

„Local Run Manager“ „Germline Variant“ analizės modulio

Darbo eigos vadovas, skirtas „NextSeq 550Dx“

NAUDOTI IN VITRO DIAGNOSTIKAI

Apžvalga	3
Mėginio informacijos įvedimas	3
Analizės metodai	6
Serijos ir mėginio duomenų peržiūra	7
Analizės ataskaita	7
Analizės išvesties failai	9
Keitimų istorija	16
Techninė pagalba	17



Šis dokumentas ir jo turinys priklauso „Illumina, Inc.“ ir jos filialams („Illumina“), jis skirtas tik klientui naudoti pagal sutartį, kiek tai susiję su čia aprašyto (-ų) produkto (-ų) naudojimu, ir jokių kitų tikslų. Šis dokumentas ir jo turinys negali būti naudojami ar platinami jokių kitų tikslų ir (arba) kitaip negali būti pateikiami, atskleidžiami ar atkuriami koku nors būdu be išankstinio rašytinio „Illumina“ sutikimo. „Illumina“ šiuo dokumentu neperduoda jokios trečiosios šalies licencijos pagal jos patentą, prekės ženklą, autorių teises, bendras teises nei panašių teisių.

Kvalifikuotas ir tinkamai išmokytas personalas turi griežtai ir aiškiai vadovautis šiame dokumente pateiktomis instrukcijomis, kad būtų užtikrintas tinkamas ir saugus šiame dokumente aprašyto (-ų) produkto (-ų) naudojimas. Prieš naudojant tokį (-ius) produktą (-us), visas šio dokumento turinys turi būti išsamiai perskaitytas ir suprastas.

JEI NEBUS PERSKAITYTOS VISOS ČIA PATEIKTOS INSTRUKCIJOS IR JOMIS NEBUS AIŠKIAI VADOVAUJAMASI, GALIMAS PRODUKTO (-Ų) PAŽEIDIMAS, NAUDOTOJO BEI KITŲ ASMENŲ SUŽEIDIMAS IR ŽALA KITAI NUOSAVYBEI, BE TO, TAI PANAIKINA PRODUKTUI (-AMS) TAIKOMOS GARANTIJOS GALIOJIMĄ.

„ILLUMINA“ NEPRISIIMA JOKIOS ATSAKOMYBĖS, JEI ČIA APRAŠOMAS (-I) PRODUKTAS (-AI) (ISKAITANT DALIS IR PROGRAMINĘ ĮRANGĄ) NAUDOJAMAS (-I) NETINKAMAI.

© 2021 m. „Illumina, Inc.“. Visos teisės saugomos.

Visi prekių ženklai yra „Illumina, Inc.“ arba atitinkamų savininkų nuosavybė. Informacijos apie konkrečius prekių ženklus ieškokite adresu www.illumina.com/company/legal.html.

Apžvalga

„Local Run Manager“ „Germline Variant“ modulis skirtas naudoti atliekant „Illumina TruSeq Custom Amplicon Kit Dx“ tyrimą ir kartu naudojant „NextSeq 550Dx“. Naudojant su „Germline Variant“ moduliu, tyrimo paskirtis yra paruošti bibliotekas, naudojamas nustatant DNR sekas iš periferinio viso kraujo bandinių.

Analizės modulis įvertina trumpus amplifikuotos DNR regionus, arba amplikonus, ieškodamas variantų. Sutelkiant dėmesį į amplikonų sekos nustatymą, užtikrinama didelė tam tikrų regionų aprėptis. Žr. „TruSeq Custom Amplicon Kit Dx“ (dokumento Nr. 1000000029772) pakuotės lapelį.

„Germline Variant“ analizės moduliui reikalingos 300 ciklų sekos nustatymo eksploatacinės medžiagos. Daugiau informacijos žr. „NextSeq 550Dx“ didelės išvesties reagentų rinkinio (2 versija) arba „NextSeq 550Dx“ didelės išvesties reagentų rinkinio (2.5 versija) pakuotės lapelyje.

Apie šį vadovą

Šiame vadove pateiktos instrukcijos, kaip nustatyti „Germline Variant“ analizės modulio sekoskaitos ir analizės serijos parametrus. Daugiau informacijos apie „Local Run Manager“ skydelį ir sistemos nuostatas žr. „NextSeq 550Dx“ prietaiso informaciniame vadove (dokumento Nr. 1000000009513).

„Local Run Manager“ peržiūra

„Local Run Manager“ sąsają galima peržiūrėti naudojant „NextSeq 550Dx“ operacinę programinę įrangą (NOS) arba žiniatinklio naršyklę. Palaikoma žiniatinklio naršyklė „Chromium“.



PASTABA

Jei naudojate nepalaikomą naršyklę, sulaukę raginimo pranešime „Confirm Unsupported Browser“ (patvirtinti nepalaikomą naršyklę) atsisiųskite palaikomą naršyklę. Norėdami atsisiųsti palaikomą „Chromium“ versiją, pasirinkite **čia**.

Peržiūra naudojant prietaiso monitorių

- 1 Jei „Local Run Manager“ sąsają norite peržiūrėti naudodami prietaiso monitorių, atlikite vieną iš toliau nurodytų veiksmų.
 - ▶ Pagrindiniame NOS ekrane pasirinkite „Local Run Manager“.
 - Viršutiniame dešiniajame kampe spustelėkite X, kad sugrįžtumėte į NOS, kai viską atliksite.
 - ▶ Pasirinkite NOS sumažinimo piktogramą, atidarykite prietaise įdiegtą žiniatinklio naršyklę „Chromium“, tada adreso juostoje įveskite **http://localhost**.
 - NOS gali sumažinti tik naudotojai, turintys administratoriaus teises.

Peržiūra naudojant prie tinklo prijungtą kompiuterį

- 1 Žiniatinklio naršyklę „Chromium“ atidarykite kompiuteryje, kuris turi prieigą prie to paties tinklo kaip prietaisas, tada prisijunkite įvedę prietaiso IP adresą arba pavadinimą. Pavyzdžiui, **http://myinstrument**.

