MSAP ANALYZER

An interactive interface for Methylation Sensitive Amplified Polymorphism (MSAP) data analysis

Help Guide

Index

1. Areas of application

MSAP analyzer was originally developed to assist with processing of high volume raw data produced following analysis of tissue-cultured plant samples with Methylation Sensitive Amplified Polymorphism (MSAP) markers. However, this technique has a wider spectrum of applications. For example it can be employed in medicine, as a diagnostic tool of cancer (Ebert, et al., 2005; Jorda and Peinado, 2010; Laird, 2003), in ecology as a means for probing putative transmission of epialleles contributing to phenotypic variability and trait inheritance (Salmon and Ainouche, 2010), in assessing whether adaptation to a specific habitat has a genetic or epigenetic origin, in molecular genetics for the construction of genetic linkage maps (Duan, et al., 2009), for genetic improvement of several species of economic interest (Baranek, et al., 2010; Bednarek, et al., 2007; Duan, et al., 2009; Menz, et al., 2002) as well as in the molecular characterization of different plant cultivars (Cervera, et al., 2003; Fang, et al., 2010; Hao, et al., 2004; Martienssen and Colot, 2001; Peredo, et al., 2009; Smykal, et al., 2007).

2. Experimental background and example dataset

For software development and for demonstration purposes, we exploited experimental MSAP profiles produced from a *Citrus* genotype (KxVolk) clonally propagated through alternate tissue-culture media each containing different plant growth regulators (PGR). The experimental aim was to rank epimutation potential of different PGR. The original maternal individual -the reference sample denoted as “KxVolk-reference”- was clonally propagated through six different tissue-culture media (coded as: KxVolk-A1, KxVolk-PY2, KxVolk-PY3, KxVolk-BE1, KxVolk-A2, KxVolk-

A3). The same coding was employed for the nomenclature of MSAP marker fingerprints produced from exposure to respective media.

For MSAP, the most frequently used pair of isoschizomers is HpaII and MspI recognising the same restriction site (5'-CCGG) but differing in their sensitivity to the methylation state of specific cytosine residues in the recognition sequence. HpaII will not digest if either of the cytosines is fully-(both strands) methylated whereas MspI will not digest if the external cytosine is fully or hemi-(single strand) methylated (McClelland, et al., 1994). Each sample was digested with two enzyme combinations (a. EcoRI + MspI and b. EcoRI + HpaII) yielding two different digests (restriction populations). In accordance with the counterpart AFLP technique each of the two digests was selectively amplified with four pair combination in total. Primers at the 5' end include one of the four EcoRI-based selective nucleotides (E-ATG, E-ACC, E-AAG or E-ACT) while on the 3' end the common HpaII / MspI-based selective nucleotides are H/M-TCAA. Since each initial sample yields two different restriction populations, each amplified with all four pair combination, a total of 8 different fingerprints were eventually produced per sample.

By comparing the two fingerprints produced from the two restriction populations of each individual treatment sample the methylation status is established for each marker locus separately. This is then compared to the methylation status of the corresponding marker locus in the reference sample. Collated quadruplet event types and deduced state changes employed by MSAP analyser, are shown in Table 1. The series of different event types (methylation status changes) are subsequently shown.

Table 1. Methylation status changes detected as quadruplet event types for isoschizomers

(A) HpaII/MspI

no	**marker presence "1" or absence "0"**				**deduced event**
	sample				
	reference		treated		
	H	M	H	M	
1.	1	1	1	1	no change
2.	0	0	1	1	
3.	1	1	0	0	
4.	1	0	1	0	*de novo* methylation
5.	1	0	1	1	
6.	1	1	1	0	
7.	0	1	0	1	demethylation
8.	0	1	1	1	
9.	1	1	0	1	
10.	0	0	1	0	other (non-classifiable)
11.	1	0	0	0	
12.	0	1	0	0	
13.	0	0	0	1	

(**B**) Acc65I/KpnI.

no	marker presence "1" or absence "0"				deduced event
	sample				
	reference		treated		
	A	K	A	K	
1.	1	1	1	1	no change
2.	0	0	1	1	
3.	1	1	0	0	
4.	1	0	1	1	*de novo* methylation
5.	0	1	1	1	demethylation
6.	0	0	0	1	sequence variation
7.	0	0	1	0	
8.	0	1	0	1	
9.	1	0	1	0	
10.	1	1	1	0	
11.	1	1	0	1	
12.	0	1	0	0	Complex events
13.	1	0	0	1	
12.	0	1	1	0	
13.	1	0	0	0	

Event types (change from reference to treated sample detected for each marker locus separately) were grouped into four different broad categories for isoschizomers HpaII/MspI : 1) no change in methylation, 2) *de novo* methylation, 3) demethylation, 4) other (non-classifiable) and into five different broad categories for isoschizomers HpaII/MspI : 1) no change in methylation, 2) *de novo* methylation, 3) demethylation, 4) Sequence variation and 5) Complex events (for detailed patterns see Table 1).

The example data set included results from 5 treatments (presence of 5 different PGRs in tissue-culture media) (KxVolk-A1, KxVolk-PY2, KxVolk-PY3, KxVolk-A2, KxVolk-A3) and from 2 different primer combinations (E-AAG, E-ACT/ HpaII- or MspI- TCAA). Note that the reference sample (in the example denoted as KxVolk-REF in Fig. 4b and referring to the maternal individual) should be labelled explicitly in one of the rows of the input file (see section 4 for details) Fig. 4b is a table of the quadruplet event types

of a single primer-enzyme combination (PEC) following comparisons between the maternal tissue KxVolk-REF (reference state) and the same tissue following exposure to tissue culture conditions/treatment "PY3" (KxVolk-PY3); as described above. The same comparative analysis is produced for all treatments (different PGRs) using the following notations as column headings (KxVolk-REF – MspI, KxVolk-X – MspI KxVolk-REF – HpaII, KxVolk-X – HpaII,) where –"X" corresponds to the different culture medium/treatment compared consecutively to the maternal tissue.

3. Getting started using MSAP analyzer

MSAP analyzer is an online tool accessible through the web page interface http://mirna.imbb.forth.gr/MSAPAnalyzer.html. MSAP allows multiple file uploads for launching the computational modules (Fig 1).

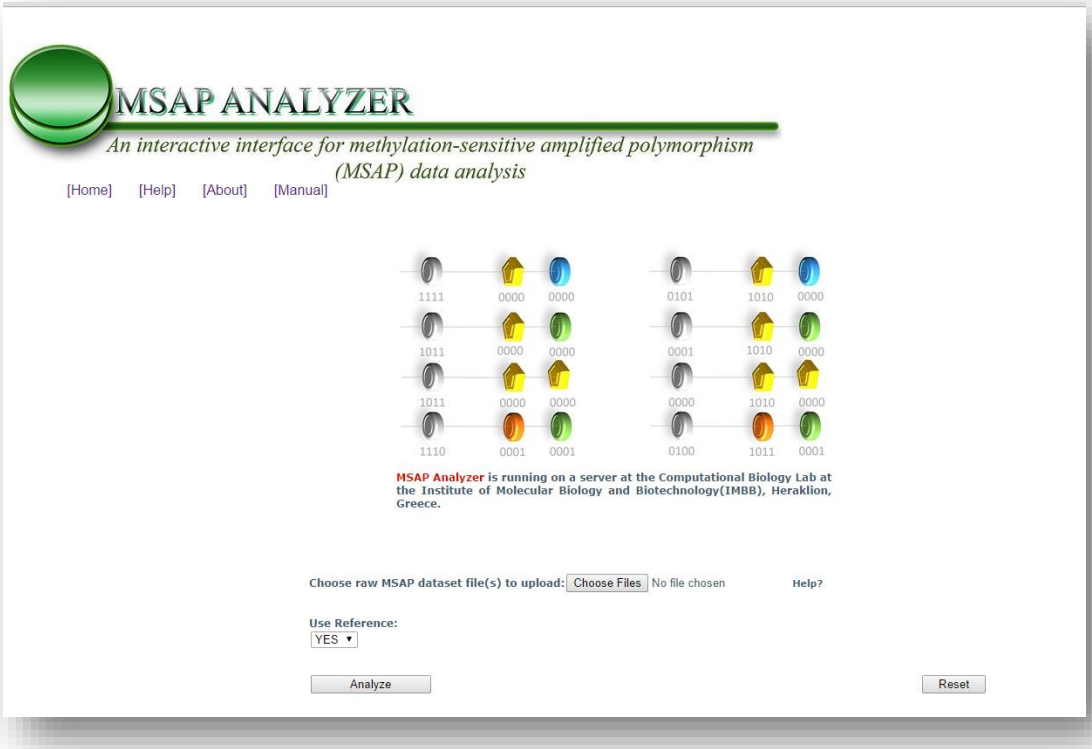


Fig. 1. MSAP analyzer interface

4. MSAP input file format and preparation

MSAP Analyzer accepts input matrix files only in tab delimited format and with the file suffix – ".txt" without size limitation. File names can be any valid Windows names. Notably, it is obligatory that, within the matrix file, field names contain no spaces and are separated by tabs. It is advised that files are prepared in any spreadsheet software and then saved as ".txt" before uploaded to MSAP analyzer. The input matrix file corresponds to a multiple genotype/primer combination sets and contains results from the entire series of treatments.

