Explanations for haplotag output files

NWG_CbyT-307-50_SNP.txt (SNP data at 50% missing level)
==
   File: NWG_CbyT-307-50_SNP.txt

   This file contains the called SNP genotypes for all three fragments of the FASTQ sequence files after processing with Haplotag. A = the first 64 bases, B the next 59 bases, and C = the last 59 bases.

   Note: The relationship between the three fragment sets was not preserved during tag generation and each fragment set (A, B, and C) was passed into the analysis pipeline as an independent data set. Therefore SNP locus A.NWG56.1.13 is unlikely to originate from the same genetic sequence locus as B.NWG56.1.13

   SNP locus name structure:
   e.g. A.NWG56.1.13
    A.
     ^ 59 bp or 64 bp fragment set
         NWG56.
         ^^^^^ Tag identifier
             1.
             ^ Haplotag paralog locus identifier
                 13
                 ^^ Position of SNP on the Tag

HTML-A.zip, HTML-B.zip, HTML-C.zip (Haplotype Locus Passport Information for Tag Clusters)
==
   Files: HTIndex-A.html, HTIndex-B.html, HTIndexC.html, which were compressed as .zip using 7-Zip.
   Folders: HTML-A, HTML-B, HTML-C

   Files HTIndex-X.html (X=A,B,C) are the index files that can be used to locate the specific tag used to call the SNPs from each of the three separate sequence fragments. These index files correspond to the .htm files located in the corresponding HTML-X (X=A,B,C) folder. SNP locus name can be used to find fragment set (X=A,B,C), tag identifier, and haplotag paralog locus identifier, and SNP position on the tag.

   Please refer to the Haplotag documentation for specific information on the Locus Passport Information format.

mergedAll-ABC.zip (Tag files)
==
   Files: mergedAll-A.txt, mergedAll-B.txt, mergedAll-C.txt

   These files contain all the tags identified by UNEAK across all taxa for each sequence fragment (A,B,C).

   Each line consists of the Tag sequence, the number of bases in the sequence, and the number of occurrences across the all taxa in that fragment set.

   Refer to the UNEAK (Tassel v 3.0 GBS Pipeline; Tassel v 3.0 UNEAK pipeline) documentation for more information on this file format.