

# DRAGEN for Illumina DNA Prep with Enrichment Dx on NextSeq 550Dx

Anwendungsb Benutzerhandbuch

Dieses Dokument und sein Inhalt sind Eigentum von Illumina, Inc. sowie deren Partner-/Tochterunternehmen („Illumina“) und ausschließlich für den bestimmungsgemäßen Gebrauch durch den Kunden in Verbindung mit der Verwendung des hier beschriebenen Produkts/der hier beschriebenen Produkte und für keinen anderen Bestimmungszweck ausgelegt. Dieses Dokument und sein Inhalt dürfen ohne schriftliches Einverständnis von Illumina zu keinem anderen Zweck verwendet oder verteilt bzw. anderweitig übermittelt, offengelegt oder auf irgendeine Weise reproduziert werden. Illumina überträgt mit diesem Dokument keine Lizenzen unter seinem Patent, Markenzeichen, Urheberrecht oder bürgerlichem Recht bzw. ähnlichen Rechten an Drittparteien.

Die Anweisungen in diesem Dokument müssen von qualifiziertem und entsprechend ausgebildetem Personal genau befolgt werden, damit die in diesem Dokument beschriebene Verwendung des Produkts/der Produkte sicher und ordnungsgemäß erfolgt. Vor der Verwendung dieser Produkte muss der Inhalt dieses Dokuments vollständig gelesen und verstanden worden sein.

FALLS NICHT ALLE HIERIN AUFGEFÜHRTEN ANWEISUNGEN VOLLSTÄNDIG GELESEN UND BEFOLGT WERDEN, KÖNNEN PRODUKTSCHÄDEN, VERLETZUNGEN DER BENUTZER UND ANDERER PERSONEN SOWIE ANDERWEITIGER SACHSCHADEN EINTRETEN UND JEDLICHE FÜR DAS PRODUKT/DIE PRODUKTE GELTENDE GEWÄHRLEISTUNG ERLISCHT.

ILLUMINA ÜBERNIMMT KEINERLEI HAFTUNG FÜR SCHÄDEN, DIE AUS DER UNSACHGEMÄSSEN VERWENDUNG DER HIERIN BESCHRIEBENEN PRODUKTE (EINSCHLIESSLICH TEILEN HIERVON ODER DER SOFTWARE) ENTSTEHEN.

© 2023 Illumina, Inc. Alle Rechte vorbehalten.

Alle Marken sind Eigentum von Illumina, Inc. bzw. der jeweiligen Eigentümer. Spezifische Informationen zu Marken finden Sie unter [www.illumina.com/company/legal.html](http://www.illumina.com/company/legal.html).

## Versionsverlauf

Dokument	Datum	Beschreibung der Änderung
200025238 Version 00	Februar 2023	Erste Version.

# Inhaltsverzeichnis

Versionsverlauf .....	iii
Überblick .....	1
Analysemethoden .....	1
Geplanten Lauf erstellen .....	5
Einstellungen .....	8
Manifestdatei .....	8
Rauschfilterung (optional) .....	9
Analyseausgaben .....	9
FASTQ-Dateien .....	11
BAM-Dateien .....	11
VCF-Dateien .....	12
Analyse erneut in die Warteschlange stellen .....	19
<b>Technische Unterstützung .....</b>	<b>20</b>

# Überblick

Mit der Anwendung DRAGEN for Illumina DNA Prep with Enrichment Dx (DRAGEN for IDPE Dx) werden sekundäre Analysen von IDPE-Dx-Bibliotheken geplant und durchgeführt, die zur Sequenzierung auf dem NextSeq 550Dx generiert wurden.

DRAGEN for IDPE Dx unterstützt die Sequenzierung zur Analyse, wenn es mit den Illumina DNA Prep with Enrichment Dx Bibliotheksvorbereitung, NextSeq 550Dx und Illumina DRAGEN Server for NextSeq 550Dx verwendet wird.

## Analysemethoden

DRAGEN for IDPE Dx führt je nach den ausgewählten Arbeitsabläufen Demultiplexing, FASTQ-Generierung, Read-Zuordnung, Alignment mit einem Referenzgenom und Calling kleiner Varianten durch:

- FASTQ-Generierung
- Germline FASTQ- und VCF-Generierung
- Somatic FASTQ- und VCF-Generierung

### HINWEIS

Die ORA-Komprimierung ist für alle drei Arbeitsabläufe verfügbar. DRAGEN ORA Compression ist eine vollständig verlustfreie Komprimierungssoftware, die eine Datei mit der Erweiterung „Original Read Archive“ (\*.ora) erstellt. Das ORA-Format ist ein referenzbasiertes Komprimierungsformat für FASTQ-Dateien und wurde für eine sehr schnelle Komprimierung/Dekomprimierung sowie ein hohes Komprimierungsverhältnis entwickelt.

### FASTQ-Generierung

Die zusammengesetzten Sequenzen werden pro Probe in FASTQ-Dateien geschrieben. FASTQ-Dateien sind Textdateien, die Sequenzierungsdaten und Qualitäts-Scores für nur eine Probe enthalten. Für jede Probe werden separate FASTQ-Dateien pro Fließzellen-Lane und Sequenzierungs-Read generiert. Der Name der Probe, wie er während der Laufeinrichtung angegeben wurde, ist im FASTQ-Dateinamen enthalten. FASTQ-Dateien sind die primären Eingabedateien für das Alignment. Der erste Schritt der FASTQ-Generierung ist das Demultiplexing. Beim Demultiplexing werden Cluster, die den Filter passieren, einer Stichprobe zugewiesen. Dazu wird jede Index-Lesesequenz mit den für den Lauf angegebenen Indexsequenzen verglichen. In diesem Schritt werden keine Qualitätswerte berücksichtigt. Index-Reads werden so identifiziert:

- Die Proben sind gemäß der Reihenfolge, in der sie für den Lauf aufgelistet sind, und mit 1 beginnend durchnummeriert.

- Die Probennummer 0 ist für Cluster reserviert, die keiner Probe zugeordnet wurden.
- Cluster werden einer Probe zugewiesen, wenn die Indexsequenz genau übereinstimmt bzw. je Index-Read maximal eine Nichtübereinstimmung festgestellt wird.

In der Software ist die ORA-Komprimierung zum Komprimieren von FASTQ-Dateien enthalten. Dieses Format kann optional aktiviert werden. Beim ORA-Format (\*.ora) bleibt die md5-Prüfsumme des FASTQ-Inhalts nach einem Komprimierungs- und Dekomprimierungszyklus für eine verlustfreie Komprimierung erhalten.