An example of an input matrix file is presented in Figs. 2 a & b in ".txt" and ".xls" formats respectively. It should be emphasized that MSAP Analyzer recognizes only tab-delimited text files and that the ".xls" file is shown here for demonstration purposes only. Moreover the user can first import and format the input data under Excel or other spreadsheet software and then save-as or copy-paste the text to a ".txt" file for input in MSAP Analyzer. This will automatically construct the tab-delimited data format required for input to MSAP Analyzer.

A more detailed explanation of Fig. 2b is presented below. Column A includes Genotype names that must contain certain user-defined information (any kind of alphabetical characters or numbers). The genotype name; is used by MSAP Analyzer during the analysis pipeline in order to describe the uploaded datasets and helps the user to distinguish datasets from different genotypes. Column B should contain a user-defined description of the treatments and Column C should provide isoschizomers enzyme used for that specific treatment and each row must alternate between first enzyme used and second enzyme used (i.e. "HpaII" on row 2 and "MspI" on row 3 - see example below). Column D provides of the name of the primer combination used (i.e. E-ATG/M-TCAA). If more than one primer combinations has been employed then

this can be indicated by providing a different primer combination name in column D (for example below treatment PY2 has been analyzed using different primer combinations and shown in rows 8-9 in Fig2B) these correspond to a unique primer combination. Column E provides the error determinant but this can be left blank if the user is not aware of any. Column F (repeats) provides the biological replicates for statistical purposes.

In case the user decides to perform a reference based analysis (seed drop down menu in main interface of MSAPAnlayzer) the specific genotype(s) should be labelled as “reference” in the treatment column (in the example case rows 10-13 show this reference samples see Fig 2B). The following columns (G-onwards) contain markers/alleles detected and listed by the user, according to their respective size in base pairs. This is the output from the process of marker binning and sizing performed in the fragment analyzer of choice. For every treatment data rows exist that contains a listing of markers produced following digestion with the first enzyme combination (in the example, in Fig 2. EcoRI + MspI is used) while the second row contains markers produced following digestion with the second enzyme combination (EcoRI + HpaII). As row values, the fields “MspI” and “HpaII” are used referring to the origin of difference between the two fingerprints from the MspI and HpaII digestion enzymes. The user must make sure that for every genotype the first row corresponds to the “MspI” and the second row corresponds to the “HpaII” isoschizomer numerical data respectively. Rows “2” and “3” contain marker data of the first treatment sample “KxVolk-PY3” while rows “6” and “7” contain data of the second treatment sample “KxVolk-PY2”. As many duplicate rows (representing different treatment samples) as necessary can be included in the same spreadsheet file to serve as input to MSAPAnalyser.

The file format may also contain data from the isoschizomer pair Acc65I/KpnI. The same formatting rules as for the HpaII/MspI pair apply as they explained above. However, the user will be asked to make a choice of the isoschizomer pair to use. It should be noted that alternate enzyme combinations are specific to each

experimental design and data sets. The presented trial data sets refer to the "HpaII/MspI" but not to the "Acc65I/KpnI" enzyme combinations (section 6).

A

Citrus_MSAP_3_Input.txt *

Genotype	Treatment	Ésoschizomer	Enzyme combination	error determinant	repeat	Allele 1	Allele 2
K x Volk	PY3 HpaII	EACT_H,MAAT	1				
K x Volk	PY3 MspI	EACT_H,MAAT	1	65		74	79
K x Volk	PY3 HpaII	EACT_H,MAAT	2 54	65	68	70 74	79
K x Volk	PY3 MspI	EACT_H,MAAT	2 54	65		70 74	79
K x Volk	PY2 HpaII	EACT_H,MAAT	1 54	65	68	70 74	79
K x Volk	PY2 MspI	EACT_H,MAAT	1 54	65		70 74	
K x Volk	PY2 MspI	2EACT_H,MAAT2	2 54	65		70 74	79
K x Volk	PY2 HpaII	2EACT_H,MAAT2	1 54	65	68	70 74	79
K x Volk	reference	HpaII EACT_H,MAAT	1	54	65	68 70	74
K x Volk	reference	MspI EACT_H,MAAT	1		65		74
K x Volk	reference	HpaII 2EACT_H,MAAT2	1	54	65	68 70	74
K x Volk	reference	MspI 2EACT_H,MAAT2	1		65		74
K x Volk	A3 HpaII	EACT_H,MAAT	1 54	65	68	70 74	79
K x Volk	A3 MspI	EACT_H,MAAT	1 54	65	68	70 74	79
K x Volk	A3 HpaII	EACT_H,MAAT	2 54	65	68	70 74	76 79

B

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	Genotype	Treatmen	Ésoschizo	Enzyme combination	error dete	repeat	Allele 1	Allele 2	Allele 3	Allele 4	Allele 5	Allele 6	Allele 7	Allele 8
2	K x Volk	PY3	HpaII	EACT_H,MAAT		1								
3	K x Volk	PY3	MspI	EACT_H,MAAT		1				65			74	
4	K x Volk	PY3	HpaII	EACT_H,MAAT		2	54			65	68	70	74	
5	K x Volk	PY3	MspI	EACT_H,MAAT		2	54			65		70	74	
6	K x Volk	PY2	HpaII	EACT_H,MAAT		1	54			65	68	70	74	
7	K x Volk	PY2	MspI	EACT_H,MAAT		1	54			65		70	74	
8	K x Volk	PY2	HpaII	2EACT_H,MAAT2		1								
9	K x Volk	PY2	MspI	2EACT_H,MAAT2		1				65			74	
10	K x Volk	reference	HpaII	EACT_H,MAAT		1	54			65	68	70	74	
11	K x Volk	reference	MspI	EACT_H,MAAT		1				65			74	
12	K x Volk	reference	HpaII	2EACT_H,MAAT2		1	54			65	68	70	74	
13	K x Volk	reference	MspI	2EACT_H,MAAT2		1				65			74	
14	K x Volk	A1	HpaII	EACT_H,MAAT		1	54			65	68	70	74	
15	K x Volk	A1	MspI	EACT_H,MAAT		1	54			65		70	74	
16	K x Volk	A1	HpaII	EACT_H,MAAT		2	54			65	68	70	74	

Fig. 2. Example of input matrix in: **A)** .txt format and in **B)** ".Excel format - note the labeling after the genotype name (i.e. KxVolk-PY3, KxVolk -PY2, etc) to indicate that these are different treatments of the same genotype.

An example file is available at:

http://mirna.imbb.forth.gr/glossaryMSAPAnalyzer.html

5. Interactive Mode of operation

Following construction of the appropriate ".txt" files (see above) the user can upload a file (corresponding to the different primer combinations employed) for analysis. To gain access to file archives and to select files for uploading the "Browse" button should be used. The user must then press the "analyze" button to launch the module. Processed files appear as a new section in the listing window (Fig. 3). Files analyzed simultaneously must contain the same number of treatments.

