

Biyoinformatik Ödev Hazırlama Kılavuzu

Ahmet Berkay Demirseçen, Emrah Kırdök

Giriş

Bu belge, size verilen ödevi nasıl hazırlamanız gerektiğini, ne tarz bir dil kullanmanızı istediğimizi ve nelere dikkat etmeniz gerektiğini anlatmak için hazırlanmıştır. Lütfen oluşturduğunuz raporları sadece bilgi olarak değil, estetik olarak da düzgün bir şekilde hazırlayın.

Belgede gösterilen kısımlar sadece nasıl hazırlamanız gerektiğine örnek olması için konulmuştur, sizin raporlarınızı daha detaylı görmek istiyoruz. Kendi raporunuzda daha farklı tarzlar ve grafikler kullanmanız beklenmektedir.

Notlandırma, yazım kurallarına uymanız, bilimsel bir dille yazmanız, verdiğiniz bilgilerin/yaptığınız yorumların doğruluğu, belgenizin düzeni üzerinden yapılmaktadır. Aynı zamanda, eğer bir yerden faydalandıysanız, kaynakça kısmında ve metnin içerisinde göstermelisiniz.

Ödeviniz basit olarak aşağıdaki kısımları içermelidir:

- Giriş kısmı: Yaptığınız çalışma ve kullandığınız veri hakkında genel bilgi
- Materyal ve Yöntem: Çalışmada kullanılan verinin ve yöntemin belirlenmesi ve anlatılması
- Sonuçlar: Yaptığınız analiz sonrasında elde ettiğiniz grafiklerin yorumu
- Tartışma: Tüm analizlerin birleşimi sonucu, veri hakkında genel bir yorum, tartışma. (Mesela, bu veriyi çalışmalarınızda kullanır mıydınız? Neden?)

Örnek Ödev

Ahmet Berkay Demirşen - 123456789123 (Numaranız)

Giriş

Bu ödevde kullandığım veri, yaşlanmanın nörodejeneratif hastalıklara nasıl etki ettiğini gösteren bir çalışmadan alınmıştır. Bu çalışmada, örnek olarak fareden (*Mus musculus*) elde edilen beyin hipokampus bölgesi kullanılmış ve genomik yöntemler kullanılarak iki uçlu (paired-end) olarak dizilenmiştir.

Elde edilen genomik veri, SRA veritabanına PRJDB18626 koduyla yüklenmiştir.

Kullanılan veride, 7 haftalık ve 108 haftalık fareler kullanılmış, yaşlanmanın nörodejeneratif hastalıklara genomik ve transkriptomik boyutta nasıl etki ettiği araştırılmıştır.

Elde edilen veri öncelikle Galaxy platformuna yüklenmiştir.

Veri üzerinde yapılan analizler şu şekildedir:

- FastQC ile önkontrol
- Cutadapt ile adaptörlerin çıkartılması
- Cutadapt işlemi sonrası FastQC ile ikinci kontrol.

*Siz yaptığınız ekstra işlemleri buraya ekleyebilirsiniz. *

Materyal ve Yöntem

Elde edilen veri, 7 ve 108 haftalık farelerin hipokampus bölgesinden izole edilmiş ve dizilenmiştir. Bu işlemler ise şu adımları içermektedir:

- Beyinler önce homojenize edilmiş ve hücre çekirdekleri izole edilmiştir.
- İzole edilen çekirdekler kullanarak kütüphaneler oluşturulmuştur.
- Oluşturulan kütüphaneler DNB-seq platformu kullanılarak, 100 baz çifti uzunluğunda dizilenmiştir.

* Siz buraya prosedürler hakkında bilgi ekleyebilirsiniz ama kaynakça göstermeye dikkat edin *

Sonuçlar

FastQC ile Önkontrol

Kullandığım veri iki uçlu (paired-end) olduğundan Forward ve Reverse okumaları ayrı ayrı yorumlayacağım.

Basic Statistics

Measure	Value
Filename	forward.gz
File type	Conventional base calls
Encoding	Sanger / Illumina 1.9
Total Sequences	21495745
Total Bases	3.2 Gbp
Sequences flagged as poor quality	0
Sequence length	151
%GC	42