## DNA-Zuordnung und -Alignment

Nach der FASTQ-Generierung werden die Reads zugeordnet und auf ein Referenzgenom ausgerichtet. In der ersten Zuordnungsphase werden Seeds aus dem Read generiert und exakte Übereinstimmungen im Referenzgenom gesucht. Diese Ergebnisse werden dann durch Ausführen eines kompletten Smith-Waterman-Alignments an den Positionen mit der höchsten Dichte an Seed-Übereinstimmungen präzisiert. Dieser gut dokumentierte Algorithmus gleicht jede Read-Position mit allen Kandidatenpositionen der Referenz ab. Dieser Abgleich entspricht einer Matrix aus möglichen Alignments zwischen Read und Referenz. Der Smith-Waterman-Algorithmus generiert für jede dieser potenziellen Alignment-Positionen Scores, anhand derer beurteilt wird, ob das beste Alignment mit einer Nukleotid-Übereinstimmung oder -Nichtübereinstimmung (diagonale Bewegung), einer Deletion (horizontale Bewegung) oder einer Insertion (vertikale Bewegung) durch diese Matrixzelle gewandert ist. Bei einer Übereinstimmung zwischen Read und Referenz wird zum Score hinzuaddiert, bei einer Nichtübereinstimmung oder einem Indel wird vom Score subtrahiert. Als Alignment wird der höchste Gesamtscore gewählt, der beim Durchwandern der Matrix erzielt wird. Der Algorithmus ist auf den feldprogrammierbaren DRAGEN Gate-Array-Karten (FPGA) hardwarebeschleunigt. Das in der Anwendung verwendete Referenzgenom wird aus dem UCSC hg19 FASTA mit der Option DRAGEN erstellt, um eine auf Liftover basierende ALT-Aware-Hash-Tabelle zu erstellen.

## DRAGEN Calling von Keimbahnvarianten

Der DRAGEN Caller für kleine Keimbahnvarianten nutzt zugeordnete und ausgerichtete DNA-Reads als Eingabe und ruft Einzelnukleotid-Polymorphismen (SNP) sowie Insertionen oder Deletionen (Indels) durch eine Kombination aus spaltenweiser Erkennung und lokaler *De-novo*-Zusammensetzung von Haplotypen ab. Zum Aktivieren des DRAGEN Caller für kleine Keimbahnvarianten wird der Keimbahnvarianten-Arbeitsablauf ausgewählt.

In der Regel wird Keimbahnvarianten-Calling für Keimbahnproben angewendet, bei denen die Ploidie bekanntermaßen 2 ist. Callfähige Referenzregionen werden zuerst mit ausreichender Alignmentabdeckung identifiziert. Innerhalb dieser Referenzregionen identifiziert ein schneller Scan der sortierten Reads aktive Regionen, die um Pile-up-Spalten mit Hinweisen auf eine Variante zentriert sind. Die aktiven Regionen werden mit genügend Kontext aufgefüllt, um wichtige, Nicht-Referenzinhalte in der Nähe abzudecken. Beim Nachweis von Indels werden die aktiven Regionen zusätzlich aufgefüllt.

Ausgerichtete Reads werden innerhalb jeder aktiven Region abgeschnitten und zu einem De Bruijn-Diagramm zusammengesetzt. Die Kanten der abgeschnittenen Reads werden nach Beobachtungszählungen gewichtet, mit der Referenzsequenz als Grundlage. Nach einiger Graphbereinigung und -vereinfachung werden alle Quelle-Senke-Pfade als Kandidaten-Haplotypen extrahiert. Jeder Haplotyp ist nach Smith-Waterman auf das Referenzgenom ausgerichtet, um die Varianten zu bestimmen, die es repräsentiert. Dieser Satz von Ereignissen kann durch einen positionsbasierten Nachweis erweitert werden. Für jedes Read-Haplotyp-Paar wird die Wahrscheinlichkeit  $P(r|H)$ , den Read zu beobachten, unter der Annahme, dass der Haplotyp das wahre Ausgangsmuster ist, mithilfe des Paar-Hidden-Markov-Modells (HMM) geschätzt.

Beim Scannen nach Referenzposition über der aktiven Region werden Kandidaten-Genotypen aus diploiden Kombinationen von Variantenereignissen (SNP oder Indels) gebildet. Für jedes Ereignis (einschließlich Referenz) wird die bedingte Wahrscheinlichkeit  $P(r|e)$  der Beobachtung jedes überlappenden Reads als das Maximum  $P(r|H)$  für Haplotypen geschätzt, die das Ereignis unterstützen. Diese werden zu der bedingten Wahrscheinlichkeit  $P(r|e_1e_2)$  für einen Genotyp (Ereignispaar) kombiniert und multipliziert, um die bedingte Wahrscheinlichkeit  $P(R|e_1e_2)$  zu erhalten, das gesamte Read-Pile-up zu beobachten. Mit der Bayes-Formel wird die A-Posteriori-Wahrscheinlichkeit  $P(e_1e_2|R)$  jedes diploiden Genotyps berechnet und die Auswahl aufgerufen.

DRAGEN for IDPE Dx wendet eine automatische Filterung an. Weitere Informationen finden Sie unter [Anmerkungen zur VCF-Datei des Keimbahnarbeitsablaufs auf Seite 14](#).

## DRAGEN Calling somatischer Varianten

Der DRAGEN Caller für kleine somatische Varianten nutzt kartierte und ausgerichtete DNA-Reads als Eingabe und ruft SNP sowie Indels durch eine lokale *De-novo*-Zusammensetzung von Haplotypen in einer aktiven Region ab. Zum Aktivieren des DRAGEN Caller für kleine somatische Varianten wird eine Anwendung für somatische Varianten ausgewählt.

In der Regel wird für Tumorproben ein Calling somatischer Varianten angewendet. Bei diesem Arbeitsablauf wird bei DRAGEN keine Ploidie vorausgesetzt. Dadurch können niederfrequente Allele nachgewiesen werden. Für Loci mit einer Coverage von bis zu 100x in der Tumorprobe hat DRAGEN einen Nachweisschwellenwert bei Variantenallelfrequenzen von 5 %. Die Grenze skaliert mit zunehmender Tiefe pro Locus und halbiert sich jedes Mal, wenn sich die Coverage über das 100-fache hinaus verdoppelt. Callfähige Referenzregionen werden zuerst mit ausreichender Alignmentabdeckung identifiziert. Innerhalb dieser Referenzregionen identifiziert ein Scan der sortierten Reads aktive Regionen, die um Pile-up-Spalten mit Hinweisen auf eine Variante in den Tumor-Reads zentriert sind. Die aktiven Regionen werden mit genügend Kontext aufgefüllt, um wichtige, Nicht-Referenzinhalte in der Nähe abzudecken. Beim Nachweis von Indels werden die aktiven Regionen zusätzlich aufgefüllt.

Ausgerichtete Reads werden innerhalb jeder aktiven Region abgeschnitten und zu einem De Bruijn-Diagramm zusammengesetzt. Die Kanten der abgeschnittenen Reads werden nach Beobachtungszählungen gewichtet, mit der Referenzsequenz als Grundlage. Nach einiger Graphbereinigung und -vereinfachung werden alle Quelle-Senke-Pfade als Kandidaten-Haplotypen

extrahiert. Jeder Haplotyp ist nach Smith-Waterman auf das Referenzgenom ausgerichtet, um die Varianten zu bestimmen, die es repräsentiert. Für jedes Read-Haplotyp-Paar wird die Wahrscheinlichkeit  $P(r|H)$ , den Read zu beobachten, unter der Annahme, dass der Haplotyp das wahre Ausgangsmuster ist, mithilfe des Paar-Hidden-Markov-Modells (HMM) geschätzt.