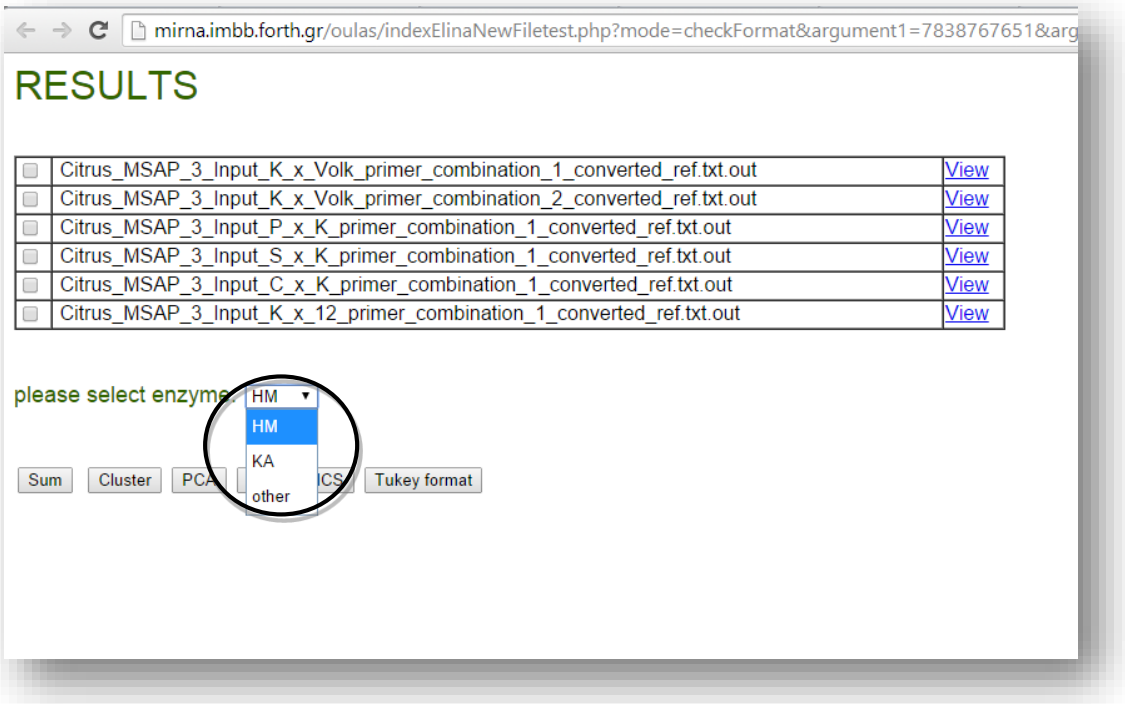


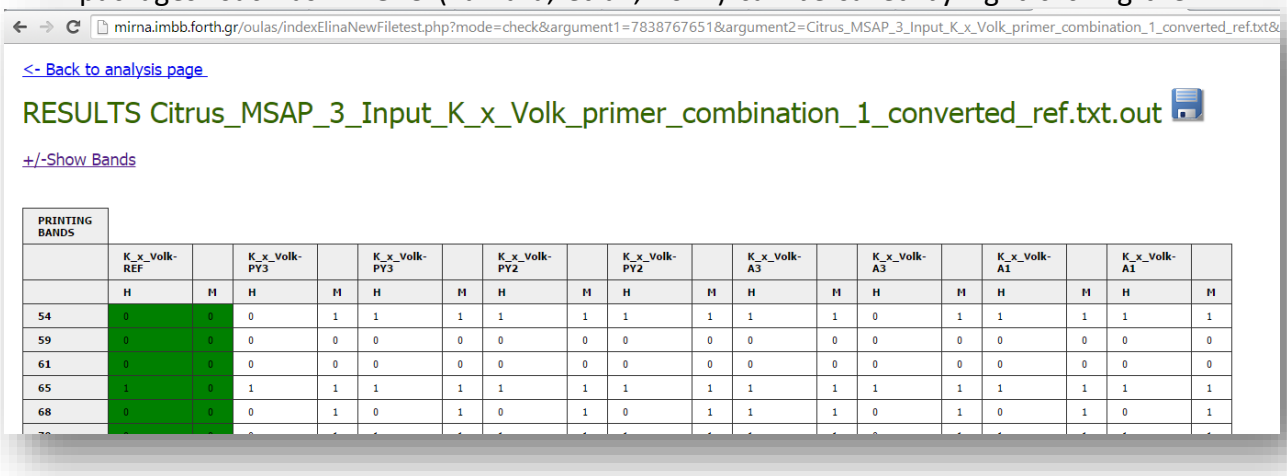
Fig. 3. First step of MSAP Analyzer calculation. The black circle indicates the enzymes menu bar providing a choice of available isoschizomers or the "other" option.

MSAP ANALYZER

An interactive interface for methylation-sensitive amplified polymorphism (MSAP) data analysis

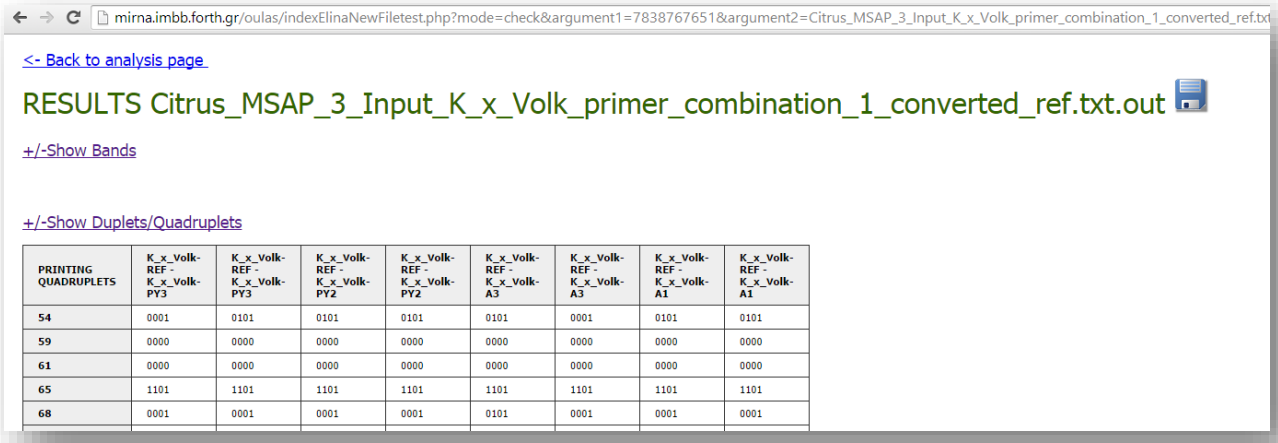
For inputted files to be submitted for analysis the left box must be selected. Then, by pressing the "View" hyperlink the user can then view the output file of transforming raw data into a binary (0, 1) matrix and then to quadruplets of 0s and 1s. The quadruplets are produced by pulling together the binary descriptions of the reference sample with each treated sample separately; they will be subsequently used for deducing methylation change events (see Fig. 4. a) & b). The "view" option is allowed for each file (PEC) separately and by selecting as many files as the user wants. Furthermore, the user has the option of selecting the pair of enzymes (HpaII/MspI – HM or Acc65I/KpnI - KA) employed for the double digest (see Fig.4). The user can subsequently advance with summarizing the different type of events, performing a cluster analysis or generating a file formatted for statistical analysis i.e. a Tukey test. Furthermore, a distance metric table which can be used by tree-drawing software packages such as MEGA5 (Tamura, et al., 2011) can be saved by right clicking the

A



PRINTING BANDS	K_x_Volk-REF		K_x_Volk-PY3		K_x_Volk-PY3		K_x_Volk-PY2		K_x_Volk-PY2		K_x_Volk-A3		K_x_Volk-A3		K_x_Volk-A1		K_x_Volk-A1	
	H	M	H	M	H	M	H	M	H	M	H	M	H	M	H	M	H	M
54	0	0	0	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1
59	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
61	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
65	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
68	0	0	0	1	0	1	0	1	0	1	1	1	0	1	0	1	0	1

B



PRINTING QUADRUPLETS	K_x_Volk-REF - K_x_Volk-PY3	K_x_Volk-REF - K_x_Volk-PY3	K_x_Volk-REF - K_x_Volk-PY2	K_x_Volk-REF - K_x_Volk-PY2	K_x_Volk-REF - K_x_Volk-A3	K_x_Volk-REF - K_x_Volk-A3	K_x_Volk-REF - K_x_Volk-A1	K_x_Volk-REF - K_x_Volk-A1
54	0001	0101	0101	0101	0101	0001	0101	0101
59	0000	0000	0000	0000	0000	0000	0000	0000
61	0000	0000	0000	0000	0000	0000	0000	0000
65	1101	1101	1101	1101	1101	1101	1101	1101
68	0001	0001	0001	0001	0101	0001	0001	0001

Fig. 4. "View" hyperlink option of the MSAP Analyzer. **A)** transformation of the raw data into 0-1 code and **B)** to quadruplets of 0s and 1s.