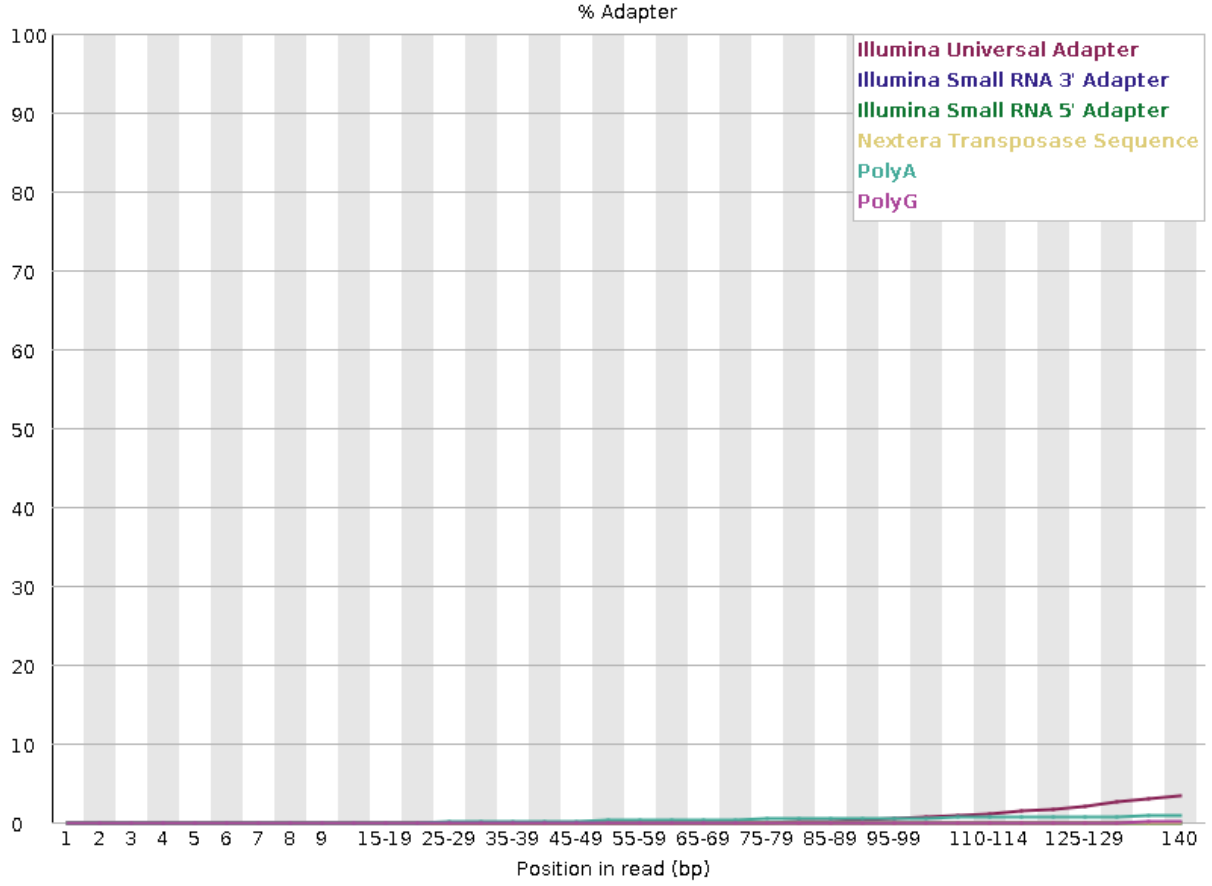
Şekil 1. Yüklenen verinin, ileri yönlü okumalarının temel istatistik tablosu

Şekil 1'deki tabloda, yüklenen verinin temel istatistikleri görülmektedir. Bu verilerden bazıları şunlardır:

- Toplam baz sayısı
- Toplam okuma sayısı
- GC oranı
- Düşük kaliteli olarak işaretlenen dizi sayısı

Bu bilgiler, bize kullandığımız verinin kalitesi hakkında bazı ön bilgiler vermektedir.

Forward okumada, hiç düşük kaliteli olan okuma bulunmamış, GC oranı ise normal olarak kabul edilen aralıktadır.

