## PLB812\_Lab1 Genomic File Download Verification

```
(base) -bash-4.2$ pwd
/mnt/home/mitch987
(base) -bash-4.2$ ls
                                                            mitch987@hpcc.msu.edu:hom
                                                                                                                    mitch987@hpcc.msu,?.edu:homemitch987minicondaSolanum
homemitch987miniconda3 Miniconda3-latest-Linux-x86 64.sh
                                                           mitch987@hpcc.msu.edu:homemitch987??[Cminiconda3Solanum mitch987@msu.edu:????????[C?[C?[C?[C?[C?[C?[C?
homemitch987mobaxterm Miniconda3-latest-Linux-x86 64.sh.1 mitch987@hpcc.msu.edu:homemitch987miniconda3Solanum
(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.repeatmasked assembly v6.1.qff3
Stuberosum 686 v6.1.annotation info.txt
                                                 Stuberosum 686 v6.1.gene exons.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$
```

```
(Wase) "Obs.N=4.29 | NEW -1.16.45.fa ←

Stuberosum 866 v6.1.cds.fa ←
>Soltu.DM.S001790 I Dacid=51125199 polypeptide=Soltu.DM.S001790.1 locus=Soltu.DM.S001790 ID=Soltu.DM.S001790.1.v6.1 annot-version=v6.1
ATGGGGAGGGGGCTTACGGGATCGTTTATAAGGAACACTTTCGAATGAAATTCATGTGCTGTAAAAATTCTCAATGA
 TTCTGAAGGAAATGGGGAAGAATTTATTAATGAAGTTGCAGCAATGGGGAATATCCATCATGTCAACGTGGTTTGCCTAG
TTGGCTTCTGTGCTGATGGACCAAAAAATATTCTGTTAGACCTGAACCCCAAGATCTCTGATTTTTGGTCTTACTAAGTTA
TGCTCTAAAGAGAAAAGTGTTGTCACAATGACTGAAGCTAGAGGAACCATGGGTTATATTGCACCAGAAGTGATATCAAG
CAATTTCGGAAAAGCGTCTCACAAATCTGATGTTTATAGCTTTGGAATGCTGCTTCTTGAAATGCGGCTGAATCAAGGAG
AAGAGTTGAAAATCAGGATTGATGAAGATGATGATGTTATAATTGTAAGGAAATTGGCTGTTATAGGAGTTTGGTGCATC
CAATGGAATGCGATAGATCGGCCATCCATCAAAGTTGTAACTCAAATGCTAGAAGGAGATGGGAGTAACCTAACTGTTCC
TTCTCCTTTTACAACAAGAAACACCACCCATGTAGGTCCTTCTACGCAAGAGTTGACGATAATCTCAGAACTAGAATGA
>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
  ⇒ Stuberosum 686 v6.1.cds primaryTranscriptOnly.fa ←
>Soltu.DM.S001790.I pacid=51125199 polypeptide=Soltu.DM.S001790.1 locus=Soltu.DM.S001790 ID=Soltu.DM.S001790.1.v6.1 annot-version=v6.1 ATGGGAGGGGGGCCTTACGGGATCGTTTATAAAGGAACACTTTCGAATGAAATTCATGTTGCTGTAAAAATTCTCAATGA
TTCTGAAGGAAATGGGGAAGAATTTATTAATGAAGTTGCAGCAATGGGGAATATCCATCATGTCAACGTGGTTTGCCTAG
TTGGCTTCTGTGCTGATGGACCAAAAAATATTCTGTTAGACCTGAACCCCAAGATCTCTGATTTTGGTCTTGCTAAGTTA
TGCTCTAAAGAGAAAAGTGTTGTCACAATGACTGAAGCTAGAGGAACCATGGGTTATATTGCACCAGAAGTGATATCAAG
CAATTTCGGAAAAGCGTCTCACAAATCTGATGTTTATAGCTTTGGAATGCTGCTTCTTGAAATGCGGCTGAATCAAGGAG
AAGAGTTGAAAATCAGGATTGATGAAGATGATGATGTTATAATTGTAAGGAAATTGGCTGTTATAGGAGTTTGGTGCATCCAATGGAATGCGATAGATCGGCATCCAACTGTTAAAGTTGTAACTCAAATGCTAGAAGGAGATGGGAGTAACCTAACTGTTCC
TTCTCCTTTTACAACAAGAACACCACCCATGTAGGTCCTTCTACGCAAGAGTTGACGATAATCTCAGAACTAGAATGA
>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
 ⇒ Stuberosum_686_v6.1.fa ←
>chr01
TAGTCCACAGACTTTTTTGTGATCTGGAATTCTAACGTATGTTTTCCCAAAAATTTTCGT
GGACGTACGTTAAGACCTTAGCTATGGAGCCAGTTAGCCCTCACGGCCACAACGTCCCAT
TTTAAAGGTAAAATGTTCCCAGAGCAGGAAAACCCCCCAATTTTGCCGATTTAAGTTTGCT
ATAGTCCATGGATTGTTTGGTGATCTGGGTTCCGACGTAAATTTTGTCAAAAATTTTCAT
AGACGTCCGTAAAGACGTCAGCTATGGAGCCAGTTAGCCCTCACGGCCAAAACGTCCCCT
```