6. Summarizing events

The "Sum" analysis, is the second step in the analysis pipeline of the MSAP analyzer. When performing the "Sum" analysis, MSAP analyzer employs the quadruplets to deduce methylation pattern changes by combining already published methylation patterns from (Hao, et al., 2004) and the methylation events from (Fiuk, et al., 2010). For example, if the distribution of the following digests: maternal plant/EcoRI + MspI digestion; in vitro originated plant/ EcoRI + MspI digestion; maternal plant/EcoRI + HpaII digestion; in vitro originated plant/EcoRI + HpaII digestion, produces the following event type notation – 1111, the corresponding CCGG locus is probably not methylated. Similarly, a distribution of fragments evaluated as 1100 indicates the fully methylated internal cytosine in CCGG loci with no changes from the tissue culture treatment. Respectively, quadruplets are being generated for the Acc65I/KpnI isoschizomers according to Fiuk et al. (2010).

If the user wants to employ a pair of restriction enzymes different than the ones offered by MSAP Analyser (i.e. HpaII/MspI or Acc65I/KpnI) he can do so by selecting the "Other" button. An example of such a pair is indicated at Fig. 6 of (Samuelsson, et al.). In this case thought MSAP Analyzer will require introduction of user-specified methylation patterns. For that, the user will be prompted to insert the methylation patterns (in the form of "0s" and "1s") which are particular to the pair of enzymes involved. This is conducted manually in text boxes provided (see Fig 5). The "other" choice is equally useful when more than four binary digits (i.e. quadruplets)

are needed for a complete description of methylation change events. This is again particular to some experimental settings such as use of more than two isoschizomers, sequential treatments of the same sample or time series experiments.

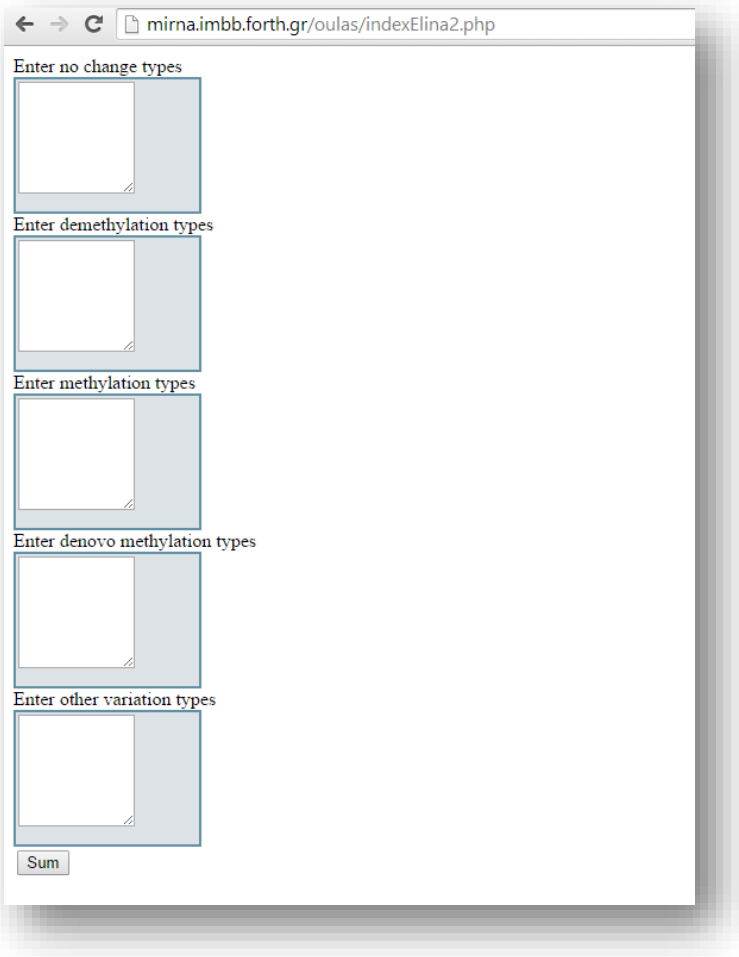


Fig. 5. Specifying unique methylation patterns. Text boxes displaying user defined methylation patterns after selecting the "other" option in the enzyme menu bar and clicking the "sum" button. Once the patterns are defined, the user can proceed by clicking the second "sum" button available on this interface.

As an output of the "Sum" operation a summing matrix is initially depicted (Fig. 6A) indicating the number of occurrences (different quadruplet types) of every event type (with respect to the known methylation patterns) for **each** individual genotype and from **all** primer combinations employed. The user is able to process as many individuals as he desires with as many different primer combinations. Every summing

matrix is then depicted as a total for all individuals and all primer combinations used (see Fig. 6B).

A

mirna.imbb.forth.gr/oulas/indexElina2.php?mode=check&argument1=7838767651&argumen

<- Back to analysis page

SUMMING PRIMER COMBINATIONS

+/-Show Event Patterns

Event Pattern	K_x_Volk-PY3	K_x_Volk-PY2	K_x_Volk-A3	K_x_Volk-A1
0000	187	175	187	193
0001	43	36	34	27
0010	0	0	0	0
0011	0	0	0	0
0100	30	36	49	36
0101	38	51	28	42
0110	0	0	2	1
0111	2	2	0	1
1000	20	14	20	18
1001	3	4	4	4
1010	0	0	0	0
1011	0	0	0	0
1100	3	5	11	9
1101	38	41	29	33
1110	0	0	1	1
1111	2	2	1	1

+/-Show Markers by Event

+/-Show Event Statistics

B

mirna.imbb.forth.gr/oulas/indexElina2.php?mode=check&argument1=7838767651&argument2=Citrus_MSAP_3_Input_K_x_Volk_primer_combination_1_converted_re

