

DRAGEN for Illumina DNA Prep with Enrichment Dx

Produktdokumentation för NovaSeq 6000Dx

TILLHÖR ILLUMINA

Dokument nr 200014776 v02

September 2022

FÖR IN VITRO-DIAGNOSTISK ANVÄNDNING

Revisionshistorik

Dokument	Datum	Ändringsbeskrivning
200014776 v02	September 2022	Korrigerat format för manifestfil från text (*.txt) till BED (*.bed) i instruktioner för att skapa körning. Korrigerade konsensus VCF-filer till VCF-filer i analysutdatasektionen.
200014776 v01	Augusti 2022	Tillagt: Avsnittet Inställningar. Avsnittet Systematisk brusfiltrering. Uppdaterade och mer detaljerade instruktioner för att skapa körningar. Rättade stavfel och grammatiska fel. Specificerat att instruktionerna är avsedda för appen när den används med NovaSeq 6000Dx-instrumentet. Uppdaterad information om VCF-utdatafilens innehåll.
200014776 v00	Mars 2022	Första utgåvan.

Dokumentet och dess innehåll tillhör Illumina, Inc. och dess dotterbolag ("Illumina") och är endast avsett för användning enligt avtal i samband med kundens bruk av produkterna som beskrivs häri. Allt annat bruk är förbjudet. Dokumentet och dess innehåll får ej användas eller distribueras i något annat syfte och/eller återges, delges eller reproduceras på något vis utan föregående skriftligt tillstånd från Illumina. I och med detta dokument överlåter Illumina inte någon licens som hör till dess patent, varumärke eller upphovsrätt, eller i enlighet med rättspraxis eller liknande tredjepartsrättigheter.

Instruktionerna i detta dokument ska följas noggrant och uttryckligen av kvalificerad och lämpligt utbildad personal för att säkerställa rätt och säker produktanvändning i enlighet med beskrivningen häri. Hela innehållet i dokumentet ska läsas och förstås i sin helhet innan produkten (produkterna) används.

UNDERLÅTENHET ATT LÄSA OCH FÖLJA ALLA INSTRUKTIONER HÄRI I SIN HELHET KAN MEDFÖRA SKADA PÅ PRODUKTEN/PRODUKTERNA, PERSONSKADA, INKLUSIVE SKADA PÅ ANVÄNDAREN/ANVÄNDARNA ELLER ANDRA PERSONER SAMT SKADA PÅ ANNAN EGENDOM, OCH LEDER TILL ATT EVENTUELL GARANTI FÖR PRODUKTEN/PRODUKTERNA BLIR OGILTIG.

ILLUMINA KAN INTE ÅLÄGGAS NÅGOT ANSVAR SOM UPPKOMMER GENOM FELAKTIG ANVÄNDNING AV PRODUKTERNA SOM BESKRIVS HÄRI (INKLUSIVE DELAR DÄRI ELLER PROGRAM).

© 2022 Illumina, Inc. Med ensamrätt.

Alla varumärken tillhör Illumina, Inc. eller respektive ägare. Specifik varumärkesinformation finns på www.illumina.com/company/legal.html.

Innehållsförteckning

Revisionshistorik	ii
Översikt	1
Analysmetoder	1
Skapa körning	4
Inställningar	5
Analysutdata	7
FASTQ-filer	8
BAM-filer	8
VCF-filer	9
Visa analysresultat	14
Teknisk hjälp	15

Översikt

DRAGEN™ för Illumina® DNA Prep with Enrichment Dx Applikationen utför demultiplexering, FASTQ-generering, läsmappning och anpassning till ett referensgenom och variantanrop beroende på det valda analysarbetsflödet.

Analysmetoder

DRAGEN for Illumina DNA Prep with Enrichment Dx utför demultiplexering, FASTQ-generering, läsmappning och anpassning till ett referensgenom beroende på det valda arbetsflödet:

- FASTQ-generering
- Germline FASTQ och VCF generation
- Somatic FASTQ- och VCF-generering

FASTQ Generation

De sammansatta sekvenserna skrivs till FASTQ-filer per prov. FASTQ-filer är textfiler som innehåller sekvenseringsdata och kvalitetsresultat för endast ett prov. För varje prov genereras separata FASTQ-filer per flödescellfil, per sekvensläsning. Namnet på provet som specificerats under körningsinstallationen ingår i FASTQ-filnamnet. FASTQ-filer är linjeringens primära indata. FDet första steget i FASTQ-genereringen är demultiplexering. Demultiplexering tilldelar kluster som skickar filter till ett sampel genom att jämföra varje indexlässekvens med de indexsekvenser som anges för körningen. Inga kvalitetsvärden beaktas i det här steget. Indexavläsningar identifieras med följande steg:

- Proverna numreras i den ordning som de anges för körningen, med början på 1.
- Provnummer 0 är reserverat för kluster som inte tilldelats till något prov.
- Kluster tilldelas ett prov när indexsekvensen matchar exakt eller när det finns upp till en enda felmatchning per indexläsning.