Mėginio informacijos įvedimas

Parametrų nustatymas

- 1 Prisijunkite prie „Local Run Manager“.
- 2 Pasirinkite **Create Run** (kurti seriją), tada pasirinkite **Germline Variant**.

- Įveskite serijos pavadinimą, pagal kurį serija bus identifikuojama nuo sekoskaitos iki analizės. Naudokite raides, skaitmenis, tarpus, pabraukimo brūkšnį ar įprastą brūkšnelį.
- (Pasirinktinai)** Įveskite serijos aprašą, padedantį identifikuoti seriją. Naudokite raides, skaitmenis, tarpus, pabraukimo brūkšnį ar įprastą brūkšnelį.
- Išskleidžiamajame sąrašė pasirinkite mėginių skaičių ir indeksų rinkinį. Rinkdamiesi atkreipkite dėmesį į toliau pateiktą informaciją.
 - Išskleidžiamajame sąrašė nurodytas mėginių skaičius indeksų rinkinyje. Pavyzdžiui, „24-Set 1“ nurodo, kad turi būti tiriami 24 mėginiai, taikant indeksus iš 1 indeksų rinkinio.
 - Indeksų rinkinio skaičiai nurodo skirtingus „i5“ indeksų rinkinius. Tiek „Set 1“ (1 rinkinys), tiek „Set 2“ (2 rinkinys) užtikrina indeksų įvairovę. Du indeksų rinkiniai siūlomi siekiant išvengti vieno rinkinio išsekimo.
 - Pasirinkite mėginių skaičių, kuo panašesnę į tiriamų mėginių skaičių. Jei tikslaus mėginių skaičiaus sąrašė nėra, pasirinkite panašiausią, bet mažesnę už tiriamų mėginių skaičių. Pavyzdžiui, jei ketinate tirti 18 mėginių, pasirinkite 16.
 - Mėginių šulinėliai ir indeksų deriniai, atitinkantys indeksų įvairovės reikalavimus, paryškina žalia spalva. Pasirinkus kitus šulinėlius bei indeksų derinius ir bandant išsaugoti seriją, bus pateikiamas pranešimas, jei nesilaikoma indeksų įvairovės reikalavimų.

Serijos deklaracijos failų importavimas

- Patikrinkite, ar importuotinos deklaracijos prieinamos pasiekiamoje tinklo vietoje arba USB atmintuke.
- Pasirinkite **Import Manifests** (importuoti deklaracijas).
- Raskite deklaracijos failą ir pasirinkite norimas įtraukti deklaracijas.



PASTABA

Kad deklaracijų failai būtų prieinami visoms serijoms, kurias vykdo „Germline Variant“ analizės modulis, deklaracijos failus įtraukite naudodami funkciją „Module Settings“ (modulio nuostatos). Norint naudotis šia funkcija, reikia turėti administratoriaus lygio leidimus. Daugiau informacijos žr. „NextSeq 550Dx“ prietaiso informaciniame vadove (dokumento Nr. 100000009513).

Mėginių priskyrimas serijai


Priskirkite mėginius serijai, pasirinkdami vieną iš parinkčių ir vykdydami tolesnius nurodymus.

- Enter samples manually** (mėginius įvesti neautomatiškai) – naudokite tuščią lentelę ekrane „Create Run“ (kurti seriją).
- Import samples** (importuoti mėginius) – suraskite išorinį kableliais atskirtų verčių (*.csv) formato failą. Ekrane „Create Run“ (kurti seriją) galima atsisiųsti šabloną.


Užpildę mėginių lentelę, mėginio informaciją galite eksportuoti į išorinį failą. Šiuo failu remkitės ruošdami bibliotekas arba importuodami failą, skirtą kitai serijai.

Neautomatinis mėginių įvedimas

- Lauke „Sample ID“ (mėginio ID) įveskite išskirtinį mėginio ID. Naudokite raides, skaitmenis, tarpus, įprastus brūkšnelius ar pabraukimo brūkšnius.
- (Pasirinktinai)** Naudodami teigiamus ar neigiamus kontrolinius mėginius, spustelėkite dešiniąją pelės klavišą ir pasirinkite kontrolinės medžiagos tipą.

- 3 **(Pasirinktinai)** Lauke „Sample Description“ (mėginio aprašas) įveskite mėginio aprašą. Naudokite raides, skaitmenis, tarpus, įprastus brūkšnelius ar pabraukimo brūkšnius. Mėginio aprašai susieti su mėginio pavadinimu. Mėginio aprašai perrašomi, jei tas pats mėginio pavadinimas dar kartą naudojamas vykdant tolesnę seriją.
- 4 „Index 1“ („i7“) (1 indeksas) išskleidžiamajame sąrašė pasirinkite 1 indekso adapterį. Jei naudojate siūlomus mėginių šulinėlius, programinė įranga automatiškai parenka „i7“ ir „i5“ indeksų adapterius, atitinkančius indeksų įvairovės reikalavimus. Jei tikslaus tiriamų mėginių skaičiaus sąrašė nėra, būtinai parinkite papildomų šulinėlių indeksų adapterius. Jei reikia pasirinkti papildomų šulinėlių indeksus arba nenaudojate rekomenduojamų indeksų adapterių derinių, prieš pasirinkdami indeksus būtinai perskaitykite skyrių *Bazių priskyrimas ir indekso įvairovė 15 psl.*
- 5 „Index 2“ („i5“) (2 indeksas) išskleidžiamajame sąrašė pasirinkite 2 indekso adapterį.
- 6 „Manifest“ (deklaracija) išskleidžiamajame sąrašė pasirinkite deklaracijos failą.
- 7 Pasirinkite atitinkamą parinktį, jei norite peržiūrėti, išspausdinti ar išsaugoti plokštelės išdėstymą, kad juo būtų galima remtis ruošiant bibliotekas.
 - ▶ Norėdami pamatyti plokštelės išdėstymą, pasirinkite piktogramą  **Print** (spausdinti). Kad plokštelės išdėstymas būtų išspausdintas, pasirinkite **Print** (spausdinti).
 - ▶ Pasirinkite **Export** (eksportuoti), kad mėginio informacija būtų eksportuota į išorinį failą. Patikrinkite, ar deklaracijos ir mėginio informacija teisinga. Netiksli informacija gali turėti įtakos rezultatams.
- 8 Pasirinkite **Save Run** (išsaugoti seriją).

