

illumina®

# DRAGEN for Illumina DNA Prep with Enrichment Dx

NovaSeq 6000Dx için Ürün Belgeleri

ILLUMINA ŞİRKETİNE ÖZEL

Belge No 200014776 v02

Eylül 2022

İN VİTRO TANI AMAÇLI KULLANIM İÇİNDİR

## Revizyon Gemiři

Belge	Tarih	Deęiřiklik Açıklaması
200014776 v02	Eylül 2022	Çalıřma oluřturma talimatında, manifest dosyası için format, metinden (*.txt) BED (*.bed) olarak düzeltilmiřtir. Analiz çıktıısı bölümünde VCF dosyaları VCF dosyaları olarak düzeltildi.
200014776 v01	Aęustos 2022	Eklenen: Ayarlar bölümü. Sistemik Gürültü Filtreleme bölümü. Çalıřma oluřturma talimatı daha fazla ayrıntı içerecek řekilde güncellendi. Yazım hatası ve dilbilgisi hataları düzeltildi. Talimatın NovaSeq 6000Dx Aletiyle kullanıldıęında uygulama için amaçlandıęı belirtildi. VCF çıktı dosyası içerięi hakkındaki bilgiler güncellendi.
200014776 v00	Mart 2022	İlk sürüm.

Bu belge ve içindekiler Illumina, Inc. ve bağılı şirketlerinin ("Illumina") mülkiyetinde olup yalnızca işbu belgede açıklanan ürünün/ürünlerin kullanımıyla bağlantılı olarak müşterisinin sözleşmeye ilişkin kullanımı içindir. Bu belge ve içindekiler Illumina'nın önceden yazılı izni olmaksızın başka hiçbir amaçla kullanılamaz veya dağıtılamaz ve/veya hiçbir şekilde iletilemez, ifşa edilemez ya da kopyalanamaz. Illumina bu belge ile patenti, ticari markası, telif hakkı veya genel hukuk hakları ya da üçüncü tarafların benzer hakları kapsamında hiçbir lisansı devretmez.

Bu belgede açıklanan ürünün/ürünlerin uygun ve güvenli bir şekilde kullanılması için nitelikli ve uygun eğitim almış çalışanlar bu belgedeki talimatları tam olarak ve açık bir şekilde uygulamalıdır. Söz konusu ürün/ürünler kullanılmadan önce bu belgedeki tüm bilgiler tam olarak okunmalı ve anlaşılmalıdır.

**BU BELGEDE YER ALAN TÜM TALİMATLARIN TAMAMEN OKUNMAMASI VE AÇIK BİR ŞEKİLDE UYGULANMAMASI, ÜRÜNÜN/ÜRÜNLERİN HASAR GÖRMESİNE, KULLANICI VEYA BAŞKALARI DAHİL OLMAK ÜZERE KİŞİLERİN YARALANMASINA VE DİĞER MALLARIN ZARAR GÖRMESİNE NEDEN OLABİLİR VE ÜRÜN/ÜRÜNLER İÇİN GEÇERLİ OLAN HER TÜRLÜ GARANTİYİ GEÇERSİZ KILACAKTIR.**

**ILLUMINA BU BELGEDE AÇIKLANAN ÜRÜNÜN/ÜRÜNLERİN (ÜRÜNÜN PARÇALARI VE YAZILIMI DAHİL) YANLIŞ KULLANIMINDAN DOĞAN DURUMLARDAN SORUMLU TUTULAMAZ.**

© 2022 Illumina, Inc. Tüm hakları saklıdır.

Tüm ticari markalar Illumina, Inc. veya ilgili sahiplerinin malıdır. Özel ticari marka bilgileri için bkz. [www.illumina.com/company/legal.html](http://www.illumina.com/company/legal.html).

# İçindekiler

Revizyon Geçmişi .....	ii
Genel Bakış .....	1
Analiz Yöntemleri .....	1
Create a Run (Bir Çalıştırma Oluştur) .....	4
Ayarlar .....	5
Analiz Çıktıları .....	7
FASTQ Dosyaları .....	8
BAM Dosyaları .....	8
VCF Dosyaları .....	9
Analiz Sonuçlarını Görüntüleme .....	14
<b>Teknik Yardım .....</b>	<b>15</b>

# Genel Bakış

Illumina® DNA Prep with Enrichment Dx Uygulaması için DRAGEN™, seçilen analiz iş akışına bağlı olarak çoğullama çözme, FASTQ oluşturma, okuma eşleme ve referans genomla hizalama ve varyant araması gerçekleştirir.

## Analiz Yöntemleri

DRAGEN for Illumina DNA Prep with Enrichment Dx seçilen iş akışına bağlı olarak çoğullama çözme, FASTQ oluşturma, okuma eşleme ve bir referans genom ile hizalama gerçekleştirir:

- FASTQ oluşturma
- Germ Hattı FASTQ ve VCF oluşturma
- Somatic FASTQ ve VCF oluşturma

## FASTQ Oluşturma

Birleştirilen sekanslar, numune başına FASTQ dosyalarına yazılır. FASTQ dosyaları, yalnızca bir numune için sıralama verilerini ve kalite skorlarını içeren metin dosyalarıdır. Her numune için, okunan dizileme uyarınca akış hücre yolu başına ayrı FASTQ dosyaları oluşturulur. Çalışma kurulumu sırasında belirtilen şekilde numunenin adı FASTQ dosya adına dahil edilir. FASTQ dosyaları hizalama için birincil girdidir. FASTQ oluşturma ilk adımı çoğullama çözmedir. Çoğullama çözme, her bir dizin okuma sekansını çalışma için belirtilen dizin sekanslarıyla karşılaştırarak bir numuneye filtreden geçen kümeler atar. Bu adım için herhangi bir kalite değeri dikkate alınmaz. Dizin okumaları aşağıdaki adımlar kullanılarak belirlenir:

- Numuneler çalıştırma için listelendikleri sıralamaya göre 1'den başlayarak numaralandırılır.
- Numune numarası 0, bir numuneye atanmamış kümeler için ayrılmıştır.
- Kümeler, dizin sekansı tam olarak eşleştiğinde veya dizin okuması başına tek bir uyumsuzluk olduğunda bir numuneye atanır.

