

Modul za analizu Germline Variant lokalnog upravljača za pokretanje

Vodič za tok rada za NextSeq 550Dx

ZA IN VITRO DIJAGNOSTIČKU UPOTREBU

Pregled	3
Unos informacija o obradi	3
Metode analize	6
Prikaz podataka o obradi i uzorku	7
Izveštaj o analizi	7
Izlazne datoteke analize	9
Istorija revizija	16
Tehnička pomoć	17



Ovaj dokument i njegov sadržaj su u vlasništvu kompanije Illumina, Inc. i njenih podružnica („Illumina“) i namenjeni su isključivo za ugovorno korišćenje njenih kupaca u vezi sa korišćenjem proizvoda koji su ovde opisani i ni za šta drugo. Ovaj dokument i njegov sadržaj ne smeju se koristiti niti distribuirati ni za koju drugu svrhu niti se smeju prenositi, otkrivati ili reprodukovati ni na koji način bez prethodnog pisanog pristanka kompanije Illumina. Illumina ne prenosi nikakvu licencu pod patentom, robnom markom, autorskim pravom ili javnim pravom niti sličnim pravima bilo kog trećeg lica prema ovom dokumentu.

Stručna i adekvatno obučena lica moraju strogo i izričito da poštuju uputstva u ovom dokumentu kako bi se obezbedila ispravna i bezbedna upotreba ovde opisanih proizvoda. Pre upotrebe tih proizvoda obavezno je u potpunosti pročitati i razumeti celokupnu sadržinu ovog dokumenta.

UKOLIKO NE PROČITATE I NE PRATITE OVO UPUTSTVO U CELOSTI, TO MOŽE DA DOVEDE DO OŠTEĆENJA PROIZVODA, POVREDA LICA, KAO ŠTO SU KORISNICI ILI DRUGA LICA, I OŠTEĆENJA DRUGE IMOVINE I TIME ĆE SE PONIŠTITI SVAKA GARANCIJA KOJA SE ODNOSI NA PROIZVODE.

KOMPANIJA ILLUMINA NE PREUZIMA NIKAKVU ODGOVORNOST USLED NEADEKVATNE UPOTREBE OVDEOPISANIH PROIZVODA (UKLJUČUJUĆI I NJIHOVE DELOVE ILI SOFTVER).

© 2021. Illumina, Inc. Sva prava zadržana.

Svi žigovi su vlasništvo kompanije Illumina, Inc. ili odgovarajućih vlasnika. Konkretno informacije o žigovima potražite na adresi www.illumina.com/company/legal.html.

Pregled

Modul Germline Variant lokalnog upravljača za pokretanje se koristi sa analizom Illumina TruSeq Custom Amplicon Kit Dx i instrumentom NextSeq 550Dx. Kada se koristi sa modulom Germline Variant, ova analiza je namenjena za pripremu biblioteka koje se koriste za sekvenciranje DNK iz perifernih uzoraka pune krvi.

Modul za analizu procenjuje kratke regione amplifikovane (umnožene) DNK, tj. amplikona. Fokusirano sekvenciranje amplikona omogućava veliku pokrivenost određenih regiona. Pogledajte uputstvo u pakovanju za komplet *TruSeq Custom Amplicon Kit Dx* (br. dokumenta 1000000029772).

Za rad modula za analizu Germline Variant potreban je potrošni materijal za sekvenciranje u 300 ciklusa. Više informacija potražite u uputstvu u pakovanju za komplet *NextSeq 550Dx High Output Reagent Kit v2* ili *NextSeq 550Dx High Output Reagent Kit v2.5*.

O ovom vodiču

Ovaj vodič navodi uputstva za podešavanje parametara obrade za sekvenciranje i analizu pomoću modula za analizu Germline Variant. Više informacija o kontrolnoj tabli lokalnog upravljača za pokretanje i podešavanjima sistema potražite u *referentnom vodiču za instrument NextSeq 550Dx* (br. dokumenta 1000000009513).

Prikaz lokalnog upravljača za pokretanje

Interfejs lokalnog upravljača za pokretanje se prikazuje u okviru softvera NextSeq 550Dx Operating Software (NOS) ili putem veb pretraživača. Podržani veb pretraživač je Chromium.



NAPOMENA

Ako koristite nepodržani pretraživač, preuzmite podržani pretraživač kada se to od vas zatraži u okviru poruke „Confirm Unsupported Browser“ (Potvrda nepodržanog pretraživača). Izaberite „**ovde**“ da biste preuzeli podržanu verziju pretraživača Chromium.

Prikaz na monitoru za instrumente

- 1 Da bi se interfejs lokalnog upravljača za pokretanje prikazivao na monitoru za instrumente, izaberite jednu od sledećih opcija:
 - ▶ Na početnom ekranu softvera NOS izaberite opciju **Local Run Manager** (Lokalni upravljač za pokretanje).
Kliknite na X u gornjem desnom uglu da biste se vratili u NOS kada završite.
 - ▶ Izaberite ikonu za umanjivanje softvera NOS, otvorite veb pretraživač Chromium koji se nalazi na instrumentu i unesite **http://localhost** u traku za adresu.
Samo administratorski korisnici mogu da umanje NOS.

Prikaz sa umreženog računara

- 1 Otvorite veb pretraživač Chromium na računaru koji ima pristup istoj mreži kao i instrument, pa se povežite koristeći IP adresu instrumenta ili naziv instrumenta. Na primer, **http://myinstrument**.

