

Analytický modul Germline Variant softvéru Local Run Manager (Správca lokálnych chodov)

Príručka postupov pre NextSeq 550Dx

NA DIAGNOSTICKÉ ÚČELY IN VITRO

Základné informácie	3
Zadanie informácií o chode	4
Metódy analýzy	6
Zobrazenie údajov o chode a vzorke	7
Správa o analýze	7
Súbory s výstupmi analýzy	9
História revízií	15
Technická pomoc	16



Tento dokument a jeho obsah sú vlastníctvom spoločnosti Illumina, Inc. a jej pridružených spoločností (ďalej len „Illumina“) a sú určené výlučne na zmluvné použitie u zákazníka v súvislosti s používaním produktu (produktov) opísaného (opísaných) v tomto dokumente a na žiadny iný účel. Tento dokument a jeho obsah sa nesmú používať ani šíriť na žiadny iný účel a/alebo inak poskytovať, zverejňovať alebo reprodukovať akýmkoľvek spôsobom bez predchádzajúceho písomného súhlasu spoločnosti Illumina. Spoločnosť Illumina týmto dokumentom neposkytuje žiadnu licenciu na základe patentu, ochrannej známky, autorských práv alebo práv podľa zvykového práva, či podobných práv tretích strán.

Pokyny v tomto dokumente musia byť prísne a výslovne dodržiavané kvalifikovaným a riadne vyškoleným personálom, aby sa zabezpečilo správne a bezpečné používanie tu popísaného produktu (produktov). Pred použitím takéhoto produktu (produktov) je nutné prečítať si celý obsah tohto dokumentu s porozumením.

NEPREČÍTANIE VŠETKÝCH POKYNOV TU OBSIAHNUTÝCH A ICH VÝSLOVNÉ NEDODRŽANIE MÔŽE MAŤ ZA NÁSLEDOK POŠKODENIE PRODUKTU (PRODUKTOV), ZRANENIE OSOBY VRÁTANE POUŽÍVATEĽOV ALEBO INÝCH OSÔB, POŠKODENIE ĎALŠIEHO MAJETKU A ZRUŠENIE PLATNOSTI ZÁRUKY VZŤAHUJÚCEJ SA NA PRODUKT (PRODUKTY).

SPOLOČNOSŤ ILLUMINA NEPREBERÁ ŽIADNU ZODPOVEDNOSŤ VYPLÝVAJÚCU Z NEBEZPEČNÉHO POUŽITIA TU POPÍSANÉHO PRODUKTU (PRODUKTOV) (VRÁTANE SÚČASTÍ ALEBO SOFTVÉRU).

© 2021 Illumina, Inc. Všetky práva vyhradené.

Všetky ochranné známky sú vlastníctvom spoločnosti Illumina, Inc. alebo príslušných vlastníkov. Informácie o konkrétnych ochranných známkach nájdete na stránke www.illumina.com/company/legal.html.

Základné informácie

Modul Germline Variant softvéru Local Run Manager (Správca lokálnych chodov) je určený na použitie so súpravou na rozbor Illumina TruSeq Custom Amplicon Kit Dx a prístrojom NextSeq 550Dx. Keď sa používa s modulom Germline Variant, je určená na prípravu knižníc používaných na sekvenovanie DNA zo vzoriek periférnej plnej krvi.

Analytický modul vyhodnocuje krátke oblasti amplifikovanej DNA, resp. amplicóny, z hľadiska variantov. Cílené sekvenovanie amplicónov umožňuje vysoké pokrytie konkrétnych oblastí. Pozrite si príbalový leták súpravy *TruSeq Custom Amplicon Kit Dx* (dokument č. 1000000029772).

Analytický modul Germline Variant si vyžaduje sekvenovací spotrebný materiál na 300 cyklov. Viac informácií nájdete v príbalovom letáku súpravy reagensii *NextSeq 550Dx High Output Reagent Kit v2* alebo *NextSeq 550Dx High Output Reagent Kit v2.5*.

Informácie o tejto príručke

V tejto príručke sú uvedené pokyny na nastavenie parametrov chodu na sekvenovanie a analýzu pre analytický modul Germline Variant. Informácie o ovládacom paneli softvéru Local Run Manager (Správca lokálnych chodov) a nastaveniach systému nájdete v referenčnej príručke prístroja *NextSeq 550Dx* (dokument č. 1000000009513).

Zobrazenie softvéru Local Run Manager (Správca lokálnych chodov)

Rozhranie softvéru Local Run Manager (Správca lokálnych chodov) sa zobrazuje v rámci operačného softvéru NextSeq 550Dx (NOS) alebo prostredníctvom webového prehliadača. Podporovaným webovým prehliadačom je Chromium.



POZNÁMKA

Ak používate nepodporovaný webový prehliadač, po zobrazení hlásenia „Confirm Unsupported Browser“ (Potvrdiť nepodporovaný prehliadač) si stiahnite podporovaný prehliadač. Podporovanú verziu prehliadača Chromium si stiahnete kliknutím na „[here](#)“ (tu).

Zobrazenie na monitore prístroja

- 1 Ak chcete zobraziť rozhranie softvéru Local Run Manager (Správca lokálnych chodov) na monitore prístroja, vyberte jednu z týchto možností:
 - ▶ Na domovskej obrazovke NOS vyberte možnosť **Local Run Manager** (Správca lokálnych chodov). Po dokončení sa do NOS vrátite kliknutím na X v pravom hornom rohu.
 - ▶ Vyberte ikonu Minimize NOS (Minimalizovať NOS), otvorte webový prehliadač Chromium, ktorý je nainštalovaný v prístroji, a do riadka adresy zadajte **http://localhost**. NOS môžu minimalizovať len používatelia s oprávneniami správcu.

Zobrazenie z počítača pripojeného k sieti

- 1 Otvorte webový prehliadač Chromium v počítači s prístupom k tej istej sieti, ku ktorej je pripojený prístroj, a pripojte sa pomocou IP adresy prístroja alebo názvu prístroja. Napríklad **http://myinstrument**.

