đích tham khảo. Thông tin tóm tắt kết quả có trong [Bảng 31](#).

Bảng 31 Tóm tắt về độ tương hợp dòng mầm

Tiêu chí	Tổng số lần quan sát (v2.5) <sup>1</sup>	Kết quả theo lần quan sát (v2.5) <sup>2</sup>	Kết quả theo lần quan sát (v2) <sup>3</sup>	Kết quả theo lần chạy (v2.5) <sup>4</sup>	Kết quả theo lần chạy (v2) <sup>4</sup>
PPA đối với SNV	1056	98,7	98,7	>99,9	>99,9
PPA với vùng chèn	1056	100	100	100	100
PPA với vùng xóa	1056	95,2	95,2	>99,9	>99,9
NPA	1056	100	100	100	100
OPA	1056	>99,9	>99,9	>99,9	>99,9

<sup>1</sup>Được tính bằng số mẫu mỗi lần chạy x số lần chạy (96 mẫu mỗi lần chạy x 11 lần chạy = 1056 lần quan sát).

<sup>2</sup>Giá trị thấp nhất quan sát được theo mẫu lặp của mẫu trên tất cả các lần chạy (dựa trên 11 lần chạy đối với NextSeq 550Dx High Output Reagent Kit v2.5).

<sup>3</sup>Giá trị thấp nhất quan sát được theo mẫu lặp của mẫu trong 1 lần chạy (tổng số 96 lần quan sát).

<sup>4</sup>Giá trị thấp nhất khi dữ liệu từ mỗi lần chạy được phân tích tổng hợp.

## Sinh dưỡng

Nghiên cứu sau đây đã được thực hiện để đánh giá độ chính xác về khả năng phát hiện biến thể của Somatic Variant Module trên NextSeq 550Dx instrument bằng NextSeq 550Dx High Output Reagent Kit v2.5 (300 chu kỳ). Mười mẫu FFPE Platinum Genome (hai mẫu có biến thể được pha loãng xuống mức VAF 0,05) đã được xét nghiệm thông qua một xét nghiệm đại diện. Tổng cộng 11 lần chạy đã được thực hiện bằng ba thiết bị NextSeq 550Dx và ba lô NextSeq 550Dx High Output Reagent Kit v2.5 (300 chu kỳ).

Độ chính xác được xác định đối với các loại đột biến SNV, vùng chèn và vùng xóa bằng cách so sánh kết quả với phương pháp tham chiếu tổng hợp có đặc điểm rõ ràng, Platinum Genome phiên bản 2016-1.0. Kết quả về độ chính xác từ một lần chạy giải trình tự đơn lẻ bằng NextSeq 550Dx High Output Reagent Kit v2 (300 chu kỳ) được cung cấp cho mục đích tham khảo. Thông tin tóm tắt kết quả có trong [Bảng 32](#).

Bảng 32 Tóm tắt về độ tương hợp sinh dưỡng

Tiêu chí	Tổng số lần quan sát (v2.5) <sup>1</sup>	Kết quả theo lần quan sát (v2.5) <sup>2</sup>	Kết quả theo lần quan sát (v2) <sup>3</sup>	Kết quả theo lần chạy (v2.5) <sup>4</sup>	Kết quả theo lần chạy (v2) <sup>4</sup>
PPA đối với SNV	528	100	100	100	100
PPA với vùng chèn	528	96,9	96,9	> 99,9	> 99,9



PPA với vùng xóa	528	100	100	100	100
NPA	528	> 99,9	> 99,9	> 99,9	> 99,9
OPA	528	> 99,9	> 99,9	> 99,9	> 99,9

<sup>1</sup>Được tính bằng số mẫu mỗi lần chạy x số lần chạy (48 mẫu mỗi lần chạy x 11 lần chạy = 528 lần quan sát).

<sup>2</sup>Giá trị thấp nhất quan sát được theo mẫu lặp của mẫu trên tất cả các lần chạy (dựa trên 11 lần chạy đối với NextSeq 550Dx High Output Reagent Kit v2.5).

<sup>3</sup>Giá trị thấp nhất quan sát được theo mẫu lặp của mẫu trong 1 lần chạy (tổng số 96 lần quan sát).