Mėginių importavimas

- 1 Pasirinkite **Import Samples** (importuoti mėginius) ir suraskite mėginių informacijos failo vietą. Galite importuoti dviejų rūšių failus.
 - ▶ Ekrane „Create Run“ (kurti seriją) pasirinkite **Template** (šablonas), kad sukurtumėte naują plokštelės šabloną. Šablono faile pateikiamos teisingos importuotinos stulpelių antraštės. Kiekviename stulpelyje įveskite serijos mėginių informaciją. Ištrinkite informaciją nenaudojamuose langeliuose ir išsaugokite failą.
 - ▶ Naudokite mėginio informacijos failą, kuris buvo eksportuotas iš „Germline Variant“ modulio, funkciją „Export“ (eksportuoti).
- 2 Norėdami pamatyti plokštelės išdėstymą, pasirinkite piktogramą  **Print** (spausdinti).
- 3 Pasirinkite **Print** (spausdinti), kad plokštelės šablonas būtų išspausdintas tam, kad juo būtų galima remtis ruošiant bibliotekas.
- 4 **(Pasirinktinai)** Pasirinkite **Export** (eksportuoti), kad mėginio informacija būtų eksportuota į išorinį failą. Patikrinkite, ar deklaracijos ir mėginio informacija teisinga. Netiksli informacija gali turėti įtakos rezultatams.
- 5 Pasirinkite **Save Run** (išsaugoti seriją).

Serijos redagavimas

Jei reikia instrukcijų, kaip redaguoti serijos informaciją prieš nustatant sekas, žr. „NextSeq 550Dx“ *prietaiso informacinį vadovą (dokumento Nr. 1000000009513).*

Analizės metodai

„Germline Variant“ analizės modulis atlieka toliau nurodytus analizės veiksmus, tada analizės išvesties failus įrašo į aplanką „Alignment“ (lygiavimas).

- ▶ Išskirsto indekso nuskaitymus.
- ▶ Generuoja FASTQ failus.
- ▶ Sulygiuoja su atskaita.
- ▶ Identifikuoja variantus.

Išskirstymas

Išskirstymo funkcija kiekvieną indeksų nuskaitymo seką palygina su serijai priskirtomis indeksų sekomis. Atliekant šį veiksmažį neatsižvelgiama į jokias kokybines vertes.

Indeksų nuskaitymai identifikuojami atliekant toliau nurodytus veiksmus.

- ▶ Mėginiai sunumeruojami pradedant nuo 1 tokia tvarka, kokia jie išdėstyti serijai skirtame sąrašė.
- ▶ Mėginys, kurio numeris yra 0, rezervuojamas sancaupoms, kurios nebuvo priskirtos mėginiui.
- ▶ Sankaupos mėginiui priskiriamos, kai indeksų seka tiksliai atitinka arba yra tik viena indeksų nuskaitymo neatitiktis.

FASTQ failo generavimas

Po išskirstymo programinė įranga generuoja tarpinės analizės failus FASTQ formatu (tekstiniu formatu, naudojamu sekoms apibūdinti). FASTQ failai apima kiekvieno mėginio nuskaitymus ir susijusius kokybės įverčius. Sankaupos, neperėjusios filtro, neįtraukiamos.

Kiekviename FASTQ faile yra tik vieno mėginio nuskaitymai. To mėginio pavadinimas įtrauktas į FASTQ failo pavadinimą. FASTQ failai yra pirminė įvestis lygiavimo funkcijai. Sugeneruojami aštuoni vieno mėginio FASTQ failai: keturi iš „Read 1“ (1 nuskaitymas) ir keturi iš „Read 2“ (2 nuskaitymas).

Lygiavimas

Lygiavimo metu juostinis Smitho-Watermano algoritmas sulygiuoja kiekvieno mėginio sancaupas pagal amplikono sekas, nurodytas deklaracijos faile.

Juostinis Smitho-Watermano algoritmas atlieka pusiau visuotinį sekos lygiavimą, kad aptiktų panašius regionus tarp dviejų sekų. Užuoat lyginęs visą seką, Smitho-Watermano algoritmas lygina visų galimų ilgių segmentus.

Kiekvienas pagal galą suporuotas nuskaitymas įvertinamas pagal jo lygiavimą su atitinkamomis to nuskaitymo zondo sekomis.

- ▶ „Read 1“ (1 nuskaitymas) įvertinamas pagal pasrautinių konkrečios srities oligonukleotidų (angl. „Downstream Locus-Specific Oligos“, DLSO) atvirkštinį komplementą.
- ▶ „Read 2“ (2 nuskaitymas) įvertinamas pagal priešrautinius konkrečios srities oligonukleotidus (angl. „Upstream Locus-Specific Oligos“, ULSO).
- ▶ Jei nuskaitymo pradžia sutampa su zondo seka (ne daugiau kaip trys skirtumai, t. y. nesutapimai arba poslinkiai dėl svarbiausių intarpų / iškritų (angl. „indel“), visas nuskaitymo ilgis lygiuojamas pagal tos sekos tikslinį amplikoną.
- ▶ Dėl tyrimo cheminių savybių DLSO ir ULSO esantys intarpai / iškritos nestebimi.

Lygiavimai išfiltruojami iš lygiavimo rezultatų, remiantis neatitiktis rodikliais arba dominančiame regione, arba visame amplikone, tai priklauso nuo amplikono ilgio. Lygiavimo failuose filtruoti lygiavimai įrašomi kaip nesulygiuoti, jie nenaudojami priskiriant variantus.

