MARCO on web Manual

English Version 1.0 December 27, 2023

The Research Center for Advanced Analysis, NARO (NAAC)
Institute of Fruit Tree and Tea Science, NARO (NIFTS)
National Agriculture and Food Research Organization (NARO) of JAPAN
Shizuoka University
National University Corporation

Hiroshi FUJII

Shizuoka University

mikan@ml.affrc.go.jp

Contents

1. Overview of MARCO on web

MARCO is a computer program that estimates offspring from parents, one parent from one parent and offspring, and parents from offspring, from polymorphism data obtained by applying co-dominant markers to a large number of breeds. It supports up to tetraploid organisms. The program is written in Python script.

2. Data available for analysis on MARCO on Web

A) Target organisms can be up to tetraploid.

B) Genotypes of codominant markers can be analyzed.

C) Genotypes should be written in ASCII characters.

D) Examples of genotypes are as follows.

String type: 'AB', numeric type such as SSR: '120/200'.

E) The input file format is CSV or xls/xlsx.

F) Data should be entered as 'Data' in the first column of the first line, and the marker name should be written to the right of the 'Data'. For details, please refer to the example input format in this manual.

G) See the sample data linked from the top page as an example input file.

3. Notes on input data preparation

A) Only codominant markers are supported, not dominant markers.

B) Does not distinguish between the letter sequence of genotypes 'AB' and 'BA'.

C) Genotypes are case-sensitive. AB' and 'Ab' are recognized as different genotypes.

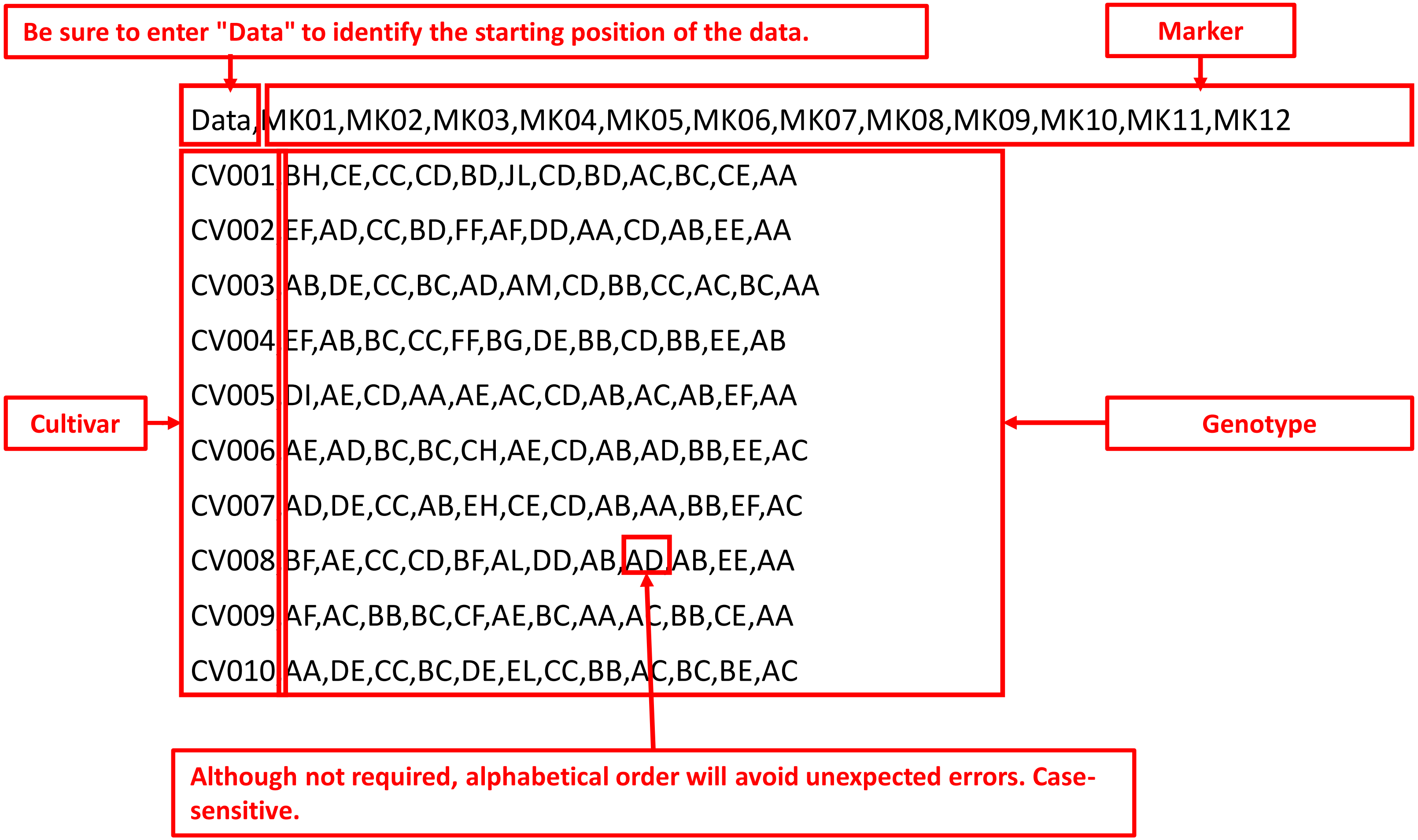
D) Case-sensitive genotype names. Unshu' and 'unshu' are treated as different genotypes. This is also the case for the cultivar names of the parent and offsprings as the arguments of the options described below.

E) Do not include spaces in the cultivar name. For example, 'Golden Delicios' should be followed by an underscore such as 'Golden_Delicious' or ' GoldenDelicios'.

4. Example of input data format:

string type marker, CSV format

Use ASCII characters. No spaces are allowed in the data



Be sure to enter "Data" to identify the starting position of the data.

Marker

```
Data,MK01,MK02,MK03,MK04,MK05,MK06,MK07,MK08,MK09,MK10,MK11,MK12
CV001 BH,CE,CC,CD,BD,JL,CD,BD,AC,BC,CE,AA
CV002 EF,AD,CC,BD,FF,AF,DD,AA,CD,AB,EE,AA
CV003 AB,DE,CC,BC,AD,AM,CD,BB,CC,AC,BC,AA
CV004 EF,AB,BC,CC,FF,BG,DE,BB,CD,BB,EE,AB
CV005 DI,AE,CD,AA,AE,AC,CD,AB,AC,AB,EF,AA
CV006 AE,AD,BC,BC,CH,AE,CD,AB,AD,BB,EE,AC
CV007 AD,DE,CC,AB,EH,CE,CD,AB,AA,BB,EF,AC
CV008 BF,AE,CC,CD,BF,AL,DD,AB,AD,AB,EE,AA
CV009 AF,AC,BB,BC,CF,AE,BC,AA,AC,BB,CE,AA
CV010 AA,DE,CC,BC,DE,EL,CC,BB,AC,BC,BE,AC
```

Cultivar

Genotype

Although not required, alphabetical order will avoid unexpected errors. Case-sensitive.

5. Example of input data format:

string type marker, xls/xlsx format

Use ASCII characters. No spaces are allowed in the data

Be sure to enter "Data" to identify the starting position of the data.