Yazılım, FASTQ dosyalarını sıkıştırmak için ORA sıkıştırması içerir. ORA (\*.ora) biçimini kullanırken, kayıpsız bir sıkıştırma sağlamak için FASTQ içeriğinin md5 sağlama toplamı sıkıştırma ve dekompresyon döngüsünden sonra korunur.

## DNA Eşleme ve Hizalama

Eşlemenin ilk aşaması, okumadan tohumlar oluşturmak ve ardından referans genomda tam eşleşmeleri aramaktır. Daha sonra bu sonuçlar, en yüksek yoğunlukta tohum eşleşmelerine sahip konumlarda tam Smith-Waterman hizalamaları çalıştırılarak rafine edilir. Bu iyi belgelenmiş algoritma, okumanın her konumunu referansın tüm aday konumlarıyla karşılaştırarak çalışır. Bu karşılaştırmalar, okuma ve referans arasındaki potansiyel hizalama matrisine karşılık gelir. Bu aday hizalama pozisyonlarının her biri

için Smith-Waterman, matris hücrelerinden geçen en iyi hizalamanın bir nükleotid eşleşmesi veya eşleşmemesi (diyagonal hareket), bir delesyon (yatay hareket) veya bir insersiyon (dikey hareket) ile ona ulaşıp ulaşmadığını değerlendirmek için kullanılan skorlar oluşturur. Okuma ve referans arasındaki bir eşleşme, puana bonus verir ve bir uyumsuzluk veya indel ceza uygular. Matris boyunca genel en yüksek skorlama yolu seçilen hizalamadır.

Bu algoritmadaki skorlar için seçilen belirli değerler, birden fazla olası yorumla hizalama için nasıl dengeleneceğini, bir veya daha fazla SNP'nin aksine bir indel olasılığını veya kesmeden hizalama tercihini gösterir. Varsayılan DRAGEN puanlama değerleri, varyant arama uygulamaları için orta uzunluktaki okumaları tüm insan referans genomuna hizalamak için makuldür. Herhangi bir Smith-Waterman puanlama parametresi seti, genomik mutasyon ve sekanslama hatalarının kesin olmayan bir modelini temsil eder. Farklı şekilde ayarlanmış hizalama puanlama değerleri, bazı uygulamalar için daha uygun olabilir.

## DRAGEN Germ Hattı Varyant Arama

DRAGEN Germ Hattı Küçük Varyant Arayıcı, eşlenmiş ve hizalanmış DNA okumalarını girdi olarak alır ve sütun bazında tespit ve haplotiplerin yerel *de novo* birleşimi yoluyla SNP'leri ve indel'leri arar.

Aranabilir referans bölgeler ilk olarak yeterli hizalama kapsamıyla tanımlanır. Bu referans bölgeler içinde, sıralanan okumaların hızlı bir taraması, varyant kanıtıyla yığın sütunların etrafında ortalanmış aktif bölgeleri tanımlar. Aktif bölgeler, yakındaki önemli, referans olmayan içeriği kapsayacak şekilde yeterli bağlamla doldurulur. İndel kanıtı varsa, aktif bölgeler ek dolgu alır.

Hizalanmış okumalar her aktif bölge içinde kırılır ve bir De Bruijn grafiğine monte edilir. Kırılan okumaların kenarları, bir omurga olarak referans sekansı ile gözlem sayılarına göre ağırlıklandırılır. Bir miktar grafik temizliği ve basitleştirmeden sonra, tüm kaynak-lavabo yolları aday haplotipler olarak çıkarılır. Her haplotip, temsil ettiği varyantları tanımlamak için referans genomu hizalanmış Smith-Waterman'dır. Bu olay kümesi, konum tabanlı bir algılama ile artırılabilir. Her okuma-haplotip çifti için, haplotipin gerçek başlangıç örneği olduğu varsayılarak okumayı gözleme olasılığı  $P(r|H)$  gizli bir Markov modeli (HMM) kullanılarak tahmin edilir.

Aktif bölge üzerinde referans pozisyona göre tarama yapan aday genotipler, varyant olayların (SNP'ler veya indel'ler) diploid kombinasyonlarından oluşturulur. Her olay için (referans dahil), her örtüşen okumayı gözleme koşullu olasılığı  $P(r|e)$ , olayı destekleyen haplotipler için maksimum  $P(r|H)$  olarak tahmin edilir. Bunlar, bir genotip (olay çifti) için koşullu olasılık  $P(r|e1e2)$  ile birleştirilir ve tüm okuma yığını gözleme koşullu olasılığı  $P(R|e1e2)$  elde etmek için çarpılır. Bayes Formülü kullanılarak, her bir diploid genotipin posterior olasılığı  $P(e1e2|R)$  hesaplanır ve kazanan aranır.

Ölçeklenebilir çok numuneli varyant arama için kullanılan gVCF modunda, DRAGEN Germ Hattı Küçük Varyant Arayıcı, bir ara genomik varyant arama dosyası (gVCF) oluşturmak için numune başına çalıştırılabilir. gVCF daha sonra birden fazla örneğin verimli eklem genotipleme için kullanılabilir, bu da numunelerin hızlı bir şekilde artımlı olarak işlenmesine ve büyük kohort boyutlarına ölçeklenmesine olanak tanır.

DRAGEN Germ Hattı Küçük Varyant Arayıcı, korelasyonlu hataları gerçek varyantlardan etkin bir şekilde ayırt edebilmesini sağlayan algoritmaları olduğundan, filtreleme kuralları çok basittir.

