

DRAGEN for Illumina DNA Prep with Enrichment Dx on NextSeq 550Dx

Notendahandbók fyrir forritið

Þetta skjal og innihald þess eru eign Illumina, Inc. og dótturfélaga þess („Illumina“) og eru eingöngu ætluð til samningsbundinnar notkunar viðskiptavina þess í tengslum við notkun þeirrar vöru/þeirra vara sem hér er lýst og í engum öðrum tilgangi. Þetta skjal og efni þess skal ekki notað eða dreift í neinum öðrum tilgangi og/eða á annan hátt miðlað, birt eða afritað án skriflegs samþykkis Illumina. Illumina veitir ekki leyfi samkvæmt einkaleyfi sínu, vörumerki, höfundarrétti eða réttindum samkvæmt venjurétti né svipuð réttindi þriðja aðila með þessu skjali.

Leiðbeiningunum í þessu skjali verður að fylgja stranglega og á skýran hátt af hæfu og vel þjálfuðu starfsfólki til að tryggja rétta og örugga notkun vörunnar/varanna sem hér er lýst. Lesa verður og skilja allt innihald þessa skjals að fullu áður en slík vara eða vörur eru notaðar.

EF EKKI ER FARIÐ NÁKVÆMLEGA EFTIR ÖLLUM LEIÐBEININGUM SEM HÉR KOMA FRAM, GETUR ÞAD LEITT TIL SKEMMDA Á VÖRU/VÖRUM, MEIÐSLA Á FÓLKI, ÞAR Á MEDAL NOTENDUM OG ÖÐRUM, AUK SKEMMDA Á ÖÐRUM EIGNUM. SLÍKT ÓGILDIR ALLA ÁBYRGÐ SEM GILDIR UM VÖRU/VÖRUR.

ILLUMINA BER EKKI ÁBYRGÐ VEGNA RANGRAR NOTKUNAR VÖRUNNAR/VARANNA SEM HÉR ER LÝST (ÞAR Á MEDAL HLUTA ÞEIRRA EÐA HUGBÚNAÐAR).

© 2023 Illumina, Inc. Allur réttur áskilinn.

Öll vörumerki eru eign Illumina, Inc. eða viðkomandi eigenda. Sérstakar upplýsingar um vörumerki er að finna á www.illumina.com/company/legal.html.

Útgáfuferill

Skjal	Dagsetning	Lýsing á breytingu
200025238 útgáfa 00	Febrúar 2023	Fyrsta útgáfa.

Efnisyfirlit

Útgáfuferill	iii
Yfirlit	1
Greiningaraðferðir	1
Búa til fyrirhugaða keyrslu	4
Stillingar	6
Upplýsingaskrá	7
Hávaðasiun (valfrjálst)	8
Greiningarúttak	8
FASTQ-skrár	9
BAM-skrár	10
VCF-skrár	10
Endurtaka greiningu	16
Tæknileg aðstoð	17

Yfirlit

DRAGEN for Illumina DNA Prep with Enrichment Dx Forritið (DRAGEN for IDPE Dx) er notað til að skipuleggja og framkvæma auka greiningu á IDPE Dx sölum sem mynduð eru fyrir raðgreiningu á NextSeq 550Dx.

DRAGEN for IDPE Dx styður raðgreiningu til greiningar þegar það er notað með Illumina DNA Prep with Enrichment Dx Library Prep, NextSeq 550Dx og Illumina DRAGEN Server for NextSeq 550Dx.

Greiningaraðferðir

DRAGEN for IDPE Dx framkvæmir affléttun, FASTQ myndun, lesvörpun og samstillingu við viðmiðunarerfðamengi og tilbrigðisköllun, allt eftir völdu vinnuflæði.

- Gerð FASTQ-skráa
- Kímlínu FASTQ og VCF myndun
- Líkamleg FASTQ og VCF myndun

ATHUGIÐ ORA þjöppun er tiltæk til notkunar með öllum þremur vinnuflæðum. DRAGEN ORA Compression er fullkomlega taplaus þjöppunarhugbúnaður sem býr til skrá með viðskeytinu Original Read Archive (*.ora). Ora-sniðið er tilvísunarbundið þjöppunarsnið fyrir FASTQ-skrár og er hannað fyrir mjög hraða þjöppun/afþjöppun og hátt þjöppunarhlutfall.

FASTQ myndun

Samsettu raðirnar eru skrifaðar í FASTQ-skrár fyrir hvert sýni. FASTQ-skrár eru textaskrár sem innihalda raðgreiningargögn og gæðastig fyrir aðeins eitt sýni. Fyrir hvert sýni eru aðskildar FASTQ-skrár búnar til fyrir hverja flæðisflögurein, fyrir hverja raðgreiningu. Nafn sýnisins eins og það var tilgreint við uppsetningu keyrslu er innifalið í FASTQ-skráarnafninu. FASTQ-skrár eru aðalinntakið fyrir samstillingu. Fyrsta skrefið í myndun FASTQ-skráa er afblöndun. Aðskiljun sýna úthlutar klösum sem fara í gegnum síu til sýna með því að bera saman hverja vísislestraröð við þær vísiraðir sem tilgreindar eru fyrir keyrsluna. Engin gæðagildi eru tekin til greina í þessu skrefi. Vísislestrar eru greindir með eftirfarandi skrefum:

- Sýnin eru númeruð frá 1 eftir því í hvaða röð þau eru skráð fyrir keyrsluna.
- Úrtaksnúmer 0 er frátekið fyrir klasa sem ekki var úthlutað úrtaki.
- Klösum er úthlutað á sýni þegar vísiröðin passar nákvæmlega saman eða þegar allt að eitt ósamræmi er í hverjum vísislestri.

Hugbúnaðurinn inniheldur ORA-þjöppun til að þjappa FASTQ-skrám. Þetta snið er hægt að virkja valfrjálst. Þegar ORA-sniðið (*.ora) er notað er md5 prófsumma FASTQ innihaldsins varðveitt eftir þjöppunar- og afþjöppunarferli til að tryggja taplausa þjöppun.