Zadanie informácií o chode

Nastavenie parametrov

- 1 Prihláste sa do softvéru Local Run Manager (Správca lokálnych chodov).
- 2 Vyberte možnosť **Create Run** (Vytvoriť chod) a potom **Germline Variant** (Germinálny variant).
- 3 Zadajte názov chodu, ktorým sa identifikuje chod od sekvenovania cez analýzu. Použite alfanumerické znaky, medzery, znaky podčiarknutia alebo pomlčky.
- 4 **[Voliteľné]** Zadajte opis chodu, čo pomôže pri identifikácii chodu. Použite alfanumerické znaky, medzery, znaky podčiarknutia alebo pomlčky.
- 5 Z rozbaľovacieho zoznamu vyberte počet vzoriek a súbor indexov. Pri výbere zvažte nasledujúce informácie.
 - ▶ V rozbaľovacom zozname sa nachádzajú počty vzoriek so súborom indexov. Napríklad 24-Set 1 znamená, že sa má otestovať 24 vzoriek s indexmi zo súboru indexov 1.
 - ▶ Čísla súboru indexov označujú rôzne súbory indexov i5. Súbor 1 aj súbor 2 predstavujú diverzitu indexov. Dva súbory indexov sa ponúkajú preto, aby sa zabránilo vyčerpaniu jedného súboru.
 - ▶ Vyberte taký počet vzoriek, aký je najbližšie k počtu vzoriek, ktoré chcete testovať. Ak sa v zozname nenachádza presný počet vzoriek, vyberte ten najbližší, no musí byť nižší ako počet vzoriek, ktoré chcete testovať. Napríklad ak chcete otestovať 18 vzoriek, vyberte 16 vzoriek.
 - ▶ Kombinácie jamiek na vzorky a indexov, ktoré spĺňajú požiadavky diverzity, sú zvýraznené nazeleno. Ak vyberiete iné kombinácie jamiek a indexov a požiadavky na diverzitu nie sú splnené, budete na to upozornení pri ukladaní chodu.

Importovanie súborov manifestov chodu

- 1 Uistite sa, že manifesty, ktoré chcete importovať, sú dostupné v sieťovom umiestnení alebo na jednotke USB.
- 2 Vyberte možnosť **Import Manifests** (Importovať manifesty).
- 3 Prejdite do súboru manifestov a vyberte manifesty, ktoré chcete pridať.



POZNÁMKA

Ak chcete sprístupniť súbory manifestov pre všetky chody využívajúce analytický modul Germline Variant, pridajte manifesty pomocou funkcie Module Settings (Nastavenia modulu). Táto funkcia vyžaduje povolenia používateľa na úrovni správcu. Viac informácií nájdete v *referenčnej príručke prístroja NextSeq 550Dx* (dokument č. 1000000009513).


Špecifikovanie vzoriek chodu

Špecifikujte vzorky chodu pomocou jednej z možností a pokynov, ktoré nasledujú.


- ▶ **Manuálne zadanie vzoriek** – použite prázdnu tabuľku na obrazovke Create Run (Vytvoriť chod).
- ▶ **Importovanie vzoriek** – prejdite do externého súboru vo formáte s hodnotami oddelenými čiarkami (*.csv). Vzor si môžete stiahnuť na obrazovke Create Run (Vytvoriť chod).

Po vyplnení tabuľky vzoriek si môžete informácie o vzorkách exportovať do externého súboru. Súbor používajte ako referenciu pri príprave knižníc alebo ho importujte na ďalší chod.

Manuálne zadávanie vzoriek

- 1 Do poľa Sample ID (ID vzorky) zadajte jedinečné ID vzorky.
Použite alfanumerické znaky, pomlčky alebo značky podčiarknutia.
- 2 **[Voliteľné]** Pri pozitívnych alebo negatívnych kontrolných vzorkách kliknite pravým tlačidlom a vyberte typ kontrolnej vzorky.
- 3 **[Voliteľné]** Do poľa Sample Description (Popis vzorky) zadajte popis vzorky.
Použite alfanumerické znaky, pomlčky alebo značky podčiarknutia.
Popisy vzorky sú pridružené k názvu vzorky. Ak sa ten istý názov vzorky znova použije v neskoršom chode, popisy vzorky sa prepíšu.
- 4 Z rozbaľovacieho zoznamu Index 1 (i7) vyberte adaptér Index 1.
Ak použijete navrhované jamky na vzorky, softvér automaticky vyplní adaptéry indexov i7 a i5, ktoré spĺňajú požiadavky na diverzitu indexov. Ak nie je v zozname uvedený presný počet vzoriek, ktoré chcete otestovať, nezabudnite vybrať adaptéry indexov pre dodatočné jamky. Ak potrebujete vybrať indexy pre dodatočné jamky alebo nepoužijete odporúčané kombinácie adaptérov indexov, pred výberom indexov si prečítajte časť *Stanovenie báz a diverzita indexov na strane 14*.
- 5 Z rozbaľovacieho zoznamu Index 2 (i5) vyberte adaptér Index 2.
- 6 Z rozbaľovacieho zoznamu Manifest vyberte súbor manifestu.
- 7 Zvoľte možnosť zobrazenia, tlače alebo uloženia rozloženia platničky ako referencie pri príprave knižníc:
 - ▶ Rozloženie platničky zobrazíte výberom ikony  **tlače**. Ak chcete vytlačiť rozloženie platničky, vyberte možnosť **Print** (Tlačiť).
 - ▶ Ak chcete exportovať informácie o vzorke do externého súboru, vyberte možnosť **Export** (Exportovať).Uistite sa, že manifest a informácie o vzorke sú správne. Nesprávne informácie môžu ovplyvniť výsledky.
- 8 Vyberte možnosť **Save Run** (Uložiť chod).