## DRAGEN Somatik Varyant Arama

DRAGEN Somatik Küçük Varyant Arayıcı, eşlenmiş ve hizalanmış DNA okumalarını girdi olarak alır ve aktif bir bölgedeki haplotiplerin yerel *de novo* montajı yoluyla SNV'lere ve indel'leri arar.

Aranabilir referans bölgeler ilk olarak yeterli hizalama kapsamıyla tanımlanır. Bu referans bölgeler içinde, sıralanan okumaların bir taraması, tümör okumalarında bir varyantın kanıtıyla yığın sütunların etrafında ortalanmış aktif bölgeleri tanımlar. Aktif bölgeler, yakındaki önemli, referans olmayan içeriği kapsayacak şekilde yeterli bağlamla doldurulur. İndel kanıtı varsa, aktif bölgeler ek dolgu alır.

Hizalanmış okumalar her aktif bölge içinde kırılır ve bir De Bruijn grafiğine monte edilir. Kırılan okumaların kenarları, bir omurga olarak referans sekansı ile gözlem sayımlarına göre ağırlıklandırılır. Bir miktar grafik temizliği ve basitleştirmeden sonra, tüm kaynak-lavabo yolları aday haplotipler olarak çıkarılır. Her haplotip, temsil ettiği varyantları tanımlamak için referans genoma hizalanmış Smith-Waterman'dır. Her bir okuma-haplotip çifti için, okumayı gözleme olasılığı  $P(r|H)$  haplotipin gerçek başlangıç örneği olduğu varsayılarak çift gizli Markov modeli (HMM) kullanılarak tahmin edilir.

TLOD skorunu belirlemek için, DRAGEN Somatik Küçük Varyant Arayıcı ilk olarak her bir aday somatik olayı ve aktif bölge üzerindeki referans olayı için referans pozisyonuna göre tarama yapar. Her örtüşen okumayı gözlemlemenin koşullu olasılığı  $P(r|e)$  olayı destekleyen haplotipler için maksimum  $P(r|H)$  olarak tahmin edilir. Bunlar, bir olay hipotezi olan E için koşullu olasılık  $P(r|E)$  ile birleştirilir, referans ve aday somatik alelin olası alel frekansları aralığında bir karışımını içerir ve tüm okuma yığını gözlemlemenin koşullu olasılığı  $P(R|E)$  elde etmek için çarpılır. Buradan, belirli bir lokusta tümör numunesinde bir ALT alelinin mevcut olduğuna dair kanıt olarak bir TLOD skoru hesaplanır.

# Create a Run (Bir Çalıştırma Oluştur)

Illumina Run Manager’de veya NovaSeq 6000Dx üzerinde veya ağa bağlı bir bilgisayarda bir tarayıcı kullanarak bir çalıştırma ayarlamak için aşağıdaki adımları izleyin. Numune verileri manuel olarak veya bir numune sayfası içe aktararak girilebilir.

## Uygulama ve Çalıştırma Ayarları

1. Çalışmalar ekranından Çalışma **Create Run** (Oluştur ögesini seçin).
2. DRAGEN for Illumina DNA Prep with Enrichment Dx uygulamasını ve ardından **Next** (Sonraki)’ni seçin.
3. Çalışma Ayarları ekranında bir çalışma adı girin. Çalıştırma adı, sekanslamadan analize kadar çalıştırmayı tanımlar.
4. **[İsteğe bağlı]** Çalıştırmayı daha fazla tanımlamak için bir çalıştırma açıklaması girin.
5. Seçilen Kitaplık Hazırlık Kitinin bir kitaplık Illumina DNA Prep with Enrichment Dx hazırlık kiti olduğundan emin olun.
6. İstenen dizin adaptör kitini seçin.
7. Okuma Uzunluğunu girin.  
Okuma 1 ve Okuma 2’nin varsayılan değeri 151 dögüdür.  
Dizin 1 ve Dizin 2, 10 dögülük sabit bir değere sahiptir.
8. **[İsteğe bağlı]** Bir kitaplık tüpü kimliği girin.
9. **Next** (Sonraki) seçeneğini belirleyin.

## Numune Verileri

Numune bilgilerini manuel olarak girmek için Numune Verileri ekranındaki tabloyu kullanın. Alternatif olarak, numune bilgilerini yüklemek için **Import Samples** (Numuneleri İçe Aktar) ögesini seçin. Numune bilgilerini içe aktarma hakkında bilgi için, [Numuneleri İçe Aktar sayfa 5](#) kısmına bakın.

### Numuneleri Manuel Olarak Girme

1. Sample ID (Numune Numarası) alanına benzersiz bir numune numarası girin.
2. Kuyu pozisyonunu seçmek için **Plate - Well Position** (Plaka - Kuyu Pozisyonu) ögesini kullanın. i7 Dizini, Dizin 1, i5 Dizini ve Dizin 2 alanları otomatik olarak doldurulur.
3. **[İsteğe bağlı]** Bir kitaplık adı girin.
4. Satırları ekleyin ve tüm numuneler tabloya eklenene kadar **1-3** adımlarını gerektiği şekilde tekrarlayın.
5. **Next** (Sonraki) seçeneğini belirleyin.



## Numuneleri İçe Aktar

Ağa bağlı bir bilgisayarda tarayıcı kullanarak bir çalışma planlarken Örnek Verileri ekranında bir Illumina Run Manager şablon (\*.csv) indirilebilir.

1. Boş bir CSV dosyası indirmek için **Download Template** (Şablonu İndir) ögesini seçin.
2. CSV dosyasından numune bilgilerini girin ve dosyayı kaydedin.  
Örnek sayfa CVS dosyası aşağıdaki veri sütunlarını içerir: Sample ID (Numune Numarası), Plate - Well Position (Plaka - Kuyu Pozisyonu), **Optional** Library Name (İsteğe Bağlı Kitaplık Adı).
3. **Import Samples** (Numuneleri İçe Aktar) ögesini seçin ve CSV dosyasının konumuna gidin.
4. **Next** (Sonraki) seçeneğini belirleyin.