DNA vörpun & samstilling

Eftir FASTQ-myndun eru lestrar varpaðir á og samstilltir við viðmiðunarerfðamengi. Fyrsta stig vörpunar er að búa til kím úr lestrinum og leita síðan að nákvæmum samsvörum í viðmiðunarerfðamenginu. Þessar niðurstöður eru síðan betrubættar með því að keyra fulla Smith-Waterman samstillingu á þeim stöðum þar sem þéttleikinn af kímsamsvörum er mestur. Þetta vel skjalfesta reiknirit virkar með því að bera saman hverja stöðu lestursins við allar mögulegar stöður viðmiðsins. Þessir samanburðir samsvara fylki mögulegra samstillinga milli lestrar og viðmiðs. Fyrir hvern af þessum mögulegu samstillingarstöðum reiknar Smith-Waterman aðferðin stig sem eru notuð til að meta hvort besta samstillingin sem fer í gegnum viðkomandi reit í fylkinu nái honum með kirkisamsvörum eða mispörun (skáhalli hreyfingu), úrfellingu (láréttri hreyfingu) eða innskoti (lóðréttri hreyfingu). Samsvörum milli lestrar og viðmiðs gefur bónus á stigatöfluna, og mispörun eða úrfelling/innskot veldur refsingu. Leiðin sem fær hæstu stig í gegnum fylkið er sú röðun sem valin er. Reikniritið er vélbúnaðarhraðað á DRAGEN field-programmable gate array (FPGA) kortum. Viðmiðunarerfðamengið sem notað er í forritinu er búið til úr UCSC hg19 FASTA með DRAGEN valkostinum til að búa til liftover-byggða alt-meðvitaða tætitöflu.

DRAGEN Germline Variant Calling

DRAGEN Germline Small Variant Caller notar vörpuð og samræmd DNA-lesgögn sem inntak og greinir einkirabreytileika (SNP) og innskot eða úrfellingar (indels) með samsettri dálkagreiningu og staðbundinni *de novo* samsetningu setraða. Til að virkja DRAGEN Germline Small Variant Caller skaltu velja vinnufæðið fyrir kímlínutilbrigði.

Kímlínutilbrigðaköllun er venjulega notuð fyrir kímlínusýni þar sem vitað er að litnun er tveir. Kallanleg viðmiðunarsvæði eru fyrst greind með nægilegri samstillingarþekju. Innan þessara viðmiðunarsvæða greinir hraðskönnun á flokkuðum lestrum virk svæði, sem eru miðjuð á uppsafnaða dálka með vísbendingum um tilbrigði. Virku svæðin eru fyllt með nægu samhengi til að hylja mikilvægt efni sem er ekki viðmið. Ef vísbendingar eru um innskot eða úrfellingar fá virk svæði viðbótarfyllingu.

Samstilltir lestrar eru klipptir innan hvers virks svæðis og settir saman í De Bruijn graf. Endar klipptu lestranna eru vigtaðir út frá fjölda athugana, með viðmiðunarröðina sem grunnstöð. Eftir nokkra hreinsun og einföldun á grafinu eru allar leiðir frá upptökum til áfangastaðar dregnar út sem mögulegar setraðir. Hver setröð er samstillt við viðmiðunarerfðamengið með Smith-Waterman aðferðinni til að bera kennsl á þau tilbrigði sem hún táknar. Þessa atburðarás má bæta við með staðsetningarbyggðri greiningu. Líkurnar $P(r|H)$ á að sjá lesturinn, að því gefnu að setröðin sé raunverulegt upphafssýni, eru fyrir hvert par af lestri og setröð áætlaðar með para-földu Markov-líkani (HMM).

Með skimun eftir viðmiðunarstöðu yfir virka svæðið eru mögulegar arfgerðir myndaðar úr tvílitna samsetningum tilbrigðisatburða (SNP eða indels). Fyrir hvern atburð (þar með talið viðmið) eru skilyrt líkindi $P(r|e)$ á að sjá hvern skarandi lestur metin sem hámarkslíkindi $P(r|H)$ fyrir þær setraðir sem styðja

viðkomandi atburð. Þau eru sameinuð í skilyrtu líkindin $P(r|e1e2)$ fyrir arfgerðina (atburðapar) og síðan margfölduð til að fá skilyrtu líkindin $P(R|e1e2)$ fyrir alla lestraruppsöfnunina. Með því að nota Bayes-formúluna eru eftirálíkur $P(e1e2|R)$ fyrir hverja tvílitna arfgerð reiknaðar, og sú arfgerð sem hefur hæstu líkur er valin.

DRAGEN for IDPE Dx beitir sjálfvirkri síun. Sjá nánari upplýsingar í [Germline Workflow VCF File Annotations á síðu 12](#).

DRAGEN Somatic Variant Calling

DRAGEN-greininirinn fyrir líkamsfrumuafbrigði notar vörpuð og samræmd DNA-lesgögn sem inntak og greinir stakar kinnabreytingar (SNV) og innskot eða úrfellingar (indels) með staðbundinni *de novo* samsetningu setraða á virku svæði. Til að virkja DRAGEN Somatic Small Variant Caller skaltu velja forrit fyrir líkamleg tilbrigði.

Greining líkamsfrumusértækra erfðabreytileika er venjulega notuð fyrir æxlissýni. Í þessu vinnuflæði gerir DRAGEN engar forsendur um litningafjölda, sem gerir kleift að greina lágtíðni arfbrigði. Fyrir genasæti með allt að 100-falda þekju í æxlissýninu hefur DRAGEN greiningarþröskuld við tíðni tilbrigðissamsæta upp á 5%. Þröskuldurinn skalar með aukinni dýpt fyrir hvert genasæti og helmingast í hvert sinn sem þekjan tvöfaldast umfram 100x. Kallanleg viðmiðunarsvæði eru fyrst greind með nægilegri samstillingarþekju. Innan þessara viðmiðunarsvæða greinir skönnun á flokkuðum lestrum virk svæði, sem eru miðjuð á uppsafnaða dálka með vísbendingum um tilbrigði í æxlislestrum. Virku svæðin eru fyllt með nægu samhengi til að hylja mikilvægt efni sem er ekki viðmið. Ef vísbendingar eru um innskot eða úrfellingar fá virk svæði viðbótarfyllingu.

