

GDR Breeders Toolbox Tutorial

Overview.....	3
Using the Toolbox.....	4
1. Breeders Toolbox starting page	4
2. Browse the Database.....	7
2.1 Browse varieties by datasets	7
2.2 Result of browsing varieties by datasets	8
3. Variety Page.....	9
4. Phenotyping Data.....	11
4.1 Select phenotype characteristics	11
4.2 Download phenotypic evaluation data.....	12
4.3 Format of phenotyping data in an Excel file.....	13
5. Search Phenotyping Data.....	15
5.1a. Search by Variety Name	15
5.1b. Results of Search by Variety Name	18
5.2a. Search by Traits.....	19
5.2b. Results of Search by Traits	21
5.3a. Search by Parentage	22
5.3b. Results of Search by Parentage	23
6. Genotyping Data	24
6.1 Download genotypic evaluation data.....	24
6.2 Format of genotyping data in Excel.....	25
7. Search Genotyping Data	26
7.1a Search by Variety Name	26
7.1b Results of Search by Variety Name	27
7.2a Search by Variety / Marker	28
7.2b Results of Search by Variety / Marker.....	30
7.3 Search by Marker / Allele	31
7.4 Marker / Allele page	32
8. Pedimap.....	35
8.1 Generation of Pedimap input file	35
8.2 Generating the Pedimap input file.....	36

8.2a Select phenotypes.....	38
9. Cross Assist.....	42
9.1 Front Page	43
9.2 Targets Page - Input.....	46
9.3 Cross Efficiency Results Page - Output.....	49

Overview

The Rosaceae "Breeder's Toolbox" is designed to house and integrate private and public phenotypic and genotypic data for apple, peach, cherry and strawberry so breeders can more accurately and efficiently assess parent and progeny performance and selection. This should enable a breeder to make more rapid and accurate progress in variety development, plus facilitate marker-trait associations through use of Pedimap input files from selected data

We provide data templates in excel to acquire phenotypic/genotypic data from breeders and associated scientists. The breeding data, then, are organized and stored in a Chado database. Chado is a relational database schema, developed under the Generic Model Organism Database (GMOD) project (<http://gmod.org>). It is a common and a well-designed schema for storing biological data and has recently been expanded to house natural diversity data.

Not all breeding data are publically available as private breeding program data should not be shared except with permission of the breeder. Therefore, it is necessary to keep certain portions of the breeding data private and create a system to control the access to public/private component of the breeding data. We have incorporated Drupal's account system and "*organic group*" module to control the access of breeding data. With "*organic group*", we are able to create a group and assign an account holder in Drupal as a member of the group. We create a group for each set of private data and the members of the group are given the permission to access the data. A Drupal account holder may become a member of one or more breeding groups. Those who do not have a Drupal account (anonymous user) become "*public users*" and they are able to access merely public portion of breeding data in the database.

Using the Toolbox

1. Breeders Toolbox starting page

Breeders Toolbox

RosBREED Crop Reference Set available [Breeders Toolbox Tutorial](#)

Choose Crop

Set Crop ① Apple

Choose Breeding Group

Set Breeding Group ② APPLE CRS

Browse Database

Browse Varieties ③

- Browse Varieties by Datasets

Search Database

Search Phenotyping Data ④

- Search by Varieties
- Search by Traits
- Search by Parentage

Search Genotyping Data ⑤

- Search by Varieties
- Search by Variety/Marker
- Search by Marker/Allele

Documentation

Descriptors ⑥

- View Descriptors
- Download Descriptors

Tutorials ⑦

- Tutorials for Public Users
- Tutorials for Group Members *(coming soon)*

Guides

- Setup Guide *(coming soon)*

Breeding Data Template ⑧

- Download Sample Template

Tools

Generate Input Files ⑨

- Input File for Pedimap
- Input Files for FlexQTL

Cross Planning Tool ⑩

- Cross Assist

Seedling Selection Tool

- Calculate Selection Recommendations

Fig. 1.3 Breeders Toolbox starting page

1. Crop selection

Select a crop from the dropdown. The dropdown lists all currently available crops.

2. Breeding group selection

Select a breeding group from the dropdown. The dropdown lists all breeding groups that the current user belongs to. The dropdown appears only if the current user belongs to more than one breeding group.

3. Browse Database (See Section 2)

Users can browse varieties by datasets to get their phenotypic evaluation data.

4. Search Phenotype Database (See Section 5)

Users can search varieties by variety names, traits and parentages to obtain their phenotypic evaluation data.

5. Search Genotype Database (See Section 7)

Users can search varieties by variety names, Variety/Marker and Marker/Allele to obtain their genotypic evaluation data.