## Analiz Ayarları

1. İstenen analiz iş akışını seçin:
  - FASTQ oluşturma
  - Bir germ hattı iş akışı için germ hattı FASTQ ve VCF oluşturma
  - Somatik bir iş akışı için Somatik FASTQ ve VCF oluşturma
2. **[İsteğe bağlı]** İstenirse, FASTQ ORA sıkıştırmasını etkinleştirmek için **Generate ORA compressed FASTQs** (ORA sıkıştırılmış FASTQ'ları Oluştur) onay kutusunu seçin.
3. **[VCF oluşturma iş akışları]** Bir manifesto dosyası seçmek için **Manifest File Selection** (Manifesto Dosyası Seçimi) açılır menüsünü kullanın.  
DRAGEN for Illumina DNA Prep with Enrichment Dx için bir manifesto dosyası gereklidir. Manifesto, hedeflenen referans bölgelerinin adlarını ve konumlarını belirten sekmeye ayrılmış bir BED (\*.bed) dosyasıdır.
4. **[Somatic FASTQ ve VCF oluşturma iş akışı]** Bir gürültü dosyası seçmek için **Noise File Selection** (Gürültü Dosyası Seçimi) açılır menüsünü kullanın.  
Sistemik gürültüyü filtrelemek için sahaya özel gürültü seviyesine sahip bir BED dosyası belirtilebilir. Daha fazla bilgi için bkz. [Gürültü Filtreleme sayfa 6](#).
5. **Next** (Sonraki) seçeneğini belirleyin.

## Çalıştırma İnceleme

1. İnceleme ekranında, Çalıştırma Ayarları, Numune Verileri ve Analiz Ayarları ekranlarına girilen bilgileri inceleyin.
2. Ardından **Save** (Kaydet) seçeneğini belirleyin.  
Çalışma, Çalışmalar ekranındaki Planlanan sekmesine kaydedilir.

## Ayarlar

Geçerli ayarları görüntülemek ve ayarları değiştirmek için Applications (Uygulamalar) ekranında uygulamayı seçin.

## Yapılandırma

Yapılandırma ekranı aşağıdaki uygulama ayarlarını görüntüler:

- **Library Prep Kits** (Kitaplık Hazırlık Kitleri) — Uygulama için varsayılan kitaplık hazırlık kitini görüntüler. Bu ayar değiştirilemez.
- **Index Adapter Kits** (Dizin Adaptör Kitleri) — Uygulama için varsayılan dizin adaptör kitini görüntüler. Bu ayar değiştirilemez.
- **Read lengths** (Okuma uzunlukları) — Okuma uzunlukları uygulama için varsayılan olarak 151 olarak ayarlanır, ancak çalışma oluşturma sırasında değiştirilebilir.
- **Manifest and Noise Files** (Manifesto ve Gürültü Dosyaları) — Manifesto ve gürültü dosyaları için ayarları yükleyin ve değiştirin.
  - Analizde kullanmak üzere dosyaları yüklemek için **Upload File** (Dosya Yükle) ögesini seçin.
  - Dosyayı, uygulama seçildiğinde çalışma oluşturma sırasında seçilen varsayılan manifesto veya gürültü dosyası olarak ayarlamak için **Default** (Varsayılan) radyo düğmesini seçin.
  - Çalışma oluşturma sırasında açılır menüde görüntülenecek dosyayı ayarlamak için **Enabled** (Etkin) onay kutusunu seçin.

## İzinler

Uygulamanın kullanıcı erişimini yönetmek için İzinler ekranındaki onay kutularını kullanın.

## Gürültü Filtreleme

Somatik iş akışı kullanılırken sistematik gürültü filtreleme kullanılabilir. Filtre Tümör Normal modunda kullanılabilir, ancak eşleşen bir normalin mevcut olmadığı Sadece Tümör çalışmaları için özellikle faydalıdır.

Sistematik gürültü YATAK normal numunelerden oluşturulmalıdır. Kitaplık hazırlama, sıralama sistemi ve panele özgü sistematik gürültü dosyaları oluşturulması önerilir. Gürültü dosyası oluşturmak için yaklaşık 50 normal örneğin kullanılması önerilir.

# Analiz Çıktıları

DRAGEN for Illumina DNA Prep with Enrichment Dx aşağıdaki bilgileri analiz klasörüne kaydeder. Yalnızca germ hattı ve somatik iş akışları PDF üretir.

- Kullanılan bildirim dosyası
- Yazılım sürümü
- Numune Kimlikleri
- Toplam hizalanmış okumalar
- Örnek başına hizalanan okumaların yüzdesi
- Örnek başına aranan SNV sayısı
- Örnek başına aranan indel sayısı
- Kapsam istatistikleri

## Analiz Çıktı Dosyaları

Aşağıdaki çıkış dosyaları uygulama tarafından oluşturulur. Oluşturulan tam dosyalar, hangi analiz iş akışının kullanıldığına bağlıdır. Çıkış dosyaları analiz klasöründe bulunur.

Çıktı Dosyası	Açıklama
FASTQ (*.fastq.gz veya *.fastq.ora)	Kaliteli puanlı baz aramaları içeren ara dosyalar. FASTQ dosyaları hizalama adımı için birincil girdidir. ORA sıkıştırması seçilirse, dosya adı bunu yansıtır.
Hizalama BAM dosyaları (*.bam)	Belirli bir numune için hizalanmış okumalar içerir.
Genom VCF dosyaları (*.gvcf.gz)	İster varyant ister referans olarak aransın, her pozisyon için genotip içerir.
VCF dosyaları (*.vcf.gz)	Her pozisyonda aranan varyantları içerir.
Çalışma metrikleri raporu (*.csv)	Toplam verim ve Q30 skoru dahil olmak üzere çalışma hakkında kalite metriklerini içerir.