Samstilltir lestrar eru klipptir innan hvers virks svæðis og settir saman í De Bruijn graf. Endar klipptu lestranna eru vigtaðir út frá fjölda athugana, með viðmiðunarröðina sem grunnstoð. Eftir nokkra hreinsun og einföldun á grafinu eru allar leiðir frá upptökum til áfangastaðar dregnar út sem mögulegar setraðir. Hver setröð er samstillt við viðmiðunarerfðamengið með Smith-Waterman aðferðinni til að bera kennsl á þau tilbrigði sem hún táknar. Fyrir hvert par af lestri og setröð eru líkurnar $P(r|H)$ á að sjá lesturinn, að því gefnu að setröðin sé raunverulegt upphafssýni, áætlaðar með para-földu Markov-líkani (HMM).

Til að ákvarða greiningarmörk æxlis (TLOD) skimar DRAGEN Somatic Small Variant Caller fyrst eftir viðmiðunarstöðu fyrir hvern mögulegan æxlisbreytileika sem og viðmiðunaratburðinn yfir virka svæðið. Skilyrt líkindi $P(r|e)$ fyrir hvern skörunarlestur eru áætluð sem hámark $P(r|H)$ fyrir þær setraðir sem styðja viðkomandi atburð. Þau eru síðan sameinuð í skilyrtu líkindin $P(r|E)$ fyrir atburðartilgátuna E. Hún felur í sér blöndu af viðmiðunar- og mögulegum samsætum yfir bil samsætutíðna. Að lokum eru þau margfölduð til að fá skilyrtu líkindin $P(r|E)$, sem lýsa heildaruppsöfnun lestrarins. Síðan eru TLOD-stig reiknuð sem sönnun þess að ALT-samsæta sé til staðar í æxlissýninu í tilteknu genasæti.

DRAGEN for IDPE Dx beitir sjálfvirkri síun. Sjá nánari upplýsingar í [Somatic Workflow VCF File Annotations á síðu 14](#).

Búa til fyrirhugaða keyrslu

Notið eftirfarandi skref til að setja upp keyrslu Illumina Run Manager annað hvort á NextSeq 550Dx eða notið vafra á nettengdri tölvu. Notið vafra á nettengdri tölvu ef óskað er eftir að flytja inn sýnisgögn. Sjá nánari upplýsingar í til að fá upplýsingar um snið vísisplátanna. Illumina Run Manager for NextSeq 550Dx Software Guide (skjal nr. 200025239) til að fá leiðbeiningar um aðgang Illumina Run Manager frá nettengdri tölvu.

Það eru tvær mismunandi leiðir til að búa til nýja áætlaða keyrslu:

- **Flytja inn keyrslu**—Notið sýnisblað úr núverandi keyrslu sem sniðmát fyrir nýja keyrslu. Sjá nánari upplýsingar í Illumina Run Manager for NextSeq 550Dx Software Guide (skjal nr. 200025239) til að fá upplýsingar um hvernig á að flytja inn keyrslu.
- **Stofna keyrslu**—Sláðu inn keyrslubreytur handvirkt. Eftirfarandi leiðbeiningar lýsa því hvernig á að búa til keyrslu.

ATHUGIÐ Nauðsynlegir innsláttarreitir í notendaviðmótinu eru merktir með stjörnu (*).

Forrit

1. Á skjánum Runs (Keyrslur), á flípanum Planned (Skipulagt), veljið **Create Run** (Stofna keyrslu).
2. Veljið DRAGEN for Illumina DNA Prep with Enrichment Dx forritið og síðan **Next** (Áfram).

Keyrslustillingar

1. Á skjánum Run Settings (Keyrslustillingar) skal slá inn einkvæmt keyrsluheiti. Keyrsluheitið auðkennir keyrsluna frá raðgreiningu til greiningar.
2. **[Valfrjálst]** Sláðu inn lýsingu á keyrslunni til að auðkenna keyrsluna betur.
3. Veljið vísislyklasett sem notuð eru við undirbúning safnsins.
4. Farið yfir lesturslengdina og breyttu ef þörf krefur. Read 1 og Read 2 hafa sjálfgefið gildi upp á 151 lotur. Index 1 og Index 2 hafa fast gildi upp á 10 lotur og er ekki hægt að breyta þeim.
5. **[Valfrjálst]** Sláðu inn auðkenni safnsrörsins.
6. Veljið **Next** (Áfram).

Sample Data (Sýnisgögn)

Sýnisgögnin innihalda sýnisauðkenni, brunnsstöðu (brunnsstaða vísisplötunnar) og heiti safns. Þegar vísar A&B eru notaðir inniheldur brunnsstaða einnig auðkenni plötu.

Það eru tvær leiðir til að slá inn sýnisgögn:

- **Flytja inn sýni**—Notið sniðmátsskrá sem hægt er að hlaða niður á skjánum Sample Data (Sýnisgögn).
- **Handvirk**— Sláðu sýnishornsgögnin beint inn í töfluna á skjánum Sample Data (Sýnisgögn).

Flytja inn sýni

Hægt er að hlaða niður sniðmáti (*.csv) á skjánum Sample Data (Sýnisgögn) þegar keyrsla er skipulögð. með því að nota vafra á nettengdri tölvu. Sniðmátsskráin er ekki tiltæk til niðurrhals þegar opnað er Illumina Run Manager í gegnum NextSeq 550Dx stýrikerfishugbúnaðinn. Til að færa inn sýnisgögn með því að nota eiginleikann Import Samples skal fylgja eftirfarandi skrefum.

ATHUGIÐ Ljúktu við skrefin í keyrslustillingum áður en þú heldur áfram.

1. Veljið **Download Template** til að hlaða niður tómri CSV-skrá.
2. Úr sniðmátsskránni skal slá inn sýnisgögn og vista síðan skrána. Heiti safns er valkvætt.

ATHUGIÐ Þegar vísar A&B eru notaðir verða gögnin fyrir dálk B að innihalda bæði staðsetningu plötunnar og brunnsins (vísir fyrir staðsetningu plötu og brunns). Dæmi: A-A01, A-A02, A-A03.

3. Veljið **Import Samples** (Flytja inn sýni) og flettið að sniðmátsskránni sem inniheldur upplýsingar um sýnishornsgögnin úr fyrra skrefi.
4. Veljið **Open, Proceed**, og síðan **Next** (Áfram).

ATHUGIÐ Að breyta sýnisauðkenninu áður en Next (Áfram) er valið, getur leitt til villu. Ljúktu við að setja upp keyrsluna áður en þú gerir breytingar, annars geta komið upp villur.