Programvaran inkluderar ORA-komprimering för att komprimera FASTQ-filer. När du använder formatet ORA (*.ora) bevaras md5-kontrollsumman för FASTQ-innehållet efter en komprimerings- och dekomprimeringscykel för att säkerställa en förlustfri komprimering.

DNA-kartläggning och justering

Det första steget i kartläggningen är att generera frön från avläsningen och sedan leta efter exakta matchningar i referensgenomet. Dessa resultat förfinas sedan genom att köra fullständiga Smith-Waterman-inriktningar på platserna med den högsta tätheten av utsäde. Denna väldokumenterade algoritm fungerar genom att jämföra varje position i avläsningen med alla kandidatpositioner i referensen. Dessa jämförelser motsvarar en matris av potentiella anpassningar mellan läsning och referens. För var och en av dessa kandidatpositioner genererar Smith-Waterman poäng som används

för att utvärdera om den bästa inriktningen som passerar genom den matricellen når den genom en nukleotidmatchning eller mismatchning (diagonal rörelse), en deletion (horisontell rörelse) eller en insertion (vertikal rörelse). En matchning mellan läsning och referens ger en bonus på poängen, och en felmatchning eller indel ger en straff. Den övergripande vägen för högsta poäng genom matrisen är den valda inriktningen.

De specifika värdena som valts för poäng i denna algoritm indikerar hur man balanserar, för en anpassning med flera möjliga tolkningar, möjligheten för en indel i motsats till en eller flera SNP, eller preferensen för en anpassning utan klippning. Standardvärdena för DRAGEN-poäng är rimliga för att anpassa avläsningar av måttlig längd till ett helt mänskligt referensgenom för variantbestämningstillämpningar. Alla Smith-Waterman-poängparametrar representerar en oprecis modell av genomisk mutation och sekvenseringsfel. Olika inställda poängvärden för justering kan vara mer lämpliga för vissa tillämpningar.

DRAGEN Bestämning av könscellvariant

DRAGEN Germline Small Variant Caller tar kartade och justerade DNA-avläsningar som indata och bestämmer SNP och indelar genom en kombination av kolumnvis detektering och lokal *de novo* sammansättning av haplotyper.

Bestämbara referensregioner identifieras först med tillräcklig inriktningstäckning. Inom dessa referensregioner identifierar en snabb skanning av de sorterade läsningarna aktiva regioner, som är centrerade kring pileup-kolonner med bevis på en variant. De aktiva regionerna är utfyllda med tillräckligt med sammanhang för att täcka betydande, icke-referensinnehåll i närheten. Om det finns bevis på indels får de aktiva regionerna ytterligare utfyllnad.

Justerade avläsningar klipps inom varje aktiv region och sätts samman till en De Bruijn-graf. Kanterna på de klippta avläsningarna viktas av observationsräkningar, med referenssekvensen som en ryggrad. Efter viss grafrensning och förenkling extraheras alla källa-till-sänkvägar som kandidathaplotyper. Varje haplotyp är Smith-Waterman anpassad till referensgenomet för att identifiera de varianter den representerar. Denna uppsättning händelser kan utökas med en positionsbaserad detektering. För varje läs-haplotyp par uppskattas sannolikheten $P(r|H)$ för att observera läsningen, förutsatt att haplotypen är det sanna startprovet, med hjälp av en par-dold Markov-modell (HMM).

Genom att skanna genom referensposition över den aktiva regionen bildas kandidatgenotyper från diploida kombinationer av varianthändelser (SNP eller indels). För varje händelse (inklusive referens) uppskattas den villkorade sannolikheten $P(r|e)$ för att observera varje överlappande avläsning som det maximala $P(r|H)$ för haplotyper som stöder händelsen. Dessa kombineras till den villkorliga sannolikheten $P(r|e_1e_2)$ för en genotyp (händelsepar) och multipliceras för att ge den villkorade sannolikheten $P(R|e_1e_2)$ för att observera hela den lästa pileupen. Med hjälp av Bayes formel beräknas den bakre sannolikheten $P(e_1e_2|R)$ för varje diploid genotyp, och vinnaren kallas.