Unos informacija o obradi

Podešavanje parametara

- 1 Prijavite se na lokalni upravljač za pokretanje.
- 2 Izaberite **Create Run** (Kreiraj obradu), pa **Germline Variant**.

- Unesite naziv obrade po kom ćete razlikovati obradu od sekvenciranja preko analize. Koristite alfanumeričke znakove, razmake, donje crte ili crtice.
- [Opcionalno]** Unesite opis obrade da biste je lakše identifikovali. Koristite alfanumeričke znakove, razmake, donje crte ili crtice.
- Na padajućoj listi izaberite broj uzoraka i skup indeksa. Prilikom izbora uzmite u obzir sledeće informacije.
 - ▶ Padajuća lista sadrži brojeve uzoraka sa skupom indeksa. Na primer, 24-Set 1 znači da će se testirati 24 uzorka sa indeksima iz skupa indeksa 1.
 - ▶ Brojevi skupova indeksa se odnose na različite skupove i5 indeksa. I Skup 1 i skup 2 omogućavaju raznovrsnost indeksa. Ponuđena su dva skupa indeksa kako bi se sprečilo iscrpljivanje jednog skupa.
 - ▶ Odaberite broj uzoraka koji je najbliži broju uzoraka koji testirate. Ako tačan broj uzoraka nije na listi, izaberite broj koji je najbliži, ali manji od broja koji testirate. Na primer, ako želite da testirate 18 uzoraka, izaberite 16 uzoraka.
 - ▶ Otvori sa uzorcima i kombinacije indeksa koji ispunjavaju uslove za raznovrsnost indeksa istaknuti su zelenom bojom. Ako izaberete druge otvore i kombinacije indeksa, nakon čuvanja obrade ćete primiti obaveštenje ako uslovi za raznovrsnost indeksa nisu ispunjeni.

Uvoz datoteka manifesta za obradu

- Uverite se da su manifesti koje želite da uvezete dostupni na pristupačnoj mrežnoj lokaciji ili na USB disku.
- Izaberite **Import Manifests** (Uvezi manifeste).
- Dodite do datoteke manifesta i izaberite manifest koji želite da dodate.



NAPOMENA

Da bi datoteke manifesta bile dostupne za sve obrade pomoću modula za analizu Germline Variant, dodajte manifeste koristeći funkciju „Module Settings“ (Podešavanja modula). Za korišćenje ove funkcije potrebne su administratorske dozvole. Više informacija potražite u *referentnom vodiču za instrument NextSeq 550Dx (br. dokumenta 1000000009513)*.

Navođenje uzoraka za obradu


Navedite uzorke za obradu pomoću jedne od opcija i sledećih smernica.

- ▶ **Enter samples manually** (Ručni unos uzoraka) – koristite praznu tabelu na ekranu „Create Run“ (Kreiraj obradu).
- ▶ **Import samples** (Uvoz uzoraka) – otvorite spoljnu datoteku u formatu vrednosti odvojenih zarezom (*.csv). Predložak možete da preuzmete na ekranu „Create Run“ (Kreiraj obradu).


Kada popunite tabelu sa uzorcima, možete da izvezete informacije o uzorcima u spoljnu datoteku. Koristite ovu datoteku kao referencu prilikom pripreme biblioteka ili je uvezite za drugu obradu.

Ručni unos uzoraka

- Unesite jedinstveni ID uzorka u polje „Sample ID“ (ID uzorka). Koristite alfanumeričke znakove, crtice ili donje crte.

- 2 **[Opcionalno]** Za pozitivne ili negativne kontrolne uzorke kliknite desnim tasterom i izaberite vrstu kontrole.
- 3 **[Opcionalno]** Unesite opis uzorka u polje „Sample Description“ (Opis uzorka).
Koristite alfanumeričke znakove, crtice ili donje crte.
Opisi uzoraka se povezuju sa nazivom uzorka. Opisi uzoraka se zamenjuju ako se isti naziv uzorka ponovo koristi u kasnijoj obradi.
- 4 Izaberite adapter za indeks 1 u padajućoj listi za indeks 1 (i7).
Kada koristite predložene otvore sa uzorcima, softver automatski popunjava i7 i i5 adaptore za indeks koji ispunjavaju uslove za raznovrsnost indeksa. Ako tačan broj uzoraka koji testirate nije na listi, obavezno izaberite adaptore indeksa za dodatne otvore. Ako treba da izaberete indekse za dodatne otvore ili ne koristite predložene kombinacije adaptera indeksa, pre izbora indeksa obavezno pročitajte odeljak *Određivanje očitanih baza i raznovrsnost indeksa na stranici 15*.
- 5 Izaberite adapter za indeks 2 na padajućoj listi za indeks 2 (i5).
- 6 Izaberite datoteku manifesta na padajućoj listi Manifest.
- 7 Izaberite opciju za prikaz, štampanje ili čuvanje rasporeda pločice kao reference za pripremu biblioteka:
 - ▶ Izaberite ikonu  **Print** (Štampaj) da biste prikazali raspored pločice. Izaberite **Print** (Štampaj) da biste štampali raspored pločice.
 - ▶ Izaberite **Export** (Izvoz) da biste izvezli informacije o uzorcima u spoljnu datoteku.Uverite se da su informacije iz manifesta i informacije o uzorcima tačne. Netačne informacije mogu da utiču na rezultate.
- 8 Izaberite **Save Run** (Sačuvaj obradu).

