**C**alculate **A**llele **S**haring **C**oefficients (CASC)

Version 1.0

**Introduction**

**C**alculate **A**llele **S**haring **C**oefficients (CASC) provides an intuitive user interface to calculate pair-wise allele sharing coefficients (Guitérrez et al. 2005) from diploid genotype data. Estimates of allele sharing (AS) between pairs of individuals are a type of similarity index, or molecular coancestry, that provide an alternative for estimating relationships among individuals when allele frequencies cannot be meaningfully calculated. Allele sharing for a pair of individuals (ASij) at a given locus *L* is calculated as $AS\_{ij};L=\frac{1}{4}\left[I\_{11}+I\_{12}+I\_{21}+I\_{22}\right]$, where *Ixy* is 1 when allele *x* on locus *L* in individual *i* and allele *y* on the same locus in individual *j* are identical and 0 if they are not. CASC is intended to be used for calculating AS coefficients from SNP data but can be used for microsatellite data if desired. AS coefficients calculated from microsatellite data are expected to have very large sampling variances that will likely limit their usefulness for most analyses.

**System Requirements**

CASC is written in the C# programming language, developed with Microsoft Visual Studio Community 2017, for the Microsoft .NET Framework (version 4) running on Windows operating systems. Users should have the latest service packs for their OS, and must also have installed the Microsoft .NET Framework, version 4. The web installer for the .NET is available at the Microsoft download center (www.microsoft.com/downloads ).

**Installation**

CASC is provided as a ready-to-run executable. The CASC folder and its associated files can be saved to any file location and the program is run by simply double-clicking the CASC.exe file. Output from CASC will be written to the output folder.

**Using CASC**

CASC requires a single input file of diploid genotype data in the genepop file format (both two and three digit genotypes are accepted). All data must be grouped under a single population (multiple “pop” delineations within the file are not allowed). Any number of header rows are allowed; CASC only begins reading data after encountering a “pop” (POP and Pop also accepted) in the file.

After an input file is selected, the CASC interface will display the number of individuals and loci in the file. If these numbers aren’t as expected, the input file has not been formatted correctly. CASC reports AS coefficients for all possible pairs of individuals as the average AS coefficient calculated over a number of user specified iterations, where the iterative AS coefficients are calculated by subsampling a portion of the genotyped loci shared by two individuals. After loading an input file, the user must enter the number of loci to subsample (with replacement) for the AS calculation and the number of iterations over which to calculate average AS values. If two individuals share fewer genotyped loci than the specified number to subsample, CASC will continue to sample available loci with replacement.

**Example Input File**

//header rows

pop

sample01, 0101 0102 0000

sample02, 0102 0101 0202

sample03, 0000 0101 0102

**Output Files**

The following four output files are created in the “output” folder:

ASmatrix: square matrix of average AS coefficients for all possible pairs of individuals

VARmatrix: square matrix of the variance associated with each average AS coefficient for all possible pairs of individuals

Nmatrix: square matrix of the number of shared, genotyped SNPs for all possible pairs of individuals

summary: information captured from the user interface is reported at the top, followed by the AS coefficients, associated variance, and number of shared, genotyped SNPs for all possible pairs of individuals in column format

**Software Citation**

There is currently no corresponding software publication. The program can be cited as: Ivy, J.A and A. Putnam. 2019. Calculate Allele Sharing Coefficients (CASC) (version 1.0). San Diego Zoo Global, San Diego, CA, USA.