

PLB812_Lab1 Genomic File Download Verification

```
bash: cd: miniconda: No such file or directory
(base) -bash-4.2$ cd miniconda3
(base) -bash-4.2$ cd Solanum
(base) -bash-4.2$ cd Genomic
(base) -bash-4.2$ ls
cds_from_genomic.fna  chrPltd.fna  genomic.gbff  genomic.gtf  ncbi_error_report.txt  protein.faa  rna.fna  unplaced.scaf.fna
(base) -bash-4.2$
```

PLB812_Lab1 Genomic File Download Verification: CDS

```
(base) -bash-4.2$ head cds_from_genomic.fna
>lcl|NW_006238927.1_cds_XP_006337956.1_1 [gene=L0C102578245] [db_xref=GeneID:102578245] [protein=universal stress protein A-like protein isoform X1] [protein_id=XP_006337956.1] [location=join(73306..73525,73674..73777,73855..73973,75507..75603)] [gbkey=CDS]
ATGGAGGCAGAGACCAGTGCAGCTAATGCTCCAGCAAGTGCAGCTAGTGATACTCCGGTTACTATGAAAAAGATGATAAT
GGTGGCTATCGATGAGAGCGAGGAGAGTTTCTATGCTCTCAAATGGGCATTAGATAATTGATCAACGATCCTTCTATTA
TCATCACTTTGATCAATGTCCAACCTTCCTTTACCCCATGGTTTATCCTGCTGGACCTGTTGTGTTGCGGCGCCTACT
GTTGTTGAAGCTGTGAAAAAAGGACAACACGAGAATGCAACTAGAACTACTGTCTCGGGCACTCCATCTCTGCAAAACAAA
GATGTTGAAGCGGAGACTCTGATCCTAGATGGGGACCCAAAAGATATGATATGCAAGTGCACATGAACTGCACGCTG
ACCTTCTAGTAGTCGGTAGCCGAGGGCTCGGCAAGATTAAAAGGGCTTCTTGGGTAGTGTGAGCGACTATTGTGTCAT
CATGTGCAATGCCCTATTCTCATTGTGAAGCCATCCAAGGAAATCCCCAAAAGCAGCTAA
>lcl|NW_006238927.1_cds_XP_015167560.1_2 [gene=L0C102578245] [db_xref=GeneID:102578245] [protein=universal stress protein A-like protein isoform X2] [protein_id=XP_015167560.1] [location=join(73306..73525,73674..73777,73855..73973,74172..74187)] [gbkey=CDS]
ATGGAGGCAGAGACCAGTGCAGCTAATGCTCCAGCAAGTGCAGCTAGTGATACTCCGGTTACTATGAAAAAGATGATAAT
(base) -bash-4.2$
```

PLB812_Lab1 Genomic File Download Verification: CHR

```
(base) -bash-4.2$ head chrPltd.fna
>NC_008096.2 Solanum tuberosum chloroplast, complete genome
CTGTCTAATATCTCATATTTATGAATTATGGGCGAACGACGGGAATTGAACCCGCGCATGGTGGATTACAAATCCACTGC
CTTGATCCACTTGGCTACATCCGCCCCCTCGCCTACTTACATTCCATTTTTACATTATTTAAATTCGAAAACAAAAGATT
CAAGTTCGAATATTTATCTTCTTTCTTATTTCAATGATATTATTATTTGATTATTATTTCAAAGATAAGAATATGAAGT
AAAAATTGTATTTTTGTGAAATGAAATAAAAAAGATATAGTAACATTAGCAACAAGAGGAACAAGTTATATTTCTACAA
TTTTAAATAAATAGAAAATAAAAAATCGAATACTCAATCATGAATAAATGCAAGCAAATACCCTCTCTTTCTTTTCGATA
ATGTAAACAAAAAAGTCTATGTCAGTAAAATACTAGGAAATTAATAAAGAAAAAAAAAAAAAGAAAGGAGCAATAGCACCT
CTTGACAGAACAAGAAAATGATTATTGCTCCTTTCTTTCAAACCTCCTATAGACTAGGCTGGGATCTTATCCATTTGT
AGATGGAGCTTCGATAGCAGCTAGGTCTAGAGGGAAGTTATGAGCATTACGTTTCATGCATAACTTCCATACCAAGGTTAG
CACGGTTGATGATATCAGCCCAAGTGTTAATTACACGACCCTGACTGTCAACTACAGATTGGTTGAAATTGAAACCATT
(base) -bash-4.2$ █
```

PLB812_Lab1 Genomic File Download Verification: GBFF

```
(base) -bash-4.2$ head genomic.gbff
LOCUS      NW_006238927      7090834 bp      DNA      linear      CON 05-JAN-2016
DEFINITION Solanum tuberosum cultivar DM 1-3 516 R44 unplaced genomic
            scaffold, SolTub_3.0 scf00001, whole genome shotgun sequence.
ACCESSION  NW_006238927
VERSION    NW_006238927.1
DBLINK     BioProject: PRJNA225997
            BioSample: SAMN02981305
            Assembly: GCF_000226075.1
KEYWORDS   WGS; RefSeq.
SOURCE     Solanum tuberosum (potato)
(base) -bash-4.2$
```