Uvoz uzoraka

- 1 Izaberite **Import Samples** (Uvezi uzorke) i idite do lokacije datoteke sa informacijama o uzorcima.
Možete da uvezete dve vrste datoteka.
 - ▶ Na ekranu „Create Run“ (Kreiraj obradu) izaberite **Template** (Predložak) da biste napravili novi raspored pločice. Datoteka predložka sadrži tačna zaglavlja kolona za uvoz. Unesite informacije o uzorku u svaku kolonu za uzorke u obradi. Izbrišite primere informacija iz ćelija koje se ne koriste, pa sačuvajte datoteku.
 - ▶ Upotrebite datoteku sa informacijama o uzorcima koja je izvezena iz modula Germline Variant pomoću funkcije „Export“ (Izvoz).
- 2 Izaberite ikonu  **Print** (Štampaj) da biste prikazali raspored pločice.
- 3 Izaberite **Print** (Štampaj) da biste odštampali raspored pločice kao referencu za pripremu biblioteka.
- 4 **[Opcionalno]** Izaberite **Export** (Izvoz) da biste izvezli informacije o uzorcima u spoljnu datoteku.
Uverite se da su informacije iz manifesta i informacije o uzorcima tačne. Netačne informacije mogu da utiču na rezultate.
- 5 Izaberite **Save Run** (Sačuvaj obradu).

Izmena obrade

Uputstva za izmenu informacija o obradi pre sekvenciranja potražite u *referentnom vodiču za instrument NextSeq 550Dx (br. dokumenta 1000000009513)*.

Metode analize

Modul za analizu Germline Variant obavlja sledeće korake analize, a zatim upisuje izlazne datoteke analize u fasciklu Alignment (Poravnanje).

- ▶ Očitavanja indeksa demultipleksiranja
- ▶ Generisanje FASTQ datoteka
- ▶ Poravnanja u odnosu na referencu
- ▶ Identifikovanje varijanti

Demultipleksiranje

Demultipleksiranje upoređuje svaku sekvencu čitanja indeksa sa sekvencama indeksa navedenim za obradu. U ovom koraku se ne razmatraju vrednosti za kvalitet.

Očitavanja indeksa se identifikuju pomoću sledećih koraka:

- ▶ Uzorci se numerišu počev od 1 na osnovu redosleda kojim su navedeni za obradu.
- ▶ Broj uzorka 0 je rezervisan za klasterne koji nisu dodeljeni uzorku.
- ▶ Klasteri se dodeljuju uzorku kada se sekvenca indeksa tačno podudara ili kada postoji najviše jedno nepodudaranje po očitavanju indeksa.

Generisanje FASTQ datoteke

Nakon demultipleksiranja softver generiše posredne datoteke analize u formatu FASTQ, što je tekstualni format koji se koristi za predstavljanje sekvenci. FASTQ datoteke sadrže očitavanja za svaki uzorak i povezane ocene kvaliteta. Klasteri koji nisu prošli filter se izostavljaju.

Svaka FASTQ datoteka sadrži očitavanja samo za jedan uzorak, a naziv tog uzorka uvršten je u naziv FASTQ datoteke. FASTQ datoteke su primarni ulaz za poravnanje. Za svaki uzorak se generiše osam FASTQ datoteka, četiri iz 1. očitavanja i četiri iz 2. očitavanja.

Poravnanje

Tokom koraka poravnanja, povezani Smit-Vatermanov algoritam poravnjava klasterne iz svakog uzorka sa sekvencama amplikona navedenim u datoteci manifesta.

Smit-Vatermanov algoritam obavlja poluglobalno poravnanje sekvenci da bi odredio slične regione između dve sekvence. Smit-Vatermanov algoritam ne poredi celu sekvencu, već segmente svih mogućih dužina.

Svako očitavanje sa uparenim krajevima se procenjuje u pogledu poravnanja sa relevantnim probnim sekvencama za to čitanje.

- ▶ Prvo očitavanje se procenjuje u odnosu na reverzni komplement silaznih oligonukleotida specifičnih za lokus (Downstream Locus-Specific Oligos, DLSO).
- ▶ Drugo očitavanje se procenjuje u odnosu na uzlazne oligonukleotide specifične za lokus (Upstream Locus-Specific Oligos, ULSO).
- ▶ Ako se početak očitavanja podudara sa probnom sekvencom uz ne više od tri razlike (nepodudaranja ili pomaci zbog vodećih indela), puna dužina očitavanja usklađuje se s amplikonskim ciljem za tu sekvencu.
- ▶ Indeli u DLSO-u i ULSO-u se ne razmatraju zbog hemije analize.

Poravnanja se filtriraju iz rezultata poravnanja na osnovu stopa nepodudaranja u regionu od interesa ili celom amplikonu, u zavisnosti od dužine amplikona. Filtrirana poravnanja se zapisuju u datoteke poravnanja kao neporavnata i ne koriste se za određivanje varijanti.

Određivanje varijanti

Pisces modul za određivanje varijanti, koji je razvila kompanija Illumina, otkriva varijante prisutne u uzorku DNK.