Sláið inn sýni handvirk

Notið töfluna á skjánum Sample Data (Sýnisgögn) til að slá inn sýnisupplýsingar handvirk.

1. Sláðu inn einkvæmt sýnisauðkenni í flípanum Sample ID.
2. Notið **Well Position** (vísir A eða vísir B) eða **Plate - Well Position** (vísir A&B) til að velja viðeigandi vísi fyrir sýnin.
Reitirnir i7 Index, Index 1, i5 Index og Index 2 fyllast út sjálfkrafa.
3. **[Valfrjálst]** Sláðu inn heiti safns.
4. Bætið við röðum og endurtakið skref 1–3 eftir þörfum þar til öllum sýnum hafa verið bætt við töfluna. Þú getur bætt við mörgum röðum í einu með því að slá fyrst inn fjölda raða sem á að bæta við og velja síðan + táknið. Þú getur líka fjarlægt raðir með því að velja reitinn við hliðina á raðnúmerinu og smella síðan á rusltáknið.
5. Veljið **Next** (Áfram).

Greiningarstillingar

1. Veljið æskilegt greiningarvinnuflæði:
 - FASTQ-myndun
 - FASTQ- og VCF-myndun fyrir kímlínuvinnuflæði (upplýsingaskrá nauðsynleg)
 - FASTQ- og VCF-myndun fyrir líkamlegt vinnuflæði (upplýsingaskrá nauðsynleg)
2. **[Valfrjálst] Generate ORA compressed FASTQs** er virkjað sjálfkrafa. FASTQ ORA tapslaus þjöppun getur minnkað FASTQ-skrár allt að fimmfalt miðað við fastq.gz. Afveljið **Generate ORA compressed FASTQs** ef óþjöppuð gögn (fastq.gz) eru æskileg.
3. Fyrir kímlínu- og líkamlegt vinnuflæði er upplýsingaskrá nauðsynleg. Notið **Manifest File Selection** (Val á yfirlitsskrá) fellivalmyndina til að búa til yfirlitsskrá. Upplýsingaskráin er BED-skrá (*.bed) með flipaskilgreiningu sem tilgreinir nöfn og staðsetningar viðmiðunarmarksvæða. Fyrir nánari upplýsingar, sjá [Upplýsingaskrá á síðu 7](#).
4. **[Valfrjálst]** Fyrir líkamlegt vinnuflæði skal nota fellivalmyndina **Noise File Selection** til að velja kerfisbundna hávaðaskrá. Hægt er að tilgreina BED-skrá (*.bed.gz) með staðarsértæku hávaðastigi til að sía út kerfisbundinn hávaða. Fyrir nánari upplýsingar, sjá [Hávaðasíun \(valfrjálst\) á síðu 8](#).
5. Veljið **Next** (Áfram).

Run (Keyrsla) Review (Skoða)

1. Á skjánum Review (Skoða) skal fara yfir upplýsingarnar fyrir Run Settings (Keyrslustillingar), Sample Data (Sýnisgögn) og Analysis Settings (Greiningarstillingar).
2. Veljið **Save** (Vista).
Keyrslan er vistuð á flípanum Planned (Skipulagt) á skjánum Runs (Keyrslur).

Stillingar

Til að skoða eða breyta DRAGEN for IDPE Dx forritastillingum skaltu fyrst velja Applications (Forrit)-táknið á aðalskjánum. Veldu síðan forritið sem þú vilt skoða eða breyta. Stjórnandareikningur er nauðsynlegur til að breyta stillingum.

Samskipan

Stillingaskjárninn sýnir eftirfarandi stillingar forritsins:

- **Library Prep Kits** (Tæki fyrir undirbúning safns) — Sýnir sjálfgefið undirbúningssett safns fyrir forritið. Þessari stillingu er ekki hægt að breyta.
- **Index Adapter Kits** (Vísilyklasett) — Sýnir sjálfgefið vísilyklasett fyrir forritið. Þessari stillingu er ekki hægt að breyta.

- **Read lengths** (Lestrarlengdir) — Lestrarlengdir eru sjálfgefið stilltar á 151 fyrir forritið en hægt er að breyta þeim við stofnun keyrslu.
- **Manifest and Noise Files** (Upplýsinga- og hávaðaskrár) — Hlaða upp og breyta stillingum fyrir upplýsinga- og hávaðaskrár.
 - Veljið **Upload File** (Hlaða upp skrá) til að hlaða upp skrá til notkunar í greiningu.
 - Veljið **Default** (Sjálfgefið) útvarpshnappinn til að stilla skrána sem sjálfgefna upplýsingaskrá eða hávaðaskrá sem valin er við stofnun keyrslu þegar forritið er valið.
 - Veljið gátreitinn **Enabled** (Virkjuð) til að stilla skrána þannig að hún birtist í fellivalmyndinni við stofnun keyrslu.

Heimildir

Notið gátreitina á Permissions (Heimildir)-skjánum til að stjórna aðgangi notenda að forritinu.

Upplýsingaskrá

Þegar notað er DRAGEN for IDPE Dx er upplýsingaskrá nauðsynleg fyrir eftirfarandi vinnuflæði:

- FASTQ og VCF myndun fyrir kímlínuvinnuflæði
- FASTQ og VCF myndun fyrir líkamlegt vinnuflæði

Upplýsingaskráin er BED-skrá (*.bed) með flípaskilgreiningu sem tilgreinir heiti og staðsetningar viðmiðunarmarksvæða. Aðalhluti upplýsingaskrárinnar er Regions-hlutinn og ætti að innihalda eftirfarandi gagnadálka:

Dálkur	Lýsing
Heiti	Einkvæmt notandatilgreint heiti fyrir markið
Litningur	Staðsetning litninga (t.d. chr10, chr5, o.s.frv.)
Start	1-basa vísir fyrir upphafsstöðu marksins
Stop	1-basa vísir fyrir stöðvunarstöðu marksins
Upstream Probe Length	Lengd uppstreymis þreifarans. Fyrir DRAGEN for IDPE Dx forritið ætti þetta að vera stillt á 0.
Lengd uppstreymisprófs	Lengd niðurstreymis þreifarans. Fyrir DRAGEN for IDPE Dx forritið ætti þetta að vera stillt á 0.

ATHUGIÐ Gilt snið fyrir upplýsingaskrá er nauðsynlegt fyrir greiningu. DRAGEN mun stöðva greiningu ef upplýsingaskráin er ógild.

