**Table S3. Input, output, and parameters of the BG7 annotation programs.**

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**Input:**

XX\_proteins\_tBLASTn.xml

XX\_Contigs.fna

*Extension\_theshold*=400

*Virus\_flag* =false

**PredictGenes.jar**

**Output:**

XX\_PredictedGenes.xml

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**Input:**

XX\_PredictedGenes.xml

**RemoveDuplicatedGenes.jar**

**Output:**

XX\_NotDuplicatedGenes.xml

XX\_Dismissed.xml

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**Input:**

XX\_NotDuplicatedGenes.xml

*Overlapping\_threshold*=102

XX\_RNA\_blastn.xml

XX\_Contigs.fna

**SolveOverlappings.jar**

**Output:**

XX\_SolvedOverlappings.xml

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**Input:**

XX\_SolvedOverlappings.xml

**GenerateFastaFiles.jar**

**Output:**

**XX\_protein\_nucleotide\_sequences.fasta**

**XX\_protein\_aminoacid\_sequences.fasta**

**XX\_Dismissed\_protein\_nucleotide\_sequences.fasta**

**XX\_Dismissed\_protein\_aminoacid\_sequences.fasta**

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**Input:**

XX\_SolvedOverlappings.xml

**FillDataFromUniprot.jar**

**Output:**

**XX\_BG7\_Annotation.xml**

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**Input:**

XX\_BG7\_Annotation.xml

**GenerateCSVFile.jar**

**Output:**

**XX\_BG7\_Annotation.tsv**

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**Input:**

XX\_BG7\_Annotation.xml

**RemoveDismissedGenes**

**Output:**

**XX\_BG7\_final\_Annotation.xml**

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**Input:**

XX\_BG7\_Annotation.xml

**GenerateGffFile.jar**

**Output:**

**XX\_BG7\_Annotation.gff**

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**Input:**

XX\_BG7\_Annotation.xml

**ExportEmblfiles.jar**

**Output:**

**XX\_BG7\_Annotation.embl**

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**Input:**

XX\_BG7\_Annotation.xml

**ExportGenBankfiles.jar**

**Output:**

**XX\_BG7\_Annotation.gbk**

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**Input:**

XX\_BG7\_Annotation.xml

**ExportGenBankfiles.jar**

**Output:**

**XX\_BG7\_Annotation\_5\_columns\_gbk.txt**

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**Input:**

XX\_BG7\_Annotation.xml

XX\_Contigs.fna

**GetIntergenicSequences.jar**

**Output:**

**XX\_Intergenic.xml**

**XX\_Intergenic.fasta**

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**Quality control**

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**Input:**

XX\_PredictedGenes.xml

XX\_Dismissed.xml

XX\_NotDuplicatedGenes.xml

**BasicQualityControl.jar**

**Output:**

XX\_BasicQualityControl.xml

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**Input:**

XX\_ReferenceProteins.fasta

XX\_proteins\_tBLASTn.xml

XX\_BG7\_Annotation.xml

**AutomaticQualityControl.jar**

**Output:**

XX\_AutomaticQualityControl.txt

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**Input:**

XX\_BG7\_Annotation.xml

XX\_BG7\_Annotation.gbk

**ControlGenBankfilesquality.jar**

**Output:**

XX\_GBK\_QC.txt

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**Input:**

XX\_BG7\_Annotation.xml

XX\_BG7\_Annotation\_5\_columns\_gbk.txt

**Control5columnsGenBankfilesquality.jar**

**Output:**

XX\_GBK\_5\_col\_QC.txt

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_