Zur Bestimmung des Tumor-Nachweislimit (TLOD)-Scores scannt der DRAGEN Somatic Small Variant Caller zunächst nach Referenzposition für jedes somatische Kandidatenereignis sowie das Referenzereignis über die aktive Region. Die bedingte Wahrscheinlichkeit  $P(r|e)$  der Beobachtung jedes überlappenden Reads als das Maximum  $P(r|H)$  wird für Haplotypen geschätzt, die das Ereignis unterstützen. Diese werden zur bedingten Wahrscheinlichkeit  $P(r|E)$  für eine Ereignishypothese  $E$  kombiniert, die eine Mischung aus dem somatischen Referenz- und Kandidaten-Allel über einen Bereich möglicher Allelhäufigkeiten umfasst, und multipliziert, um die bedingte Wahrscheinlichkeit  $P(R|E)$  aus der Beobachtung des gesamten Read-Pile-up zu erhalten. Daraus wird ein TLOD-Score als Nachweis dafür berechnet, dass ein ALT-Allel in der Tumorprobe an einem bestimmten Locus vorhanden ist.

DRAGEN for IDPE Dx wendet eine automatische Filterung an. Weitere Informationen finden Sie unter [Anmerkungen zur VCF-Datei des somatischen Arbeitsablaufs auf Seite 17](#).



# Geplanten Lauf erstellen

Richten Sie mit diesen Schritten einen Lauf im Illumina Run Manager entweder auf dem NextSeq 550Dx oder in einem Browser auf einem vernetzten Computer ein. Wenn Sie Probanden importieren wollen, nutzen Sie einen Browser auf einem vernetzten Computer. Weitere Anweisungen zum Zugreifen auf den Illumina Run Manager über einen vernetzten Computer finden Sie im Illumina Run Manager for NextSeq 550Dx Softwarehandbuch (Dokument-Nr. 200025239).

Es gibt zwei verschiedene Möglichkeiten, einen neuen geplanten Lauf zu erstellen:

- **Import Run** (Lauf importieren): Verwenden Sie ein Probenblatt aus einem vorhandenen Lauf als Matrize für einen neuen Lauf. Weitere Informationen zum Importieren eines Laufs finden Sie im Illumina Run Manager for NextSeq 550Dx Softwarehandbuch (Dokument-Nr. 200025239).
- **Create Run** (Lauf erstellen): Geben Sie die Laufparameter manuell ein. Richten Sie einen Lauf mit diesen Anweisungen ein.

**HINWEIS** Die erforderlichen Eingabefelder in der Benutzeroberfläche sind mit einem Sternchen (\*) gekennzeichnet.

## Anwendung

1. Wählen Sie auf dem Bildschirm „Runs“ (Läufe) in der Registerkarte „Planned“ (Geplant) die Option **Create Run** (Lauf erstellen) aus.
2. Wählen Sie die Anwendung DRAGEN for Illumina DNA Prep with Enrichment Dx und anschließend **Next** (Weiter) aus.

## Laufeinstellungen

1. Geben Sie auf dem Bildschirm „Run Settings“ (Laufeinstellungen) einen eindeutigen Laufnamen ein. Der Laufname ist der Name, der den Lauf von der Sequenzierung bis hin zur Analyse identifiziert.
2. **[Optional]** Geben Sie eine Laufbeschreibung ein, um den Lauf näher zu bestimmen.
3. Wählen Sie das/die Indexadapterkit(s) aus, das/die während der Vorbereitung der Bibliothek verwendet werden.
4. Überprüfen Sie die Read-Länge und ändern Sie sie bei Bedarf. Read 1 und Read 2 haben einen Standardwert von 151 Zyklen. Index 1 und Index 2 haben einen festen Wert von 10 Zyklen und können nicht geändert werden.
5. **[Optional]** Geben Sie eine Bibliotheksröhrchen-ID ein.
6. Wählen Sie **Next** (Weiter) aus.

## Probendaten

Zu den Probendaten gehören die Proben-ID, die Well-Position (Position des Indexplatten-Wells) und der Bibliotheksname. Bei Verwendung von Index A&B enthält die Well-Position auch die Plattenkennung.

Es gibt zwei Möglichkeiten, um Probendaten einzugeben:

- **Import Samples** (Proben importieren): Verwenden Sie eine Matrizen-datei, die Sie auf dem Bildschirm „Sample Data“ (Probendaten) herunterladen können.
- **Manually** (Manuell): Geben Sie die Probendaten auf dem Bildschirm „Sample Data“ (Probendaten) direkt in die Tabelle ein.

## Importieren von Proben

Wenn Sie einen Sequenzierungslauf in einem Browser auf einem vernetzten Computer planen, steht auf dem Bildschirm „Sample Data“ (Probendaten) eine Matrizen-datei (\*.csv) zum Herunterladen zur Verfügung. Die Matrizen-datei steht beim Zugreifen auf Illumina Run Manager über die NextSeq 550Dx Betriebssystemsoftware nicht zum Herunterladen zur Verfügung. Geben Sie die Probendaten mit der Funktion „Import Samples“ (Proben importieren) mit den folgenden Schritten ein.

**HINWEIS** Bevor Sie fortfahren, führen Sie die Schritte „Run Settings“ (Laufeinstellungen) aus.

1. Wählen Sie zum Herunterladen einer leeren CSV-Datei **Download Template** (Matrize herunterladen) aus.
2. Geben Sie in der Matrizen-datei Probendaten ein und speichern Sie die Datei. Ein Bibliotheksname ist optional.

**HINWEIS** Bei Verwendung von Index A&B müssen die Daten für Spalte B sowohl die Platten- als auch die Well-Position (Position des Indexplatten-Wells) enthalten. Beispiel: A-A01, A-A02, A-A03.

3. Wählen Sie **Import Samples** (Proben importieren) aus, und wechseln Sie zur Matrizen-datei mit den Probendateninformationen aus dem vorherigen Schritt.
4. Wählen Sie **Open** (Öffnen), **Proceed** (Fortfahren) und dann **Next** (Weiter) aus.

**HINWEIS** Wenn Sie die Proben-ID ändern, bevor Sie „Next“ (Weiter) auswählen, kann das zu einem Fehler führen. Bevor Sie Änderungen vornehmen, beenden Sie die Einrichtung des Laufs, um Fehler zu vermeiden.

## Manuelles Eingeben der Proben

Geben Sie Probanddaten anhand der Tabelle auf dem Bildschirm „Sample Data“ (Probanddaten) manuell ein.

1. Geben Sie im Feld „Sample ID“ (Proben-ID) eine eindeutige Proben-ID ein.
2. Wählen Sie den zugehörigen Index für die Proben mit **Well Position** (Well-Position) (Index A oder Index B) oder **Plate - Well Position** (Platten-Well-Position) (Index A&B) aus.  
Die Felder „i7-Index“, „Index 1“, „i5-Index“ und „Index 2“ werden automatisch ausgefüllt.
3. **[Optional]** Geben Sie einen Bibliotheksnamen ein.
4. Fügen Sie Zeilen hinzu und wiederholen Sie die Schritte **1** bis **3** nach Bedarf, bis alle Proben zur Tabelle hinzugefügt wurden. Sie können mehrere Zeilen gleichzeitig hinzufügen. Geben Sie dazu zuerst die Anzahl der hinzuzufügenden Zeilen ein und wählen Sie dann das Symbol „+“ aus. Sie können Zeilen auch entfernen. Markieren Sie dazu das Kästchen neben der Zeilennummer und klicken Sie dann auf das Papierkorb-Symbol.
5. Wählen Sie **Next** (Weiter) aus.