<sup>4</sup>Giá trị thấp nhất khi dữ liệu từ mỗi lần chạy được phân tích tổng hợp.

## Độ chụm

### Germline

Độ chụm của NextSeq 550Dx High Output Reagent Kit v2.5 (300 chu kỳ) với Germline Variant Module đã được đánh giá bằng cách sử dụng các mẫu Platinum Genome và một xét nghiệm đại diện. Xét nghiệm bao gồm một lần chuẩn bị thư viện duy nhất bằng TruSeq Custom Amplicon Kit Dx và gồm 12 mẫu được xử lý với tám mẫu lặp ở mỗi mẫu. Các thư viện được giải trình tự với ba lô NextSeq 550Dx High Output Reagent Kit v2.5 (300 chu kỳ) và ba NextSeq 550Dx instruments cho tổng số chín lần chạy giải trình tự.

Các mẫu có biến thể dị hợp tử được sử dụng để xác định xem liệu sự biến thiên vốn có của xét nghiệm có ảnh hưởng đến việc phát hiện kiểu gen hay không (N = 153 biến thể dị hợp tử duy nhất). Cx đã được xác định cho cả hai ngưỡng của Germline Variant Module (0,2 đối với kiểu gen dị hợp tử và 0,7 đối với kiểu gen đồng hợp tử), trong đó x là tỷ lệ các xét nghiệm lặp lại vượt quá ngưỡng. Với ngưỡng dưới là VAF 0,2, biến thể có Cx tối thiểu đối với NextSeq 550Dx Reagent Kit v2.5 (300 chu kỳ) > 99,9%, cho thấy rằng > 99,9% biến thể dị hợp tử sẽ được phát hiện là dị hợp tử. Với ngưỡng trên là VAF 0,7, biến thể có Cx tối đa đối với NextSeq 550Dx Reagent Kit v2.5 (300 chu kỳ) < 1,5%, cho thấy rằng ≤ 1,5% biến thể dị hợp tử sẽ được phát hiện là đồng hợp tử. [Bảng 33](#) tóm tắt kết quả theo loại biến thể. Giá trị Cx từ một lần chạy giải trình tự đơn lẻ bằng NextSeq 550Dx High Output Reagent Kit v2 (300 chu kỳ) được cung cấp cho mục đích tham khảo.

Bảng 33 Giá trị Cx dòng mầm cho các biến thể dị hợp tử

Loại biến thể	N	Ngưỡng ở VAF 0,2		Ngưỡng ở VAF 0,7	
		Cx tối thiểu (v2.5) <sup>1</sup>	Cx tối thiểu (v2) <sup>2</sup>	Cx tối đa (v2.5) <sup>1</sup>	Cx tối đa (v2) <sup>2</sup>
SNV	94	>99,9%	>99,9%	1,5%	1,0%
Vùng chèn	24	100%	100%	0%	<0,1%
Vùng xóa	35	100%	>99,9%	<0,1%	<0,1%

<sup>1</sup>Giá trị Cx dựa trên ước tính Tổng độ lệch chuẩn từ phân tích thành phần phương sai.

<sup>2</sup>Giá trị Cx dựa trên Độ lệch chuẩn mẫu.

## Sinh dưỡng

Độ chụm của NextSeq 550Dx High Output Reagent Kit v2.5 (300 chu kỳ) với Somatic Variant Module đã được đánh giá bằng cách sử dụng các mẫu FFPE Platinum Genome và một xét nghiệm đại diện. Xét nghiệm bao gồm một lần chuẩn bị thư viện duy nhất bằng TruSeq Custom Amplicon Kit Dx và gồm hai mẫu lặp ở mỗi mẫu. Các thư viện được giải trình tự bằng ba lô NextSeq 550Dx High Output Reagent Kit v2.5 (300 chu kỳ) và ba NextSeq 550Dx instruments cho tổng số chín lần chạy giải trình tự.