I gVCF-läget som används för skalbar flerprovsvariantbestämning kan DRAGEN Germline Small Variant Caller köras per prov för att generera en mellanliggande genomisk variantbestämning (gVCF). gVCF kan sedan användas för effektiv gemensam genotypning av flera prover, vilket möjliggör snabb inkrementell bearbetning av prover och skalning till stora kohortstorlekar.

Eftersom DRAGEN Germline Small Variant Caller har algoritmer som gör det möjligt att effektivt särskilja korrelerade fel från sanna varianter är filtreringsreglerna mycket enkla.

DRAGEN Somatisk variantbestämning

DRAGEN Somatic Small Variant Caller tar mappade och justerade DNA-avläsningar som indata och bestämmer SNV och indels genom lokal *de novo*-sammansättning av haplotyper i en aktiv region.

Bestämbara referensregioner identifieras först med tillräcklig inriktningstäckning. Inom dessa referensregioner identifierar en skanning av de sorterade läsningarna aktiva regioner, som är centrerade kring pileup-kolonner med bevis på en variant i tumöravläsningarna. De aktiva regionerna är utfyllda med tillräckligt med sammanhang för att täcka betydande, icke-referensinnehåll i närheten. Om det finns bevis på indels får de aktiva regionerna ytterligare utfyllnad.

Justerade avläsningar klipps inom varje aktiv region och sätts samman till en De Bruijn-graf. Kanterna på de klippta avläsningarna viktas av observationsräkningar, med referenssekvensen som en ryggrad. Efter viss grafrensning och förenkling extraheras alla källa-till-sänkvägar som kandidathaplotyper. Varje haplotyp är Smith-Waterman anpassad till referensgenomet för att identifiera de varianter den representerar. För varje läs-haplotyp uppskattas sannolikheten $P(r|H)$ för att observera läsningen med hjälp av en pardold Markov-modell (HMM) under antagande att haplotypen är det sanna startprovet.

För att bestämma TLOD-poängen, skannar DRAGEN Somatic Small Variant Caller först efter referensposition för varje somatisk kandidathändelse såväl som referenshändelsen över den aktiva regionen. Den villkorade sannolikheten $P(r|e)$ för att observera varje överlappande läsning uppskattas som maximala $P(r|H)$ för haplotyper som stödjer händelsen. Dessa kombineras till den villkorliga sannolikheten $P(r|E)$ för en händelsehypotes, E, som involverar en blandning av referens- och kandidatallelen över ett intervall av möjliga allelfrekvenser och multipliceras för att ge den villkorade sannolikheten $P(R|E)$ att observera hela lässtackeln. Därifrån beräknas en TLOD-poäng som bevis på att en ALT-allel finns i tumörprovet på ett givet lokus.

Skapa körning

Använd följande steg för att ställa in Illumina Run Manager inkörning antingen på NovaSeq 6000Dx eller med en webbläsare på en nätverksansluten dator. Provdatabas kan matas in manuellt eller genom att importera ett provark.

Applikations- och körinställningar

1. Från skärmen Körningar väljer du **Create Run** (Skapa körning).
2. Välj DRAGEN for Illumina DNA Prep with Enrichment Dx-appen och sedan **Next** (Nästa).
3. På skärmen Run Settings (Kör inställningar) anger du ett körnamn. Körningsnamnet identifierar körningen från sekvensering till analys.
4. **[Valfritt]** Ange en körbeskrivning för att ytterligare identifiera körningen.
5. Se till att det valda biblioteksförberedande satsen är en Illumina DNA Prep with Enrichment Dx biblioteksförberedande sats.
6. Välj önskad indexadaptersats.
7. Ange läslängden.
Read 1 (Avläsning 1) och Read 2 (Avläsning 2) har ett standardvärde på 151 cykler.
Index 1 och Index 2 har ett fast värde på 10 cykler.
8. **[Valfritt]** Ange ett biblioteksror-ID.
9. Välj **Next** (Nästa).

Provdatabas

Använd tabellen på skärmen Sample Data (Provdatabas) för att ange provinformation manuellt. Alternativt, välj **Import Samples** (Importera Prover) för att ladda upp exempelinformation. För information om att importera exempelinformation, se [Importera Prover på sidan 5](#).

Ange prover manuellt

1. Ange ett unikt prov-ID i fältet Sample ID (Prov-ID).
2. Använd **Plate - Well Position** (platta - brunnposition) för att välja brunnposition.
Fälten i7 Index, Index 1, i5 Index och Index 2 fylls i automatiskt.
3. **[Valfritt]** Ange ett biblioteksnamn.
4. Lägg till rader och upprepa steg [1–3](#) efter behov tills alla prover har lagts till i tabellen.
5. Välj **Next** (Nästa).