## Analyseeinstellungen

1. Wählen Sie den gewünschten Analysearbeitsablauf aus:
  - FASTQ-Generierung
  - FASTQ- und VCF-Generierung für einen Keimbahnarbeitsablauf (Manifestdatei erforderlich)
  - FASTQ- und VCF-Generierung für einen somatischen Arbeitsablauf (Manifestdatei erforderlich)
2. **[Optional] Generate ORA compressed FASTQs** (ORA-komprimierte FASTQ generieren) ist standardmäßig aktiviert. Die FASTQ-ORA-Komprimierung komprimiert FASTQ-Dateien verlustfrei bis zu 5-mal im Vergleich zu fastq.gz. Deaktivieren Sie **Generate ORA compressed FASTQs** (ORA-komprimierte FASTQ generieren), wenn unkomprimierte Daten (fastq.gz) bevorzugt werden.
3. Für Keimbahn- und somatische Arbeitsabläufe ist eine Manifestdatei erforderlich. Wählen Sie aus dem Dropdown-Menü **Manifest File Selection** (Auswahl der Manifestdatei) eine Manifestdatei aus. Das Manifest ist eine tabulatorgetrennte BED-Datei (\*.bed), in der die Namen und Positionen der Target-Referenzregionen definiert sind. Weitere Informationen finden Sie unter [Manifestdatei auf Seite 8](#).
4. **[Optional]** Wählen Sie eine Datei mit systematischem Rauschen für somatische Arbeitsabläufe aus dem Dropdown-Menü **Noise File Selection** (Auswahl der Rauschdatei) aus.  
Zum Herausfiltern von systematischem Rauschen kann eine BED-Datei (\*.bed.gz) mit positionsspezifischem Rauschpegel vorgegeben werden. Weitere Informationen finden Sie unter [Rauschfilterung \(optional\) auf Seite 9](#).
5. Wählen Sie **Next** (Weiter) aus.

## Lauf- Überprüfung

1. Überprüfen Sie auf dem Bildschirm „Review“ (Überprüfen) die Informationen für „Run Settings“ (Laufeinstellungen), „Sample Data“ (Probanddaten) und „Analysis Settings“ (Analyseeinstellungen).

2. Wählen Sie **Save** (Speichern) aus.

Der Lauf wird im Bildschirm „Runs“ (Läufe) in der Registerkarte „Planned“ (Geplant) gespeichert.

## Einstellungen

Wählen Sie zum Aufrufen oder Ändern der DRAGEN for IDPE Dx Anwendungseinstellungen auf dem Hauptbildschirm zunächst das Symbol „Applications“ (Anwendungen) aus. Wählen Sie dann die aufzurufende oder zu ändernde Anwendung aus. Zum Ändern der Einstellungen ist ein Administratorkonto erforderlich.

### Konfiguration

Im Bildschirm „Configuration“ (Konfiguration) werden diese Anwendungseinstellungen angezeigt:

- **Library Prep Kits** (Bibliotheksvorb.-Kits): Zeigt das standardmäßige Bibliotheksvorb.-Kit für die Anwendung an. Diese Einstellung kann nicht geändert werden.
- **Index Adapter Kits** (Index-Adapter-Kits): Zeigt das standardmäßige Index-Adapter-Kit für die Anwendung an. Diese Einstellung kann nicht geändert werden.
- **Read lengths** (Read-Längen): Die Read-Längen sind für die App standardmäßig auf 151 festgelegt, können jedoch während der Lauferstellung geändert werden.
- **Manifest and Noise Files** (Manifest- und Rauschdateien): Laden Sie die Einstellungen für Manifest- und Rauschdateien hoch und ändern Sie sie.
  - Wählen Sie **Upload File** (Datei hochladen) aus, um Dateien für die Analyse hochzuladen.
  - Wählen Sie das Optionsfeld **Default** (Standard) aus, um die Datei als Standardmanifest oder Rauschdatei festzulegen, die während der Lauferstellung ausgewählt wird, wenn die Anwendung ausgewählt wird.
  - Markieren Sie das Kästchen **Enabled** (Aktiviert), um festzulegen, dass die Datei während der Lauferstellung im Dropdown-Menü angezeigt wird.

### Berechtigungen

Verwalten Sie über die Kästchen auf dem Bildschirm „Berechtigungen“ den Benutzerzugriff für die App.

## Manifestdatei

Wenn Sie DRAGEN for IDPE Dx anwenden, ist für diese Arbeitsabläufe eine Manifestdatei erforderlich:

- FASTQ- und VCF-Generierung für einen Keimbahnarbeitsablauf
- FASTQ- und VCF-Generierung für einen somatischen Arbeitsablauf

Die Manifestdatei ist eine tabulatorgetrennte Datei im BED-Format (\*.bed), die die Namen und Positionen der Target-Referenzregionen definiert. Der Hauptabschnitt der Manifestdatei ist der Abschnitt „Regions“ (Regionen) und muss diese Datenspalten enthalten:

Spalte	Beschreibung
Name	Eindeutiger benutzerdefinierter Name für das Target
Chromosom	Chromosomenlage (z. B. chr10, chr5 usw.)
Anfang	1-basierter Index für die Startposition des Targets
Ende	1-basierter Index für die Stopposition des Targets
Länge der stromaufwärts gelegenen Sonde	Die Länge der stromaufwärts gelegenen Sonde. Für die Anwendung DRAGEN for IDPE Dx muss dieser Wert auf „0“ festgelegt werden.
Länge der stromabwärts gelegenen Sonde	Die Länge der stromabwärts gelegenen Sonde. Für die Anwendung DRAGEN for IDPE Dx muss dieser Wert auf „0“ festgelegt werden.

**HINWEIS** Für die Analyse ist ein gültiges Manifestdateiformat erforderlich. DRAGEN stoppt die Analyse, wenn die Manifestdatei ungültig ist.

## Rauschfilterung (optional)

Der systematische Rauschfilter ist für das Calling somatischer Varianten verfügbar. Dadurch können falsch positive Calls durch Berücksichtigung standortspezifisches Rauschen reduziert werden. Die Datei mit systematischem Rauschen wird erzeugt, indem zuerst ca. 50 normale Proben entnommen werden (vorzugsweise spezifisch für Panel, Bibliotheksvorbereitung und Sequenzer). Dann wird die Summe der Allelfrequenzen unter 30 % an jedem Standort mit ausreichender Coverage durch die Gesamtzahl der Proben geteilt (Allelfrequenzen über 30 % werden als Keimbahnvarianten und nicht als Rauschen angenommen). Sobald die Rauschwerte generiert wurden, werden an dieser Stelle erkannte somatische Varianten gefiltert.