<- Back to analysis page

SUMMING PRIMER COMBINATIONS

+/-Show Event Patterns

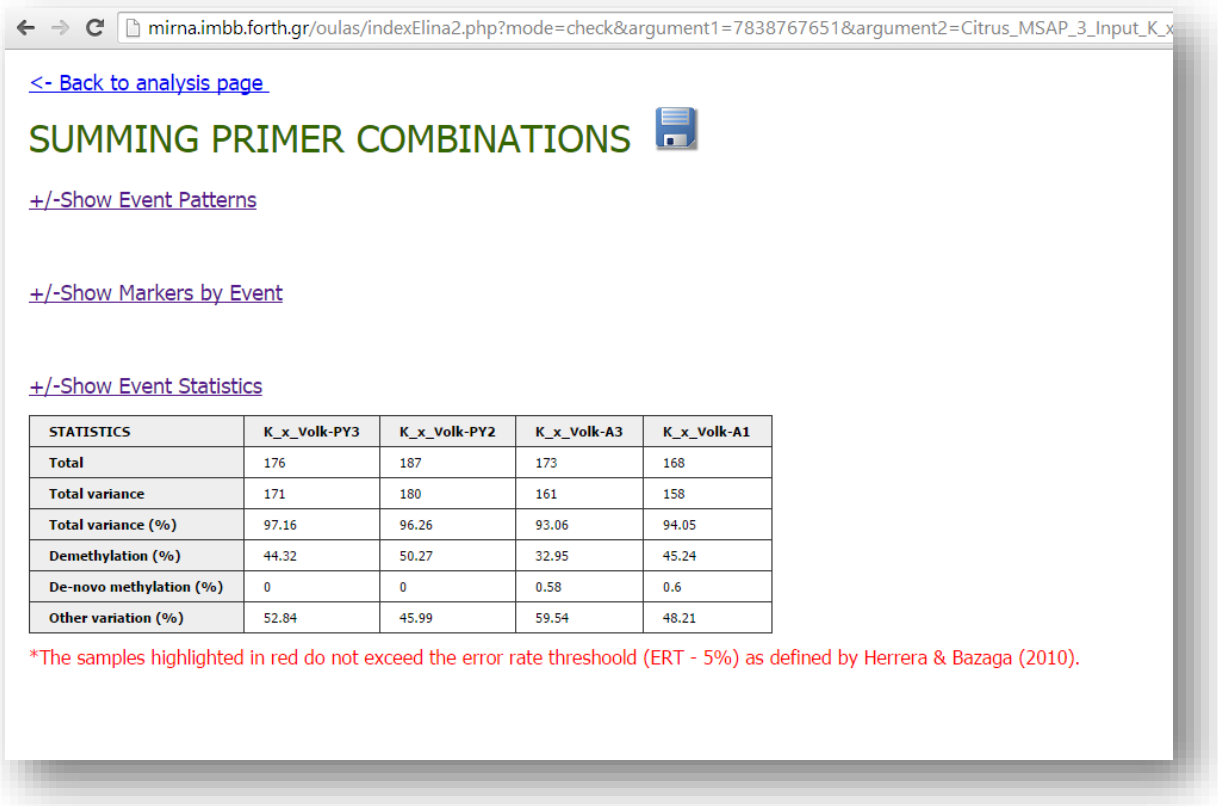
+/-Show Markers by Event

EVENT TYPES	Event	K_x_Volk-PY3	Citrus_MSAP_3_Input_K_x_Volk_primer_combination_2_converted_ref	Citrus_MSAP_3_Input_K_x_Volk_primer_combination_1_converted_ref	K_x_Volk-PY2
1111	No change	2	[378]	[378]	2
0011	No change	0	no_bands	no_bands	0
1100	No change	3	[227]	[227]	5
1010	De-novo methylation	0	no_bands	no_bands	0
1011	De-novo methylation	0	no_bands	no_bands	0
1110	De-novo methylation	0	no_bands	no_bands	0
0101	Demethylation	38	[79, 98, 115, 141, 143, 156, 174, 252, 261, 275, 287, 292, 315, 324, 336, 373, 387, 396, 426, 481]	[79, 98, 115, 141, 143, 156, 174, 252, 261, 275, 287, 292, 315, 324, 336, 373, 387, 396, 426, 481]	51
0111	Demethylation	2	[367]	[367]	2
1101	Demethylation	38	[65, 74, 83, 93, 124, 134, 153, 164, 168, 189, 192, 196, 206, 310, 330, 346, 375, 409, 455, 468]	[65, 74, 83, 93, 124, 134, 153, 164, 168, 189, 192, 196, 206, 310, 330, 346, 375, 409, 455, 468]	41

MSAP ANALYZER

An interactive interface for methylation-sensitive amplified polymorphism (MSAP) data analysis

C



STATISTICS	K_x_Volk-PY3	K_x_Volk-PY2	K_x_Volk-A3	K_x_Volk-A1
Total	176	187	173	168
Total variance	171	180	161	158
Total variance (%)	97.16	96.26	93.06	94.05
Demethylation (%)	44.32	50.27	32.95	45.24
De-novo methylation (%)	0	0	0.58	0.6
Other variation (%)	52.84	45.99	59.54	48.21

*The samples highlighted in red do not exceed the error rate threshoold (ERT - 5%) as defined by Herrera & Bazaga (2010).

Fig. 6. Second step of the analysis of MSAP Analyzer. **A)** Summing matrix of every event type (with respect to the known methylation patterns) for **each** individual separately and from **all** pair combinations used, **B)** Total summing matrix of all individuals and all pair combinations used, **C)** Analysis of MSAP Analyzer. Evaluation of the occurrences of methylation change or not of CCGG loci and relative distributions given known methylation patterns.

Furthermore, in the "sum" step (Fig. 6C), MSAP Analyzer calculates the total number of event types of **all** individuals and from **all** primer combinations used (Total). Next the total value of variation is computed (Total variance), which corresponds to

all types of variation (de-novo methylation, demethylation, other variation) divided by the total number of event types. Finally, the percentage of methylation patterns with respect to the total number of event types is calculated (Total variance, %). The percentage variation corresponding to every event type (De-novo methylation %, Demethylation %, Other variance %) is also calculated with respect to the total number of events. For example, the number of de novo methylation events divided by the total number of events identified relative to 100 results in the percentage value of de novo methylation events. These calculations are performed separately for every individual across all primer combinations (see Fig. 6C).

7. Cluster Analysis

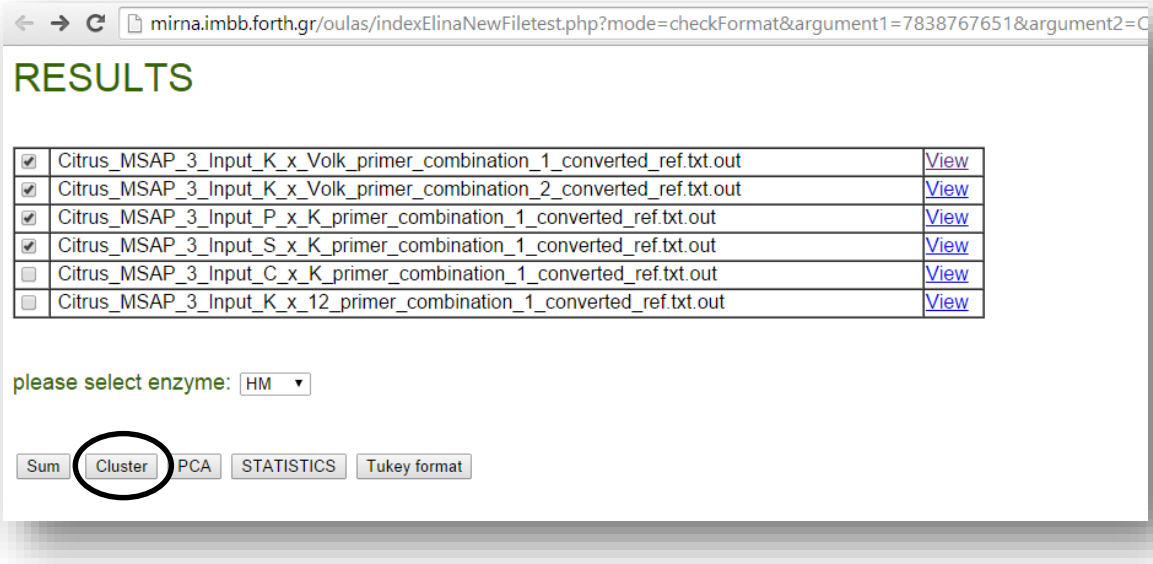


Fig. 7. Cluster option of MSAP Analyzer. The black circle indicates the cluster button.

The user can select one of the distance measures offered as well as the different clustering algorithms (see Fig. 8A) and execute the cluster analysis in order to view methylation effects in the form of a dendrogram (see Fig. 8B). MSAP Analyzer gives also the option, to print the dendrogram of his preference, by clicking on “print tree” option (see Fig.9B). For better printing results print as .pdf and then send the .pdf file to the printer. Finally, the user can click on a link to save the cluster matrix for

use in other clustering application. This saves a file with the extension suffix .clust and can be opened for viewing with any text editor.

A **B**

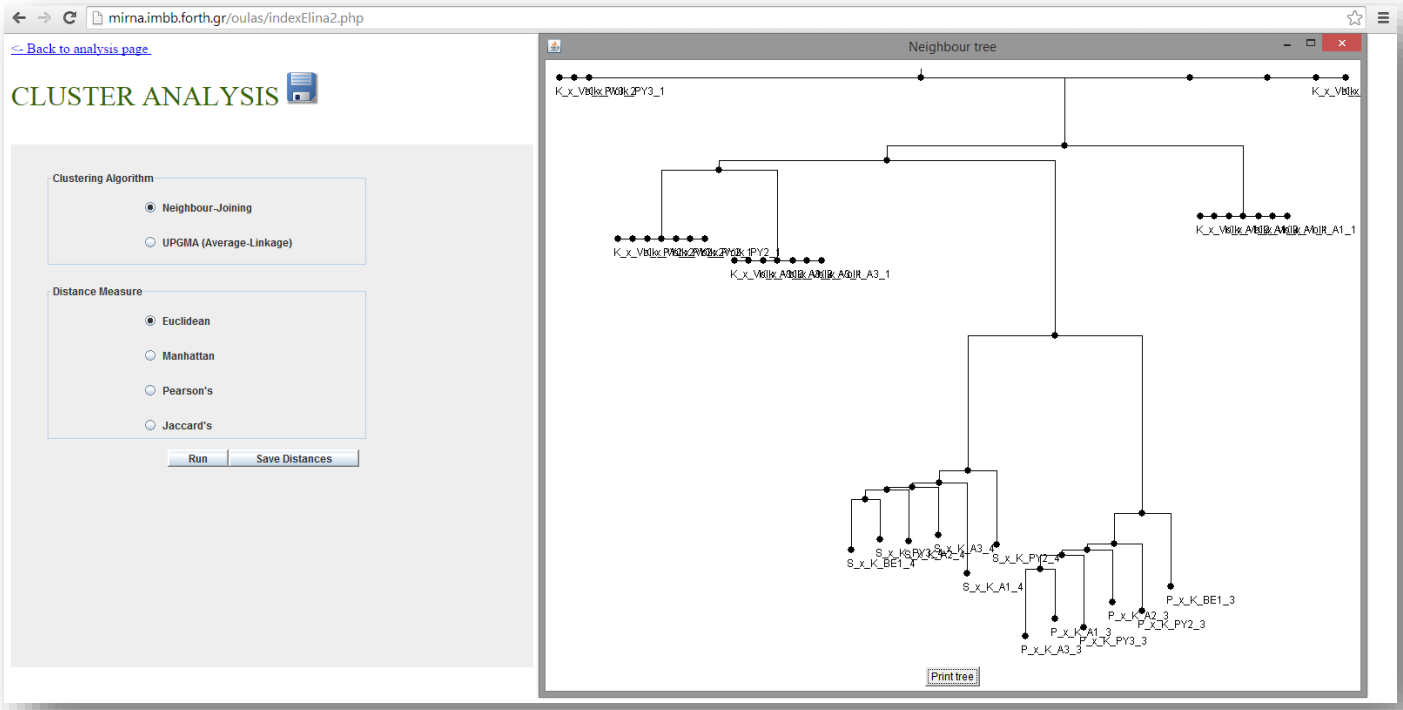


Fig. 8. Cluster analysis A) cluster choices offered by MSAP Analyzer in order to construct different types of dendrograms. **B)** one of the dendrograms generated by the "MSAP Analyzer".