Marker

Data	MK01	MK02	MK03	MK04	MK05	MK06	MK07	MK08	MK09	MK10
CV001	BH	CE	CC	CD	BD	JL	CD	BD	AC	BC
CV002	EF	AD	CC	BD	FF	AF	DD	AA	CD	AB
CV003	AB	DE	CC	BC	AD	AM	CD	BB	CC	AC
CV004	EF	AB	BC	CC	FF	BG	DE	BB	CD	BB
CV005	DI	AE	CD	AA	AE	AC	CD	AB	AC	AB
CV006	AE	AD	BC	BC	CH	AE	CD	AB	AD	BB
CV007	AD	DE	CC	AB	EH	CE	CD	AB	AA	BB
CV008	BF	AE	CC	CD	BF	AL	DD	AB	AD	AB
CV009	AF	AC	BB	BC	CF	AE	BC	AA	AC	BB
CV010	AA	DE	CC	BC	DE	EL	CC	BB	AC	BC

Cultivar

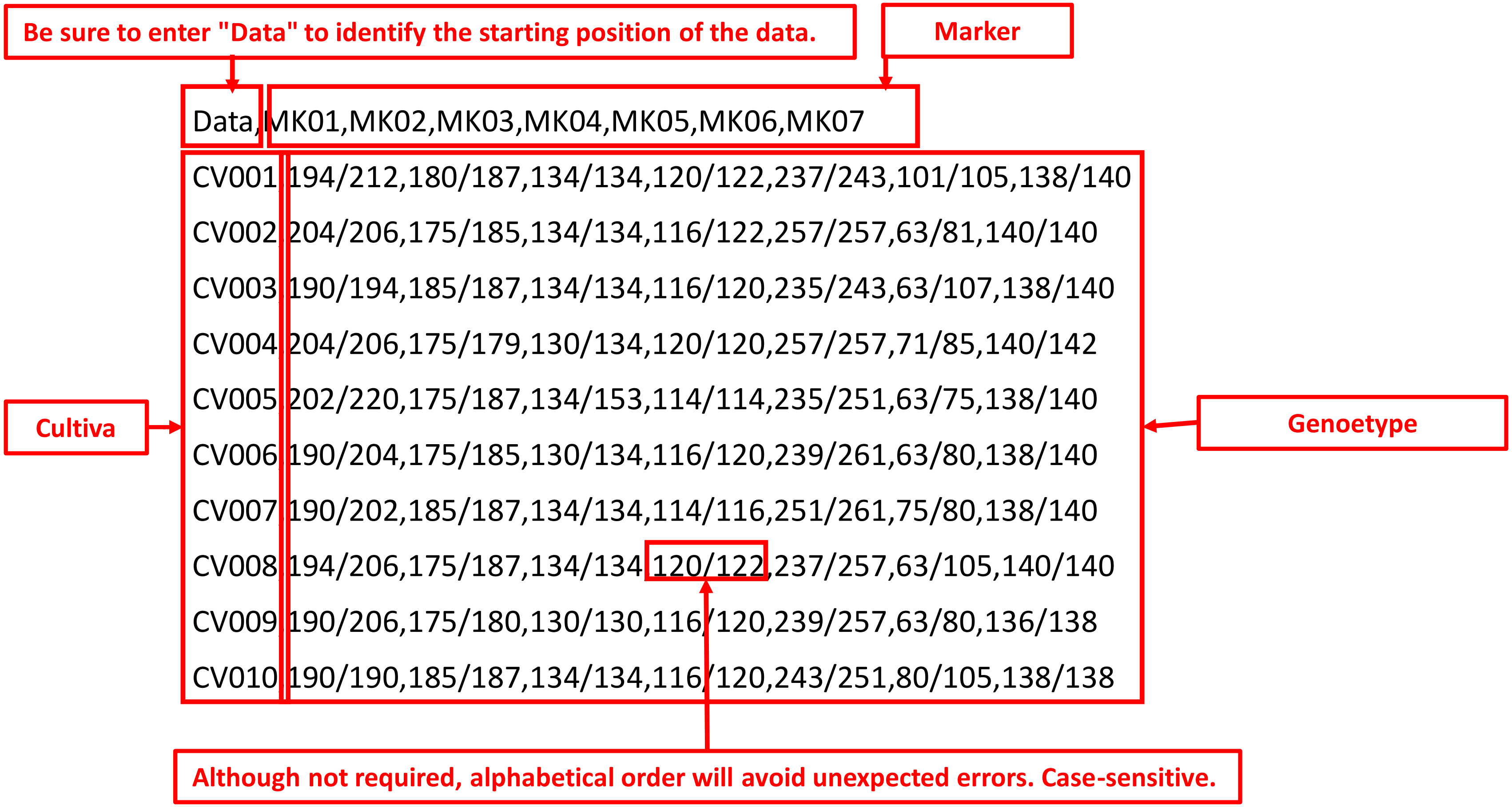
Genotype

Although not required, alphabetical order will avoid unexpected errors. Case-sensitive.

6. Example numerical type marker:

CSV formatof input data format:

Use ASCII characters. Put a "/" (slash) between numbers. No spaces are allowed in the data.



Be sure to enter "Data" to identify the starting position of the data.

Marker

```
Data,MK01,MK02,MK03,MK04,MK05,MK06,MK07
CV001,194/212,180/187,134/134,120/122,237/243,101/105,138/140
CV002,204/206,175/185,134/134,116/122,257/257,63/81,140/140
CV003,190/194,185/187,134/134,116/120,235/243,63/107,138/140
CV004,204/206,175/179,130/134,120/120,257/257,71/85,140/142
CV005,202/220,175/187,134/153,114/114,235/251,63/75,138/140
CV006,190/204,175/185,130/134,116/120,239/261,63/80,138/140
CV007,190/202,185/187,134/134,114/116,251/261,75/80,138/140
CV008,194/206,175/187,134/134,120/122,237/257,63/105,140/140
CV009,190/206,175/180,130/130,116/120,239/257,63/80,136/138
CV010,190/190,185/187,134/134,116/120,243/251,80/105,138/138
```

Cultiva

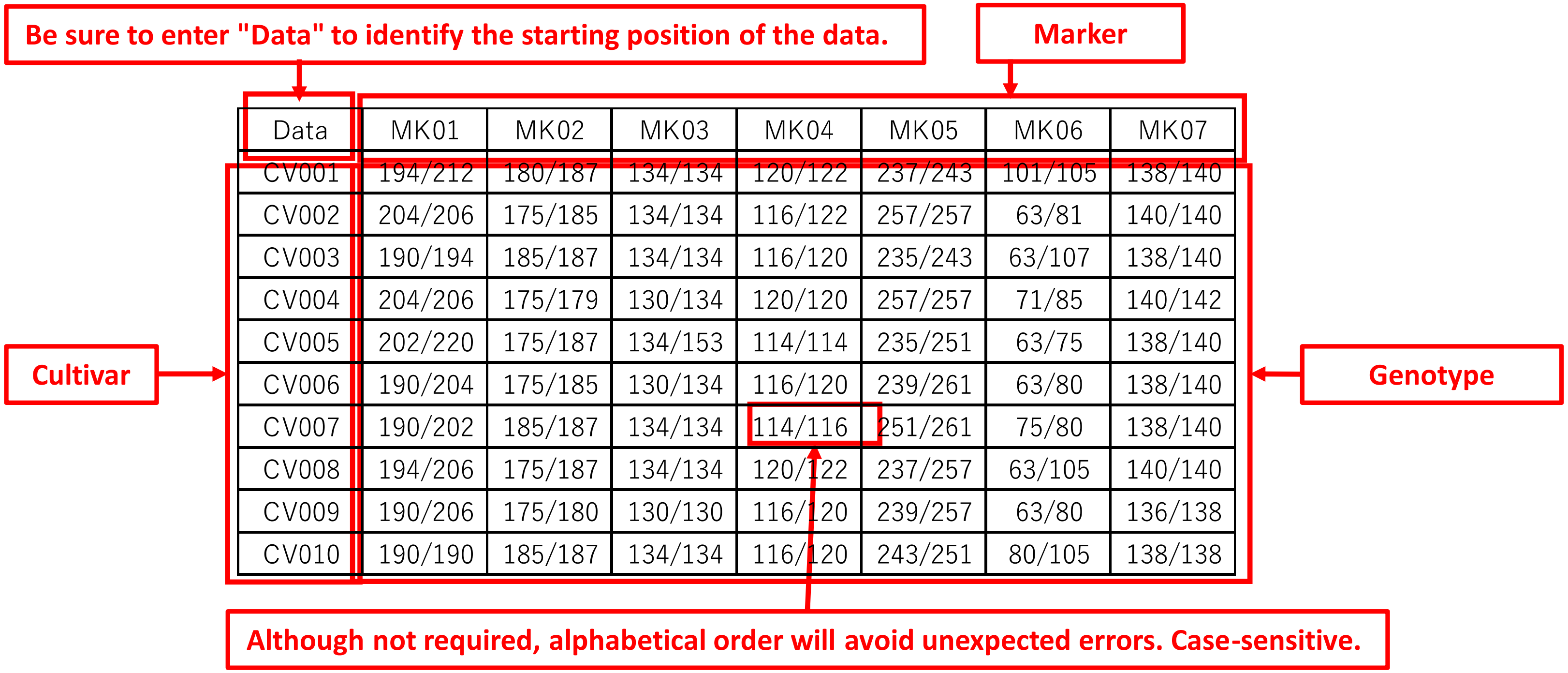
Genoetype

Although not required, alphabetical order will avoid unexpected errors. Case-sensitive.

