

Tarkvara Local Run Manager somaatilise variandi analüüsimoodul

Töövoog juhend seadmele NextSeq 550Dx

IN VITRO DIAGNOSTILISEKS KASUTAMISEKS

Ülevaade	3
Käituse teabe sisestamine	4
Analüüsimeetodid	6
Käituse ja proovi andmete kuvamine	7
Analüüsiaruanne	7
Analüüsi väljundfailid	9
Aluste nimetamine ja indeksite mitmekesisus	15
Muudatuste ajalugu	16
Tehniline tugi	17



See dokument ja selle sisu kuuluvad ettevõttele Illumina, Inc. ja selle tütarettevõtetele („Illumina“) ning on mõeldud kasutamiseks ainult ettevõtte lepingulistele klientidele seoses selles dokumendis kirjeldatud toote (toodete) kasutamisega ega ole mõeldud mitte mingiks muuks otstarbeks. Seda dokumenti ega selle sisu ei tohi mis tahes viisil kasutada ega muul eesmärgil levitada ega edastada, avaldada või reprodutseerida ilma Illumina eelneva kirjaliku nõusolekuta. Illumina ei anna selle dokumendiga kolmandale isikule oma patendi-, kaubamärgi-, autori-, tava- või muu sarnase õiguse alusel mitte ühtegi litsentsi.

Kvalifitseeritud ja asjakohase koolituse saanud töötajad peavad selles dokumendis kirjeldatud suuniseid järgima rangelt ja üksikasjalikult, et tagada siin kirjeldatud toote (toodete) õige ja ohutu kasutusviis. Kogu siinse dokumendi sisu tuleb enne nimetatud toote (toodete) kasutamist täielikult läbi lugeda ja endale selgeks teha.

SELLES DOKUMENDIS KIRJELDATUD SUUNISTE LUGEMATA JÄTMINE JA EBATÄPNE JÄRGIMINE VÕIB KAHJUSTADA TOODET (TOOTEID), VIGASTADA INIMESI (SH KASUTAJAID VÕI TEISI) JA KAHJUSTADA MUUD VARA NING NIMETATUD JUHUL EI KEHTI ÜKSKI TOOTELE (TOODETELE) ANTUD GARANTII.

ILLUMINA EI VASTUTA SELLES DOKUMENDIS KIRJELDATUD TOOTE (TOOTETE) (SEALHULGAS TOOTE OSAD VÕI TARKVARA) VÄÄRKASUTUSE EEST.

© 2021 Illumina, Inc. Kõik õigused on kaitstud.

Kõik kaubamärgid kuuluvad ettevõttele Illumina, Inc. või nende vastavatele omanikele. Kaubamärgi kohta lisateabe saamiseks vt www.illumina.com/company/legal.html.

Ülevaade

Tarkvara Local Run Manager somaatilise variandi moodul on mõeldud kasutamiseks koos seadme Illumina TruSeq kohandatud amplikoni komplekti Dx testiga ja seadmega NextSeq 550Dx. Kui seda kasutatakse koos somaatilise variandi mooduliga, on see analüüs ette nähtud teekide ettevalmistamiseks, mida kasutatakse DNA sekveneerimiseks formaliiniga fikseeritud, parafiinis (FFPE) koest. Analüüs tuvastab madalatel variantide sagedustel somaatilised mutatsioonid.

Analüüsimoodul hindab variantide jaoks amplifitseeritud DNA lühikesi piirkondi ehk amplikone. Amplikonide fokuseeritud sekveneerimine võimaldab kindlate piirkondade kõrget katvust paljude proovide ulatuses. Analüüsimoodul viib läbi sekundaaranalüüsi ja aruannete genereerimise sekveneerimiskäitustest, kasutades kaheaheelist lähenemist, mis hõlmab eesmist ja tagumist oligokogumit. Vt pakendi teabelehte *Seadme TruSeq kohandatud amplikoni komplekt Dx (dokument nr 1000000029772)*.

Somaatilise variandi analüüsimoodul nõuab 300-tsüklilise sekveneerimise tarvikuid. Lisateavet vt pakendi infolehel *Seadme NextSeq 550Dx suure väljundiga reaktiivikomplekt v2* või *Seadme NextSeq 550Dx suure väljundiga reaktiivikomplekt v2.5*.

Teave selle juhendi kohta

Selles juhendis on toodud juhised sekveneerimise ja analüüsimise käitusparameetrid analüüsimooduli Somatic Variant jaoks. Lisateavet tarkvara Local Run Manager juhtpaneeli ja süsteemi sätete kohta vt *Seadme NextSeq 550Dx viitejuhend (dokument nr 1000000009513)*.

Tarkvara Local Run Manager kuvamine

Tarkvara Local Run Manager liides kuvatakse operatsioonisüsteemi NextSeq 550Dx (NOS) või veebibrauseri kaudu. Toetatud veebibrauser on Chromium.



MÄRKUS.

Kui kasutate toetuseta brauserit, laadige toetatud brauser alla, kui kuvatakse teade „Kinnita toetuseta brauser“. Chromiumi toetatud versiooni allalaadimiseks klõpsake [siin](#).

Kuvamine seadme monitoril

- 1 Tarkvara Local Run Manager liidese kuvamiseks seadme monitoril valige üks järgmistest võimalustest.
 - ▶ Valige NOS-i avakuval **Local Run Manager**.
NOS-i naasmiseks klõpsake paremas ülanurgas X-i.
 - ▶ Valige ikoon Minimeeri NOS, avage seadmel Chromiumi veebibrauser ja sisestage aadressiribale **http://localhost**.
Ainult administraatorid saavad NOS-i minimeerida.