8. Principal Component Analysis (PCA)

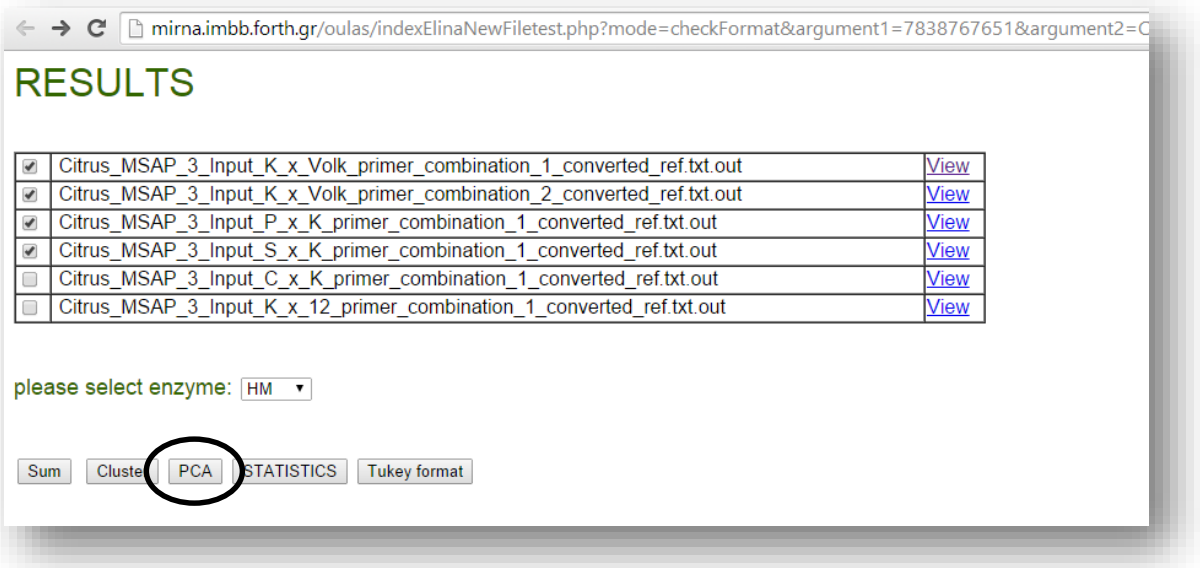


Fig. 9. PCA analysis output

The user can select to perform PCA analysis of selected samples (see Fig. 9,10) in order to perform dimensionality reduction an visualize samples in a 2D PCA plot

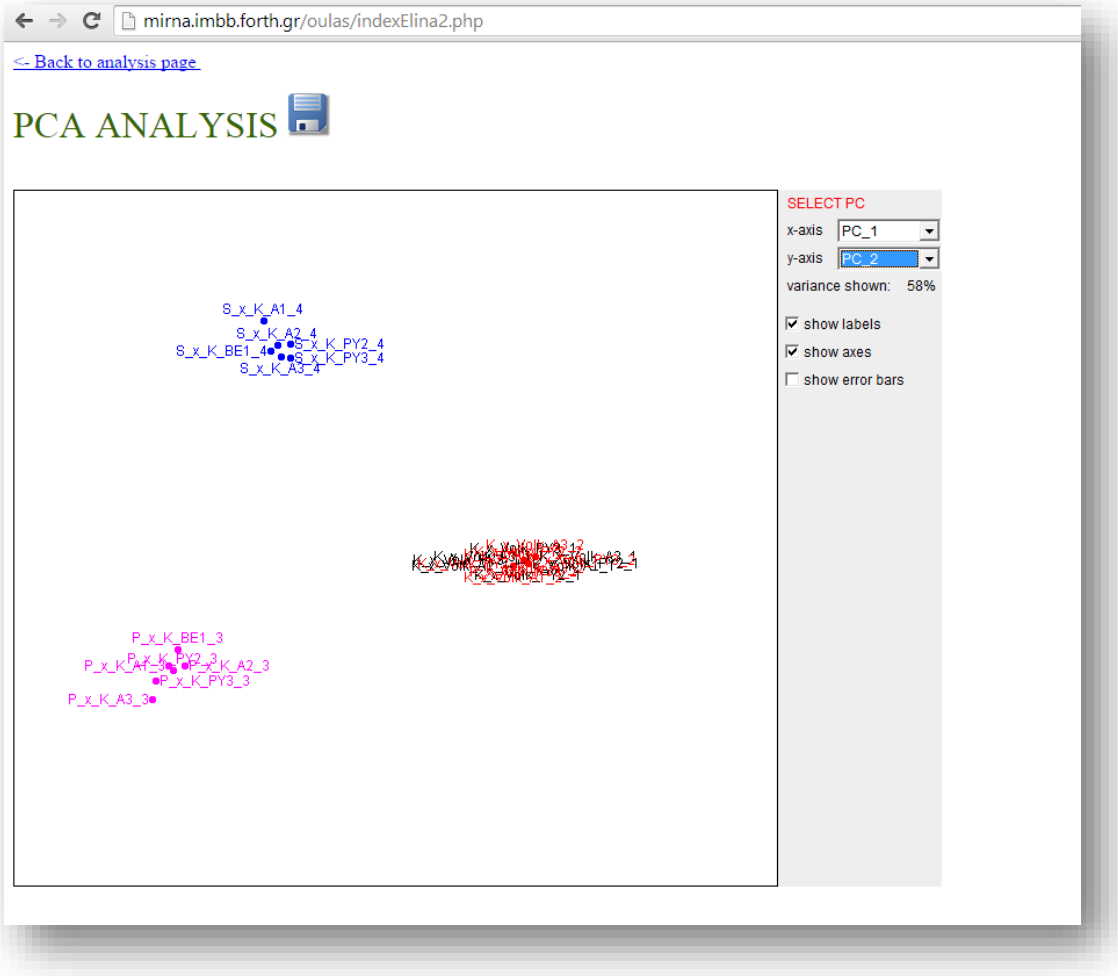
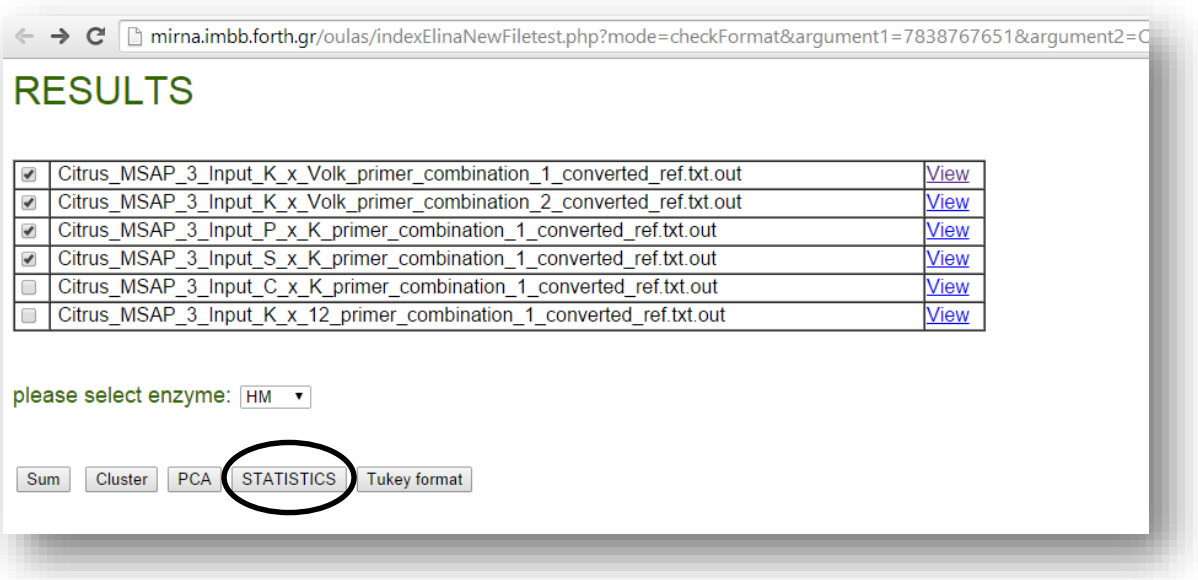