Importovanie vzoriek

- 1 Vyberte možnosť **Import Samples** (Importovať vzorky) a vyhľadajte umiestnenie súboru s informáciami o vzorke. Importovať môžete dva typy súborov.
 - ▶ Ak chcete vytvoriť nové rozloženie platničky, na obrazovke Create Run (Vytvoriť chod) vyberte možnosť **Template** (Vzor). Vzorový súbor obsahuje správne nadpisy stĺpcov na importovanie. Do každého stĺpca zadajte informácie o vzorkách v chode. Vymažte ukázkové informácie v políčkach, ktoré nepoužijete, a potom súbor uložte.
 - ▶ Použite súbor s informáciami o vzorke, ktorý sa exportoval z modulu Germline Variant pomocou funkcie Export (Exportovať).
- 2 Rozloženie platničky zobrazíte výberom ikony  **tlače**.
- 3 Ak chcete rozloženie platničky vytlačiť ako referenciu pri príprave knižníc, vyberte možnosť **Print** (Tlačiť).
- 4 **[Voliteľné]** Ak chcete exportovať informácie o vzorke do externého súboru, vyberte možnosť **Export** (Exportovať).
Uistite sa, že manifest a informácie o vzorke sú správne. Nesprávne informácie môžu ovplyvniť výsledky.
- 5 Vyberte možnosť **Save Run** (Uložiť chod).

Úprava chodu

Pokyny na úpravu informácií vo vašom chode pred sekvenovaním nájdete v *referenčnej príručke prístroja NextSeq 550Dx (dokument č. 1000000009513)*.

Metódy analýzy

Analytický modul Germline Variant vykonáva nasledujúce kroky analýzy a potom zapisuje súbory s výstupmi analýzy do priečinka Alignment (Zarovnanie).

- ▶ Demultiplexuje čítania indexov.
- ▶ Generuje súbory FASTQ.
- ▶ Zarovnáva s referenciou.
- ▶ Identifikuje varianty.

Demultiplexovanie

Demultiplexovanie porovnáva každú sekvenciu čítania indexu so sekvenciami indexov špecifikovaných pre chod. V tomto kroku sa neberú do úvahy žiadne kvalitatívne hodnoty.

Čítania indexov sa identifikujú týmito krokmi:

- ▶ Vzorky sa očísľujú od 1 podľa poradia, v akom sú uvedené na účely chodu.
- ▶ Číslo vzorky 0 je vyhradené pre klastre, ktoré neboli priradené ku vzorke.
- ▶ Klastre sa ku vzorke priradujú vtedy, keď sa sekvencia indexov presne zhoduje alebo keď existuje maximálne jedna nezhoda na čítanie indexu.

Generovanie súborov FASTQ

Po demultiplexovaní softvér vygeneruje súbory predbežnej analýzy vo formáte FASTQ, čo je textový formát používaný na znázornenie sekvencií. Súbory FASTQ obsahujú čítania každej vzorky a súvisiace kvalitatívne skóre. Klastre, ktoré nespĺňajú požiadavky filtra, sú vylúčené.

Každý súbor FASTQ obsahuje čítania len jednej vzorky, pričom názov tejto vzorky je uvedený v názve súboru FASTQ. Súbory FASTQ predstavujú primárny vstup na zarovnanie. Na vzorku sa vygeneruje osem súborov FASTQ: štyri z čítania 1 a štyri z čítania 2.

Zarovnanie

Počas kroku zarovnania pásmový Smithov-Watermanov algoritmus zarovnáva klastre z každej vzorky oproti sekvenciám amplicónov špecifikovaných v súbore manifestu.

Pásmový Smithov-Watermanov algoritmus vykonáva semiglobálne zarovnania sekvencií s cieľom určiť podobné oblasti medzi dvoma sekvenciami. Namiesto porovnania celej sekvencie Smithov-Watermanov algoritmus porovnáva segmenty všetkých možných dĺžok.

Čítanie oboch koncov (tzv. paired-end) sa vyhodnocuje v súvislosti s jeho zarovnaním voči relevantným sekvenciám prób pre dané čítanie.

- ▶ Čítanie 1 sa vyhodnocuje oproti reverznému komplementu nižšie položených oligonukleotidov, ktoré sú špecifické pre lokus (DLSO).
- ▶ Čítanie 2 sa vyhodnocuje oproti vyššie položeným oligonukleotidom, ktoré sú špecifické pre lokus (ULSO).

- ▶ Ak sa začiatok čítania zhoduje so sekvenciou próby a existujú maximálne tri rozdiely (nezhody alebo posuny v dôsledku hlavných indelov), celá dĺžka čítania sa zarovná s cieľovým amplicónom tejto sekvencie.
- ▶ Indely v rámci DLSO a ULSO sa vzhľadom na chemické vlastnosti analýzy nezaznamenávajú.

Na základe frekvencie nezhôd buď v oblasti záujmu, alebo celom amplicóne (v závislosti od dĺžky amplicónu) sa zarovnanie filtrujú z výsledkov zarovnaní. Vyfiltrované zarovnanie sa zapisujú do súborov zarovnaní a nepoužívajú sa pri stanovení variantov (tzv. calling).

Stanovenie variantov

Program Pisces Variant Caller, ktorý vyvinula spoločnosť Illumina, identifikuje varianty prítomné vo vzorke DNA.



Program Pisces Variant Caller identifikuje SNV, MNV a malé indely v 3 krokoch:

- ▶ Samostatne zvažuje každú pozíciu v referenčnom genóme.
- ▶ Počíta bázy v danej pozícii pre zarovnané čítania, ktoré presahujú pozíciu.
- ▶ Vypočítava skóre variantu, ktoré meria kvalitu stanovenia pomocou Poissonovho modelu. Varianty s kvalitatívnym skóre nižším ako Q20 sa vylúčia.

Ak variant vyhovuje všetkým požiadavkám filtra, variant sa vo VCF označí ako PASS (Úspešný).

Viac informácií nájdete na stránke github.com/Illumina/Pisces/wiki.

