

PLB812_Lab1 Genomic File Download Verification

```
(base) -bash-4.2$ pwd
/mnt/home/mitch987
(base) -bash-4.2$ ls
Documents          miniconda3
homemitch987miniconda3  Miniconda3-latest-Linux-x86_64.sh
homemitch987mobaxterm  Miniconda3-latest-Linux-x86_64.sh.1
(base) -bash-4.2$ cd miniconda3/
(base) -bash-4.2$ ls
bin  compiler_compat  condabin  conda-meta  envs  etc  include  lib  LICENSE.txt  pkgs  share  shell  Solanum  ssl  x86_64-conda_cos6-linux-gnu  x86_64-conda-linux-gnu
(base) -bash-4.2$ cd Solanum/
(base) -bash-4.2$ ls
Genomic  RNA_Seq
(base) -bash-4.2$ cd Genomic/
(base) -bash-4.2$ ls
Stuberosum_686_v6.1.annotation_info.txt  Stuberosum_686_v6.1.gene_exons.gff3  Stuberosum_686_v6.1.repeatmasked_assembly_v6.1.gff3
Stuberosum_686_v6.1.cds.fa                Stuberosum_686_v6.1.gene.gff3         Stuberosum_686_v6.1.softmasked.fa
Stuberosum_686_v6.1.cds_primaryTranscriptOnly.fa  Stuberosum_686_v6.1.hardmasked.fa    Stuberosum_686_v6.1.transcript.fa
Stuberosum_686_v6.1.defline.txt            Stuberosum_686_v6.1.protein.fa        Stuberosum_686_v6.1.transcript_primaryTranscriptOnly.fa
Stuberosum_686_v6.1.fa                    Stuberosum_686_v6.1.protein_primaryTranscriptOnly.fa
(base) -bash-4.2$
```

PLB812_Lab1 Genomic File Download Verification (.fa) screenshots

```
(base) ~-bash-4.2$ head *.fa
=> Stuberousm_686_v6.1.cds.fa <==
>Soltu.DM.S001790.1 pacid=51125199 polypeptide=Soltu.DM.S001790.1 locus=Soltu.DM.S001790 ID=Soltu.DM.S001790.1.v6.1 annot-version=v6.1
ATGGGAGAGGGAGCTTACGGGATCGTTTATAAAGGAACACTTTCGAATGAAATTCATGTTTGGCTGTAATAATCTCTCAATGA
TTCTCGAAGGAATGGGAGGAATTTATTAATGAAGTTGCGAGCAATGGGGAATATCCATCATGTCAACGTGGTTTGGCTAG
TTGGGTTCTGTGCTGATGAGCAACAAAATAATCTGTTAGACCTGGAACCCCAAGATCTCTGATTTTGGCTTGGCTAAGTTA
TGCTCTAAAGAGAAAGTGTGTGACAAATGACTGAAGCTAGAGGAACCATGGGTTATATTGCCAGAGATGATATCAAG
CAATTTCCGAAAAGCGCTCACAAATCTGATGTTTATAGCTTTGGAATGCTGCTTCTTGAATGCGGCTGAATCAAGGAG
AAGAGTTGAAAATCAGGATTTGATGAAGATGATGATGTTATAATTGTAAGGAAATTTGGCTGTTATAGGAGTTTGGTGATC
CAATGGAATGCGATAGATCGGCCATCCATCAAAGTTGTAACTCAAATGCTAGAAGGAGATGGGAGTAACCTAACTGTTCC
TTCTCCTTTTACAACAAGAAACACCACCATGATAGGTCCTTACGCAAGAGTTGACGATAATCTCAGAACTAGAATGA
>Soltu.DM.S001810.1 pacid=51125200 polypeptide=Soltu.DM.S001810.1 locus=Soltu.DM.S001810 ID=Soltu.DM.S001810.1.v6.1 annot-version=v6.1

=> Stuberousm_686_v6.1.cds.primaryTranscriptOnly.fa <==
>Soltu.DM.S001790.1 pacid=51125199 polypeptide=Soltu.DM.S001790.1 locus=Soltu.DM.S001790 ID=Soltu.DM.S001790.1.v6.1 annot-version=v6.1