Kuvamine võrguühendusega arvutis

- 1 Avage arvutis Chromiumi veebibrauser, millel on juurdepääs samale võrgule kui seadmel, ja ühendage seade, kasutades seadme IP-aadressi või nime. Näiteks **http://myinstrument**.

Käituse teabe sisestamine

Parameetrite määramine

- 1 Logige sisse rakendusse Local Run Manager.
- 2 Valige **Loo käitus** ja seejärel **Somaatiline variant**.
- 3 Sisestage käituse nimi, mis identifitseerib käituse sekveneerimisest analüüsini. Kasutage tähtnumbrilisi märke, tühikuid, kriipse või allkriipse.
- 4 **[Valkuline]** Sisestage käituse kirjeldus, et aidata käitust identifitseerida. Kasutage tähtnumbrilisi märke, tühikuid, kriipse või allkriipse.
- 5 Valige rippmenüüst proovide arv ja indeksikomplekt. Mõelge valiku tegemisel järgmisele teabele.
 - ▶ Rippmenüü sisaldab proovide arvu, millel on indeksikomplekt. Näiteks 24-komplekt 1 tähistab 24 testitavat proovi koos indeksikomplekti 1 indeksitega.
 - ▶ Indeksikomplekti numbrid viitavad erinevatele i5 indeksite komplektidele. Komplekt 1 ja komplekt 2 pakuvad indeksite mitmekesisust. Pakutakse kahte indeksikomplekti, mis aitavad vältida ühe komplekti ammendumist.
 - ▶ Valige proovide arv, mis on lähim testitavate proovide arvule. Kui loendis täpset proovide arvu pole, valige testimiseks vajalikule lähim, kuid proovide arvust väiksem arv. Näiteks kui soovite testida 18 proovi, valige 16 proovi.
 - ▶ Indeksi mitmekesisuse nõuetele vastavad proovide kaevud ja indeksikombinatsioonid on esile tõstetud rohelisega. Kui valite muud kaevud ja indeksikombinatsioonid, saate käituse salvestamisel teate, kui indeksite mitmekesisuse nõuded pole täidetud.

Manifestifailide importimine käituse jaoks

- 1 Veenduge, et imporditavad manifestid oleksid kättesaadavad juurdepääsetavas võrgus või USB -draivil.
- 2 Valige **Impordi manifestid**.
- 3 Liikuge manifestifaili juurde ja valige manifestid, mille soovite lisada.



MÄRKUS.

Manifestifailide kättesaadavaks tegemiseks kõigi käituste jaoks, kasutades somaatilise variandi analüüsimoodulit, lisage manifestid funktsiooni Mooduli sätteid abil. See funktsioon nõuab administraatori taseme õigusi. Lisateavet vt *Seadme NextSeq 550Dx viitejuhendist (dokument nr 1000000009513)*.


Proovide määramine käituse jaoks

Määrake käituse proovid, kasutades ühte järgmistest valikutest ja suunistes.


- ▶ **Proovide sisestamine käsitsi** – kasutage tühja tabelit kuvall Loo käitus.
- ▶ **Proovide importimine** – liikuge komaeraldusega väärtuste (*.csv) vormingus välise faili juurde. Malli saab alla laadida kuvall Käituse loomine.

Pärast proovide tabeli täitmist saate proovide teabe eksportida välisesse faili. Kasutage faili viitena teekide ettevalmistamisel või teiseks käituseks importimisel.

Proovide sisestamine käsitsi

- 1 Sisestage väljale Proovi nimi kordumatu proovi nimi.
Kasutage tähtnumbrilisi märke, kriipse või allkriipse.
Proovi nimi täidab vastava kaevu automaatselt teises kogumis.
- 2 **[Valikuline]** Positiivsete või negatiivsete kontrollproovide jaoks paremklõpsake ja valige kontrollitüüp.
Ühes proovikaevus sisalduv kontroll täidab teise kogumi vastava kaevu sama kontrolliga.
- 3 **[Valikuline]** Sisestage väljale Proovi kirjeldus proovi kirjeldus.
Kasutage tähtnumbrilisi märke, kriipse või allkriipse.
Proovi kirjeldus täidab vastava kaevu automaatselt teises kogumis.
Proovi kirjeldused on seotud proovi ID-ga. Proovikirjeldused kirjutatakse üle, kui sama proovi ID-d kasutatakse hiljem uuesti.
- 4 Valige indeksi 1 adapter rippmenüüst Indeks 1 (i7).
Soovitatud proovide kaevude kasutamisel täidab tarkvara i7 ja i5 indeksiadapterid, mis vastavad indeksi mitmekesisuse nõuetele. Kui loendis pole täpset testitavate proovide arvu, valige lisakaevude jaoks kindlasti indeksiadapterid. Kui peate lisakaevude jaoks valima indeksid või kui te ei kasuta soovitatud indeksiadapteri kombinatsioone, lugege enne indeksite valimist kindlasti dokumendi *Aluste nimetamine ja indeksite mitmekesisus leheküljel 15*.
- 5 Valige indeksi 2 adapter rippmenüüst Indeks 2 (i5).
- 6 Valige rippmenüüst Manifest manifestifail.
A-kogumi proovid nõuavad teistsugust manifesti kui proovid B-kogumist.
- 7 Valige suvand plaadi kujunduse vaatamiseks, printimiseks või salvestamiseks viitena teekide ettevalmistamisel.
 - ▶ Plaadi kujunduse kuvamiseks valige ikoon  **Prindi**. Plaadi kujunduse printimiseks valige **Prindi**.
 - ▶ Valige **Ekspordi**, et eksportida proovi teave välisesse faili.
Veenduge, et manifesti ja proovide teave oleks õige. Vale teave võib tulemusi mõjutada.
- 8 Valige **Salvesta käitus**.