Zobrazenie údajov o chode a vzorke

- 1 Na ovládacom paneli softvéru Local Run Manager (Správca lokálnych chodov) kliknite na názov chodu.
- 2 Na karte Run Overview (Prehľad chodu) skontrolujte metriku sekvenovacieho chodu.
- 3 **[Voliteľné]** Ak chcete skopírovať dráhu priečinka s výstupom chodu, kliknite na ikonu **kopírovania do schránky** .
- 4 Kliknutím na kartu Sequencing Information (Informácie o sekvenovaní) skontrolujte parametre chodu a informácie o spotrebnom materiáli.
- 5 Ak chcete zobraziť umiestnenie správy o analýze, kliknite na kartu Samples and Results (Vzorky a výsledky).
 - ▶ Ak sa analýza opakovala, otvorte rozbaľovací zoznam Select Analysis (Vybrať analýzu) a vyberte príslušnú analýzu.
- 6 Ak chcete skopírovať dráhu priečinka s analýzou, kliknite na ikonu **kopírovania do schránky** .

Viac informácií o kartách Run Overview (Prehľad chodu) a Sequencing Information (Informácie o sekvenovaní) a o tom, ako znova zaradiť analýzu do frontu, nájdete v *referenčnej príručke prístroja NextSeq 550Dx* (dokument č. 1000000009513).

Správa o analýze

Výsledky analýzy sú zhrnuté na karte Samples and Results (Vzorky a výsledky) a ako súhrnná správa v priečinku Alignment (Zarovnanie). Pre každú vzorku je taktiež k dispozícii správa vo formáte súboru PDF.

Informácie o karte Samples and Results (Vzorky a výsledky)

1 Kliknutím na vzorku v zozname zobrazíte správu o vzorke.

Tabuľka 1 Informácie o chode a vzorke

Hlavička stĺpca	Popis
Run Status (Stav chodu)	Uvádza, či bol sekvenovací chod úspešný alebo neúspešný.
Total Yield (GB) (Celková výťažnosť (GB))	Počet báz stanovených v sekvenovacom chode. Zobrazuje prahovú hodnotu úspešnosti a úspešný alebo neúspešný stav.
% ≥ Q30	Percento čítaní v sekvenovacom chode so skóre kvality 30 (Q30) alebo vyšším. Zobrazuje prahovú hodnotu úspešnosti a úspešný alebo neúspešný stav.
Sample ID (ID vzorky)	ID vzorky poskytnuté pri vytváraní chodu.
Total PF Reads (Celkový počet čítaní PF)	Celkový počet čítaní, ktoré splnili požiadavky filtra.
Read 1% ≥ Q30 (Čítanie 1 % ≥ Q30)	Percento čítaní v čítaní 1 so skóre kvality vzorky 30 (Q30) alebo vyšším.
Read 2% ≥ Q30 (Čítanie 2 % ≥ Q30)	Percento čítaní v čítaní 2 so skóre kvality vzorky 30 (Q30) alebo vyšším.
Autosome Call Rate (Miera stanovenia autozómov)	Počet genomických pozícií medzi autozómami (chromozómy 1 až 22), ktoré spĺňajú vopred definovanú prahovú hodnotu spoľahlivosti, vydelený celkovým počtom vyšetrovaných autozomálnych genómových pozícií. Miera stanovenia sa opisuje po jednotlivých vzorkách a nahlasuje sa ako percento, ktoré sa vypočítava ako 1 mínus (počet autozomálnych pozícií s nekompletnými stanoveniami vydelený celkovým počtom sekvenovaných autozomálnych pozícií).

Tabuľka 2 Informácie o správe o vzorke

Hlavička stĺpca	Popis
Sample (Vzorka)	ID vzorky poskytnuté pri vytváraní chodu.
Report Date (Dátum správy)	Dátum vytvorenia správy.
Sample Information (Informácie o vzorke)	ID vzorky poskytnuté pri vytváraní chodu, celkový počet čítaní vo vzorke, ktoré vyhoveli požiadavkám filtra, percento čítaní vzorky so skóre kvality 30 (Q30) alebo vyšším a miera autozomálnych stanovení.
Amplicon Summary (Zhrnutie amplicónov)	Celkový počet sekvenovaných oblastí amplicónov a celková dĺžka v bazových pároch sekvenovaných amplicónov v cieľových oblastiach pre vzorku a súbor manifestu. Súbor manifestu určuje referenčný genóm a cieľové referenčné oblasti použité v kroku zarovnania.
Read Level Statistics (Štatistika úrovne čítania)	Počet a percento čítaní vzorky, ktoré pokrývajú každú referenčnú pozíciu čítania 1 a čítania 2.
Variants Summary (Súhrn variantov)	Počet SNV, inzercíí a delácií detegovaných pre vzorku, ktoré splnili navrhnuté hodnoty s cieľom určiť, či sa kvalitatívne výsledky nachádzajú v rámci prijateľného rozsahu.
Coverage Summary (Súhrn pokrytia)	Celkový počet zarovnaných báz vydelený veľkosťou cieľovej oblasti a percento oblastí amplicónov s hodnotami pokrytia väčšími ako dolná prahová hodnota pokrytia (0,2 * priemerné pokrytie amplicónu) vzorky.
Coverage Plots (Grafy pokrytia)	Graf pokrytia podľa oblasti amplicónov znázorňuje pokrytie v rámci oblastí amplicónov vzorky. Oblasti s hodnotami pokrytia nižšími ako je prahová hodnota pokrytia sa zvýraznia načerveno. Priemer všetkých hodnôt označuje oranžová čiara.
Software Versions (Verzie softvéru)	Verzie softvéru pri sekvenovaní vzorky. Zahŕňa verziu operačného softvéru NextSeq 550Dx (NOS), softvéru Local Run Manager (Správca lokálnych chodov), softvéru RTA a modulu Germline Variant.

Súbory s výstupmi analýzy

Pre analytický modul Germline Variant sa generujú nasledujúce súbory s výstupmi analýzy, ktoré poskytujú výsledky analýzy na zarovnanie a stanovenie variantov. Súbory s výstupmi analýzy sa nachádzajú v priečinku Alignment (Zarovnanie).

