Methylation-sensitive Amplified Polymorphism as a Tool to Analyze Wild Potato Hybrids

Nicolás Cara, Carlos F. Marfil*, María V. Bertoldi and Ricardo W. Masuelli

Universidad Nacional de Cuyo, Facultad de Ciencias Agrarias. Instituto de Biología Agrícola de Mendoza (IBAM) Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET). Almirante Brown 500, Chacras de Coria. Mendoza. M5528AHB, Argentina

*For correspondence: cmarfil@fca.uncu.edu.ar

[Abstract] Methylation-Sensitive Amplification Polymorphism (MSAP) is a versatile marker for analyzing DNA methylation patterns in non-model species. The implementation of this technique does not require a reference genome and makes it possible to determine the methylation status of hundreds of anonymous loci distributed throughout the genome. In addition, the inheritance of specific methylation patterns can be studied. Here, we present a protocol for analyzing DNA methylation patterns through MSAP markers in potato interspecific hybrids and their parental genotypes.

Keywords: MSAP, DNA Methylation Pattern Inheritance, Epigenetic changes, *Solanum*, Interspecific hybridization, Epialleles, R Script

[Background] Nucleotide sequences are not the only form of genomic information: DNA methylation, histone proteins, enzymes that modify histones and nucleotide residues on DNA and even RNAs, influence gene activity and provide another layer of instructions to the cell. Epigenetic changes, also called *epimutations*, can be inherited and have important phenotypic consequences. In plants, methylation reactions modify cytosine residues into 5-methylcytosines. This epigenetic mechanism is essential to maintain the genomic integrity and contributes to regulate gene expression in developmental processes and in response to biotic and abiotic stresses. In addition, changes in DNA methylation are triggered by genomic shocks like hybridization and polyploidization, two vital phenomena in plant evolution (Cara *et al.*, 2019).

There are different alternatives to study changes in DNA methylation. Global cytosine methylation can be assessed by using high-performance liquid chromatography (HPLC), analytical methods that allow to quantify cytosines and 5-methylcytosines and to calculate the percentage of methylated residues in the genome. For studying the DNA methylation at specific positions on the genome two alternatives can be mentioned. One is to use isoschizomers with different sensitivities to methylation at cytosines of the restriction site. For example, Methylation-Sensitive Amplification Polymorphism (MSAP) markers characterize the methylation pattern at anonymous 5′-CCGG sequences from random genomic DNA. This is an adaptation of the original AFLP protocol (Vos *et al.*, 1995) substituting the frequent cutter enzyme *Mse*I by *Hpa*II and *Msp*I. These enzymes recognize the same tetranucleotide restriction site (5′-CCGG), but *Hpa*II is sensitive to full methylation (both strands methylated) and cleaves the hemi-methylated external cytosine, whereas *Msp*I is sensitive only to methylation of the external cytosines of the restriction site. Another possibility to study site-specific methylation status is

by using a bisulphite sequencing methodology. Bisulphite treatments on DNA convert cytosines into uracils while 5-methylcytosines remain unchanged. Then, by sequencing amplicons (*i.e.*, a target gene or promoter) of bisulphited and control DNA, it is possible to distinguish between methylated and unmethylated cytosines. With the development of next generation sequencing technologies, whole-genome bisulphite sequencing, or WGBS, can be implemented to infer the position for all 5-methylcytosines in a genome. However, this approach requires a high-quality reference genome to perform epigenetic analyses. Although the increasing number of draft genomes and the reduction in sequencing costs offer possibilities to implement massive methylation analyses in non-model species, the use of MSAP markers continues to be a valuable tool in many laboratories.

Wild potatoes (*Solanum*, section *Petota*) are a group of species related to the cultivated potato *Solanum tuberosum* L. Internal breeding barriers can be incomplete, thus, interspecific hybridization occurs in areas of sympatry (Camadro *et al.*, 2012). Epigenetic changes in response to interspecific hybridization have been documented in synthetic and natural hybrids of wild potato species (Cara *et al.*, 2019). *Solanum* x *rechei* H. & H. is a hybrid species that grows in sympatry with its wild progenitors, *Solanum kurtzianum* B. & W. and *Solanum microdontum* B. Here, we present a protocol for analyzing DNA methylation patterns through MSAP markers in hybrids and their parental genotypes. Using an R script, fragments present in the synthetic hybrids are categorized as *S. microdontum* or *S. kurtzianum* species-specific if they are present on the parental genotypes, *S.* x *rechei* species-specific, if they are present in at least one of the *S.* x *rechei* evaluated genotypes or as novel if they are only observed in the synthetic hybrids.