Importera Prover

En mall (*.csv) finns tillgänglig för nedladdning på skärmen Sample Data (Prov Data) när du planerar en körning Illumina Run Manager med en webbläsare på en nätverksansluten dator.

1. Välj **Download Template** (Ladda ned mall) för att ladda ner en tom CSV-fil.
2. Från CSV-filen anger du exempelinformationen och sparar filen.
Exempelarkets CVS-fil innehåller följande datakolumner: Prov-ID, platta - brunnposition, **valfritt** biblioteksnamn.
3. Välj **Import Samples** (Importera prover) och bläddra till platsen för CSV-filen.
4. Välj **Next** (Nästa).

Analysinställningar

1. Välj önskat analysarbetsflöde:
 - FASTQ-generering
 - Germline FASTQ och VCF generation för ett könscellarbetsflöde
 - Somatisk FASTQ- och VCF-generering för ett somatiskt arbetsflöde
2. [**Valfritt**] Om så önskas, markera **Generate ORA compressed FASTQs** (kryssrutan Generera ORA-komprimerade FASTQs) för att aktivera FASTQ ORA-komprimering.
3. [**VCF-genereringsarbetsflöden**] Använd **Manifest File Selection** (val av manifestfil) rullgardinsmenyn för att välja en manifestfil.
En manifestfil krävs inmatning för DRAGEN for Illumina DNA Prep with Enrichment Dx. Manifestet är en tabbavgränsad BED-fil (*.bed) som anger namn och platser för riktade referensregioner.
4. [**Somatic FASTQ och VCF generation workflow**] Använd rullgardinsmenyn **Noise File Selection** (Val av brusfil) för att välja en brusfil.
En BED-fil med platsspecifik ljudnivå kan specificeras för att filtrera bort systematiskt brus. För mer information, se [Brusfiltrering på sidan 6](#).
5. Välj **Next** (Nästa).

Körning Granska

1. På Review (gransknings)skärmen granskar du informationen som angetts på skärmarna Run Settings (Körinställningar), Sample Data (Provdatabas) och Analysis Settings (Analysinställningar).
2. Välj **Save** (Spara).
Körningen sparas på fliken Planned (Planerad) på skärmen Runs (Körningar).

Inställningar

Välj appen på skärmen Applications (Applikationer) för att se aktuella inställningar och ändra inställningar.

Konfiguration

Konfigurationsskärmen visar följande programinställningar:

- **Library Prep Kits** (Biblioteksförberedande sats)— Visar appens standardsats för biblioteksförberedelse. Denna inställning kan inte ändras.
- **Index Adapter Kits** (Indexadaptersatser)— Visar appens standardindexadaptersats. Denna inställning kan inte ändras.
- **Read lengths** (Läslängder)— Läslängder är inställda på 151 för appen som standard men kan ändras när körningen skapas.
- **Manifest and Noise Files** (Manifest- och brusfiler)— Ladda upp och ändra inställningar för manifest- och brusfiler.
 - Välj **Upload File** (Ladda upp fil) för att ladda upp filer för användning i analys.
 - Välj knappen **Default** (Standard) för att ställa in filen som standardmanifest eller brusfil som valts under körningen när programmet väljs.
 - Markera **Enabled** (Aktiverad) för att ställa in filen så att den visas i rullgardinsmenyn under körningen.

Behörigheter

Använd kryssrutorna på skärmen Permissions (Behörigheter) för att hantera användaråtkomst för appen.

Brusfiltrering

Systematisk brusfiltrering är tillgänglig när du använder det somatiska arbetsflödet. Filtret kan användas i tumör-normalt läge men är särskilt användbart för körningar med endast tumör där en matchad normal inte är tillgänglig.

Den systematiska buller-BED bör genereras från normala prover. Det rekommenderas att bygga systematiska brusfiler som är specifika för biblioteksförberedelser, sekvenseringssystem och paneler. Det rekommenderas att använda cirka 50 normala sampel för generering av brusfiler.

Analysutdata

DRAGEN for Illumina DNA Prep with Enrichment Dx sparar följande information i analysmappen. Endast könszell- och somatiska arbetsflöden producerar en PDF.

- Manifestfil använd
- Mjukvaruversion
- Sample ID (Prov-ID)
- Totalt justerade läsningar
- Procent av justerade avläsningar per prov
- Antal bestämda SNV per prov
- Antal bestämda indels per prov
- Täckningsstatistik

Analysutdatafiler

Följande utdatafiler genereras av programmet. Exakt vilka filer som genereras beror på vilket analysarbetsflöde som används. Utdatafiler finns i analysmappen.