Proovide importimine

- 1 Valige **Impordi proovid** ja sirvige proovi teabefaili asukohta. Imporditavaid faile on kahte tüüpi.
 - ▶ Uue plaadikujunduse tegemiseks valige **Mall** kuval Käituse loomine. Mallifail sisaldab impordiks õigeid veerupäiseid. Sisestage igas veerus käituses olevate proovide kohta proovide teave. Kustutage näidisteave kasutamata lahtritest ja seejärel salvestage fail.
 - ▶ Kasutage somaatilise variandi moodulist ekspordifunktsiooni abil eksporditud proovide teabega faili.
- 2 Plaadi kujunduse kuvamiseks valige ikoon  **Prindi**.
- 3 Valige **Prindi**, et printida plaadi kujundus teekide ettevalmistamise näidisena.
- 4 **[Valikuline]** Valige **Ekspordi**, et eksportida proovide teave välisesse faili.
Veenduge, et manifesti ja proovide teave oleks õige. Vale teave võib tulemusi mõjutada.
- 5 Valige **Salvesta käitus**.

Käituse redigeerimine

Käituses oleva teabe redigeerimise suunised enne järjestamist vt *Seadme NextSeq 550Dx viitejuhend (dokument nr 1000000009513)*.

Analüüsimeetodid

Analüüsimoodul Somatic Variant teeb järgmised analüüsietapid ja kirjutab seejärel analüüsi väljundfailid kausta Joondamine.

- ▶ Demultipleksib indeksi lugemid
- ▶ Genereerib FASTQ-failid
- ▶ Joondab etaloniga
- ▶ Tuvastab variandid

Demultipleksimine

Demultipleksimine võrdleb iga indeksi lugemi järjestust käituse jaoks määratud indeksite järjestustega. Selles etapis ei arvestata kvaliteediväärtustega.

Indeksite lugemid identifitseeritakse järgmiste sammudega.

- ▶ Proovid nummerdatakse alates 1-st, lähtudes nende järjestusest käituse jaoks.
- ▶ Valimi number 0 on reserveeritud klastritele, mida proovile ei määratud.
- ▶ Klastrid määratakse proovile siis, kui indeksi järjestus vastab täpselt või kui indeksi lugemi kohta on kuni üks mittevastavus.

FASTQ-failide genereerimine

Pärast demultipleksimist genereerib tarkvara vaheanalüüsi failid FASTQ-vormingus, mis on tekstivorming, mida kasutatakse järjestuste esitamiseks. FASTQ-failid sisaldavad kõikide proovide lugemeid ja seonduvaid kvaliteediskoore. Klastrid, mis ei läbinud filtrit, on välistatud.

FASTQ-fail sisaldab ainult ühe proovi lugemeid ja selle proovi nimi sisaldub FASTQ-faili nimes. FASTQ-failid on joondamise peamine sisend. Oligokogumi kohta moodustatakse proovi kohta kaheksa FASTQ-faili, neli lugemist 1 ja neli lugemist 2, mille tulemuseks on proovi kohta kokku 16 FASTQ-faili.

Joondamine

Joondamisetapi ajal joondab Smith-Watermani ribaalgoritm iga proovi klastrid manifestifailis täpsustatud amplikonjäärjärejstuste kohaselt.

Smith-Watermani ribaalgoritm teostab semiglobaalseid järjestuste joondusi, et määrata sarnased piirkonnad kahele järjestusele. Kogu järjestuse võrdlemise asemel võrdleb Smith-Watermani algoritm kõigi võimalike pikkustega segmente.

Iga paaristulemusega lugemit hinnatakse selle joondamise suhtes selle lugemi asjakohaste sondijärjestustega.

- ▶ Lugemist 1 hinnatakse allavoolu lookusekohaste oligode (DLSO) pöördkomplementi suhtes.
- ▶ Lugemist 2 hinnatakse ülesvoolu lookusekohaste oligode (ULSO) suhtes.
- ▶ Kui lugemi algus ühtib sondijärjestusega, milles pole rohkem kui kolm erinevust (mittevastavus või nihked juhtivate indelide tõttu), joondatakse lugemist täispikkus selle järjestuse amplikoni sihtmärgi suhtes.
- ▶ Analüüsikeemiat arvestades DLSO-s ja ULSO-s indeleid ei täheldata.

Joondused filtreeritakse joondamistulemuste põhjal, lähtudes mittevastavuse määrast kas huvipakkuvas piirkonnas või kogu amplikonis, olenevalt amplikoni pikkusest. Filtreeritud joondused kirjutatakse joondusfailidesse joondamata ja neid ei kasutata variantide nimetamisel.

Variandi nimetamine



Illumina väljatöötatud variandinimetaja Pisces identifitseerib DNA proovis madalal sagedusel esinevad variandid.

Variandinimetaja Pisces identifitseerib SNV-d, MNV-d ja väiksed indelid kolme etapiga:

- ▶ Vaatleb referentsgenoomi iga positsiooni eraldi
- ▶ Loendab alused selles positsioonis joondatud lugemite korral, mis kattuvad positsiooniga
- ▶ Arvutab Poissoni mudeli abil nimetamise kvaliteeti mõõtvana variandiskoori. Variandid, mille kvaliteediskoor on alla Q30, on välistatud.

Variante nimetatakse kõigepealt iga kogumi jaoks eraldi. Seejärel võrreldakse iga kogumi variante ja ühendatakse need üheks väljundfailiks. Kui variant on olemas mõlemas kogumis ja see läbib kõik filtrid, mis on loetletud *VCF-faili annotatsioonid leheküljel 11*, on variandi nimetamise failis (VCF) variant märgitud kui PASS.