Materials and Reagents

A. Consumables
  1. Microfuge tubes (DELTALAB, catalog numbers: 4095.5N, 4095.9N)
  2. Pipette tips (DELTALAB, catalog numbers: 200072, 200016)
  3. PCR microplate 96-well (Axygen, catalog number: 32165051)
  4. Plastic pestle (Sigma, catalog number: Z359947)

B. Chemicals
  1. Tris base (BIOPACK, catalog number: 2000166800)
  2. Ethylenediaminetetraacetic acid (EDTA) (BIOPACK, catalog number: 2000964500)
  3. Sodium Chloride (BIOPACK, catalog number: 2000164606)
  4. Cetyltrimethyl ammonium bromide (CTAB) (Bio Basic INC, catalog number: DB0108)
  5. β-mercaptoethanol (BIOPACK, catalog number: 2000954500)
  6. Chloroform (BIOPACK, catalog number: 2000165100)
  7. Isoamyl alcohol (BIOPACK, catalog number: 2000972500)
  8. Ethyl alcohol (BIOPACK, catalog number: 2000165400)
  9. Sodium acetate (BIOPACK, catalog number: 2000168000)

10. RNase A (Thermo Fisher Scientific, catalog number: EN0531)

11. Lambda DNA/EcoRI+HindIII Marker (Promega, catalog number: G1731)

12. Glycerol (BIOPACK, catalog number: 2000162000)

13. Bromophenol blue (BIOPACK, catalog number: 2000962200)

14. Xylene cyanol FF (SIGMA, catalog number: X4126)

15. Boric acid (BIOPACK, catalog number: 2000935900)

16. Agarose (TransGen, catalog number: GS201)

17. UltraPure^{TM} Ethidium Bromide (Thermo Fisher Scientific, catalog number: 15585011)

18. *Eco*RI (New England Biolabs, catalog number: R0101S)

19. Bovine Serum Albumin (BSA) (Promega, catalog number: R3961)

20. *Hpa*II (New England Biolabs, catalog number: R0171S)

21. *Msp*I (New England Biolabs, catalog number: R0106)

22. T4 DNA Ligase (Promega, catalog number: M1801)

23. Oligonucleotides (Table 1) (IDT, Integrated DNA Technologies, Inc., Iowa, USA)

Table 1. Sequences of adaptors and primers used

Oligonucleotides	Sequence (5′-3′)
Adaptors	
EcoRI	CTCGTAGACTGCGTACC
	AATTGGTACGCAGTCTAC
HpaII/MspI	GACGATGAGTCTCGAT
	CGATCGAGACTCATC
Pre-amplification primers	
EcoRI +0	GACTGCGTACCAATTC
HpaII/MspI +1	ATGAGTCTCGATCGGA
Amplification primers	
EcoRI +3	6-FAM-GACTGCGTACCAATTC +3
HpaII/MspI +3	ATGAGTCTCGATCGGA +3

24. *Taq* DNA Polymerase (Thermo Fisher Scientific, catalog number: 11615044)

25. $MgCl_2$ (Thermo Fisher Scientific, catalog number: 11615044)

26. dNTP Set (100 mM) (Thermo Fisher Scientific, catalog number: 10297018)

27. Hi-Di formamide (Thermo Fisher Scientific, catalog number: 4311320)

28. GeneScan 500HD Rox size standard (Thermo Fisher Scientific, catalog number: 401734)

29. ddH_2O (sterile)

30. Extraction buffer (see Recipes)

31. TBE (Tris-Borate-EDTA) buffer (see Recipes)

32. 6x DNA loading buffer (see Recipes)

Equipment

1. Pipettes (BOECO, Wheaton SOCOREX and Finnpipette)
2. Vortex (IKA, model: MS1)
3. UltraCentrifuge (Eppendorf, model: 5804R)
4. Thermostatic bath (Vicking, model: Masson)
5. Agarose gel electrophoresis system (Bio-Rad Laboratories, model: 1645056)
6. Microwave oven (HITPLUS, model: CM203M)
7. Spectrophotometer (AmpliQuant, model: AQ-07)
8. Gel Imager (Bio-Rad Laboratories, model: Gel Doc 1000)
9. Incubator (SANYO, model: MIR 262)
10. Thermocycler Veriti 96 well (Applied Biosystems, model: 9902)
11. Genetic Analyzer (Invitrogen, model: ABI PRISM 3130)
12. Ultra-low Freezer -80 °C (Forma Scientific, model: 8270)

Software

1. GeneMapper v3.7. (Applied Biosystems, Foster City, CA, USA)
2. R v2.15.1. (R Foundation for Statistical Computing, Vienna, Austria)
3. RStudio v1.1.442 (RStudio Inc., Boston, MA, USA)