## FASTQ Dosyaları

FASTQ (\*.fastq.gz, \*.fastq.ora), temel aramaları ve okuma başına kalite değerlerini içeren metin tabanlı bir dosya biçimidir. Her dosya aşağıdaki bilgileri içerir:

- Numune tanımlayıcı
- Sekans
- Bir artı işareti (+)
- ASCII + 33 kodlanmış formatta Phred kalite skorları

Numune tanımlayıcısı aşağıdaki gibi biçimlendirilir.

```
@Instrument:RunID:FlowCellID:Lane:Tile:X:Y
ReadNum:FilterFlag:0:SampleNumber
Örnek:
@SIM:1:FCX:1:15:6329:1045 1:N:0:2
TCGCACTCAACGCCCTGCATATGACAAGACAGAATC
+
<>;##=><9=AAAAAAAAAAA9#:<#<;<<<????#=#
```

## BAM Dosyaları

BAM dosyası (\*.bam), 128 Mb'ye kadar hizalanmış dizileri temsil etmek için kullanılan SAM (sekans hizalama haritası) dosyasının sıkıştırılmış ikili sürümüdür. BAM dosyaları `SampleName_S#.bam` dosya adlandırma formatını kullanır. # çalışma için numunelerin listelenme sırası ile belirlenen numune numarasıdır. Çok düğümlü modda S#, numunenin sırasına bakılmaksızın S1 olarak ayarlanır.

BAM dosyaları bir başlık bölümü ve bir hizalama bölümü içerir:

- Başlık: Örnek adı, numune uzunluğu ve hizalama yöntemi gibi tüm dosya hakkında bilgiler içerir. Hizalamalar bölümündeki hizalamalar, başlık bölümündeki belirli bilgilerle ilişkilidir.
- Hizalamalar—Okuma adı, okuma sekansı, okuma kalitesi, hizalama bilgileri ve özel etiketler içerir. Okunan ad kromozomu, başlangıç koordinatını, hizalama kalitesini ve eşleşme tanımlayıcı dizesini içerir.

Hizalamalar bölümü, her okuma veya okuma çifti için aşağıdaki bilgileri içerir:

- AS: Eşleştirilmiş uç hizalama kalitesi.
- RG: Belirli bir numune için okuma sayısını belirten okuma grubu.
- BC: Okuma ile ilişkili demultipleks numune kimliğini gösteren barkod etiketi.
- SM: Tek uçlu hizalama kalitesi.
- XC: Tanımlayıcı dizeyi eşleştirin.

- XN: Okunan BAM dizin dosyaları (\*.bam.bai) ile ilişkili amplicon kimliğini kaydeden amplicon isim etiketi, ilgili BAM dosyasının bir dizinini sağlar.

## VCF Dosyaları

Varyant arama formatı (\*.vcf) dosyaları, bir referans genomdaki spesifik konumlarda bulunan varyantlara ilişkin bilgileri içerir.

VCF dosya başlığı, VCF dosya biçimi sürümünü ve varyant arayıcısı sürümünü içerir ve dosyanın geri kalanında kullanılan notları listeler. VCF başlığı ayrıca referans genom dosyasını ve BAM dosyasını da içerir. Başlıktaki son satır, veri satırları için sütun başlıklarını içerir. VCF dosyası veri satırlarının her biri tek bir varyant hakkında bilgi içerir.

Tablo 1 VCF Dosya Başlıkları

Başlık	Açıklama
CHROM	Referans genomun kromozomu. Kromozomlar referans FASTA dosyasıyla aynı sırada görünür.
POS	Referans kromozomdaki varyantın tek baz konumu. Tek nükleotid varyantları (SNV'ler) için bu pozisyon, varyantın referans bazıdır. İndeller için bu pozisyon, varyanttan hemen önceki referans bazıdır.
Kimlik	Varsa dbSNP.txt'den elde edilen SNP için rs (referans SNP) numarası. Bu konumda birden fazla rs numarası varsa, liste noktalı virgülle sınırlandırılır. Bu pozisyonda bir dbSNP girişi yoksa, eksik bir değer işareti ('.') kullanılır.
REF	Referans genotip. Örnek olarak, tek bir T'nin silinmesi referans TT ve alternatif T olarak temsil edilir. Bir A'dan T'ye tek nükleotid varyantı referans A ve alternatif T olarak temsil edilir.
ALT	Referanstan farklı olan aleller okundu. Örneğin, tek bir T'nin insersiyonu referans A ve alternatif AT olarak temsil edilir. Bir A ile T tek nükleotid varyantı referans A ve alternatif T olarak temsil edilir.
İKİLİ	Varyant arayıcı tarafından atanan Phred ölçekli kalite puanı. Daha yüksek skorlar varyanta daha fazla güven ve daha düşük hata olasılığına işaret eder. Q kalite puanı için, tahmini hata olasılığı $10^{-(Q/10)}$ 'dur. Örneğin, Q30 aramaları kümesinin %0,1 hata oranı vardır. Birçok varyant arayıcı, gözlemlenen hata oranına göre yüksek olan istatistiksel modellerine dayalı olarak kalite skorları atar.