Der Filter kann im Tumor-Normal-Modus verwendet werden, ist aber besonders nützlich für Nur-Tumor-Läufe, bei denen kein übereinstimmender Normalwert verfügbar ist. Die Datei mit systematischem Rauschen muss eine BED-Datei mit einer Dateierweiterung (\*.bed.gz) verwenden und vier Spalten enthalten: Chromosom, Anfang, Ende und standortspezifische Rauschpegel für jede Zeile. Die systematische Rauschfilterung ist optional.

## Analyseausgaben

Laufende Läufe werden in der Registerkarte „Active“ (Aktiv) angezeigt. Abgeschlossene Läufe werden auf der Registerkarte „Completed“ (Abgeschlossen) angezeigt. DRAGEN for IDPE Dx erstellt für jede Analyse einen eindeutig benannten Analyseordner, der vom Ordner mit Sequenzierungsdaten getrennt ist. Im Analyseordner sind diese Informationen enthalten:

- Verwendete Manifestdatei

- Softwareversion
- Proben-ID
- Ausgerichtete Reads insgesamt
- Prozentsatz ausgerichteter Reads pro Probe
- Anzahl aufgerufener SNV pro Probe
- Anzahl aufgerufener Indels pro Probe
- Abdeckungsstatistik

## Analyse-Ausgabedateien

Der Speicherort des Analyseordners wird durch die Einstellung „External Storage for Analysis Results“ (Externer Speicher für Analyseergebnisse) festgelegt. Weitere Informationen zur Einstellung „External Storage for Analysis Results“ (Externer Speicher für Analyseergebnisse) finden Sie im Illumina Run Manager for NextSeq 550Dx Softwarehandbuch (Dokument-Nr. 200025239).

Auf dem Bildschirm „Run Details“ (Laufdetails) enthält das Feld „External Location“ (Externer Speicherort) den Pfad für Sequenzierungsdaten. Der eindeutige Name des Analyseordners wird auf dem Bildschirm „Run Details“ (Laufdetails) im Feld „Analysis Output Folder“ (Analyseausgabeordner) angegeben. Die genauen generierten Dateien hängen davon ab, welcher Analyseablauf angewendet wird. Die folgenden Analyseausgabedateien werden von der Anwendung generiert.

**HINWEIS** Wenn beim Zugriff auf Analyseausgabedateien ein Fehler bei der maximalen Dateipfadlängenbegrenzung auftritt, versuchen Sie, die Datei an einen kürzeren Pfad zu verschieben oder die Datei mit einer anderen Methode zu öffnen.

Ausgabedatei	Beschreibung
Zusammenfassender Variantenbericht (*.pdf)	Enthält eine Zusammenfassung der Dateiinformationen, Softwareversionen, Probeninformationen, Statistikwerte auf Read-Ebene und SNV-, Insertionen-, Deletionen- und Coverage-Zusammenfassungen. Nur im Keimbahn-Arbeitsablauf und im somatischen Arbeitsablauf wird ein Variantenbericht erstellt.
FASTQ (*.fastq.gz oder *.fastq.ora)	Temporäre Dateien mit qualitativ benoteten Base-Calls. FASTQ-Dateien sind die primären Eingabedateien für den Alignment-Schritt. Wenn ORA-Komprimierung ausgewählt ist, wird die Dateierweiterung *.fastq.ora verwendet.
Alignment-BAM-Dateien (*.bam)	Enthält ausgerichtete Reads für eine bestimmte Probe.

Ausgabedatei	Beschreibung
Genom-VCF-Dateien (* .gvcf.gz)	Enthält den Genotyp für jede Position, ob als Variante oder als Referenz aufgerufen.
VCF-Dateien (* .vcf.gz)	Enthält an jeder Position aufgerufene Varianten.
Laufmetrikbericht (* .csv)	Enthält Qualitätsmetriken zum Lauf, einschließlich nicht indizierte Gesamtausbeute und Q30-Score.

## FASTQ-Dateien

FASTQ (\*.fastq.gz, \*.fastq.ora) ist ein textbasiertes Dateiformat mit den Base-Calls und Qualitätswerten pro Read. Jeder Eintrag enthält die folgenden Informationen:

- den Probenbezeichner
- die Sequenz
- die Phred-Qualitäts-Scores in einem ASCII + 33-codierten Format

Der Probenbezeichner ist so formatiert:

```
@Instrument:RunID:FlowCellID:Lane:Tile:X:Y
ReadNum:FilterFlag:0:SampleNumber
Beispiel:
@SIM:1:FCX:1:15:6329:1045 1:N:0:2
TCGCACTCAACGCCCTGCATATGACAAGACAGAATC
+
<>;##=><9=AAAAAAAAAA9#:<#<;<<<????#=#
```

## BAM-Dateien

Eine BAM-Datei (\*.bam) ist die komprimierte Binärversion einer SAM-Datei Sequenz-Alignment-Zuordnung, mit der ausgerichtete Sequenzen bis zu 128 Mb dargestellt werden. BAM-Dateien verwenden das Dateibenennungsformat `SampleName_S#.bam`, wobei # die Probennummer ist, die durch die Reihenfolge bestimmt wird, in der die Proben für den Lauf aufgelistet werden. Im Multinode-Modus wird S# unabhängig von der Reihenfolge der Probe auf S1 festgelegt.

BAM-Dateien verfügen über einen Kopfzeilen- und einen Alignment-Abschnitt:

- **Kopfzeile:** Enthält Informationen über die gesamte Datei, z. B. den Probennamen, die Probenlänge und die Alignment-Methode. Alignments im Abschnitt „Alignments“ (Ausrichtungen) sind bestimmten Informationen im Abschnitt „Header“ (Kopfzeile) zugeordnet.
- **Ausrichtungen:** Enthält Read-Name, Read-Sequenz, Read-Qualität, Informationen zu den Alignments und benutzerdefinierte Tags. Der Read-Name enthält das Chromosom, die Startkoordinate, die Alignment-Qualität und den Match-Deskriptor-String.

Im Abschnitt „Alignments“ (Ausrichtungen) sind diese Informationen für jeden Read oder jedes Read-Paar enthalten:

- AS: Paired-End-Alignment-Qualität.
- RG: Read-Gruppe, die die Anzahl der Reads für eine bestimmte Probe angibt.
- BC: Barcode-Tag, der die demultiplexierte Proben-ID angibt, die dem Read zugeordnet ist.
- SM: Single-End-Alignment-Qualität.
- XC: Match-Deskriptor-String.
- XN: Amplikon-Namens-Tag, das die dem Read zugeordnete Amplikon-ID aufzeichnet.

BAM-Indexdateien (\*.bam.bai) zeigen einen Index der entsprechenden BAM-Datei.

## VCF-Dateien

Varianten-Call-Format-Dateien (\*.vcf) enthalten Informationen über Varianten, die an spezifischen Positionen in einem Referenzgenom gefunden wurden.