Variantų priskyrimas

„Illumina“ sukurta priskyrimo priemonė „Pisces Variant Caller“ identifikuoja DNR mėginyje esančius variantus.



„Pisces Variant Caller“ identifikuoja SNV, MNV ir mažus intarpus / iškritas, atlikdamas 3 veiksmus.

- ▶ Atskirai atsižvelgia į kiekvieną padėtį atskaitiniame genome.
- ▶ Suskaičiuoja bazes konkrečioje sulygiuotų nuskaitymų, kurie persidengia su ta padėtimi, padėtyje.
- ▶ Apskaičiuoja variantų įvertį, kuris išmatuoja priskyrimo kokybę, taikant Puasono modelį. Variantai, kurių kokybės įvertis nesiekia Q20, neįtraukiami.

Jei variantas pereina visus filtrus, naudojant VCF jis pažymimas kaip PASS (perėjo).

Jei reikia daugiau informacijos, žr. github.com/Illumina/Pisces/wiki.

Serijos ir mėginio duomenų peržiūra

- 1 „Local Run Manager“ skydelyje spustelėkite serijos pavadinimą.
- 2 Skirtuke „Run Overview“ (serijos apžvalga) peržiūrėkite sekos nustatymo serijos metriką.
- 3 **(Pasirinktinai)** Spustelėkite piktogramą **Copy to Clipboard**  (kopijuoti į iškarpinę), jei norite nukopijuoti serijos išvesties aplanko kelią.
- 4 Norėdami peržiūrėti serijos parametrus ir informaciją apie eksploataavimo medžiagas, spustelėkite skirtuką „Sequencing Information“ (sekos nustatymo informacija).
- 5 Spustelėkite skirtuką „Samples and Results“ (mėginiai bei rezultatai), kad sužinotumėte, kur yra analizės ataskaita.
 - ▶ Jei analizė buvo pakartota, išplėskite išskleidžiamąjį sąrašą „Select Analysis“ (pasirinkti analizę) ir pasirinkite atitinkamą analizę.
- 6 Spustelėkite piktogramą **Copy to Clipboard**  (kopijuoti į iškarpinę), jei norite nukopijuoti kelią iki aplanko „Analysis“ (analizė).

Daugiau informacijos apie skirtukus „Run Overview“ (serijos apžvalga) ir „Sequencing Information“ (informacija apie sekos nustatymą) ir tai, kaip analizę vėl įtraukti į eilę, žr. „NextSeq 550Dx“ *prietaiso informaciniame vadove (dokumento Nr. 100000009513)*.

Analizės ataskaita

Analizės rezultatų suvestinė pateikta skirtuke „Samples and Results“ (mėginiai ir rezultatai). Šią apibendrintą ataskaitą rasite aplanke „Alignment“ (lygiavimas). Kiekvieno mėginio ataskaita taip pat generuojama PDF formatu.

Skirtuko „Samples and Results“ (mėginiai ir rezultatai) informacija

1 Norėdami pamatyti mėginio ataskaitą, spustelėkite mėginį sąrašė.

1 lent. Serijos ir mėginio informacija

Stulpelio antraštė	Aprašas
„Run Status“ (serijos būseną)	Nurodo, ar sekos nustatymo serija sėkminga, ar nepavyko.
„Total Yield (GB)“ (bendrasis kiekis (GB))	Per sekos nustatymo seriją priskirtų bazių skaičius. Nurodo perėjimo slenkstinę vertę ir būseną („pass“ (perėjo) arba „fail“ (neperėjo)).
% ≥ Q30	Per sekos nustatymo seriją įvykusių nuskaitymų, kurių kokybės įvertis yra 30 (Q30) arba didesnis, procentinė vertė. Nurodo perėjimo slenkstinę vertę ir būseną („pass“ (perėjo) arba „fail“ (neperėjo)).
„Sample ID“ (mėginio ID)	Mėginio ID, pateiktas kuriant seriją.
„Total PF Reads“ (bendrasis filtrą perėjusių nuskaitymų skaičius)	Bendrasis nuskaitymų, perėjusių filtrą, skaičius.
„Read 1% ≥ Q30“ (1 nuskaitymų % ≥ Q30)	Per „Read 1“ (1 nuskaitymas) įvykusių nuskaitymų, kurių mėginio kokybės įvertis yra 30 (Q30) arba didesnis, procentinė vertė.
„Read 2% ≥ Q30“ (2 nuskaitymų % ≥ Q30)	Per „Read 2“ (2 nuskaitymas) įvykusių nuskaitymų, kurių mėginio kokybės įvertis yra 30 (Q30) arba didesnis, procentinė vertė.
„Autosome Call Rate“ (autosomų priskyrimo dažnis)	Padėčių genomo autosomose (1–22 chromosomose), kurios atitinka iš anksto nustatytą pasiklojimo slenkstinę vertę, skaičius, padalytas iš bendrojo analizuojamų padėčių genomo autosomose skaičiaus. Priskyrimo rodiklis nustatomas vieno mėginio lygiu, jis pateikiamas kaip procentinė vertė, apskaičiuota kaip minus 1 (padėčių genomo autosomose su nevisiškais priskyrimais skaičius, padalytas iš bendrojo padėčių autosomose, kurių seka nustatyta, skaičiaus).