PLB812_Lab1 Genomic File Download Verification: RNA

```
(base) -bash-4.2$ head rna.fna
>NM_001287825.1 Solanum tuberosum MADS transcriptional factor (STMADS11), mRNA
ATGGTGAGACAAAAATCCAGATCAAGAAGATAGACAATTTGACAGCAAGACAAGTGACATTTTCAAAGAGAAGAAGAGG
GCTTTTAAAGAAAGCTCAAGAGCTTTCAACTCTTGTGATGCTGATATTGGACTCATTGTTTCTCTGCTACTGGAAAAC
TTTTTGAGTATTCAAGCTCCAGTATGATGCAACTGATTGAGAAACACAAGATGCAGTCAGAAAGGGATAGTATGGATAAT
CCAGAACAGCTGCATTCTTCTAACCTTCTGAGCGAGAAGAAAACCCATGCAATGCTAAGCAGGGACTTTGTGGAGAAGAA
TCGGGAGTTAAGGCAACTACATGGAGAAGAGCTGCAAGGACTTGGCTTGGATGATTTAATGAAATTGGAGAAATTAGTCG
AGGGAGGAATAAGTCGTGTCCTCAGAATTAAGGGTGACAAGTTTATGAAAGAGATCAGTTCCTCAAGAAAAAGGAAGCT
CAACTCCAGGAAGAGAATTCACAGTTGAAACAGCAATCACAAGCAAGATTGAATGAAGAAGGGCAAAATGTAATTGAGCA
AGGACATTCAGCAGACTCCATCACCAACAATCGCAGCTTAGTCAATAGTCATCAGGACTACAATGACTCTGATACAAGCC
TCAAGTTATGCTTGGCTTTTCCTTAATGAGCTCGATCAACATGAAATGGGTGATATATAGTTACTGAAACAAGGGCGAG
```

PLB812_Lab1 Genomic File Download Verification: SCAF

```
(base) -bash-4.2$ head unplaced.scaf.fna
>NW_006238927.1 Solanum tuberosum cultivar DM 1-3 516 R44 unplaced genomic scaffold, SolTub_3.0 scf00001, whole genome shotgun sequence
ATTTATTTCCCATTTTATTTATGTATTATTAATGTTTCTGAATAATACCTACCATGATGTTCTATCGTATTGCCATCAT
ACCTATATTTTAATATTCTCCATAACTTTTTCACTCGCATCTAATTCTATTATATAAGTAAATTTTTGTTTCATCT
GGAAAGTATAGTTTAGGTAAGTTTTACATAACAATTTACCTTTTGTAATTTGTAATTTGGTCTCTTTTATTAAGTAT
ATTCTATATCTTTATTTAATTTTTGTTGAATGGTTTTAAATCTCTGCTACTTTTGGTATATATTCTTCTACTTGATT
ACTAGTCCTAAGAATGATTGAATTTCTTTTCGTATCTAGTTGTTCAATTTGAGTTAATTATTTTGTACTATATGTGCT
TGCATTTTTATTCCATTTTATCTATTTGAATTCCTAAAAATTCGATTTGCTTTTTCATTATGCTGCTTTCATTTTACT
TAGACTTATGCCTGAATACTCTATAATATGTATAAATTTTTCTAATAATCTTATGTGTTTCATCTTGTGTTTGGAAATACA
GTAATATATCATCAATATATACTACGCAATTTCTAGTTGTTTAAATAGTTATCCATAAAGTGTGGTATCTACCTGGT
GCATTTTATATCCAAATGGTAATACATTCCATTTATAAAATCCTTGCGGTACCGTAAATGTTGTATATTATTAGATTC
(base) -bash-4.2$
```