Hávaðasíun (valfrjálst)

Kerfisbundin hávaðasía er tiltæk fyrir köllun tilbrigða í líkamsfrumum og má nota til að draga úr fölskum jákvæðum köllum með því að taka tillit til staðbundins hávaða. Kerfisbundna hávaðaskráin er búin til með því að safna um það bil 50 eðlilegum sýnum (helst sértækum fyrir spjaldið, undirbúning safnsins og raðgreiningartækið). Síðan er samtala samsætutíðni undir 30% á hverjum stað með nægilegri þekju deilt með heildarfjölda sýna, (Gert er ráð fyrir að samsætutíðni yfir 30% endurspegli kímlínutilbrigði en ekki hávaða). Þegar hávaðagildin hafa verið búin til verða líkamleg tilbrigði sem greinast á þeim stað síuð.

Síuna má nota í venjulegri æxlisstillingu, en hún er sérstaklega gagnleg fyrir æxliskeyrslur þegar samsvarandi eðlileg gildi eru ekki tiltæk. Kerfisbundna hávaðaskráin verður að nota BED-skrá sem hefur skráarendinguna (*.bed.gz) og verður að innihalda fjóra dálka: Litningur, upphafs-, endi- og staðarbundin hávaðastig fyrir hverja röð. Kerfisbundin hávaðasíun er valfrjáls.

Greiningarúttak

Keyrslur sem eru í gangi eru birtar í flípanum Active (Virkar). Loknar keyrslur birtast í „Completed“-flípanum (Lokið). DRAGEN for IDPE Dx býr til greiningarmöppu með einkvæmu nafni fyrir hverja greiningu, sem er aðskilin frá möppunni sem inniheldur raðgreiningargögnin. Greiningarmappan hefur að geyma eftirfarandi upplýsingar:

- Upplýsingaskrá notuð
- Hugbúnaðarútgáfa
- Sýnisauðkenni
- Heildarfjöldi samræmdra lestra
- Hlutfall samstilltra lestra á hvert sýni
- Fjöldi kallaðra SNV á hvert sýni
- Fjöldi kallaðra innskota og úrfellinga á hvert sýni
- Þekjutölfræði

Úttaksskrár greiningar

Staðsetning greiningarmöppunnar er tilgreind með stillingunni External Storage for Analysis Results (Ytri geymsla fyrir greiningarniðurstöður). Sjá nánari upplýsingar í Illumina Run Manager for NextSeq 550Dx Software Guide (skjal nr. 200025239) fyrir frekari upplýsingar um stillinguna External Storage for Analysis Results (Ytri geymsla fyrir greiningarniðurstöður).

Á skjánum Run Details (Keyrsluupplýsingar) sýnir reiturinn External Location (Ytri staðsetning) slóðina fyrir raðgreiningargögnin. Einkvæmt heiti greiningarmöppunnar er gefið upp í reitnum Analysis Output Folder (Greiningarúttaksmappa) á skjánum Run Details (Keyrsluupplýsingar). Nákvæmlega hvaða skrár eru búnar til fer eftir því hvaða greiningarvinnuflæði er notað. Eftirfarandi úttaksskrár úr greiningu eru myndaðar af forritinu.

ATHUGIÐ Ef villa kemur upp varðandi takmörkun á hámarkslengd skráarslóðar þegar reynt er að opna greiningarúttaksskrár skaltu reyna að færa skrána á styttri slóð eða nota aðra aðferð til að opna skrána.

Úttaksskrá	Lýsing
Yfirlitsskýrsla um tilbrigði (*.pdf)	Inniheldur samantekt á skráarupplýsingum, hugbúnaðarútgáfum, upplýsingum um sýni, tölfræði um lestrarstig og SNV, innskot, úrfellingar og samantektir á þekju. Aðeins kímlínu- og líkamsvinnuflæði búa til tilbrigðisskýrslu.
FASTQ (*.fastq.gz eða *.fastq.ora)	Milliskrár sem innihalda basaköll með gæðaskori. FASTQ-skrár eru aðalinntakið fyrir samstillingarskrefið. Þegar ORA -þjöppun er valin er nafnaukinn *.fastq.ora notaður.
Samstillingar-BAM skrár (*.bam)	Inniheldur samstillta lestra fyrir tiltekið sýni.
Genome VCF-skrár (*.gvcf.gz)	Inniheldur arfgerðina fyrir hverja stöðu, hvort sem hún er kölluð sem tilbrigði eða viðmið.
VCF-skrár (*.vcf.gz)	Inniheldur kölluð tilbrigði á hverri stöðu.
Skýrsla um mæligildi keyrslu (*.csv)	Inniheldur gæðamælikvarða um keyrsluna, þar á meðal heildarafrakstur án vísa og Q30 skor.

FASTQ-skrár

FASTQ (*.fastq.gz, *.fastq.ora) er textabundið skráarsnið sem inniheldur basaköll og gæðagildi fyrir hvern lestur. Skráin inniheldur eftirfarandi upplýsingar.

- Auðkenni sýnisins
- Röðin
- Gæðastig Phred í ASCII + 33 kóðuðu sniði

Auðkenni sýnis er sniðið á eftirfarandi hátt.

```
@Instrument:Run ID:FlowCell ID:Lane:Tile:X:Y
ReadNum:FilterFlag:0:SampleNumber
Dæmi:
@SIM:1:FCX:1:15:6329:1045 1:N:0:2
TCGCACTCAACGCCCTGCATATGACAAGACAGAATC
+
<>;##=><9=AAAAAAAAAA9#:<#<;<<<????#=#
```

BAM-skrár

BAM-skrá (*.bam) er þjöppuð tvíundarskrá af SAM-skrá (sequence alignment map) sem er notuð til að tákna samstilltar raðir allt að 128 Mb. BAM skrár nota skráarnafnssniðið `SampleName_S#.bam.` # er sýnisnúmerið sem ákvarðast af þeirri röð sem sýnin eru skráð í fyrir keyrsluna. Í fjölhnútastillingu er S# stillt á S1, óháð röð sýnisins.