Käituse ja proovi andmete kuvamine

- 1 Klõpsake Local Run Manageri juhtpaneelil käituse nime.
- 2 Vaadake vahekaardil Käituse ülevaade sekveneerimiskäituse mõõdikud üle.
- 3 **[Valikuline]** Käituse väljundkausta tee kopeerimiseks klõpsake ikooni **Kopeeri lõikelauale** .
- 4 Käitusparameetrite ja tarvikute teabe vaatamiseks klõpsake vahekaarti Sekveneerimise teave.
- 5 Analüüsiaruande asukoha vaatamiseks klõpsake vahekaarti Proovid ja tulemused.
 - ▶ Kui analüüsi korrati, laiendage rippmenüüd Valige analüüs ja valige sobiv analüüs.
- 6 Analüüsikausta tee kopeerimiseks klõpsake ikooni **Kopeeri lõikelauale** .

Lisateavet vahekaartide Käituse ülevaade ja Sekveneerimise teave ja analüüsi taotlemise kohta vt *Seadme NextSeq 550Dx viitejuhend (dokument nr 1000000009513)*.

Analüüsiaruanne

Analüüsitulemused võetakse kokku vahekaardil Proovid ja tulemused ning koondaruandena kaustas Joondamine. Proovide aruanded on saadaval ka vastava proovi PDF-failis.

Teave vahekaardil Proovid ja tulemused

1 Proovi aruande nägemiseks klõpsake loendis proovil.

Tabel 1 Käituse ja proovi teave

Veeru pealkiri	Kirjeldus
Käituse olek	Näitab, kas sekveneerimiskäitus õnnestus või ebaõnnestus.
Kogusaagis (GB)	Sekveneerimiskäitusel nimetatud aluste arv. Näitab künnise ületamist ja õnnestumise või ebaõnnestumise olekut.
% ≥ Q30	Kvaliteediskooriga 30 (Q30) või suuremaga sekveneerimiskäituse lugemite protsent. Näitab künnise ületamist ja õnnestumise või ebaõnnestumise olekut.
Proovi nimi	Proovi nimi, mis anti käituse loomisel.
PF-lugemite koguarv	Filtrit läbivate lugemite koguarv.
Lugemi 1% ≥ Q30	Lugemite protsent lugemis 1 proovi kvaliteediskooriga 30 (Q30) või rohkem.
Lugemi 2% ≥ Q30	Lugemite protsent lugemis 2 proovi kvaliteediskooriga 30 (Q30) või rohkem.
Autosoomi nimetuse määr	Autosoomide (kromosoomid 1–22) genoomsete positsioonide arv, mis vastab varem kindlaksmääratud usaldusväärtuse künnisele, jagatuna nõutud autosoomsete genoomsete positsioonide koguarvuga. Nimetamismäär kirjeldatakse proovipõhiselt ja see esitatakse protsendina, mis arvutatakse 1 miinus autosoomsete positsioonide arv koos mittetäielike nimetustega jagatuna sekveneeritud autosoomsete positsioonide koguarvuga.

Tabel 2 Prooviaruande teave

Veeru pealkiri	Kirjeldus
Proov	Proovi nimi, mis anti käituse loomisel.
Aruande kuupäev	Aruande loomise kuupäev.
Proovi teave	Käituse loomisel esitatud proovi ID, proovis filtri läbinud lugemite koguarv, proovi lugemite protsent kvaliteediskooriga 30 (Q30) või rohkem ja autosoomne nimetamise määr.
Amplikoni kokkuvõte	Sekveneeritud ampliconi piirkondade koguarv ja sekveneeritud ampliconide kogupikkus aluspaarides sihtpiirkondades A- ja B-kogumi proovi kohta ning iga kogumi jaoks kasutatud manifestifaili kohta. Manifestifail määrab etalongenoomi ja sihitud etalonpiirkonnad, mida joondamisetapis kasutatakse.
Lugemi taseme statistika	Etalonis iga positsiooni katva proovi lugemite arv ja protsent lugemi 1 ja lugemi 2 kohta kogumis A ja kogumis B.
Variantide kokkuvõte	Soovitavad väärtused ületanud tuvastatud SNV-de, sisestuste ja kustutuste arv proovi kohta, et teha kindlaks, kas kvaliteeditulemused on vastuvõetavas vahemikus.
Katvuse kokkuvõte	Joondatud aluste koguarv jagatuna sihtpiirkonna suurusel ja ampliconi piirkondade protsent, mille katvuse väärtus on suurem kui madala katvuse künnis 0,2 * ampliconi keskmine katvus proovi jaoks kogumis A ja kogumis B.
Katvusdiagrammid	Diagramm Katvus ampliconi piirkondade kaupa näitab proovi ampliconi piirkondade katvust. Piirkonnad, mille katvuse väärtus on väiksem kui katvuse künnis, on esile tõstetud punasega. Kõigi väärtuste keskmine tähistatakse oranži joonega. Kogumi A ja kogumi B katvus on näidatud diagrammil.
Tarkvaraversioonid	Tarkvaraversioonid proovi sekveneerimise ajal. Sisaldab NextSeq 550Dx operatsioontarkvara (NOS), Local Run Manageri tarkvara, RTA tarkvara ja somaatilise variandi mooduli versiooni.

Analüüsi väljundfailid

Järgmised analüüsi väljundfailid genereeritakse somaatilise variandi analüüsimooduli jaoks ja need pakuvad analüüsi tulemusi joondamise ja variantide nimetamise kohta. Analüüsi väljundfailid asuvad kaustas Joondamine.

