

Local Run Manager Germline Variant-analysemodul

Arbeidsflytveiledning for MiSeqDx

TIL IN VITRO-DIAGNOSTISK BRUK

Oversikt	3
Oppgi kjøringsinformasjon	3
Analysemetoder	5
Vise kjørings- og prøvedata	6
Analyserapport	6
Analyseutgangsfiler	8
Teknisk hjelp	14



Dette dokumentet og dets innhold er opphavsrettslig beskyttet for Illumina, Inc. og tilknyttede selskaper («Illumina»), og er ment utelukkende for kontraktbruk av sin kunde i forbindelse med bruk av produktet (produktene) beskrevet her, og for intet annet formål. Dette dokumentet og dets innhold skal ikke brukes eller distribueres til andre formål og/eller på annen måte kommuniseres, fremlegges eller reproduseres på noen måte uten forutgående, skriftlig samtykke fra Illumina. Illumina fører ikke noen lisens under sin patent, varemerke, opphavsrett eller sedvanerett eller lignende rettigheter til tredjeparter ved dette dokumentet.

Instruksjonene i dette dokumentet skal være strengt og tydelig fulgt av kvalifisert og tilfredsstillende utdannet personell for å sikre riktig og sikker bruk av produktet (produktene) som er beskrevet i dette dokumentet. Alt innhold i dette dokumentet skal leses fullt ut og være forstått før produktet (produktene) brukes.

HVIS DET UNNLATES Å FULLSTENDIG LESE OG UTTRYKkelig FØLGE ALLE INSTRUKSJONENE I DETTE DOKUMENTET, KAN DETTE FØRE TIL SKADE PÅ PRODUKTET (PRODUKTENE), SKADE PÅ PERSONER, INKLUDERT BRUKERE ELLER ANDRE, OG SKADE PÅ ANNEN EIENDOM.

ILLUMINA PÅTAR SEG IKKE ANSVAR SOM FØLGE AV FEIL BRUK AV PRODUKTET (PRODUKTENE) SOM ER BESKREVET I DETTE DOKUMENTET (INKLUDERT DELER AV DETTE ELLER PROGRAMVARE).

© 2021 Illumina, Inc. Med enerett.

Illumina, MiSeqDx, og streamingbase-designen er registrerte eller ventende varemerker som tilhører Illumina, Inc. og/eller tilknyttede selskaper i USA og/eller andre land. Alle andre navn, logoer og andre varemerker tilhører deres respektive eiere.

Oversikt

Local Run Manager Germline Variant-analysemodulen er til bruk med Illumina TruSeq Custom Amplicon Kit Dx-analysen. Når analysen brukes med Germline Variant-modulen, er den ment for klargjøring av biblioteker som anvendes til sekvensering av DNA fra perifere fullblodprøver.

Analysemodulen evaluerer korte regioner av amplifisert DNA, eller PCR-produkter, for varianter. Fokuset på sekvensering av PCR-produkter muliggjør høy dekning av bestemte regioner. Se pakningsvedlegget *TruSeq tilpasset PCR-produktsett-Dx* (dokumentnr. 1000000029772).

Germline Variant-analysemodulen krever sekvenseringsforbruksmaterialer fra MiSeqDx Reagent Kit v3. Se pakningsvedlegget *MiSeqDx Reagent Kit v3* (dokumentnr. 1000000030849).

Om denne veiledningen

Denne veiledningen gir instruksjoner for å konfigurere kjøringsparametere for sekvensering og analyse for analysemodulen Germline Variant. Informasjon om dashbord- og systeminnstillinger for Local Run Manager finnes i *Referansehåndbok for Local Run Manager-programvare for MiSeqDx* (dokumentnr. 1000000011880).

Visning for Local Run Manager

Oppgi kjøringsinformasjon

Angi parametere

- 1 Logg på Local Run Manager.
- 2 Klikk på **Create Run** (Opprett kjøring), og velg **Germline Variant**.
- 3 Skriv inn et kjøringsnavn som identifiserer kjøringen fra sekvensering til og med analyse. Bruk alfanumeriske tegn, mellomrom, understreker eller bindestreker.
- 4 [Valgfritt] Angi en kjøringsbeskrivelse for å identifisere kjøringen. Bruk alfanumeriske tegn, mellomrom, understreker eller bindestreker.