BAM-skrár innihalda hausahluta og samstillingarhluta:

- **Síðuhaus**—Inniheldur upplýsingar um alla skrána, svo sem nafn sýnis, lengd sýnis og samstillingaraðferð. Samstillingar í samstillingarhlutanum eru tengdar ákveðnum upplýsingum í hausahlutanum.
- **Samstillingar**—Inniheldur lestursheiti, lestursröð, lestursgæði, upplýsingar um samstillingu og sérsniðin merki. Lestrarheitið inniheldur litninginn, upphafshnit, gæði samstillingar og lýsingarstreng samsvörunar.

Samstillingarhlutinn inniheldur eftirfarandi upplýsingar fyrir hvern lestur eða lestrapar:

- AS: Gæði paraðra enda samstillinga.
- RG: Lesturshópur, sem gefur til kynna fjölda lestra fyrir tiltekið sýni.
- BC: Strikamerki sem gefur til kynna affléttuð sýnisauðkenni sem tengist lestrinum.
- SM: Gæði samstillingar með einum enda.
- XC: Lýsingarstrengur samsvörunar.
- XN: Nafnmerki fyrir amplicon, sem skráir auðkenni ampliconsins sem tengist lestrinum

BAM-vísisskráin (*.bam.bai) er vísir fyrir samsvarandi BAM-skrá.

VCF-skrár

Skrár á tilbrigðakallssniði (*.vcf) innihalda upplýsingar um tilbrigði sem finnast á tilteknum stöðum í viðmiðunarerfðamengi.

Síðuhaus VCF-skráarinnar inniheldur útgáfu VCF-skráarsniðsins, útgáfu tilbrigðakallarans og lista yfir skýringar sem notaðar eru í afgangnum af skránni. Síðuhaus VCF-skráarinnar inniheldur einnig viðmiðunarerfðamengisskrána og BAM-skrána. Síðasta línan í síðuhausunum inniheldur dálkafyrirsagnir fyrir gagnalínurnar. Hver gagnalína VCF-skráarinnar inniheldur upplýsingar um eitt tilbrigði.

Tafla 1 Síðuhausar VCF-skráa

Yfirskrift	Lýsing
CHROM	Litningur viðmiðunarerfðamengisins. Litningar birtast í sömu röð og í viðmiðunar FASTA-skránni.

Yfyrskrift	Lýsing
POS	Einbasa staðsetning tilbrigðisins í viðmiðunarlitningnum. Fyrir einkirnitilbrigði (single nucleotide variants, SNV) er þessi staðsetning viðmiðunarbásinn með tilbrigðinu. Fyrir innskot og úrfellingar (indels) er þessi staðsetning viðmiðunarbásinn rétt á undan tilbrigðinu.
ID	rs (viðmiðunar-SNP)-númerið fyrir SNP sem fengið er úr <code>dbSNP.txt</code> , ef við á. Ef það eru fleiri en eitt rs-númer á þessum stað er listinn aðskilinn með semíkómmu. Ef engin dbSNP-færsla er til staðar á þessum stað er notað merki fyrir gildi sem vantar ('.').
REF	Viðmiðunarfgerð. Til dæmis er úrfelling á einu T táknuð þannig að viðmiðunarrunan sé TT og vararunan T. Tilbrigði á einu kirni frá A í T er táknað sem viðmiðunarkirnið A og varakirnið T.
ALT	Samsæturnar sem eru frábrugðnar viðmiðunarlestrinum. Til dæmis er innskot á einu T táknað þannig að viðmiðunarrunan er A og vararunan AT. Tilbrigði af stöku kirni frá A til T er táknað sem viðmiðunar A og vara T.
QUAL	Gæðastig sem er úthlutað af afbrigðakallaranum, út frá Phred-kvarða. Hærri stig gefa til kynna meiri öryggi í tilbrigðinu og minni líkur á villum. Fyrir gæðastig Q eru áætlaðar líkur á villu 10 ^{-Q} (Q/10). Til dæmis hefur mengi Q30-kallanna 0,1% villutíðni. Margir tilbrigðakallarar úthluta gæðastigum á grundvelli tölfræðilegra líkana sinna, og eru þau oft há miðað við þá villutíðni sem sést.

Tafla 2 Germline Workflow VCF File Annotations

Yfyrskrift	Lýsing
FILTER	Ef allar síur eru taldar gildar, þá er skrifað PASS í síudálkinn. Mögulegar FILTER-færslur eru meðal annars: • DRAGENsnpHardQUAL—Notað ef QUAL-stig SNP tilbrigðisins nær ekki þröskuldi • DRAGENindelHardQUAL—Notað ef QUAL-stig indel-tilbrigðisins nær ekki þröskuldi • LowDepth—Staður síaður út vegna þess að þekjudýpt nær ekki tilskildum þröskuldi • LowGQ—Staður síaður út vegna þess að þekjudýpt nær ekki tilskildum þröskuldi • PloidyConflict—Arfgerðarkall frá tilbrigðakallara er ekki í samræmi við litnun litninga • base_quality—Staður var síaður þar sem miðgildi grunngæða í alt-lestrum á þessum stað uppfyllti ekki viðmiðunarmörk • filtered_reads—Staður síaður út vegna þess að of stór hluti lestra hefur verið síaður út • fragment_length—Staður síaður út vegna þess að alger munur á miðgildi brotlengdar alt lestra og miðgildi brotlengdar ref lestra í þessu genasæti fer yfir þröskuld. • low_depth—Staður síaður út vegna þess að lestrardýptin er of lítil • low_frac_info_reads—Staður síaður út vegna þess að hlutfall upplýsandi lestra er undir þröskuldi • low_normal_depth—Staður síaður út vegna þess að lestrardýpt eðlilegra sýna er of lítil • long_indel—Staður síaður út vegna þess að indel-lengd er of mikil • mapping_quality—Staður síaður út vegna þess að miðgildi vörpunargæða alt lestra í þessu genasæti nær ekki þröskuldi • multiallelic—Staður síaður út vegna þess að fleiri en tvær alt samsætur komast í gegnum æxlis-LOD • non_homref_normal—Staður síaður út vegna þess að arfgerð eðlilegs sýnis er ekki arfhreint viðmið • no_reliable_supporting_read—Þessi genastaður var síaður út þar sem enginn áreiðanlegur stuðningslestur fyrir lífsfrumusértækan breytileika fannst. • panel_of_normals—Sést í að minnsta kosti einu sýni í spjaldi eðlilegra vcf • read_position—Staður síaður út vegna þess að miðgildi fjarlægð milli upphafs/loka lestars og þessa genasætis er undir þröskuldi • RMxNRepeatRegion—Staður síaður út vegna þess að öll eða hluti tilbrigðissamsætunnar er endurtekning á viðmiðinu. • strand_artifact—Staður síaður út vegna mikillar skekkju í práðum • str_contraction—Set síað vegna gruns um PCR-villu þar sem alt samsætan er einni endurtekningareiningu minni en viðmiðunarsamsætan • too_few_supporting_reads—Staður síaður út vegna þess að of fáir stuðningslestrar eru í æxlissýninu. • weak_evidence—Stig líkamstilbrigðis ná ekki þröskuldi

