Removing invariant sites

Silvia Justi

9/18/2017

Using phrynomics to remove invariant sites from SNP alignments.
Install and load the following packages

```
#install.packages("devtools")
#devtools::install_github("bbanbury/phrynomics")
library(phrynomics)
```

Sample names must have 9 or less characters
Input must look like this
Read your formated alignment

```
snpdata <-ReadSNP("input_file.txt")
head(snpdata, n=10)
```

```
## $data
##                                                  locus1
## 15_b_T   NNNNNNNNNNNACCTTGGCCCTGAWACGCAGCGNNCTCTTACGANN
## 16_a_T   NNNNNNNNNNNACCTTGGCCCTGANACGCAGCGNNCTCTTACGANN
## 16_b_T   NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
## 59aCR    TGTGACGNNNNAACTTGGCCCTAANACGCAGCGNGCTCTTACTANN
## A10050   NNNNNNCNNNNNNNNNNNNNNNNNNACGCAGCGNNCTCTTACGCAA
## A10051   TAGGACCNNNNNNNNNNNNNNNNNNNGCAGCGNNCTCTTACGCNN
## A10052   NNNNNNCNNNNNNNNNNNNNNNNNNACGCAGCGNNNNNNNNNNAA
## A10053Ab NNNNNNCNNNNNNNNNNNNNNNNNNACGCAGCGNNCTCTTACGCAA
## A10053L  NNNNNNCNNNNNNNNNNNNNNNNNNACGCAGCGNNCTCTTACGCAA
## A10054   TAGGACNNNNNNNNNNNNNNNNNNNACGCAGCGNNNNNNNNNNAA
## A10055   TAGGACNNNNNNNNNNNNNNNNNNNGCAGCGNNCTCTTACGCAA
##
## $ntax
## [1] 11
##
## $nloci
## [1] 1
##
## $nsites
## [1] 46
```

```
summary(snpdata)
```

```
##
## SNP dataset: snpdata
##
## SNP dataset contains 11 taxa, 1 loci, and 46 sites
##
##  Number of taxa: 11
##  Number of loci: 1
##  Number of sites: 46
##
##  Average number of sites per locus: 46
```

```
##  Minimum number of sites per locus: 46
##  Maximum number of sites per locus: 46
```

Removing invariant sites

```
only_variants <-RemoveInvariantSites(snpdata)
summary(only_variants)
```

```
##
## SNP dataset: only_variants
##
## SNP dataset contains 11 taxa, 1 loci, and 7 sites
##
##  Number of taxa: 11
##  Number of loci: 1
##  Number of sites: 7
##
##  Average number of sites per locus: 7
##  Minimum number of sites per locus: 7
##  Maximum number of sites per locus: 7
```

Write your output to phylip format

```
WriteSNP(only_variants, file = "only_var.phy", format = "phylip", missing = "N")
```