# Bioinformatics Overview 

From Raw Sequence to Gene Counts

## Rna-Seq Applications

- Differential Expression
- Transcriptome
- Genome Annotation
- "Bargain Exome"
- SNPs
- Gene Fusions


## In a Nutshell

- Input: RNA Sequence Fragments
- Output: Gene Counts of Fragments


# Rna-Seq Analysis 

A Biblio-Archeologic Analogy

## Fragments of a Library

ng on nights, it's us; we're going to set you free." Jim only had time to grab us by the hand and s of this agreement. See paragraph 1.C below. There are a lot of things you can do with Project Guten with his head!"' 'How dreadfully savage!' exclaimed Alice. 'And ever since that,' the Hatter went g, that chateau of Monsieur the Marquis, with a large stone courtyard before it, and two stone sweep he nursery. And he was indeed crying. She heard him and hastened. But the faster she went, the loude $h$ to intensify the influence of her gaze. "Yes, they draw away all the sap and give a false appeara him. I could not look at him without feeling sorry for him. We both know him. He's good-hearted, ed, under various disguises of Art, through the portraits of every Drinking Age. "You are a little eupon Punch Costello dinged with his fist upon the board and would sing a bawdy catch _Staboo Stabel him to act in anything without her quiet aid), and the day passed quickly. Early in the evening he e ed. 'Give your evidence,' said the King; 'and don't be nervous, or I'll have you executed on the sp harge a fee for access to, viewing, displaying, performing, copying or distributing any Project Gute ef secretary. She knew that he usually went out quickly to his office, and she wanted to see him bef onsiderable. Then he said he must start in and "'terpret" it, because it was sent for a warning. ? ides her body's flaws calling under her brown shawl from an archway where dogs have mired. Her fancy recatory manner, "the anguish of his daughter, which must be a dreadful anguish to him!" "I have ob uld be like, but it puzzled her too much, so she went on: 'But why did they live at the bottom of a Indivisible. Liberty, Equality, Fraternity, or Death! Who could that be with Mr. Lorry--the owner ns in general, and I believe that's just why philanthropic institutions always give such poor resu here were no proper cupboards for their clothes; what cupboards there were either would not close at he detected all around him, walked from one to another. The first was the best room, and in it were $t$. All her arrangements had to be modified because they could not be carried out, and they were modi er hand shook more violently, but she did not take her eyes off him, watching how he would take it. s." It was done. "Well?" "Monseigneur, it is nothing. The trees and the night are all that are he spectacles offered by our streets, hideous publicity posters, religious ministers of all denominatio sore, but I couldn't see no way out of the trouble. After all this long journey, and after all we she got to pumping me about England, and blest if I didn't think the ice was getting mighty thin so the morning all sound and right. So she had to be satisfied. But she said she'd set up for him prison--seemed to strike across the earth, messieurs, to where the sky rests upon it!" The hungry m oys running, playing at horses. Seryozha! And I'm losing everything and not getting him back. Yes,

## Mapped Fragments

na on niahts it's sus: we're aoina to set vou free " lim onlv had time to arab us hy the hand and s.--471971 Huckleberrv Finn of this agreement. See paragraph 1.C below. There are a lot of things you can do with Project Guten -- ?????????????

With his head! How dreadfulty savage! exclaimed Alice. And ever since that, the Hatter went - $/ 8968$ Allce's Adventures $g$, that chateau of Monsieur the Marquis, with a large stone courtyard before it, and two stone sweep -- 232648 A Tale of Two Cities he nursery. And he was indeed crying. She heard him and hastened. But the faster she went, the loude -- 1897819 Anna Karenina $h$ to intensify the influence of her gaze. "Yes, they draw away all the sap and give a false appeara -- 1186098 Anna Karenina him. I could not look at him without feeling sorry for him. We both know him. He's good-hearted, -- 165528 Anna Karenina ed, under various disguises of Art, through the portraits of every Drinking Age. "You are a little -- 168896 A Tale of Two Cities eupon Punch Costello dinged with his fist upon the board and would sing a bawdy catch _Staboo Stabel -- 803339 Ulysses him to act in anything without her quiet aid), and the day passed quickly. Early in the evening he e-- 490073 A Tale of Two Cities ed. 'Give your evidence, said the King: 'and don't be nervous, or I'll have you executed on the sp -- 127701 Alice's Adventures harge a fee for access to, viewing, displaying, performing, copying or distributing any Project Gute -- ?????????????
'
onsiderable. Then he said he must start in and "'terpret" it, because it was sent for a warning. ? -- 176566 Huckleberry Finn ides her body's flaws calling under her brown shawl from an archway where dogs have mired. Her fancy -- 90509 Ulysses recatory manner, "the anguish of his daughter, which must be a dreadful anguish to him!" "I have ob -- 682808 A Tale of Two Cities uld be like, but it puzzled her too much, so she went on: 'But why did they live at the bottom of a -- 80787 Alice's Adventures Indivisible. Liberty, Equality, Fraternity, or Death! Who could that be with Mr. Lorry--the owner -- 562830 A Tale of Two Cities ns in general, and I believe that's just why philanthropic institutions always give such poor resu -- 1699588 Anna Karenina here were no proper cupboards for their clothes; what cupboards there were either would not close at -- 634360 Anna Karenina he detected all around him, walked from one to another. The first was the best room, and in it were -- 182459 A Tale of Two Cities $t$. All her arrangements had to be modified because they could not be carried out, and they were modi -- 1252025 Anna Karenina er hand shook more violently, but she did not take her eyes off him, watching how he would take it. -- 455094 Anna Karenina s." It was done. "Well?" "Monseigneur, it is nothing. The trees and the night are all that are he -- 235828 A Tale of Two Cities spectacles offered by our streets, hideous publicity posters, religious ministers of all denominatio -- 870689 Ulysses
sore, but I couldn't see no way out of the trouble. After all this long journey, and after all we -- 422461 Huckleberry Finn she got to pumping me about England, and blest if I didn't think the ice was getting mighty thin so -- 345525 Huckleberry Finn the morning all sound and right. So she had to be satisfied. But she said she'd set up for him -- 558004 Huckleberry Finn prison--seemed to strike across the earth, messieurs, to where the sky rests upon it!" The hungry m -- 340478 A Tale of Two Cities oys running, playing at horses. Seryozha! And I'm losing everything and not getting him back. Yes, -- 1837882 Anna Karenina