Yfyrskrift	Lýsing
INFO	Mögulegar INFO-færslur eru meðal annars: • AC—Fjöldi samsæta í arfgerðum fyrir hverja ALT samsætu, í sömu röð og talið er upp. • AF—Tíðni samsæta fyrir hverja ALT samsætu, í sömu röð og talið er upp. • AN—Heildarfjöldi samsæta í kölluðum arfgerðum. • DB—dbSNP aðild. • FS—Phred-kvarðað p-gildi með því að nota nákvæmnispróf Fishers til að greina þráðskekku. • QD—Tilbrigðisöryggi/gæði eftir dýpt. • R2_5P_bias—Stig byggð á pörunarskekku og fjarlægð frá 5. aðalenda. • SOR—Samhverft gagnlíkindahlutfall 2x2 tengslatöflu til að greina skekkju í þráðum. • DP—Áætluð lestrardýpt (upplýsandi og óupplýsandi); ákveðnir lestrar kunna að hafa verið fjarlægðir vegna MAPQ-gilda o.fl. • END—Stöðvunarstaða bilsins. • FractionInformativeReads—Hlutfall upplýsandi lestra af heildarfjölda lestra. • MQ—RMS vörpunargæði. • MQRankSum—Z-gildi úr Wilcoxon-prófi á summu sætistalna fyrir eiginleika Alt, borið saman við Ref-lestrarvörpun. • ReadPosRankSum—Z-stig úr Wilcoxon-prófi á sætisröðun Alt- og Ref-lestra til að meta skekkju í lestrarstöðu. • SOMATIC—Að minnsta kosti eitt tilbrigði á þessari stöðu er líkamlegt.
SNID	Í sniðsdálknum eru reitir taldir upp og aðskildir með tvípunktum. Til dæmis, GT:GQ. Tiltækir reitir eru meðal annars: • AD—Samsæt dýpt (aðeins upplýsandi lestrar taldir af heildarfjölda lestra) fyrir ref og alt samsætur í þeirri röð sem þær eru taldar upp. • AF—Samsæt brot fyrir allar samsætur í þeirri röð sem þær eru taldar upp. • DP—Áætluð lestrardýpt (lestur með MQ=255 eða með slæmum pörunum er síaður út). • F1R2—Fjöldi lestra í F1R2 parastefnu sem styður hverja samsætu. • F1R2—Fjöldi lestra í F2R1 parastefnu sem styður hverja samsætu. • GT—Arfgerð. 0 samsvarar viðmiðunarbasanum, 1 samsvarar fyrstu færslunni í ALT-dálknum og svo framvegis. Skástrikið (/) gefur til kynna að engar upplýsingar um fösun eru tiltækar. • MB—Tölfræði fyrir hvert sýni til að greina pörunarskekku. • PS—Upplýsingar um fösunarauðkenni, þar sem hvert einkvæmt auðkenni innan tiltekins sýnis (en ekki á milli sýna) tengir saman færslur innan fösunarhóps. • SB—Tölfræði um búta á hvert sýni, sem felur í sér nákvæmnispróf Fishers til að greina skekkju í þráðum. • SQ—Líkamleg gæði.
SÝNI	Sample-dálkurinn sýnir þau gildi sem tilgreind eru í FORMAT-dálkinum.

Tafla 3 Somatic Workflow VCF File Annotations

Yfyrskrift	Lýsing
FILTER	Ef allar síur eru taldar gildar, þá er skrifað PASS í síudálkinn. Mögulegar FILTER-færslur eru meðal annars: • base_quality—Staður var síaður þar sem miðgildi grunngæða í alt-lestrum á þessum stað uppfyllti ekki viðmiðunarmörk • filtered_reads—Staður síaður út vegna þess að of stór hluti lestra hefur verið síaður út • fragment_length—Staður síaður út vegna þess að alger munur á miðgildi brotlengdar alt lestra og miðgildi brotlengdar ref lestra í þessu genasæti fer yfir þröskuld. • low_depth—Staður síaður út vegna þess að lestrardýptin er of lítil • low_frac_info_reads—Staður síaður út vegna þess að hlutfall upplýsandi lestra er undir þröskuldi • low_normal_depth—Staður síaður út vegna þess að lestrardýpt eðlilegra sýna er of lítil • long_indel—Staður síaður út vegna þess að indel-lengd er of mikil • mapping_quality—Staður síaður út vegna þess að miðgildi vörpunargæða alt lestra í þessu genasæti nær ekki þröskuldi • multiallelic—Staður síaður út vegna þess að fleiri en tvær alt samsætur komast í gegnum æxlis-LOD • non_homref_normal—Staður síaður út vegna þess að arfgerð eðlilegs sýnis er ekki arfhreint viðmið • no_reliable_supporting_read—Þessi genastaður var síaður út þar sem enginn áreiðanlegur stuðningslestur fyrir lífsfrumusértækan breytileika fannst. • panel_of_normals—Sést í að minnsta kosti einu sýni í spjaldi eðlilegra vcf • read_position—Staður síaður út vegna þess að miðgildi fjarlægða milli upphafs/loka lestrurs og þessa genasætis er undir þröskuldi • RMxNRepeatRegion—Staður síaður út vegna þess að öll eða hluti tilbrigðissamsætunnar er endurtekning á viðmiðinu. • strand_artifact—Staður síaður út vegna mikillar skekkju í þráðum • str_contraction—Set síað vegna gruns um PCR-villu þar sem alt samsætan er einni endurtekningareiningu minni en viðmiðunarsamsætan • too_few_supporting_reads—Staður síaður út vegna þess að of fáir stuðningslestrar eru í æxlissýninu. • weak_evidence—Stig líkamstilbrigðis ná ekki þröskuldi • systematic_noise—Set síað eftir vísbendingum um kerfisbundinn hávaða í eðlilegum gildum