Importere manifestfiler for kjøringen

- 1 Kontroller at manifestene du vil importere, er tilgjengelige på en tilgjengelig nettverksplassering eller på en USB-stasjon.
- 2 Klikk på **Import Manifests** (Importer manifeste).
- 3 Naviger til manifestfilen, og velg manifestene du vil legge til.



MERK

For å gjøre manifestfiler tilgjengelige for alle kjøring ved hjelp av Germline Variant-analysemodulen legger du til manifestene ved hjelp av modulinnstillingsfunksjonen. Denne funksjonen krever administratorrettigheter. Se *Referansehåndbok for Local Run Manager-programvare for MiSeqDx* (dokumentnr. 1000000011880) for mer informasjon.

Angi prøver for kjøringen

Angi prøver for kjøringen ved å bruke ett av alternativene og instruksjonene som følger.

- ▶ **Enter samples manually** (Angi prøver manuelt) – Bruk den tomme tabellen i skjermbildet Create Run (Opprett kjøring).
- ▶ **Import samples** (Importer prøver) – Naviger til en ekstern fil i et format med kommaseparerte verdier (*.csv). En mal er tilgjengelig for nedlasting i skjermbildet Create Run (Opprett kjøring).

Etter at du har fylt ut prøvetabellen, kan du eksportere prøveinformasjonen til en ekstern fil og bruke filen som referanse når du klargjør bibliotek, eller importere filen til en annen kjøring.

Angi prøver manuelt

- 1 Velg antall prøver og indekssett fra rullegardinlisten.
- 2 Skriv inn en unik prøve-ID i feltet Sample ID (Prøve-ID).
Bruk alfanumeriske tegn, bindestreker eller understreker.
- 3 [Valgfritt] For positive eller negative kontrollprøver: Høyreklikk og velg kontrolltypen.
- 4 [Valgfritt] Skriv inn en prøvebeskrivelse i feltet Sample Description (Prøvebeskrivelse).
Bruk alfanumeriske tegn, bindestreker eller understreker.
Prøvebeskrivelser er knyttet til et prøvenavn. Prøvebeskrivelser overskrives hvis samme prøvenavn brukes igjen i en senere kjøring.
- 5 Velg en indeks 1-adapter fra rullegardinlisten Index 1 (i7) (Indeks 1 (i7)).
- 6 Velg en indeks 2-adapter fra rullegardinlisten Index 2 (i5) (Indeks 2 (i5)).
- 7 Velg en manifestfil fra rullegardinlisten Manifest.
- 8 Velg et alternativ for å vise, skrive ut eller lagre plateoppsettet som en referanse for klargjøring av biblioteker:
 - ▶ Klikk på ikonet  **Print** (Skriv ut) for å vise plateoppsettet. Velg **Print** (Skriv ut) for å skrive ut plateoppsettet.
 - ▶ Klikk på **Export** (Eksporter) for å eksportere prøveinformasjon til en ekstern fil.
Kontroller at manifest- og prøveinformasjon er riktig. Feil informasjon kan påvirke resultatene.
- 9 Klikk på **Save Run** (Lagre kjøring).

Importere prøver

- 1 Klikk på **Import Samples** (Importer prøver), og bla til plasseringen av prøveinformasjonsfilen. Det finnes to typer filer du kan importere.
 - ▶ Klikk på **Template** (Mal) i skjermbildet Create Run (Opprett kjøring) for å lage et nytt plateoppsett. Malfilen inneholder de riktige kolonneoverskriftene for import. Skriv inn prøveinformasjon i hver kolonne for prøvene i kjøringen. Slett eksempelinformasjon i ubrukte celler, og lagre deretter filen.
 - ▶ Bruk en fil med prøveinformasjon som ble eksportert fra Germline Variant-modulen ved hjelp av funksjonen Export (Eksport).
- 2 Klikk på ikonet  **Print** (Skriv ut) for å vise plateoppsettet.
- 3 Velg **Print** (Skriv ut) for å skrive ut plateoppsettet som en referanse for klargjøring av biblioteker.
- 4 [Valgfritt] Klikk på **Export** (Eksporter) for å eksportere prøveinformasjon til en ekstern fil.
Kontroller at manifest- og prøveinformasjon er riktig. Feil informasjon kan påvirke resultatene.
- 5 Klikk på **Save Run** (Lagre kjøring).

Redigere en kjøring

Du finner instruksjoner om redigering av informasjonen i kjøringen før sekvensering i *Referansehåndbok for Local Run Manager-programvare for MiSeqDx (dokumentnr. 1000000011880)*.