7. Example of input data format:

numerical type marker, xls/xlsx format

Use ASCII characters. Put a "/" (slash) between numbers. No spaces are allowed in the data. .



Be sure to enter "Data" to identify the starting position of the data.

Marker

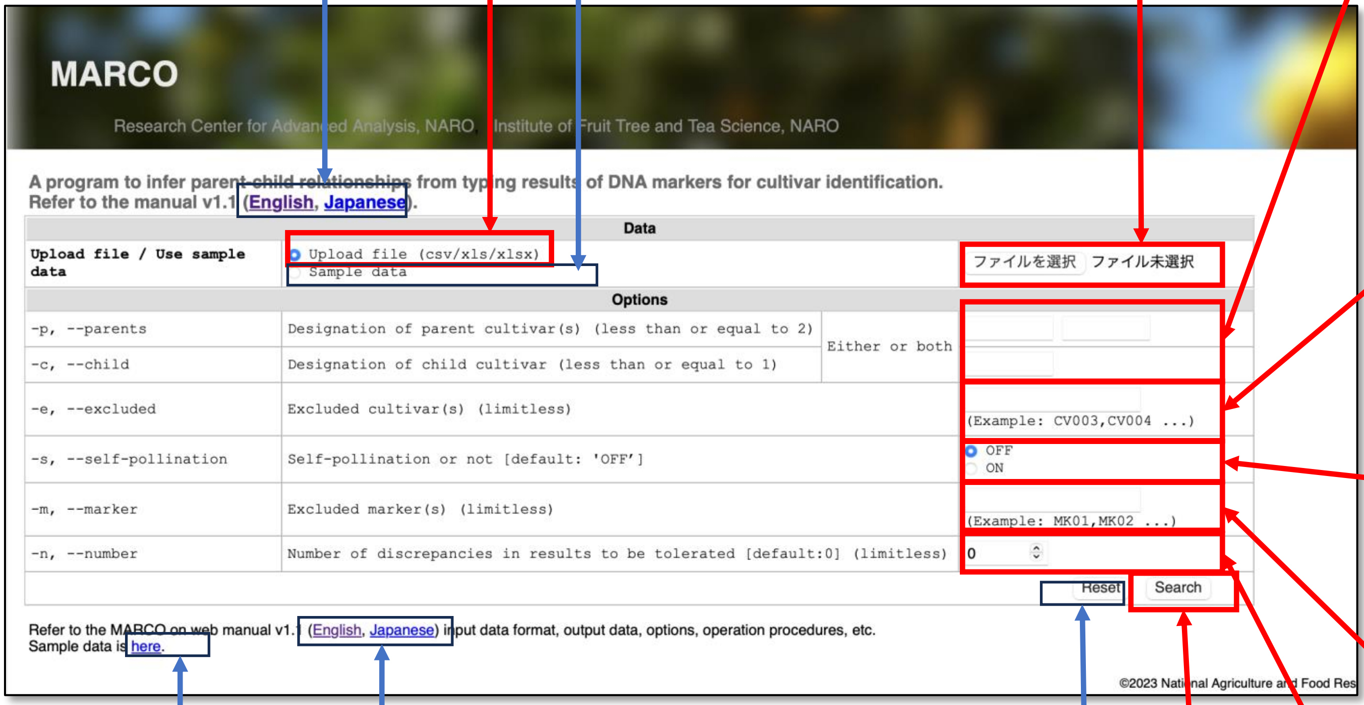
Cultivar

Genotype

Data	MK01	MK02	MK03	MK04	MK05	MK06	MK07
CV001	194/212	180/187	134/134	120/122	237/243	101/105	138/140
CV002	204/206	175/185	134/134	116/122	257/257	63/81	140/140
CV003	190/194	185/187	134/134	116/120	235/243	63/107	138/140
CV004	204/206	175/179	130/134	120/120	257/257	71/85	140/142
CV005	202/220	175/187	134/153	114/114	235/251	63/75	138/140
CV006	190/204	175/185	130/134	116/120	239/261	63/80	138/140
CV007	190/202	185/187	134/134	114/116	251/261	75/80	138/140
CV008	194/206	175/187	134/134	120/122	237/257	63/105	140/140
CV009	190/206	175/180	130/130	116/120	239/257	63/80	136/138
CV010	190/190	185/187	134/134	116/120	243/251	80/105	138/138

Although not required, alphabetical order will avoid unexpected errors. Case-sensitive.

8. Operation procedures



MARCO

Research Center for Advanced Analysis, NARO Institute of Fruit Tree and Tea Science, NARO

A program to infer parent-child relationships from typing results of DNA markers for cultivar identification.
Refer to the manual v1.1 (English, Japanese).

Data			
Upload file / Use sample data	● Upload file (csv/xls/xlsx) ○ Sample data		ファイルを選択 ファイル未選択
Options			
-p, --parents	Designation of parent cultivar(s) (less than or equal to 2)	Either or both	
-c, --child	Designation of child cultivar (less than or equal to 1)		
-e, --excluded	Excluded cultivar(s) (limitless)		(Example: CV003,CV004 ...)
-s, --self-pollination	Self-pollination or not [default: 'OFF']		● OFF ○ ON
-m, --marker	Excluded marker(s) (limitless)		(Example: MK01,MK02 ...)
-n, --number	Number of discrepancies in results to be tolerated [default:0] (limitless)		0

Refer to the MARCO on web manual v1.1 (English, Japanese) input data format, output data, options, operation procedures, etc.
Sample data is here.

1. Select "Upload File" radio button

2. Select upload file

3. Specify parents or offspring
There are three combinations of designations
(1) Find offsprings from both parents
(2) Find one parent from one parent and a offspring
(3) Find parents from foosprings

4. Specify cultivars to be excluded from the analysis:
Specify the name of the cultivar. If there are multiple varieties, separate them with a comma.

5. Self-pollination: ON/OFF

6. Specify markers to be excluded from the analysis: Specify the name of the marker(s). If there are multiple markers, separate them with a comma.