## Counting

ng on nights, it's us; we're going to set you free." Jim only had time to grab us by the hand and s -- 471971 with his head!"' 'How dreadfully savage!' exclaimed Alice. 'And ever since that,' the Hatter went -- 78968 $g$, that chateau of Monsieur the Marquis, with a large stone courtyard before it, and two stone sweep -- 232648 A Tale of Two Cities he nursery. And he was indeed crying. She heard him and hastened. But the faster she went, the loude -- 1897819 Anna Karenina $h$ to intensify the influence of her gaze. "Yes, they draw away all the sap and give a false appeara -- 1186098 Anna Karenina him. I could not look at him without feeling sorry for him. We both know him. He's good-hearted, -- 165528 Anna Karenina ed, under various disguises of Art, through the portraits of every Drinking Age. "You are a little -- 168896 A Tale of Two Cities eupon Punch Costello dinged with his fist upon the board and would sing a bawdy catch _Staboo Stabel -- 803339 him to act in anything without her quiet aid), and the day passed quickly. Early in the evening he e -- 490073 ed. 'Give your evidence,' said the King; 'and don't be nervous, or I'll have you executed on the sp -- 127701 ef secretary. She knew that he usually went out quickly to his office, and she wanted to see him bef -- 782396 onsiderable. Then he said he must start in and "'terpret" it, because it was sent for a warning. ? -- 176566 ides her body's flaws calling under her brown shawl from an archway where dogs have mired. Her fancy -- 90509 recatory manner, "the anguish of his daughter, which must be a dreadful anguish to him!" "I have ob -- 682808 uld be like, but it puzzled her too much, so she went on: 'But why did they live at the bottom of a -- 80787 Indivisible. Liberty, Equality, Fraternity, or Death! Who could that be with Mr. Lorry--the owner -- 562830 ns in general, and I believe that's just why philanthropic institutions always give such poor resu -- 1699588 here were no proper cupboards for their clothes; what cupboards there were either would not close at -- 634360 he detected all around him, walked from one to another. The first was the best room, and in it were -- 182459 $t$. All her arrangements had to be modified because they could not be carried out, and they were modi -- 1252025 er hand shook more violently, but she did not take her eyes off him, watching how he would take it. -- 455094 s." It was done. "Well?" "Monseigneur, it is nothing. The trees and the night are all that are he -- 235828 spectacles offered by our streets, hideous publicity posters, religious ministers of all denominatio -- 870689 sore, but I couldn't see no way out of the trouble. After all this long journey, and after all we -- 422461 she got to pumping me about England, and blest if I didn't think the ice was getting mighty thin so -- 345525 the morning all sound and right. So she had to be satisfied. But she said she'd set up for him -- 558004 prison--seemed to strike across the earth, messieurs, to where the sky rests upon it!" The hungry m -- 340478 oys running, playing at horses. Seryozha! And I'm losing everything and not getting him back. Yes, -- 1837882 Anna Karenina

## Counting

ed, under various disguises of Art, through the portraits of every Drinking Age. "You are a little -- 168896 A Tale of Two Cities he detected all around him, walked from one to another. The first was the best room, and in it were -- 182459 A Tale of Two Cities g, that chateau of Monsieur the Marquis, s." It was done. "Well?" "Monseigneur prison--seemed to strike across the eart him to act in anything without her quiet Indivisible. Liberty, Equality, Fratern recatory manner, "the anguish of his dau onsiderable. Then he said he must start she got to pumping me about England, an sore, but I couldn't see no way out of ng on nights, it's us; we're going to se the morning all sound and right. So sh
with his head!"' 'How dreadfully savag uld be like, but it puzzled her too much ed. 'Give your evidence,' said the King
him. I could not look at him without fe er hand shook more violently, but she di here were no proper cupboards for their ef secretary. She knew that he usually $w$ $h$ to intensify the influence of her gaze $t$. All her arrangements had to be modifi ns in general, and I believe that's just oys running, playing at horses. Seryozha he nursery. And he was indeed crying. Sh
ides her body's flaws calling under her
 eupon Punch costed dinged with his fise upplicity posters, religious ministers of all denominatio -- 870689 Ulysses

## Short Reads

```
ng on nights, it's us; we're going to set you free." Jim only had time to grab us by the hand and s
of this agreement. See paragraph 1.C below. There are a lot of things you can do with Project Guten
with his head!"" "How dreadfulLy savage!" exclaimed Alice. "And ever since that," the Hatter went
g, that chateau of Monsieur the Marquis, with a Large stone courtyard before it, and two stone sweep
he nursery. And he was indeed crying. She heard him and hastened. But the faster she went, the loude
h to intensify the influence of her gaze. "Yes, they draw away alL the sap and give a false appeara
    him. I could not look at him without feeling sorry for him. We both know him. He's good-hearted,
ed, under various disguises of Art, through the portraits of every Drinking Age. "You are a Little
eupon Punch CosteLLo dinged with his fist upon the board and would sing a bawdy catch _Staboo Stabel
him to act in anything without her quiet aid), and the day passed quickly. Early in the evening he e
ed. "Give your evidence," said the King; "and don't be nervous, or I'LL have you executed on the sp
harge a fee for access to, viewing, displaying, performing, copying or distributing any Project Gute
ef secretary. She knew that he usually went out quickly to his office, and she wanted to see him bef
onsiderable. Then he said he must start in and" "terpret" it, because it was sent for a warning. ?
ides her body's flaws calLing under her brown shawl from an archway where dogs have mired. Her fancy
recatory manner, "the anguish of his daughter, which must be a dreadful anguish to him!" "I have ob
uld be Like, but it puzzled her too much, so she went on: "But why did they live at the bottom of a
    Indivisible. Liberty, Equality, Fraternity, or Death! Who could that be with Mr. Lorry--the owner
ns in general, and I believe that's just why philanthropic institutions always give such poor resu
here were no proper cupboards for their clothes; what cupboards there were either would not close at
he detected alL around him, walked from one to another. The first was the best room, and in it were 
t. All her arrangements had to be modified because they could not be carried out, and they were modi
er hand shook more violently, but she did not take her eyes off him, watching how he would take it.
s." It was done. "WeLL?" "Monseigneur, it is nothing. The trees and the night are alL that are he
spectacles offered by our streets, hideous publicity posters, religious ministers of all denominatio
    sore, but I couldn't see no way out of the trouble. After all this long journey, and after all we
    she got to pumping me about England, and blest if I didn't think the ice was getting mighty thin so
    the morning alL sound and right. So she had to be satisfied. But she said she"d set up for him
prison--seemed to strike across the earth, messieurs, to where the sky rests upon it!" The hungry m
oys running, playing at horses. Seryozha! And I'm losing everything and not getting him back. Yes,
```


## Paired-End Reads

ng on nights, it's us; we're going to set you free." Jim only had time to grab us by the hand and s
of this agreement. See paragraph 1.C below. There are a lot of things you can do with Project Guten
with his head!"' 'How dreadfully savage!' exclaimed Alice. 'And ever since that,' the Hatter went
g, that chateau of Monsieur the Marquis, with a large stone courtyard before it, and two stone sweep
he nursery. And he was indeed crying. she heard him and mastened. but the faster she went, the loude
h to intensify the influence of her gaze. "Yes, they draw away all the sap and give a false appeara
him. I could not look at him without feeling sorry for him. We both know him. He's good-hearted,
ed, under various disguises of Art, through the portraits of every Drinking Age. ..... ttle
eupon Punch Costello dinged with his fist upon the board and would sing a bawdy catch _Staboo tabel
him to act in anything without her quiet aid), and the day passed quickly. Early in the evening he e
ed. 'Give your evidence,' said the King; 'and don't be nervous, or I'Ll have you executed on the sp
harge a fee for access to, viewing, displaying, performing, copying or distributing any Project Gute
ef secretary. She knew that he usually went out quickly to his office, and she wanted to see him bef
onsiderable. Then he said he must start in an ..... g. ?
ides her body's flaws calling under her brown shawl from an archway where dogs have mired. Her fancy
recatory manner, "the anguish of his daughter, which must be a dreadful anguish to him! ..... ve ob
uld ..... of $a$
Indivisible. Liberty, Equality, Fraternity, or Death ..... wner
ns in general, and I believe that's just why philanthropic institutions always give such poor ..... esu
here
he detected all around him, walked from one to anothe ..... were
t. All modi
er han but she d ..... it.
s." It was done. "Well?" "Monseigneur, it is nothing. The trees and the night are all that are he
spectacles offered by our streets, hideous publicity posters, religious ministers of all denominatio
sore, but I couldn't see no way out of the trouble. After all this long journey, and after all we
she got to pumping me about England, and blest if I didn't think the ice was getting mighty thin so
the morning all sound and right. So she had to be satisfied. But she said she'd set up for him
prison--seemed to strike across the earth, messieurs, to where the sky rests upon it!" The hungry m oys running, playing at horses. Seryozha! And I'm losing everything and not getting him back. Yes,