6. Phenotypic descriptors

Users can view definitions of phenotypic descriptors of the selected database online (Fig. 1.4). Users can also download the phenotypic descriptors of database in an excel file.



Phenotype	Definition
BITTERPIT	bitterpit presense
BLUSTRICOL	Type of red color
CALYXO	degree of opening of fruit calyx,average of 5 fruit
COREO	core opening around seeds in equatorial slice
CRACK	cracking of fruit
DIAM	Diameter of fruit at widest point (inches)
GREASE	tackiness or greasiness of skin
GRNDCOL	Color of skin before/under blush/ripe
INTBROWN	internal browning not due to bruises
MOLDYCORE	moldy core
OVRCOL	Color
PERCOVRCOL	% red/overcolor color of skin
PERCRUSS	Amount of russet
RUSSET_LOCATION	location of russetting on apple skin
RUSSLOC1	presence/absence of stem cavity russet
RUSSLOC2	presence/absence of shoulder russet

Fig. 1.4 Definitions of phenotypic descriptors

7. Tutorials

Users can download this tutorial.

8. Breeding Data Template

Users can download a sample of Breeding Data Template in an excel format. "Info" sheet at the first sheet of the excel explains data sheets. Example data colored in red are provided for each data sheet.

9. Generate input files for Pedimap (See Section 8)

Users can generate an input file for the Pedimap software. Pedimap is designed to graphically present genetic information in pedigrees.

10. Cross Assist (See Section 9)

Cross Assist is a decision-support tool for breeders to plan crosses. User can use this tool for their decision making with crosses.

2. Browse the Database

Select a dataset to access the breeding data. From the search results, you can access the record of the individual varieties in the dataset or download the evaluation data of the entire dataset.

2.1 Browse varieties by datasets

Breeding Datasets Choose a destination... ▾

Select a dataset to view the breeding data. From the search results, you can access the record of the individual variety in the dataset or download the evaluation data of the entire dataset.

① **Choose Crop** ▾

② **Choose Breeding Group** ▾

③

Dataset	Description
Sweet_Cherry_CRS_phenotyping_2011	MSU
	Pear Acres
	Roza
Sweet_Cherry_CRS_phenotyping_2010	Pear Acres
	Roza

Fig. 2.1 Breeding datasets

1. Crop selection

Select a crop from the dropdown. The dropdown lists all currently available crops.

2. Breeding group selection

Select a breeding group from the dropdown. The dropdown lists all breeding groups that the current user belongs to. The dropdown appears only if the current user belongs to more than one breeding group.

3. Datasets table

The table lists all the datasets in the selected database. Click a description of a dataset to view all varieties in the dataset (See Section 2.2).

2.2 Result of browsing varieties by datasets

The screenshot shows a web interface with a yellow background. At the top right, there is a link: "Go Back to Browse Datasets with Sites".

① The dataset you selected returned data from '494' varieties.
Click the variety name to access the individual record.

Click "Select phenotypic characters" to bring up a screen from which to select the traits or variables which you would like to have downloaded.

Click on "Download search results" to start downloading a file containing the results of your variety and trait selection request.

After downloading one set of search results, you can return to this page to select different sets of phenotypic characters to download.

②

0407-0001	0407-0005	0407-0006	0407-0007
0407-0008	0408-0002	0408-0003	0408-0004
0408-0005	0408-0007	0408-0008	0408-0010
0408-0011	0408-0014	0408-0015	0408-0019

③ Select phenotypic characters

④ Download search results

⑤ Download help

1 2 3 4 5 6 7 8 9 ... next › last »

Fig. 2.2 Variety table

1. Summary of the selected dataset

It shows the number of varieties in the dataset.

2. Results table

The table lists all varieties in the selected dataset. The variety names are linked to the variety page to access the individual records (see Section 3).

3. Select phenotypic characteristics

See section 4.1 "*Select phenotype characteristics*"

4. Download the result in an Excel file

See section 4.2 "*Download phenotypic evaluation data*"

5. Download help

Click the "*Download help*" button to view download help

3. Variety Page

The "Variety page" shows the detailed information about an individual variety.

The screenshot displays the variety page for AE0211-036. It features four numbered callouts: 1. A table with variety information: Alias (n/a), Mother (GingerGold), Father (n/a), and Cross Number (n/a). 2. Evaluation data section with download links for Phenotyping Data (phenotypic characters) and Genotyping Data. 3. Associated Markers / Alleles table listing CH05c06 (120), Hi04e04 (228, 232), and Md-Exp7 (202, 212). 4. A photo of three apples, with the text "image not available" below it.