Các biến thể sinh dưỡng có mức VAF dự kiến  $\leq 0,10$  ( $N = 131$  biến thể duy nhất) được sử dụng để đánh giá độ biến thiên của thiết bị gần ngưỡng VAF của Somatic Variant Module (các biến thể sinh dưỡng có mức VAF  $\geq 0,026$  được phát hiện là dương tính với biến thể này). Giá trị C95 được xác định cho từng biến thể sinh dưỡng. Giá trị C95 đại diện cho VAF mà tại đó xác suất lớn hơn ngưỡng VAF của Somatic Variant Module là 95%. Giá trị C95 cao nhất theo loại biến thể được báo cáo trong [Bảng 34](#). Kết quả C95 từ một lần chạy giải trình tự đơn lẻ bằng NextSeq 550Dx High Output Reagent Kit v2 (300 chu kỳ) được cung cấp cho mục đích tham khảo.

Bảng 34 Tóm tắt giá trị C95 sinh dưỡng

Loại biến thể	Số biến thể được đánh giá	C95 (v2.5) <sup>1</sup>	C95 (v2) <sup>2</sup>
SNV	74	0,064	0,063
Vùng chèn	24	0,062	0,061
Vùng xóa	33	0,060	0,060

<sup>1</sup>Giá trị C95 dựa trên ước tính Tổng độ lệch chuẩn từ phân tích thành phần phương sai.

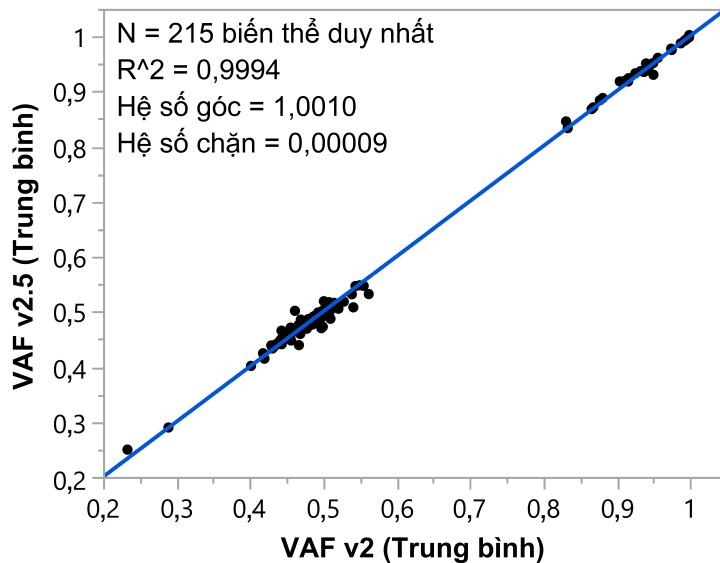
<sup>2</sup>Giá trị C95 dựa trên Độ lệch chuẩn mẫu.

## So sánh phương pháp (Bộ kit thuốc thử)

### Germline

VAF trung bình từ 215 biến thể duy nhất đã được đánh giá trên NextSeq 550Dx High Output Reagent Kit v2 (300 chu kỳ) và NextSeq 550Dx High Output Reagent Kit v2.5 (300 chu kỳ), sử dụng kết quả được tạo từ Germline Variant Module. Giá trị trung bình của VAF được tính toán từ 11 lần chạy giải trình tự (v2.5) và một lần chạy giải trình tự (v2). Ít nhất tám mẫu lặp đã được sử dụng để tính giá trị trung bình cho mỗi biến thể. [Hình 3](#) minh họa mối tương quan về VAF giữa hai bộ kit thuốc thử. Dựa trên mối tương quan tuyến tính chặt chẽ về VAF và sự tương đồng về kết quả giữa các bộ kit thuốc thử, các đặc điểm hiệu suất ban đầu từng được xác minh và xác nhận với NextSeq 550Dx High Output Reagent Kit v2 (300 chu kỳ) qua Germline Variant Module được xác định là có thể áp dụng cho NextSeq 550Dx High Output Reagent Kit v2.5 (300 chu kỳ).