## RNA-Seq for <br> Differential Expression Analysis

| Step | Input | Output | Output Format |
| :---: | :---: | :---: | :---: |
| Pre-processing <br> and QC | Raw Reads | Processed <br> Reads | FASTQ |
| Mapping | Reads | Mapped <br> Reads | SAM/BAM |
| Counting | Mapped Reads | Count Table | CSV |


| Step | Input | Output | Output Format |
| :---: | :---: | :---: | :---: |
| Pre-processing <br> and QC | Raw Reads | Processed <br> Reads | FASTQ |
| Mapping | Reads | Mapped <br> Reads | SAM/BAM |
| Counting | Mapped Reads | Count Table | CSV |

## FASTQ

- FASTA with Quality
- https://en.wikipedia.org/wiki/FASTQ format


## FASTQ: Filenames

"What's in a name? that which we call a FASTQ
By any other name would smell as sweet"

- Romeo and Juliet, William Shakespeare


## FASTQ: Filenames

1C_TAAGGCGA_L006_R1_001.fastq.gz 1C_TAAGGCGA_L006_R1_002.fastq.gz 1C_TAAGGCGA_L006_R1_003.fastq.gz 1C_TAAGGCGA_L006_R1_004.fastq.gz 1C_TAGGCATG_L003_R1_001.fastq.gz 1C_TAGGCATG_L005_R1_001.fastq.gz 2C_CGTACTAG_L003_R1_001.fastq.gz

## FASTQ: Filenames

| Sample | Barcode | Lane |
| :---: | :---: | :---: | :---: | :---: |
| Name |  | Read |
| $\#$ | File <br> Part |  |

1C_TAAGGCGA_L006_R1_001.fastq.gz 1C_TAAGGCGA_L006_R1_002.fastq.gz 1C_TAAGGCGA_L006_R1_003.fastq.gz 1C_TAAGGCGA_L006_R1_004.fastq.gz 1C_TAGGCATG_L003_R1_001.fastq.gz 1C_TAGGCATG_L005_R1_001.fastq.gz
2C_CGTACTAG_L003_R1_001.fastq.gz

## FASTQ: Format

## - FASTA Format

$>$ C_007779.1 Escherichia coli str. K-12 substr. W3110, complete genome AGCTTTTCATTCTGACTGCAACGGGCAATATGTCTCTGTGTGGATTAAAAAAAGAGTGTCTGATAGCAGCTTCTGAACTG GTTACCTGCCGTGAGTAAATTAAAATTTTATTGACTTAGGTCACTAAATACTTTAACCAATATAGGCATAGCGCACAGAC

## Header <br> Sequence

 AGATAAAAATTACAGAGTACACAACATCCATGAAACGCATTAGCACCACCATTACCACCACCATCACCATTACCACAGGT AACGGTGCGGGCTGACGCGTACAGGAAACACAGAAAAAAGCCCGCACCTGACAGTGCGGGCTTTTTTTTTCGACCAAAGG TAACGAGGTAACAACCATGCGAGTGTTGAAGTTCGGCGGTACATCAGTGGCAAATGCAGAACGTTTTCTGCGTGTTGCCG ATATTCTGGAAAGCAATGCCAGGCAGGGGCAGGTGGCCACCGTCCTCTCTGCCCCCGCCAAAATCACCAACCACCTGGTG GCGATGATTGAAAAAACCATTAGCGGCCAGGATGCTTTACCCAATATCAGCGATGCCGAACGTATTTTTGCCGAACTTTT GACGGGACTCGCCGCCGCCCAGCCGGGGTTCCCGCTGGCGCAATTGAAAACTTTCGTCGATCAGGAATTTGCCCAAATAA AACATGTCCTGCATGGCATTAGTTTGTTGGGGCAGTGCCCGGATAGCATCAACGCTGCGCTGATTTGCCGTGGCGAGAAA
## - FASTQ Format

@N00698:36:000000000-AFBEL:1:1101:16483:1412 1:N:0:0
CTGCCAGTTGAACGACGGCGAGCAGTTATAAGCCAGCAGTTTGCCCGGATATTTCGCGTGGATAGCTTGTGCAAAGCGACGCGCCAGTTCCAGATCCGGCG
ÄAABBFFFFFFFGGGGGGGGGGGGHHHHHHHHGHGHGHHHHHGHHHGGGGGHHHHGGGGGGGHHHGHHFFHHHHHGHGGGGGGGGGGHHHHHHHHHHGGG

