

Örnek RNA-Seq Analizi

Kaynak Makale: Functional transcriptomic analysis of the role of MAB-5/Hox in Q neuroblast migration in *Caenorhabditis elegans*. (Tamayo ve ark. 2013)

Gastrulasyon basamağından sonra görülmeye başlanan yöndendirilmiş hücre göçü, sağlıklı ve kanserli hücrelerin gelişimi açısından oldukça önemli bir basamaktır. Merkezi sinir sistemini oluşturan önemli hücrelerinden olan nöroblastlar farklı işlev ve görevleri olan hücrelere dönüşürler. Bu çalışmada Q nöroblastları çalışılmıştır. Q nöroblastlarının embriyodaki ilkin haline bakacak olursak embriyonun posterior kısmında bulunduklarını söyleyebiliriz. Normal embriyolarda bu şekilde posterior yönelim görülse de, mab-5 mutantlarında bunun tersi şekilde Q nöroblastlarının anterior kısma doğru yönelim gösterdikleri bilinmektedir. Hücrelerin bu yönelimleri florasan boyama teknikleriyle takip edilebilmektedir. Bu çalışmada yabancı *C. elegans* ile mab-5 mutantlarında tüm transkript sekansları karşılaştırılarak Q hücrelerinin göçünü etkileyen transkripsiyon faktörleri tespit edilmiştir. Kullanılan tipler aşağıda görülebilir:

Yabancı Form: LE2544

Fonksiyon Kazanmış Form: LE3037

Fonksiyon Kaybetmiş Form: LE2962, LE2963

Yöntem:

Temel Linux Komutları:

Bilgisayarlar üretildiğinden günümüze kadar olan süreç boyunca bilgisayar gücü ciddi bir artış göstermiştir. Eskiden bir oda büyüklüğünde olan bilgisayar sistemleri artık küçük bir usb belleğe sığabilecek kadar küçülmüştür ve kapasiteleri eski bilgisayarları katlamıştır. Genomik düzeydeki veriler de küçük gen bölgelerinin Sanger sekansları çalışılarak başlamışken, günümüzde çok daha gelişmiştir. Genomun içerisindeki bütün gen bölgeleri dizilenebilmektedir. Ancak kullandığımız kişisel bilgisayarın kapasiteleri bütün genomun çalışılabilmesi için yeterli elverişliliğe sahip değildir. Bu nedenle küme bilgisayar sistemleri biyoinformatik analizler için kullanılmaktadır. Bu sistemler de genel olarak Linux altyapısıyla çalışmaktadır. Bu nedenle linuxtaki temel bazı komutların bilinmesi gerekmektedir.

Windows 10 bilgisayarlarınızda linux alt sistemini çalıştırmak için <https://mobaxterm.mobatek.net/download.html> adresindeki yazılımı bilgisayarınıza yükleyiniz.

Macintosh bilgisayarlar zaten linux altsistemini çalıştırdığı için Terminal'i açarak uygun şekilde kullanabilirsiniz.

1. pwd - Terminali ilk açışınızda bilgisayarınızın ev dizininde olduğunuzdur. Bu komut hangi dizinde bulunduğunuz hakkında bilgi vermektedir.

```
[Bioasus@bioasus ~]$ pwd
/home/Bioasus
```