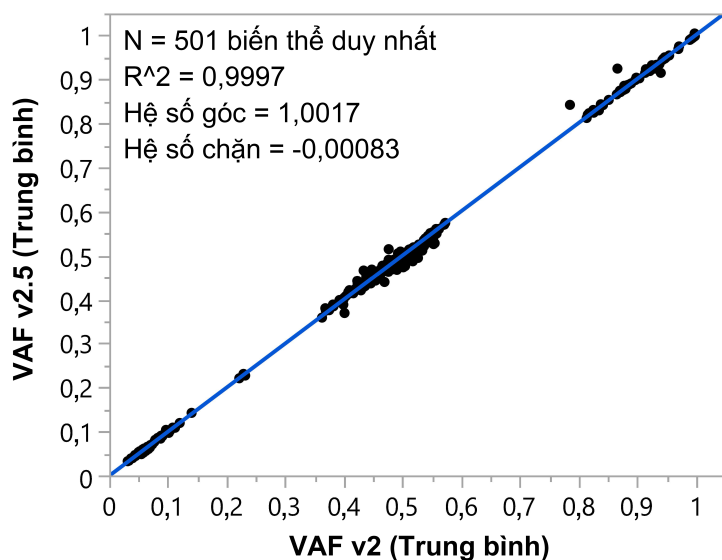
Hình 3 Mối tương quan về Tần số alen biến thể (VAF) khi sử dụng Germline Variant Module giữa NextSeq 550Dx High Output Reagent Kit v2 (300 chu kỳ) và NextSeq 550Dx High Output Reagent Kit v2.5 (300 chu kỳ).



## Sinh dưỡng

VAF trung bình cho 501 biến thể duy nhất đã được đánh giá trên NextSeq 550Dx High Output Reagent Kit v2 (300 chu kỳ) và NextSeq 550Dx High Output Reagent Kit v2.5 (300 chu kỳ), sử dụng kết quả được tạo từ Somatic Variant Module. Giá trị trung bình của VAF được tính toán từ 11 lần chạy giải trình tự (v2.5) và một lần chạy giải trình tự (v2). Ít nhất ba mẫu lặp đã được sử dụng để tính giá trị trung bình cho mỗi biến thể duy nhất. [Hình 4](#) minh họa mối tương quan về VAF giữa hai bộ kit thuốc thử. Dựa trên mối tương quan về VAF và sự tương đồng về kết quả giữa các bộ kit thuốc thử, các đặc điểm hiệu suất từng được xác minh và xác nhận với NextSeq 550Dx High Output Reagent Kit v2 (300 chu kỳ) qua Somatic Variant Module được xác định là có thể áp dụng cho NextSeq 550Dx High Output Reagent Kit v2.5 (300 chu kỳ).

Hình 4 Mối tương quan về Tần số alen biến thể (VAF) khi sử dụng Somatic Variant Module giữa NextSeq 550Dx High Output Reagent Kit v2 (300 chu kỳ) và NextSeq 550Dx High Output Reagent Kit v2.5 (300 chu kỳ).



## Lịch sử sửa đổi

Tài liệu	Ngày	Mô tả thay đổi
Tài liệu số 1000000030326 v07	Tháng 8 năm 2023	Đã thêm nội dung để hỗ trợ Illumina DRAGEN Server for NextSeq 550Dx. Đã cập nhật số hiệu bộ phận Bộ lọc khí. Đã cập nhật số biến thể FFPE (Somatic) được sử dụng để xác định tần số alen biến thể. Tóm tắt cập nhật về bảng độ tương hợp dòng mầm Đã thêm cảnh báo về các mối đe dọa an ninh mạng.
Tài liệu số 1000000030326 v06	Tháng 5 năm 2022	Đã thực hiện cập nhật để sửa nội dung vô tình thêm vào từ phần mềm nguồn.
Tài liệu số 1000000030326 v05	Tháng 11 năm 2021	Bổ sung tuyên bố về việc báo cáo sự cố nghiêm trọng vào phần Cảnh báo và biện pháp phòng ngừa. Bổ sung nội dung xác định rõ đối tượng sử dụng vào phần Nguyên tắc của quy trình. Xóa phần đề cập tới High Output Reagent Kit v2 (300 chu kỳ). Bổ sung phần đề cập tới High Output Reagent Kit v2.5 (75 chu kỳ).

Tài liệu	Ngày	Mô tả thay đổi
Tài liệu số 1000000030326 v04	Tháng 8 năm 2021	Bổ sung bảng Lịch sử sửa đổi. Cập nhật địa chỉ của Đại diện được ủy quyền tại châu Âu.