Názov súboru	Popis
Demultiplexovanie (*.txt)	Predbežné súbory obsahujúce súhrnné výsledky demultiplexovania.
FASTQ (*.fastq.gz)	Predbežné súbory obsahujúce stanovené bázy s kvalitatívnym skóre. Súbory FASTQ predstavujú primárny vstup na krok zarovnanie.
Súbory zarovnaní vo formáte BAM (*.bam)	Obsahujú zarovnané čítania danej vzorky.
Súbory so stanovenými variantmi v genóme vo formáte VCF (*.genome.vcf.gz)	Obsahujú genotyp pre každú pozíciu, či už stanovenú ako variant alebo ako referencia.
Súbory so stanovenými variantmi vo formáte VCF (*.vcf.gz)	Obsahujú všetky varianty stanovené v cieľovej oblasti.
AmpliconCoverage_M1.tsv	Obsahuje informácie o pokrytí na amplicón pre každý poskytnutý manifest. MČ. predstavuje číslo manifestu.

Formát súboru demultiplexovania

Pri procese demultiplexovania sa čítajú sekvencie indexov pripojené ku každému klastru s cieľom určiť, z ktorej vzorky klaster pochádza. Mapovanie medzi klastrami a číslom vzorky sa zapisuje do súboru demultiplexovania (*.demux) pre každú dlaždicu prietokovej komôrky.

Formát pomenovania súboru demultiplexovania je **s_1_X.demux**, kde X je číslo dlaždice.

Súbory demultiplexovania sa začínajú nadpisom:

- ▶ verzia (4-bajtové celé číslo), aktuálne 1,
- ▶ počet klasterov (4-bajtové celé číslo).

Zvyšok súboru tvoria čísla vzoriek pre každý klaster z dlaždice.

Po dokončení kroku demultiplexovania softvér vygeneruje súbor demultiplexovania s názvom **DemultiplexSummaryF1L1.txt**.

- ▶ **F1** v názve súboru predstavuje číslo prietokovej komôrky.
- ▶ **L1** v názve súboru predstavuje číslo pruhu.
- ▶ Výsledky demultiplexovania v tabuľke s 1 riadkom na dlaždicu a 1 stĺpcom na vzorku vrátane vzorky 0.
- ▶ Najčastejšie sa vyskytujúce sekvencie v čítaniach indexov.

Formát súborov FASTQ

FASTQ je textový formát súborov, ktorý obsahuje stanovené bázy a kvalitatívne hodnoty na čítanie. Každý záznam obsahuje 4 riadky:

- ▶ identifikátor,
- ▶ sekvencia,
- ▶ znamienko plus (+),
- ▶ Phredovo skóre kvality v ASCII + 33 kódovanom formáte.

Identifikátor je naformátovaný ako:

@Prístroj:ID chodu:ID prietokovej komôrky:Pruh:Dlaždica:X:Y Č. čítania:Príznak filtra:0:Číslo vzorky

Príklad:

```
@SIM:1:FCX:1:15:6329:1045 1:N:0:2
TCGCACTCAACGCCCTGCATATGACAAGACAGAATC
+
<>;##=><9=AAAAAAAAAA9#:<#<;<<<????#=#
```

Formát súborov BAM

Súbor BAM (*.bam) je komprimovaná binárna verzia súboru SAM, ktorá sa používa na zobrazenie zarovnaných sekvencií do 128 Mb. Formáty SAM a BAM sú podrobne opísané na stránke samtools.github.io/hts-specs/SAMv1.pdf.

Súbory BAM využívajú formát pomenovania **NázovVzorky_S#.bam**, kde # je číslo vzorky určené podľa poradia, v ktorom sú vzorky uvedené na účely chodu.

Súbory BAM obsahujú časť s hlavičkou a časť so zarovnaním:

- ▶ **Hlavička** – obsahuje informácie o celom súbore (napr. názov vzorky, dĺžka vzorky a metóda zarovnaní). Zarovnaní v časti so zarovnaním sú spojené so špecifickými informáciami v časti s hlavičkou.
- ▶ **Zarovnanie** – obsahuje názov čítania, sekvenciu čítania, informácie o zarovnaní a prispôbené príznaky. Názov čítania zahŕňa chromozóm, počiatkové súradnice, kvalitu zarovnaní a reťazec deskriptora zhody.

V časti so zarovnaním sú uvedené nasledujúce informácie pre každé čítanie alebo čítaný pár:

- ▶ **AS:** kvalita zarovnaní na oboch koncoch,
- ▶ **RG:** skupina čítania, ktorá indikuje počet čítaní špecifickej vzorky,
- ▶ **BC:** štítok (tzv. tag) čiarového kódu, ktorý indikuje ID demultiplexovanej vzorky súvisiacej s čítaním,
- ▶ **SM:** kvalita zarovnaní na jednom konci,
- ▶ **XC:** reťazec deskriptora zhody,
- ▶ **XN:** štítok (tzv. tag) názvu amplikónu, v ktorom je zaznamenané ID amplikónu súvisiaceho s čítaním.

Súbory indexov BAM (*.bam.bai) poskytujú index zodpovedajúceho súboru BAM.

Formát súborov VCF

Formát stanovenia variantov (VCF) je bežný formát súborov vyvinutých v genómovej odbornej komunite. Obsahuje informácie o variantoch zistených v špecifických pozíciách v referenčnom genóme. Súbory VCF majú príponu .vcf.

V hlavičke súboru VCF sa uvádza verzia formátu súboru VCF a verzia prekladača (tzv. caller) variantu a tiež poznámky použité v zostávajúcej časti súboru. V hlavičke súboru VCF je tiež uvedený súbor referenčného genómu a súbor BAM. V poslednom riadku hlavičky sú uvedené nadpisy stĺpcov na riadky s údajmi. V každom riadku s údajmi súboru VCF sú uvedené informácie o jednom variante.