AE0211-036	
Alias	n/a
Mother	GingerGold
Father	n/a
Cross Number	n/a

①

Evaluation Data

Phenotyping Data	download <i>phenotypic characters</i>
Genotyping Data	download

②

Associated Markers / Alleles

CH05c06	120
Hi04e04	228, 232
Md-Exp7	202, 212

③

④

image not available

Fig. 3.1 Variety page

1. Variety information.

The top table shows alias, parent and crossing information about the selected variety. Variety name of parents are linked to their variety pages.

2. Evaluation data

Both phenotyping and genotyping evaluation data of the variety can be downloaded. Click "*phenotypic characters*" link to bring up "*Phenotype Descriptor Selector*" to select which evaluation data to be included in the Excel file.

*See Section 4.2 "*Download phenotypic evaluation data*" and Section 6.1 "*Download genotypic evaluation data*" for downloading evaluation data.

3. Associated marker / allele

The marker / allele table lists all the markers that have been used for genotyping the variety and alleles that have been detected. All the associated markers are listed in the leftmost column and alleles are listed in the next column. Marker names are linked to GDR marker page.

4. Variety photo images

It shows photo image(s) of a variety if available. When a variety image is clicked, it shows the enlarged image and slideshow (Fig. 3.2).

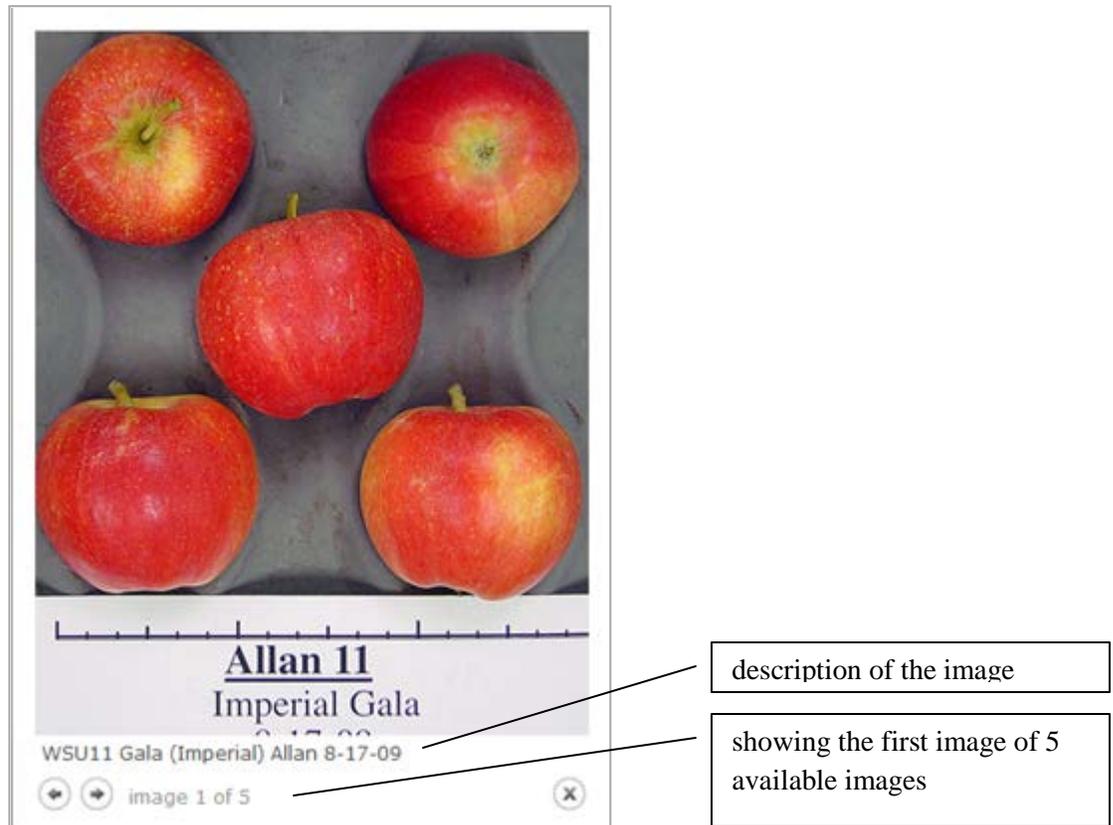


Fig. 3.2 Slideshow of variety images

4. Phenotyping Data

Users can download the phenotyping evaluation data for varieties in an Excel file. However, evaluation data may be too large to be downloaded online. In such a case, an Excel file is generated offline and stored in our server. Upon creation of an Excel file, an email notification along with the online file location will be sent to the requested user for downloading

4.1 Select phenotype characteristics

Before downloading the result in an Excel file, a user can select which phenotypic evaluation data to be included in a downloadable Excel file. By clicking the "Select phenotypic characters" button brings up "Phenotype Descriptor Selector" (Fig. 4.1).