Procedure

A. DNA extraction procedure
  1. Collect 50-100 mg of leaf tissue in a 1.5 ml centrifuge microtube. Flash-freeze the tissue in liquid nitrogen and grind it with a plastic pestle.
  2. Add 300 µl of extraction buffer and mix it vigorously by vortexing (5-10 s) to homogenize the tissue. Incubate the homogenate at 65 °C for 30 min in a thermostatic bath.
  3. Add 100 µl of chloroform:isoamyl alcohol (24:1) and mix the sample gently by inversion.
  4. Centrifuge the homogenate at 20,800 x g at 4 °C for 5 min.
  5. By using a 200 µl micropipette, carefully pipette the aqueous phase and transfer it to a new microtube.
  6. Add 500 µl of ice-cold ethanol:acetate (24:1) and mix the sample by inverting the microtube 2-3 times to precipitate the DNA, and let it settle at -20 °C for 30 min.
  7. Centrifuge the sample at 20,800 x g at 4 °C for 10 min to form a DNA pellet at the bottom of the microtube.
  8. Discard the supernatant and clean the pellet with 100 µl of 70% ethanol. Vortex the tube for 1 min to wash the pellet.
  9. Centrifuge the sample at 15,300 x g at room temperature for 5 min and let the pellet dry by

inverting the microtube onto a paper towel for 30 min until the ethanol evaporates completely or by placing it in a drying chamber for 15 min.

10. Resuspend the DNA in 50 µl of ultrapure water with RNase A (50 µg·ml^{-1}). Incubate at room temperature for 30 min.

11. Store the DNA sample at 4 °C until use. If the DNA will not be used for a long time, store it at -20 °C.

12. Check the integrity and purity of the extraction by loading 2 µl of the DNA solution on a 0.8% agarose gel and running it at 90 V for 30 min. It should show a thick band of high molecular weight, above 21 Kbp, with no smear (Figure 1).

13. Measure the concentration of DNA samples in a spectrophotometer and prepare working solutions of ~100 ng·µl^{-1}.
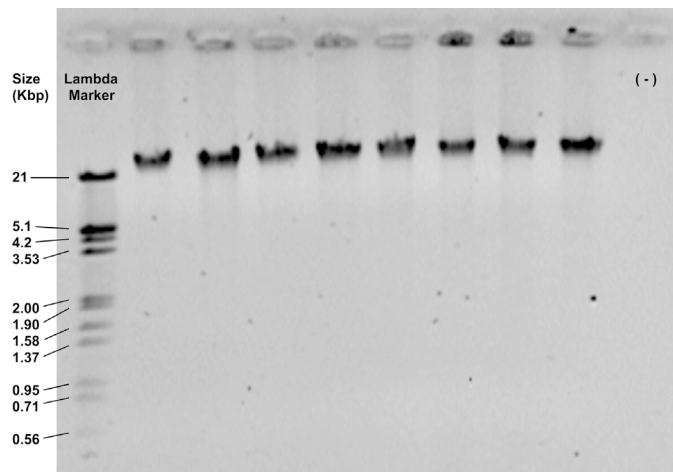

Figure 1. Representative image of DNA samples visualized in a 0.8% agarose gel

B. Digestion with *EcoR*I

1. Calculate the volume of the sample that corresponds to 700 ng of DNA, according to its concentration. However, consider that DNA concentration should be higher than 40.5 ng·µl^{-1}, otherwise it will exceed the reaction's final volume.

2. Calculate the amount of each reagent needed for the total number of samples based on Table 2. To compensate pipetting errors, it is convenient to consider an additional sample to the total number of analyzed samples.

Table 2. Master mix for the enzymatic digestion with *EcoR*I

Reagents	µl per sample
DNA (700 ng)	X
Buffer 10x	2
*EcoR*I 20 U/µl (NEB)	0.5
BSA 100x	0.2
H_2O (ultrapure)	$17.3 - x$
Final volume	20

3. Prepare the Master mix, homogenize by vortexing (2-3 s) and centrifuge for 10 s.
4. Distribute the Master mix in individual tubes and add the DNA. Homogenize gently by inversion or by flicking the tubes (do not vortex) and centrifuge for 10 s.
5. Incubate at 37 °C overnight in an incubator.
6. Place the tubes on ice if you continue with the next step. Otherwise, store them in freezer at -20 °C until use.

C. Digestion with *Hpa*II and *Msp*I

1. For each tube from the previous reaction (B), distribute the content into two tubes (10 µl in each tube) and label the new tubes with the sample code, adding the letter "H" to one and the letter "M" to the other. From now on you will work with two tubes for each analyzed sample.
2. Two Master mixes will be prepared, only differing on the restriction enzyme, one with *Hpa*II and the other with *Msp*I. For the two reactions, calculate the required amount of each reagent for the total number of samples plus one, based on Table 3.