7. Specify the number of contradictions allowed in the parentage search.

8. Reset specified files and options

Manual Download

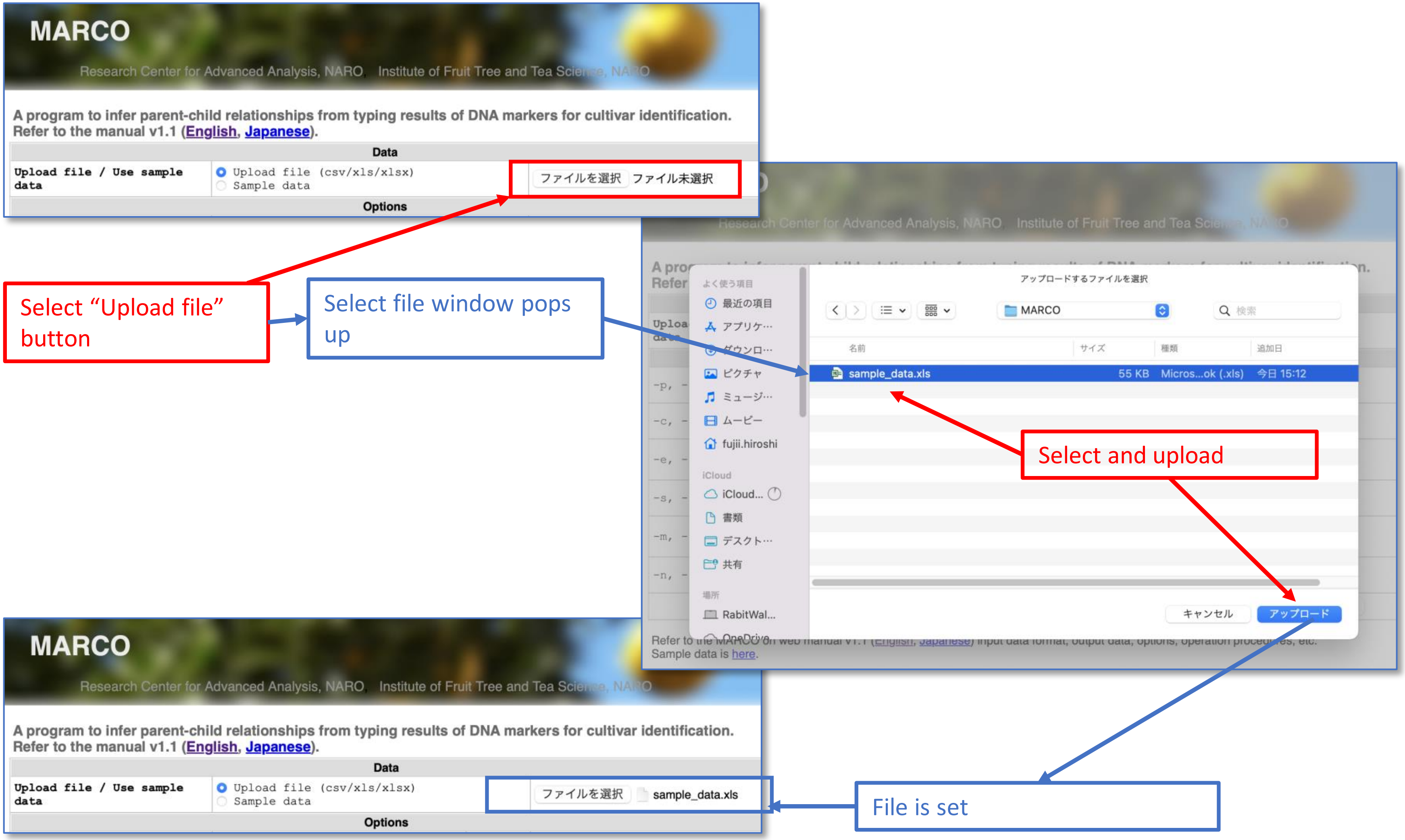
Select sample data

Sample data download

Manual Download

Rest

9. Specify a input file



MARCO

Research Center for Advanced Analysis, NARO Institute of Fruit Tree and Tea Science, NARO

A program to infer parent-child relationships from typing results of DNA markers for cultivar identification.
Refer to the manual v1.1 (English, Japanese).

Data		
Upload file / Use sample data	● Upload file (csv/xls/xlsx) ○ Sample data	ファイルを選択 ファイル未選択
Options		

Select "Upload file" button

Select file window pops up

名前	サイズ	種類	追加日
sample_data.xls	55 KB	Micros...ok (.xls)	今日 15:12

Select and upload

MARCO

Research Center for Advanced Analysis, NARO Institute of Fruit Tree and Tea Science, NARO

A program to infer parent-child relationships from typing results of DNA markers for cultivar identification.
Refer to the manual v1.1 (English, Japanese).

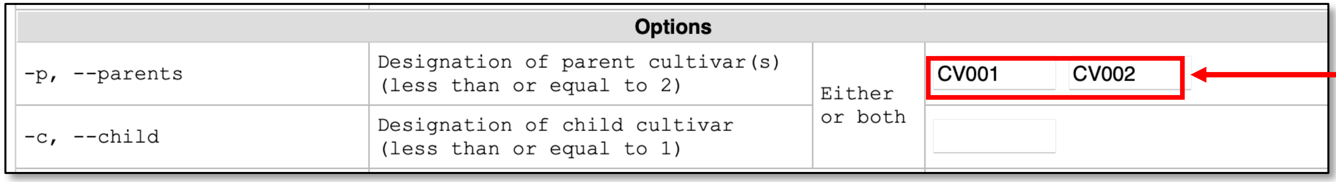
Data		
Upload file / Use sample data	● Upload file (csv/xls/xlsx) ○ Sample data	ファイルを選択 sample_data.xls
Options		

File is set

10. Parent and offspring specification

The cultivar name specified must exactly match the input data.

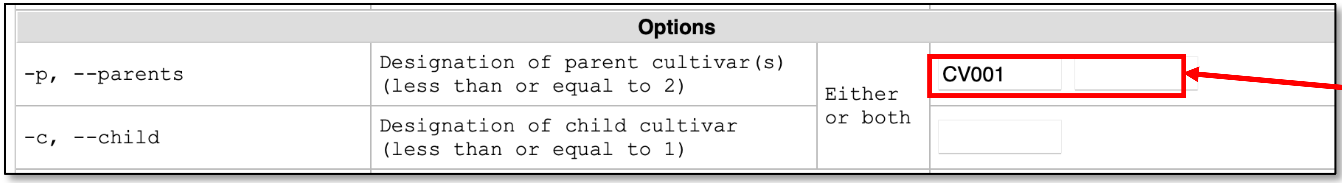
A) When specifying parents and presuming a offspring



Options			
-p, --parents	Designation of parent cultivar(s) (less than or equal to 2)	Either or both	CV001 CV002
-c, --child	Designation of child cultivar (less than or equal to 1)		

Specify the names of the parents listed in the input data

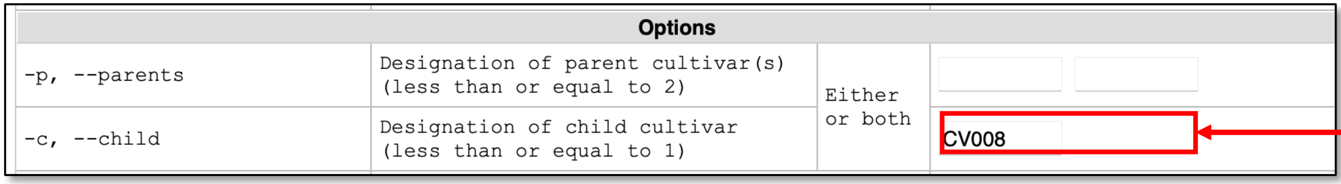
B) When one parent is specified, and one parent and offspring are presumed



Options			
-p, --parents	Designation of parent cultivar(s) (less than or equal to 2)	Either or both	CV001
-c, --child	Designation of child cultivar (less than or equal to 1)		

Specify the name of the single-parent cultivar name listed in the input data.

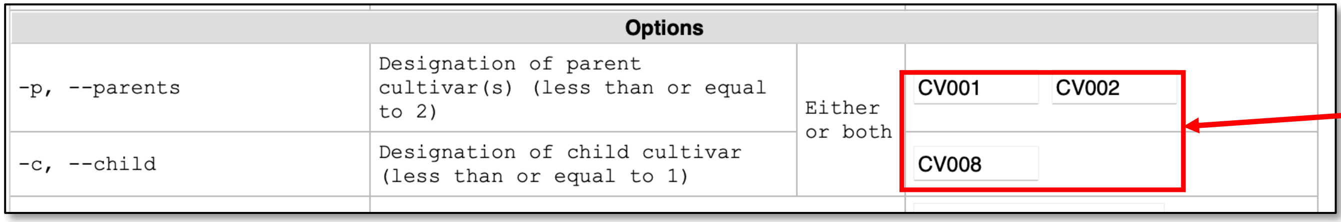
C) When specifying a offspring and presuming parents



Options			
-p, --parents	Designation of parent cultivar(s) (less than or equal to 2)	Either or both	
-c, --child	Designation of child cultivar (less than or equal to 1)		CV008

Specify the name of the offspring cultivar name listed in the input data.

