msap (v. 1.2.0.9002) - User's Guide

Andres Perez-Figueroa

March 15, 2019

1 Introduction

msap provides a deep analysis of epigenetic variation starting from a binary data matrix indicating the presence or absence of EcoRI-HpaII and EcoRI-MspI fragments, typical of MSAP technique. After comparing the data from both enzyme combinations, the program determines if each fragment is susceptible to methylation (representative of epigenetic variation) or if there is no evidence of methylation (representative of genetic variation). Different analyses of the variation and differentiation (genetic and epigenetic) among user-defined groups of samples are then performed, as well as the classification of the methylation occurrences in those groups. A comprehensive report of the analyses and several useful plots could help researchers to asses the epigenetic variation in their experiments using MSAP. Standard AFLP data is also suitable to be analyzed. All analyses follow a band-based strategy (**?**). There are several examples in the literature of MSAP experiments using some of the analyses provided by *msap* (see **?????**) so the package could be useful in those kind of approaches.

The package is intended to be easy to use even for those people non-familiar to the R environment. Advanced users could take advantage of available source code to adapt *msap* for more complex analyses.

2 R basics. All you need to know about R to run *msap*

The only knowledge required for installing and running *msap* is about how to open an R session in your computer. R is a statistical programming language that provides many built-in functions for performing statistical analysis and is also flexible to allow users to write their own functions.

R can be downloaded and installed for free from the website http://cran.r-project.org where detailed instructions for installing R on any operating system are provided. Accessing R is different for every operating system. For windows users, simply double click the R icon that is created after installation. For Mac users, you can double click the R icon under your Applications menu. On Linux,

in the terminal window simply type R at the command prompt and R will be opened within the terminal window.

When you open R, no matter the operating system you are using, you will see the command prompt symbol `>` which simply means that R is waiting for you to give it a command. To quit R, simply type `q()` in the command prompt and R will ask you if you want to save the workspace before quitting. And that's all.

3 Installing *msap*

You can install *msap* automatically from a R session. To install the last stable version from CRAN:

```
> install.packages("msap")
```

To get the last daily development version from R-Forge:

```
>  install.packages("msap", repos= c("http://R-Forge.R-project.org", getOption("repos")))
```

The above instructions should install *msap* and all required dependecies.

4 Preparation of data

In order to use *msap* for analyzing your results from a MSAP experiment, you need to provide a data file with a binary matrix (1/0) indicating the presence or absence of EcoRI-HpaII and EcoRI-MspI fragments in a bunch of samples of two or more populations/groups. Data file should be a .csv file with markers as columns and two rows by sample, one for each isoschizomer reaction. File could be edited in the a spreadsheet of your choice (see Figure **??**) and then saved as csv (with ',' as field separator). The final text file should look like Figure **??** if opened in a text editor.

The first row should include the markers name/references, these should be ordered by primer combination (if applicable) as analysis will be seprated for them. The first column should provide the label for the group where the sample is included, with the aim to make comparisons between different groups. Second column is reserved for an arbitrary label (i.e. to name the sample). Third column should identify the isoschizomer with 'HPA' or 'MSP'. From version 1.1.0 the fragment classification is done within the different primer combinations using their own error rate. Marker list (columns fourth to end) should be ordered by primer combinations (take note of the number of markers per primer combination, as they will be required as argument later).