Tablo 2 VCF Dosya Açıklamaları

Başlık	Açıklama
FİLTRE	<p>Tüm filtreler geçilirse filtre sütununa GEÇTİ yazılır. Germ Hattı iş akışı olası FİLTRE girişleri şunları içerir:</p> <ul style="list-style-type: none"> <li>• <b>DRAGENSnpHardQUAL</b>—SNSP varyantı QUAL skoru eşiği karşılamıyorsa uygulanır</li> <li>• <b>DRAGENIndelHardQUAL</b>—Indel varyantı QUAL skoru eşiği karşılamıyorsa uygulanır</li> <li>• <b>DüşükDerinlik</b>—Kapsam derinliği eşiği karşılamadığı için çalışma merkezi filtrelendi</li> <li>• <b>Düşük GQ</b>—Genotip kalitesi eşiği karşılamadığı için çalışma merkezi filtrelendi</li> <li>• <b>PloidyConflict</b>—Varyant arayıcıdan genotip araması kromozom ploidi ile tutarlı değil</li> <li>• <b>base_quality</b>—Bu lokusta alt okumalarının medyan baz kalitesi eşiği karşılamadığı için çalışma merkezi filtrelendi</li> <li>• <b>filter_reads</b>—Okumaların çok büyük bir kısmı filtrelendiğinden çalışma merkezi filtrelendi</li> <li>• <b>fragment_length</b>—Bu lokusta alt okumaların medyan fragman uzunluğu ile ref okumaların medyan fragman uzunluğu arasındaki mutlak fark eşiği aştığından, alan filtrelendi</li> <li>• <b>low_Derinlik</b>—Okuma derinliği çok düşük olduğu için çalışma merkezi filtrelendi</li> <li>• <b>low_frac_info_reads</b>—Bilgilendirici okumaların fraksiyonu eşiğin altında olduğu için çalışma merkezi filtrelendi</li> <li>• <b>low_normal_derinlik</b>—Normal numune okuma derinliği çok düşük olduğu için çalışma merkezi filtrelendi</li> <li>• <b>long_indel</b>—Merkez, indel uzunluğu çok uzun olduğu için filtrelendi</li> <li>• <b>map_quality</b>—Bu lokusta alt okumaların medyan haritalama kalitesi eşiği karşılamadığı için çalışma merkezi filtrelendi</li> <li>• <b>multiallelik</b>—İkiden fazla alt alel tümör LOD'sini geçtiği için çalışma merkezi filtrelendi</li> <li>• <b>non_homref_normal</b>—Normal numune genotipi homozigot referans olmadığı için çalışma merkezi filtrelendi</li> <li>• <b>no_reliable_supporting_read</b>—Güvenilir destekleyici somatik okuma mevcut olmadığından site filtrelendi</li> <li>• <b>panel_of_normals</b>—Normaller panelinde en az bir numunede görülür vcf</li> <li>• <b>read_position</b>—Çalışma merkezi, okumanın başlangıcı/bitışı ile bu lokus arasındaki mesafelerin medyanının eşiğin altında olması nedeniyle filtrelendi</li> <li>• <b>RMxNRepeatRegion</b>—Varyant alelinin tümü veya bir kısmı referansın tekrarı olduğu için çalışma merkezi filtrelendi</li> <li>• <b>strand_artifact</b>—Şiddetli zincir biası nedeniyle çalışma merkezi filtrelendi</li> <li>• <b>str_contraction</b>—Alt alelin referanstan bir tekrar birimi daha az olduğu şüpheli PCR hatası nedeniyle çalışma merkezi filtrelendi</li> </ul>

Başlık	Açıklama
FİLTRE (devamı)	<ul style="list-style-type: none"> <li>• <b>too_few_supporting_reads</b>—Tümör numunesinde çok az destekleyici okuma olduğu için çalışma merkezi filtrelendi</li> <li>• <b>weak_evidence</b>—Somatik varyant skoru eşiği karşılamıyor</li> </ul> <p>Somatik iş akışı olası FİLTRE girişleri şunları içerir:</p> <ul style="list-style-type: none"> <li>• <b>base_quality</b>—Bu lokusta alt okumalarının medyan baz kalitesi eşiği karşılamadığı için çalışma merkezi filtrelendi</li> <li>• <b>filtered_reads</b>—Okumaların çok büyük bir kısmı filtrelendiğinden çalışma merkezi filtrelendi</li> <li>• <b>fragment_length</b>—Bu lokusta alt okumaların medyan fragman uzunluğu ile ref okumaların medyan fragman uzunluğu arasındaki mutlak fark eşiği aştığından, alan filtrelendi</li> <li>• <b>low_Derinlik</b>—Okuma derinliği çok düşük olduğu için çalışma merkezi filtrelendi</li> <li>• <b>low_frac_info_reads</b>—Bilgilendirici okumaların fraksiyonu eşiğin altında olduğu için çalışma merkezi filtrelendi</li> <li>• <b>low_normal_derinlik</b>—Normal numune okuma derinliği çok düşük olduğu için çalışma merkezi filtrelendi</li> <li>• <b>long_indel</b>—Merkez, indel uzunluğu çok uzun olduğu için filtrelendi</li> <li>• <b>map_quality</b>—Bu lokusta alt okumaların medyan haritalama kalitesi eşiği karşılamadığı için çalışma merkezi filtrelendi</li> <li>• <b>multiallelik</b>—İkiden fazla alt alel tümör LOD'sini geçtiği için çalışma merkezi filtrelendi</li> <li>• <b>non_homref_normal</b>—Normal numune genotipi homozigot referans olmadığı için çalışma merkezi filtrelendi</li> <li>• <b>no_reliable_supporting_read</b>—Güvenilir destekleyici somatik okuma mevcut olmadığından site filtrelendi</li> <li>• <b>panel_of_normals</b>—Normaller panelinde en az bir numunede görülür vcf</li> <li>• <b>read_position</b>—Çalışma merkezi, okumanın başlangıcı/bitışı ile bu lokus arasındaki mesafelerin medyanının eşiğin altında olması nedeniyle filtrelendi</li> <li>• <b>RMxNRepeatRegion</b>—Varyant alelinin tümü veya bir kısmı referansın tekrarı olduğu için çalışma merkezi filtrelendi</li> <li>• <b>strand_artifact</b>—Şiddetli zincir biası nedeniyle çalışma merkezi filtrelendi</li> <li>• <b>str_contraction</b>—Alt alelin referanstan bir tekrar birimi daha az olduğu şüpheli PCR hatası nedeniyle çalışma merkezi filtrelendi</li> <li>• <b>too_few_supporting_reads</b>—Tümör numunesinde çok az destekleyici okuma olduğu için çalışma merkezi filtrelendi</li> <li>• <b>weak_evidence</b>—Somatik varyant skoru eşiği karşılamıyor</li> <li>• <b>systematic_noise</b>—Normallerdeki sistematik gürültü kanıtına göre çalışma merkezi filtrelendi</li> </ul>