2. ls - Bulunduğunuz dizinin içinde hangi öğelerin bulunduğunu öğrenmek amacıyla kullanılmaktadır. Yanına -l, -t, -a gibi ekstra parametreler ekleyerek nasıl listemek istediğinizi ayarlayabilirsiniz. Ekstra parametreler hakkında daha fazla bilgi için şu adresi kullanabilirsiniz: http://linuxcommand.org/lc3_man_pages/ls1.html.

```
[Bioasus@bioasus ~]$ ls
'Calibre Library'  Desktop  Documents  Downloads  eduroam  gsl  igv  Music
```

```
[Bioasus@bioasus ~]$ ls -ltrh
total 476K
drwxr-xr-x.  2 Bioasus Bioasus 4.0K Oct 23 2018 Templates
drwxrwxr-x.  3 Bioasus Bioasus 4.0K Oct 25 2018 wisdom.aiFiles
drwxr-xr-x.  3 Bioasus Bioasus 4.0K Nov  7 2018 Public
drwxrwxr-x.  2 Bioasus Bioasus 4.0K Nov  8 2018 'VirtualBox VMs'
drwxrwxr-x.  2 Bioasus Bioasus 4.0K Feb  4 10:37 eduroam
drwxrwxr-x. 11 Bioasus Bioasus 4.0K Feb 15 14:04 'Calibre Library'
drwxrwxr-x.  3 Bioasus Bioasus 4.0K Feb 25 16:57 igv
drwxr-xr-x.  4 Bioasus Bioasus 4.0K Apr 30 14:04 Projects
drwxr-xr-x.  3 Bioasus Bioasus 4.0K May  1 17:19 Music
-rw-r--r--.  1 Bioasus Bioasus 380K May  6 14:06 symbology-style.db
drwxrwxr-x.  5 Bioasus Bioasus 4.0K May 20 20:58 perl5
drwxr-xr-x.  3 Bioasus Bioasus 4.0K May 21 11:11 gsl
drwxrwxr-x.  4 Bioasus Bioasus 4.0K May 21 11:39 R
drwxr-xr-x.  7 Bioasus Bioasus 4.0K May 28 17:00 Zotero
drwxr-xr-x. 15 Bioasus Bioasus 4.0K Jun 11 19:37 Documents
drwxr-xr-x.  6 Bioasus Bioasus 4.0K Jun 13 08:51 Desktop
drwxr-xr-x. 17 Bioasus Bioasus 4.0K Jun 15 15:58 Videos
drwxrwxr-x.  2 Bioasus Bioasus 4.0K Jun 16 21:58 vpn
drwxr-xr-x.  4 Bioasus Bioasus 12K Jun 17 14:17 Downloads
drwxr-xr-x.  3 Bioasus Bioasus 12K Jun 17 15:04 Pictures
```

3. cd - Yeni bir dizinin içerisine girmek için bu komut kullanılır. Örneğin Downloads dizininin içinde aşağıdaki gibi girebilirsiniz. Gireceğiniz dizinin isminde boşluk varsa boşluk yerine \ işaretini kullanmanız gerekmektedir.

```
[Bioasus@bioasus ~]$ cd Downloads/
```

```
[Bioasus@bioasus ~]$ cd bosluk\ varsa
```

4. mkdir – Yeni bir dizin oluşturmak için kullanılır. Yanlışlıkla oluşturduğunuz içi boş dizini silmek için rmdir komutu kullanılır.

```
[Bioasus@bioasus ~]$ mkdir yenedizin
```

```
[Bioasus@bioasus ~]$ rmdir yenedizin
```

5. rm – Dosya silmek için kullanılır. Dikkatli bir şekilde kullanılmalıdır. Çünkü sildikten sonra geri dönüşü yoktur. Klasör silmek için rm -r şeklinde kullanılmalıdır.

```
[Bioasus@bioasus ~]$ rm dosya
```

```
[Bioasus@bioasus ~]$ rm -r dizin
```

6. touch – Yeni bir dosya oluşturmak için kullanılır. Herhangi bir dosya formatında olabilir.

```
[Bioasus@bioasus Dizin]$ touch dosya.txt
```

```
[Bioasus@bioasus Dizin]$ ls
```

```
dosya.txt
```

7. man & --help & -h – Kullanılan fonksiyonun kitapçığını/yardım sayfasını görüntülemek için kullanılır. Bazen yalnızca man, bazen -help, bazense -h kullanılabilir. man yeni bir sayfa üzerinden açılır ve çıkmak için q'ya basılır.

```
[Bioasus@bioasus Dizin]$ man mkdir
```

```
[Bioasus@bioasus Dizin]$ mkdir --help
```

```
Usage: mkdir [OPTION]... DIRECTORY...
```

```
Create the DIRECTORY(ies), if they do not already exist.
```

```
Mandatory arguments to long options are mandatory for short options too.
```

```
-m, --mode=MODE    set file mode (as in chmod), not a=rwx - umask
```

```
-p, --parents       no error if existing, make parent directories as needed
```

```
-v, --verbose       print a message for each created directory
```

```
-Z                 set SELinux security context of each created directory  
to the default type
```

```
--context[=CTX]   like -Z, or if CTX is specified then set the SELinux  
or SMACK security context to CTX
```

```
--help            display this help and exit
```

```
--version         output version information and exit
```

```
GNU coreutils online help: <https://www.gnu.org/software/coreutils/>
```

```
Full documentation <https://www.gnu.org/software/coreutils/mkdir>
```

```
or available locally via: info '(coreutils) mkdir invocation'
```

```
[Bioasus@bioasus Dizin]$ mkdir -h
```

```
mkdir: invalid option -- 'h'
```

```
Try 'mkdir --help' for more information.
```