Modul za određivanje varijanti Pisces identifikuje jednonukleotidne varijante (Single Nucleotide Variant, SNV), višenuklotidne varijante (Multiple Nucleotide Variant, MNV) i male indele, u sledeća 3 koraka:

- ▶ Uzima u obzir svaku poziciju u referentnom genomu zasebno
- ▶ Broji baze na datoj poziciji i traži poravnata očitavanja koja se preklapaju sa tom pozicijom
- ▶ Izračunava ocenu varijante kojom se meri kvalitet otkrivanja na osnovu Poasonovog modela. Varijante čija je ocena kvaliteta manja od Q20 se izuzimaju.

Ako varijanta prođe sve filtere, ona se označava sa PASS (ZADOVOLJAVA) u VCF datoteci.

Za više informacija, pogledajte github.com/Illumina/Pisces/wiki.

Prikaz podataka o obradi i uzorku

- 1 Na kontrolnoj tabli lokalnog upravljača za pokretanje kliknite na naziv obrade.
- 2 Na kartici „Run Overview“ (Pregled obrade) pregledajte pokazatelje obrade sekvenciranjem.
- 3 **[Opcionalno]** Kliknite na ikonu **Copy to Clipboard**  (Kopiraj u međumemoriju) da biste kopirali putanju izlazne fascikle obrade.
- 4 Kliknite na karticu „Sequencing Information“ (Informacije o sekvenciranju) da biste pregledali parametre obrade i informacije o potrošnom materijalu.
- 5 Kliknite na karticu „Samples and Results“ (Uzorci i rezultati) da biste videli lokaciju izveštaja o analizi.
 - ▶ Ako je analiza ponovljena, proširite padajuću listu „Select Analysis“ (Izaberi analizu) i izaberite odgovarajuću analizu.
- 6 Kliknite na ikonu **Copy to Clipboard**  (Kopiraj u međumemoriju) da biste kopirali putanju fascikle „Analysis“.

Da biste videli više informacija o karticama „Run Overview“ (Pregled obrade) i „Sequencing Information“ (Informacije o sekvenciranju), kao i o tome kako da ponovo stavite analizu u red, pročitajte *referentni vodič za instrument NextSeq 550Dx (br. dokumenta 1000000009513)*.

Izveštaj o analizi

Rezime rezultata analize je prikazan na karticama „Samples“ (Uzorci) i „Results“ (Rezultati), dok se u fascikli „Alignment“ (Poravnanje) nalazi objedinjeni izveštaj. Pored toga, za svaki uzorak je dostupan izveštaj u formatu PDF datoteke.

Informacije na karticama Samples (Uzorci) i Results (Rezultati)

1 Kliknite na uzorak na listi da biste videli izveštaj o uzorku.

Tabela 1 Informacije o obradi i uzorku

Zaglavlje kolone	Opis
Run Status (Status obrade)	Pokazuje da li je obrada sekvenciranjem uspjela ili nije.
Total Yield (GB) (Ukupan prinos (GB))	Broj baza otkrivenih u obradi sekvenciranjem. Pokazuje prag prolaznosti i uspješan ili neuspješan status.
% ≥ Q30	Procenat očitavanja u obradi sekvenciranjem čija je ocena kvaliteta 30 (Q30) ili veća. Pokazuje prag prolaznosti i uspješan ili neuspješan status.
Sample ID (ID uzorka)	ID uzorka koji je naveden prilikom pravljenja obrade.
Total PF Reads (Ukupan broj očitavanja sa statusom PF)	Ukupan broj očitavanja koja su prošla filter
Read 1% ≥ Q30 (% ≥ Q30 za 1. očitavanje)	Procenat očitavanja u 1. očitavanju za uzorak čija je ocena kvaliteta 30 (Q30) ili veća.
Read 2% ≥ Q30 (% ≥ Q30 za 2. očitavanje)	Procenat očitavanja u 2. očitavanju za uzorak čija je ocena kvaliteta 30 (Q30) ili veća.
Autosome Call Rate (Stopa otkrivanja autozoma)	Broj genomskih pozicija u autozomima (od 1. do 22. hromozoma) koje su dostigle unapred definisanu graničnu vrednost za pouzdanost, podeljen ukupnim brojem ispitivanih autozomnih genomskih pozicija. Stopa otkrivanja se opisuje za svaki uzorak i prikazuje kao procenat. Izračunava se po formuli 1 minus (broj autozomnih pozicija sa nepotpunim određivanjem podeljen ukupnim brojem sekvenciranih autozomnih pozicija).

Tabela 2 Informacije u izveštaju o uzorku

Zaglavlje kolone	Opis
Sample (Uzorak)	ID uzorka koji je naveden prilikom pravljenja obrade.
Report Date (Datum izveštaja)	Datum generisanja izveštaja.
Sample Information (Informacije o uzorku)	ID uzorka koji je naveden prilikom pravljenja obrade, ukupan broj očitavanja u uzorku koja su prošla filter, procenat očitavanja za uzorak čija je ocena kvaliteta 30 (Q30) ili veća, stopa otkrivanja autozoma.
Amplicon Summary (Rezime amplikona)	Ukupan broj sekvenciranih regiona amplikona i ukupna dužina u parovima baza sekvenciranih amplikona u ciljnim regionima, za uzorak i datoteku manifesta. Datoteka manifesta određuje referentni genom i ciljne referentne regione koji se koriste u koraku poravnanja.
Read Level Statistics (Statistički podaci na nivou očitavanja)	Broj i procenat očitavanja za uzorak koji pokrivaju svaku poziciju u referenci, za 1. očitavanje i 2. očitavanje.
Variants Summary (Rezime varijanti)	Broj SNV-ova, umetanja (inercija) i brisanja (delecija) prepoznatih za uzorak koji je premašio predložene vrednosti pomoću kojih se određuje da li su rezultati kvaliteta u prihvatljivom opsegu.
Coverage Summary (Rezime pokrivenosti)	Ukupan broj poravnatih baza podeljen veličinom ciljnog regiona, kao i procenat amplikonskih regiona čije su vrednosti pokrivenosti veće od niže granične vrednosti za pokrivenost od 0,2 * srednja pokrivenost amplikona, za uzorak.
Coverage Plots (Grafikoni pokrivenosti)	Grafikon pokrivenosti prema amplikonskom regionu prikazuje pokrivenost u amplikonskim regionima za uzorak. Regioni sa vrednostima pokrivenosti manjim od granične vrednosti za pokrivenost istaknuti su crvenom bojom. Prosek svih vrednosti je naznačen narandžastom crtom.