If you want to analyze a standard AFLP dataset (or any other dominant markers coded by a 1/0 matrix) the datafile format is the same, but the program will ignore content of third column and treat all rows as independent samples.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T
1				m1	m2	m3	m4	m5	m6	m7	m8	m9	m10	m11	m12	m13	m14	m15	m16	m17
2	Pop1	a1	HPA	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
3	Pop1	a1	MSP	0	0	0	1	1	0	0	1	1	1	0	0	0	0	0	0	1
4	Pop1	a2	HPA	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
5	Pop1	a2	MSP	0	0	1	0	0	1	1	1	1	1	1	0	0	1	0	1	0
6	Pop1	a3	HPA	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
7	Pop1	a4	MSP	1	1	0	1	0	0	0	1	1	1	0	0	0	0	0	0	1
8	Pop1	a5	HPA	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
9	Pop1	a5	MSP	0	0	1	0	0	0	1	0	1	1	1	0	0	1	0	1	0
10	Pop1	a6	HPA	1	1	0	1	0	0	0	1	1	1	0	0	0	0	0	0	1
11	Pop1	a6	MSP	0	0	0	0	0	0	0	1	1	1	1	0	0	1	0	0	0
12	Pop1	a7	HPA	0	1	0	1	0	0	0	1	1	1	1	0	0	1	0	0	0
13	Pop1	a7	MSP	0	0	1	0	0	0	0	1	1	0	0	0	0	0	0	0	0
14	Pop2	b1	HPA	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
15	Pop2	b1	MSP	0	0	0	0	1	0	0	0	1	1	0	1	0	0	0	0	0
16	Pop2	b2	HPA	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
17	Pop2	b2	MSP	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
18	Pop2	b3	HPA	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0
19	Pop2	b3	MSP	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0
20	Pop2	b4	HPA	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
21	Pop2	b4	MSP	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
22	Pop2	b5	HPA	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0
23	Pop2	b5	MSP	0	0	0	0	0	0	0	0	1	1	0	0	0	1	0	0	0
24	Pop2	b6	HPA	0	1	0	0	0	0	0	1	1	1	1	0	0	0	0	0	0
25	Pop2	b6	MSP	0	1	0	0	1	1	0	1	1	0	1	0	0	1	0	0	0
26	Pop3	c1	HPA	0	1	0	1	1	0	0	1	1	1	0	0	0	1	0	0	1
27	Pop3	c1	MSP	0	0	1	1	0	0	0	1	1	1	1	0	0	0	0	0	0
28	Pop3	c2	HPA	0	1	0	1	1	0	0	1	1	1	0	0	0	0	0	1	0
29	Pop3	c2	MSP	0	0	0	1	1	0	0	1	1	1	1	0	0	1	0	0	1
30	Pop3	c3	HPA	1	1	0	1	1	0	0	1	1	1	1	0	0	0	1	0	1
31	Pop3	c3	MSP	1	1	0	1	0	0	1	1	1	1	0	0	0	0	1	0	1

Figure 1: Data format as seen in a spreadsheet for edition

```
,,,m1,m2,m3,m4,m5,m6,m7,m8,m9,m10,m11,m12,m13,m14,m15,m16,m17,m18,m19,m20,m
Pop1,a1,HPA,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,1,1,0,0,0,0,0,0,0,0
Pop1,a1,MSP,0,0,0,1,1,0,0,1,1,1,0,0,0,0,0,0,1,0,0,0,0,1,0,0,0,0,0,1,1,0,0,0
Pop1,a2,HPA,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,1,1,0,0,0,0,0,0,0,0,0,0
Pop1,a2,MSP,0,0,1,0,0,1,1,1,1,1,1,0,0,1,0,1,0,1,0,0,0,1,0,0,0,0,0,1,0,0,0,0
Pop1,a3,HPA,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,1,1,0,0,0,0,0,0,0,0,0,0
Pop1,a4,MSP,1,1,0,1,0,0,0,1,1,1,0,0,0,0,0,0,1,1,0,0,0,1,0,0,0,0,0,0,1,0,0,0
Pop1,a5,HPA,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,1,1,0,0,0,0,1,0,0,0,0,0
Pop1,a5,MSP,0,0,1,0,0,0,1,0,1,1,1,0,0,1,0,1,0,1,0,0,0,1,0,0,0,0,0,1,0,0,0,0
Pop1,a6,HPA,1,1,0,1,0,0,0,1,1,1,0,0,0,0,0,0,1,0,1,0,0,0,0,0,0,0,0,0,1,0,0,0
Pop1,a6,MSP,0,0,0,0,0,0,0,0,1,1,1,1,0,0,1,0,0,0,1,0,0,0,1,0,0,0,0,0,0,0,0,0
Pop1,a7,HPA,0,1,0,1,0,0,0,1,1,1,1,1,0,0,1,0,0,0,1,0,0,0,1,0,0,0,0,0,1,1,0,0,0
Pop1,a7,MSP,0,0,1,0,0,0,0,0,1,1,0,0,0,0,0,0,0,0,0,0,0,0,1,0,1,0,0,0,0,0,0,0,0
Pop2,b1,HPA,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,1,1,0,0,0,0,0,0,0,0,0,0
Pop2,b1,MSP,0,0,0,0,1,0,0,0,1,1,0,1,0,0,0,0,0,0,0,0,1,1,0,0,0,0,0,0,0,0,1
Pop2,b2,HPA,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,1,1,0,0,0,0,0,0,0,0,0,0
Pop2,b2,MSP,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,1,1,0,0,0,0,1,0,0,0,0,0
Pop2,b3,HPA,0,0,0,0,0,0,0,0,1,1,0,0,0,0,0,0,0,0,0,0,1,1,0,0,0,0,0,0,0,0,0,0
Pop2,b3,MSP,0,0,0,0,0,0,0,0,1,1,0,0,0,0,0,0,0,0,0,0,1,1,0,0,0,0,0,0,0,0,0,0
Pop2,b4,HPA,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,1,1,0,0,0,0,1,0,0,0,0,0
Pop2,b4,MSP,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0
Pop2,b5,HPA,0,0,0,0,0,0,0,0,1,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0
Pop2,b5,MSP,0,0,0,0,0,0,0,0,1,1,0,0,0,1,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0
Pop2,b6,HPA,0,1,0,0,0,0,0,1,1,1,1,0,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0
Pop2,b6,MSP,0,1,0,0,1,1,0,1,1,0,1,0,0,1,0,0,0,0,0,1,0,1,1,1,1,1,0,0,0,1,1,0
```