8. cp – Dosyaları yeni bir dizine kopyalamak için kullanılır. Bu komuttan sonra kopyalanacak dosya adı ve sonrasında kopyalanacağı dizin yazılmalıdır. Klasörleri kopyalamak için peşine -r parametresi eklenmelidir.

```
[Bioasus@bioasus Dizin]$ cp dosya.txt Gonderilecek/Dizin
```

```
[Bioasus@bioasus Dizin]$ cp -r Klasor/ Gonderilecek/Dizin
```

9. mv – Dosyayı bir yerden kesip başka bir yere yapıştırmak için kullanılır. Komutun ardından dosyanın ismi sonrasındaysa gönderilecek dizin yazılır. Gönderilecek dizin yerine başka bir dosya ismi yazılırsa dosyanın ismini değiştirir.

```
[Bioasus@bioasus Dizin]$ ls
Dizin dosya.txt
[Bioasus@bioasus Dizin]$ mv dosya.txt Dizin/
[Bioasus@bioasus Dizin]$ ls
Dizin
[Bioasus@bioasus Dizin]$ ls Dizin/
dosya.txt
```

```
[Bioasus@bioasus Dizin]$ mv dosya.txt yenedosta.txt
[Bioasus@bioasus Dizin]$ ls
yenedosta.txt
```

10. echo – Dosyaya veri yazmak için kullanılır. Genellikle metin dosyaya aktarılır. Komuttan sonra yazılacak metin, daha sonra >> işareti ve dosya ismi yazılmalıdır.

```
[Bioasus@bioasus Dizin]$ echo Temel linux komutlarını öğrendim. >> dosya.txt
```

11. cat – Dosyanın içeriğini görüntülemek için kullanılır.

```
[Bioasus@bioasus Dizin]$ cat dosya.txt
Temel linux komutlarını öğrendim.
```

Makalede Kullanılan Yöntemlerin Tekrarı:

RNAseq sonrası elde edilen sekansları indirmek için makalede verilen SRP017621 proje numarasını NCBI SRA veritabanı üzerinden aratmanız gerekmektedir. “<https://www.ncbi.nlm.nih.gov/sra>” internet adresinde arama kısmına proje numarasını yazarak çalışmada kullanılan bütün sekans verilerinin ham hallerine ulaşabilirsiniz. Bu sekans verilerini indirmek için SRA Toolkit’in bir komutu olan “fastq-dump” komutunu kullanabilirsiniz (SRA Toolkit’i yüklemek için “https://trace.ncbi.nlm.nih.gov/Traces/sra/sra.cgi?view=toolkit_doc&f=std” adresindeki yönergeleri takip edebilirsiniz.) Bu komutun nasıl çalıştığı hakkında bilgi edinmek için aşağıdaki komutu kullanabilirsiniz (-bash-4.2\$’dan sonrasını komut olarak girmeniz gerekmektedir.):

```
(base) -bash-4.2$ fastq-dump -h
Usage:
  /truba/home/celverici/Admin/Tools/sratools/sratoolkit.2.9.1-1-centos_linux64/bin/fastq-dump [options] <path> [<path>...]
  /truba/home/celverici/Admin/Tools/sratools/sratoolkit.2.9.1-1-centos_linux64/bin/fastq-dump [options] <accession>

INPUT
  -A|--accession <accession>      Replaces accession derived from <path> in
                                   filename(s) and defines (only for single
                                   table dump)
  --table <table-name>            Table name within cSRA object, default is
                                   "SEQUENCE"

PROCESSING
Read Splitting                    Sequence data may be used in raw form or
  --split-spot                    split into individual reads
                                   Split spots into individual reads
```

Verileri indirmeden önce lütfen kendinize uygun bir klasör belirleyip bu klasörün içerisinde çalışmanızı yapınız. Klasör açmak için “mkdir” komutunu, klasör içerisine girmek için “cd” komutunu ve içerisindeki öğeleri listelemek için “ls” komutunu kullanabilirsiniz.

Veritabanında görebileceğiniz üzere LE544, LE3037, LE2962 ve LE2963 isimli suşlar ve yapılan farklı tekrarların sonuçları toplamda 26 farklı veri olarak veritabanına girilmiştir. Bunları indirmek için SRR etiketlerini kullanmanız gerekmektedir.