## Quality <br> Score

CTGCCAGTTGAACGACGGCGAGCAGTTATAAGCCAGCAGTTTGCCCGGATATTTCGCGTGGATAGCTTGTGCAAAGCGACGCGCCAGTTCCAGATCCGGCG
AAABBFFFFFFFGGGGGGGGGGGGHHHHHHHGHGHGHHHHHGHHHGGGGGHHHHGGGGGGGHHHGHHFFHHHHHGHGGGGGGGGGHHHHHHHHHHGGG @M00698:36:000000000-AFBEL:1:1101:15928:1413 1:N:0:0
GTAAAGTCCTGAGTGATACCGGCAACTTTTACCCCCAGTCCCACTTTCGAACCGGCAAACATATCGGCAAAAGAGGCCGTGCCTGATTTTAAAGCCGTAGGT
$+$
1>AAAFFFFFFFGGGGGGGGCEGGGHHHHHFHHHHGGGHHHHHHHHGHFA0FEGGGGGHGGHHHHEECEEHFEFH/E/E/>EGGHBHHHDGHHHHC<?E/1
 ABCCCCDFFFFEGGGGGGGGGHHHHHGHHC GGFFHHI HGGHG HGGGGGHF HHHH HHHH HGGGGGG HHGGGGGG<C HGHGHEH DGGF
@M00698:36:000000000-AFBEL:1:11d:1534d 1413 1: 10
CCACTAACAAACTAGCCTGATTAAGTTTTAACGCTTCAACCCCAGGCAGGGCTTCCACGCGATCTCTTTTGGGTTTGACCTCTCTTGATCCCCGTCCTAAG
CCACTAACAAACTAGCCTGATTAAGTTTTAACGCTTCAACCCCAGGCAGGGCTTCCACGCGATCTCTTTTTGGGTTTGACCTCTCTTGATCCCCGTCCTAAG
+
AABABFFFFFFCGGGGGGGGGGEHHGGFHHGFHGGGGGHHHHGGGGFHGGGFGGHHHCFGFFEFCGHHFHHHHGGAFGGFFFFHHHHHHHHHHHGEEGGGHHB
AABABFFFFFFCGGGGGGCGABEL:1:1101:16045:1413 1:N:0:0
GTAGCATTATCAGAGAGTTGCCATTCACGCATTGGCTTAACCGCGCGCAGACCATCAACAGTCACTTTGGCGTCAAAGACATTAGGCGTGCAGTATTTTTT
$+$
?ABCCFFFFFFFGGGGGGGGGGHHHHHHHGGGHHHHHHHHHGGGGGGGGGHHHHHHHHHHHHHHHHHHHGGGGGHHHHHHHHHHGGGGHHGHHHHHHG @M00698:36:000000000-AFBEL:1:1101:17191:1414 1:N:0:0
TCGGCACCAATATAGGTAACGCATGGTTCACCGTCTTCAGCTACGGCGGCGATTTTGGTCAGGATCGGTGCTACCAATTCATAGGCTTCTTTCTGGCCAC
$+$
ABAABBBBFBFFGGGGGGGGGGGGGHHHHHHHGHGHGHHHHHHHGGGGGGGGGHGHGGHHHHHHHGGGGGHHHHHHHHHFHGHHHHHHHHBFHHG @M00698:36:000000000-AFBEL:1:1101:16186:1414 1:N:0:0
TCAGGGTCGTCCGTGGAGCAAACATAGCCCTGAGGCTTATTGAGCATGAAGTAACGTGGACCGTGTTGCTGCGCCAGCGGGTTGCCATCGTAAGCGACAT $+$
BCCCCFFFCCCCGGGGGGGGGGHHHHGHHHHHHHGGHHHHHHHHHHHHHHHHHHHGHHGGHHGGHHHHHHHHGGGGGGGGGGGGHHHHHGHHHFHGGGGG @M00698:36:000000000-AFBEL:1:1101:15394:1415 1:N:0:0
ACTTCTTGAAGGTACGTTGCAGGCTGGCACTTCTGCCGTTTCTGATAAGTTGCTTGATTTGGTTGGACTTGGTGGCAAGTCTGCCGCTGATAAAGGAAAGG
$+$
@BBCCFFFFFFFGGGGGGGGGGGGHHDGHHHHHHHHHGEGHGHHHGHHHHHHHHHEHHHHHHHGGHHHHHHHEGGHHGHHHHHHGGGGFHHHHHHHHHHH @M00698:36:000000000-AFBEL:1:1101:14326:1417 1:N:0:0
TGGCGCAACTAACAGAACGTCTTGCGTTTTGTTGGCGAAGCCGCTGGTGTTTGTAAATTTATTAGTGATCGGCGTAGCGTTACGGGTTTTCACCGTAGTTCG $+$
ABCCCCCCCFFFGGGGGGGGGGGGHHGGGGGHHHHHHGGGGGHGGGGGFGFFHHHHHHHGHHHHHHHHHGHGGGGGFFAE>EGGGGGEGGHHHHGGHGGHHG @M00698:36:000000000-AFBEL:1:1101:15479:1417 1:N:0:0
GCCCCCCTCACTCTGACTTTAGTGTGCGCCTTTTGCCTGCGCCACGATCTCCTCGACATTTTTCCGGCATTGGTTCATACGGCGTTTCCTTTCGCCGTTGG
+
+
+
@BCCCCCCCFCFGGGGGGGGGGGHHHHGGGGGHHHHHHHHGGGGGGGHHHHHHGGGGGHHHHHGGGGGHHHHHHHHHHGGGGGGHHHHGGGGGGGH @M00698:36:000000000-AFBEL:1:1101:16850:1418 1:N:0:0
CGTTTTCTTCATCGCGCTCTTTGCTGCCTAACAGCGTGCGCCAGCCTGCTTCAGCAAGAAAACGCGCTTTAGCGACAAATTTGCCTTTGGCAATGTCCAGT $+$
AAACCCFFFFFFGGGGGGGGGGGHHHHHHHHHHHGGGGGGGGGGGHHHHHHHHHHHHHHHHGGGGGGGHGHGGGGGFHHHHHHHGHHHHHHHHHHG @M00698:36:000000000-AFBEL:1:1101:16255:1418 1:N:0:0
CCGGCTTGCTGGTTGCAGCCGTTGCTGTACTTGATGTCAGGCGTGCCGGTGCCGTATTTTCAAACGGTGATGCCGGACGCGACTCTACGGGTTTGACATCG $+$
AAABCCCCCFFFGGGGGGGGGGGGGHHHHHHHHHHHHHHHHGGGGHGGGGGHGFGHHHHHHHHHGGGHGHHHHGGGGGGGGGGHHHGGGCGGGHHHHH? @M00698:36:000000000-AFBEL:1:1101:17071:1419 1:N:0:0
CACGCCACCTTTATCCAGTTTGCGGCTTTGCGAAGTGGCGAGCCACAGCGCGTTTTCTTGCTGGGCTATAAGCCATTTCGTAGGCACCTTTACCTACCGCTT
$+$
AABBBBBBFFFFGGGGGGGGGGGHGGGGGHHHGGGGFHHGGGGGGHHHHHGGGFGGGHH H $H H H H$ H $H H H H H H H H H H H H H G G H G H G H H F H H H H H G G G G$ @M00698:36:000000000-AFBEL:1:1101:15606:1420 1:N:0:0
CGGGTTTTTAACTTTCAGCCACGGGCCACCGTCGATCAGTTCACCGCCAAACTCTTCACGCGCCAGCTGGTAGCCCCAGTCTTTAAACGCTCCTTCGGTGA
BBCDBBCCDCFFGGGGGGGGGGGGGGGGGHGGGGGGHGHHHHHHHGGGGGHHHHHHHHHHGGGGGGHHHHHHHHHGGGHHHHHHHHGGGGHHHHGFGG @M00698:36:000000000-AFBEL:1:1101:15945:1421 1:N:0:0
TCATTGTTGCGCGGTATTCGCGCCCGTTGGTCGAGTAGCAGAAGGGGATTTTAAACCGTTGTTTGCCGCTGGTGTCCTGCCAGCTGGTTTCATACTCTGG
$+$
?ABABFFFFBBBFGEEGGGGGGGGGGGHGHCHGCGGGHGHHHHHGGGGHGHHHHGHGEGDGHHHHHGGG?EB2GGHGHHHGHGHHHHGHHFHHHHHGGH @M00698:36:000000000-AFBEL:1:1101:16329:1421 1:N:0:0
GCCCCGACAGCTGTATGCATAGCGATAAATTCCAGCAGGCCGGGACGCCGGTCTATTTTTGCCCCAGAGTGTCAATGTTAGACTTGACGGACATTGTGCAG
$+$
3ABCCCCCCCCFGGGGGGGGGHGGHGGHHHHHHHHHHHGGGGGGGGGGGGGGHHHHHHHHHGGGHHHHHHHGHHHHHHHHHHHHHGGGGGHHHHHHH

## Quality Scores

- Go to Notebook:

HTS2018-notebooks/bioinformatics/quality_scores.ipynb

## FASTQ: Read Files

## Combined R1.fasta.gz

@M00698:36:000000000-AFBEL:1:1101:14738:1412 1:N:0:0 TTACGCTAACAGGCGGTAGCCTGGCAGGGTCAGGAAATCAATTAACTCATCGGAAGTGGTGATCTGTTCCATCAAGCGTGCGGCATCGTCAAAACGCCC $+$
ABBBABBBAFFFGGGGGGGGGGHGGHGGGCG2GF3FFGHHHHHHGGFGHEHHGGGEHHHHAGGHHGHHHFFDHFHHHGEGGGG@F@H?GHH/GBEFGGG @M00698:36:000000000-AFBEL:1:1101:16483:1412 1:N:0:0
CTGCCAGTTGAACGACGGCGAGCAGTTATAAGCCAGCAGTTTGCCCGGATATTTCGCGTGGATAGCTTGTGCAAAGCGACGCGCCAGTTCCAGATCCGGCG $+$
AAABBFFFFFFFGGGGGGGGGGGGHHHHHHHHGHGHGHHHHHGHHHGGGGHHHHGGGGGGGHHHGHHFFHHHHHGHGGGGGGGGGGHHHHHHHHHHGGG