ATGGGAGAGGGAGCTTACGGGATCGTTTATAAAGGAACACTTTCGAATGAAATTCATGTTTGGCTGTAATAATCTCTCAATGA
TTCTCGAAGGAATGGGAGGAATTTATTAATGAAGTTGCGAGCAATGGGGAATATCCATCATGTCAACGTGGTTTGGCTAG
TTGGGTTCTGTGCTGATGAGCAACAAAATAATCTGTTAGACCTGGAACCCCAAGATCTCTGATTTTGGCTTGGCTAAGTTA
TGCTCTAAAGAGAAAGTGTGTGACAAATGACTGAAGCTAGAGGAACCATGGGTTATATTGCCAGAGATGATATCAAG
CAATTTCCGAAAAGCGCTCACAAATCTGATGTTTATAGCTTTGGAATGCTGCTTCTTGAATGCGGCTGAATCAAGGAG
AAGAGTTGAAAATCAGGATTTGATGAAGATGATGATGTTATAATTGTAAGGAAATTTGGCTGTTATAGGAGTTTGGTGATC
CAATGGAATGCGATAGATCGGCCATCCATCAAAGTTGTAACTCAAATGCTAGAAGGAGATGGGAGTAACCTAACTGTTCC
TTCTCCTTTTACAACAAGAAACACCACCATGATAGGTCCTTACGCAAGAGTTGACGATAATCTCAGAACTAGAATGA
>Soltu.DM.S001810.1 pacid=51125200 polypeptide=Soltu.DM.S001810.1 locus=Soltu.DM.S001810 ID=Soltu.DM.S001810.1.v6.1 annot-version=v6.1

=> Stuberousm_686_v6.1.hardmasked.fa <==
<chr01
TAGTCCACAGACTTTTTTTGTGATCTGGAATCTAACGTATGTTTCCCAAAAATTTTCGT
GGACGTAGCTTAAGCATTAGCTATGGAGCCAGTTAGCCCTCAGCGGCACACAGTCCCAT
TTTAAAGGTTAAATGTTCCAGAGCAGGAAACCCCAATTTTGGCGATTTAAGTTTGTCT
ATAGTCCATGGATTGTTTGGTGATCTGGGTTCCGACCTAAATTTTGTCAAAAATTTTCAT
AGACGTCCTTAAGACGTCAGCTATGGAGCCAGTTAGCCCTCAGCGGCACAAACGTCCTCT
TTAAAAGGTCAAATGTGACCCAGAGTTGAAAAACCCCAATTTTGGCGATTTCGTGCGGT
ATAGTCCATGGACTTTTGGTGAAGTGAATCCGACAAAATTTTGTGTTAAAAATTTTCG
TGGATGTCGGTTAAGACATTAGCTACGGAGCCAGTTAGCCCTCAGACCTAAACATCTCA
TTTTGAAAGTCAAATCTGCCCGGAGACGTTAAACCCCCATTTTGGCGAATATCGTGTG

=> Stuberousm_686_v6.1.hardmasked.fa <==
<chr01
NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
```

```
=> Stuberousm_686_v6.1.transcript.primaryTranscriptOnly.fa <==
>Soltu.DM.S000060.1 pacid=51152479 locus=Soltu.DM.S000060 ID=Soltu.DM.S000060.1.v6.1 annot-version=v6.1
ACTAAACAAAGTCCTTATAATTTCTCTATTTTATACCAAAGTTTAACTAAACTTCAATTT
CTTTTCAATTAATTTTCTGTTGTTAATTCATCTCCATTTTGTGCAACAATGGTTCTCTC
AAAAACTTCTCGGAGTCTGATGTTTATGTACACTCCACCTTTGCTTCTCGCTATGTTTCG
AGCTTCACTTCCAAGGTTTGAGATGCCAGAAAATTCATCCCAAAAGAGGCAGCATACCA
GATTAATTAAGATGATGAGTTGATGCTTGATGGCAATCCAAGGTTGAATTTGGCATCATTTGT
GACAACATGGATGGAGCCAGAATGTGATAAGCTTATGATGTTCTTCAATAAACAGAACATA
TGTGATATGGATGAATACCTGTTTACCACCTGAGCTTCAGAATCGATGTGTGAATATCAT
AGCACGTTTTATTCAACGCACCTTTGAAGAGGGAAGAAATAGCAATTTGGTGTGGCCACAGT
GGGCTCATCAGAGGCCATAATGTTAGCGGGCTAGCCTTCAAGAGGAACTGGCAAAACAA
(base) ~-bash-4.2$ █
```

```
=> Stuberousm_686_v6.1.protein.fa <==
>Soltu.DM.S001790.1 pacid=51125199 transcript=Soltu.DM.S001790.1 locus=Soltu.DM.S001790 ID=Soltu.DM.S001790.1.v6.1 annot-version=v6.1
MGEAGYGIYVKGTLISNEIHVAVKILNDSEGNIEEINEVAAMGNHHVNVVCLVGFCDGPKNILLDNPKISDFGLAKL
CSKEKSVVTMTTEARGTMGYIAPEVISSNFGKASHKSDVYSFGMLLLEMLRNLNOGEELKIRIDEDDDVIIVRKLAIVGWCI