In der Kopfzeile der VCF-Datei sind die VCF-Dateiformatversion und die Varianten-Caller-Version enthalten und die Anmerkungen angegeben, die im Rest der Datei verwendet werden. In der VCF-Kopfzeile sind auch die Referenzgenomdatei und die BAM-Datei enthalten. In der letzten Zeile in der Kopfzeile sind die Spaltenüberschriften für die Datenzeilen enthalten. In allen Datenzeilen der VCF-Datei sind Informationen zu einer Variante enthalten.

Tabelle 1 Kopfzeilen in VCF-Dateien

Kopfzeile	Beschreibung
CHROM	Das Chromosom des Referenzgenoms. Chromosomen erscheinen in der gleichen Reihenfolge wie die Referenz-FASTA-Datei.
POS	Die Single-Base-Position der Variante im Referenzchromosom. Für Einzelnukleotidvarianten (SNV) ist diese Position die Referenz-Base mit der Variante. Für Indels ist diese Position die Referenzbasis unmittelbar vor der Variante.
ID	Die rs-Nummer (Referenz-SNP) für den SNP, erhalten aus <code>dbSNP.txt</code> , falls zutreffend. Wenn es mehrere rs-Nummern an dieser Position gibt, wird die Liste durch Strichpunkte getrennt. Wenn an dieser Position kein dbSNP-Eintrag existiert, wird eine Kennung für einen fehlenden Wert ('.') verwendet.
REF	Der Referenzgenotyp. Beispielsweise wird die Deletion eines einzelnen T als Referenz TT und alternatives T dargestellt. Die Einzelnukleotidvarianten A bis T werden als Referenz A und alternatives T dargestellt.



Kopfzeile	Beschreibung
ALT	Die Allele, die sich vom Referenz-Read unterscheiden. Beispielsweise wird eine Insertion eines einzelnen T als Referenz A und alternatives AT dargestellt. Die Einzelnukleotidvarianten A bis T werden als Referenz A und alternatives T dargestellt.
QUAL	Ein vom Varianten-Caller zugewiesener Phred-skaliertes Qualitäts-Score. Höhere Scores weisen auf eine höhere Zuverlässigkeit der Variante und eine geringere Fehlerwahrscheinlichkeit hin. Bei einem Qualitäts-Score von Q beträgt die geschätzte Fehlerwahrscheinlichkeit $10^{-(Q/10)}$ . Beispielsweise hat die Reihe der Q30-Calls eine Fehlerrate von 0,1 %. Viele Varianten-Caller weisen Qualitäts-Scores auf Basis ihrer statistischen Modelle zu, die im Verhältnis zur beobachteten Fehlerrate hoch sind.

Tabelle 2 Anmerkungen zur VCF-Datei des Keimbahnarbeitsablaufs

Kopfzeile	Beschreibung
FILTER	<p>Wenn alle Filter erfolgreich sind, wird in die Filterspalte „PASS“ (ERFOLGREICH) geschrieben. Mögliche FILTER-Einträge sind:</p> <ul style="list-style-type: none"> <li>• <b>DRAGENSnpHardQUAL:</b> Wird angewendet, wenn der QUAL-Score der SNP-Variante den Schwellenwert nicht erreicht.</li> <li>• <b>DRAGENIndelHardQUAL:</b> Wird angewendet, wenn der QUAL-Score der Indel-Variante den Schwellenwert nicht erreicht.</li> <li>• <b>LowDepth:</b> Position gefiltert, da die Abdeckungstiefe den Schwellenwert nicht erreicht.</li> <li>• <b>LowGQ:</b> Position gefiltert, da die Qualität des Genotyps den Schwellenwert nicht erreicht.</li> <li>• <b>PloidyConflict:</b> Genotyp-Call vom Varianten-Caller stimmt nicht mit Chromosomen-Ploidie überein.</li> <li>• <b>base_quality:</b> Position gefiltert, da die mittlere Basisqualität von Alt-Reads an diesem Locus den Schwellenwert nicht erreicht.</li> <li>• <b>filtered_reads:</b> Position gefiltert, da ein zu großer Teil der Reads herausgefiltert wurde.</li> <li>• <b>fragment_length:</b> Position gefiltert, da der absolute Unterschied zwischen der mittleren Fragmentlänge von Alt-Reads und der mittleren Fragmentlänge von Ref-Reads an diesem Locus den Schwellenwert überschreitet.</li> <li>• <b>low_depth:</b> Position gefiltert, da die Read-Tiefe zu gering ist.</li> <li>• <b>low_frac_info_reads:</b> Position gefiltert, da der Anteil informativer Reads unter dem Schwellenwert liegt.</li> <li>• <b>low_normal_depth:</b> Position gefiltert, da die Read-Tiefe normaler Proben zu gering ist.</li> <li>• <b>long_indel:</b> Position gefiltert, da die Indel-Länge zu lang ist.</li> <li>• <b>mapping_quality:</b> Position gefiltert, da die mediane Zuordnungsqualität von Alt-Reads an diesem Locus den Schwellenwert nicht erreicht.</li> <li>• <b>multiallelic:</b> Position gefiltert, da mehr als zwei alt-Allele die Tumor-LOD passieren.</li> <li>• <b>non_homref_normal:</b> Position gefiltert, da der normale Probengenotyp keine homozygote Referenz ist.</li> <li>• <b>no_reliable_supporting_read:</b> Position gefiltert, da kein zuverlässiger unterstützender somatischer Read vorhanden ist.</li> <li>• <b>panel_of_normals:</b> In mindestens einer Probe der Normalpanel-VCF festgestellt.</li> <li>• <b>read_position:</b> Position gefiltert, da der Median der Reichweiten zwischen Start/Ende des Reads und diesem Locus unter dem Schwellenwert liegt.</li> <li>• <b>RMxNRepeatRegion:</b> Position gefiltert, da das Varianten-Allel oder ein Teil davon eine Wiederholung der Referenz ist.</li> <li>• <b>strand_artifact:</b> Position gefiltert aufgrund starker Strang-Verzerrung.</li> <li>• <b>str_contraction:</b> Position gefiltert aufgrund eines vermuteten PCR-Fehlers, bei dem das alt-Allel eine Wiederholungseinheit weniger als die Referenz ist.</li> </ul>