## Combined R2.fasta.gz

@M00698:36:000000000-AFBEL:1:1101:14738:1412 2:N:0:0
GGAAGATGCGGCGACGGCTGAAATTTCCCGTACCTCGATCTGGCAGTGGATCCATCATCAAAAAACGTTGAGCAATGGCAAACCGGTGACCAAAGCCTTGT $+$
ABBABFFFFDBDGCG??FFGGGHGHFEG3EAEGGFHAE3GFBGGHGGGHHCFGHFGBGHFHHDFEGGHFHEFHHHH3BFGF0GFEGGGGGHHA/FGHFHHH @M00698:36:000000000-AFBEL:1:1101:16483:1412 2:N:0:0
GCTTCTTCCGTACTCATGCGGGCATTGAGCAAGCGATCAGCCGTGGCCTGGCGTATGCGCCATATGCTGACCTGGTCTGGTGTGAAACCTCCACGCCGGAT $+$
CCCCCFFFFBFFGGGGGGGGCECGHHHHHHHHHHGGHGGGGHGGCGCHHGFHGGGGHHGGGGGHHHHHHGHHHHHHHHHHHHHEHGHHHGHHHHGGGGGGG

```
Combined I1.fastq.gz
@M00698:36:000000000-AFBEL:1:1101:14738:1412 1:N:0:0
AGTTCC
+
CCCCDF
@M00698:36:000000000-AFBEL:1:1101:16483:1412 1:N:0:0
CCTGTC
+
A11>>1
```

| Step | Input | Output | Output Format |
| :---: | :---: | :---: | :---: |
| Pre-processing <br> and QC | Raw Reads | Processed <br> Reads | FASTQ |
| Mapping | Reads | Mapped <br> Reads | SAM/BAM |
| Counting | Mapped Reads | Count Table | CSV |

## Pre-processing

1. Demultiplex [often done at sequencing facility]
2. Quality Control
3. Trimming (adapter and quality)
4. Filtering [must keep reads in sync]

## Demultiplexing

## Undetermined S0 L001 R1 001.fastq.gz

@M00698:36:000000000-AFBEL:1:1101:14738:1412 1:N:0:0
TTACGCTAACAGGCGGTAGCCTGGCAGGGTCAGGAAATCAATTAACTCATCGGAAGTGGTGATCTGTTCCATCAAGCGTGCGGCATCGTCAAAACGCCC $+$
ABBBABBBAFFFGGGGGGGGGGHGGHGGGCG2GF3FFGHHHHHHGGFGHEHHGGGEHHHHAGGHHGHHHFFDHFHHHGEGGGG@F@H?GHH/GBEFGGG @M00698:36:000000000-AFBEL:1:1101:16483:1412 1:N:0:0
CTGCCAGTTGAACGACGGCGAGCAGTTATAAGCCAGCAGTTTGCCCGGATATTTCGCGTGGATAGCTTGTGCAAAGCGACGCGCCAGTTCCAGATCCGGCG
$+$
AAABBFFFFFFFGGGGGGGGGGGGHHHHHHHHGHGHGHHHHHGHHHGGGGGHHHHGGGGGGGHHHGHHFFHHHHHGHGGGGGGGGGGHHHHHHHHHHGGG

## Undetermined S0 L001 R2 001.fasta.gz

@M00698:36:000000000-AFBEL:1:1101:14738:1412 2:N:0:0
GGAAGATGCGGCGACGGCTGAAATTTCCCGTACCTCGATCTGGCAGTGGATCCATCATCAAAAAACGTTGAGCAATGGCAAACCGGTGACCAAAGCCTTGT +
ABBABFFFFDBDGCG??FFGGGHGHFEG3EAEGGFHAE3GFBGGHGGGHHCFGHFGBGHFHHDFEGGHFHEFHHHH3BFGF0GFEGGGGGHHA/FGHFHHH @M00698:36:000000000-AFBEL:1:1101:16483:1412 2:N:0:0
GCTTCTTCCGTACTCATGCGGGCATTGAGCAAGCGATCAGCCGTGGCCTGGCGTATGCGCCATATGCTGACCTGGTCTGGTGTGAAACCTCCACGCCGGAT $+$
CCCCCFFFFBFFGGGGGGGGCECGHHHHHHHHHHGGHGGGGHGGCGCHHGFHGGGGHHGGGGGHHHHHHGHHHHHHHHHHHHHEHGHHHGHHHHGGGGGGG

Sample Barcode
7A_K AGTCAA 7B_K AGTTCC 7C_N CGTACG 8A_N GAGTGG 8B_N CCTGTC 8C_N ATTCCT 7A_P CGATGT

## Undetermined S0 L001 I1 001.fasta.gz

@M00698:36:000000000-AFBEL:1:1101:14738:1412 1:N:0:0
AGTTCC
$+$
CCCCDF
@M00698:36:000000000-AFBEL:1:1101:16483:1412 1:N:0:0
CCTGTC
+
A11>>1

## Demultiplexing

```
7B K R1 001.fasta.gz
@M00698:36:000000000-AFBEL:1:1101:14738:1412 1:N:0:0
TTACGCTAACAGGCGGTAGCCTGGCAGGGTCAGGAAATCAATTAACTCATCGGAAGTGGTGATCTGTTCCATCAAGCGTGCGGCATCGTCAAAACGCCC
\(+\)
ABBBABBBAFFFGGGGGGGGGGHGGHGGGCG2GF3FFGHHHHHHGGFGHEHHGGGEHHHHAGGHHGHHHFFDHFHHHGEGGGG@F@H?GHH/GBEFGGG
```


## 7B K R2 001.fastq.gz

@M00698:36:000000000-AFBEL:1:1101:14738:1412 2:N:0:0
GGAAGATGCGGCGACGGCTGAAATTTCCCGTACCTCGATCTGGCAGTGGATCCATCATCAAAAAACGTTGAGCAATGGCAAACCGGTGACCAAAGCCTTGT
$+$
ABBABFFFFDBDGCG??FFGGGHGHFEG3EAEGGFHAE3GFBGGHGGGHHCFGHFGBGHFHHDFEGGHFHEFHHHH3BFGF0GFEGGGGGHHA/FGHFHHH

## 8B N R1 001.fasta.gz

@M00698:36:000000000-AFBEL:1:1101:16483:1412 1:N:0:0
CTGCCAGTTGAACGACGGCGAGCAGTTATAAGCCAGCAGTTTGCCCGGATATTTCGCGTGGATAGCTTGTGCAAAGCGACGCGCCAGTTCCAGATCCGGCG $+$
AAABBFFFFFFFGGGGGGGGGGGGHHHHHHHHGHGHGHHHHHGHHHGGGGGHHHHGGGGGGGHHHGHHFFHHHHHGHGGGGGGGGGGHHHHHHHHHHHGGG

## 8B N R2 001.fastq.gz

## @M00698:36:000000000-AFBEL:1:1101:16483:1412 2:N:0:0

GCTTCTTCCGTACTCATGCGGGCATTGAGCAAGCGATCAGCCGTGGCCTGGCGTATGCGCCATATGCTGACCTGGTCTGGTGTGAAACCTCCACGCCGGAT
$+$
CCCCCFFFFBFFGGGGGGGGCECGHHHHHHHHHHGGHGGGGHGGCGCHHGFHGGGGHHGGGGGHHHHHHGHHHHHHHHHHHHHEHGHHHGHHHHGGGGGGG