Zaglavljje kolone	Opis
Software Versions (Verzije softvera)	Verzije softvera kada je uzorak sekvenciran. Obuhvata softver NextSeq 550Dx Operating Software (NOS), softver Local Run Manager (Lokalni upravljač za pokretanje), softver RTA i verziju Germline Variant modula.

Izlazne datoteke analize

Sledeće izlazne datoteke analize se generišu za modul za analizu Germline Variant i sadrže rezultate analize za poravnanje i određivanje varijanti. Izlazne datoteke analize se nalaze u fascikli „Alignment“ (Poravnanje).

Naziv datoteke	Opis
Demultipleksiranje (*.txt)	Sporedne datoteke koje sadrže rezime rezultata demultipleksiranja.
FASTQ (*.fastq.gz)	Posredne datoteke koje sadrže određivanja očitanih baza za koja postoji ocena kvaliteta. FASTQ datoteke su primarne ulazne datoteke za korak poravnanja.
Datoteke poravnanja u formatu BAM (*.bam)	Sadrže poravnata očitavanja za dati uzorak.
Datoteke određivanja varijanti u genomskom formatu VCF (*.genome.vcf.gz)	Sadrže genotip za svaku poziciju, bilo da je određen kao varijanta ili kao referenca.
Datoteke određivanja varijanti, u formatu VCF (*.vcf.gz)	Sadrži sve određene varijante u celom ciljnom regionu.
AmpliconCoverage_M1.tsv	Sadrže informacije o pokrivenosti prema amplikonu i uzorku za svaki navedeni manifest. M# predstavlja broj manifesta.

Format datoteka za demultipleksiranje

Proces demultipleksiranja očitava sekvence indeksa priključene svakom klasteru kako bi se utvrdilo iz kog uzorka je potekao klaster. Mapiranje između klastera i broja uzorka se upisuje u datoteku za demultipleksiranje (*.demux) za svaku pločicu ćelije toka.

Format za imenovanje datoteke za demultipleksiranje je `s_1_X.demux`, gde je X broj pločice.

Datoteka za demultipleksiranje počinje zaglavljem:

- ▶ Verzija (ceo broj od 4 bajta), trenutno 1
- ▶ Broj klastera (ceo broj od 4 bajta)

Ostatak datoteke sastoji se od brojeva uzoraka za svaki klaster sa pločice.

Kada se korak demultipleksiranja završi, softver generiše datoteku za demultipleksiranje sa nazivom `DemultiplexSummaryF1L1.txt`.

- ▶ U nazivu datoteke **F1** predstavlja broj ćelije toka.
- ▶ U nazivu datoteke **L1** predstavlja broj trake.
- ▶ Rezultati demultipleksiranja u tabeli sa 1 redom po pločici i 1 kolonom po uzorku, uključujući uzorak 0.
- ▶ Sekvence koje se najčešće pojavljuju u očitavanjima indeksa.

Format datoteke FASTQ

FASTQ je format tekstualne datoteke koji sadrži određivanja očitanih baza i vrednosti za kvalitet po očitavanju. Svaki zapis se sastoji od 4 reda:

- ▶ identifikatora

- ▶ sekvence
- ▶ znaka plus (+)
- ▶ Phred ocene kvaliteta u formatu kodiranom prema standardu ASCII + 33

Identifikator je formatiran kao:

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

Primer:

```
@SIM:1:FCX:1:15:6329:1045 1:N:0:2
TCGCACTCAACGCCCTGCATATGACAAGACAGAATC
+
<>;##=><9=AAAAAAAAAAA9#:<#<;<<<????#=#
```

Format datoteke BAM

BAM datoteka (*.bam) je komprimovana binarna verzija SAM datoteke koja se koristi za predstavljanje poravnatih sekvenci do 128 Mb. Formati SAM i BAM su detaljno opisani na samtools.github.io/hts-specs/SAMv1.pdf.

BAM datoteke koriste format imenovanja **NazivUzorka_S#.bam**, gde je # broj uzorka koji se određuje redosledom kojim su uzorci navedeni za obradu.

BAM datoteke sadrže odeljak zaglavlja i odeljak za poravnanja:

- ▶ **Zaglavlje** — Sadrži informacije o celoj datoteci, kao što su naziv uzorka, dužina uzorka i metod poravnanja. Poravnanja u odeljku za poravnanja su povezana sa konkretnim informacijama u odeljku zaglavlja.
- ▶ **Poravnanja** — Sadrži naziv očitavanja, sekvencu očitavanja, kvalitet očitavanja, informacije o poravnanju i prilagođene oznake. Naziv očitavanja obuhvata hromozom, početnu koordinatu, kvalitet poravnanja i nisku opisa podudaranja.