D) When designating parents and offsprings to verify parentage



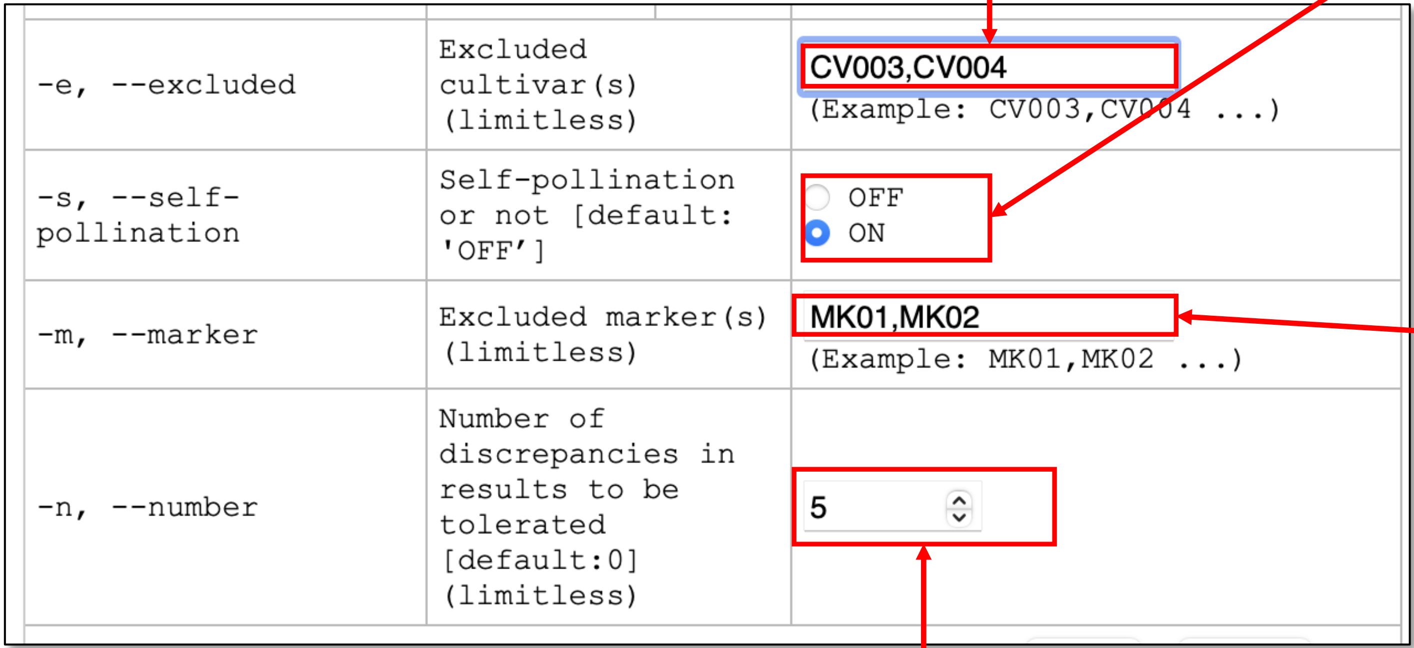
Options			
-p, --parents	Designation of parent cultivar(s) (less than or equal to 2)	Either or both	CV001 CV002
-c, --child	Designation of child cultivar (less than or equal to 1)		CV008

Specify the names of the parents and the offspring breeds listed in the input data

11. Options

Enter cultivars to be excluded from the parent-offspring estimation, separated by commas.

When turned ON, the calculation includes the case of self-pollination.

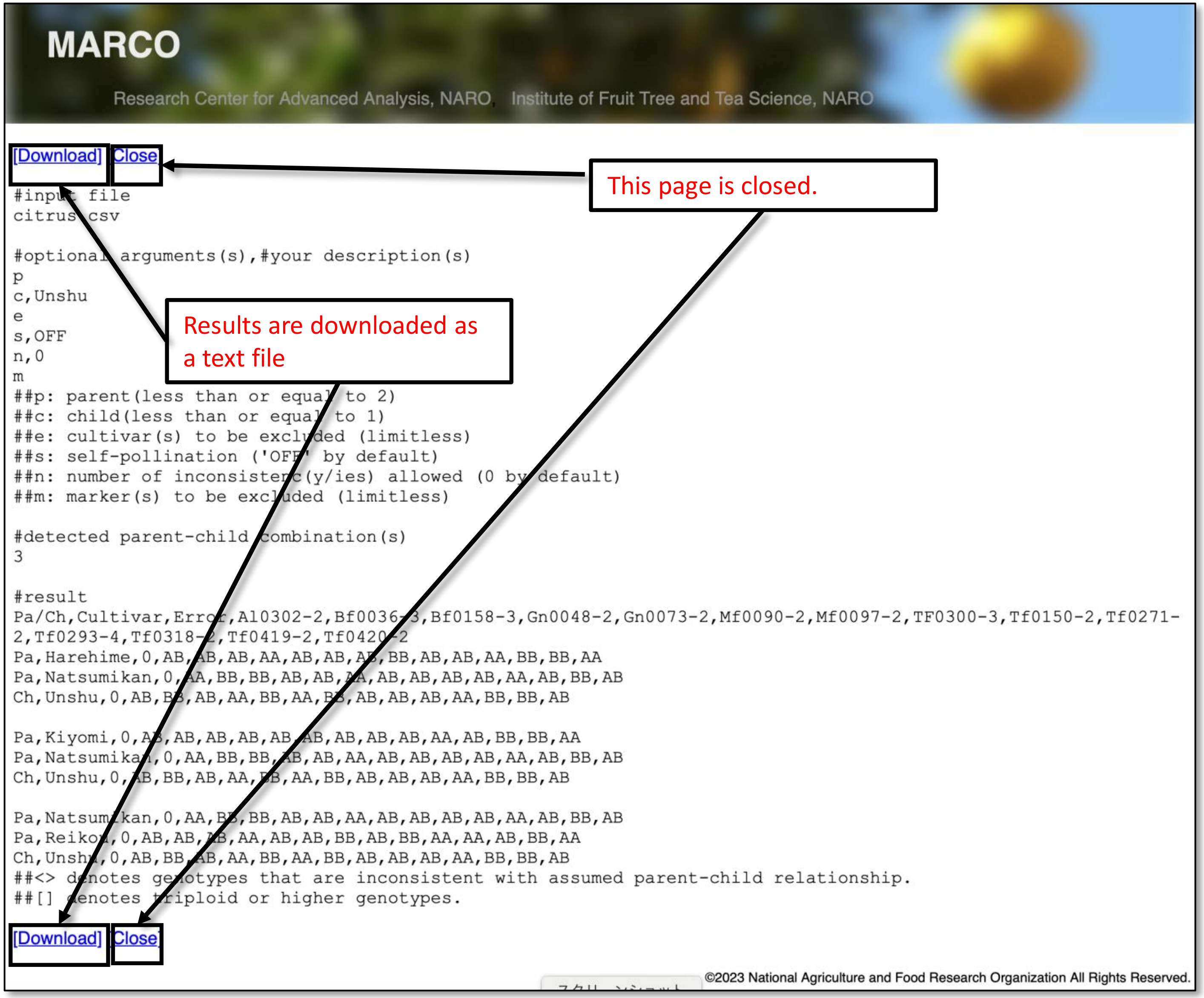


-e, --excluded	Excluded cultivar(s) (limitless)	CV003,CV004 (Example: CV003,CV004 ...)
-s, --self-pollination	Self-pollination or not [default: 'OFF']	○ OFF ● ON
-m, --marker	Excluded marker(s) (limitless)	MK01,MK02 (Example: MK01,MK02 ...)
-n, --number	Number of discrepancies in results to be tolerated [default:0] (limitless)	5

Enter the markers you wish to exclude from the parent- estimation, separated by commas. For example, if a marker shows a contradiction in a known parentage, you can assume that there is a problem with the marker and exclude it from the calculation.

Even a correct parentage can be inconsistent if there are problems with the markers. This option allows a specified number of contradictions to be allowed in order to estimate the parentage.

12. Example of output upon successful completion



MARCO

Research Center for Advanced Analysis, NARO Institute of Fruit Tree and Tea Science, NARO

This page is closed.

Results are downloaded as a text file

```
#input file
citrus.csv

#optional arguments(s),#your description(s)
p
c,Unshu
e
s,OFF
n,0
m
##p: parent(less than or equal to 2)
##c: child(less than or equal to 1)
##e: cultivar(s) to be excluded (limitless)
##s: self-pollination ('OFF' by default)
##n: number of inconsistenc(y/ies) allowed (0 by default)
##m: marker(s) to be excluded (limitless)

#detected parent-child combination(s)
3

#result
Pa/Ch,Cultivar,Error,Al0302-2,Bf0036-3,Bf0158-3,Gn0048-2,Gn0073-2,Mf0090-2,Mf0097-2,TF0300-3,Tf0150-2,Tf0271-
2,Tf0293-4,Tf0318-2,Tf0419-2,Tf0420-2
Pa,Harehime,0,AB,AB,AB,AA,AB,AB,AB,BB,AB,AB,AA,BB,BB,AA
Pa,Natsumikan,0,AA,BB,BB,AB,AB,AA,AB,AB,AB,AB,AA,AB,BB,AB
Ch,Unshu,0,AB,BB,AB,AA,BB,AA,BB,AB,AB,AB,AA,BB,BB,AB

Pa,Kiyomi,0,AB,AB,AB,AB,AB,AB,AB,AB,AB,AA,AB,BB,BB,AA
Pa,Natsumikan,0,AA,BB,BB,AB,AB,AA,AB,AB,AB,AB,AA,AB,BB,AB
Ch,Unshu,0,AB,BB,AB,AA,BB,AA,BB,AB,AB,AB,AA,BB,BB,AB

Pa,Natsumikan,0,AA,BB,BB,AB,AB,AA,AB,AB,AB,AB,AA,AB,BB,AB
Pa,Reikou,0,AB,AB,AB,AA,AB,AB,BB,AB,BB,AA,AA,AB,BB,AA
Ch,Unshu,0,AB,BB,AB,AA,BB,AA,BB,AB,AB,AB,AA,BB,BB,AB
##<> denotes genotypes that are inconsistent with assumed parent-child relationship.
##[] denotes triploid or higher genotypes.
```