- ☐ [C. elegans loss-of-function strain LE2962 \(mab-5\(e1239\) III; lqls80 IV; lqls58 V\), early L1 larvae transcriptome \(replicate library 3, lane 1\)](#)
1 ILLUMINA (Illumina Genome Analyzer IIx) run: 33.6M spots, 2G bases, 1Gb downloads
Accession: SRX212107
- ☐ [C. elegans gain-of-function strain LE3037 \(mab-5\(e1751\) III; lqls80 IV; lqls58 V\), early L1 larvae transcriptome \(replicate library 3, lane 2\)](#)
1 ILLUMINA (Illumina Genome Analyzer IIx) run: 30.2M spots, 1.8G bases, 912.5Mb downloads
Accession: SRX212122
- ☐ [C. elegans gain-of-function strain LE3037 \(mab-5\(e1751\) III; lqls80 IV; lqls58 V\), early L1 larvae transcriptome \(replicate library 3, lane 1\)](#)
1 ILLUMINA (Illumina Genome Analyzer IIx) run: 29M spots, 1.7G bases, 894.1Mb downloads
Accession: SRX212121
- ☐ [C. elegans gain-of-function strain LE3037 \(mab-5\(e1751\) III; lqls80 IV; lqls58 V\), early L1 larvae transcriptome \(replicate library 2, lane 2\)](#)
1 ILLUMINA (Illumina Genome Analyzer IIx) run: 31.7M spots, 1.9G bases, 949.9Mb downloads
Accession: SRX212120
- ☐ [C. elegans gain-of-function strain LE3037 \(mab-5\(e1751\) III; lqls80 IV; lqls58 V\), early L1 larvae transcriptome \(replicate library 2, lane 1\)](#)
1 ILLUMINA (Illumina Genome Analyzer IIx) run: 34.8M spots, 2.1G bases, 1.1Gb downloads
Accession: SRX212119

SRX212107: C. elegans loss-of-function strain LE2962 (mab-5(e1239) III; lqls80 IV; lqls58 V), early L1 larvae transcriptome (replicate library 3, lane 1)
1 ILLUMINA (Illumina Genome Analyzer IIx) run: 33.6M spots, 2G bases, 1Gb downloads

Submitted by: University of Kansas

Study: Using RNA-seq to identify transcriptional targets of MAB-5 in C. elegans

[PRJNA184024](#) • [SRP017621](#) • [All experiments](#) • [All runs](#)

[show Abstract](#)

Sample: Caenorhabditis elegans strain LE2962; Genotype mab-5(e1239) III; lqls80 IV; lqls58 V; Loss-of-function mab-5 strain; Transcriptome from L1 larvae

[SAMN01831652](#) • [SRS380992](#) • [All experiments](#) • [All runs](#)

Organism: [Caenorhabditis elegans](#)

Library:

Instrument: Illumina Genome Analyzer IIx

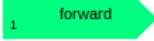
Strategy: RNA-Seq

Source: TRANSCRIPTOMIC

Selection: RANDOM

Layout: SINGLE

Spot descriptor:



Runs: 1 run, 33.6M spots, 2G bases, [1Gb](#)

Run	# of Spots	# of Bases	Size	Published
SRR639089	33,628,057	2G	1Gb	2015-07-22

SRR etiketi yukarıdaki şekilde gördüğünüz SRR639089 numaralı kısımdır. İndirmek için ise;

```
(base) -bash-4.2$ fastq-dump -I SRR639105
```

komutunu kullanabilirsiniz. Bu komut ile LE2962 suşunu indirmektesiniz. Kalan suşlar için kendiniz uygun etiketleri seçip indiriniz. İndirmeleriniz sonucunda aşağıdaki gibi dosyalarınız olmalıdır. Lütfen indirmiş olduğunuz bu fastq dosyalarının hangi suşlara ait olduğunu not alınız. Örneğin aşağıda SRR639097 – LE3037 suşuna aittir. Diğerleri için de suşları belirleyiniz.

```
(base) -bash-4.2$ ls -ltrh
total 25G
-rw-rw-r-- 1 celverici celverici 6.6G Jun 17 11:36 SRR639097.fastq
-rw-rw-r-- 1 celverici celverici 6.6G Jun 17 11:37 SRR639082.fastq
-rw-rw-r-- 1 celverici celverici 6.9G Jun 17 11:39 SRR639091.fastq
-rw-rw-r-- 1 celverici celverici 4.8G Jun 17 11:39 SRR638942.fastq
```

Zamandan ve veri büyüklüğünden tasarruf etmek için, bu çalışmada her bir örnek için bir tane örnek kullanacağız. Ancak makalede veritabanına yüklenen 26 verinin hepsi kullanılmıştır. Bu durumda biz 4 farklı örnek kullanacağız.