Odeljak za poravnanja obuhvata sledeće informacije za svako očitavanje ili par očitavanja:

- ▶ **AS:** kvalitet poravnanja uparenog kraja
- ▶ **RG:** grupa očitavanja, koja pokazuje broj očitavanja za određeni uzorak.
- ▶ **BC:** oznaka barkoda, koja pokazuje demultipleksirani ID uzorka koji je povezan sa očitavanjem.
- ▶ **SM:** kvalitet poravnanja jednog kraja.
- ▶ **XC:** niska opisa podudaranja
- ▶ **XN:** oznaka imena amplicona, koja beleži ID amplicona povezanog sa očitavanjem

BAM indeksne datoteke (*.bam.bai) pružaju indeks odgovarajuće BAM datoteke.

Format datoteke VCF

Variant Call Format (VCF, format otkrivanja varijante) je uobičajen format datoteke koji je razvila genomska naučna zajednica. Sadrži informacije o varijantama pronađenim na određenim pozicijama u referentnom genomu. VCF datoteke se završavaju sufiksom .vcf

Zaglavlje VCF datoteke sadrži verziju formata datoteke VCF, verziju modula za određivanje varijanti i listu oznaka koje se koriste u ostatku datoteke. VCF zaglavlje obuhvata i datoteku referentnog genoma i BAM datoteku. Poslednji red zaglavlja sadrži naslove kolona za redove podataka. Svaki red podataka u VCF datoteci sadrži informacije o jednoj varijanti.

Zaglavlja VCF datoteka

Zaglavlje	Opis
CHROM	Hromozom referentnog genoma. Hromozomi se prikazuju istim redosledom kao u referentnoj FASTQ datoteci.
POS	Pozicija varijante sa jednom bazom u referentnom hromozomu. Za SNP-ove, ova pozicija je referentna baza u varijanti; za indele ili brisanja, ova pozicija je referentna baza neposredno pre varijante.
ID	Rs broj varijante dobijen iz datoteke dbSNP.txt, ako je primenljivo. Ako na tom mestu postoji više rs brojeva, lista je razdvojena tačkom i zarezom. Ako na toj poziciji nema dbSNP unosa, koristi se marker za vrednost koja nedostaje ('.').
REF	Referentni genotip. Na primer, brisanje jednog T predstavljeno je kao referentni TT i alternativni T. Jedna varijanta nukleotida A na T predstavljena je kao referentni A i alternativni T.
ALT	Aleli koji se razlikuju od referentnog očitavanja. Na primer, umetanje jednog T predstavljeno je kao referentni A i alternativni AT. Jedna varijanta nukleotida A na T predstavljena je kao referentni A i alternativni T.
QUAL	Phred ocena kvaliteta koju dodeljuje modul za određivanje varijanti. Veće ocene ukazuju na veću pouzdanost varijante i manju mogućnost grešaka. Za ocenu kvaliteta Q, procenjena verovatnoća greške iznosi $10^{-(Q/10)}$. Na primer, skup od Q30 otkrivanja ima stopu grešaka od 0,1%. Mnogi moduli za određivanje varijanti dodeljuju ocene kvaliteta na osnovu svojih statističkih modela. Te ocene su visoke u odnosu na uočene stope grešaka.

Oznake u VCF datoteci

Zaglavlje	Opis
FILTER	<p>Ako je datoteka prošla sve filtere, u koloni filtera piše PASS (ZADOVOLJAVA).</p> <ul style="list-style-type: none"> • LowDP – odnosi se na mesta za koja je dubina pokrivenosti manja od 150x. Za pozicije amplikona koje su pokrivene i očitavanjem unapred i reverznim očitavanjem, to je ekvivalentno broju od 300 preklapljenih očitavanja uparenih krajeva. • q20 – ocena kvaliteta < 20. • MultiAllelicSite – varijanta nije usaglašena sa diploidnim modelom. • R5x9 – broj susednih ponavljanja (dužine od 1 do 5 bp) za otkrivanja varijanti ≥ 9. • SB – otklon lanca je veći od date granične vrednosti.
INFO	<p>Mogući unosi u koloni INFO obuhvataju sledeće:</p> <ul style="list-style-type: none"> • AC – broj alela u genotipovima za svaki ALT alel, istim redosledom kojim su navedeni. • AF – frekvencija alela za svaki ALT alel, istim redosledom kojim su navedeni. • AN – ukupan broj alela u otkrivenim genotipovima. • CD – oznaka da se SNP pojavljuje u regionu kodiranja barem 1 unosa RefGene. • DP – dubina (broj otkrivanja baza poravnatih u odnosu na poziciju i korišćenih u otkrivanju varijanti). • Exon – lista očitavanja regiona egzona, odvojenih zarezom, iz datoteke RefGene. • FC – funkcionalna posledica. • GI – lista očitavanja ID-ova gena, odvojenih zarezom, iz datoteke RefGene. • QD – pouzdanost varijante i kvalitet prema dubini. • TI – lista očitavanja ID-ova transkripta, odvojenih zarezom, iz datoteke RefGene.
FORMAT	<p>U koloni Format su navedena polja koja su odvojena pomoću dve tačke. Na primer, GT:GQ. Navedena lista polja zavisi od korišćenog modula za određivanje varijanti. Dostupna su sledeća polja:</p> <ul style="list-style-type: none"> • AD – unos u obliku X,Y, gde je X broj referentnih otkrivanja, a Y broj alternativnih otkrivanja. • DP – približna dubina čitanja; filtriraju se očitavanja za koja je MQ=255 ili koja sadrže loše parove. • GQ – kvalitet genotipa. • GT – genotip. 0 odgovara referentnoj bazi, 1 odgovara prvom unosu u koloni ALT itd. Kosa crta (/) ukazuje da nema dostupnih informacija o fazama. • NC – deo baza koje nisu otkrivene ili kod kojih je kvalitet otkrivanja baze ispod minimalnog praga. • NL – nivo šuma; procena šuma u otkrivanju baze na ovoj poziciji. • SB – otklon lanca na ovoj poziciji. Veće negativne vrednosti ukazuju na manji otklon; vrednosti blizu 0 ukazuju na veći otklon. • VF – frekvencija varijante; procenat očitavanja koja podržavaju alternativni alel.
SAMPLE (Uzorak)	Kolona SAMPLE sadrži vrednosti navedene u koloni FORMAT.