13. Output description 1

Example of checking the parentage by specifying the parents as CV001 and CV002 and the offspring as CV008 with default options

```
#input file
marco_20231212104729.csv
#optional arguments(s),#your description(s)
p,CV001,CV002
c,CV008
e
s,OFF
n,10
m
##p: parent(less than or equal to 2)
##c: child(less than or equal to 1)
##e: cultivar(s) to be excluded (limitless)
##s: self-pollination ('OFF' by default)
##n: number of inconsistenc(y/ies) allowed (0 by default)
##m: marker(s) to be excluded (limitless)

#detected parent-child combination(s)
1
#result
Pa/Ch,Cultivar,Error,MK01,MK02,MK03,MK04,MK05,MK06,MK07,MK08,MK09,MK10,MK11,MK12,MK13,MK14,
MK15,MK16,MK17,MK18,MK19,MK20,MK21,MK22
Pa,CV001,0,BH,CE,CC,CD,BD,JL,CD,BD,AC,BC,CE,AA,EE,CC,AB,AA,CD,AD,BD,CC,BE,AE
Pa,CV002,0,EF,AD,CC,BD,FF,AF,DD,AA,CD,AB,EE,AA,BD,CC,AA,AA,CC,AB,AE,CD,BD,AE
Ch,CV008,0,BF,AE,CC,CD,BF,AL,DD,AB,AD,AB,EE,AA,DE,CC,AB,AA,CC,BD,DE,CC,BB,AE
##<> denotes genotypes that are inconsistent with assumed parent-child relationship.
##[] denotes triploid or higher genotypes.
```

Input file name

Specified parents and offspring

Specified options

Option description

Result

Number of parentage detected

Symbol description

14. Output description 2

All parentage that satisfy the condition are output, so more than one parent-offspring combination may be output.

Result 1

```
#result
Pa/Ch,Cultivar,Error,MK01,MK02,MK03,MK04,MK05,MK06,MK07,MK08,MK09,MK10,MK11,MK
12,MK13,MK14,MK15,MK16,MK17,MK18,MK19
Pa,CV008,0,BF,AE,CC,CD,BF,AL,DD,AB,AD,AB,EE,AA,DE,CC,AB,AA,CC,BD,DE
Pa,CV013,0,AA,DE,CC,BC,DH,AC,CC,AA,AC,BC,EF,AC,BD,BC,AB,DD,DD,BC,AD
Ch,_Smp02,0,AB,AE,CC,CD,BD,[ACL],CD,AA,AD,BC,EF,AC,[BDE],CC,AB,AD,CD,BC,AE
```

Result 2

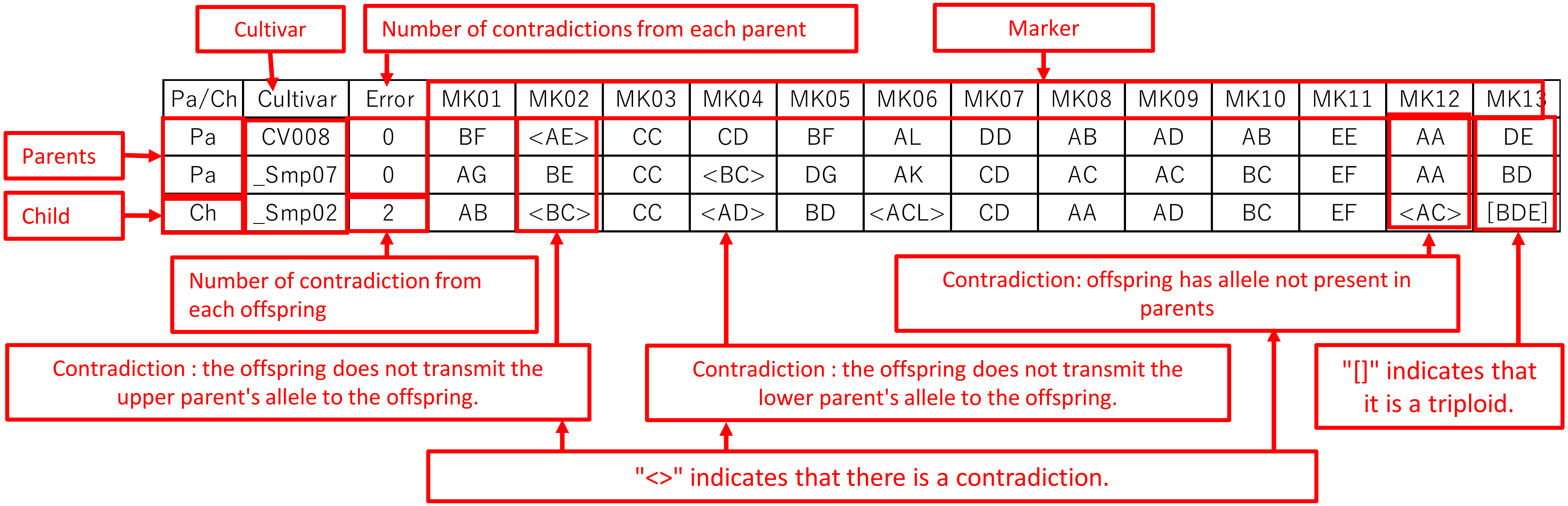
```
Pa,CV008,0,BF,AE,CC,CD,BF,AL,DD,AB,AD,AB,EE,AA,DE,CC,AB,AA,CC,BD,DE
Pa,_Smp07,0,AG,BE,CC,BC,DG,AK,CD,AC,AC,BC,EF,AA,BD,BC,AB,AD,CD,BC,AD
Ch,_Smp02,2,AB,AE,CC,CD,BD,<ACL>,CD,AA,AD,BC,EF,<AC>,[BDE],CC,AB,AD,CD,BC,AE
```

15. Output description 3

Result

```
#result
Pa/Ch,Cultivar,Error,MK01,MK02,MK03,MK04,MK05,MK06,MK07,MK08,MK09,MK10,MK11,MK12,MK13,MK1
4,MK15,MK16
Pa,CV008,0,BF,AE,CC,CD,BF,AL,DD,AB,AD,AB,EE,AA,DE
Pa,_Smp07,0,AG,BE,CC,BC,DG,AK,CD,AC,AC,BC,EF,AA,BD
Ch,_Smp02,2,AB,AE,CC,CD,BD,<ACL>,CD,AA,AD,BC,EF,<AC>,[BDE]
```

Output results converted to CSV and displayed in Excel.



Pa/Ch	Cultivar	Error	MK01	MK02	MK03	MK04	MK05	MK06	MK07	MK08	MK09	MK10	MK11	MK12	MK13
Pa	CV008	0	BF	<AE>	CC	CD	BF	AL	DD	AB	AD	AB	EE	AA	DE
Pa	_Smp07	0	AG	BE	CC	<BC>	DG	AK	CD	AC	AC	BC	EF	AA	BD
Ch	_Smp02	2	AB	<BC>	CC	<AD>	BD	<ACL>	CD	AA	AD	BC	EF	<AC>	[BDE]

Cultivar

Number of contradictions from each parent

Marker

Parents

Child

Number of contradiction from each offspring

Contradiction: offspring has allele not present in parents

Contradiction : the offspring does not transmit the upper parent's allele to the offspring.

Contradiction : the offspring does not transmit the lower parent's allele to the offspring.