PLB812_Lab1 Genomic File Download Verification: GTF

```
(base) -bash-4.2$ head genomic.gtf
#gtf-version 2.2
#!genome-build SolTub_3.0
#!genome-build-accession NCBI_Assembly:GCF_000226075.1
#!annotation-source NCBI Solanum tuberosum Annotation Release 101
NW_006238927.1 Gnomon gene 73207 75755 . + . gene_id "LOC102578245"; db_xref "GeneID:102578245"; gbkey "Gene"; gene "LOC102578245"; gene_biotype "protein_coding";
NW_006238927.1 Gnomon exon 73207 73525 . + . gene_id "LOC102578245"; transcript_id "XM_015312074.1"; db_xref "GeneID:102578245"; gbkey "mRNA"; gene "LOC102578245"; model_evidence "Supporting evidence includes similarity to: 5 ESTs, and 100% coverage of the annotated genomic feature by RNAseq alignments, including 20 samples with support for all annotated introns"; product "universal stress protein A-like protein, transcript variant X2"; exon_number "1";
NW_006238927.1 Gnomon exon 73674 73777 . + . gene_id "LOC102578245"; transcript_id "XM_015312074.1"; db_xref "GeneID:102578245"; gbkey "mRNA"; gene "LOC102578245"; model_evidence "Supporting evidence includes similarity to: 5 ESTs, and 100% coverage of the annotated genomic feature by RNAseq alignments, including 20 samples with support for all annotated introns"; product "universal stress protein A-like protein, transcript variant X2"; exon_number "2";
NW_006238927.1 Gnomon exon 73855 73973 . + . gene_id "LOC102578245"; transcript_id "XM_015312074.1"; db_xref "GeneID:102578245"; gbkey "mRNA"; gene "LOC102578245"; model_evidence "Supporting evidence includes similarity to: 5 ESTs, and 100% coverage of the annotated genomic feature by RNAseq alignments, including 20 samples with support for all annotated introns"; product "universal stress protein A-like protein, transcript variant X2"; exon_number "3";
NW_006238927.1 Gnomon exon 74172 74488 . + . gene_id "LOC102578245"; transcript_id "XM_015312074.1"; db_xref "GeneID:102578245"; gbkey "mRNA"; gene "LOC102578245"; model_evidence "Supporting evidence includes similarity to: 5 ESTs, and 100% coverage of the annotated genomic feature by RNAseq alignments, including 20 samples with support for all annotated introns"; product "universal stress protein A-like protein, transcript variant X2"; exon_number "4";
NW_006238927.1 Gnomon CDS 73306 73525 . + 0 gene_id "LOC102578245"; transcript_id "XM_015312074.1"; db_xref "GeneID:102578245"; gbkey "CDS"; gene "LOC102578245"; product "universal stress protein A-like protein isoform X2"; protein_id "XP_015167560.1"; exon_number "1";
(base) -bash-4.2$
```

PLB812_Lab1 Genomic File Download Verification: Protein

```
(base) -bash-4.2$ head protein.faa
>NP_001274754.1 MADS transcriptional factor [Solanum tuberosum]
MVRQKIQIKKIDNLTARQVTFKRRRGLFKKAQELSTLCDADIGLIVFSATGKLFEYSSSSMMQLIEKHKMQSERDSMDN
PEQLHSSNLLSEKKTHAMLSRDFVEKNRELQRLHGEELQGLGLDDLMKLEKLVEGGISRVLRIGDKFMKEISSLKKKEA
QLQEENSQLKQQSQARLNEEGQNVIEQGHSAISITNNRSLVNSHQDYNDSDTSLKLCLAFP
>NP_001274757.1 ethylene receptor homolog [Solanum tuberosum]
MESDCIEALLPTGDLLVKYQYLSDFFIAYVFSIPLELIYFVHKSACFPYRWVLMQFGAFIVLCGATHFISLWTFMHS
KTVAVVMTISKMLTAAVSCITALMLVHIIPDLLSVKTRELFLKTRAEELDKEMGLIIRQEETGRHVRMLTHEIRSTLDRH
TILKTTLVELGRTDLAECALWMPQCQGLTLQLSHNLNNLIPLGSTVPINLPIINEIFSSPEAIQIPHTNPLARMRNTVG
RYIPPEVVAVRVPLLHLSSFTNDWAELESTRSYAVMVLVLPNGLRKWREHELELVQVVADQVAVALSHAAILED SMRAHD
QLMEQNNALDDARQEAEMAIRARNDFLAVMNHMRTPMHAVIALCSLLLETDLTPEQRVMIETILKSSNLLATLINDVLD
(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$
```

PLB812_Lab1 SRR Download Verification

```
(base) -bash-4.2$ head SRR5448235_1.fastq
@SRR5448235.1 FCC3MNJACXX:6:1101:1499:2149# length=100
GTTCCGGAAGCATGCTTCTTCTTCAGAGCATCAATCAGATCATTTGCATAGAGGTAAGGATCGGTAGGCATCCCAATCACTCCAGCACCTCCATGATCAC
+SRR5448235.1 FCC3MNJACXX:6:1101:1499:2149# length=100
CCCCFFFFHHHFIJIIJJJJJJJJJIGHGHIJJJGHIGIIJJIIJJIGIJJJJIFFHIIJIHGI;CBHHHHHHFFDFEEEEEDDDDDDDDDCDDDDDED
@SRR5448235.2 FCC3MNJACXX:6:1101:1529:2072# length=100
CNCCCAGAAAGTTTGGAGTTGTTGCTGCAATTGTTGCTGCTGCTGCTGATGAATGTGCTGATAAGCGAGCTGATGTTGTGCCATCTGGGCTGACGAAGCA
+SRR5448235.2 FCC3MNJACXX:6:1101:1529:2072# length=100
C#1ADDFFGHHFFIEIHHHHJIIJJJJIIJJJEHIJJIGIIJJJJJJGHJGIHIHIHEIGGII:FHEGHEE@DEFBEFCCECCDDDDDBDDDDDB?=?B>
@SRR5448235.3 FCC3MNJACXX:6:1101:1625:2080# length=100
GCCACACTCATTTGTTCTCTCCTAATCTTGAAGTATCCATCTTCACCCAGCCTCTATTCCAAGAATTAACAATGAGCCAATAGTCCTCTCCTTGCTCA
(base) -bash-4.2$
```