## Bảng sáng chế và nhãn hiệu

Tài liệu này và nội dung trong đó thuộc quyền sở hữu của Illumina, Inc. và các công ty liên kết của Illumina, Inc. ("Illumina") và chỉ dành cho việc sử dụng theo hợp đồng với khách hàng của Illumina liên quan đến việc sử dụng (các) sản phẩm được mô tả trong tài liệu này và không dành cho mục đích nào khác. Tài liệu này và nội dung trong đó sẽ không được sử dụng hay phân phối vì bất kỳ mục đích nào khác và/hoặc không được truyền tải, tiết lộ hay sao chép dưới bất kỳ hình thức nào khác mà không có sự cho phép trước bằng văn bản của Illumina. Illumina không chuyển nhượng bất kỳ giấy phép nào theo các bằng sáng chế, nhãn hiệu, bản quyền hoặc các quyền theo thông luật cũng như các quyền tương tự của bất kỳ bên thứ ba nào thông qua tài liệu này.

Các hướng dẫn nêu trong tài liệu này phải được tuân thủ nghiêm ngặt và rõ ràng bởi cá nhân được đào tạo phù hợp và có đủ trình độ nhằm đảm bảo sử dụng an toàn và đúng cách (các) sản phẩm được mô tả trong tài liệu này. Phải đọc và hiểu hoàn toàn tất cả nội dung của tài liệu này trước khi sử dụng (các) sản phẩm đó.

VIỆC KHÔNG ĐỌC TOÀN BỘ VÀ TUÂN THỦ RÕ RÀNG TẤT CẢ CÁC HƯỚNG DẪN NÊU TRONG TÀI LIỆU NÀY CÓ THỂ DẪN ĐẾN GÂY HƯ HỎNG (CÁC) SẢN PHẨM, GÂY TỔN THƯƠNG CHO CON NGƯỜI, BAO GỒM NGƯỜI DÙNG HOẶC NHỮNG NGƯỜI KHÁC VÀ GÂY THIẾT HẠI TÀI SẢN KHÁC, VÀ SẼ LÀM MẤT HIỆU LỰC BẢO HÀNH ÁP DỤNG CHO (CÁC) SẢN PHẨM ĐÓ.

ILLUMINA KHÔNG CHỊU BẤT KỲ TRÁCH NHIỆM NÀO PHÁT SINH TỪ VIỆC SỬ DỤNG KHÔNG ĐÚNG CÁCH (CÁC) SẢN PHẨM ĐƯỢC MÔ TẢ TRONG TÀI LIỆU NÀY (BAO GỒM CẢ CÁC BỘ PHẬN CỦA SẢN PHẨM HOẶC PHẦN MỀM).

© 2023 Illumina, Inc. Bảo lưu mọi quyền.

Tất cả các nhãn hiệu đều là tài sản của Illumina, Inc. hoặc các chủ sở hữu tương ứng. Để biết thông tin cụ thể về nhãn hiệu, hãy tham khảo trang [www.illumina.com/company/legal.html](http://www.illumina.com/company/legal.html).

## Thông tin liên hệ



Illumina, Inc.  
5200 Illumina Way  
San Diego, California 92122 U.S.A.  
+1.800.809.ILMN (4566)  
+1.858.202.4566 (ngoài khu vực Bắc Mỹ)  
techsupport@illumina.com  
www.illumina.com



Illumina Netherlands B.V.  
Steenoven 19  
5626 DK Eindhoven  
The Netherlands

**Nhà bảo trợ tại Úc**  
Illumina Australia Pty Ltd  
Nursing Association Building  
Level 3, 535 Elizabeth Street  
Melbourne, VIC 3000  
Úc

## Thông tin trên nhãn sản phẩm

Để có thông tin tham chiếu đầy đủ về các ký hiệu xuất hiện trên bao bì và nhãn sản phẩm, hãy tham khảo bản chú giải ký hiệu tại địa chỉ [support.illumina.com](http://support.illumina.com), trên tab *Documentation* (Tài liệu) cho bộ kit của bạn.