Figure 2: Final data format in the .csv file

5 Running *msap*

We start by loading the msap package into an R session[1].

```
> library(msap)
```

It is highly recommended to change the working directory to that where datafile is located. Windows users can use the menu item 'File>Change dir' and choose the appropiate folder. To change the working directory within an R console run the command `setwd(dir)` where *dir* is the absolute path to the directory. The output files created by *msap* will be saved in that working directory.

Once we are in the right working directory with an appropiate data file, we can run all analyses of *msap* with a single command (change "example.csv" by the name of your datafile, keeping the quotes, and change "Example" by a custom name, keeping quotes, to identify your data):

```
> msap("example.csv", name = "Example")
```

Those users familiar with R would prefer to store the returning list (with several useful data for further analysis) of msap:

```
> myList <- msap("example.csv", name = "Example")
```

On execution, *msap* will run all analyses and will show an on screen text report with the results:

```
msap  1.2.0.9002  - Statistical analysis for Methylation-Sensitive Amplification Polimorphis

Reading  example.csv ..... Ok!
Number of loci:  701
Number of samples/individuals:  38
Number of groups/populations:  4
Number of primer combinations:  1
Loci per primer combinations  701
Error rates per primer combination:  0.05
Primer:  1
--Number of Methylation-Susceptible Loci (MSL):  659
--Number of No Methylated Loci (NML):  24

All combinations:
Number of Methylation-Susceptible Loci (MSL):  659
Number of No Methylated Loci (NML):  24

Number of polymorphic MSL:  367  ( 56 % of total MSL)
Number of polymorphic NML:  12  ( 50 % of total NML)
```

[1] Windows users have a step-by-step guide in section **??**

```
- Saving transformed matrix for MSL in file:  Example-MSL-transformed.csv
- Saving transformed matrix for NML in file:  Example-NML-transformed.csv

Shannon's Diversity Index
MSL: I =  0.532922   (SD:  0.1437365 )
NML: I =  0.1873695   (SD:  0.1597695 )
Wilcoxon rank sum test with continuity correction :  W  =  4073  ( P < 0.0001 )


*****************************
Analysis of MSL
Report of methylation levels
                                     pop1   pop2   pop3
HPA+/MSP+ (Unmethylated)           0.1478 0.1275 0.1247
HPA+/MSP- (Hemimethylated)         0.1369 0.1127 0.1291
HPA-/MSP+ (Internal C methylation) 0.1675 0.1085 0.1763
HPA-/MSP- (Uninformative)          0.5478 0.6514 0.5698
                                     pop4
HPA+/MSP+ (Unmethylated)           0.1278
HPA+/MSP- (Hemimethylated)         0.1050
HPA-/MSP+ (Internal C methylation) 0.1223
HPA-/MSP- (Uninformative)          0.6449


- Saving clustering tree figure for MSL in file:  Example-MSL-NJ.png .....- Saving PCoA figu

Performing AMOVA
AMOVA TABLE         d.f.          SSD                 MSD                 Variance
among groups         3          566.8          188.9          13.48
within groups         34          2081          61.21          61.21
Total                 37          2648          71.57

Phi_ST =  0.1805   (P<0.0001)


*****************************
Analysis of NML

- Saving clustering tree figure for NML in file:  Example-NML-NJ.png .....- Saving PCoA figu

Performing AMOVA
AMOVA TABLE         d.f.          SSD                 MSD                 Variance
among groups         3          8          2.667          0.2365
within groups         34          14.5          0.4265          0.4265
Total                 37          22.5          0.6081
```

```
Phi_ST =  0.3567   (P<0.0001)
Done!
```