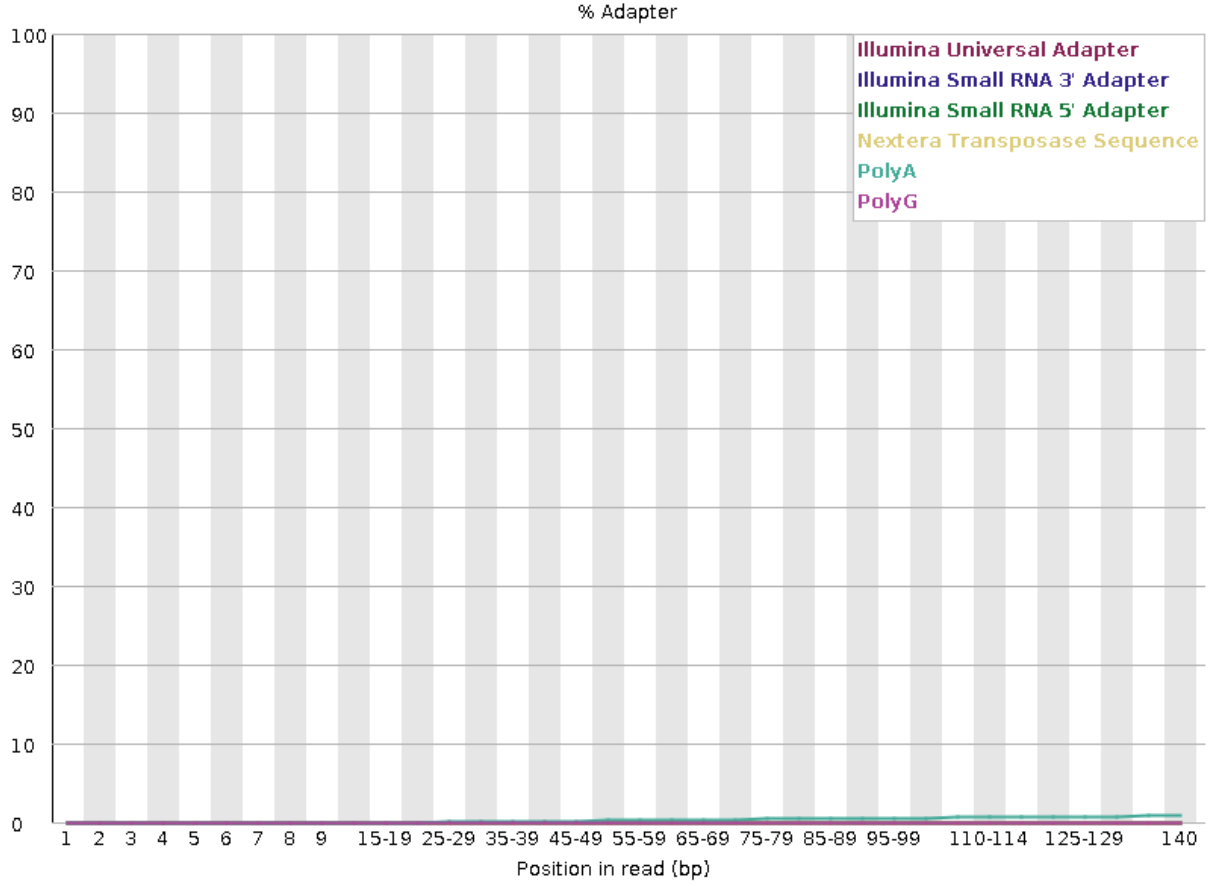


Şekil 2. Baz çiftlerine göre adaptör oranları görülmektedir. Çizgilerin renkleri ise hangi adaptörün olduğunu göstermektedir.

Şekil 2’de dizinin sonlarına doğru adaptör oranının yükseldiği görülmektedir. Bu yükselme, adaptörlerin genelde sonlara bağlandığından olabilir.

Bu adımlar sonrasında, veride bulunan adaptör dizilerini temizlemek için cutadapt isimli aracı kullandım, adaptör dizisi olarak ise, Şekil 2’de de görünen Illumina Universal Adapter dizisini kullandım.

Adaptör dizisi çıkartma işleminden sonra veriden toplamda 532bin adet baz çifti kesildi ve cutadapt’ten sonraki kalite kontrol grafikleri aşağıda görülebilir.



Şekil 3. Cutadapt işleminden sonra elde edilen adaptör içeriği grafiği

Şekil 3'te, dizinin sonuna bakıldığında, Illumina Universal Adapter dizilerinin kaldırıldığını görebiliriz. Bu, cutadapt işleminin başarıyla gerçekleştiğini göstermektedir.

Tartışma/Genel Yorum

Yapılan çalışmada, dizileme için çekirdek izolasyonu yapılmıştır. Çekirdek izolasyonu, hem genomik hem de transkriptomik analizler yapmaya yarayan ve başka yöntemlere göre daha fazla verim veren bir yöntemdir (Li Pengfei, vd., 2025).

Şekil 2'de adaptör dizilerinin yükseldiğini görülebilir, dizileme sırasında, DNA parçalarının uçlarına adaptör dizileri bağlanır. Adaptörler, dizileme makinesinin çalışması için gereklidir. Okuma bittiğinde ise bu adaptörlerin veriden çıkarılması gerekir. Adaptör dizileri veride hâlâ duruyorsa, doğru sonuçlar alınamayabilir. Bu adaptörlerin analiz edilecek veriden çıkartılmaması ise hatalı sonuçlar doğurabilir. Bu işlem için çokça araç geliştirilmiştir (Lindgreen S., vd., 2016).

Yapılan analizler ve adaptör dizilerinin kesilmesi sonucu, bu veride ileriki analizler için kullanılmaya engel bir durum görülmemiştir, ortalama baz kalitesi 30'un üstündedir. GC oranı ortalama düzeyde ve %41 civarlarındadır. Tüm bu belirtiler, verinin kaliteli olduğunu göstermektedir.

* Siz buraya, elde ettiğiniz grafiklerin yorumunu eklemelisiniz, daha detaylı tartışmalı ve hata varsa olası sebepleri, kaynak göstererek yazmalısınız. Kapanış cümlesi olarak da tüm veri hakkında yorumunuzu eklemelisiniz. *

Kaynakça

Li, P., Zhang, J., Liu, X., Wu, Z., Kang, Y. J., & Zhang, W. (2025). An improved cell nuclear isolation method. *Biology methods & protocols*, 10(1), bpaf007. <https://doi.org/10.1093/biomethods/bpaf007>

Schubert, M., Lindgreen, S. & Orlando, L. AdapterRemoval v2: rapid adapter trimming, identification, and read merging. *BMC Res Notes* 9, 88 (2016). <https://doi.org/10.1186/s13104-016-1900-2>