Utdatafil	Beskrivning
FASTQ (*.fastq.gz eller *.fastq.ora)	Intermediära filer som innehåller basbestämningar med kvalitetspoäng. FASTQ-filer är linjeringsstegets primära indata. Om ORA-komprimering väljs återspeglar filnamnet detta.
Justering av BAM-filer (*.bam)	Innehåller justerade läsningar för ett givet prov.
Genome VCF-filer (*.gvcf.gz)	Innehåller genotypen för varje position, oavsett om den bestäms som en variant eller som en referens.
VCF-filer (*.vcf.gz)	Innehåller varianter som kallas vid varje position.
Kör statistikrapport (*.csv)	Innehåller kvalitetsstatistik om körningen, inklusive total avkastning och Q30-poäng.

FASTQ-filer

FASTQ (*.fastq.gz, *.fastq.ora) är ett textbaserat filformat som innehåller basbestämningar och kvalitetsvärden per läsning. Varje fil innehåller följande information:

- Proidentifierare
- Sekvensen
- Ett plustecken (+)
- Phred-kvalitetspoäng i ett ASCII + 33-kodat format

Proidentifieraren formateras enligt följande.

```
@Instrument:Körnings-ID:Flödescell-ID:Spår:Ruta:X:Y
Avläsningsnummer:Filterflagga:0:Provnummer
Exempel:
@SIM:1:FCX:1:15:6329:1045 1:N:0:2
TCGCACTCAACGCCCTGCATATGACAAGACAGAATC
+
<>;##=><9=AAAAAAAAAAA9#:<#<;<<<????#=#
```

BAM-filer

En BAM-fil (*.bam) är den komprimerade binära versionen av en SAM-fil (sequence alignment map) som används för att representera justerade sekvenser upp till 128 Mb. BAM-filer använder filnamnsformatet `SampleName_S#.bam`. # är provnumret som bestäms av den ordning som proverna listas för körningen. I multinodsläge är S# satt till S1, oavsett ordningen på samplet.

BAM-filer innehåller en rubriksektion och en inriktningssektion:

- Rubrik—Innehåller information om hela filen, som exempelnamn, provlängd och justeringsmetod. Justeringar i inriktningssektionen är associerade med specifik information i rubriksektionen.
- Justeringar—Innehåller läsnamn, lässekvens, läskvalitet, anpassningsinformation och anpassade taggar. Läsnamnet inkluderar kromosomen, startkoordinaten, justeringens kvalitet och matchningsbeskrivningssträngen.

Justeringssektionen innehåller följande information för varje läs- eller läspar:

- AS: Inriktningskvalitet med parad ände.
- RG: Läsgrupp, som anger antalet läsningar för ett specifikt prov.
- BC: Streckkodstagg, som indikerar det demultiplexade prov-ID som är kopplat till läsningen.
- SM: Ensidig inriktningskvalitet.
- XC: Matcha deskriptorsträng.

- XN: Amplicon-namnetikett, som registrerar amplicon-ID som är associerat med de lästa BAM-indexfilerna (*.bam.bai) tillhandahåller ett index för motsvarande BAM-fil.

VCF-filer

Filer med variantbestämningsformat (*.vcf) innehåller information om varianter som finns på specifika positioner i ett referensgenom.

VCF-filhuvudet inkluderar VCF-filformatversionen och variantbestämningsversionen och listar de anteckningar som används i resten av filen. VCF-huvudet inkluderar också referensgenomfilen och BAM-filen. Den sista raden i huvudet innehåller kolumnhuvudet för datalinjerna. Var och en av VCF-filens datarader innehåller information om en enda variant.

Tabell 1 VCF-filrubriker

Rubrik	Beskrivning
CHROM (Kromosom)	Referensgenomets kromosom. Kromosomerna visas i samma ordning som referens-FASTA-filen.
POS	Variantens enkelbasposition i referenskromosomen. För enkelnukleotidvarianter (SNV) är denna position referensbasen med varianten. För indels är denna position referensbasen omedelbart före varianten.
ID	rs (referens SNP) numret för SNP erhållen från <code>dbSNP.txt</code> , om tillämpligt. Om flera rs-nummer finns på den här platsen avgränsas listan med semikolon. Om en dbSNP-post inte finns på denna position används en saknad-markör ('.').
REF	Referensgenotypen. Till exempel representeras en deletion av ett enda T som referens TT och alternativ T. En enkel nukleotidvariant A till T representeras som referens A och alternativ T.
ALT	Allelerna som skiljer sig från referensen läses. Till exempel representeras en insertion av ett enda T som referens A och alternativ AT. En enkelnukleotidvariant A till T representeras som referens A och alternativ T.
QUAL	En Phred-skalad kvalitetspoäng tilldelad av variantbestämmaren. Högre poäng indikerar högre tilltro till varianten och lägre sannolikhet för fel. För kvalitetspoäng Q är den uppskattade sannolikheten för ett fel $10^{-(Q/10)}$. Till exempel har uppsättningen Q30-bestämningar en felfrekvens på 0,1 %. Många variantbestämmare tilldelar kvalitetspoäng baserat på deras statistiska modeller, som är höga i förhållande till den observerade felfrekvensen.