From version 1.1.2, msap returns a list with data useful for further analysis. These are the data slots stored in the list:

- groups A factor with the name of the group of every individual analysed
- patterns A list showing the MSAP patterns (11, 10, 01 and 00 coded as u, h, i, f) in all groups
- transformed.MSL A data frame including the binary (1: unmmethylated, 2: methylated) values for those loci classified as MSL
- transformed.NML A data frame including the binary (1: unmmethylated, 2: methylated) values for those loci classified as NML
- DM.MSL A distance matrix object between all individuals for those polymorphic loci classified as MSL
- DM.NML A distance matrix object between all individuals for those polymorphic loci classified as NML
- DM.AFLP A distance matrix object between all individuals when analysing AFLP data

In addition, *msap* also produces some exploratory figures that are directly saved into .png files:

- A plot of Principal Coordinate Analysis (PCoA), see Figure **??**, showing the first two axis. They are saved as '<*name*>-MSL.png' and '<*name*>-NML.png' for MSL and NML respectively, where <*name*> represents the name passed as argument in the calling to msap function.
- A neighbor-joining tree of all samples (see Figure **??**) saved as '<*name*>-MSL-NJ.png' and '<*name*>-NML-NJ.png' for MSL and NML respectively

Furthermore, some files (.csv format) with useful data are produced, in the working directory, to be processed with external programs:

- A couple of .csv files with the transformed data matrices for MSL and NML. They are saved as '<*name*>-MSL-transformed.csv' and '<*name*>-NML-transformed.csv' for MSL and NML respectively, where <*name*> represents the name passed as argument in the calling to msap function.
- A table with the PCoA coordinates in all axis for each sample. It is saved as '<*name*>-MSL-PCoA.coor.csv' and '<*name*>-NML-PCoA.coor.csv' for MSL and NML respectively.
- A table with the eigenvalues for all axis obtained by PCoA saved as '<*name*>-MSL-PCoA.eige.csv' and '<*name*>-NML-PCoA.eige.csv' for MSL and NML respectively

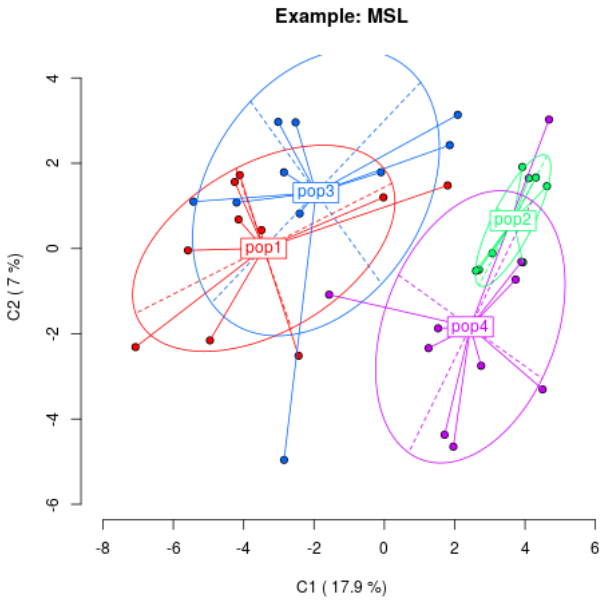

Figure 3: Representation of Principal Coordinate Analysis (PCoA) for epigenetic (MSL) differentiation between groups. The first two coordinates (C1 and C2) are shown with the percentage of variance explained by them. Different point types represent individuals from different groups. Group labels show the centroid for the points cloud in each group. Ellipses represent the average dispersion of those poins around their centre. The long axis of the ellipse shows the direction of maximum dispersion and the short axis, the direction of minimum dispersion.