Analysemetoder

Germline Variant-analysemodulen utfører følgende analysetrinn og skriver deretter analyseutgangsfiler til mappen Alignment (Innretting).

- ▶ Demultiplekser indeksavlesninger
- ▶ Genererer FASTQ-filer
- ▶ Innretter etter en referanse
- ▶ Identifiserer varianter

Demultipleksing

Demultipleksing sammenligner hver indeksavlesningssekvens med indekssekvensene som er angitt for kjøringen. Ingen kvalitetsverdier vurderes på dette trinnet.

Indeksavlesninger identifiseres ved hjelp av følgende trinn:

- ▶ Prøvene er nummerert fra 1 basert på rekkefølgen de er oppført i for kjøringen.
- ▶ Prøvenummer 0 er reservert for klynger som ikke ble tildelt noen prøve.
- ▶ Klynger tilordnes en prøve når indekssekvensen samsvarer nøyaktig, eller når det er opptil ett misforhold per indeksavlesning.

FASTQ-filgenerering

Etter demultipleksing genererer programvaren mellomliggende analysefiler i FASTQ-formatet, som er et tekstformat som brukes til å representere sekvenser. FASTQ-filer inneholder avlesninger for hver prøve og tilhørende kvalitetsscore. Klynger som ikke passerte filter, utelukkes.

Hver FASTQ-fil inneholder avlesninger for bare én prøve, og navnet på denne prøven er inkludert i FASTQ-filnavnet. FASTQ-filer er primære inndata for innretting. To FASTQ-filer genereres per prøve, én fra Read 1 og én fra Read 2.

Innretting

Under innrettingstrinnet innretter den bundne Smith-Waterman-algoritmen klynger fra hver prøve med PCR-produktsekvenser spesifisert i manifestfilen.

Den bundne Smith-Waterman-algoritmen utfører halvglobale sekvensinnrettinger for å bestemme lignende regioner mellom 2 sekvenser. I stedet for å sammenligne den totale sekvensen, sammenligner Smith-Waterman-algoritmen segmenter med alle mulige lengder.

Hver paired-end-avlesning blir evaluert med hensyn til sin innretting med de relevante probesekvensene for denne avlesningen.

- ▶ Read 1 evalueres mot reverskomplementet av nedstrøms lokusspesifikke oligoer (DLSO).
- ▶ Read 2 blir vurdert opp mot oppstrøms lokusspesifikke oligoer (ULSO).
- ▶ Hvis starten på en avlesning samsvarer med en probesekvens med ikke mer enn 3 ulikheter (misforhold eller forskyvninger på grunn av ledende indeler), innrettes den fulle lengden av avlesningen mot PCR-produktmålet for den sekvensen.

- Indeler innenfor DLSO og ULSO blir ikke observert på grunn av analysekjemien.

Innrettinger filtreres fra innrettingsresultater basert på antall misforhold over enten interesseområdet eller hele PCR-produktet, avhengig av PCR-produktlengden. Filtrerte innrettinger registreres i innrettingsfiler som uinnrettet og brukes ikke i variantbetegnelse.

Variantbetegnelse

Pisces Variant Caller er utviklet av Illumina og identifiserer varianter som er til stede i DNA-prøven.

Pisces variantbetegner identifiserer SNV-er, MNV-er og små indeler i tre trinn:

- Vurderer hver posisjon i referansegenomet separat
- Teller baser på den angitte posisjonen for innrettede avlesninger som overlapper posisjonen
- Beregner en variantscore som måler kvaliteten på betegnelsen ved hjelp av Poisson-modellen. Varianter med kvalitetsscore under Q20 utelates.

Hvis en variant passerer alle filtre, merkes varianten som PASS i VCF.

Se github.com/Illumina/Pisces/wiki for mer informasjon.

Vise kjørings- og prøvedata

- 1 Fra Local Run Manager-dashbordet klikker du på kjøringsnavnet.
- 2 Fra fanen Run Overview (Kjøringsoversikt) kan du gjennomgå sekvenseringskjøringsmetrikk.
- 3 [Valgfritt] Klikk på ikonet **Copy to Clipboard**  (Kopier til utklippstavle) for å kopiere banen til utgangskjøringsmappen.
- 4 Klikk på fanen Sequencing Information (Sekvenseringsinformasjon) for å se informasjon om kjøringsparametere og forbruksvarer.
- 5 Klikk på fanen Samples and Results (Prøver og resultater) for å vise analyserapportens plassering.
 - Hvis analysen ble gjentatt, utvider du rullegardinlisten Select Analysis (Velg analyse) og velger den aktuelle analysen.
- 6 Klikk på ikonet **Copy to Clipboard**  (Kopier til utklippstavle) for å kopiere banen til mappen Analysis (Analyse).