-bash-4.2\$ head \*.fa

TTAAAAGGTCAAATGTGACCCAGAGTTGGAAAACCCTAATTTTGCCGATTTTCGTGCGGT ATAGTCCATGGACTTTTTGGTGAACTAGAATTCCGACAAAATTTTTGTTAAAAATTTTCG

TGGATGTCCGTTAAGACATTAGCTACGGAGCCAGTTAGCCCTCACGACCTAAACATCTCA

TTTTGAAAGTCAAATCTGCCCCGGAACAGGTAAACCCCCCATTTTGCCGAATATCGTGTG

⇒ Stuberosum\_686\_v6.1.hardmasked.fa ←

# ⇒ Stuberosum\_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 ACTAAACAAAGTCCTTATAATTTCCTCTATTTTATACCAAAGTTTAACTAAACTTCAATT CTTTTCAATTAATTTTCTTGTTGTTATTTATGTACTCCCATTTTTGCAACAATGGTTCTCC AAAAAACTTCCTCGGAAGTCTGATGTTTATGTACACTCCACCCTTTGCTTTCTGCTATGTTCG AGCTTCACTTCCAAGGTTTGAGGTGCCAGAAAATTCTATCCCAAAAGAGGCAGCATACCA AATAATTAATGATGAGTTGATGCTTGATGGCAATCCAAGGTTGAATTTGGCATCATTTGT GACAACATGGATGGACCAGAAATGTGAAAGCTTATAGTGTTCTAATAAACAAGAACTA TGTTGATATGGATGACCCAGAATGTGATAAGCTTATCAAATCAAGAACTA TGTTGATATGGATGACCCCTTTTGAAAGAAGAAAATAGCAATTGGTGTTCAAATAACAAGAACTA GGCTCATCAGAAGGCCCATATAGGTTGAAAGAAGAAAATAGCAATTGGTGTGGAATATCAT AGCACCGTTTATTCAACGCACCCTTTGAAAGAAGAAAATAGCAATTGGTGTGGACACAT GGGCTCATCAGAAGGCCTATAATGTTAGCGGGCCTCAGACCTTCAAGAGGAAAACAA (base) -bash-4.2\$

## PLB812\_Lab1 Genomic File Download Verification (.fa) screenshots

```
-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
MGEGAYGIVYKGTLSNEIHVÄVKILNDSEGNGEEFINEVAAMGNIHHVNVVCLVGFCADGPKNILLDLNPKISDFGLAKL
CSKEKSVYTMTEARGTMGYIAPEVISSNFGKASHKSDVYSFGMLLLEMRLNQGEELKIRIDEDDDVIIVRKLAVIGVWCI
OWNAIDRPSIKVVTOMLEGDGSNLTVPSPFTTRNTTHVGPSTQELTIISELE*
MGAVNELGNGIGQKLNKGDDSPSGSKDAHTEYDKELFAVSMRRFLYIRCYDSFEFLATLKTLHFQLQKEKEAHGSGVYLL
MIDSIGAFYWMDRASPSIPPGATTGKPAVSATYLLAGKGPLLVKHCKRSLSLOSVSDIAVOEIOKILVVHPMLVLTTKAA
SLODKETSREVIRSSTNAOLYREYMPSVWOSEVSHRIEVRPDDSKYOKOPIYLSEWLLPSLKLSDKSOSMMVHTLYRET
->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
 ⇒ Stuberosum_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
MGEGAYGIVYKGTLSNEIHVÄVKILNDSEGNGEEFINEVAAMGNIHHVNVVCLVGFCADGPKNILLDLNPKISDFGLAKL
 SKEKSVVTMTEARGTMGYTAPEVTSSNEGKASHKSDVYSEGMLLLEMRI NOGEELKTRIDEDDDVITVRKI AVIGWCI
 WNAIDRPSIKVVTQMLEGDGSNLTVPSPFTTRNTTHVGPSTQELTIISELE*
 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
MGAVNELGNGIGQKLNKGDDSPSGSKDAHTEYDKELFAVSMRRFLYIRCYDSFEFLATLKTLHFQLQKEKEAHGSGVYLL
MIDSIGAFYWMDRASPSIPPGATTGKPAVSATYLLAGKGPLLVKHCKRSLSLOSVSDIAVOFIOKILVVHPMLVLTTKAA
 .
IQDKFTSREVIRSSTNAQLYREYMPSVWQSFVSHRIFVRPDDSKYQKQPIYLSEWLLPSLKLSDKSQSMMMVHTLYREI
 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
 ⇒ Stuberosum_686_v6.1.softmasked.fa ←
tagtccacagacttttttgtgatctggaattctaacgtatgttttcccaa
aaattttcgtggacgtacgttaagaccttagctatggagccagttagccc
tcacggccacaacgtcccattttaaaggtaaaatgttcccagagcaggaa
  cccccaattttgccgatttaagtttgctatagtccatggattgtttgg
 gatctgggttccgacgtaaattttgtcaaaaattttcatagacgtccgt
aaagacgtcagctatggagccagttagccctcacggccaaaacgtcccct
ttaaaaggtcaaatgtgacccagagttggaaaaccctaattttgccgatt
ttcgtgcggtatagtccatggactttttggtgaactagaattccgacaaa
atttttgttaaaaattttcgtggatgtccgttaagacattagctacggag
 ⇒ Stuberosum_686_v6.1.transcript.fa ←
GTTGTTAATTCATCTCCATTTTTGCAACAATGGTTCTCTCAAAAACTTCCTCGGAGTCTGATGTTTATGTACACTCCACC
  TGCTTCTCGCTATGTTCGAGCTTCACTTCCAAGGTTTGAGATGCCAGAAAATTCTATCCCAAAAGAGGCAGCATACCA
AATGTGATAAGCTTATGATGTCTTCAATAAACAAGAACTATGTTGATATGGATGAATACCCTGTTACCACTGAGCTTCAG
```