Başlık	Açıklama
BİLGİ	<p>Germ Hattı iş akışı olası INFO girişleri şunları içerir:</p> <ul style="list-style-type: none"> <li>• <b>AC</b>—Her ALT aleli için genotiplerde listelenenle aynı sırada alel sayısı.</li> <li>• <b>AF</b>—Her ALT aleli için Alel Frekansı, listelenenle aynı sırada.</li> <li>• <b>AN</b>—Aranan genotiplerde bulunan alelin toplam sayısı.</li> <li>• <b>DB</b>—dbSNP Üyeliği.</li> <li>• <b>FS</b>—Zincir biasını tespit etmek için Fisher'ın kesin testi kullanılarak Phred ölçeğinde ölçülen p değeri.</li> <li>• <b>QD</b>—Derinliğe Göre Varyant Güvenirliği/Kalitesi.</li> <li>• <b>R2_5P_bias</b>—Eşleşme biasına ve 5 ana uçtan uzaklığa dayalı puan.</li> <li>• <b>SOR</b>—Zincir sapmasını tespit etmek için 2x2 beklenmedik durum tablosunun Simetrik Oranları.</li> <li>• <b>DP</b>—Yaklaşık okuma derinliği (bilgilendirici ve bilgilendirici olmayan); bazı okumalar mapq vb.'ye göre filtrelenmiş olabilir.</li> <li>• <b>BİTİŞ</b>—Aralığın durma konumu.</li> <li>• <b>FractionInformativeReads</b>—Toplam okumalardan bilgilendirici okumaların fraksiyonu.</li> <li>• <b>MQ</b>—RMS Eşleme Kalitesi.</li> <li>• <b>MQRankSum</b>—Wilcoxon sıralama toplamı Alt ve Ref okuma eşleştirme kalitelerinin testinden Z-skoru.</li> <li>• <b>ReadPosRankSum</b>—Wilcoxon sıralama toplamı Alt ve Ref okuma konumu bias testinden Z-skoru.</li> <li>• <b>SOMATIC</b>—Bu pozisyonda en az bir varyant somatiktir.</li> </ul> <p>Somatik iş akışı olası INFO girişleri şunları içerir:</p> <ul style="list-style-type: none"> <li>• <b>DP</b>—Yaklaşık okuma derinliği (bilgilendirici ve bilgilendirici olmayan); bazı okumalar mapq vb.'ye göre filtrelenmiş olabilir.</li> <li>• <b>BİTİŞ</b>—Aralığın durma konumu.</li> <li>• <b>FractionInformativeReads</b>—Toplam okumalardan bilgilendirici okumaların fraksiyonu.</li> <li>• <b>MQ</b>—RMS Eşleme Kalitesi.</li> <li>• <b>MQRankSum</b>—Wilcoxon sıralama toplamı Alt ve Ref okuma eşleştirme kalitelerinin testinden Z-skoru.</li> <li>• <b>ReadPosRankSum</b>—Wilcoxon sıralama toplamı Alt ve Ref okuma konumu bias testinden Z-skoru.</li> <li>• <b>AQ</b>—Sistemik gürültü puanı.</li> <li>• <b>hotspot</b>—Aramada güveni artırmak için kullanılan, bilinen somatik çalışma merkezi.</li> <li>• <b>SOMATIC</b>—Bu pozisyonda en az bir varyant somatiktir.</li> </ul>