Faili nimi	Kirjeldus
Demultipleksimine (*.txt)	Vahefailid, mis sisaldavad kokkuvõtlike tulemuste demultipleksimist.
FASTQ (*.fastq.gz)	Vahefailid, mis sisaldavad kvaliteediskooriga aluste nimetusi. FASTQ-failid on joondamisetapi peamine sisend.
Joendusfailid BAM-vormingus (*.bam)	Sisaldab selle proovi joondatud lugemeid.
Kogumikohased variantide nimetuste failid VCF-vormingus (*.vcf)	Sisaldab variante, mida nimetatakse igas positsioonis kas eesmisest või tagumisest kogumist.
Variantide nimetuste failid genoomi VCF-vormingus (*.genome.vcf.gz)	Sisaldab iga positsiooni jaoks genotüüpi, olenemata sellest, kas seda nimetatakse variandiks või etaloniks.
Konsensuse variantide nimetamise failid VCF-vormingus (*.vcf.gz)	Sisaldab variante, mida nimetatakse mõlemas kogumis igas positsioonis.
AmpliconCoverage_M1.tsv	Sisaldab teavet iga pakutava manifesti katvuse kohta ampliconi ja proovi osas. M# tähistab manifesti numbrit.

Failivormingu demultipleksimine

Demultipleksimisprotsess loeb igale klastrile lisatud indeksjärjestuse, et teha kindlaks, millisest proovist klaster pärineb. Klastre ja proovinumbri vastendus kirjutatakse läbivooluküveti iga paani demultipleksimise (*.demux) faili.

Demultipleksimise faili nimevorming on **s_1_X.demux**, kus X on paani number.

Demultipleksimise failid algavad päisega.

- ▶ Versioon (4 baiti täisarv), praegu 1
- ▶ Klastre arv (4 baiti täisarv)

Ülejäänud fail koosneb paani klastre proovide numbritest.

Kui demultipleksimise etapp on lõpule jõudnud, genereerib tarkvara demultipleksimise faili **DemultiplexSummaryF1L1.txt**.

- ▶ **F1** failinimes tähistab läbivooluküveti numbrit.
- ▶ **L1** failinimes tähistab raja numbrit.
- ▶ Demultipleksimise tulemuseks on tabel, kus on 1 rida paani kohta ja 1 veerg proovi kohta, sealhulgas proov 0.
- ▶ Kõige sagedamini esinevad järjestused indeksi lugemites.

FASTQ-failivorming

FASTQ on tekstipõhine failivorming, mis sisaldab aluste nimetusi ja kvaliteedi väärtusi ühe lugemi kohta. Iga kirje sisaldab 4 rida:

- ▶ Identifikaator
- ▶ Järjestus
- ▶ Plussmärk (+)

- ▶ Phredi kvaliteediskoorid ASCII + 33 kodeeritud vormingus

Identifikaator on vormindatud järgmiselt:

@Instrument:RunID:FlowCellID:Lane:Tile:X:Y ReadNum:FilterFlag:0:SampleNumber

Näide:

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

BAM-failivorming

BAM-fail (*.bam) on SAM-faili tihendatud binaarversioon, mida kasutatakse kuni 128 Mb joondatud järjestuste kuvamiseks. SAM- ja BAM-vorminguid on üksikasjalikult kirjeldatud failis samtools.github.io/hts-specs/SAMv1.pdf.

BAM-failid kasutavad failinimede vormingut **SampleName_S#.bam**, kus # on proovi number, mis on määratud loendi alusel, mille järgi proovid on kätuse jaoks järjestatud.

BAM-failid sisaldavad päisejaotist ja joondusejaotist.

- ▶ **Päis** – sisaldab teavet kogu faili kohta, näiteks proovi nimi, proovi pikkus ja joondusmeetod. Joondusejaotises seostatakse joondused päisejaotises oleva konkreetse teabega.
- ▶ **Joondused** – sisaldab lugemise nime, lugemise järjestust, lugemise kvaliteeti, joondamisteavet ja kohandatud silte. Lugemi nimi sisaldab kromosoomi, alguskoordinaati, joondamiskvaliteeti ja kattuvuse deskriptorstringi.

Joondusejaotises on iga lugemi või lugemipaari kohta järgmine teave:

- ▶ **AS:** paaristulemusega joondamiskvaliteet
- ▶ **BC:** vötkoodi silt, mis näitab lugemiga seotud demultipleksitud proovi ID-d.
- ▶ **SM:** üksiktulemusega joondamiskvaliteet.
- ▶ **XC:** kattuvuse deskriptorstring
- ▶ **XN:** amplikoni nimesilt, mis salvestab lugemiga seotud amplikoni ID

BAM-i indeksfailid (*.bam.bai) sisaldavad vastava BAM-faili indeksit.

VCF-failivorming

Variant Call Format (VCF) on levinud failivorming, mille on välja töötanud genoomika teadusringkonnad. See sisaldab teavet variantide kohta, mis on leitud referentsgenoomi kindlatel positsioonidel. VCF-failid lõpevad laiendiga .vcf

VCF-faili päis sisaldab VCF-failivormingu versiooni ja variandi nimetaja versiooni ning loetleb faili ülejäänud osas kasutatud annotatsioonid. VCF-i päis sisaldab ka referentsgenoomi faili ja BAM-faili. Päise viimane rida sisaldab andmeridade veerupäiseid. Kõik VCF-failide andmerekad sisaldavad teavet ühe variandi kohta.