Hlavičky súborov VCF

Hlavička	Popis
CHROM	Chromozóm referenčného genómu. Chromozómy sa zobrazujú v rovnakom poradí ako v referenčnom súbore FASTQ.
POS	Jednobázová pozícia variantu v referenčnom chromozóme. Pre SNP je táto pozícia referenčnou bázou s variantom; pre indely a delécie je táto pozícia referenčnou bázou bezprostredne pred variantom.
ID	Číslo rs pre variant získaný z dbSNP.txt (ak je to uplatniteľné). Ak sa na tomto mieste nachádza viacero čísel rs, zoznam je oddelený bodkočiarkami. Ak pre túto pozíciu neexistuje žiadny dbSNP vstup, použije sa značka chýbajúcej hodnoty („.“).
REF	Referenčný genotyp. Napríklad delécia jedného T je vyjadrená ako referenčný TT alebo alternatívny T. Variant s jedným nukleotidom z A na T je vyjadrený ako referenčný A a alternatívny T.
ALT	Alely, ktoré sa líšia od referenčného čítania. Napríklad inzercia jediného T je vyjadrená ako referenčný A a alternatívny AT. Variant s jedným nukleotidom z A na T je vyjadrený ako referenčný A a alternatívny T.
QUAL	Skóre kvality podľa Phredovej škály priradené k prekladaču (tzv. caller) variantu. Vyššie skóre označuje vyššiu istotu vo variante a nižšiu pravdepodobnosť chýb. Pre skóre kvality Q je odhadovaná pravdepodobnosť chyby $10^{-(Q/10)}$. Napríklad súbor stanovení Q30 má mieru výskytu chyby 0,1 %. Mnoho prekladačov variantov prideluje skóre kvality na základe svojich štatistických modelov, ktoré sú vysoké v porovnaní s pozorovanou mierou výskytu chýb.

Poznámky v súboroch VCF

Hlavička	Popis
FILTER	Ak je vyhovené všetkým filtrom, do stĺpca na filter sa zapíše PASS (Úspešné). <ul style="list-style-type: none"> • LowDP – platí pre pracoviská s hĺbkou pokrytia pod 150x. Pri pozíciách amplikónov pokrytých priamym aj reverzným čítaním je to ekvivalent 300 prekrývajúcich sa čítaní oboch koncov. • q20 – skóre kvality < 20. • MultiAllelicSite – variant nezodpovedá diploidnému modelu. • R5x9 – počet susediacich opakovaní (dĺžky 1 až 5 bp) k stanoveným variantom ≥ 9. • SB – odchýlka vlákna je vyššia ako je daná prahová hodnota.
INFO (INFORMÁCIE)	V stĺpci INFO sú možné tieto záznamy: <ul style="list-style-type: none"> • AC – počet alel v genotypoch pre každú ALT alelu v tom istom poradí, ako sú uvedené. • AF – frekvencia alel pre každú ALT alelu v takom istom poradí, ako sú uvedené. • AN – celkový počet alel v stanovených genotypoch. • CD – príznak, ktorý označuje, že sa SNP vyskytuje v oblasti kódovania aspoň 1 záznamu referenčného génu (RefGene). • DP – hĺbka (počet stanovení báz zarovnaných k pozícii a použitých pri stanovení variantov). • Exon – čiarkami oddelený zoznam oblastí exónov prečítaných z referenčného génu. • FC – funkčný dôsledok. • GI – čiarkami oddelený zoznam ID génov prečítaných z referenčného génu. • QD – spoľahlivosť/kvalita variantov podľa hĺbky. • TI – čiarkami oddelený zoznam ID transkriptov prečítaných z referenčného génu.

Hlavička	Popis
FORMAT (FORMÁT)	<p>V stĺpci formátu sú políčka oddelené dvojbodkami. Napríklad GT:GQ. Zoznam poskytnutých políčok závisí od použitého prekladača (tzv. caller) variantov. K dispozícii sú tieto políčka:</p> <ul style="list-style-type: none"> • AD – záznam v tvare X, Y, kde X je počet referenčných stanovení a Y je počet alternatívnych stanovení. • DP – približná hĺbka čítania; čítanie s MQ = 255 alebo zlými väzbami sú odfiltrované. • GQ – kvalita genotypu. • GT – genotyp. 0 zodpovedá referenčnej báze, 1 zodpovedá prvému záznamu v stĺpci ALT atď. Lomka (/) znamená, že k dispozícii nie sú žiadne informácie o fázovaní. • NC – frakcia báz, ktoré neboli stanovené alebo mali kvalitu stanovenia báz nižšiu ako minimálna prahová hodnota. • NL – úroveň šumu; odhad šumu stanovenia báz v tejto pozícii. • SB – odchýlka vlákna v tejto pozícii. Väčšie negatívne hodnoty označujú menšiu odchýlku; hodnoty blízko 0 označujú väčšiu odchýlku. • VF – frekvencia variantov; percento čítaní podporujúcich alternatívnu alelu.
SAMPLE (VZORKA)	Stĺpec vzorky udáva hodnoty špecifikované v stĺpci FORMAT (FORMÁT).

Súbory VCF genómu

Súbory VCF genómu (gVCF) sú súbory typu VCF v4.1, ktoré sa riadia súborom konvencií na vyjadrenie všetkých miest v rámci genómu v primerane kompaktnom formáte. Súbory gVCF (*.genome.vcf.gz) obsahujú všetky miesta v oblasti záujmu v jedinom súbore pre každú vzorku.

Súbor gVCF zobrazuje absenciu stanovenia v pozíciách, ktoré nespĺnili všetky požiadavky filtrov. Štítok genotypu (GT) ./. označuje žiadne stanovenie.

Viac informácií nájdete na stránke sites.google.com/site/gvcftools/home/about-gvcf.

Súbor pokrytia amplikónov

Súbor pokrytia amplikónov sa generuje pre každý súbor manifestu. MČ. v názve súboru predstavuje číslo manifestu.

V každom súbore je riadok s hlavičkou, kde sú uvedené ID vzoriek súvisiacich s manifestom. V súbore sú uvedené nasledujúce informácie.