2 lent. Mėginio ataskaitos informacija

Stulpelio antraštė	Aprašas
„Sample“ (mėginys)	Mėginio ID, pateiktas kuriant seriją.
„Report Date“ (ataskaitos data)	Ataskaitos generavimo data.
„Sample Information“ (mėginio informacija)	Mėginio ID, pateiktas kuriant seriją, bendrasis mėginio nuskaitymų, perėjusių filtrą, skaičius, mėginio nuskaitymų, kurių kokybės įvertis yra 30 (Q30) ar didesnis, procentinė vertė ir autosomų priskyrimo koeficientas.
„Amplicon Summary“ (amplikonų suvestinė)	Bendrasis amplikonų regionų, kurių sekos nustatytos, skaičius, taip pat bendrasis ilgis tikslinių regionų amplikonų, kurių sekos nustatytos, bazių poromis (skirta mėginio ir deklaracijos failui). Deklaracijos failas nurodo atskaitinį genomą ir tikslinius atskaitinius regionus, naudojamus atliekant lygiavimo veiksmą.
„Read Level Statistics“ (nuskaitymo lygio statistiniai duomenys)	Mėginio nuskaitymų skaičius ir procentas, apimant kiekvieną atskaitinę padėtį (skirta „Read 1“ (1 nuskaitymas) ir „Read 2“ (antras nuskaitymas)).
„Variants Summary“ (variantų suvestinė)	Vieno nukleotido variantų (SNV), tarpų ir iškritų, aptiktų mėginyje, kurios atitinka siūlomas vertes, siekiant nustatyti, ar kokybės rezultatai patenka į priimtą intervalą, skaičius.
„Coverage Summary“ (aprėpties suvestinė)	Bendrasis sulygiuotų bazių skaičius, padalytas iš tikslinio regiono dydžio, taip pat amplikonų regionų, kurių aprėpties vertės didesnės už mažos aprėpties slenkstinę vertę (0,2 * vidutinė amplikonų aprėptis), procentinė vertė mėginyje.

Stulpelio antraštė	Aprašas
„Coverage Plots“ (aprėpties diagramos)	Diagramoje „Coverage by Amplicon Region“ (aprėptis pagal amplikonų regionus) parodyta aprėptis mėginio amplikonų regionuose. Regionai, kurių aprėpties vertės mažesnės nei aprėpties slenkstinė vertė, paryškinti raudona spalva. Visų verčių vidurkį nurodo oranžinė linija.
„Software Versions“ (programinės įrangos versijos)	Programinės įrangos versijos, kurios buvo įdiegtos, kai buvo nustatomos sekos mėginyje. Apima „NextSeq 550Dx“ operacinę programinę įrangą (NOS), „Local Run Manager Software“, RTA programinę įrangą ir „Germline Variant“ modulio versijas.

Analizės išvesties failai

Generuojami toliau išvardyti analizės išvesties failai, skirti „Germline Variant“ analizės moduliui. Juose pateikiami lygiavimo ir variantų priskyrimo veiksmams reikalingi analizės rezultatai. Analizės išvesties failai saugomi aplanke „Alignment“ (lygiavimas).

Failo pavadinimas	Aprašas
Išskirstymas (*.txt)	Tarpiniai failai, kuriuose pateikiami išskirstymo suvestinės rezultatai.
FASTQ (*.fastq.gz)	Tarpiniai failai, kuriuose pateikiami kokybės įverčių bazės priskyrimai. FASTQ failai yra pirminė įvestis lygiavimo veiksmui.
Lygiavimo failai BAM formatu (*.bam)	Pateikiami konkretaus mėginio sulygiuoti nuskaitymai.
Variantų priskyrimo failai genomo VCF formatu (*.genome.vcf.gz)	Pateikiamas kiekvienos padėties genotipas, nesvarbu, ar priskirtas kaip variantas, ar kaip atskaitos pagrindas.
Variantų priskyrimo failai VCF formatu (*.vcf.gz)	Pateikiami visi variantai, priskirti tiksliniame regione.
AmpliconCoverage_M1.tsv	Pateikiama informacija apie kiekvienos pateiktos deklaracijos atskiro mėginio vieno amplikono aprėptį. M# nurodo deklaracijos numerį.

Demultiplekso failų formatas

Demultiplekso procesas nuskaityto prie kiekvienos sankaupos pritvirtintą indeksavimo seką, kad nustatytų, iš kurio mėginio susidarė ši sankaupa. Sankaupų ir mėginio numerio saitas įrašomas į pratekamosios kiuvetės kiekvienos išklotinės demultiplekso (*.demux) failą.

Demultiplekso failų pavadinimo formatas yra **s_1_X.demux** (X yra išklotinės numeris).

Demultiplekso failai prasideda antrašte:

- ▶ versija (4 baitų sveikasis skaičius), šiuo metu tai yra 1;
- ▶ sankaupų skaičius (4 baitų sveikasis skaičius).

Likusią failo dalį sudaro kiekvienos sankaupos išklotinėje mėginių numeriai.

Pasibaigus demultiplekso veiksmui, programinė įranga generuoja demultiplekso failą, pavadintą **DemultiplexSummaryF1L1.txt**.

- ▶ Failo pavadinime **F1** atitinka pratekamosios kiuvetės numerį.
- ▶ Failo pavadinime **L1** atitinka takelio numerį.
- ▶ Išskirstant gaunama lentelė (po 1 eilutė išklotinei ir po 1 stulpelį mėginiui, įskaitant 0 mėginį).
- ▶ Dažniausiai pasitaikančios sekos indekso nuskaitymuose.

FASTQ failų formatas

FASTQ – tai tekstinių failų formatas. Šiuose failuose pateikiami nuskaitymo bazių priskyrimai ir kokybės vertės. Kiekviename įrašė yra 4 eilutės.

- ▶ Identifikatorius
- ▶ Seka
- ▶ Pluso ženklas (+)
- ▶ „Phred“ kokybės įverčiai ASCII + 33 užkoduotu formatu

Identifikatorių formatas nurodytas toliau.

@Prietaisas:SerijosID:PratekamosiosKiuvetėsID:Takelis:Išklotinė:X:Y
NuskaitymoNumeris:FiltroVėliavėlė:0:MėginioNumeris

Pavyzdys:

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

BAM failų formatas

BAM failas (*.bam) – tai suglaudinta dvejetainė SAM failo versija, naudojama sulygiuotoms sekoms iki 128 Mb atvaizduoti. SAM ir BAM formatai išsamiai apibūdinti čia: samtools.github.io/hts-specs/SAMv1.pdf.

BAM failų pavadinimo formatas yra **SampleName_S#.bam** (# yra mėginio numeris pagal eilės tvarką, kuria išvardyti serijos mėginiai).