Kopfzeile	Beschreibung
FILTER	<ul style="list-style-type: none"> <li>• <b>too_few_supporting_reads</b>: Position gefiltert, da zu wenige unterstützende Reads in der Tumorprobe vorhanden sind.</li> <li>• <b>weak_evidence</b>: Score der somatischen Variante erreicht Schwellenwert nicht.</li> </ul>
INFO	<p>Mögliche INFO-Einträge sind:</p> <ul style="list-style-type: none"> <li>• <b>AC</b>: Allelanzahl in Genotypen für jedes ALT-Allel, in derselben Reihenfolge wie angegeben.</li> <li>• <b>AF</b>: Allelhäufigkeit für jedes ALT-Allel, in derselben Reihenfolge wie angegeben.</li> <li>• <b>AN</b>: Gesamtzahl der Allele in aufgerufenen Genotypen.</li> <li>• <b>DB</b>: dbSNP-Mitgliedschaft.</li> <li>• <b>FS</b>: Phred-skaliertes p-Wert anhand des exakten Fisher-Tests zur Erkennung von Strang-Verzerrungen.</li> <li>• <b>QD</b>: Variantenzuverlässigkeit/-qualität nach Tiefe.</li> <li>• <b>R2_5P_bias</b>: Score basierend auf Mate-Verzerrung und Reichweite von 5 Prime Ends.</li> <li>• <b>SOR</b>: Symmetrisches Chancenverhältnis der 2x2-Kontingenztafel zur Erkennung von Strang-Verzerrungen.</li> <li>• <b>DP</b>: Ungefähre Read-Tiefe (informativ und nicht informativ); einige Reads wurden möglicherweise basierend auf Mapq usw. gefiltert.</li> <li>• <b>END</b>: Endposition des Intervalls.</li> <li>• <b>FractionInformativeReads</b>: Der Anteil informativer Reads an den Gesamt-Reads.</li> <li>• <b>MQ</b>: RMS-Zuordnungsqualität.</li> <li>• <b>MQRankSum</b>: Z-Score aus dem Wilcoxon-Rangsummentest von Alt- vs. Ref-Read-Zuordnungsqualitäten.</li> <li>• <b>ReadPosRankSum</b>: Z-Score aus dem Wilcoxon-Rangsummentest von Alt- vs. Ref-Read-Positionsverzerrungen.</li> <li>• <b>SOMATIC</b>: Mindestens eine Variante an dieser Position ist somatisch.</li> </ul>

Kopfzeile	Beschreibung
FORMAT	<p>In der Formatspalte werden durch Doppelpunkte getrennte Felder angegeben. Beispiel: GT:GQ. Verfügbare Felder sind:</p> <ul style="list-style-type: none"> <li>• <b>AD</b>: Alleltiefen (wobei nur informative Reads von den Gesamt-Reads gezählt werden) für die ref- und alt-Allele in der angegebenen Reihenfolge.</li> <li>• <b>AF</b>: Allelfraktionen für Alt-Allele in der angegebenen Reihenfolge.</li> <li>• <b>DP</b>: Ungefähre Read-Tiefe (Reads mit MQ = 255 oder schlechten Mates werden gefiltert).</li> <li>• <b>F1R2</b>: Anzahl der Reads in F1R2-Paarorientierung, die jedes Allel unterstützen.</li> <li>• <b>F2R1</b>: Anzahl der Reads in F2R1-Paarorientierung, die jedes Allel unterstützen.</li> <li>• <b>GT</b>: Genotyp. 0 entspricht der Referenzbasis, 1 entspricht dem ersten Eintrag in der Spalte „ALT“ und so weiter. Der Schrägstrich (/) gibt an, dass keine Phaseninformationen verfügbar sind.</li> <li>• <b>MB</b>: Komponentenstatistiken pro Probe zur Erkennung von Mate-Verzerrungen.</li> <li>• <b>PS</b>: Physische Phasen-ID-Informationen, bei denen jede eindeutige ID innerhalb einer bestimmten Probe (aber nicht über Proben hinweg) Datensätze innerhalb einer Phasengruppe verbindet.</li> <li>• <b>SB</b>: Komponentenstatistiken pro Probe mit dem exakten Fisher-Tests zur Erkennung von Strang-Verzerrungen.</li> <li>• <b>SQ</b>: Somatische Qualität.</li> </ul>
SAMPLE	Die Probenspalte enthält die Werte, die in der Spalte „FORMAT“ angegeben sind.

Tabelle 3 Anmerkungen zur VCF-Datei des somatischen Arbeitsablaufs

Kopfzeile	Beschreibung
FILTER	<p>Wenn alle Filter erfolgreich sind, wird in die Filterspalte „PASS“ (ERFOLGREICH) geschrieben. Mögliche FILTER-Einträge sind:</p> <ul style="list-style-type: none"> <li>• <b>base_quality</b>: Position gefiltert, da die mittlere Basisqualität von Alt-Reads an diesem Locus den Schwellenwert nicht erreicht.</li> <li>• <b>filtered_reads</b>: Position gefiltert, da ein zu großer Teil der Reads herausgefiltert wurde.</li> <li>• <b>fragment_length</b>: Position gefiltert, da der absolute Unterschied zwischen der mittleren Fragmentlänge von Alt-Reads und der mittleren Fragmentlänge von Ref-Reads an diesem Locus den Schwellenwert überschreitet.</li> <li>• <b>low_depth</b>: Position gefiltert, da die Read-Tiefe zu gering ist.</li> <li>• <b>low_frac_info_reads</b>: Position gefiltert, da der Anteil informativer Reads unter dem Schwellenwert liegt.</li> <li>• <b>low_normal_depth</b>: Position gefiltert, da die Read-Tiefe normaler Proben zu gering ist.</li> <li>• <b>long_indel</b>: Position gefiltert, da die Indel-Länge zu lang ist.</li> <li>• <b>mapping_quality</b>: Position gefiltert, da die mediane Zuordnungsqualität von Alt-Reads an diesem Locus den Schwellenwert nicht erreicht.</li> <li>• <b>multiallelic</b>: Position gefiltert, da mehr als zwei alt-Allele die Tumor-LOD passieren.</li> <li>• <b>non_homref_normal</b>: Position gefiltert, da der normale Probengenotyp keine homozygote Referenz ist.</li> <li>• <b>no_reliable_supporting_read</b>: Position gefiltert, da kein zuverlässiger unterstützender somatischer Read vorhanden ist.</li> <li>• <b>panel_of_normals</b>: In mindestens einer Probe der Normalpanel-VCF festgestellt.</li> <li>• <b>read_position</b>: Position gefiltert, da der Median der Reichweiten zwischen Start/Ende des Reads und diesem Locus unter dem Schwellenwert liegt.</li> <li>• <b>RMxNRepeatRegion</b>: Position gefiltert, da das Varianten-Allel oder ein Teil davon eine Wiederholung der Referenz ist.</li> <li>• <b>strand_artifact</b>: Position gefiltert aufgrund starker Strang-Verzerrung.</li> <li>• <b>str_contraction</b>: Position gefiltert aufgrund eines vermuteten PCR-Fehlers, bei dem das alt-Allel eine Wiederholungseinheit weniger als die Referenz ist.</li> <li>• <b>too_few_supporting_reads</b>: Position gefiltert, da zu wenige unterstützende Reads in der Tumorprobe vorhanden sind.</li> <li>• <b>weak_evidence</b>: Score der somatischen Variante erreicht Schwellenwert nicht.</li> <li>• <b>systematic_noise</b>: Position gefiltert aufgrund von Anzeichen systematischen Rauschens bei Normalen.</li> </ul>