Genomske VCF datoteke

Genomske VCF (gVCF) datoteke su VCF datoteke verzije 4.1 koje prate niz dogovorenih pravila za predstavljanje svih mesta u genomu u relativno kompaktnom obliku. gVCF (*.genome.vcf.gz) datoteke obuhvataju sva mesta u regionu od interesa u jednoj datoteci za svaki uzorak.

gVCF datoteka ne pokazuje otkrivanja na pozicijama koje ne prođu sve filtere. Genotipska (GT) oznaka ./ pokazuje da nema otkrivanja.

Više informacija potražite na sites.google.com/site/gvcftools/home/about-gvcf.

Datoteka pokrivenosti amplikona

Za svaku datoteku manifesta se generiše datoteka pokrivenosti amplikona. M# u nazivu datoteke predstavlja broj manifesta.

Svaka datoteka sadrži red zaglavlja koji obuhvata ID-ove uzoraka povezane s manifestom. Ta datoteka sadrži sledeće informacije.

- ▶ ID cilja onako kako je naveden u manifestu.
- ▶ Dubina pokrivenosti očitavanja koja su prošla filter.

Dodatne izlazne datoteke

Sledeće izlazne datoteke sadrže dodatne informacije ili rezime rezultata obrade i grešaka u analizi.

Te datoteke nisu potrebne za procenu rezultata analize, ali mogu da se koriste za rešavanje problema.

Ako nije drugačije navedeno, sve datoteke se nalaze u fascikli „Alignment“ (Poravnanje).

Naziv datoteke	Opis
AnalysisLog.txt	Evidencija obrade koja opisuje svaki korak obavljen tokom analize trenutne fascikle obrade. Ova datoteka ne sadrži poruke o greškama. Nalazi se u fascikli „Alignment“ (Poravnanje).
AnalysisError.txt	Evidencija obrade koja navodi sve greške do kojih je došlo tokom analize. Ova datoteka će biti prazna ako nije bilo grešaka. Nalazi se u fascikli „Alignment“ (Poravnanje).
DemultiplexSummaryF1L1#.txt	Izveštava o rezultatima demultipleksiranja u tabeli sa 1 redom po pločici i 1 kolonom po uzorku. # predstavlja trake 1, 2, 3 ili 4 ćelije toka. Nalazi se u fascikli „Alignment“ (Poravnanje).
AmpliconRunStatistics.xml	Sadrži rezime statističkih podataka specifičnih za obradu. Nalazi se u fascikli „Alignment“ (Poravnanje).

Fascikla „Analysis“ (Analiza)

Fascikla sa analizom sadrži datoteke koje generiše softver lokalnog upravljača za pokretanje.

Rezime odnosa između izlazne fascikle i fascikle sa analizom prikazan je u nastavku:

- ▶ Tokom sekvenciranja, analiza u realnom vremenu (Real-Time Analysis – RTA) popunjava izlaznu fasciklu datotekama generisanim tokom analize slike, određivanja očitanih baza i ocenjivanja kvaliteta.
- ▶ RTA u realnom vremenu kopira datoteke u fasciklu sa analizom. Kada RTA dodeli ocenu kvaliteta svakoj bazi iz svakog ciklusa, softver upisuje datoteku RTAComplete.txt u obe fascikle.
- ▶ Kada je prisutna datoteka RTAComplete.txt, počinje analiza.
- ▶ Kako se analiza nastavlja, lokalni upravljač za pokretanje upisuje izlazne datoteke u fasciklu sa analizom, a zatim te datoteke ponovno kopira u izlaznu fasciklu.

Fascikle „Alignment“ (Poravnanje)

Svaki put kada se analiza ponovo zakaže, lokalni upravljač za pokretanje pravi fasciklu poravnanja pod nazivom **Alignment_N** (Poravnanje_N), gde je N redni broj.