Tabell 2 VCF-filanteckningar

Rubrik	Beskrivning
FILTER	<p>Om alla filter är godkända skrivs PASS i filterkolumnen.</p> <p>Möjliga FILTER-poster i Germline-arbetsflödet inkluderar:</p> <ul style="list-style-type: none"> • DRAGENSnpHardQUAL—Tillämpas om SNP-variant QUAL-poäng inte når tröskeln • DRAGENIndelHardQUAL—Tillämpas om indel variant QUAL poäng inte når tröskeln • LowDepth—Webbplatsen filtrerades eftersom täckningsdjupet inte når tröskeln • LowGQ—Platsen filtrerades eftersom genotypkvaliteten inte når tröskeln • PloidyConflict—Genotypbestämning från variantbestämare överensstämmer inte med kromosomploidi • base_quality—Platsen filtrerades eftersom medianbaskvaliteten för alt-läsningar på detta lokus inte når tröskeln • filtered_reads—Platsen har filtrerats eftersom en för stor del av läsningarna har filtrerats bort • fragment_length—Platsen filtrerades eftersom den absoluta skillnaden mellan medianfragmentlängden för alt-läsningar och medianfragmentlängden för ref-läsningar på detta ställe överskrider tröskeln • low_depth—Platsen har filtrerats eftersom läsdjupet är för lågt • low_frac_info_reads—Platsen filtrerades eftersom andelen informativa läsningar ligger under tröskeln • low_normal_depth—Platsen filtreras eftersom det normala provets läsdjup är för lågt • long_indel—Platsen filtreras eftersom indellängden är för lång • mapping_quality—Platsen filtreras eftersom medianmappingskvaliteten för alt-läsningar på detta lokus inte når tröskeln • multiallelic—Platsen filtreras eftersom mer än två alt-alleler passerar tumör-LOD • non_homref_normal—Platsen filtrerades eftersom den normala provgenotypen inte är homozygot referens • no_reliable_supporting_read—Platsen filtrerades eftersom det inte finns någon pålitlig stödjande somatisk läsning • panel_of_normals—Ses i minst ett prov i panelen av normaler vcf • read_position—Platsen har filtrerats eftersom medianen för avstånden mellan start/slut av läsning och detta lokus ligger under tröskeln • RMxNRepeatRegion—Platsen filtrerades eftersom hela eller delar av variantallelen är en upprepning av referensen • strand_artifact Platsen har filtrerats på grund av allvarlig strängbias • str_contraction—Platsen filtrerad på grund av misstänkt PCR-fel där alt-allelen är en upprepad enhet mindre än referensen • too_few_supporting_reads—Platsen filtrerades eftersom det finns för få stödjande avläsningar i tumörprovet • weak_evidence—Somatisk variantpoäng når inte tröskeln

Rubrik	Beskrivning
FILTER (fortsättning)	<p>Möjliga FILTER-poster för somatiskt arbetsflöde inkluderar:</p> <ul style="list-style-type: none"> • base_quality—Platsen filtrerades eftersom medianbaskvaliteten för alt-läsningar på detta lokus inte når tröskeln • filtered_reads—Platsen har filtrerats eftersom en för stor del av läsningarna har filtrerats bort • fragment_length—Platsen filtrerades eftersom den absoluta skillnaden mellan medianfragmentlängden för alt-läsningar och medianfragmentlängden för ref-läsningar på detta ställe överskrider tröskeln • low_depth—Platsen har filtrerats eftersom läsdjupet är för lågt • low_frac_info_reads—Platsen filtrerades eftersom andelen informativa läsningar ligger under tröskeln • low_normal_depth—Platsen filtreras eftersom det normala provets läsdjup är för lågt • long_indel—Platsen filtreras eftersom indellängden är för lång • mapping_quality—Platsen filtreras eftersom medianmappningskvaliteten för alt-läsningar på detta lokus inte når tröskeln • multiallelic—Platsen filtreras eftersom mer än två alt-alleler passerar tumör-LOD • non_homref_normal—Platsen filtrerades eftersom den normala provgenotypen inte är homozygot referens • no_reliable_supporting_read—Platsen filtrerades eftersom det inte finns någon pålitlig stödjande somatisk läsning • panel_of_normals—Ses i minst ett prov i panelen av normaler vcf • read_position—Platsen har filtrerats eftersom medianen för avstånden mellan start/slut av läsning och detta lokus ligger under tröskeln • RMxNRepeatRegion—Platsen filtrerades eftersom hela eller delar av variantallelen är en upprepning av referensen • strand_artifact Platsen har filtrerats på grund av allvarlig strängbias • str_contraction—Platsen filtrerad på grund av misstänkt PCR-fel där alt-allelen är en upprepning av enhet mindre än referensen • too_few_supporting_reads—Platsen filtrerades eftersom det finns för få stödjande avläsningar i tumörprovet • weak_evidence—Somatisk variantpoäng når inte tröskeln • systematic_noise—Platsen filtreras baserat på tecken på systematiskt brus i normala