BAM failuose yra antraštės skirsnis ir lygiavimo skirsnis.

- ▶ **Header** (antraštė) – pateikiama informacija apie visą failą, pvz., mėginio pavadinimas, mėginio ilgis ir lygiuotės būdas. Lygiuotės atitinkame skirsnyje susijusios su konkrečia informacija, pateikta antraštės skirsnyje.
- ▶ **Alignments** (lygiuotės) – pateikiamas nuskaitymo pavadinimas, nuskaitymo seka ir kokybė, lygiuotės informacija bei tinkintos žymės. Nuskaitymo pavadinimą sudaro chromosoma, pradžios koordinatė, lygiuotės kokybė ir atitikties deskriptoriaus eilutė.

Lygiuotės skirsnyje pateikiama toliau nurodyta informacija apie kiekvieną nuskaitymą ar jų porą.

- ▶ **AS:** suporuotos pagal galą lygiuotės kokybė
- ▶ **RG:** nuskaitymo grupė, nurodanti konkretaus mėginio nuskaitymų skaičių
- ▶ **BC:** brūkšninio kodo žymė, nurodanti išskirstyto mėginio ID, susietą su nuskaitymu
- ▶ **SM:** vieno galo lygiuotės kokybė
- ▶ **XC:** atitikties deskriptoriaus eilutė
- ▶ **XN:** amplikono pavadinimo žymė, įrašanti su nuskaitymu susietą amplikono ID

BAM indekso failuose (*.bam.bai) pateikiamas atitinkamo BAM failo indeksas.

VCF failų formatas

Variantų priskyrimo formatas (angl. „Variant Call Format“, VCF) yra įprastas failų formatas, sukurtas genų tyrimus atliekančių mokslininkų. Šio formato failuose pateikiama informacija apie variantus, aptiktus konkrečios atskaitos genomo padėtyse. VCF failai baigiasi .vcf.

VCF failų antraštė apima VCF failo formato versiją, variantų priskyrimo priemonės versiją ir likusioje failo dalyje naudojamų pastabų sąrašą. VCF antraštė taip pat apima atskaitos genomo failą ir BAM failą. Paskutinėje antraštės eilutėje yra duomenų eilučių stulpelių antraštės. Kiekvienoje VCF failo duomenų eilutėje pateikiama informacija apie vieną variantą.

VCF failų antraštės

Antraštė	Aprašas
CHROM (chromosoma)	Atskaitos genomo chromosoma. Chromosomos rodomos ta pačia eilės tvarka kaip atskaitos FASTQ faile.
POS (padėtis)	Varianto vienos bazės padėtis atskaitos chromosomoje. SNP atveju ši padėtis yra atskaitos bazė su variantu; tarpų / iškritų arba iškritų atveju – atskaitos bazė prieš pat variantą.
ID	Tai yra varianto RS numeris, gautas iš dbSNP.txt, jei taikytina. Jei šioje vietoje yra keli RS numeriai, sąrašo elementai atskirti kabliataškiais. Jei šioje vietoje dbSNP įvesties nėra, naudojamas trūkstamos vertės žymiklis ('.').
REF	Atskaitinis genotipas. Pavyzdžiui, vienos T iškritą atitinka atskaitos TT ir alternatyvusis T. A–T vieno nukleotido variantą atitinka atskaitinis A ir alternatyvusis T.
ALT	Aleliai, kurie skiriasi nuo atskaitinio nuskaitymo. Pavyzdžiui, vieno T tarpą atitinka atskaitinis A ir alternatyvusis AT. A–T vieno nukleotido variantą atitinka atskaitinis A ir alternatyvusis T.
QUAL	Phred kokybės įvertis, kurį priskyrė variantų priskyrimo priemonė. Didesni įverčiai atitinka didesnį variantų patikimumą ir mažesnę klaidų tikimybę. Kokybės įvertis Q reiškia, kad numatoma klaidos tikimybė yra $10^{-(Q/10)}$. Pavyzdžiui, Q30 priskyrimų rinkinio klaidų koeficientas 0,1 %. Daugelis variantų priskyrimo priemonių kokybės įverčius priskiria pagal savo statistinius modelius, kurie glaudžiai susiję su stebimu klaidų koeficientu.