Şimdi elimizde 4 farklı suşa ait okumalar var. Ancak yeni jenerasyon sekanslamada bu dizileri haritalayacağımız referans genomu ihtiyacımız var. Bu nedenle *C. elegans*'a ait referans genomu indirmemiz gerekmektedir. Referans genomu indirmek için şu adresi kullanabilirsiniz:

ftp://igenome:G3nom3s4u@usd-ftp.illumina.com/Caenorhabditis_elegans/NCBI/WS195/Caenorhabditis_elegans_NCBI_WS195.tar.gz .

Bu dosyayı kümeye indirmek için ise wget komutunu kullanabilirsiniz.

```
(base) -bash-4.2$ wget ftp://igenome:G3nom3s4u@usd-ftp.illumina.com/Cae
```

Dosya indikten sonra elinizde sıkıştırılmış formatta tar.gz uzantılı bir dosya olacak. Bu sıkıştırılmış dosyayı açmak için tar komutunu aşağıdaki parametreleri kullanarak çalıştırabilirsiniz.

```
(base) -bash-4.2$ tar -xvf Caenorhabditis_elegans_NCBI_WS195.tar.gz
```

Bu dosyayı açıp, oluşan klasörün içine girdikten sonra ../Sequence/Bowtie2Index klasörünün içine giriniz. Bu klasörün içinde aşağıdaki dosyaları görmelisiniz.

```
(base) -bash-4.2$ ls -l
total 146924
-rwxrwxr-x 1 celverici celverici 37617153 Jun 17 12:00 genome.1.bt2
-rwxrwxr-x 1 celverici celverici 25066916 Jun 17 12:00 genome.2.bt2
-rwxrwxr-x 1 celverici celverici 62 Jun 17 12:00 genome.3.bt2
-rwxrwxr-x 1 celverici celverici 25066909 Jun 17 12:00 genome.4.bt2
lrwxrwxrwx 1 celverici celverici 29 Jun 17 12:00 genome.fa -> ../WholeGenomeFasta/genome.fa
-rw-rw-r-- 1 celverici celverici 153 Jun 17 12:00 genome.fa.fai
-rwxrwxr-x 1 celverici celverici 37617153 Jun 17 12:00 genome.rev.1.bt2
-rwxrwxr-x 1 celverici celverici 25066916 Jun 17 12:00 genome.rev.2.bt2
```

Şimdi bu dosyaları SRRXXXXXX.fastq dosyalarınızın olduğu klasöre taşıyınız. Bunun için mv komutunu kullanabilirsiniz (* işareti bütün dosyaları seçmenizi sağlamaktadır.). Ardından genomları referans genom üzerine haritalamak için tophat programını kullanacağız.

Tophat programını kurmak için programlarınızı yüklediğiniz klasöre gelin ve wget komutuyla programı kümeye indirin (Bu yazının yazıldığı tarihte program şuradan indirilmekteydi: https://ccb.jhu.edu/software/tophat/downloads/tophat-2.1.1.Linux_x86_64.tar.gz).

```
(base) -bash-4.2$ wget https://ccb.jhu.edu/software/tophat/downloads/tophat-2.1.1.Linux_x86_64.tar.gz
```

İndirdiğiniz dosyayı tar komutuyla, yukarıda yaptığınıza benzer şekilde, açabilirsiniz. Konsola tophat yazdığınız zaman programın çalışmasını istiyorsanız, cd komutuyla kök klasörünüze dönüp .bashrc dosyasını değiştirmeniz gerekmektedir. Bunun için emacs .bashrc komutunu kullanabilirsiniz. Emacs programı açıldığı zaman kendi programınızın yolağını aşağıdakine benzer şekilde belirttiğinizde;

```
alias tophat="/truba/home/celverici/Admin/Tools/tophat-2.1.1/tophat"
```

ardından da “ctrl x” + “ctrl s” ve “ctrl x” + “ctrl c” yaptığınızda kaydınızı yapmış olursunuz. Sonrasındaysa “source .bashrc” komutuyla .bashrc’yi çağırabilirsiniz.