## Quality Control

## Trim:

## Trimming

- Phred <20
- N's
- Adapter:
- AAACCTCCACGCCGGATACCAGTCTAATCGGCTAGTC

```
SampleX R1.fastq.gz
@M00698:36:000000000-AFBEL:1:1101:14738:1412 1:N:0:0
TTACGCTAACAGGCGGTAGCCTGGCAGGGTCAGGAAATCAATTAACTCATCGGAAGTGGTGATCTGTTCCATCAAGCGTGCGGCATCGTCAAAACGCCC
+
ABBBABBBAFFFGGGGGGGGGGHGGHGGGCG2GF3FFGHHHHHHGGFGHEHHGGGEHHHHAGGHHGHHHFFDHFHHHGEGGGG@F@H?GHH/GBEFGGG
@M00698:36:000000000-AFBEL:1:1101:16483:1412 1:N:0:0
CTGCCAGTTGAACGACGGCGAGCAGTTATAAGCCAGCAGTTTGCCCGGATATTTCGCGTGGATAGCTTGTGCAAAGCGACGCGCCAGTTCCAGATCCGGCG
+
AAABBFFFFFFFGGGGGGGGGGGGHHHHHHHHGHGHGHHHHHGHHHGGGGGHHHHGGGGGGGHHHGHHFFHHHHHGHGGGGGGGGGGHHHHHHHHHHHGGG
```


## SampleX R2.fastq.gz

```
@M00698:36:000000000-AFBEL:1:1101:14738:1412 2:N:0:0
GGAAGATGCGGCGACGGCTGAAATTTCCCGTACCTCGATCTGGCAGTGGATCCATCATCAAAAAACGTTGAGCAATGGCAAACCGGTGACCAAAGCCTTGT \(+\)
ABBABFFFFDBDGCG??FFGGGHGHFEG3EAEGGFHAE3GFBGGHGGGHHCFGHFGBGHFHHDFEGGHFHEFHHHH3BFGF0GFEGGGGGHHA/FGHFHHH @M00698:36:000000000-AFBEL:1:1101:16483:1412 2:N:0:0
GCTTCTTCCGTACTCATGCGGGCATTGAGCAAGCGATCAGCCGTGGCCTGGCGTATGCGCCATATGCTGACCTGGTCTGGTGTGAAACCTCCACGCCGGAT
\(+\)
CCCCCFFFFBFFGGGGGGGGCECGHHHHHHHHHHGGHGGGGHGGCGCHHGFHGGGGHHGGGGGHHHHHHGHHHHHHHHHHHHHEHGHHHGHHHHGGGGGGG
```


## Filter:

## Filtering

- Length < 40 - Averge Phred <20
- N's > 3

```
SampleX R1.fastq.gz
@M00698:36:000000000-AFBEL:1:1101:14738:1412 1:N:0:0
TTACGCTAACAGGCGGTAGCCTGGCAGGGTCAGGAAATCAATTAACTCATCGGAAGTGGTGATCTGTTCCATCAAGCGTGCGGCATCGTCAAAACGCCC
+
ABBBABBBAFFFGGGGGGGGGGHGGHGGGCG2GF3FFGHHHHHHGGFGHEHHGGGEHHHHAGGHHGHHHFFDHFHHHGEGGGG@F@H?GHH/GBEFGGG
@M00698:36:000000000-AFBEL:1:1101:16483:1412 1:N:0:0
CTGCCAGTTGAACGACGGCGAGCAGTTATAAGCCAGCAGTTTGCCCGGATATTTCGCGTGGATAGCTTGTGCAAAGCGACGCGCCAGTTCCAGATCCGGCG
+
AAABBFFFFFFFGGGGGGGGGGGGHHHHHHHHGHGHGHHHHHGHHHGGGGGHHHHGGGGGGGHHHGHHFFHHHHHGHGGGGGGGGGGHHHHHHHHHHHGGG
```


## SampleX R2.fastq.gz

@M00698:36:000000000-AFBEL:1:1101:14738:1412 2:N:0:0
GGAAGATGCGGCGACGGCTGAAATTTCCCGTACCTCGATCTGGCAGTGGATCCATCATCAAAAAACGTTGAGCAATGGCAAACCGGTGACCAAAGCCTTGT $+$
ABBABFFFFDBDGCG??FFGGGHGHFEG3EAEGGFHAE3GFBGGHGGGHHCFGHFGBGHFHHDFEGGHFHEFHHHH3BFGF0GFEGGGGGHHA/FGHFHHH @M00698:36:000000000-AFBEL:1:1101:16483:1412 2:N:0:0
GCTTCTTCCGTACTCATGCGGGCATTGAGCAAGCGATCAGCCGTGGCCTGGCGTATGCGCCATATGCTGACCTGGTCTGGTGTGAAACCTCCACGCCGGAT $+$
CCCCCFFFFBFFGGGGGGGGCECGHHHHHHHHHHGGHGGGGHGGCGCHHGFHGGGGHHGGGGGHHHHHHGHHHHHHHHHHHHHEHGHHHGHHHHGGGGGGG

## Maintaining Synchrony

```
SampleX R1.fastq.gz
@M00698:36:000000000-AFBEL:1:1101:14738:1412 1:N:0:0
TTACGCTAACAGGCGGTAGCCTGGCAGGGTCAGGAAATCAATTAACTCATCGGAAGTGGTGATCTGTTCCATCAAGCGTGCGGCATCGTCAAAACGCCC
+
ABBBABBBAFFFGGGGGGGGGGHGGHGGGCG2GF3FFGHHHHHHGGFGHEHHGGGEHHHHAGGHHGHHHFFDHFHHHGEGGGG@F@H?GHH/GBEFGGG
@M00698:36:000000000-AFBEL:1:1101:16483:1412 1:N:0:0
CTGCCAGTTGAACGACGGCGAGCAGTTATAAGCCAGCAGTTTGCCCGGATATTTCGCGTGGATAGCTTGTGCAAAGCGACGCGCCAGTTCCAGATCCGGCG
+
AAABBFFFFFFFGGGGGGGGGGGGHHHHHHHHGHGHGHHHHHGHHHGGGGGHHHHGGGGGGGHHHGHHFFHHHHHGHGGGGGGGGGGHHHHHHHHHHHGGG
```


## SampleX R2.fastq.gz

```
@M00698:36:000000000-AFBEL:1:1101:14738:1412 2:N:0:0
GGAAGATGCGGCGACGGCTGAAATTTCCCGTACCTCGATCTGGCAGTGGATCCATCATCAAAAAACGTTGAGCAATGGCAAACCGGTGACCAAAGCCTTGT \(+\)
ABBABFFFFDBDGCG??FFGGGHGHFEG3EAEGGFHAE3GFBGGHGGGHHCFGHFGBGHFHHDFEGGHFHEFHHHH3BFGF0GFEGGGGGHHA/FGHFHHH @M00698:36:000000000-AFBEL:1:1101:16483:1412 2:N:0:0
GCTTCTTCCGTACTCATGCGGGCATTGAGCAAGCGATCAGCCGTGGCCTGGCGTATGCGCCATATGCTGACCTGGTCTGGTGTGAAACCTCCACGCCGGAT \(+\)
CCCCCFFFFBFFGGGGGGGGCECGHHHHHHHHHHGGHGGGGHGGCGCHHGFHGGGGHHGGGGGHHHHHHGHHHHHHHHHHHHHEHGHHHGHHHHGGGGGGG
```


## Maintaining Synchrony

```
SampleX R1.fasta.gz
@M00698:36:000000000-AFBEL:1:1101:16483:1412 1:N:0:0
CTGCCAGTTGAACGACGGCGAGCAGTTATAAGCCAGCAGTTTGCCCGGATATTTCGCGTGGATAGCTTGTGCAAAGCGACGCGCCAGTTCCAGATCCGGCG
+
AAABBFFFFFFFGGGGGGGGGGGGHHHHHHHHGHGHGHHHHHGHHHGGGGGHHHHGGGGGGGHHHGHHFFHHHHHGHGGGGGGGGGGHHHHHHHHHHHGGG
```