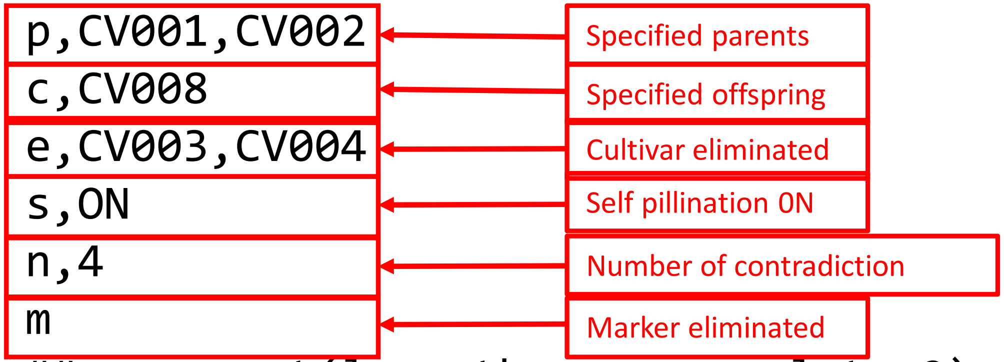
"[]" indicates that it is a triploid.

"<>" indicates that there is a contradiction.

“AB”, “BC”, etc. in the table indicate genotypes

16. Output description 4

Option description



```
#optional arguments(s),#your description(s)
p,CV001,CV002
c,CV008
e,CV003,CV004
s,ON
n,4
m
##p: parent(less than or equal to 2)
##c: child(less than or equal to 1)
##e: cultivar(s) to be excluded (limitless)
##s: self-pollination ('OFF' by default)
##n: number of inconsistenc(y/ies) allowed (0
by default)
##m: marker(s) to be excluded (limitless)
```

Specified parents

Specified offspring

Cultivar eliminated

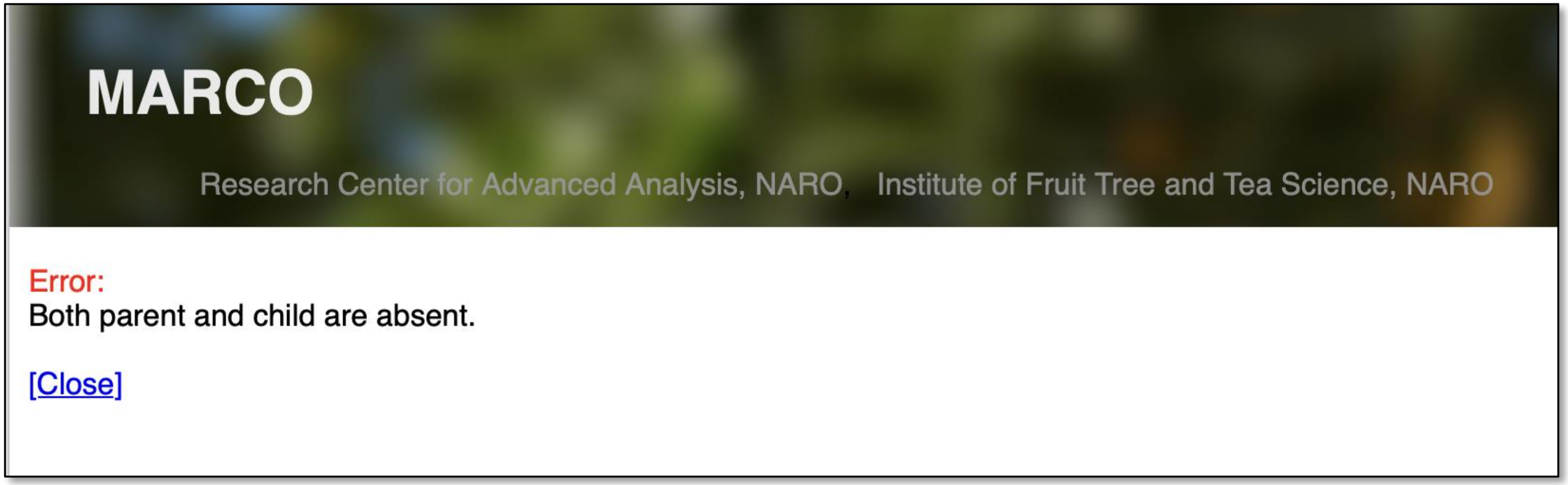
Self pillination ON

Number of contradiction

Marker eliminated

17. Error indication 1

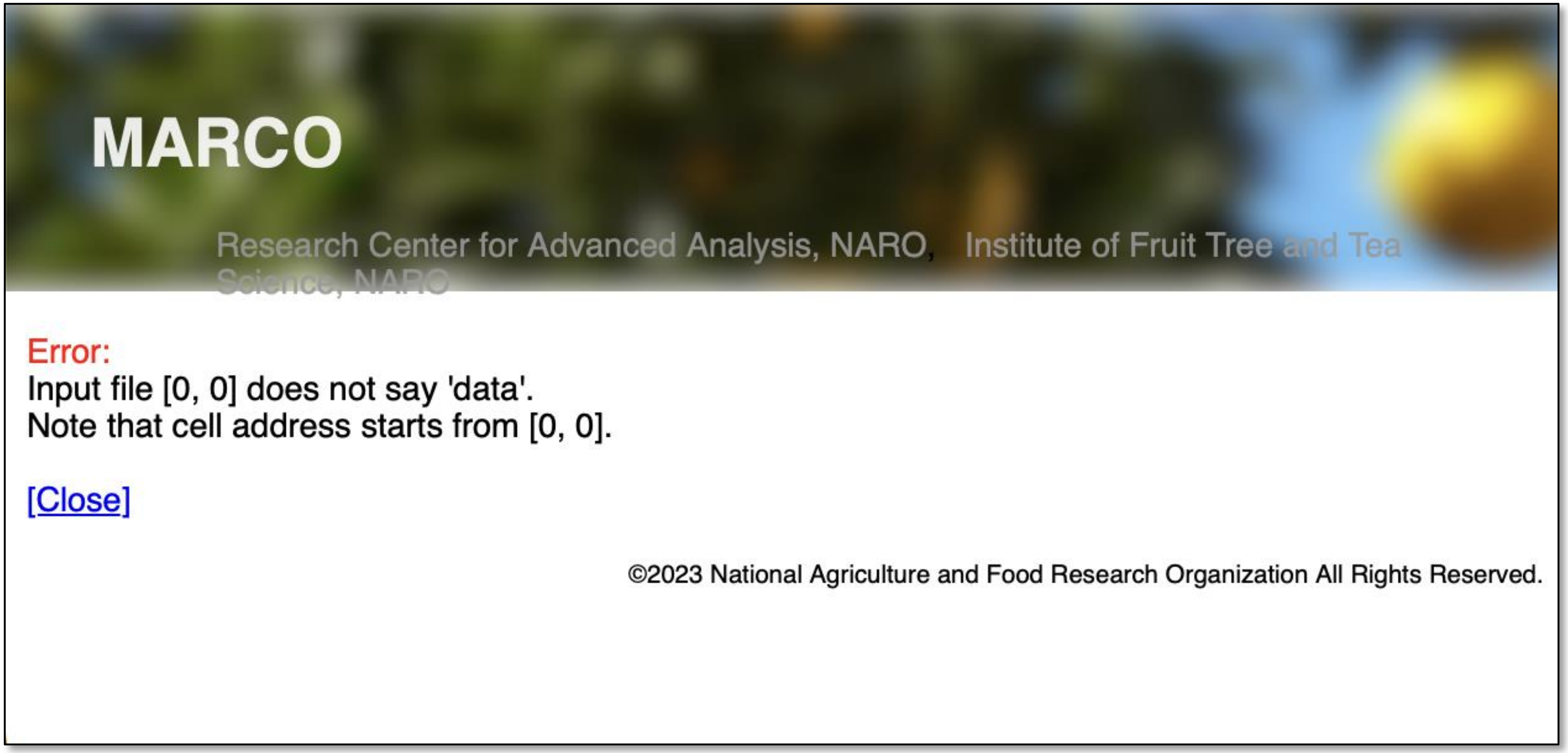
If neither parents nor offspring were specified



Error:
Both parent and child are absent.