Table 3. Master mixes for enzymatic digestions with *Hpa*II y *Msp*I

Master mix H	
Reagents	µl per sample
*EcoR*I digestion	10
Buffer 1 10x	2
*Hpa*II 10 U/µl (NEB)	1
BSA 100x	0.2
H$_2$O (ultrapure)	6.8
Final volume	20
Master mix M	
Reagents	µl per sample
*EcoR*I digestion	10
Buffer 4 10x	2
*Msp*I 20 U/µl (NEB)	0.5
BSA 100x	0.2
H$_2$O (ultrapure)	7.3
Final volume	20

3. Prepare the Master mixes, homogenize by vortexing (2-3 s) and centrifuge for 10 s.

4. Distribute the Master mix H in the tubes labeled with the letter "H" and the Master mix M in the tubes labeled with the letter "M". From now on you will have two tubes per sample. Homogenize gently by inversion or by flicking the tubes (do not vortex) and centrifuge for 10 s.

5. Incubate at 37 °C overnight in an incubator.

6. (Optional) Examine the efficiency of both digestions by running 6 µl of the restriction products on a 1.2% agarose gel, at 90 V for 30 min. It should show a subtle smear from 5,100 to 100 bp (Figure 2).
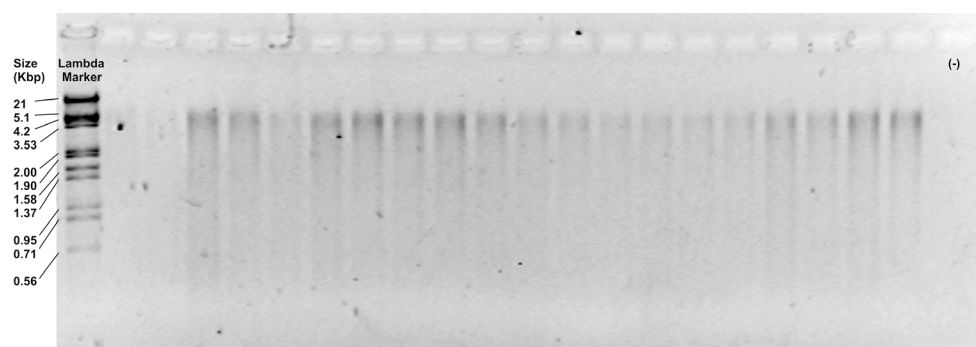

Figure 2. Representative image of restriction products visualized in a 1.2% agarose gel

7. Keep the tubes on ice if you will continue with the next step. Otherwise, store them in the freezer at -20 °C until use.

D. Adaptors ligation

1. Prepare the double stranded *Eco*RI and *Msp*I/*Hpa*II adaptors from the single strand oligonucleotides by adding equal amounts of each oligo in a tube (Table 1). Place the tubes in the thermocycler and run the ADAPTORS program (Table 4) with a hot lid.

Table 4. PCR programs

ADAPTORS			PREAMPLIFICATION			AMPLIFICATION			DENATURATION		
Step	Time	°T	Step	Time	°T	Step	Time	°T	Step	Time	°T
1	5 min	95 °C	1	30 s	94 °C	1	30 s	94 °C	1	3 min	90 °C
2	1 min	94 °C	2	1 min	56 °C	2	30 s	65 °C	2	∞	4 °C
3	Go to step 2 x69, decreasing 1 °C per cycle		3	1 min	72 °C	3	1 min	72 °C			
			4	Go to step 1 x19		4	Go to step 1 x13, decreasing 0.7 °C per cycle				
4	∞	4 °C	5	∞	4 °C						
						5	30 s	94 °C			
						6	30 s	56 °C			
						7	1 min	72 °C			
						8	Go to step 5 x22				
						9	∞	4 °C			

2. Based on Table 5, calculate the amount of each reagent for the total number of samples plus one.

Table 5. Master mix for the ligation reaction

Reagents	µl per sample
Digestion	10
Adaptor EcoRI	1.25
Adaptor MspI/HpaII	1.25
10x ligase buffer	2
T4 ligase 3 U/µl (Promega)	0.25
H_2O (ultrapure)	5.25
Final volume	20

3. Prepare the Master mix, homogenize by vortexing (2-3 s) and centrifuge for 10 s.
4. Distribute the Master mix in individual tubes and add 10 µl of digestions. Homogenize gently by inversion or by flicking the tubes (do not vortex) and centrifuge for 10 s.
5. Incubate at room temperature for 3 h.
6. Keep the tubes on ice if you will continue with the next step. Otherwise, store them in a freezer at -20 °C until use.