## SampleX R2.fasta.gz

```
@M00698:36:000000000-AFBEL:1:1101:14738:1412 2:N:0:0
GGAAGATGCGGCGACGGCTGAAATTTCCCGTACCTCGATCTGGCAGTGGATCCATCATCAAAAAACGTTGAGCAATGGCAAACCGGTGACCAAAGCCTTGT
\(+\)
ABBABFFFFDBDGCG??FFGGGHGHFEG3EAEGGFHAE3GFBGGHGGGHHCFGHFGBGHFHHDFEGGHFHEFHHHH3BFGF0GFEGGGGGHHA/FGHFHHH @M00698:36:000000000-AFBEL:1:1101:16483:1412 2:N:0:0
GCTTCTTCCGTACTCATGCGGGCATTGAGCAAGCGATCAGCCGTGGCCTGGCGTATGCGCCATATGCTGACCTGGTCTGGTGTGAAACCTCCACGCCGGAT \(+\)
CCCCCFFFFBFFGGGGGGGGCECGHHHHHHHHHHGGHGGGGHGGCGCHHGFHGGGGHHGGGGGHHHHHHGHHHHHHHHHHHHEHGHHHGHHHHGGGGGGG
```

| Step | Input | Output | Output Format |
| :---: | :---: | :---: | :---: |
| Pre-processing <br> and QC | Raw Reads | Processed <br> Reads | FASTQ |
| Mapping | Reads | Mapped <br> Reads | SAM/BAM |
| Counting | Mapped Reads | Count Table | CSV |

## Fragments of a Library

ng on nights, it's us; we're going to set you free." Jim only had time to grab us by the hand and s of this agreement. See paragraph 1.C below. There are a lot of things you can do with Project Guten with his head!"' 'How dreadfully savage!' exclaimed Alice. 'And ever since that,' the Hatter went g, that chateau of Monsieur the Marquis, with a large stone courtyard before it, and two stone sweep he nursery. And he was indeed crying. She heard him and hastened. But the faster she went, the loude $h$ to intensify the influence of her gaze. "Yes, they draw away all the sap and give a false appeara him. I could not look at him without feeling sorry for him. We both know him. He's good-hearted, ed, under various disguises of Art, through the portraits of every Drinking Age. "You are a little eupon Punch Costello dinged with his fist upon the board and would sing a bawdy catch _Staboo Stabel him to act in anything without her quiet aid), and the day passed quickly. Early in the evening he e ed. 'Give your evidence,' said the King; 'and don't be nervous, or I'll have you executed on the sp harge a fee for access to, viewing, displaying, performing, copying or distributing any Project Gute ef secretary. She knew that he usually went out quickly to his office, and she wanted to see him bef onsiderable. Then he said he must start in and "'terpret" it, because it was sent for a warning. ? ides her body's flaws calling under her brown shawl from an archway where dogs have mired. Her fancy recatory manner, "the anguish of his daughter, which must be a dreadful anguish to him!" "I have ob uld be like, but it puzzled her too much, so she went on: 'But why did they live at the bottom of a Indivisible. Liberty, Equality, Fraternity, or Death! Who could that be with Mr. Lorry--the owner ns in general, and I believe that's just why philanthropic institutions always give such poor resu here were no proper cupboards for their clothes; what cupboards there were either would not close at he detected all around him, walked from one to another. The first was the best room, and in it were $t$. All her arrangements had to be modified because they could not be carried out, and they were modi er hand shook more violently, but she did not take her eyes off him, watching how he would take it. s." It was done. "Well?" "Monseigneur, it is nothing. The trees and the night are all that are he spectacles offered by our streets, hideous publicity posters, religious ministers of all denominatio sore, but I couldn't see no way out of the trouble. After all this long journey, and after all we she got to pumping me about England, and blest if I didn't think the ice was getting mighty thin so the morning all sound and right. So she had to be satisfied. But she said she'd set up for him prison--seemed to strike across the earth, messieurs, to where the sky rests upon it!" The hungry m oys running, playing at horses. Seryozha! And I'm losing everything and not getting him back. Yes,

## Mapped Fragments

na on niahts it's sus: we're aoina to set vou free " lim onlv had time to arab us hy the hand and s.--471971 Huckleberrv Finn of this agreement. See paragraph 1.C below. There are a lot of things you can do with Project Guten -- ?????????????

With his head! How dreadfulty savage! exclaimed Alice. And ever since that, the Hatter went - $/ 8968$ Allce's Adventures $g$, that chateau of Monsieur the Marquis, with a large stone courtyard before it, and two stone sweep -- 232648 A Tale of Two Cities he nursery. And he was indeed crying. She heard him and hastened. But the faster she went, the loude -- 1897819 Anna Karenina $h$ to intensify the influence of her gaze. "Yes, they draw away all the sap and give a false appeara -- 1186098 Anna Karenina him. I could not look at him without feeling sorry for him. We both know him. He's good-hearted, -- 165528 Anna Karenina ed, under various disguises of Art, through the portraits of every Drinking Age. "You are a little -- 168896 A Tale of Two Cities eupon Punch Costello dinged with his fist upon the board and would sing a bawdy catch _Staboo Stabel -- 803339 Ulysses him to act in anything without her quiet aid), and the day passed quickly. Early in the evening he e-- 490073 A Tale of Two Cities ed. 'Give your evidence, said the King: 'and don't be nervous, or I'll have you executed on the sp -- 127701 Alice's Adventures harge a fee for access to, viewing, displaying, performing, copying or distributing any Project Gute -- ?????????????
'
onsiderable. Then he said he must start in and "'terpret" it, because it was sent for a warning. ? -- 176566 Huckleberry Finn ides her body's flaws calling under her brown shawl from an archway where dogs have mired. Her fancy -- 90509 Ulysses recatory manner, "the anguish of his daughter, which must be a dreadful anguish to him!" "I have ob -- 682808 A Tale of Two Cities uld be like, but it puzzled her too much, so she went on: 'But why did they live at the bottom of a -- 80787 Alice's Adventures Indivisible. Liberty, Equality, Fraternity, or Death! Who could that be with Mr. Lorry--the owner -- 562830 A Tale of Two Cities ns in general, and I believe that's just why philanthropic institutions always give such poor resu -- 1699588 Anna Karenina here were no proper cupboards for their clothes; what cupboards there were either would not close at -- 634360 Anna Karenina he detected all around him, walked from one to another. The first was the best room, and in it were -- 182459 A Tale of Two Cities $t$. All her arrangements had to be modified because they could not be carried out, and they were modi -- 1252025 Anna Karenina er hand shook more violently, but she did not take her eyes off him, watching how he would take it. -- 455094 Anna Karenina s." It was done. "Well?" "Monseigneur, it is nothing. The trees and the night are all that are he -- 235828 A Tale of Two Cities spectacles offered by our streets, hideous publicity posters, religious ministers of all denominatio -- 870689 Ulysses
sore, but I couldn't see no way out of the trouble. After all this long journey, and after all we -- 422461 Huckleberry Finn she got to pumping me about England, and blest if I didn't think the ice was getting mighty thin so -- 345525 Huckleberry Finn the morning all sound and right. So she had to be satisfied. But she said she'd set up for him -- 558004 Huckleberry Finn prison--seemed to strike across the earth, messieurs, to where the sky rests upon it!" The hungry m -- 340478 A Tale of Two Cities oys running, playing at horses. Seryozha! And I'm losing everything and not getting him back. Yes, -- 1837882 Anna Karenina

| Step | Input | Output | Output Format |
| :---: | :---: | :---: | :---: |
| Pre-processing <br> and QC | Raw Reads | Processed <br> Reads | FASTQ |
| Mapping | Reads | Mapped <br> Reads | SAM/BAM |
| Counting | Mapped Reads | Count Table | CSV |