Rubrik	Beskrivning
INFO	<p>Möjliga INFO-poster för Germline arbetsflöde inkluderar:</p> <ul style="list-style-type: none"> • AC—Alleler i genotyper för varje ALT-allel, i samma ordning som anges. • AF—Allelfrekvens för varje ALT-allel, i samma ordning som anges. • AN—Det totala antalet alleler i kallade genotyper. • DB—dbSNP-medlemskap. • FS—Phred-skalat p-värde med Fishers exakta test för att upptäcka strängförspänning. • QD—Varianttillförlitlighet/-kvalitet efter djup. • R2_5P_bias—Poäng baserat på parningspartiskhet och avstånd från 5 primeslut. • SOR—Symmetrisk oddskvot på 2x2 beredskapstabell för att upptäcka strängförspänning. • DP—Ungefärligt läsdjup (informativt och icke-informativt); vissa läsningar kan ha filterats baserat på mapq etc. • END—Stopp position för intervallet. • FractionInformativeReads—Bråkdelen av informativa läsningar av det totala antalet läsningar. • MQ—RMS Kartläggningens kvalitet. • MQRankSum—Z-poäng Från Wilcoxon rangsummetest av Alt vs Ref läs mappningskvaliteter. • ReadPosRankSum—Z-poäng från Wilcoxon rangsummetest av Alt vs. Ref läs positionsbias. • SOMATIC—Minst en variant på denna position är somatisk. <p>Möjliga INFO-poster i somatiskt arbetsflöde inkluderar:</p> <ul style="list-style-type: none"> • DP—Ungefärligt läsdjup (informativt och icke-informativt); vissa läsningar kan ha filterats baserat på mapq etc. • END—Stopp position för intervallet. • FractionInformativeReads—Bråkdelen av informativa läsningar av det totala antalet läsningar. • MQ—RMS Kartläggningens kvalitet. • MQRankSum—Z-poäng Från Wilcoxon rangsummetest av Alt vs Ref läs mappningskvaliteter. • ReadPosRankSum—Z-poäng från Wilcoxon rangsummetest av Alt vs. Ref läs positionsbias. • AQ—Systematisk bruspoäng. • hotspot—Känd somatisk sida, används för att öka bestämningens tillförlitlighet. • SOMATIC—Minst en variant på denna position är somatisk.

Rubrik	Beskrivning
FORMAT	<p>Formatkolumnen visar fält separerade med kolon. Till exempel GT:GQ.</p> <p>Tillgängliga fält för Germline arbetsflöde inkluderar:</p> <ul style="list-style-type: none"> • AD—Alleldjup (räknas endast informativa läsningar av de totala läsningarna) för ref- och alt-allelerna i den ordning som anges. • AF—Allelfraktioner för alt-alleler i den ordning som anges. • DP—Ungefärligt läsdjup (avläsningar med MQ=255 eller med dåliga kompisar filtreras). • F1R2—Antal avläsningar i F1R2-parorientering som stöder varje allel. • F2R1—Antal avläsningar i F2R1-parorientering som stöder varje allel. • GP—Phred-skalade posteriora sannolikheter för genotyper enligt definitionen i VCF-specifikationen. • GQ—Genotyp kvalitet. • GT—Genotyp. 0 motsvarar referensbasen, 1 motsvarar den första posten i ALT-kolumnen och så vidare. Snedstreck (/) indikerar att ingen fasinformation är tillgänglig. • MB—Per-prov komponentstatistik för att upptäcka parningspartiskhet. • PL—Normaliserade, Phred-skalade sannolikheter för genotyper enligt definitionen i VCF-specifikationen. • PRI—Phred-skalade tidigare sannolikheter för genotyper. • PS—Fysisk fas-ID-information, där varje unikt ID inom ett givet prov (men inte över prover) kopplar samman poster inom en fasningsgrupp. • SB—Per-prov komponentstatistik som omfattar Fishers Exact Test för att upptäcka strängbias. • SQ—Somatisk kvalitet. <p>Tillgängliga fält för somatiskt arbetsflöde inkluderar:</p> <ul style="list-style-type: none"> • AD—Alleldjup (räknas endast informativa läsningar av de totala läsningarna) för ref- och alt-allelerna i den ordning som anges. • AF—Allelfraktioner för alt-alleler i den ordning som anges. • DP—Ungefärligt läsdjup (avläsningar med MQ=255 eller med dåliga kompisar filtreras). • F1R2—Antal avläsningar i F1R2-parorientering som stöder varje allel. • F2R1—Antal avläsningar i F2R1-parorientering som stöder varje allel. • GT—Genotyp. 0 motsvarar referensbasen, 1 motsvarar den första posten i ALT-kolumnen och så vidare. Snedstreck (/) indikerar att ingen fasinformation är tillgänglig. • MB—Per-prov komponentstatistik för att upptäcka parningspartiskhet. • PS—Fysisk fas-ID-information, där varje unikt ID inom ett givet prov (men inte över prover) kopplar samman poster inom en fasningsgrupp. • SB—Per-prov komponentstatistik som omfattar Fishers Exact Test för att upptäcka strängbias. • SQ—Somatisk kvalitet.