VCF-faili pealkirjad

Pealkiri	Kirjeldus
CHROM	Referentsgenoomi kromosoom. Kromosoomid ilmuvad samas järjekorras kui viitefailis FASTQ.
POS	Variandi ühe-aluseline positsioon referentskromosoomis. SNP-de puhul on see positsioon referentsalus variandiga; indelide või kustutuste puhul on see positsioon referentsalus vahetult enne varianti.
ID	Vajaduse korral saidilt dbSNP.txt saadud variandi rs-number. Kui selles asukohas on mitu rs-numbrit, on loend semikoolonitega eraldatud. Kui sellel positsioonil ei ole dbSNP-kirjet, kasutatakse puuduva väärtuse markerit ('.').
REF	Referentsgenotüüp. Näiteks ühe T kustutamine on kujutatud referentsina TT ja alternatiivina T. A-T üksiknukleotiidi variant on kujutatud referentsina A ja alternatiivina T.
ALT	Alleelid, mis erinevad referentslugemist. Näiteks on üksiku T sisestamine esindatud kui referents A ja alternatiiv AT. Üksiknukleotiidi variant A-T on esindatud kui referents A ja alternatiiv T.
QUAL	Phred-skaalaga kvaliteediskoor, mille määrab variandi nimetaja. Kõrge skoor näitab suuremat usaldust variandi suhtes ja väiksemat vigade tõenäosust. Q kvaliteediskoori korral on vea hinnanguline tõenäosus $10^{-(Q/10)}$. Näiteks on Q30 nimetusekomplekti veamäär 0,1%. Paljud variantide nimetajad määravad oma statistiliste mudelite põhjal kvaliteediskoori, mis on täheldatud veamäär suhtes kõrge.

VCF-faili annotatsioonid

Pealkiri	Kirjeldus
FILTER	Kui kõik filtrid on läbitud, kirjutatakse filtriveergu PASS . <ul style="list-style-type: none"> • LowDP – rakendatakse saitidele, mille katvuse sügavus on mõlemas kogumis alla 450x. Nii edasi- kui ka tagasilugemiga kaetud amplikoni positsioonide korral võrdub see 900x ühekordse lugemisega. • LowGQ – genotüpiseerimise kvaliteet (GQ) jääb alla piirväärtuse. • q30 – kvaliteediskoor < 30. • LowVariantFreq – variandi sagedus on väiksem kui antud künnis. • PB – sondikogumi nihe. Varianti ei leitud või leiti madala sagedusega ühes või kahes sondikogumis. • R3x6 – külgnevate korduste arv (pikkusega 1 kuni 3 bp) variandi nimetustele ≥ 6. • SB – ahela nihe on suurem kui antud künnis.
TEAVE	Veeru TEAVE võimalikud kirjed on järgmised. <ul style="list-style-type: none"> • AC – alleelide arv genotüüpides iga ALT-alleeli jaoks, samas järjekorras nagu loetletud. • AF – alleelide sagedus iga ALT-alleeli jaoks, samas järjekorras nagu loetletud. • AN – alleelide koguarv nimetatud genotüüpides. • CD – lipp, mis näitab, et SNP esineb vähemalt ühe RefGene'i kirje kodeerivas piirkonnas. • DP – sügavus (aluse nimetuste arv, mis on joondatud positsiooniga ja mida kasutatakse variandi nimetamiseks). • Exon – RefGene'ist loetud komaeraldusega eksonipiirkondade loend. • FC – funktsionaalne tagajärg. • GI – RefGene'ist loetud komaeraldusega geeni ID-de loend. • QD – variandi usaldatavus/kvaliteet sügavuse järgi. • TI – RefGene'ist loetud komaeraldusega transkriptsiooni ID-de loend.

Pealkiri	Kirjeldus
VORMING	<p>Vormindatud veerus on koolonitega eraldatud väljad. Nt GT:GQ. Väljade loend oleneb kasutatavast nimetaja variandist. Saadaval on järgmised väljad.</p> <ul style="list-style-type: none"> • AD – vormi X,Y sisestamine, kus X on etalonnimetuste arv ja Y on alternatiivnimetuste arv. • DP – ligikaudne lugemisügavus; lugemid MQ = 255 või halbade paarilistega filtreeritakse. • GQ – genotüübi kvaliteet. • GQX – genotüübi kvaliteet. GQX on GQ väärtuse ja veeru QUAL miinimum. Üldiselt on need väärtused sarnased; miinimumi võtmine muudab GQX-i genotüübi kvaliteedi konservatiivsemaks näitajaks. • GT – genotüüp. 0 vastab etalonlusele, 1 vastab veeru ALT esimesele kirjele ja nii edasi. Kalkkriips (/) näitab, et faasimise teave pole saadaval. • NL – müratase; hinnang aluse nimetamise müra kohta sellel positsioonil. • PB – sondikogumi nihe. 0-le lähemal olevad väärtused näitavad suuremat nihet ühe sondikogumi suunas ja vähem usaldatavust variandi nimetuse suhtes. • SB – aluse nihe sellel positsioonil. Suuremad negatiivsed väärtused näitavad väiksemat nihet; 0 lähedal olevad väärtused näitavad suuremat nihet. • VF – variandi sagedus; alternatiivset alleeli toetavate lugemite protsent.
PROOV	Proovi veerg annab veerus VORMING määratud väärtused.

Genoomi VCF-failid

Genoomi VCF-failid (gVCF) on VCF v4.1 failid, mis järgivad genoomi kõigi saitide mõistlikult kompaktses vormis esitamise tavasid. GVCF (*.genome.vcf.gz) failid sisaldavad ühes failis kõiki huvipakkuvas piirkonnas asuvaid saite iga proovi jaoks.

GVCF-fail näitab nimetuse puudumist positsioonidel, mis ei läbi kõiki filtreid. Genotüübi (GT) silt ./.. näitab nimetuse puudumist.