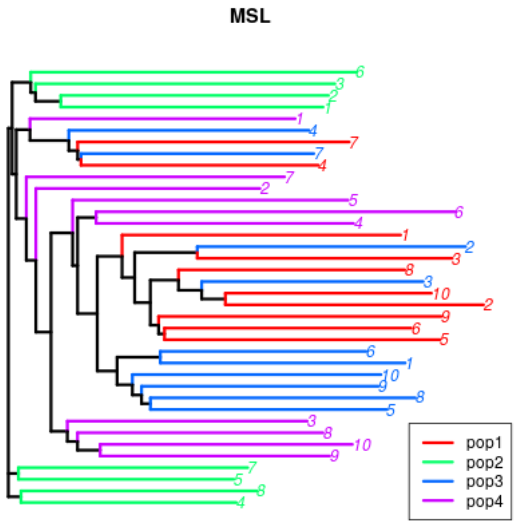

Figure 4: Neighbor-Joining tree of all samples (numbered labels at the tips) for epigenetic (MSL) distances. Colors represent different groups/populations.

5.1 Further options

In the previous section, the basic use of *msap* was described. However, it is possible to set some different options in the program if passed as arguments to the msap() function.

Here is the full usage of msap() function including all the arguments and their default values (if applicable). Except for the 'datafile' that is required, all other arguments are optional.

```
msap(datafile, name=datafile, pattern=c(1,2,2,NA),
                  nDec=4, meth=TRUE, rm.redundant=TRUE,
                  rm.monomorphic=TRUE, do.pcoa=TRUE, do.shannon=TRUE, do.amova=TRUE,
                  do.pairwisePhiST=FALSE, do.cluster=TRUE, use.groups=NULL, do.mantel=FALSE,
                  np.mantel=1000, loci.per.primer=NULL, error.rate.primer=NULL,
                  enz1="HPA", enz2="MSP",
                  threshold.poly.MSL=1, threshold.poly.NML=1,
                  no.bands=NULL, uninformative=NULL)
```

datafile String containing the url of the csv file with the data. Required.

name a name for the dataset to be included in the output files. By default, the name of the given datafile is used.

pattern vector of methylation states (1 - unmethylated, 2- methylated, NA - missing data) for the four combinations of both enzymes enz1+/enz2+, enz1+/enz2-, enz1-/enz2+, enz1-/enz2-. By default c(1,2,2,NA) corresponding with standard MSAP using HPA/MSP considering -/- as uninformative status. Note that the value of this argument could be modified if using the deprecated arguments no.bands or uninformative.

nDec number of digits of precision for floating point output.

meth Logical value switching between MSAP ('TRUE') and standard AFLP ('FALSE') analysis. The difference lies in that for AFLP (or any other dominant marker coded by a 1/0 matrix) the 'enzyme' column is ignored and every row in data represent an independent sample, without combination of data. This is useful to compare MSAP (epigentic) and AFLP (genetic) differentiation using the same analyses.

rm.redundant Not implemented yet.

rm.monomorphic Logical value switching between the removal ('TRUE', by default) of monomorphic fragments (defined as those with only one state or just one ocurrence of the second state across the whole dataset) after data transformation.

do.pcoa Option switcher for doing a Principal Coordinate Analysis for variation between groups. TRUE by default.

do.shannon Option switcher for Shannon's Diversity Index comparison between MSL and NML.

do.amova Option switcher for doing an AMOVA for differentiation between groups. TRUE by default.

do.pairwisePhiST Logical value switching between the calculation of the pairwise Φ_{st} between pairs of groups/populations ('TRUE' by default) or skip it ('FALSE').

do.cluster Calculates and plots a Neighbour-Joining tree ('TRUE' by default) or skip it ('FALSE').

do.mantel Performs a Mantel test to obtain correlation between MSL and NML ('TRUE') or skip it ('FALSE' by default).