# PLB812\_Lab1 Genomic File Download Verification (.gff) screenshot

```
(base) -bash-4.2$ head *.gff3

⇒ Stuberosum 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
                                432461 449161 .
                                                                       ID=Soltu.DM.01G000010.v6.1;Name=Soltu.DM.01G000010
chr01 phytozomev13
                                       449161 .
                                                                       ID=Soltu.DM.01G000010.4.v6.1; Name=Soltu.DM.01G000010.4; pacid=51152355; longest=0; Parent=Soltu.DM.01G00
0010.v6.1
chr01 phytozomev13
                                448964 449161
                                                                       ID=Soltu.DM.01G000010.4.v6.1.exon.1; Parent=Soltu.DM.01G000010.4.v6.1; pacid=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;pacid=51152355
 ⇒ Stuberosum_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
                               432461 449161 .
                                                                       ID=Soltu.DM.01G000010.v6.1; Name=Soltu.DM.01G000010
chr01 phytozomev13
                                                                       ID=Soltu.DM.01G000010.4.v6.1; Name=Soltu.DM.01G000010.4; pacid=51152355; longest=0; Parent=Soltu.DM.01G00
                               432509 449161 .
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;pacid=51152355
       phytozomev13
                               448775 448858 .
                                                                       ID=Soltu.DM.016000010.4.v6.1.CDS.1; Parent=Soltu.DM.016000010.4.v6.1; pacid=51152355

⇒ Stuberosum_686_v6.1.repeatmasked_assembly_v6.1.gff3 ←
##gff-version 3
chr01
      lcMask similarity
                                                                       ID=2434818.1; Name=rpt1; length=200619
chr01
       lcMask similarity
                                200621 250848 .
                                                                       ID=2434818.2; Name=rpt2; length=50228
       lcMask similarity
                                250868
                                       274935 .
                                                                       ID=2434818.3; Name=rpt3; length=24068
chr01
                                274943 294476 .
                                                                       ID=2434818.4; Name=rpt4; length=19534
chr01
        lcMask similarity
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
        lcMask similarity
                                                                       ID=2434818.8; Name=rpt8; length=48401
chr01
                                340193
                                       388593
                                                                       ID=2434818.9; Name=rpt9; length=22699
chr01
       lcMask similarity
                                388597 411295 .
(base) -bash-4.2$
```

### SRR download/fastq-dump Verification

```
(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
                                                                SRR5448228 2.fastq.gz SRR5448230 1.fastq.gz SRR5448232
                                                                                                                                 SRR5448233 2.fastq.gz SRR5448235 1.fastq.gz
                                                                                      SRR5448230_2.fastq.gz
                                          SRR5448227 1.fastq.qz
SRR5448223 1.fastq.qz
                                                                                                           SRR5448232 1.fastq.gz
                                                                                                                                                       SRR5448235 2.fastq.qz
                                           SRR5448227 2.fastq.gz
                                                                                                            SRR5448232_2.fastq.gz
SRR5448223 2.fastq.gz SRR5448226
                                                                                     SRR5448231
                     SRR5448226 1.fastq.gz SRR5448228
                                                                SRR5448229 2.fastq.gz
                                                                                     SRR5448231 1.fastq.gz SRR5448233
                                                                                                                                 SRR5448234 2.fastq.gz
                     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

SRR5448223_1.fastq 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$
```