Lisateavet leiate asukohast sites.google.com/site/gvcftools/home/about-gvcf.

Kogumipõhised ja konsensuslikud VCF-failid

Somaatilise variandi töövoog genereerib 2 komplekti variandi nimetuste faile.

- ▶ **Kogumipõhised VCF-failid** – sisaldab variante, mida nimetatakse kas eesmisses või tagumises kogumis. Kogumipõhised failid kirjutatakse kausta VariantCallingLogs.
- ▶ **Konsensuslikud VCF-failid** – sisaldavad mõlemast kogumist nimetatud variante. Konsensusfailid kirjutatakse kausta Joondamine.

Kogumipõhised ja konsensuslikud VCF-failid sisaldavad nii VCF- (*.vcf) kui ka gVCF- (*.genome.vcf) faile ning kasutavad järgmist nimetamisviisi, kus S# tähistab järjestust, milles proov on loetletud käituse jaoks.

- ▶ **Kõigi saitide aruanded** – SampleName_S#.genome.vcf
- ▶ **Ainult variantide aruanded** – SampleName_S#.vcf

Tarkvara võrdleb kogumipõhiseid VCF-faile ja ühendab igas positsioonis olevad andmed, et luua proovi jaoks konsensuslik VCF-fail.

Variandinimetused igast kogumist liidetakse konsensuslikeks VCF-failideks järgmiste kriteeriumide alusel.

Kriteeriumid	Tulemus
Referentsnimetus igas kogumis	Referentsnimetus
Referentsnimetus ühes ja variandinimetused teises kogumis	Filtreeritud variandinimetused
Ühtivad variandinimetused igas kogumis sarnaste sagedustega	Variandinimetused
Ühtivad variandinimetused igas kogumis oluliselt erinevate sagedustega	Filtreeritud variandinimetused
Mitteühtivad variandinimetused igas kogumis	Filtreeritud variandinimetused

Iga kogumi näitajad ühendatakse järgmiste väärtuste abil.

Mõõdik	Väärtus
Sügavus	Sügavuste lisamine mõlemast kogumist
Variandi sagedus	Variantide koguarv jagatuna kogu katvuse sügavusega
Q-skoor	Mõlema kogumi miinimumväärtus

Amplikoni katvusfail

Iga manifestifaili jaoks luuakse ampliconi katvusfail. Failinimes olev M# tähistab manifesti numbrit.

Iga fail sisaldab päiserida, mis sisaldab manifestiga seotud proovi ID-sid. Fail sisaldab järgmist teavet.

- ▶ Sihtmärgi ID sellisena, nagu see on manifestis toodud.
- ▶ Filtrit läbiva lugemi katvussügavus.

Täiendavad väljundfailid

Järgmised väljundfailid pakuvad lisateavet või võtavad kokku käituse tulemused ja analüüsivead. Kuigi neid faile pole analüüsitulemuste hindamiseks vaja, saab neid tõrkeotsingu eesmärkidel kasutada. Kõik failid asuvad kaustas Joondamine, kui pole täpsustatud teisiti.

Faili nimi	Kirjeldus
AnalysisLog.txt	Töötuslogi, mis kirjeldab kõiki toiminguid, mis toimusid praeguse käituse kausta analüüsimisel. See fail ei sisalda veateateid. Asub kaustas Joondamine.
AnalysisError.txt	Töötuslogi, mis loetleb analüüsimisel ilmnunud vead. Kui vigu ei esinenud, on see fail tühi. Asub kaustas Joondamine.
DemultiplexSummaryF1L1#.txt	Annab demultipleksimise tulemused tabelis, milles on 1 rida paani ja 1 veerg proovi kohta. # tähistab läbivooluküveti rada 1, 2, 3 või 4. Asub kaustas Joondamine.
AmpliconRunStatistics.xml	Sisaldab käitusele omast kokkuvõtlikku statistikat. Asub kaustas Joondamine.

Analüüsikaust

Analüüsikaust sisaldab tarkvara Local Run Manager loodud faile.

Väljundkausta ja analüüsikausta seos on kokku võetud järgmiselt:

- ▶ Järjestamise ajal täidab reaajas analüüs (RTA) väljundkausta piltide analüüsi, aluste nimetamise ja kvaliteedi hindamise käigus loodud failidega.
- ▶ RTA kopeerib failid analüüsikausta reaajas. Pärast seda, kui RTA määrab kõigile tsükli alustele kvaliteediskoori, kirjutab tarkvara faili RTAComplete.txt mõlemasse kausta.
- ▶ Kui fail RTAComplete.txt on olemas, algab analüüs.
- ▶ Analüüsi jätkudes kirjutab Local Run Manager väljundfailid analüüsikausta ja kopeerib seejärel failid tagasi väljundkausta.

Joonduste kaustad

Iga kord, kui seda analüüsi nõutakse, loob Local Run Manager joonduskausta nimega **Alignment_N**, kus N on järjekorranumber.

Kaustastruktuur

📁 **Joondamine** – sisaldab *.bam-, *.vcf-, FASTQ-faile ning analüüsimoodulile omaseid faile.

📁 **Kuupäeva ja kellaaja tempel** – analüüsi kuupäeva ja kellaaja tempel YYYYMMDD_HHMMSS

📄 AnalysisError.txt

📄 AnalysisLog.txt

📄 aggregate.report.html

📄 aggregate.report.pdf

📄 aggregate.summary.csv

📄 AmpliconCoverage_M#.tsv

📄 AmpliconRunStatistics.xml

📄 Sample1.genome.vcf.gz

📄 Sample1.coverage.csv

📄 Sample1.report.pdf

📄 Sample1.summary.csv

📄 Sample1.vcf.gz

📄 Sample1.bam

📁 **FASTQ**

📁 **Sample1**

📄 Sample1_L001_R1_001_fastq.gz

📁 **Stats**

📄 DemuxSummaryF1L1.txt

📄 FastqSummaryF1L1.txt

📁 **Data**

📁 **Intensities**

📁 **BaseCalls**

📁 **L001** – sisaldab *.bcl-faile.