Dersom du ønsker mer informasjon om fanene Run Overview (Kjøringsoversikt) og Sequencing Information (Sekvenseringsinformasjon) og hvordan du skal stille en analyse tilbake i køen, se *Referansehåndbok for Local Run Manager-programvare for MiSeqDx* (dokumentnr. 1000000011880).

Analyserapport

Analyseresultater oppsummeres på fanen Samples and Results (Prøver og resultater) og som en samlet rapport i mappen Alignment (Innretting). En rapport for hver prøve er også tilgjengelig i PDF-filformat for hver prøve.

Informasjon om fanen Sample and Details (Prøve og detaljer)

- 1 Klikk på en prøve i listen for å se prøverapporten.

Tabell 1 Tabellen Run Information (Kjøringsinformasjon)

Kolonneoverskrift	Beskrivelse
Run Status (Kjøringsstatus)	Angir om sekvenseringskjøringen var godkjent eller mislyktes.
Total Yield (GB) (Total produksjon (GB))	Antall baser betegnet i sekvenseringskjøringen. Viser grenseverdi for godkjenning, og status (godkjent eller mislykket).
% ≥ Q30	Prosentandelen av avlesninger i sekvenseringskjøringen med en kvalitetsscore på 30 (Q30) eller høyere. Viser grenseverdi for godkjenning, og status (godkjent eller mislykket).
Sample Name (Prøvenavn)	Prøvenavnet som ble gitt da kjøringen ble opprettet.
Total PF Reads (PF-avlesninger totalt)	Totalt antall avlesninger som passerer filteret.
Read 1% ≥ Q30	Prosentandelen av avlesninger i Read 1 med en kvalitetsscore på 30 (Q30) eller høyere for prøven.
Read 2% ≥ Q30	Prosentandelen av avlesninger i Read 2 med en kvalitetsscore på 30 (Q30) eller høyere for prøven.
Autosome call rate (Betegnelsesfrekvens for autosomer)	Antallet genomiske posisjoner blant autosomene (kromosom 1 til 22) som oppfyller en forhåndsdefinert konfidensverditerskel, dividert med det totale antallet autosomale genomiske posisjoner som behandles. Betegnelsesfrekvensen på per prøve-basis og rapportert som prosentandel som er beregnet som 1 minus (antall autosomale posisjoner med ufullstendige betegnelser, dividert med totalt antall sekvenserte autosomale posisjoner).

Tabell 2 Tabellen Sample Reports (Prøverapporter)

Kolonneoverskrift	Beskrivelse
Sample (Prøve)	Prøve-ID-en som ble gitt da kjøringen ble opprettet.
Report Date (Rapportdato)	Datoen rapporten ble generert.
Sample Information (Prøveinformasjon)	Prøve-ID-en som ble oppgitt da kjøringen ble opprettet, totalt antall avlesninger som passerte filteret i prøven, prosentandelen av avlesninger for prøven med en kvalitetsscore på 30 (Q30) eller høyere, og den autosomale betegnelsesfrekvensen.
Amplicon Summary (PCR-produktsammendrag)	Totalt antall sekvenserte PCR-produktregioner, og total lengde i basepar av sekvenserte PCR-produkter i målregionene, for prøven og manifestfilen. Manifestfilen angir referansegenomet og målrettede referanseregioner som brukes i innrettingstrinnet.
Read Level Statistics (Statistikk på avlesningsnivå)	Antall og prosent av avlesninger for prøven som dekker hver posisjon i referansen, for Read 1 og Read 2.
Variants Summary (Variantsammendrag)	Antall SNV-er, innsettinger og slettinger oppdaget for prøven som passerte foreslåtte verdier, for å avgjøre om kvalitetsresultatene er innenfor et akseptabelt område.
Coverage Summary (Dekningssammendrag)	Det totale antallet justerte baser dividert med målrettet regionstørrelse og prosentandelen av PCR-produktregioner med dekningverdier som er større enn den lave dekningsterskelen på 0,2* PCR-produktdekning, for prøven.
Coverage Plots (Dekningsplott)	Dekningen av PCR-produktregionplottet viser dekingen over PCR-produktregionene for prøven. Regioner med dekningverdier som er lavere enn dekningsterskelen, er markert i rødt. Gjennomsnittet av alle verdiene er angitt med en oransje linje.
Software Versions (Programwareversjoner)	Programwareversjoner da prøven ble sekvensert. Inkluderer MiSeq-driftsprogramvare (MOS), Local Run Manager Software, RTA Software, og Germline Variant Modul-versjonene.

