# Population Genetics in BioPerl HOWTO

Jason Stajich, Dept Molecular Genetics and Microbiology, Duke University
`<jason-at-bioperl-dot-org>`

2005-03-1

Revision History
| Revision 0.1 | 2004-06-28 | JES |
|---|---|---|
| | First draft | |
| Revision 0.2 | 2004-02-22 | JES |
| | Updated method docs | |
| Revision 0.3 | 2005-03-05 | JES |
| | Expanded to cover coalescent and others | |

## Table of Contents

A Marker is the name given to a polymorphic region of the genome. Mark

```
                );
```

A population is a collecion of indi

a flag called -no_header can be supplied which specifies there is no header line in the report and that the object should assign arbitrary marker names in the form 'Marker1', 'Marker2' ... etc.

Pretty Base format

Phase and hapmap format

# 5. Allele data from Alignments using `Bio::AlignIO` and `Bio::PopGen::Utilities`

Often one doesn't already have data in SNP format but want to determine the polymorphisms from an alignment of sequences from many individuals. To do this we can read in an alignment and process each column of the alignment determine if it is polymorphic in the individuals assayed. Of course this will not work properly if the alignment is bad or with very distantly related species. It also may not properly work for gapped or indel columns so we might need to recode these as Insertion or Deletion depending on the questions one is asking.

The modules to parse alignments are part of the `Bio::AlignIO` system. To parse a clustalw or clustalw-like output one uses the following code to get an alignment which is a `Bio::SimpleAlign` object.


```
use Bio::AlignIO;
my $in G8>new(-data in=> '61 Tm (w95-fi51 => 'fi51.aln')Bio::AlignIO;)Tj Tj  0 1 71 49.a
```

# Bibliography

[1] " Disentangling the effects of demograph