VCF failų komentarai

Antraštė	Aprašas
FILTER (filtras)	<p>Jei visi filtrai pereinami, filtro stulpelyje įrašoma PASS (perėjo).</p> <ul style="list-style-type: none"> • LowDP (maža DP) – taikoma zonoms, kuriose aprėpties gylis nesiekia 150x. Amplikonų padėtyse, kurias aprėpia tiek nuskaitymas į priekį, tiek atgalinis nuskaitymas, tai atitinka 300 persidengiančių pagal galą suporuotų nuskaitymų. • q20 – kokybės įvertis < 20. • MultiAllelicSite (kelių alelių zona) – variantas neatitinka diploidinio modulio. • R5x9 – gretimų nuskaitymų (1–5 bazių porų ilgio) skaičius iki variantų priskyrimo ≥ 9. • SB – gijos nukrypimas didesnis už naudojamą slenkstinę vertę.
INFO (informacija)	<p>Toliau nurodytos galimos įvestys stulpelyje INFO (informacija).</p> <ul style="list-style-type: none"> • AC – alelių skaičius genotipuose (kiekvieno ALT alelio) taikant tą pačią eilės tvarką kaip sąrašė. • AF – alelių dažnumas (kiekvieno ALT alelio), taikant tą pačią eilės tvarką kaip sąrašė. • AN – bendrasis alelių skaičius priskirtuose genotipuose. • CD – vėliavėlė, nurodanti, kad SNP pasireiškia bent 1 „RefGen“ įvesties kodavimo regione. • DP – gylis (bazių priskyrimų, sulygiuotų su padėtimi ir naudojamų priskiriant variantus, skaičius). • Exon (egzonas) – kableliais atskirtas egzonų regionų, nuskaitytų iš „RefGen“, sąrašas. • FC – funkcinis nuoseklumas. • GI – kableliais atskirtas genų ID, nuskaitytų iš „RefGen“, sąrašas. • QD – variantų pasiklovimas / kokybė pagal gylį. • TI – kableliais atskirtas transkriptonų ID, nuskaitytų iš „RefGen“, sąrašas
FORMAT (formatas)	<p>Formato stulpelyje pateikiamas dvitaškiais atskirtų laukų sąrašas. Pavyzdžiui, GT:GQ. Pateiktas laukų sąrašas skiriasi priklausomai nuo to, kokia variantų priskyrimo priemonė naudojama. Toliau išvardyti galimi laukai.</p> <ul style="list-style-type: none"> • AD – X,Y formos įvestis (X yra atskaitos priskyrimų skaičius, Y – alternatyvių priskyrimų skaičius). • DP – apytikslis nuskaitymo gylis; filtruojami nuskaitymai, kurių MQ = 255 arba kurie turi blogų atitikmenų. • GQ – genotipo kokybė. • GT – genotipas. 0 atitinka atskaitos bazę, 1 – pirmąją įvestį ALT stulpelyje ir t. t. Dešininis pasvirasis brūkšnylis (/) nurodo, kad fazių informacija nepasiekama. • NC – bazių, kuriuos nebuvo priskirtos arba kurių priskyrimo kokybė neatitinka minimalios slenkstinės vertės, dalis. • NL – triukšmo lygis (bazių priskyrimo triukšmo šioje padėtyje įvertinimas). • SB – gijos nukrypimas šioje padėtyje. Didesnės neigiamos vertės nurodo mažesnį nukrypimą, o vertės, artimesnės 0, nurodo didesnį nukrypimą. • VF – varianto dažnumas (nuskaitymų, nurodančių alternatyvų alelį, procentinė vertė).
SAMPLE (mėginys)	Mėginio stulpelyje pateiktos vertės, nurodytos stulpelyje FORMAT (formatas).

Genomų VCF failai

Genomų VCF (gVCF) failai – tai VCF 4.1 versijos failai, kuriuose visoms vietoms genome reprezentuoti kompaktišku formatu naudojamas taisyklių rinkinys. gVCF (*.genome.vcf.gz) failai apima visas vietas dominančioje srityje (atskiras failas kiekvienam mėginiui).

gVCF faile nurodomi nepriskyrimo atvejai vietose, kurios nepereina visų filtrų. Genotipo (GT) žymė ./, nurodo nepriskyrimą.

Jei reikia daugiau informacijos, žr. sites.google.com/site/gvcftools/home/about-gvcf.

Amplikono aprėpties failas

Generuojamas su kiekvienu deklaracijos failu susietas amplikono aprėpties failas. Failo pavadinime M# nurodo deklaracijos numerį.

Kiekviename faile yra antraštės eilutė, joje pateikiami mėginio ID, susieti su deklaracija. Faile pateikiama toliau nurodyta informacija.

- ▶ „Target ID“ (taikinio ID) pagal sąrašą deklaracijoje.
- ▶ Filtrą perėjusių nuskaitymų aprėpties gylys.

Papildomi išvesties failai

Toliau nurodytuose išvesties failuose pateikiama papildomos informacijos arba serijos rezultatų ir analizės klaidų santrauka. Nors šie failai nėra būtini norint gauti analizės rezultatus, jie gali būti naudojami šalinant triktis. Visi failai yra aplanke „Alignment“ (lygiavimas), nebent nurodyta kitaip.

Failo pavadinimas	Aprašas
AnalysisLog.txt	Apdorojimo žurnalas, apibūdinantis kiekvieną veiksmą, atliktą analizuojant dabartinės serijos aplanką. Šiame faile klaidų pranešimų nėra. Saugomas aplanke „Alignment“ (lygiavimas).
AnalysisError.txt	Apdorojimo žurnalas, kuriame pateikiamas visų analizės klaidų sąrašas. Jei klaidų nepasitaikė, šis failas gali būti tuščias. Saugomas aplanke „Alignment“ (lygiavimas).
DemultiplexSummaryF1L1#.txt	Pateikia išskirstymo rezultatų lentelę (po 1 eilutę išsklotinei ir po 1 stulpelį mėginiui). Simbolis „#“ nurodo pratekamosios kiuvetės 1, 2, 3 arba 4 takelį. Saugomas aplanke „Alignment“ (lygiavimas).
AmpliconRunStatistics.xml	Pateikiami statistiniai konkrečios serijos suvestinės duomenys. Saugomas aplanke „Alignment“ (lygiavimas).

Aplankas „Analysis“

Aplanke „Analysis“ (analizė) saugomi „Local Run Manager“ programinės įrangos sugeneruoti failai.

Išvesties aplanko ir analizės aplanko ryšys apibendrintas toliau.

- ▶ Nustatant seką, „Real-Time Analysis“ (RTA) į išvesties aplanką surenka failus, sugeneruotus atliekant vaizdų analizę, bazių priskyrimą ir kokybės įvertinimą.
- ▶ RTA failus į analizės aplanką nukopijuoja realiuoju laiku. Po to, kai RTA priskiria kokybės įvertį kiekvieno ciklo kiekvienai bazei, programinė įranga failą RTAComplete.txt įrašo į abu aplankus.
- ▶ Kai atsiranda failas RTAComplete.txt, prasideda analizė.
- ▶ Vykstant analizei, „Local Run Manager“ išvesties failus įrašo į analizės aplanką, o tada failus vėl nukopijuoja į išvesties aplanką.

Lygiavimo aplankai

Kiekvieną kartą, kai analizė vėl įtraukiama į eilę, „Local Run Manager“ sukuria lygiavimo aplanką pavadinimu **Alignment_N** (N yra priskirtas eilės numeris).

Aplanko struktūra

Alignment (lygiavimas) – čia saugomi *.bam, *.vcf, FASTQ failai ir konkretaus analizės modulio failai.

Date and Time Stamp (datos ir laiko žyma) – tai žyma „Data_laikas“ formatu MMMMmmDD_VVMMSS.