Kopfzeile	Beschreibung
INFO	<p>Mögliche INFO-Einträge sind:</p> <ul style="list-style-type: none"> <li>• <b>DP</b>: Ungefähre Read-Tiefe (informativ und nicht informativ); einige Reads wurden möglicherweise basierend auf Mapq usw. gefiltert.</li> <li>• <b>END</b>: Endposition des Intervalls.</li> <li>• <b>FractionInformativeReads</b>: Der Anteil informativer Reads an den Gesamt-Reads.</li> <li>• <b>MQ</b>: RMS-Zuordnungsqualität.</li> <li>• <b>MQRankSum</b>: Z-Score aus dem Wilcoxon-Rangsummentest von Alt- vs. Ref-Read-Zuordnungsqualitäten.</li> <li>• <b>ReadPosRankSum</b>: Z-Score aus dem Wilcoxon-Rangsummentest von Alt- vs. Ref-Read-Positionsverzerrungen.</li> <li>• <b>AQ</b>: Score für systematisches Rauschen.</li> <li>• <b>hotspot</b>: Bekannte somatische Position, mit der die Zuverlässigkeit des Calls erhöht wird.</li> <li>• <b>SOMATIC</b>: Mindestens eine Variante an dieser Position ist somatisch.</li> </ul>
FORMAT	<p>In der Formatspalte werden durch Doppelpunkte getrennte Felder angegeben. Beispiel: GT:GQ. Verfügbare Felder sind:</p> <ul style="list-style-type: none"> <li>• <b>AD</b>: Alleltiefen (wobei nur informative Reads von den Gesamt-Reads gezählt werden) für die ref- und alt-Allele in der angegebenen Reihenfolge.</li> <li>• <b>AF</b>: Allelfraktionen für Alt-Allele in der angegebenen Reihenfolge.</li> <li>• <b>DP</b>: Ungefähre Read-Tiefe (Reads mit MQ = 255 oder schlechten Mates werden gefiltert).</li> <li>• <b>F1R2</b>: Anzahl der Reads in F1R2-Paarorientierung, die jedes Allel unterstützen.</li> <li>• <b>F2R1</b>: Anzahl der Reads in F2R1-Paarorientierung, die jedes Allel unterstützen.</li> <li>• <b>GP</b>: Phred-skalierte Posterior-Wahrscheinlichkeiten für Genotypen gemäß der Definition in der VCF-Spezifikation.</li> <li>• <b>GQ</b>: Genotypqualität.</li> <li>• <b>GT</b>: Genotyp. 0 entspricht der Referenzbasis, 1 entspricht dem ersten Eintrag in der Spalte „ALT“ und so weiter. Der Schrägstrich (/) gibt an, dass keine Phaseninformatoren verfügbar sind.</li> <li>• <b>MB</b>: Komponentenstatistiken pro Probe zur Erkennung von Mate-Verzerrungen.</li> <li>• <b>PL</b>: Normalisierte, Phred-skalierte Wahrscheinlichkeiten für Genotypen gemäß der Definition in der VCF-Spezifikation.</li> <li>• <b>PRI</b>: Phred-skalierte Prior-Wahrscheinlichkeiten für Genotypen.</li> <li>• <b>PS</b>: Physische Phasen-ID-Informationen, bei denen jede eindeutige ID innerhalb einer bestimmten Probe (aber nicht über Proben hinweg) Datensätze innerhalb einer Phasengruppe verbindet.</li> <li>• <b>SB</b>: Komponentenstatistiken pro Probe mit dem exakten Fisher-Tests zur Erkennung von Strang-Verzerrungen.</li> <li>• <b>SQ</b>: Somatische Qualität.</li> </ul>
SAMPLE	Die Probenspalte enthält die Werte, die in der Spalte „FORMAT“ angegeben sind.

## Genom-VCF-Dateien

Genom-VCF-Dateien (\*.gvcf.gz) folgen einer Reihe von Konventionen zur Darstellung aller Positionen innerhalb des Genoms in einem angemessen kompakten Format. Die gVCF-Dateien enthalten alle Positionen innerhalb der interessierenden Region in einer einzigen Datei für jede Probe. Die gVCF-Datei zeigt No-Calls an Positionen, die nicht alle Filter bestehen. Ein Genotyp(GT)-Tag von ./. weist auf einen No-Call hin.

## Analyse erneut in die Warteschlange stellen

Sie können eine Analyse erneut in die Warteschlange stellen, wenn sie angehalten wurde, nicht erfolgreich war oder wenn Sie einen Lauf mit anderen Einstellungen neu analysieren möchten. Stellen Sie die Analyse mit diesen Schritten in die Warteschlange:

1. Wählen Sie auf dem Bildschirm „Run“ (Ausführen) die Registerkarte „Completed“ (Abgeschlossen) und dann den zu analysierenden Laufnamen aus.  
Wenn zuvor „Requeue Analysis“ (Analyse erneut in die Warteschlange stellen) durchgeführt wurde, wählen Sie den Laufnamen des „Parent Run“ (Übergeordneter Lauf) aus.
2. Wählen Sie auf dem Bildschirm „Run Details“ (Laufdetails) nach „Sequencing Information“ (Sequenzierungsinformationen) die Option **Requeue Analysis** (Analyse erneut in die Warteschlange stellen) aus.
3. Wählen Sie eine Option aus:
  - Analyse ohne Änderungen wieder in die Warteschlange stellen
  - Laufeinstellungen bearbeiten und Analyse erneut in die Warteschlange stellen
  - Analyse mit einer anderen Anwendung wieder in die Warteschlange stellen
4. Bestätigen Sie, dass der Speicherort, an dem sich die Sequenzierungsdaten derzeit befinden, im Feld **Sequencing data file path** (Sequenzierungsdatendateipfad) angegeben ist.

### HINWEIS

Der Pfad zu den Sequenzierungsdaten muss mit dem Pfad in der Einstellung „External Storage for Analysis Results“ (Externer Speicher für Analyseergebnisse) übereinstimmen. Weitere Informationen zum Ändern des externen Speicherpfads finden Sie im Illumina Run Manager for NextSeq 550Dx Softwarehandbuch (Dokument-Nr. 200025239).

5. Geben Sie einen „Reanalysis Reason“ (Grund für erneute Analyse) ein.
6. Wählen Sie **Requeue Analysis** (Analyse erneut in die Warteschlange stellen).
7. Bearbeiten Sie die gewünschten Änderungen an den „Run Settings“ (Laufeinstellungen), „Sample Data“ (Probendaten) und „Analysis Settings“ (Analyseeinstellungen).
8. Wählen Sie **Save** (Speichern) aus. Die Analyse wird mit den aktuellen Analyseparametern durchgeführt.

# Technische Unterstützung

Wenn Sie technische Unterstützung benötigen, wenden Sie sich an den technischen Support von Illumina.

**Website:** [www.illumina.com](http://www.illumina.com)

**E-Mail:** [techsupport@illumina.com](mailto:techsupport@illumina.com)

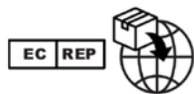
**Sicherheitsdatenblätter (SDS)** sind auf der Illumina-Website unter [support.illumina.com/sds.html](http://support.illumina.com/sds.html) verfügbar.

Die **Produktdokumentation** steht unter [support.illumina.com](http://support.illumina.com) zum Herunterladen zur Verfügung.





Illumina  
5200 Illumina Way  
92122 San Diego, Kalifornien, USA  
+1.800.809.ILMN (4566)  
+1.858.202.4566 (außerhalb von Nordamerika)  
techsupport@illumina.com  
www.illumina.com



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

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

FÜR IN-VITRO-DIAGNOSTIK.

© 2023 Illumina, Inc. Alle Rechte vorbehalten.

**illumina®**