np.mantel Gives the number of permutations for the above Mantel test (1000 by default) or skip it ('FALSE').

use.groups Gives the groups/populations/treatments of the datafile to be analysed. By default all groups are considered for the analysis. To provide a subset of the groups a vector should be passed with the names of groups to be included. For example, in a datafile with 5 groups (Control, pop1, pop2, pop3 and pop4) we are interested only in Control and pops 1 and 3. Then, msap should be called with 'use.groups=c('Control','pop1','pop3')'.

loci.per.primer Vector providing the number of loci/fragements obtained per primer combination. Fragment classsification is performed independenty for each primer combination. These fragment should be ordered in the datafilein the sameway as specified here. If this is nor provided (by default) then all fragments should be analyzed as they come from a single combination. For example, if there are three primer combinations with 135, 234 and 210 loci each, then msap should be called with 'loci.per.primer=c(135,234,210)'

error.rate.primer Gives the repeatibility value of MSAP assays for each primer combination. It provides a threshold to consider methylation events as genotyping errors.

enz1 String for the label used in the datafile for the first enzyme. By default, as it considers standard MSAP with HPA/MSP, the value is "HPA". Note that this is case-sensitive and it should match the label given in the datafile.

enz2 String for the label used in the datafile for the second enzyme. By default, as it considers standard MSAP with HPA/MSP, the value is "MSP". Note that this is case-sensitive and it should match the label given in the datafile.

threshold.poly.MSL Minumum number of occurrences of each state (methylated/unmethylated) in a MSL to be defined as polymorphic. The default is 1, that gives the máximum number of polymorphic sites, as only on occurrence of any of the states is needed.

threshold.poly.NML Minumum number of occurrences of each state (band/no band) in a NML to be defined as polymorphic. The default is 1, that gives the máximum number of polymorphic sites, as only on occurrence of any of the states is needed.

no.bands This is a deprecated argument kept for compatilibility porpouses with previous versions. Use pattern.

uninformative This is a deprecated argument kept for compatilibility porpouses with previous versions. Use pattern.

6 Frequently Asked Questions

These are some questions that users have made about the use of *msap*. Here I try to answer them with step-by-step procedures.

6.1 I'm very new into R and I run it under Windows, what steps must I follow to analyse my MSAP data?

Please follow these instructions carefully:

1. Open R by double clicking its icon in your program list or shortcut in desktop.
2. Check the working directory by typing in the R console: *getwd()* This would return the current working directory (i.e: *[1] "C:/Documents and Settings/andres/Mis documentos"*), if this is not the desired (i.e. that with the input file) then you have to change it (or, alternatively, move your datafile to that folder)
3. Change the working directory. You have two alternatives to do this:
   - By typing: *setwd("C:/Documents and Settings/andres/Mis documentos/TheFolderWithMyData/")* changing the route to that folder where your datafile lies.
   - By selecting the folder from the R menu item File > Change dir. The selected folder (that containing your datafile) will be the working directory now.
4. Optionally check that change was right by typing *getwd()* again and check if your datafile is in the working directory by typing *dir()* that will list all files in the working directory.

5. Now R is ready. Load the package (that you have installed before) by typing: *library(msap)* and run the analyses by typing: *msap("yourdatafile.csv", name="YourAnalysis")* or adding further option parameters (within the parenthesis and separated by commas) as described in the Further Options section. With some experience you would skip steps 2 and 4.

6. A text report will appear in the R window (that you can copy-paste) and some files will be created (starting by "YourAnalysis-" and described in the user guide) in the working directory/folder.

6.2 Can I analyze AFLP data? How?

Yes, it is possible to analyse standard AFLP data by skipping all data transformation and classification of MSL/NML and going directly to diversity/differentiation analysis. To do this, there is just a couple of steps:

1. The file format is the same that for MSAP data, but for AFLP the third column (enzyme label) will be ignored AND each row is an individual sample (recall that for MSAP there were 2 rows per sample). It sould be something like that:

```
,,,m1,m2,m3,m4.....
Pop1,i1,ANYTHING,1,1,0,1.....
Pop1,i2,ANYTHING,0,0,1,1.....
Pop1,i3,ANYTHING,1,0,0,0.....
Pop2,i4,ANYTHING,0,0,1,1.....
Pop2,i6,ANYTHING,1,0,0,0.....
Pop2,i6,ANYTHING,1,1,1,1.....
```

2. Run the program with the option meth=FALSE (so it skips all methylation-related stuff):

```
msap("MyAFLPdatafile.csv", name="AFLPanalysis", meth=FALSE)
```

This is useful for those experiments that combine MSAP profiles and AFLP profiles to compare epigenetic vs. genetic variation, as both can be analysed with the same program.