- 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“ (statistiniai duomenys)**
 - DemuxSummaryF1L1.txt
 - FastqSummaryF1L1.txt

Data (duomenys)

Intensities (intensyvumas)

BaseCalls (bazių priskyrimai)

L001 – čia saugomi *.bcl failai.

L001 – čia saugomi *.locs failai.

RTA Logs (RTA žurnalai) – čia saugomi RTA programinės įrangos analizės žurnalo failai.

InterOp (vidinės operacijos) – saugomi dvejetainiai failai, naudojami sekoskaitos metrikos ataskaitoms.

Logs (žurnalai) – žurnalo failai, apibūdinantis sekoskaitos veiksmus.

- RTAComplete.txt
- RunInfo.xml
- RunParameters.xml

Bazių priskyrimas ir indekso įvairovė

Kai mėginiuose esančios sekos nustatomos naudojant prietaisą „NextSeq 550Dx“, bazių priskyrimas nustato bazę (A, C, G arba T) kiekvienai konkrečios išsklotinės sankaupai arba vaizdavimo srityje pratekamojoje kiuvetėje, vykstant konkrečiam ciklui. Prietaisas „NextSeq 550Dx“ atlieka dviejų kanalų sekos nustatymą, todėl keturių DNR bazių duomenims koduoti tereikia dviejų vaizdų – vieno iš raudono kanalo ir kito iš žalio kanalo.

Bazių priskyrimo indekso nuskaitymo procesas skiriasi nuo bazių priskyrimo kitų nuskaitymų metu.

Pirmieji du indekso nuskaitymo ciklai turi prasidėti nuo bent vienos ne G bazės. Jei indekso nuskaitymas prasideda dviem G bazės priskyrimais, nebus sugeneruotas signalo intensyvumas. Siekiant užtikrinti išskirstymo našumą, signalas turi būti viename iš pirmųjų dviejų ciklų.

Jei indeksai pasirenkami kuriant seriją ir neatitinka įvairovės reikalavimų, pasirodys mažos įvairovės įspėjimas. Kad nebūtų sugeneruotas mažos įvairovės įspėjimas, pasirinkite indeksų sekas, kurios pateikia signalą abiejuose kiekvieno ciklo kanaluose.

- ▶ Raudonas kanalas – A arba C
- ▶ Žalias kanalas – A arba T

Šis bazių priskyrimo procesas užtikrina tikslumą analizuojant žemos fragmentacijos mėginius. Daugiau informacijos apie indeksų sekas žr. „*TruSeq Custom Amplicon Kit Dx*“ (dokumento Nr. 1000000029772) pakuotės lapelyje.

Kai kuriate seriją naudodami „Local Run Manager“, turite pasirinkti tirtinų mėginių skaičių. Programinė įranga automatiškai įveda indeksų derinius, kurie atitinka indeksų įvairovės kriterijus. Nors neprivalote naudoti siūlomų indeksų derinių, juos naudoti rekomenduojama.

Keitimų istorija

Dokumentas	Data	Keitimo aprašymas
Dokumento Nr. 1000000030329, 04 versija	2021 m. rugpjūčio mėn.	Atnaujintas įgaliotojo atstovo ES adresas.
Dokumento Nr. 1000000030329, 03 versija	2019 m. gruodžio mėn.	Atnaujintas įgaliotojo atstovo ES adresas. Atnaujintas rėmėjo Australijoje adresas.
Dokumento Nr. 1000000030329, 02 versija	2019 m. sausio mėn.	Pridėta informacija apie 2.5 versijos reagentų rinkinius.
Dokumento Nr. 1000000030329, 01 versija	2018 m. rugpjūčio mėn.	Atnaujintos reglamentavimo žymės.
Dokumento Nr. 1000000030329, 00 versija	2017 m. lapkričio mėn.	Pirmasis leidimas.

Techninė pagalba

Dėl techninės pagalbos kreipkitės į „Illumina“ techninės pagalbos tarnybą.

Interneto svetainė www.illumina.com
El. p. techsupport@illumina.com

„Illumina“ klientų aptarnavimo telefono numeriai

Regionas	Nemokamas telefono numeris	Regioninis
Šiaurės Amerika	+1.800.809.4566	
Airija	+353 1800936608	+353 016950506
Australija	+1.800.775.688	
Austrija	+43 800006249	+43 19286540
Belgija	+32 80077160	+32 34002973
Danija	+45 80820183	+45 89871156
Hongkongas	800960230	
Ispanija	+34 911899417	+34 800300143
Italija	+39 800985513	+39 236003759
Japonija	0800.111.5011	
Jungtinė Karalystė	+44 8000126019	+44 2073057197
Kinija	400.066.5835	
Naujoji Zelandija	0800.451.650	
Nyderlandai	+31 8000222493	+31 207132960
Norvegija	+47 800 16836	+47 21939693
Prancūzija	+33 805102193	+33 170770446
Singapūras	+1.800.579.2745	
Suomija	+358 800918363	+358 974790110
Švedija	+46 850619671	+46 200883979
Šveicarija	+41 565800000	+41 800200442
Taivanas	00806651752	
Vokietija	+49 8001014940	+49 8938035677
Kitos šalys	+44 1799 534000	

Saugos duomenų lapai (SDL) prieinami „Illumina“ interneto svetainėje support.illumina.com/sds.html.

Produkto dokumentaciją PDF formatu galima atsisiųsti iš „Illumina“ interneto svetainės. Apsilankykite support.illumina.com, pasirinkite gaminį, tada pasirinkite **Documentation & Literature** (dokumentacija ir literatūra).



„Illumina“
5200 Illumina Way
San Diego, California 92122 JAV
+1.800.809.ILMN (4566)
+1.858.202.4566 (ne Šiaurės Amerikoje)
techsupport@illumina.com
www.illumina.com

CE



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

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

NAUDOTI IN VITRO DIAGNOSTIKAI

© 2021 „Illumina, Inc.“. Visos teisės saugomos.

illumina®