E. First amplification (Preamplification)

For this PCR reaction, oligonucleotides with the same sequences as the adaptors plus 1 selective nucleotide are used. These oligonucleotides are identified as Primer EcoRI +1 and Primer H/M +1 (Table 1).

1. Calculate the necessary amount of each reagent needed for the total samples based on Table 6, considering an additional sample to compensate for pipetting errors.

Table 6. Master mix for the pre-amplification

Reagents	µl per sample
Ligation	2
Buffer 10x	2
$MgCl_2$ 50 mM	0.6
dNTPs 2 mM	1
Primer EcoRI +0 20 µM	0.2
Primer H/M +1 20 µM	0.2
Taq Polymerase 5 U/µl (Invitrogen)	0.2
H_2O (ultrapure)	13.8
Final volume	20

2. Prepare the Master mix, homogenize by vortexing (2-3 s) and centrifuge for 10 s.

3. Distribute 18 µl of the Master mix in each tube and add 2 µl of the ligations. Add 2 µl of water to the remaining mix to use it as a negative reaction control. Homogenize gently by inversion or flicking (do not vortex) and centrifuge for 10 s.

4. Place the samples plus the negative control in the thermocycler and run the PREAMPLIFICATION program (Table 4) with a hot lid.

5. Check the efficiency of the reaction by loading 5 µl of the pre-amplifications on a 1.2% agarose gel and running at 90 V for 30 min. It should show an intense smear with few or some diffuse bands from 1,500 to 100 bp (Figure 3).
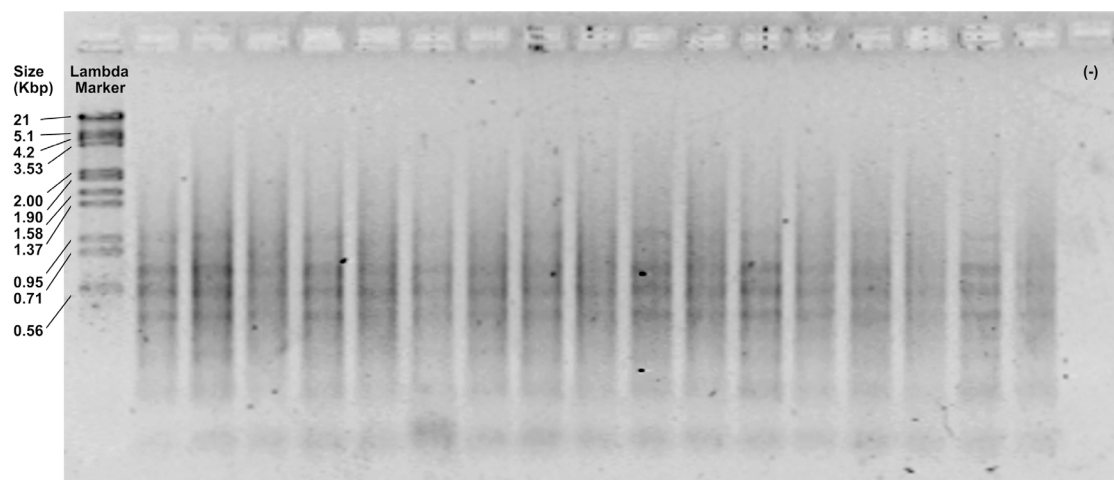

Figure 3. Representative image of pre-amplification products visualized in a 1.2% agarose gel

6. Dilute the pre-amplifications 1:3. Keep on ice if you continue with the next step or store in a freezer at -20 °C. Preamplifications can be safely stored in a freezer for several months.

F. Second amplification (Selective amplification)

    For this PCR reaction, oligonucleotides with the same sequences as the adaptors plus 3 selective nucleotides are used. These oligonucleotides are identified as Primer EcoRI +3 and Primer H/M +3 (Table 1).

1. Calculate the necessary amount of each reagent needed for the total number of samples plus one, based on Table 7.

Table 7. Master mix for the amplification reaction

Reagents	µl per sample
Pre-amplification 1:3	1
Buffer 10x	1
MgCl₂ 50 mM	0.3
dNTPs 2 mM	0.5
Primer EcoRI +3 4 µM	0.5
Primer H/M +3 20 µM	0.1
Taq Polymerase 5 U/µl (Invitrogen)	0.05
H₂O (ultrapure)	6.55
Final volume	10

2. Prepare the Master mix, homogenize by vortexing (2-3 s) and centrifuge for 10 s.

3. Distribute 19 µl of the Master mix in each tube and add 1 µl of 1:3 diluted preamplifications. To the remaining mix add 1 µl of water to use it as a negative reaction control. Homogenize gently by inversion or flicking (do not vortex) and centrifuge for 10 s.

