Gene locations

C:\Users\nesoltis\Documents\Projects\BcSolGWAS\data\genome\WGS\new\_gene\_annot2.tar.gz

Fasta sequences for each isolate: C:\Users\nesoltis\Documents\Projects\BcSolGWAS\data\genome\WGS\87\_iso\_no\_organic.fasta.tar.gz

**SNPdat to annotate SNPs**

<https://code.google.com/archive/p/snpdat/>

<https://github.com/agdoran/snpdat>

package and manual is now in Programs folder

need Perl to run.

* Options for Windows: linux emulator (Cygwin) / strawberry perl/ activestate perl
* Strawberry and activestate look similar. Trying strawberry <http://www.perlmonks.org/?node_id=1060809>

Save as data/SNPdat\_Annotate/Final\_annots/…\_FORPERL.txt with the following format:

chromosome\_id position mutation

Chromosome1 12045 A

Chromosome1 51226 A

Strawberry Perl

> perl Programs\SNPdat\_package\_ v1.0.5\SNPdat\_v1.0.5.pl

Usage: perl SNPdat –i Input\_file –f Fasta\_file –g Gene\_transfer\_file

–o

The version for Domestication:

perl Programs/SNPdat\_package\_v1.0.5/SNPdat\_v1.0.5.pl -i Projects/BcSolGWAS/data/SNPdat\_Annotate/Domestication\_TopSNPs\_SegLong\_trueMAF20\_10NA\_FORPERL.txt -f Projects/BcSolGWAS/data/Annotate/suziT4.fasta -g Projects/BcSolGWAS/data/SNPdat\_Annotate/genes\_Chromosome.gtf

For IndPlants:

Input file from "data/GWAS\_files/05\_annotation/12Plants\_Top1000SNPs\_SegWide\_trueMAF20\_10NA.csv"

perl Programs/SNPdat\_package\_v1.0.5/SNPdat\_v1.0.5.pl -i Projects/BcSolGWAS/data/Annotate/Domestication\_TopSNPs\_SegLong\_trueMAF20\_10NA\_FORPERL.txt -f Projects/BcSolGWAS/data/Annotate/suziT4.fasta -g Projects/BcSolGWAS/data/Annotate/genes\_Chromosome.gtf

For High overlap SNPs (HOSNPs)

perl Programs/SNPdat\_package\_v1.0.5/SNPdat\_v1.0.5.pl -i Projects/BcSolGWAS/data/SNPdat\_Annotate/MyAnnots/plant12snp.HO.FORPERL.txt -f Projects/BcSolGWAS/data/SNPdat\_Annotate/suziT4.fasta -g Projects/BcSolGWAS/data/SNPdat\_Annotate/genes\_Chromosome.gtf

Optional usage: -o output\_file

<https://code.google.com/archive/p/snpdat/>

**-i**

tab-delimited input file: chromosome\_id position mutation

must remove headers from file \*\*\*

format: “chr1” in gtf, “Chromosome1” in fasta

From: Projects\BcSolGWAS\data\GWAS\_files\05\_annotation\12Plants\_Top1000SNPs\_SegWide\_trueMAF20\_20NA\_forPERL.csv

In:

Projects/BcSolGWAS/data/Annotate/12Plants\_Top1000SNPs\_SegWide\_trueMAF20\_20NA\_forPERL.csv

**-f**

From:

the T4 reference fasta file. FULL GENOME.

C:\Users\nesoltis\Documents\Projects\BcSolGWAS\data\genome\WGS\ suziT4.fasta

In:

Projects/BcSolGWAS/data/Annotate/suziT4.fasta

**-g**

**The gtf file**

From: C:\Users\nesoltis\Documents\Projects\BcSolGWAS\data\genome\WGS\genes.gtf

In:

Projects/BcSolGWAS/data/Annotate/genes.gtf

**Circos to draw circular genome plots**

For botrydial primer development for PCR genotyping:

Email: botrydial/ botcynalide sequences