QWNAIDRPSIKVVTQMLEGDSNLTVPSPFTTRNTTHVGPSTQELTIISEL*
>Soltu.DM.S001810.1 pacid=51125200 transcript=Soltu.DM.S001810.1 locus=Soltu.DM.S001810 ID=Soltu.DM.S001810.1.v6.1 annot-version=v6.1
MGAVNELNGIGQKLNKGDSPSGSKDAHEYOKELFAVSMRRFLYIRCYDSFEFLATLKLHLFQLOKEKEAHGSGVYLL
MIDISGAFYMMDRASPSIPPGATTGKPAVSATYLLAGKGPLLKVHCKRSLSLQSVSDIAVQEIQKILVVHPMLVLTAKAA
SLQDKFTSREVIRSSNTAQLYREYMPVWQSFVSHRIFVRPDDSKYQKQPIYLSWLLPSLKLSDKSQSMHMMVHTLYREI
QMILDLWLFST*
>Soltu.DM.S001780.1 pacid=51125201 transcript=Soltu.DM.S001780.1 locus=Soltu.DM.S001780 ID=Soltu.DM.S001780.1.v6.1 annot-version=v6.1

=> Stuberousm_686_v6.1.protein.primaryTranscriptOnly.fa <==
>Soltu.DM.S001790.1 pacid=51125199 transcript=Soltu.DM.S001790.1 locus=Soltu.DM.S001790 ID=Soltu.DM.S001790.1.v6.1 annot-version=v6.1
MGEAGYGIYVKGTLISNEIHVAVKILNDSEGNIEEINEVAAMGNHHVNVVCLVGFCDGPKNILLDNPKISDFGLAKL
CSKEKSVVTMTTEARGTMGYIAPEVISSNFGKASHKSDVYSFGMLLLEMLRNLNOGEELKIRIDEDDDVIIVRKLAIVGWCI
QWNAIDRPSIKVVTQMLEGDSNLTVPSPFTTRNTTHVGPSTQELTIISEL*
>Soltu.DM.S001810.1 pacid=51125200 transcript=Soltu.DM.S001810.1 locus=Soltu.DM.S001810 ID=Soltu.DM.S001810.1.v6.1 annot-version=v6.1
MGAVNELNGIGQKLNKGDSPSGSKDAHEYOKELFAVSMRRFLYIRCYDSFEFLATLKLHLFQLOKEKEAHGSGVYLL
MIDISGAFYMMDRASPSIPPGATTGKPAVSATYLLAGKGPLLKVHCKRSLSLQSVSDIAVQEIQKILVVHPMLVLTAKAA
SLQDKFTSREVIRSSNTAQLYREYMPVWQSFVSHRIFVRPDDSKYQKQPIYLSWLLPSLKLSDKSQSMHMMVHTLYREI
QMILDLWLFST*
>Soltu.DM.S001780.1 pacid=51125201 transcript=Soltu.DM.S001780.1 locus=Soltu.DM.S001780 ID=Soltu.DM.S001780.1.v6.1 annot-version=v6.1

=> Stuberousm_686_v6.1.softmasked.fa <==
<chr01
tagtccacagacttttttgtgatctggaattctaacgtatgttttcccaa
aaatttttgtagcgtacgttaagaccttagctatggagccagttagcgc
tcacggccacaacgtccattttaaaggtaaaattgtccagagcaggaa
aaccctcaattttgccgatttaagtttgctatagtcctgaggttttgg
tgatctgggttccgacgtaaatttgcataaaatttcatagacgtccgt
aaagacgtcagctatggagccagtttagcctcaccggccaaaacgtccct
ttaaaaggctcaattgtgacccagagttggaanaacctaattttgccgatt
ttctgcggtatagtcacgtgacatttttggtgaactagaattccgacaaa
attttgttaaaatttttgcgtatgtccgttaagacatttagctacggag

=> Stuberousm_686_v6.1.transcript.fa <==
>Soltu.DM.S000060.1 pacid=51152479 locus=Soltu.DM.S000060 ID=Soltu.DM.S000060.1.v6.1 annot-version=v6.1
ACTAAACAAGTCTTATAATTTGCTCTATTTATACCAAGTTTAACTAAAGTCAATTTTCAATTAATTTTCTT
GTTGTTAATCTATCTCCATTTTTCACAAAGTTGCTCTCAAAACTTCTCGAGTCTGATGTTTATGTACACTCCACC
TTTGTCTTCTGCTATGTTGACGCTTCACTTCCAAAGTTTGAATGCCAGAAATTTCTATCCAAAGAGGCAGCATACA
AATAAATTAATGATGAGTTGATGCTTGAATGCCAATCAAGGTTGAATTTGGCATCATTTGTGACAACTGGATGGAGCCAG