Struktura fascikle

- 📁 **Alignment** (Poravnanje) – Sadrži *.bam, *.vcf, FASTQ datoteke i datoteke specifične za modul za analizu.
 - 📁 **Date and Time Stamp** (Oznaka datuma i vremena) – Oznaka datum_vreme analize u formatu GGGGMMDD_ČČMMSS
 - 📄 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** – Sadrži *.bcl datoteke.
 - 📁 **L001** – Sadrži *.locs datoteke.
 - 📁 **RTA Logs** (RTA evidencija) – Sadrži datoteke evidencije nastale analizom softvera RTA.
 - 📁 **InterOp** – Sadrži binarne datoteke koje se koriste za izveštavanje o pokazateljima obrade sekvenciranjem.
 - 📁 **Logs** (Evidencija) – Sadrži datoteke evidencije koje opisuju korake obavljene tokom sekvenciranja.
 - 📄 RTAComplete.txt
 - 📄 RunInfo.xml
 - 📄 RunParameters.xml

Određivanje očitanih baza i raznovrsnost indeksa

Kada se uzorci sekvenciraju na instrumentu NextSeq 550Dx, određivanjem očitanih baza se definiše baza (A, C, G ili T) za svaki klaster date pločice ili područje snimanja na ćeliji toka u određenom ciklusu. Instrument NextSeq 550Dx koristi dvokanalno sekvenciranje koje pomoću samo dve slike šifruje podatke za četiri baze DNK, jednu iz crvenog i jednu iz zelenog kanala.

Postupak čitanja indeksa tokom određivanja očitanih baza razlikuje se od određivanja očitanih baza tokom drugih očitavanja.

Očitavanje indeksa moraju početi sa bar jednom bazom koja nije G u jednom od prvih dva ciklusa.

Ako očitavanje indeksa počinje sa dva određivanja očitanih baza G, ne generiše se intenzitet signala.

Signal mora biti prisutan u nekom od prvih dva ciklusa da bi se obezbedilo obavljanje demultipleksiranja.

Kada se tokom pravljenja obrade biraju indeksi, pojaviće se upozorenje na malu raznovrsnost ako indeksi ne ispunjavaju uslove za raznovrsnost. Da biste sprečili pojavu upozorenja na malu raznovrsnost, odaberite sekvence indeksa koje daju signal u oba kanala za svaki ciklus.

- ▶ Crveni kanal – A ili C
- ▶ Zeleni kanal – A ili T

Ovaj postupak određivanja očitanih baza osigurava tačnost prilikom analize uzoraka manje složenosti (low-plex). Više informacija o sekvencama indeksa potražite u uputstvu u pakovanju za komplet *TruSeq Custom Amplicon Kit Dx* (br. dokumenta 1000000029772).

Tokom pravljenja obrade u lokalnom upravljaču za pokretanje odabraćete broj uzoraka koji će biti testiran. Softver automatski popunjava predložene kombinacije indeksa koje ispunjavaju uslove za raznovrsnost indeksa. Korišćenje predloženih kombinacija indeksa nije obavezno, ali se preporučuje.

Istorija revizija

Dokument	Datum	Opis promene
Br. dokumenta 1000000030329 v04	Avgust 2021.	Ažurirana je adresa ovlašćenog predstavnika za EU.
Br. dokumenta 1000000030329 v03	Decembar 2019.	Ažurirana je adresa ovlašćenog predstavnika za EU. Ažurirana je adresa australijskog sponzora.
Br. dokumenta 1000000030329 v02	Januar 2019.	Dodate su informacije o v2.5 kompleta reagensa.
Br. dokumenta 1000000030329 v01	Avgust 2018.	Ažurirane su regulatorne oznake.
Br. dokumenta 1000000030329 v00	Novembar 2017.	Početno izdanje.

Tehnička pomoć

Za tehničku pomoć, obratite se tehničkoj podršci kompanije Illumina.

Veb-sajt: www.illumina.com
E-adresa: techsupport@illumina.com

Telefoni korisničke podrške kompanije Illumina

Region	Besplatan broj	Regionalni
Severna Amerika	+1.800.809.4566	
Australija	+1.800.775.688	
Austrija	+43 800006249	+43 19286540
Belgija	+32 80077160	+32 34002973
Danska	+45 80820183	+45 89871156
Finska	+358 800918363	+358 974790110
Francuska	+33 805102193	+33 170770446
Holandija	+31 8000222493	+31 207132960
Hongkong	800960230	
Irska	+353 1800936608	+353 016950506
Italija	+39 800985513	+39 236003759
Japan	0800.111.5011	
Kina	400.066.5835	
Nemačka	+49 8001014940	+49 8938035677
Norveška	+47 800 16836	+47 21939693
Novi Zeland	0800.451.650	
Singapur	+1.800.579.2745	
Španija	+34 911899417	+34 800300143
Švajcarska	+41 565800000	+41 800200442
Švedska	+46 850619671	+46 200883979
Tajvan	00806651752	
Ujedinjeno Kraljevstvo	+44 8000126019	+44 2073057197
Ostale zemlje	+44.1799.534000	

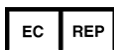
Safety data sheets (SDS) (Bezbednosno-tehnički listovi) — Dostupni su na veb-sajtu kompanije Illumina na adresi support.illumina.com/sds.html.

Dokumentacija o proizvodu — Dostupna za preuzimanje u PDF formatu sa veb-sajta kompanije Illumina. Idite na support.illumina.com, izaberite proizvod, pa zatim **Documentation & Literature** (Dokumentacija i literatura).



Illumina
5200 Illumina Way
San Diego, California 92122 SAD
+1.800.809.ILMN (4566)
+1.858.202.4566 (van Severne Amerike)
techsupport@illumina.com
www.illumina.com

CE



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

Australijski sponzor

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

ZA IN VITRO DIJAGNOSTIČKU UPOTREBU

© 2021. Illumina, Inc. Sva prava zadržana.

illumina®