📁 **L001** – sisaldab *.locs-faile.

📁 **RTA logid** – sisaldab RTA tarkvara analüüsi logifaile.

📁 **InterOp** – sisaldab binaarfaile, mida kasutatakse sekveneerimise käituse näitajate aruandluseks.

📁 **Logid** – sisaldab logifaile, mis kirjeldavad sekveneerimise ajal läbitud etappe.

📄 RTAComplete.txt

📄 RunInfo.xml

📄 RunParameters.xml

Aluste nimetamine ja indeksite mitmekesisus

Kui proovid sekveneeritakse seadmel NextSeq 550Dx, määrab aluse nimetus kindlas tsükli aluse (A, C, G või T) antud paani iga klasteri jaoks või kujutise ala läbivooluküvetil. Seade NextSeq 550Dx kasutab kahekanalilist sekveneerimist, mis nõuab nelja DNA-aluse andmete kodeerimiseks ainult kahte kujutist, üks punasest ja teine rohelisest kanalist.

Indeksi lugemite aluste nimetamise protsess erineb aluste nimetamisest teiste lugemite ajal.

Indeksi lugemid peavad algama kahes esimeses tsükli vähemalt ühe G-st erineva alusega. Kui indeksi lugem algab kahe G-aluse nimetamisega, siis signaali intensiivsust ei genereerita. Demultipleksimise toimivuse tagamiseks peab signaal olema ühes kahest esimesest tsüklist.

Indeksite valimisel käituse loomise ajal ilmub vähese mitmekesisuse hoiatus, kui indeksid ei vasta mitmekesisuse nõuetele. Vähese mitmekesisuse hoiatuse vältimiseks valige indeksitsükliid, mis annavad signaali mõlemas kanalid iga tsükli jaoks.

- ▶ Punane kanal – A või C
- ▶ Roheline kanal – A või T

See aluste nimetamise protsess tagab madala pleksiga proovide analüüsimisel täpsuse. Lisateavet indeksite sekveneerimise kohta leiate pakendi teabelehel *Seadme TruSeq kohandatud amplikonikomplekt Dx (dokument nr 1000000029772)*.

Tarkvaras Local Run Manager käituse loomise ajal valite testitavate proovide arvu. Soovitatud indeksikombinatsioonid, mis vastavad indeksite mitmekesisuse nõuetele, täidab tarkvara automaatselt. Kuigi te ei pea soovitatud indeksikombinatsiooni kasutama, on see siiski soovitatav.

Muudatuste ajalugu

Dokument	Kuupäev	Muudatuse kirjeldus
Dokument nr 1000000030330 v04	August 2021	ELi volitatud esindaja uuendatud aadress.
Dokument nr 1000000030330 v03	Detsember 2019	ELi volitatud esindaja uuendatud aadress. Austraalia sponsori uuendatud aadress.
Dokument nr 1000000030330 v02	Jaanuar 2019	Lisatud teave v2.5 reaktiivikomplektide kohta.
Dokument nr 1000000030330 v01	August 2018	Uuendatud regulatiivsed märgised.
Dokument nr 1000000030330 v00	November 2017	Esialgne väljalase.

Tehniline tugi

Tehnilise abi saamiseks pöörduge Illumina tehnilise toe poole.

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

Illumina klienditoe telefoninumbrid

Piirkond	Tasuta	Piirkondlik
Põhja-Ameerika	+1 800 809 4566	
Austraalia	+1 800 775 688	
Austria	+43 800 006 249	+43 1928 6540
Belgia	+32 8007 7160	+32 3400 2973
Hiina	40 0066 5835	
Hispaania	+34 911 899 417	+34 800 300 143
Holland	+31 800 022 2493	+31 207 132 960
Hongkong	800960230	
Iirimaa	+353 1800 936 608	+353 016950506
Itaalia	+39 800 985 513	+39 236 003 759
Jaapan	080 0111 5011	
Norra	+47 8001 6836	+47 2193 9693
Prantsusmaa	+33 805 102 193	+33 170 770 446
Rootsi	+46 850 619 671	+46 200 883 979
Saksamaa	+49 800 101 4940	+49 893 803 5677
Singapur	+1 800 579 2745	
Soome	+358 800 918 363	+358 974 790 110
Šveits	+41 565 800 000	+41 800 200 442
Taani	+45 8082 0183	+45 8987 1156
Taiwan	008 0665 1752	
Uus-Meremaa	08 0045 1650	
Ühendkuningriik	+44 800 012 6019	+44 207 305 7197
Muud riigid	+44 179 953 4000	

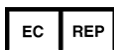
Ohutuse andmelehed (SDS-id): saadaval Illumina veebisaidil support.illumina.com/sds.html.

Toote dokumendid: allalaaditavad PDF-vormingus Illumina veebisaidilt. Minge veebisaidile support.illumina.com, valige toode ning seejärel **Dokumendid ja kirjandus**.



Illumina
5200 Illumina Way
San Diego, California 92122 USA
+ 1 800 809 ILMN (4566)
+1 85 8202 4566 (väljaspool Põhja-Ameerikat)
techsupport@illumina.com
www.illumina.com

CE



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

Sponsor Austraalias

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

IN VITRO DIAGNOSTILISEKS KASUTAMISEKS

© 2021 Illumina, Inc. Kõik õigused on kaitstud.

illumina®