6.3 What if I need another kind of analysis for my MSAP data?

The analyses currently provided by *msap* are limited but this does not mean that further analysis could not be added in the future. I try to keep the package updated to allow exploratory assays of epigenetic diversity and differentiation, focused on the field of evolutionary ecology. If you are trying to analyse your MSAP data and options in *msap* do not fit your requirements, then you have three alternatives to get your tasks done by *msap*:

- If you have programming skills or experience with R, then you can get the source code of *msap* and adapt it to your needs. That is the main advantage of open source software!
- If you have programming skills or experience with R, and want to collaborate with me to expand *msap*, then do not hesitate to contact me for joining the development team.
- If you are not familiar with R or do not feel confident to make code yourself, then use the support tracker (in msap website) to request new features in *msap*. I'll try to implement them as soon as possible.

References

Bonin, A., Ehrich, D., and Manel, S. (2007). Statistical analysis of amplified fragment length polymorphism data: a toolbox for molecular ecologists and evolutionists. *Molecular ecology*, 16(18), 3737-58.

Chwedorzewska, K. J., and Bednarek, P. T. (2012). Genetic and epigenetic variation in a cosmopolitan grass Poa annua from Antarctic and Polish populations. *Polish Polar Research*, 33(1), 63-80.

Gupta, V., Bijo, J., Kumar, M., Reddy, C. R. K., and Jha, B. (2012). Detection of Epigenetic Variations in the Protoplast-Derived Germlings of Ulva reticulata Using Methylation Sensitive Amplification Polymorphism (MSAP). *Marine biotechnology* DOI: 10.1007/s10126-012-9434-7

Herrera, C. M., and Bazaga, P. (2010). Epigenetic differentiation and relationship to adaptive genetic divergence in discrete populations of the violet Viola cazorlensis. *The New phytologist*, 187(3), 867-76.

Moran, P., and Perez-Figueroa, A. (2011). Methylation changes associated with early maturation stages in the Atlantic salmon. *BMC genetics*, 12(1), 86.

Rodriguez, C. M., Moran, P., Lago, F., Beckmann, M., and Consuegra, S. (2012). Detection and quantification of tissue of origin in salmon and veal products using methylation sensitive AFLPs, *Food Chemistry* 131, 1493-1498.

7 Session Info

This document was created using the following:

```
> sessionInfo()

R version 3.5.2 (2018-12-20)
Platform: x86_64-pc-linux-gnu (64-bit)
Running under: Ubuntu 18.04.2 LTS
```

```
Matrix products: default
BLAS: /usr/lib/x86_64-linux-gnu/blas/libblas.so.3.7.1
LAPACK: /usr/lib/x86_64-linux-gnu/lapack/liblapack.so.3.7.1

locale:
 [1] LC_CTYPE=en_US.UTF-8       LC_NUMERIC=C
 [3] LC_TIME=es_ES.UTF-8        LC_COLLATE=en_US.UTF-8
 [5] LC_MONETARY=es_ES.UTF-8    LC_MESSAGES=en_US.UTF-8
 [7] LC_PAPER=es_ES.UTF-8       LC_NAME=C
 [9] LC_ADDRESS=C               LC_TELEPHONE=C
[11] LC_MEASUREMENT=es_ES.UTF-8 LC_IDENTIFICATION=C

attached base packages:
[1] stats     graphics  grDevices utils     datasets
[6] methods   base

other attached packages:
[1] msap_1.2.0.9002

loaded via a namespace (and not attached):
 [1] MASS_7.3-51.1   compiler_3.5.2  parallel_3.5.2
 [4] tools_3.5.2     Rcpp_1.0.0      nlme_3.1-137
 [7] ape_5.2         grid_3.5.2      ade4_1.7-13
[10] lattice_0.20-38
```