4. Place in the thermocycler and run the AMPLIFICATION program (Table 4) with a hot lid.

5. Check the efficiency of the reaction by loading 5 µl of the amplification products on a 1.2% agarose gel and running at 90 V for 30 min. Depending of the primer combination, it should show a smear with diffuse bands from 700 to 100 bp (Figure 4).
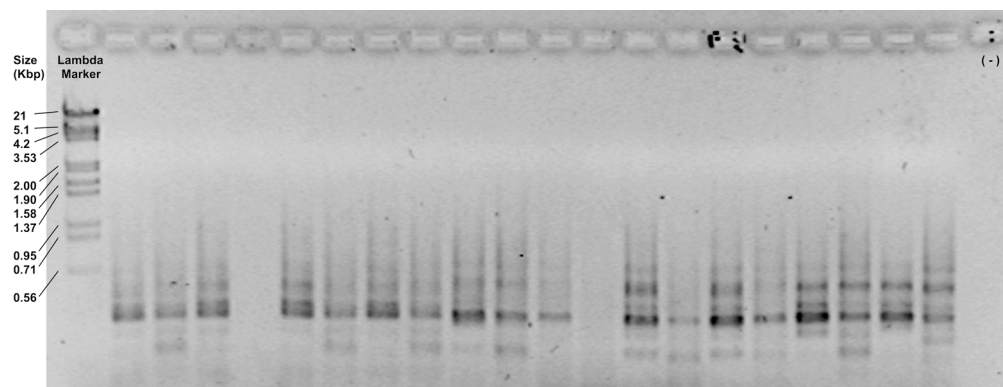

Figure 4. Representative image of amplification products visualized in a 1.2% agarose gel

6. Keep the tubes on ice if you continue with the next step or store them in a freezer at -20 °C. Fluorescence-labeled amplifications can be safely stored at least for 6 months at -20 °C maintaining their light emission.

G. Electrophoresis

1. Calculate the necessary amount of each reagent needed for the total samples based on Table 8, considering an extra sample to compensate for pipetting errors.

Table 8. Master mix for the amplification reaction

Reagents	µl per sample
Amplification	1
Hi-Di formamide	8.5
GeneScan 500HD ROX (Applied Biosystems)	0.5
Final volume	10

2. Prepare the Master mix, homogenize by vortexing (2-3 s) and centrifuge for 10 s.

3. Distribute 9 µl of the Master mix in each well of the 96-well plate and add 1 µl of amplification products.

4. Place the plate in the thermocycler and run the program DENATURATION (Table 4) with a hot lid.

5. Take the plate out of the thermocycler and put it on ice.

6. Place the plate in the ABI Prism 3,130 DNA sequencer (Applied Biosystems) and run.

Data analysis

A. Allele calling and matrix generation

1. Semi-automated scoring was performed on the resulting electronic profiles using GeneMapper v3.7 (Applied Biosystems). The length range of analyzed fragments was from 100 to 500 bp. Full specifications of the parameters used in GeneMapper are summarized in Cara *et al.* (2014). GeneMapper's Report Manager tool was used to retrieve only sample names and allele columns into a Comma Separated Values (.csv) file.

2. Patterns of presence/absence between the *EcoR*I/*Hpa*II and the *EcoR*I/*Msp*I digests were codified from 0 to 3; then, this codification was converted into a binary matrix for either presence (1) or absence (0) of patterns 3, 2 or 1 (Figure 5). Pattern 0 was not codified in the binary matrix because the absence of fragments in *EcoR*I/*Hpa*II and *EcoR*I/*Msp*I digests might be due to either methylation of external cytosines or variations in the nucleotide sequence. The following R script was developed to perform this analysis.