- ▶ Cieľové ID tak, ako je uvedené v manifeste.
- ▶ Hĺbka pokrytia čítaní, ktoré splnili požiadavky filtra.

Doplnkové súbory s výstupmi

Nasledujúce súbory s výstupmi obsahujú doplnujúce informácie alebo sumarizujú výsledky chodu a chyby analýzy. Hoci sa tieto súbory nevyžadujú na hodnotenie výsledkov analýzy, môžu sa použiť na riešenie problémov. Pokiaľ nie je uvedené inak, všetky súbory sa nachádzajú v priečinku Alignment (Zarovnanie).

Názov súboru	Popis
AnalysisLog.txt	<p>Protokol spracovania, v ktorom sú opísané jednotlivé kroky, ktoré sa vyskytli počas analýzy aktuálneho priečinka chodu. V tomto súbore nie sú uvedené chybové hlásenia.</p> <p>Nachádza sa v priečinku Alignment (Zarovnanie).</p>
AnalysisError.txt	<p>Protokol spracovania, v ktorom sú uvedené všetky chyby, ktoré sa vyskytli počas analýzy. Ak sa nevyskytli žiadne chyby, tento súbor bude prázdny.</p> <p>Nachádza sa v priečinku Alignment (Zarovnanie).</p>

Názov súboru	Popis
DemultiplexSummaryF1L1#.txt	Tu sú uvedené výsledky demultiplexovania v tabuľke s 1 riadkom na dlaždicu a 1 stĺpcom na vzorku. Znak # predstavuje pruh 1, 2, 3 alebo 4 prietokovej komôrky. Nachádza sa v priečinku Alignment (Zarovnanie).
AmpliconRunStatistics.xml	Obsahuje štatistické zhrnutie, ktoré je špecifické pre chod. Nachádza sa v priečinku Alignment (Zarovnanie).

Priečinko analýzy

V priečinku analýzy sa nachádzajú súbory, ktoré vygeneroval softvér Local Run Manager (Správca lokálnych chodov).


Vzťah medzi priečinkom s výstupmi a priečinkom analýzy je zhrnutý takto:


- ▶ Počas sekvenovania zaplní softvér Real-Time Analysis (RTA) priečinko s výstupmi súbormi, ktoré sa vygenerovali počas analýzy snímok, stanovenia báz a kvalitatívneho skórovania.
- ▶ Softvér RTA skopíruje súbory do priečinka analýzy v reálnom čase. Po tom, ako softvér RTA priradí skóre kvality každej báze každého cyklu, softvér zapíše súbor RTAComplete.txt do oboch priečinkov.
- ▶ Analýza sa začne vtedy, keď je prítomný súbor RTAComplete.txt.
- ▶ V priebehu analýzy softvér Local Run Manager (Správca lokálnych chodov) zapisuje súbory s výstupmi do priečinka analýzy a potom kopíruje súbory späť do priečinka s výstupmi.

Priečinky Alignment (Zarovnanie)

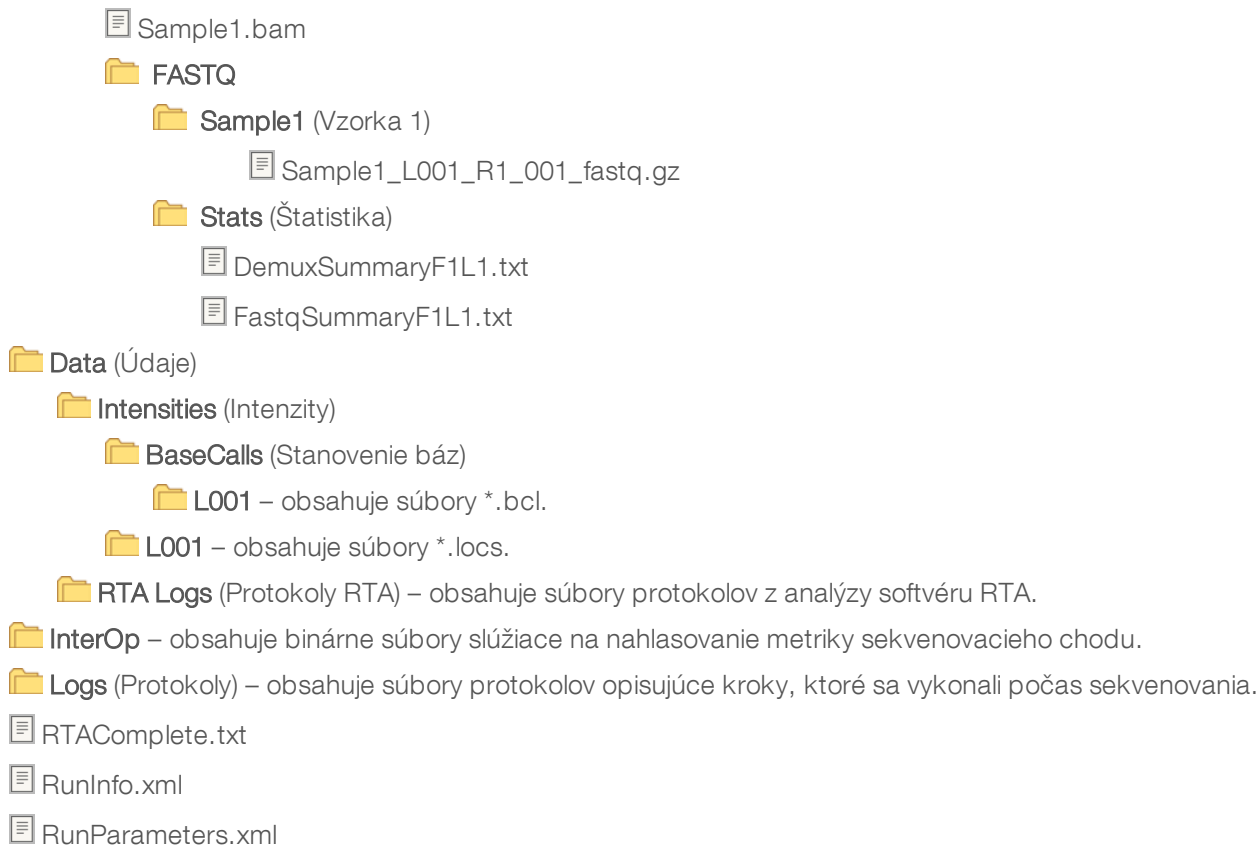
Pri každom opätovnom zaradení analýzy do frontu vytvorí softvér Local Run Manager (Správca lokálnych chodov) priečinko Alignment (Zarovnanie) s názvom **Alignment_N**, kde N je poradové číslo.