Yfyrskrift	Lýsing
INFO	Mögulegar INFO-færslur eru meðal annars: • DP—Áætluð lestrardýpt (upplýsandi og óupplýsandi); ákveðnir lestrar kunna að hafa verið fjarlægðir vegna MAPQ-gilda o.fl. • END—Stöðvunarstaða bilsins. • FractionInformativeReads—Hlutfall upplýsandi lestra af heildarfjölda lestra. • MQ—RMS vörpunargæði. • MQRankSum—Z-gildi úr Wilcoxon-prófi á summu sætistalna fyrir eiginleika Alt, borið saman við Ref-lestrarvörpun. • ReadPosRankSum—Z-stig úr Wilcoxon-prófi á sætisröðun Alt- og Ref-lestra til að meta skekkju í lestrarstöðu. • AQ—Kerfisbundin hávaðastig. • hotspot—Þekkt líkamsset, notað til að auka öryggi í köllum. • SOMATIC—Að minnsta kosti eitt tilbrigði á þessari stöðu er líkamlegt.
SNID	Í sniðsdálknum eru reitir taldir upp og aðskildir með tvípunktum. Til dæmis, GT:GQ. Tiltækir reitir eru meðal annars: • AD—Samsæt dýpt (aðeins upplýsandi lestrar taldir af heildarfjölda lestra) fyrir ref og alt samsætur í þeirri röð sem þær eru taldar upp. • AF—Samsæt brot fyrir allar samsætur í þeirri röð sem þær eru taldar upp. • DP—Áætluð lestrardýpt (lestur með MQ=255 eða með slæmum pörunum er síður út). • F1R2—Fjöldi lestra í F1R2 parastefnu sem styður hverja samsætu. • F2R1—Fjöldi lestra í F2R1 parastefnu sem styður hverja samsætu. • GP—Phred-kvörðuð eftirálíkindi fyrir arfgerðir, eins og þau eru skilgreind í VCF-forskriftinni. • GQ—Gæði arfgerðar. • GT—Arfgerð. 0 samsvarar viðmiðunarbaskanum, 1 samsvarar fyrstu færslunni í ALT-dálknum og svo framvegis. Skástrikið (/) gefur til kynna að engar upplýsingar um fösun eru tiltækar. • MB—Tölfræði fyrir hvert sýni til að greina pörunarskekkju. • PL—Staðlaðar, Phred-kvarðaðar líkur á arfgerðum eins og þær eru skilgreindar í VCF-forskriftinni. • PRI—Phred-kvörðuð fyrirframlíkindi fyrir arfgerðir. • PS—Upplýsingar um fösunarauðkenni, þar sem hvert einkvæmt auðkenni innan tiltekins sýnis (en ekki á milli sýna) tengir saman færslur innan fösunarhóps. • SB—Tölfræði um búta á hvert sýni, sem felur í sér nákvæmnispróf Fishers til að greina skekkju í þráðum. • SQ—Líkamleg gæði.
SÝNI	Sample-dálkurinn sýnir þau gildi sem tilgreind eru í FORMAT-dálkinum.

Genome VCF-skrár

Genome VCF (.gvcf.gz) skrár fylgja ákveðnum venjum til að tákna öll sæti í erfðamenginu á tiltölulega samþjöppuðu sniði. gVCF-skrárnar innihalda öll set innan svæðisins sem um ræðir í einni skrá fyrir hvert sýni. gVCF-skráin sýnir engin köll (no-calls) á stöðum sem fara ekki í gegnum allar síur. Arfgerðarmerkingin (GT) ./ gefur til kynna ekkert kall.

Endurtaka greiningu

Þú getur endurtekið greininguna ef henni var hætt, ef hún mistókst eða ef þú vilt keyra hana aftur með öðrum stillingum. Til að endurtaka greiningu skaltu framkvæma eftirfarandi skref:

1. Á Run (Keyrsla) skjánum skaltu velja flipann Completed (Lokið) og síðan velja heiti keyrslunnar til að endurgreina.
Ef Requeue Analysis hefur verið keyrt áður, veldu heiti upphaflegrar keyrslu (Parent Run).
2. Á skjánum Run Details (Keyrsluupplýsingar), á eftir Sequencing Information (Raðgreiningarupplýsingar), veldu **Requeue Analysis**.
3. Veljið valkost:
 - Endurtaka greiningu án breytinga
 - Breyta keyrslustillingum og endurtaka greiningu
 - Endurtaka greiningu með öðru forriti
4. Staðfestið að staðsetningin þar sem raðgreiningargögnin eru nú geymd sé tilgreind í reitnum **Sequencing data file path** (Slóð að raðgreiningargögnum).

ATHUGIÐ Slóðin að raðgreiningargögnunum ætti að passa við slóðina í stillingunni External Storage for Analysis Results (Ytri geymsla fyrir greiningarniðurstöður). Sjá nánari upplýsingar í Illumina Run Manager for NextSeq 550Dx Software Guide (skjal nr. 200025239) til að fá leiðbeiningar um hvernig breyta má slóð ytri geymslu.

5. Sláið inn ástæðu fyrir endurgreiningu.
6. Veljið **Endurtaka greiningu**.
7. Gerðu þær breytingar sem þú vilt á Run Settings (Keyrslustillingar), Sample Data (Sýnisgögn) og Analysis Settings (Greiningarstillingar).
8. Veljið **Save** (Vista). Greining hefst með því að nota núverandi greiningarbreytur.

Tæknileg aðstoð

Fyrir tæknilega aðstoð hafið samband við Illumina tæknilega aðstoð.

Vefsíða: www.illumina.com

Netfang: techsupport@illumina.com

Öryggisupplýsingum (SDS)— Fáanleg á vefsíðu Illumina á support.illumina.com/sds.html.

Vöruupplýsingar— Hægt að sækja á support.illumina.com.



Illumina
5200 Illumina Way
San Diego, California 92122 U.S.A.
+1.800.809.ILMN (4566)
+1.858.202.4566 (utan Norður-Ameríku)
techsupport@illumina.com
www.illumina.com



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

Ástralskur bakhjarl

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

FYRIR IN VITRO GREININGARNOTKUN.

© 2023 Illumina, Inc. Allur réttur áskilinn.

illumina®