R Script for generating an MSAP binary matrix:

```
##### R Commands used to generate an MSAP binary matrix from a Genemapper
csv file


## The following command makes sure that the working memory is clear
```

```
rm(list = ls(all = TRUE))


## Set working directory (PATH will be specific to where data are)
setwd("C:/...")


## Read in dataset. Input file should contain Samples by rows and Loci
by columns, including header (loci names) and sample names in the first
column
## Each sample digested by HpaII and MspI is identified as 'SampleName_H'
and 'SampleName_M', one below the other.
Data = read.csv("MSAP_Genemapper.csv", header=T, row.names = 1,
check.names=F)


## Create objects
out <- c()
MSAP_patterns <- matrix(nrow = dim(Data)[2], ncol = 0)


## This is a double loop for j hybrids and i loci. It generates a matrix
of MSAP patterns (0, 1, 2 or 3).
for (j in 1:dim(Data)[1]) if (j %% 2 ==1)
{
  for (i in 1:dim(Data)[2])
  {
    out <- append(out, ifelse(Data[j,i]+Data[j+1,i]==0, '0',
              ifelse(Data[j,i]+Data[j+1,i]==2, '1',
              ifelse(Data[j,i]==1, '2', '3'))))
  }
  MSAP_patterns <- cbind(MSAP_patterns, matrix(out, dimnames =
list(colnames(Data), substr(row.names(Data)[j], 1,
nchar(as.character(row.names(Data)[j]))-2))))
  out <- c()
}


##This is a loop that creates a binary matrix of presence/absence of each
pattern (1, 2 or 3) for i loci.
dummies <- matrix(nrow = dim(MSAP_patterns)[2], ncol = 0)
for (i in 1:dim(MSAP_patterns)[1])
{
  dummies <- cbind(dummies, matrix(as.numeric(MSAP_patterns[i,] == 1),
dimnames = list(colnames(MSAP_patterns),
```

```
paste(row.names(MSAP_patterns)[i], "1", sep = "_"))),
matrix(as.numeric(MSAP_patterns[i,] == 2), dimnames =
list(colnames(MSAP_patterns), paste(row.names(MSAP_patterns)[i], "2",
sep = "_"))), matrix(as.numeric(MSAP_patterns[i,] == 3), dimnames =
list(colnames(MSAP_patterns), paste(row.names(MSAP_patterns)[i], "3",
sep = "_"))))
}
binary_matrix <- dummies[,colSums(dummies != 0) != 0]


## Save output to working directory
write.csv(binary_matrix, file = "MSAP_binary_matrix.csv")
```
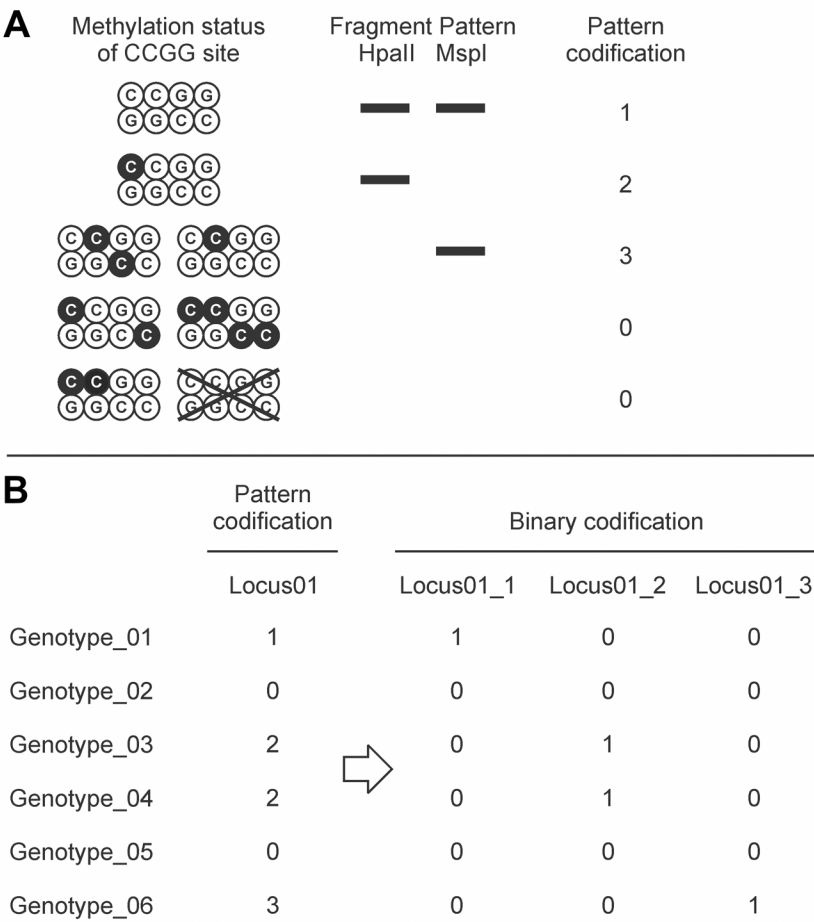

Figure 5. Methylation-sensitive Amplified Polymorphism (MSAP) pattern analysis. A. Inference of methylation status of CCGG restriction sites from the fragment amplification profiles and DNA methylation pattern codification. Black C represents methylated cytosines. B. Binary matrix obtained by codifying the presence (1) or absence (0) of the particular methylation pattern.