Analyseutgangsfiler

Følgende analyseutdatafiler genereres for Germline Variant-analysemodulen og gir analyseresultater for innretting og variantbetegnelse. Analyseutgangsfiler finnes i mappen Alignment (Innretting).

Filnavn	Beskrivelse
Demultipleksing (*.txt)	Mellomliggende filer som inneholder sammendrag av resultater fra demultipleksing.
FASTQ (*.fastq.gz)	Mellomliggende filer som inneholder kvalitetsscorede basebetegnelser. FASTQ-filer er primære inndata for innrettingstrinnet.
Innrettingsfiler i BAM-format (*.bam)	Inneholder innrettede avlesinger for en gitt prøve.
Variantbetegnelsesfiler i genom-VCF-format (*.genome.vcf.gz)	Inneholder genotypen for hver posisjon, enten den betegnes som en variant eller som en referanse.
Variantbetegnelsesfiler i VCF-formatet (*.vcf.gz)	Inneholder alle varianter betegnet over målområdet.
AmpliconCoverage_M1.tsv	Inneholder informasjon om dekning per PCR-produkt per prøve for hvert oppgitte manifest. M# angir manifesttallet.

Filformat for demultipleksing

Demultipleksingprosessen leser indekssekvensen for hver klynge for å bestemme hvilken prøve klyngen oppsto fra. Kartleggingen mellom klynger og prøvenummer skrives til en demultipleksingsfil (*.demux) for hver plate av strømningscellen.

Filnavnet for demultipleksingsfiler er s_1_X.demux, hvor X er platenummeret.

Demultipleksingsfiler starter med en overskrift:

- ▶ Versjon (4 byte heltall), for tiden 1
- ▶ Klyngetall (4 byte heltall)

Resten av filen består av prøvetall for hver klynge fra platen.

Når demultipleksingstrinnet er fullført, genererer programvaren en demultipleksingsfil som heter DemultiplexSummaryF1L1.txt.

- ▶ I filnavnet representerer **F1** strømningscellenummeret.
- ▶ I filnavnet representerer **L1** banenummeret.
- ▶ Demultipleksingsresultater i en tabell med én rad per plate og én kolonne per prøve, inkludert prøve 0.
- ▶ De vanligste sekvensene i indeksavlesninger.

FASTQ-filformat

FASTQ er et tekstbasert filformat som inneholder basebetegnelser og kvalitetsverdier per avlesning. Hver oppføring inneholder 4 linjer:

- ▶ Identifikatoren
- ▶ Sekvensen
- ▶ Et plusstegn (+)
- ▶ Phred-kvalitetsresultatene i et ASCII + 33-kodet format

Identifikatoren har formatet:

@Instrument:Kjørings-ID:Strømningscelle-ID:Bane:Plate:X:Y Avlesningsnr.:FilterFlagg:0:Prøvenummer

Eksempel:

```
@SIM:1:FCX:1:15:6329:1045 1:N:0:2
TCGCACTCAACGCCCTGCATATGACAAGACAGAATC
+
<>;##=><9=AAAAAAAAA9#:<#<;<<<????#=#
```

BAM-filformat

En BAM-fil (*.bam) er en komprimert binær versjon av en SAM-fil som brukes til å angi innrettede sekvenser opptil 128 Mb. SAM- og BAM-formater beskrives i detalj på samtools.github.io/hts-specs/SAMv1.pdf.

BAM-filer bruker filnavngivningsformatet SampleName_#.#.bam, der # er prøvenummeret som bestemt av rekkefølgen av prøvene som er oppført for kjøringen.