18. Error indication 2

If the input data format is incorrect

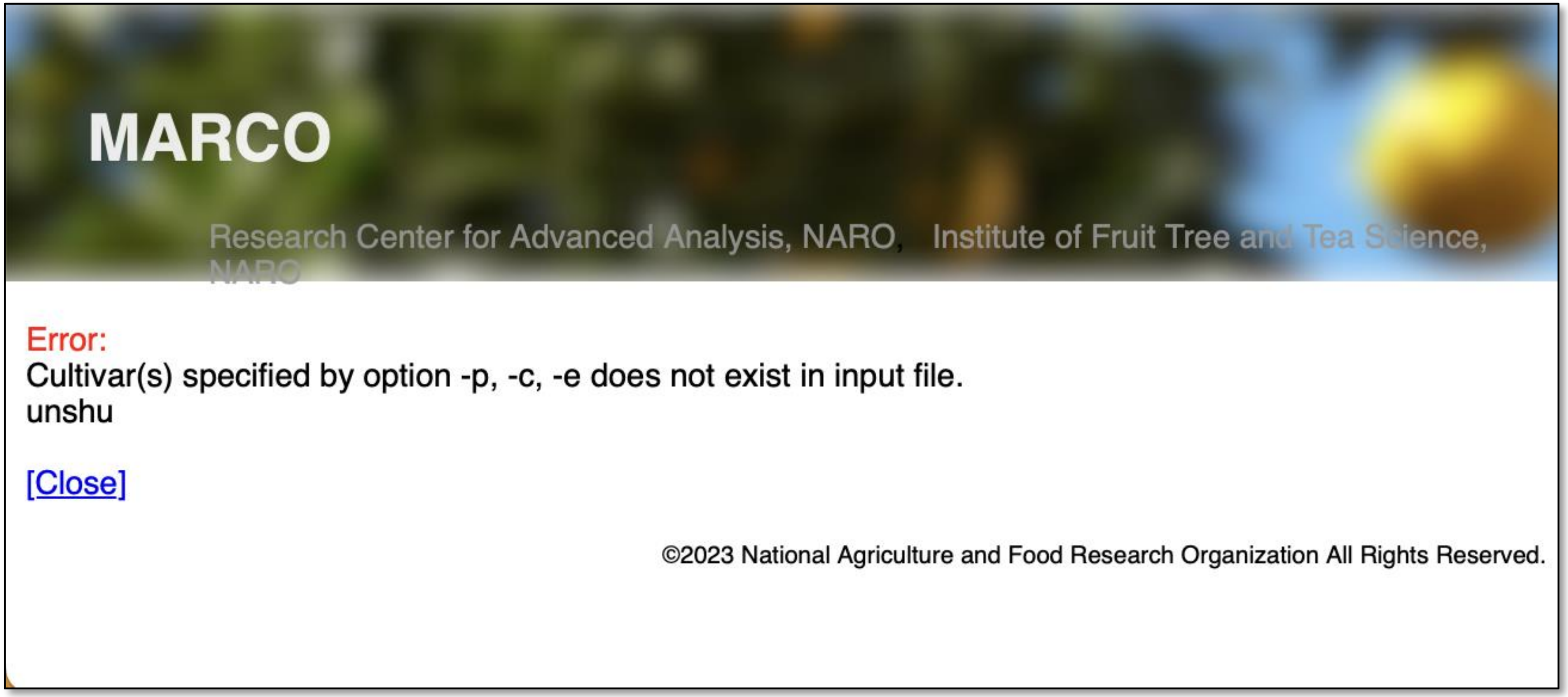


Error:
Input file [0, 0] does not say 'data'.
Note that cell address starts from [0, 0].

Please correct the input data by referring to the input format.

19. Error indication 3

If the parent or offspring cultivar name entered in the options does not match the cultivar name in the input data.



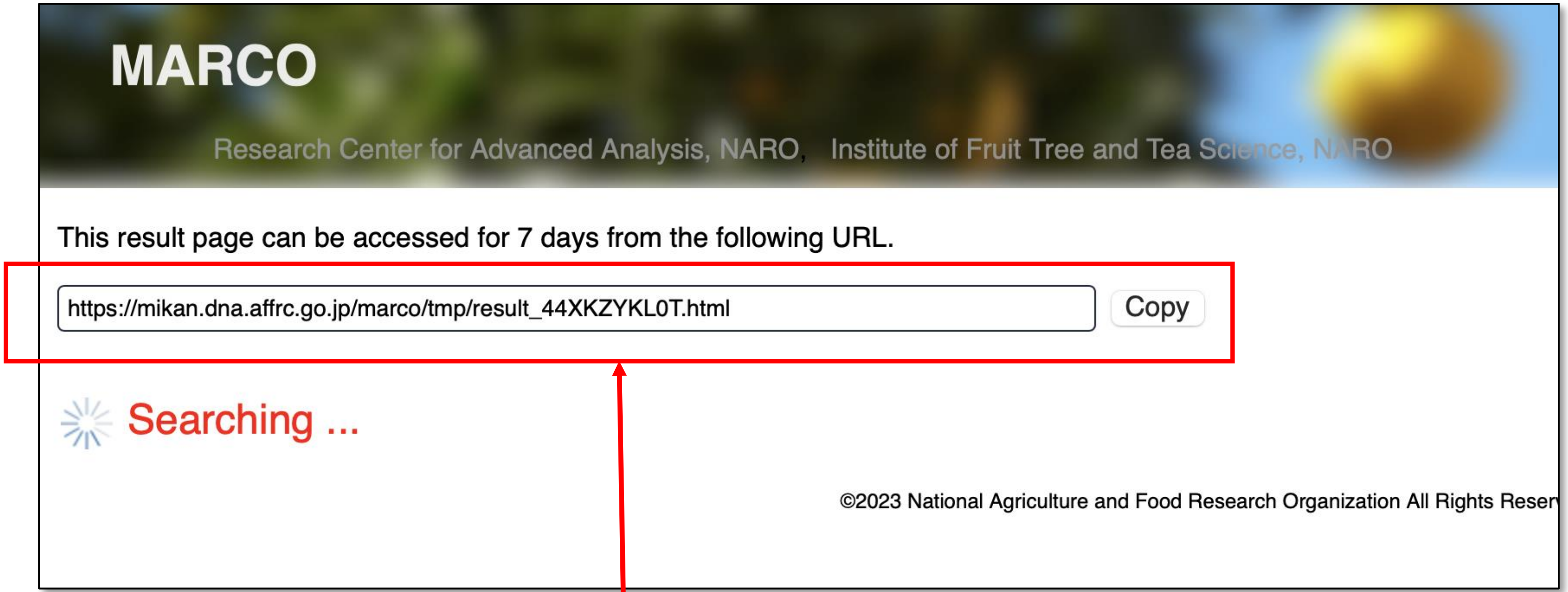
Error:
Cultivar(s) specified by option -p, -c, -e does not exist in input file.
unshu

[Close]

©2023 National Agriculture and Food Research Organization All Rights Reserved.

Enter the name of the cultivar exactly as it appears in the input data.

20. When "Searching" does not end



MARCO

Research Center for Advanced Analysis, NARO, Institute of Fruit Tree and Tea Science, NARO

This result page can be accessed for 7 days from the following URL.

https://mikan.dna.affrc.go.jp/marco/tmp/result_44XKZYKL0T.html

Searching ...

©2023 National Agriculture and Food Research Organization All Rights Reser

Click the "Copy" button on the right to copy the URL to your clipboard. Accessing this URL will display the calculation results. Calculation results are stored for 7 days. You can view the calculation results even if you close your browser, shut down your PC, or from another terminal.

21. Citation

When publishing an academic paper using the results obtained using the MARCO on web, please describe the use of MARCO in the paper and cite the following literature..

- MiGD

  KAWAHARA Yoshihiro, ENDO Tomoko, OMURA Mitsuo, TERAMOTO Yumiko, ITHO Tsuyoshi, FUJII Hiroshi, SHIMADA Takehiko (2020) Mikan Genome Database (MiGD): integrated database of the genome annotation, genomic diversity, and CAPS marker information for mandarin molecular breeding. Breeding Science.

  doi: 10.1270/jsbbs.19097

22. Copyright

The computer program MARCO is protected by copyright. Therefore, except as expressly permitted, no portion of this software may be distributed or reproduced by any means, or in any form, without NARO’s prior written permission.

COPYRIGHT NOTICE

23. Disclaimer

MARCO on web provided by National Agriculture and Food Research Organization (NARO) is available for use subject to the following disclaimer.

・MARCO on web makes no warranty of any kind regarding the output results.

・The author will try to find and fix bugs in the program on MARCO on web. But is not responsible for any kind of damage or loss caused by the use or misuse of MARCO on web.

・The author is under no obligation to provide support, service, fixes, or upgrades for MARCO on web.

24. Contact

mikan@ml.affrc.go.jp

25. History of manual

English version 1.0	December 27, 2023