Genom üzerine haritalama yapmak için tophat programını aşağıdaki gibi çalıştırabilirsiniz. Her okuma sonucunda oluşan tophat_out klasörünün ismini sonradan karışıklık oluşturmaması için değiştirmeniz gerekmektedir.

```
(base) -bash-4.2$ tophat --no-novel-juncs genome SRR639082.fastq
```

Program uzun bir süre çalıştıktan sonra oluşturduğu yeni klasörde şu sonuçları göstermektedir. Sonuçları bu şekilde göremiyorsanız bir hata olmuş olabilir.

```
(base) -bash-4.2$ ls -ltrh
total 1.1G
-rw-rw-r-- 1 celverici celverici 70 Jun 17 13:10 prep_reads.info
-rw-rw-r-- 1 celverici celverici 275M Jun 17 14:48 unmapped.bam
drwxrwxr-x 2 celverici celverici 4.0K Jun 17 14:48 logs
-rw-rw-r-- 1 celverici celverici 52 Jun 17 14:48 junctions.bed
-rw-rw-r-- 1 celverici celverici 770K Jun 17 14:48 insertions.bed
-rw-rw-r-- 1 celverici celverici 993K Jun 17 14:48 deletions.bed
-rw-rw-r-- 1 celverici celverici 199 Jun 17 14:48 align_summary.txt
-rw-rw-r-- 1 celverici celverici 770M Jun 17 14:54 accepted_hits.bam
```

Gen ifadeleri arasındaki farklılıkları bulmak için ise Cufflinks programının bir parçası olan Cuffdiff komutunu kullanabilirsiniz. Program hakkında detaylı bilgi için <http://cole-trapnell-lab.github.io/cufflinks/cuffdiff/> adresini kullanabilirsiniz. Cufflinks programını kümeye yüklemek için wget komutuyla aşağıdaki adresten paketi indirebilirsiniz:

http://cole-trapnell-lab.github.io/cufflinks/assets/downloads/cufflinks-2.2.1.Linux_x86_64.tar.gz

Tophat programını yüklerken izlediğiniz aşamaları takip ederek Cufflinks programını da kümeye yükleyebilirsiniz. Son kısımda .bashrc'yi düzenlerken cuffdiff alias'ı vermeyi unutmamanız gerekmektedir.

Cuffdiff fonksiyonunu kullanmadan önce kullanacağınız dosyaları bulunduğunuz klasöre taşımanız gerekmektedir. Kullanacağımız dosyalar referans genom içinde bulunan genes.gtf dosyası, önceden klasörümüze taşıdığımız genome.fa dosyası ve Tophat'in oluşturduğu accepted_hits.bam dosyalarıdır. genes.gtf dosyasını referans'ın içinden annotation klasöründen bulabilirsiniz. Buradan çalıştığınız klasöre kopyalayabilirsiniz (cp fonksiyonu). Tophat'in oluşturduğu bütün output dosyalarında aynı accepted_hits.bam dosyası oluşmuş durumdadır. Bu yüzden her birini tekrardan isimlendirmeniz gerekmektedir. Başlarına suş numaralarını yazıp isimlerini değiştirebilirsiniz (Örneğin LE2544_accepted_hits.bam). Yeniden isimlendirdikten sonra bulunduğunuz klasöre kopyalamanız gerekmektedir. Bunların dışında cuffdiff fonksiyonunun çıktılarını yollayacağı klasörü de önceden belirleyebilirsiniz. Bu amaçla mkdir komutuyla CD_out veya kendi belirlediğiniz bir isimde klasör oluşturabilirsiniz.

Cuffdiff fonksiyonunu çalıştırırken bulunduğunuz klasörde aşağıdaki dosyaların bulunması gerekmektedir:

```
(base) -bash-4.2$ ls
CD_out      genome.1.bt2  genome.3.bt2  genome.fa      genome.rev.1.bt2  LE2544_accepted_hits.bam  LE2963_accepted_hits.bam
genes.gtf   genome.2.bt2  genome.4.bt2  genome.fa.fai  genome.rev.2.bt2  LE2962_accepted_hits.bam  LE3037_accepted_hits.bam
```

Cuffdiff fonksiyonunu ise aşağıdaki şekilde çalıştırıp makalede ulaşılan sonuçlara ulaşabilirsiniz.

```
cuffdiff -b genome.fa -u genes.gtf LE2544_accepted_hits.bam LE2963_accepted_hits.bam LE2963_accepted_hits.bam LE3037_accepted_hits.bam -o CD_out
```