BAM-filer inneholder et overskriftsavsnitt og et innrettingsavsnitt:

- ▶ **Header** (Overskrift) – Inneholder informasjon om hele filen, som prøvenavn, prøvelengde og innrettingsmetode. Innrettinger i innrettingsavsnittet er forbundet med spesifikk informasjon i overskriftsavsnittet.
- ▶ **Alignments** (Innrettinger) – Inneholder avlesningsnavn, avlesningssekvens, avlesningskvalitet, innrettingsinformasjon og tilpassede faner. Avlesningsnavnet inkluderer kromosomet, startkoordinat, innrettingskvalitet og samsvarsindikatorstreng.

Innrettingsavsnittet inneholder følgende informasjon for hver avlesing eller hvert avlesningspar:

- ▶ **AS:** Paired-end innrettingskvalitet
- ▶ **RG:** Avlesingsgruppe, som indikerer antall avlesinger for en bestemt prøve.
- ▶ **BC:** Strekkode-fane, som indikerer den demultipleksede prøve-ID-en som er knyttet til avlesningen.
- ▶ **SM:** Single-end innrettingskvalitet.
- ▶ **XC:** Samsvarsindikatorstreng
- ▶ **XN:** Navnefaner for PCR-produkt, som registrerer PCR-produkt-ID-en som er knyttet til avlesningen

BAM-indeksfiler (*.bam.bai) gir en indeks av den tilsvarende BAM-filen.

VCF-filformat

Variant Call Format (VCF) er et vanlig filformat utviklet av det genomikkvitenskapelige samfunnet.

Det inneholder informasjon om varianter funnet i spesifikke posisjoner i et referansegеном. VCF-filer slutter med .vcf-suffikset

VCF-filoverskriften inneholder VCF-filformatversjonen og variantbetegnerversjonen og angir kommentarene som brukes i resten av filen. VCF-overskriften inneholder også referansegеномfilen og BAM-filen. Den siste linjen i overskriften inneholder kolonneoverskriftene for datalinjene. Hver av VCF-fildatalinjene inneholder informasjon om én variant.

VCF-filoverskrifter

Overskrift	Beskrivelse
KROM	Kromosomet på referanseggenomet. Kromosomer vises i samme rekkefølge som referanse-FASTQ-filen.
POS	Enkeltbaseposisjonen til varianten i referansekromosomet. I SNP-er er denne posisjonen referansebasen med varianten, i indeler eller slettinger er denne posisjonen referansebasen umiddelbart før varianten.
ID	rs-nummeret for varianten, hentet fra dbSNP.txt, hvis dette er aktuelt. Hvis det finnes flere rs-numre på denne plasseringen, blir listen avgrenset med semikolon. Hvis det ikke eksisterer noen dbSNP-oppføring i denne posisjonen, brukes en markør for manglende verdi ('.').
REF	Referansegenotypen. Eksempel: en sletting av en enkel T angis som referanse-TT og alternativ T. En enkel A til T-nukleotidvariant er representert som referanse-A og alternativ T.
ALT	Allelene som er forskjellige fra referanseavlesingen. Eksempel: en innsetting av en enkel T angis som referanse-A og alternativ AT. En enkel A til T-nukleotidvariant er representert som referanse-A og alternativ T.
QUAL	En Phred-skalert kvalitetsscore tildelt av variantbetegneren. Høyere score angir høyere konfidens i varianten og lavere sannsynlighet for feil. I en kvalitetsscore på Q er den anslåtte sannsynligheten for en feil $10^{-(Q/10)}$. Eksempel: settet med Q30-betegnelser har en feilrate på 0,1 %. Mange variantbetegnere tildeler kvalitetsscorer basert på de tilhørende statistikkmodellene, som er høye i forhold til den observerte feilraten.

VCF-filkommentarer

Overskrift	Beskrivelse
FILTER	Hvis alle filtrene godkjennes, blir PASS skrevet i filterkolonnen. • LowDP (Lav DP) – Gjelder steder der dekningsdybden er under 150x. For PCR-produktposisjoner som dekkes av avlesning både fremover og reversert, svarer dette til 300x enkeltlesningsdekning. • q20 – Kvalitetsscore < 20. • MultiAllelicSite – Varianten stemmer ikke overens med diploid-modellen. • R5x9 – Antall tilstøtende repetisjoner (med lengde 1 til 5 bp) til variantbetegnelsene ≥ 9. • SB – Strengsavviket er mer enn gitt grenseverdi.
INFO	Mulige oppføringer i INFO-kolonnen omfatter: • AC – Alleltelling i genotyper for hver ALT-allel, i samme rekkefølge som oppført. • AF – Allelfrekvens for hver ALT-allel, i samme rekkefølge som oppført. • AN – Det totale antallet alleler i betegnede genotyper. • CD – Et flagg som indikerer at SNP opptrer innenfor kodingsområdet med minst 1 RefGene-oppføring. • DP – Dybden (antall basebetegnelser som er innrettet til en posisjon og brukt i variantbetegnelse). • Exon – En kommaseparert liste over eksonregioner avlest fra RefGene. • FC – Funksjonell konsekvens. • GI – En kommaseparert liste over gen-ID-er avlest fra RefGene. • QD – Variantkonfidens / kvalitet etter dybde. • TI – En kommaseparert liste over transkripsjons-ID-er avlest fra RefGene.