Rubrik	Beskrivning
PROV	Exempelkolumnen ger de värden som anges i FORMAT-kolumnen.

Genome VCF-filer

Genome VCF-filer (*.gvcf.gz) följer en uppsättning konventioner för att representera alla platser inom genomet i ett relativt kompakt format. gVCF-filerna inkluderar alla platser inom området av intresse i en enda fil för varje prov. gVCF-filen visar no-calls på positioner som inte klarar alla filter. En genotyp-tag (GT) av ./ indikerar ett no-call.

Visa analysresultat

Pågående körningar visas på fliken Aktiv. Slutförda körningar visas på fliken Slutförda. Se [NovaSeq 6000Dx produktdokumentation \(dokument nr 200010105\)](#) för mer information om visning av resultat.

Teknisk hjälp

För teknisk hjälp, kontakta Illumina Teknisk Support.

Webbplats: www.illumina.com

E-post: techsupport@illumina.com

Illumina Telefonnummer för teknisk Support

Region	Avgiftsfritt	Internationellt
Australien	+61 1800 775 688	
Österrike	+43 800 006249	+43 1 9286540
Belgien	+32 800 77 160	+32 3 400 29 73
Kanada	+1 800 809 4566	
Kina		+86 400 066 5835
Danmark	+45 80 82 01 83	+45 89 87 11 56
Finland	+358 800 918 363	+358 9 7479 0110
Frankrike	+33 8 05 10 21 93	+33 1 70 77 04 46
Tyskland	+49 800 101 4940	+49 89 3803 5677
Hongkong, Kina	+852 800 960 230	
Indien	+91 8006500375	
Indonesien		0078036510048
Irland	+353 1800 936608	+353 1 695 0506
Italien	+39 800 985513	+39 236003759
Japan	+81 0800 111 5011	
Malaysia	+60 1800 80 6789	
Nederländerna	+31 800 022 2493	+31 20 713 2960
Nya Zeeland	+64 800 451 650	
Norge	+47 800 16 836	+47 21 93 96 93
Filippinerna	+63 180016510798	
Singapore	1 800 5792 745	
Sydkorea	+82 80 234 5300	
Spanien	+34 800 300 143	+34 911 899 417

Region	Avgiftsfritt	Internationellt
Sverige	+46 2 00883979	+46 8 50619671
Schweiz	+41 800 200 442	+41 56 580 00 00
Taiwan, Kina	+886 8 06651752	
Thailand	+66 1800 011 304	
Storbritannien	+44 800 012 6019	+44 20 7305 7197
USA	+1 800 809 4566	+1 858 202 4566
Vietnam	+84 1206 5263	

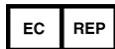
Säkerhetsdatablad (SDS)—Finns på Illuminas webbplats på support.illumina.com/sds.html.

Produktdokumentation—Kan hämtas på support.illumina.com.



Illumina
5200 Illumina Way
San Diego, California 92122 USA
+1.800.809.ILMN (4566)
+1 858 202 4566 (utanför Nordamerika)
techsupport@illumina.com
www.illumina.com

CE



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

Australisk sponsor

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

FÖR IN VITRO-DIAGNOSTISK ANVÄNDNING

© 2022 Illumina, Inc. Med ensamrätt.

illumina[®]