Fig. 10. PCA analysis output

9. Statistical evaluation of data

MSAP analyzer performs statistical analysis (see Fig 11A) using AMOVA testing for all genotypes uploaded in the system, It calculates statistical measures such as SQUARED DEVIATIONS (SSD, MSD) among and within groups as well as shannon index and overall information with respect to loci and primer combinations (see Fig 11B)

A



B

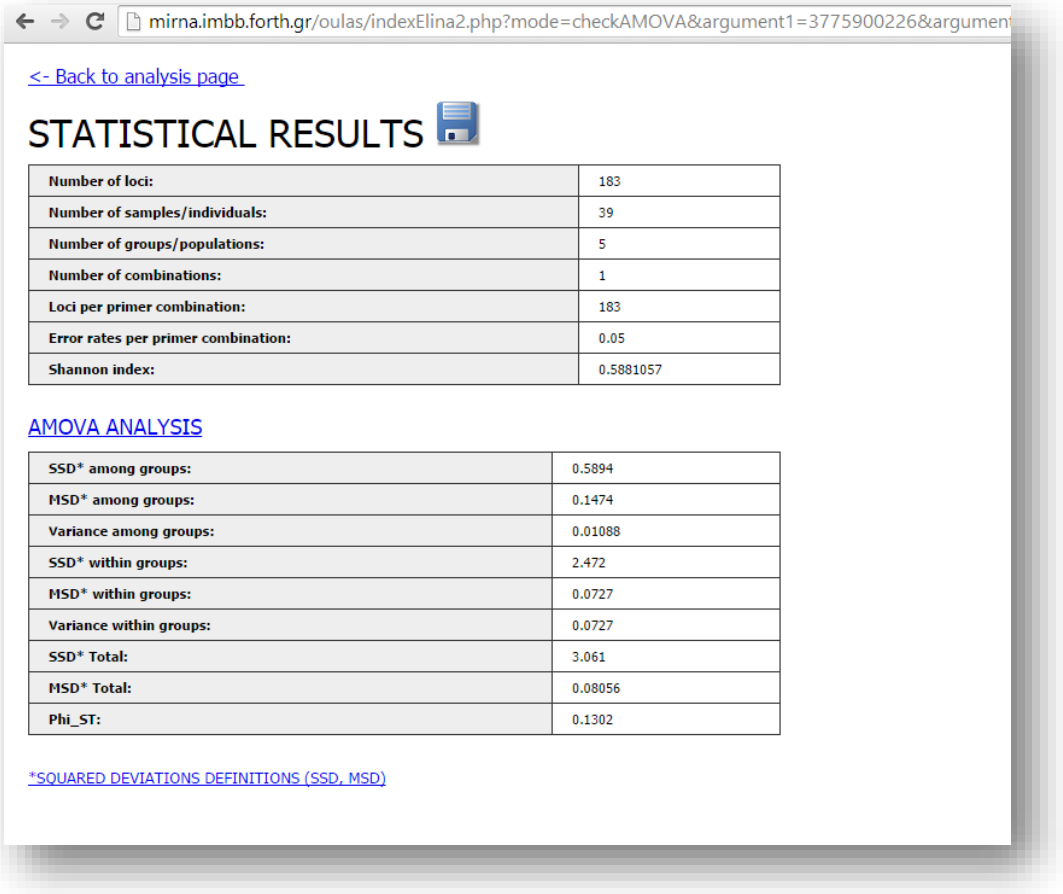


Fig 11 A) Statistical analysis button shown by black circle B) Analysis form AMOVA test on test dataset.

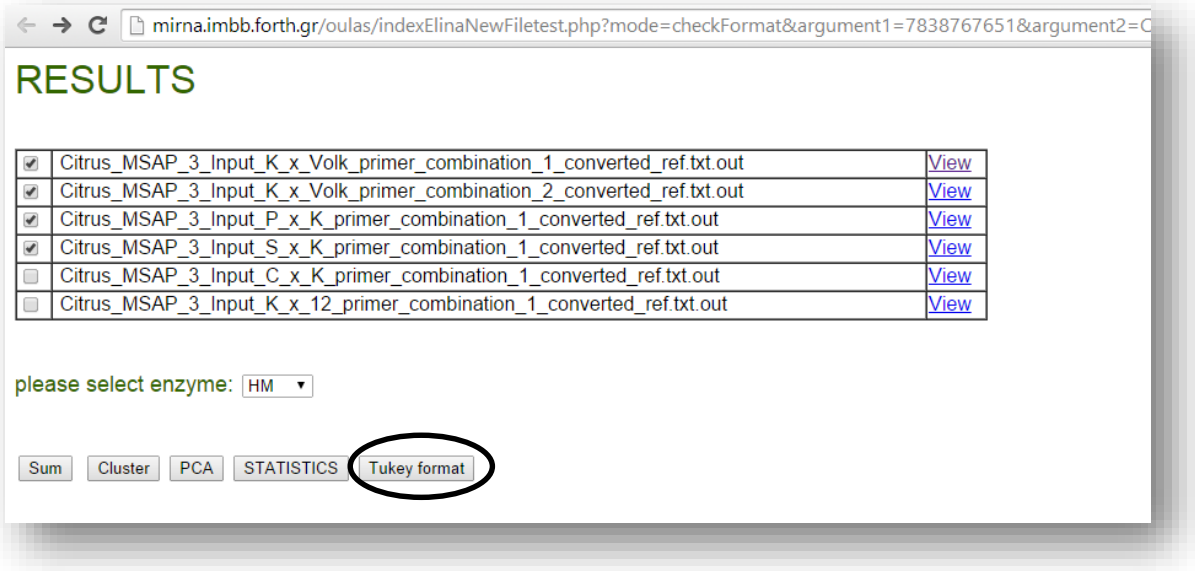
10.Generating Tukey test format file

Another possibility offered by MSAP Analyzer is generating matrices which can be used for any post-hoc comparisons i.e. additional statistical analyses such as Tukey test (see Fig 12 A & B).

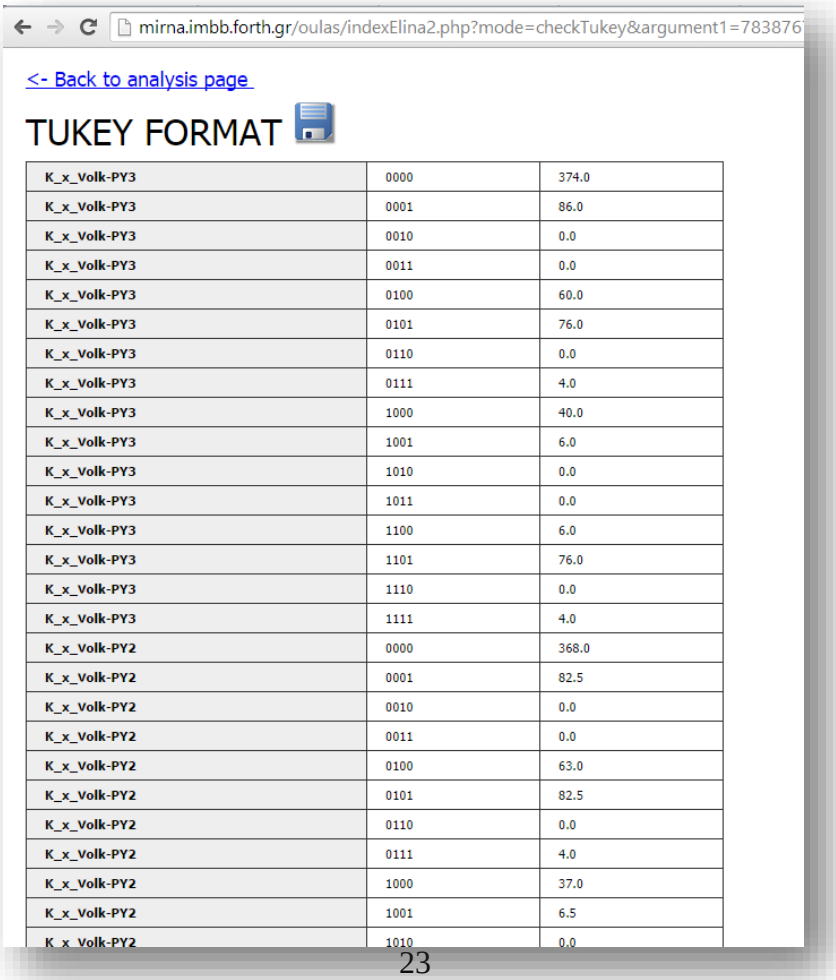
MSAP ANALYZER

An interactive interface for methylation-sensitive amplified polymorphism (MSAP) data analysis