Overskrift	Beskrivelse
FORMAT	Formatkolonnen viser felt adskilt med kolon. For eksempel GT:GQ. Listen over viste felter avhenger av variantbetegneren som ble brukt. Følgende felt er tilgjengelige: • AD – Oppføring i skjemaet X, Y, der X er antallet referansebetegnelser, og Y er antallet alternative betegnelser. • DP – Omtrentlig avlesingsdybde, avlesinger med MQ=255 eller med feil tilordninger blir filtrert. • GQ – Genotypekvalitet. • GT – Genotype. 0 tilsvarer referansebasen, 1 tilsvarer den første oppføringen i ALT-kolonnen, osv. Skråstrek (/) angir at ingen fasingsinformasjon er tilgjengelig. • NC – Fraksjon av baser som var ubetegnet eller med basebetegnelseskvalitet under minimumsgrensen. • NL – Støynivå; et estimat av basebetegnelsesstøy i denne posisjonen. • SB – Strengavvik på denne posisjonen. Større negative verdier indikerer mindre avvik; verdier nær 0 angir mer avvik. • VF – Variantfrekvens, prosentandelen av avlesingene som støtter den alternative allelen.
PRØVE	I prøvekolonnen oppgis verdiene som angis i FORMAT-kolonnen.

Genome VCF-filer

Genome VCF (gVCF)-filer er VCF v4.1-filer som følger et sett med konvensjoner for å representere alle områder i genomet i et rimelig kompakt format. gVCF-filene (*.genome.vcf.gz) inneholder alle områdene i interesseregionen i en enkelt fil for hver prøve.

gVCF-filen viser ikke-betegnelser på posisjoner som ikke passerer alle filtre. En genotype (GT)-fane med ./ indikerer en ikke-betegnelse.

Se sites.google.com/site/gvcftools/home/about-gvcf for mer informasjon.

Fil for PCR-produktdekning

En fil for PCR-produktdekning genereres for hver manifestfil. M-nummeret i filnavnet angir manifestnummeret som er oppført i prøvetabellen for kjøringen.

Hver fil omfatter en overskriftsrad som inneholder prøve-ID-ene forbundet med manifestet. Under overskriftsraden er det 3 kolonner som viser følgende informasjon:

- ▶ Mål-ID-en slik den er oppført i manifestet.
- ▶ Dekningsdybden av avlesninger som passerer filteret.
- ▶ Den totale dekningsdybden.

Tilleggsutgangsfiler

Følgende utgangsfiler gir tilleggsinformasjon, eller oppsummerer kjøringsresultater og analysefeil. Selv om disse filene ikke er påkrevd for å vurdere analyseresultater, kan de brukes for feilsøkningsformål. Alle filene ligger i mappen Alignment (Innretting) med mindre annet er angitt.

Filnavn	Beskrivelse
AnalysisLog.txt	Behandlingslogg som beskriver hvert trinn som oppstod under analysen av gjeldende kjøringsmappe. Denne filen inneholder ingen feilmeldinger. Ligger i mappen Alignment (Innretting).
AnalysisError.txt	Behandlingsloggen som har opplistet eventuelle feil som har oppstått under analyse. Denne filen vil være tom hvis det ikke oppsto noen feil. Ligger i mappen Alignment (Innretting).

Filnavn	Beskrivelse
DemultiplexSummaryF1L1#.txt	Rapporterer demultipleksingsresultater i en tabell med 1 rad per plate og 1 kolonne per prøve. # representerer bane 1, 2, 3 eller 4 av strømningscellen. Ligger i mappen Alignment (Innretting).
AmpliconRunStatistics.xml	Inneholder oppsummeringsstatistikk som er spesifikk for kjøringen. Ligger i mappen Alignment (Innretting).