Başlık	Açıklama
FORMAT	<p>Format sütunu, iki nokta üst üste işaretleriyle ayrılan alanları listeler. Örneğin, GT:GQ.</p> <p>Germ Hattı iş akışı kullanılabilir alanları şunları içerir:</p> <ul style="list-style-type: none"> <li>• <b>AD</b>—Listelenen sıradaki ref ve alt aleller için alel derinlikleri (toplam okumalardan sadece bilgilendirici okumaları sayarak).</li> <li>• <b>AF</b>—Listelenen sırada alt aleller için alel fraksiyonları.</li> <li>• <b>DP</b>—Yaklaşık okuma derinliği (MQ=255 veya kötü eşler ile okumalar filtrelenir).</li> <li>• <b>F1R2</b>—Her bir aleli destekleyen F1R2 çift oryantasyonundaki okumaların sayısı.</li> <li>• <b>F2R1</b>—Her bir aleli destekleyen F2R1 çift oryantasyonunda okuma sayısı.</li> <li>• <b>GP</b>—VCF spesifikasyonunda tanımlanan genotipler için Phred ölçekli posterior olasılıklar.</li> <li>• <b>GQ</b> — Genotip kalitesi.</li> <li>• <b>GT</b> — Genotip. 0 referans bazına karşılık gelir, 1 ALT sütunundaki ilk girişe karşılık gelir, vb. İleri kesme işareti (/), faz bilgisinin mevcut olmadığını gösterir.</li> <li>• <b>MB</b> — Eşleşme biasını tespit etmek için numune başına bileşen istatistikleri.</li> <li>• <b>PL</b> — VCF spesifikasyonunda tanımlandığı gibi genotipler için normalleştirilmiş, Phred ölçekli olasılıklar.</li> <li>• <b>PRI</b> — Genotipler için önceden belirlenmiş olasılıklar.</li> <li>• <b>PS</b> — Fiziksel faz kimliği bilgileri; burada belirli bir numunedeki (ancak numuneler arasında değil) her bir benzersiz kimlik, bir faz grubu içindeki kayıtları bağlar.</li> <li>• <b>SB</b> — Zincir biasını saptamak için Fisher'ın Kesin Testini oluşturan numune başına bileşen istatistikleri.</li> <li>• <b>SQ</b> — Somatik kalite.</li> </ul> <p>Somatik iş akışı kullanılabilir alanları şunları içerir:</p> <ul style="list-style-type: none"> <li>• <b>AD</b>—Listelenen sıradaki ref ve alt aleller için alel derinlikleri (toplam okumalardan sadece bilgilendirici okumaları sayarak).</li> <li>• <b>AF</b>—Listelenen sırada alt aleller için alel fraksiyonları.</li> <li>• <b>DP</b>—Yaklaşık okuma derinliği (MQ=255 veya kötü eşler ile okumalar filtrelenir).</li> <li>• <b>F1R2</b>—Her bir aleli destekleyen F1R2 çift oryantasyonundaki okumaların sayısı.</li> <li>• <b>F2R1</b>—Her bir aleli destekleyen F2R1 çift oryantasyonunda okuma sayısı.</li> <li>• <b>GT</b> — Genotip. 0 referans bazına karşılık gelir, 1 ALT sütunundaki ilk girişe karşılık gelir, vb. İleri kesme işareti (/), faz bilgisinin mevcut olmadığını gösterir.</li> <li>• <b>MB</b> — Eşleşme biasını tespit etmek için numune başına bileşen istatistikleri.</li> <li>• <b>PS</b> — Fiziksel faz kimliği bilgileri; burada belirli bir numunedeki (ancak numuneler arasında değil) her bir benzersiz kimlik, bir faz grubu içindeki kayıtları bağlar.</li> <li>• <b>SB</b> — Zincir biasını saptamak için Fisher'ın Kesin Testini oluşturan numune başına bileşen istatistikleri.</li> <li>• <b>SQ</b> — Somatik kalite.</li> </ul>
NUMUNE	Örnek sütunu, FORMAT sütununda belirtilen değerleri verir.

## Genom VCF Dosyaları

Genome VCF (\*.gvcf.gz) dosyaları, genom içindeki tüm çalışma merkezlerini makul ölçüde kompakt bir biçimde temsil etmek için bir dizi kurala uyar. gVCF dosyaları, her bir numune için tek bir dosyada ilgili bölge içindeki tüm çalışma merkezlerini içerir. gVCF dosyası tüm filtreleri geçmeyen pozisyonlarda no call gösterir. ./ genotip (GT) etiketi no call anlamına gelir.

## Analiz Sonuçlarını Görüntüleme

Şu anda devam etmekte olan çalışmalar Aktif sekmesinde görüntülenir. Tamamlanan çalışmalar Tamamlandı sekmesinde görüntülenir. Sonuçları görüntüleme hakkında daha fazla bilgi için [NovaSeq 6000Dx Ürün Belgeleri \(doküman no. 200010105\)](#)'e bakın.

# Teknik Yardım

Teknik yardım için Illumina Teknik Destek bölümüyle iletişim kurun.

Web sitesi: [www.illumina.com](http://www.illumina.com)  
E-posta: [techsupport@illumina.com](mailto:techsupport@illumina.com)

## Illumina Teknik Destek Birimi Telefon Numaraları

Bölge	Ücretsiz	Uluslararası
Avustralya	+61 1800 775 688	
Avusturya	+43 800 006249	+43 1 9286540
Belçika	+32 800 77 160	+32 3 400 29 73
Kanada	+1 800 809 4566	
Çin		+86 400 066 5835
Danimarka	+45 80 82 01 83	+45 89 87 11 56
Finlandiya	+358 800 918 363	+358 9 7479 0110
Fransa	+33 8 05 10 21 93	+33 1 70 77 04 46
Almanya	+49 800 101 4940	+49 89 3803 5677
Hong Kong, Çin	+852 800 960 230	
Hindistan	+91 8006500375	
Endonezya		0078036510048
İrlanda	+353 1800 936608	+353 1 695 0506
İtalya	+39 800 985513	+39 236003759
Japonya	+81 0800 111 5011	
Malezya	+60 1800 80 6789	
Hollanda	+31 800 022 2493	+31 20 713 2960
Yeni Zelanda	+64 800 451 650	
Norveç	+47 800 16 836	+47 21 93 96 93
Filipinler	+63 180016510798	
Singapur	1 800 5792 745	
Güney Kore	+82 80 234 5300	
İspanya	+34 800 300 143	+34 911 899 417

Bölge	Ücretsiz	Uluslararası
İsveç	+46 2 00883979	+46 8 50619671
İsviçre	+41 800 200 442	+41 56 580 00 00
Tayvan, Çin	+886 8 06651752	
Tayland	+66 1800 011 304	
Birleşik Krallık	+44 800 012 6019	+44 20 7305 7197
Amerika Birleşik Devletleri	+1 800 809 4566	+1 858 202 4566
Vietnam	+84 1206 5263	

**Güvenlik veri sayfaları (SDS'ler)**—Illumina web sitesinde [support.illumina.com/sds.html](https://support.illumina.com/sds.html) adresinde mevcuttur.

**Ürün belgeleri**—[support.illumina.com](https://support.illumina.com) adresinden indirilebilir.



Illumina

5200 Illumina Way  
San Diego, California 92122 ABD  
+1.800.809.ILMN (4566)  
+1.858.202.4566 (Kuzey Amerika dışından)  
techsupport@illumina.com  
www.illumina.com

CE

IVD

EC REP

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

**Avustralya Sponsoru**

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

İN VİTRO TANI AMAÇLI KULLANIM İÇİNDİR

© 2022 Illumina, Inc. Tüm hakları saklıdır.

illumina®