B. Determination of species-specific and novel fragments

1. The amplification patterns observed in the evaluated natural hybrids were compared with the patterns obtained in the synthetic hybrids and their parental genotypes. Fragments present in the synthetic hybrids were categorized as *S. microdontum* or *S. kurtzianum* species-specifics. In addition, the presence of novel amplification fragments in the synthetic hybrids (that is present in the hybrids but absent in the genotypes of the parental species *S. microdontum* and *S. kurtzianum*) was assessed in the natural hybrids. Novel fragments were classified as *S. x rechei* species-specific if they were also present in at least one of the *S. x rechei* evaluated genotypes or as novel if they were only observed in the synthetic hybrids. The following R script was developed to perform this analysis.

R Script for analyzing MSAP epiloci inheritance in hybrids:

```
#### R Commands used to analyze MSAP epiloci inheritance in Hybrids

## The following command makes sure that the working memory is clear
rm(list = ls(all = TRUE))

## Set working directory (PATH will be specific to where data are)
setwd("C:/...")

## Read in dataset. Data is organized as Samples by rows and Loci by columns
Data = read.csv("MSAP_binary_matrix.csv", header=T, row.names = 1,
check.names=F)

## Create objects
out <- c()
results <- matrix(nrow = 4, ncol = 0)

## This is a double loop for j hybrids and i loci. Rows are as follows:
1-5 Hybrids, 6 S. kurtzianum (Parental1), 7 S. microdontum (Parental2),
8-15 S. x rechei (Natural hybrid). Any modifications in these numbers should
be applied to the following commands.
for (j in 1:5)
{
  for (i in 1:dim(Data)[2])
  {
    out <- append(out, ifelse(Data[j,i]>0 & Data[6,i]>0 & Data[7,i]==0 &
sum(Data[8:15,i])==0, 'A',    ## Fragments shared only by hybrids and S.
kurtzianum

                      ifelse(Data[j,i]>0 & Data[6,i]==0 & Data[7,i]>0 &
```

```
sum(Data[8:15,i])==0, 'B',    ## Fragments shared only by hybrids and S.
microdontum

                            ifelse(Data[j,i]>0 & Data[6,i]==0 & Data[7,i]==0
& sum(Data[8:15,i])>0, 'C',    ## Fragments shared only by hybrids and S.
x rechei

                            ifelse(Data[j,i]>0 & Data[6,i]==0 & Data[7,i]==0
& sum(Data[8:15,i])==0, 'D', 'E')))))    ## Fragments specific to hybrids
  }

  results <- cbind(results, matrix(table(out, exclude = "E"),dimnames =
list(c("Kurtzianum", "Microdontum", "Rechei", "Hybrid"),
row.names(Data)[j])))
  out <- c()
}

## Transform counts by hybrid into percentages
percentages <- t(results)/colSums(results)*100

## Save output to working directory
write.csv(percentages, file = "MSAP_Hybrids.csv")

## A basic graphical representation
barplot(t(percentages), horiz = TRUE, las=1, col = c("red", "blue",
"green", "yellow"))
```

Recipes

1. Extraction buffer
   100 mM Tris-HCl (pH 8.0)
   20 mM EDTA (pH 8.0)
   1.4 M NaCl
   0.4% (v/v) β-mercaptoethanol
   2% (w/v) CTAB

2. 1x TBE (Tris-Borate-EDTA) gel running buffer
   89 mM Tris base
   89 mM Boric acid
   2 mM EDTA (pH 8.0)

3. 6x DNA loading buffer
   30% (v/v) Glycerol
   0.25% (w/v) Bromophenol blue

0.25% (w/v) Xylene cyanol FF

Notes

Prepare all Master mixes by first adding the larger volumes and then the smaller ones, leaving the enzymes to the end.

Acknowledgments

Competing interests

The authors have not competing interest.

References

1. Camadro, E. L., Erazzu, L. E., Maune, J. F. and Bedogni, M. C. (2012). A genetic approach to the species problem in wild potato. *Plant Biol (Stuttg)* 14(4): 543-554.
2. Cara, N., Ferrer, M. S., Masuelli, R. W., Camadro, E. L. and Marfil, C. F. (2019). Epigenetic consequences of interploidal hybridisation in synthetic and natural interspecific potato hybrids. *New Phytol* 222(4): 1981-1993.
3. Cara, N., Marfil, C. F., García Lampasona, S. C., Masuelli, R. W. (2014). Comparison of two detection systems to reveal AFLP markers in plants. *Botany* 92: 607-610.
4. Vos, P., Hogers, R., Bleeker, M., Reijans, M., van de Lee, T., Hornes, M., Frijters, A., Pot, J., Peleman, J., Kuiper, M. and *et al.* (1995). AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Res* 23(21): 4407-4414.