Analysemappe

Analysemappen inneholder filene som genereres av Local Run Manager-programvaren.

Forholdet mellom utgangsmappen og analysemappen er oppsummert som følger:

- ▶ Under sekvensering fyller Sanntidsanalyse (RTA) utgangsmappen med filer som genereres under bildeanalyse, basebetegnelse og kvalitetsscoring.
- ▶ RTA kopierer filer til analysemappen i sanntid. Etter at RTA har tilordnet en kvalitetsscore til hver base for hver syklus, skriver programvaren filen RTAComplete.txt til begge mappene.
- ▶ Når filen RTAComplete.txt er til stede, begynner analysen.
- ▶ Mens analysen fortsetter, skriver Local Run Manager utgangsfiler til analysemappen og kopierer deretter filene til utgangsmappen.

Innrettingsmapper

Hver gang denne analysen settes tilbake i kø, oppretter Local Run Manager en innrettingsmappe som heter **Alignment_N** (Innretting_N), hvor N er et sekvensielt tall.

Mappestruktur

 **Alignment** – Inneholder *.bam-, *.vcf-, FASTQ-filer, og filer som er spesifikke for analysemodulen.

 **Dato og klokkeslettstempel** – Date_time-stempel av analyse som ÅÅÅÅMMDD_TTMMSS

 AnalysisError.txt

 AnalysisLog.txt

 AmpliconRunStatistics.xml

 Sample1.genome.vcf.gz

 Sample1.coverage.csv

 Sample1.report.pdf

 Sample1.summary.csv

 Sample1.vcf.gz

 Sample1.bam

 **FASTQ**

 **Sample1**

 **Stats**

 DemuxSummaryF1L1.txt

 FastqSummaryF1L1.txt

 Data

 Intensities

 BaseCalls

 **L001** – Inneholder en undermappe per syklus, som hver inneholder *.bcl-filer.

 **L001** – Inneholder *.locs-filer, 1 for hver plate.

 **RTA Logs** – Inneholder loggfiler fra RTA-programvareanalyse.

 **InterOp** – Inneholder binære filer som brukes til å rapportere sekvenseringskjøringsmetrikk.

 **Logs** – Inneholder loggfiler som beskriver trinnene som utføres under sekvensering.

 RTAComplete.txt

 RunInfo.xml

 runParameters.xml

Teknisk hjelp

Kontakt Illuminas tekniske støtteavdeling for teknisk hjelp.

Nettsted: www.illumina.com
E-post: techsupport@illumina.com

Illumina Telefonnumre til kundestøtte

Region	Gratis	Regionalt
Nord-Amerika	+1.800.8094566	
Australia	+1.800.775.688	
Belgia	+32 80077160	+32 34002973
Danmark	+45 80820183	+45 89871156
Finland	+358 800918363	+358 974790110
Frankrike	+33 805102193	+33 170770446
Hongkong	800960230	
Irland	+353 1800936608	+353 016950506
Italia	+39 800985513	+39 236003759
Japan	0800.111.5011	
Kina	400.635.9898	
Nederland	+31 8000222493	+31 207132960
New Zealand	0800.451.650	
Norge	+47 800 16836	+47 21939693
Singapore	+1.800.579.2745	
Spania	+34 911899417	+34 800300143
Storbritannia	+44 8000126019	+44 2073057197
Sveits	+41 565800000	+41 800200442
Sverige	+46 850619671	+46 200883979
Taiwan	00806651752	
Tyskland	+49 8001014940	+49 8938035677
Østerrike	+43 800006249	+43 19286540
Andre land	+44.1799.534000	

Sikkerhetsdatablad – Tilgjengelige på Illuminas nettsider på support.illumina.com/sds.html.

Produktdokumentasjon – Tilgjengelig for nedlasting i PDF-format fra Illuminas nettsider. Gå til support.illumina.com, velg et produkt, og velg deretter **Documentation & Literature** (Dokumentasjon og litteratur).



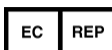
Illumina

5200 Illumina Way
San Diego, California, 92122 USA
+1.800.809.ILMN (4566)
+1.858.202.4566 (utenfor Nord-Amerika)
techsupport@illumina.com
www.illumina.com

TIL IN VITRO-DIAGNOSTISK BRUK

© 2021 Illumina, Inc. Med enerett.

CE



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

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

illumina®