AATGTGATAGACTTATGATGCTTCAATAAACAGAACTATGTTGATATGGATGAATACCTGTTTACCACCTGAGCTTCAG
```

PLB812_Lab1 Genomic File Download Verification (.gff) screenshot

```
(base) -bash-4.2$ head *.gff3
=> Stuberoseum_686_v6.1.gene_exons.gff3 <=
##gff-version 3
##assembly-version v6.1
##assembly-site MSU
##annot-version v6.1
##annot-site MSU
##species Solanum tuberosum
chr01 phytozomev13 gene 432461 449161 . - . ID=Soltu.DM.01G000010.v6.1;Name=Soltu.DM.01G000010
chr01 phytozomev13 mRNA 432509 449161 . - . ID=Soltu.DM.01G000010.4.v6.1;Name=Soltu.DM.01G000010.4;pac id=51152355;longest=0;Parent=Soltu.DM.01G00
0010.v6.1
chr01 phytozomev13 exon 448964 449161 . - . ID=Soltu.DM.01G000010.4.v6.1.exon.1;Parent=Soltu.DM.01G000010.4.v6.1;pac id=51152355
chr01 phytozomev13 five_prime_UTR 448964 449161 . - . ID=Soltu.DM.01G000010.4.v6.1.five_prime_UTR.1;Parent=Soltu.DM.01G000010.4.v6.1;pac id=51152355

=> Stuberoseum_686_v6.1.gene.gff3 <=
##gff-version 3
##assembly-version v6.1
##assembly-site MSU
##annot-version v6.1
##annot-site MSU
##species Solanum tuberosum
chr01 phytozomev13 gene 432461 449161 . - . ID=Soltu.DM.01G000010.v6.1;Name=Soltu.DM.01G000010
chr01 phytozomev13 mRNA 432509 449161 . - . ID=Soltu.DM.01G000010.4.v6.1;Name=Soltu.DM.01G000010.4;pac id=51152355;longest=0;Parent=Soltu.DM.01G00
0010.v6.1
chr01 phytozomev13 five_prime_UTR 448964 449161 . - . ID=Soltu.DM.01G000010.4.v6.1.five_prime_UTR.1;Parent=Soltu.DM.01G000010.4.v6.1;pac id=51152355
chr01 phytozomev13 CDS 448775 448858 . - 0 ID=Soltu.DM.01G000010.4.v6.1.CDS.1;Parent=Soltu.DM.01G000010.4.v6.1;pac id=51152355

=> Stuberoseum_686_v6.1.repeatmasked_assembly_v6.1.gff3 <=
##gff-version 3
chr01 lcMask similarity 1 200619 . + . ID=2434818.1;Name=rpt1;length=200619
chr01 lcMask similarity 200621 250848 . + . ID=2434818.2;Name=rpt2;length=50228
chr01 lcMask similarity 250868 274935 . + . ID=2434818.3;Name=rpt3;length=24068
chr01 lcMask similarity 274943 294476 . + . ID=2434818.4;Name=rpt4;length=19534
chr01 lcMask similarity 294522 298483 . + . ID=2434818.5;Name=rpt5;length=3962
chr01 lcMask similarity 298491 333370 . + . ID=2434818.6;Name=rpt6;length=34880
chr01 lcMask similarity 333379 340185 . + . ID=2434818.7;Name=rpt7;length=6807
chr01 lcMask similarity 340193 388593 . + . ID=2434818.8;Name=rpt8;length=48401
chr01 lcMask similarity 388597 411295 . + . ID=2434818.9;Name=rpt9;length=22699
(base) -bash-4.2$
```


SRR download/fastq-dump Verification

```
(base) -bash-4.2$ cd mitch987
(base) -bash-4.2$ cd miniconda3