A



B



K_x_Volk-PY3	0000	374.0
K_x_Volk-PY3	0001	86.0
K_x_Volk-PY3	0010	0.0
K_x_Volk-PY3	0011	0.0
K_x_Volk-PY3	0100	60.0
K_x_Volk-PY3	0101	76.0
K_x_Volk-PY3	0110	0.0
K_x_Volk-PY3	0111	4.0
K_x_Volk-PY3	1000	40.0
K_x_Volk-PY3	1001	6.0
K_x_Volk-PY3	1010	0.0
K_x_Volk-PY3	1011	0.0
K_x_Volk-PY3	1100	6.0
K_x_Volk-PY3	1101	76.0
K_x_Volk-PY3	1110	0.0
K_x_Volk-PY3	1111	4.0
K_x_Volk-PY2	0000	368.0
K_x_Volk-PY2	0001	82.5
K_x_Volk-PY2	0010	0.0
K_x_Volk-PY2	0011	0.0
K_x_Volk-PY2	0100	63.0
K_x_Volk-PY2	0101	82.5
K_x_Volk-PY2	0110	0.0
K_x_Volk-PY2	0111	4.0
K_x_Volk-PY2	1000	37.0
K_x_Volk-PY2	1001	6.5
K_x_Volk-PY2	1010	0.0

Fig. 12. A) Selecting for Tukey test option, **B)** formatted for Tukey test.

11.Saving files

The user can select to save results from any intermediate analysis step i.e the Tukey test format file, the user has to only press blue disk hyperlink (Fig. 13). This saves a file with the extension suffix .xls and can be opened as an excel spread sheet.

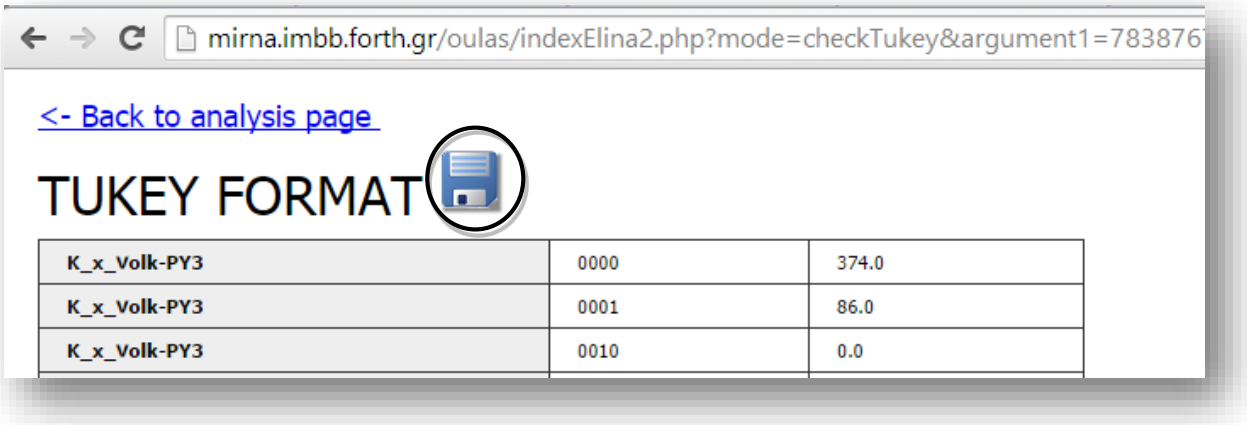


Fig. 13. How to save the "summing output" file or the Tukey test format file. The black circle shows the option to select after right clicking on the link.

12.Bibliography

Baranek, M., Krizan, B., Ondrusikova, E. and Pidra, M. (2010) DNA-methylation changes in grapevine somaclones following in vitro culture and thermotherapy, *Plant Cell Tiss Org*, **101**, 11-22.

Bednarek, P.T., Orlowska, R., Koebner, R.M. and Zimny, J. (2007) Quantification of the tissue-culture induced variation in barley (Hordeum vulgare L.), *BMC Plant Biol*, **7**, 10.

Cervera, M.T., Ruiz-Garcia, L. and Martinez-Zapater, J.M. (2003) Analysis of DNA methylation in Arabidopsis thaliana based on methylation-sensitive AFLP markers (vol 268, pg 543, 2002), *Mol Genet Genomics*, **268**, 832-833.

Duan, Y., Qian, J., Sun, Y., Yi, Z. and Yan, M. (2009) Construction of methylation linkage map based on MSAP and SSR markers in Sorghum bicolor (L.), *IUBMB Life*, **61**, 663-669.

Ebert, M.P., Mooney, S.H., Tonnes-Priddy, L., Lograsso, J., Hoffmann, J., Chen, J., Rocken, C., Schulz, H.U., Malfertheiner, P. and Lofton-Day, C. (2005) Hypermethylation of the TPEF/HPP1 gene in primary and metastatic colorectal cancers, *Neoplasia*, **7**, 771-778.

Fang, J.G., Song, C.N., Qian, J.L., Zhang, X.Y., Shangguan, L.F., Yu, H.P. and Wang, X.C. (2010) Variation of cytosine methylation in 57 sweet orange cultivars, *Acta Physiol Plant*, **32**, 1023-1030.

Fiuk, A., Bednarek, P.T. and Rybczynski, J.J. (2010) Flow Cytometry, HPLC-RP, and metAFLP Analyses to Assess Genetic Variability in Somatic Embryo-Derived Plantlets of Gentiana pannonica Scop, *Plant Mol Biol Rep*, **28**, 413-420.

Hao, Y.J., Wen, X.P. and Deng, X.X. (2004) Genetic and epigenetic evaluations of citrus calluses recovered from slow-growth culture, *Journal of Plant Physiology*, **161**, 479-484.

Jorda, M. and Peinado, M.A. (2010) Methods for DNA methylation analysis and applications in colon cancer, *Mutat Res*, **693**, 84-93.

Laird, P.W. (2003) The power and the promise of DNA methylation markers, *Nat Rev Cancer*, **3**, 253-266.

Martienssen, R.A. and Colot, V. (2001) DNA methylation and epigenetic inheritance in plants and filamentous fungi (vol 293, pg 1070, 2001), *Science*, **293**, 1997-1997.

McClelland, M., Nelson, M. and Raschke, E. (1994) Effect of site-specific modification on restriction endonucleases and DNA modification methyltransferases, *Nucleic Acids Res*, **22**, 3640-3659.

Menz, M.A., Klein, R.R., Mullet, J.E., Obert, J.A., Unruh, N.C. and Klein, P.E. (2002) A high-density genetic map of Sorghum bicolor (L.) Moench based on 2926 AFLP, RFLP and SSR markers, *Plant Mol Biol*, **48**, 483-499.

Peredo, E.L., Arroyo-Garcia, R. and Revilla, M.A. (2009) Epigenetic changes detected in micropropagated hop plants, *J Plant Physiol*, **166**, 1101-1111.

Salmon, A. and Ainouche, M.L. (2010) Polyploidy and DNA methylation: new tools available, *Mol Ecol*, **19**, 213-215.

Samuelsson, J.K., Alonso, S., Yamamoto, F. and Perucho, M. DNA fingerprinting techniques for the analysis of genetic and epigenetic alterations in colorectal cancer, *Mutat Res*, **693**, 61-76.

Smykal, P., Valledor, L., Rodriguez, R. and Griga, M. (2007) Assessment of genetic and epigenetic stability in long-term in vitro shoot culture of pea (Pisum sativum L.), *Plant Cell Rep*, **26**, 1985-1998.

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M. and Kumar, S. (2011) MEGA5: Molecular Evolutionary Genetics Analysis Using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony Methods, *Mol Biol Evol*, **28**, 2731-2739.