## SAM/BAM

- Sequence Alignment/Map
- Specification: http://samtools.github.io/hts-specs/SAMv1.pdf
- BAM: Binary version of SAM


## SAM/BAM

- Per read information:
- Name
- Position in genome
- Quality of mapping
- Insertions? Deletions?
- Calculated Insert Length (pair-end only)
- Sequence
- Sequence Quality
- Etc


## SAM/BAM: Format

| Coor ref | 123456789012345678901234567890123456789012345 AGCATGTTAGATAA**GATAGCTGTGCTAGTAGGCAGTCAGCGCCAT |  |
| :---: | :---: | :---: |
| +r001/1 | TTAGATAAAGGATA*CTG | Read Pair |
| +r002 | aaaAGATAA*GGATA |  |
| +r003 | acctaAGCTAA | Chimeric |
| +r004 | ATAGCT. . . . . . . . . . . . .TCAGC | ea |
| -r003 | ttagctTAGGC | lignment |
| -r001/2 | CAGCGGCAT |  |

```
@HD VN:1.5 SO:coordinate
@SQ SN:ref LN:45
r001 99 ref 7 30 8M2I4M1D3M = 37 39 TTAGATAAAGGATACTG *
r002 0 ref 9 30 3S6M1P1I4M * 0 0 AAAAGATAAGGATA *
r003 0 ref 9 30 5S6M * 0 0 GCCTAAGCTAA * SA:Z:ref,29,-,6H5M,17,0;
r004 0 ref 16 30 6M14N5M * 0 0 ATAGCTTCAGC *
r003 2064 ref 29 17 6H5M * 0 0 TAGGC * SA:Z:ref,9,+,5S6M,30,1;
r001 147 ref 37 30 9M = 7 -39 CAGCGGCAT * NM:i:1
```


## STAR


(b)

(c)


| Step | Input | Output | Output Format |
| :---: | :---: | :---: | :---: |
| Pre-processing <br> and QC | Raw Reads | Processed <br> Reads | FASTQ |
| Mapping | Reads | Mapped <br> Reads | SAM/BAM |
| Counting | Mapped Reads | Count Table | CSV |

## Counting

ng on nights, it's us; we're going to set you free." Jim only had time to grab us by the hand and s -- 471971 with his head!"' 'How dreadfully savage!' exclaimed Alice. 'And ever since that,' the Hatter went -- 78968 $g$, that chateau of Monsieur the Marquis, with a large stone courtyard before it, and two stone sweep -- 232648 A Tale of Two Cities he nursery. And he was indeed crying. She heard him and hastened. But the faster she went, the loude -- 1897819 Anna Karenina $h$ to intensify the influence of her gaze. "Yes, they draw away all the sap and give a false appeara -- 1186098 Anna Karenina him. I could not look at him without feeling sorry for him. We both know him. He's good-hearted, -- 165528 Anna Karenina ed, under various disguises of Art, through the portraits of every Drinking Age. "You are a little -- 168896 A Tale of Two Cities eupon Punch Costello dinged with his fist upon the board and would sing a bawdy catch _Staboo Stabel -- 803339 him to act in anything without her quiet aid), and the day passed quickly. Early in the evening he e -- 490073 ed. 'Give your evidence,' said the King; 'and don't be nervous, or I'll have you executed on the sp -- 127701 ef secretary. She knew that he usually went out quickly to his office, and she wanted to see him bef -- 782396 onsiderable. Then he said he must start in and "'terpret" it, because it was sent for a warning. ? -- 176566 ides her body's flaws calling under her brown shawl from an archway where dogs have mired. Her fancy -- 90509 recatory manner, "the anguish of his daughter, which must be a dreadful anguish to him!" "I have ob -- 682808 uld be like, but it puzzled her too much, so she went on: 'But why did they live at the bottom of a -- 80787 Indivisible. Liberty, Equality, Fraternity, or Death! Who could that be with Mr. Lorry--the owner -- 562830 ns in general, and I believe that's just why philanthropic institutions always give such poor resu -- 1699588 here were no proper cupboards for their clothes; what cupboards there were either would not close at -- 634360 he detected all around him, walked from one to another. The first was the best room, and in it were -- 182459 $t$. All her arrangements had to be modified because they could not be carried out, and they were modi -- 1252025 er hand shook more violently, but she did not take her eyes off him, watching how he would take it. -- 455094 s." It was done. "Well?" "Monseigneur, it is nothing. The trees and the night are all that are he -- 235828 spectacles offered by our streets, hideous publicity posters, religious ministers of all denominatio -- 870689 sore, but I couldn't see no way out of the trouble. After all this long journey, and after all we -- 422461 she got to pumping me about England, and blest if I didn't think the ice was getting mighty thin so -- 345525 the morning all sound and right. So she had to be satisfied. But she said she'd set up for him -- 558004 prison--seemed to strike across the earth, messieurs, to where the sky rests upon it!" The hungry m -- 340478 oys running, playing at horses. Seryozha! And I'm losing everything and not getting him back. Yes, -- 1837882 Anna Karenina

## Counting

ed, under various disguises of Art, through the portraits of every Drinking Age. "You are a little -- 168896 A Tale of Two Cities he detected all around him, walked from one to another. The first was the best room, and in it were -- 182459 A Tale of Two Cities g, that chateau of Monsieur the Marquis, s." It was done. "Well?" "Monseigneur prison--seemed to strike across the eart him to act in anything without her quiet Indivisible. Liberty, Equality, Fratern recatory manner, "the anguish of his dau onsiderable. Then he said he must start she got to pumping me about England, an sore, but I couldn't see no way out of ng on nights, it's us; we're going to se the morning all sound and right. So sh
with his head!"' 'How dreadfully savag uld be like, but it puzzled her too much ed. 'Give your evidence,' said the King
him. I could not look at him without fe er hand shook more violently, but she di here were no proper cupboards for their ef secretary. She knew that he usually $w$ $h$ to intensify the influence of her gaze $t$. All her arrangements had to be modifi ns in general, and I believe that's just oys running, playing at horses. Seryozha he nursery. And he was indeed crying. Sh
ides her body's flaws calling under her
 eupon Punch costed dinged with his fise upplicity posters, religious ministers of all denominatio -- 870689 Ulysses

## Counting



| Step | Input | Output | Output Format |
| :---: | :---: | :---: | :---: |
| Pre-processing <br> and QC | Raw Reads | Processed <br> Reads | FASTQ |
| Mapping | Reads | Mapped <br> Reads | SAM/BAM |
| Counting | Mapped Reads | Count Table | CSV |

## CSV:

# Comma-separated values 

gene990,1
gene991,2
gene992,20
gene993,1
gene994,3
gene995,0
gene996,14
gene997,9
gene998,45
gene999,0

## TSV:

## Tab-separated values

## gene990 1 <br> gene991 2

gene992 20
gene993 1
gene994 3
gene995 0
gene996 14
gene997 9
gene998 45
gene999 0

## STAR Counts

Coding Antisense
Unstranded Strand Strand

| N_unmapped | 3386 | 3386 | 3386 |
| :--- | :--- | :--- | :--- |
| N_multimapping | 149388 | 149388 | 149388 |
| N_noFeature | 472714 | 2327552 | 503625 |
| N_ambiguous | 138341 | 1266 | 1354 |
| CNAG_04548 | 1 | 0 | 1 |
| CNAG_07303 | 0 | 0 | 0 |
| CNAG_07304 | 9 | 0 | 9 |
| CNAG_00001 | 0 | 0 | 0 |