(base) -bash-4.2$ pwd
/mnt/home/mitch987/miniconda3
(base) -bash-4.2$ ls
bin  compiler_compat  condabin  conda-meta  envs  etc  include  lib  LICENSE.txt  pkgs  share  shell  Solanum  ssl  x86_64-conda_cos6-linux-gnu  x86_64-conda-linux-gnu
(base) -bash-4.2$ cd Solanum
(base) -bash-4.2$ ls
Genomic  RNA_Seq
(base) -bash-4.2$ cd RNA_Seq/
(base) -bash-4.2$ ls
SRR5448223  SRR5448225_1.fastq.gz  SRR5448227  SRR5448228_2.fastq.gz  SRR5448230_1.fastq.gz  SRR5448232  SRR5448233_2.fastq.gz  SRR5448235_1.fastq.gz
SRR5448223_1.fastq.gz  SRR5448225_2.fastq.gz  SRR5448227_1.fastq.gz  SRR5448229  SRR5448230_2.fastq.gz  SRR5448232_1.fastq.gz  SRR5448234  SRR5448235_2.fastq.gz
SRR5448223_2.fastq.gz  SRR5448226  SRR5448227_2.fastq.gz  SRR5448229_1.fastq.gz  SRR5448231  SRR5448232_2.fastq.gz  SRR5448234_1.fastq.gz
SRR5448224  SRR5448226_1.fastq.gz  SRR5448228  SRR5448229_2.fastq.gz  SRR5448231_1.fastq.gz  SRR5448233  SRR5448234_2.fastq.gz
SRR5448225  SRR5448226_2.fastq.gz  SRR5448228_1.fastq.gz  SRR5448230  SRR5448231_2.fastq.gz  SRR5448233_1.fastq.gz  SRR5448235
(base) -bash-4.2$ █
```

fastq.gz unzip Verification

```
(plb812) -bash-4.2$ ls
fastqc_data  SRR5448225_1.fastq  SRR5448227_1.fastq  SRR5448229_1.fastq  SRR5448231_1.fastq  SRR5448233_1.fastq  SRR5448235_2.fastq
SRR5448223_2.fastq  SRR5448225_2.fastq  SRR5448227_2.fastq  SRR5448229_2.fastq  SRR5448231_2.fastq  SRR5448233_2.fastq
SRR5448224_1.fastq  SRR5448226_1.fastq  SRR5448228_1.fastq  SRR5448230_1.fastq  SRR5448232_1.fastq  SRR5448234_1.fastq
SRR5448224_2.fastq  SRR5448226_2.fastq  SRR5448228_2.fastq  SRR5448230_2.fastq  SRR5448232_2.fastq  SRR5448235_1.fastq
(plb812) -bash-4.2$ █
```

```
(base) -bash-4.2$ head SRR5448235_1.fastq
@SRR5448235.1 FCC3MNJACXX:6:1101:1499:2149# length=100
GTTCCGGAAGCATGCTTCTTCTCAGAGCATCAATCAGATCATTTGCATAGAGGTAAGGATCGGTAGGCATCCCAATCACTCCAGCACCTCCATGATCAC
+SRR5448235.1 FCC3MNJACXX:6:1101:1499:2149# length=100
CCCCFFFFHHHHFIJJJJJJJJJJGHHGHIJJJGHHGIIJJIIJJIGIJJJJIFFHIIJIHGI;CBHHHHHHFFDFEEEEEDDDDDDDDDCDDDDDED
@SRR5448235.2 FCC3MNJACXX:6:1101:1529:2072# length=100
CNCCCAGAAAGTTTGGAGTTGTTGCTGCAATTGTTGCTGCTGCTGATGAATGTGCTGATAAGCGAGCTGATGTTGTGCCATCTGGGCTGACGAAGCA
+SRR5448235.2 FCC3MNJACXX:6:1101:1529:2072# length=100
C#1ADDFFGHHFFIEIHHHHJJJJJJIIJJJJJEHIJJIGIJJJJJJGHHGIIHHEIGGII:FHEGHEE@DEFBEFCCECCDDDDDBDDDDDB?=?B>
@SRR5448235.3 FCC3MNJACXX:6:1101:1625:2080# length=100
GCCCACTCATTTGTTCTCTCCTAATCTTGAAGTATCCATCTTACCCAGCCTCTATTCCAAGAATTAACAATGAGCCAATAGTCCTCTCCTTGCTCA
(base) -bash-4.2$ █
```