Štruktúra priečinkov

 **Alignment** (Zarovnanie) – obsahuje súbory *.bam, *.vcf, FASTQ a súbory, ktoré sú špecifické pre analytický modul.

 **Date and Time Stamp** (Označenie dátumu a času) – označenie dátumu a času analýzy vo formáte RRRRMMDD_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



Stanovenie báz a diverzita indexov

Pri sekvenovaní vzoriek v prístroji NextSeq 550Dx sa stanovením báz určuje báza (A, C, G alebo T) každého klastra danej dlaždice alebo snímacej oblasti v prietokovej komôrke v konkrétnom cykle. Prístroj NextSeq 550Dx používa dvojkanálové sekvenovanie, ktoré si na kódovanie údajov štyroch DNA báz vyžaduje len dve snímky: jednu z červeného kanála a druhú zo zeleného kanála.

Proces čítaní indexov stanovenia báz sa odlišuje od stanovenia báz počas iných čítaní.

Čítania indexov sa musia začínať minimálne jednou bázou inou ako G v ktoromkoľvek z prvých dvoch cyklov. Ak sa čítanie indexov začne dvoma stanovenými bázami G, nevygeneruje sa žiadna intenzita signálu. Signál musí byť v ktoromkoľvek z prvých dvoch cyklov prítomný, aby sa zabezpečilo fungovanie demultiplexovania.

Ak indexy nespĺňajú požiadavky na diverzitu, počas výberu indexov pri vytváraní chodu sa zobrazí varovanie týkajúce sa nízkej diverzity. Ak chcete predísť varovaniu týkajúcemu sa nízkej intenzity, vyberte sekvencie indexov, ktoré poskytujú signál v oboch kanáloch pre každý cyklus.

- ▶ Červený kanál – A alebo C
- ▶ Zelený kanál – A alebo T

Tento proces stanovenia báz zabezpečuje správnosť pri analyzovaní malého počtu vzoriek. Viac informácií o sekvenciách vašich indexov nájdete v príbalovom letáku súpravy *TruSeq Custom Amplicon Kit Dx* (dokument č. 1000000029772).

Počas vytvárania chodu v softvéri Local Run Manager (Správca lokálnych chodov) vyberiete počet vzoriek, ktoré sa majú testovať. Navrhované kombinácie indexov, ktoré spĺňajú požiadavky na diverzitu, automaticky doplní softvér. Hoci nie je použitie navrhovaných kombinácií indexov povinné, odporúča sa.

História revízií

Dokument	Dátum	Popis zmeny
Dokument č. 1000000030329 v04	August 2021	Aktualizovaná adresa splnomocneného zástupcu EÚ.
Dokument č. 1000000030329 v03	Apríl 2020	Aktualizovaná adresa splnomocneného zástupcu EÚ. Aktualizovaná adresa austrálskeho zadávateľa.
Dokument č. 1000000030329 v02	Január 2019	Pridané informácie o súpravách reagencií v2.5.
Dokument č. 1000000030329 v01	August 2018	Aktualizované regulačné označenia.
Dokument č. 1000000030329 v00	November 2017	Úvodné vydanie.

Technická pomoc

Technickou pomoc vám poskytne technická podpora společnosti Illumina.

Webová lokalita: www.illumina.com
E-mail: techsupport@illumina.com

Telefonní čísla oddělení zákaznické podpory společnosti Illumina

Región	Bezplatné	Regionálně
Severná Amerika	+1.800.809.4566	
Austrálie	+1.800.775.688	
Belgicko	+32 80077160	+32 34002973
Čína	400.066.5835	
Dánsko	+45 80820183	+45 89871156
Fínsko	+358 800918363	+358 974790110
Francúzsko	+33 805102193	+33 170770446
Holandsko	+31 8000222493	+31 207132960
Hongkong	800960230	
Írsko	+353 1800936608	+353 016950506
Japonsko	0800.111.5011	
Nemecko	+49 8001014940	+49 8938035677
Nórsko	+47 800 16836	+47 21939693
Nový Zéland	0800.451.650	
Rakúsko	+43 800006249	+43 19286540
Singapur	+1.800.579.2745	
Španielsko	+34 911899417	+34 800300143
Spojené kráľovstvo	+44 8000126019	+44 2073057197
Švajčiarsko	+41 565800000	+41 800200442
Švédsko	+46 850619671	+46 200883979
Taiwan	00806651752	
Taliansko	+39 800985513	+39 236003759
Ostatné krajiny	+44 1799 534000	

Karty bezpečnostných údajov (Safety data sheets, SDS) – k dispozícii na webovej lokalite spoločnosti Illumina na stránke support.illumina.com/sds.html.

Produktová dokumentácia – k dispozícii na prevzatie vo formáte PDF z webovej lokality spoločnosti Illumina. Prejdite na stránku support.illumina.com, vyberte produkt a potom položku **Documentation & Literature** (Dokumentácia a literatúra).



Illumina
5200 Illumina Way
San Diego, California 92122
U.S.A.
+1.800.809.ILMN (4566)
+1.858.202.4566 (okrem Severnej
Ameriky)
techsupport@illumina.com
www.illumina.com

CE

IVD

EC REP

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

Austrálsky zadávateľ

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

NA DIAGNOSTICKÉ ÚČELY I N VITRO

© 2021 I Illumina, I nc. Všetky práva vyhradené.

illumina®