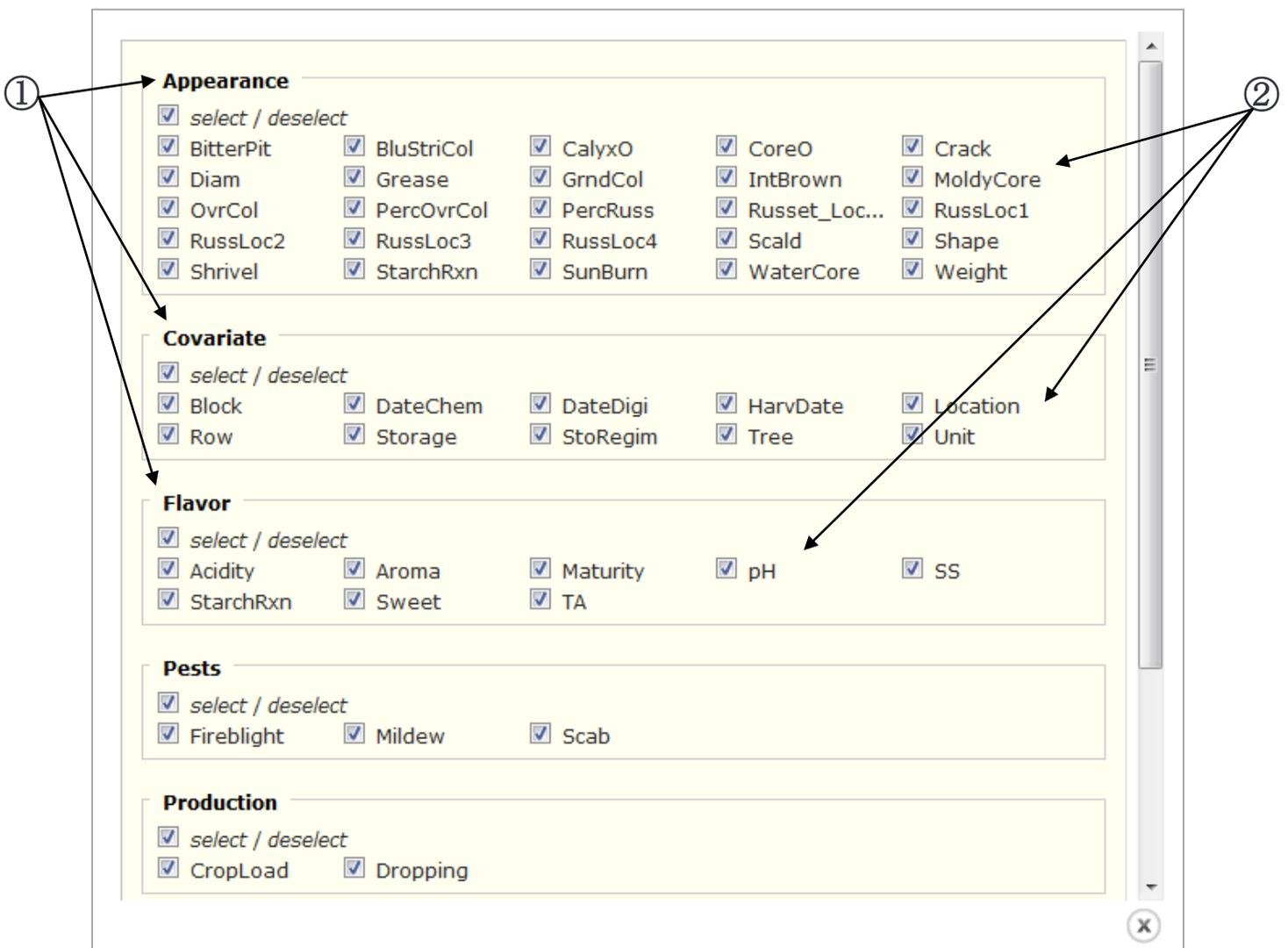


Fig. 4.1 Phenotype descriptor selector

1. Phenotypic categories

Phenotypic descriptors are organized by the categories that a breeder has provided.

2. Descriptor checkboxes

Check all the descriptor checkboxes you want to include in a downloadable Excel file.

4.2 Download phenotypic evaluation data

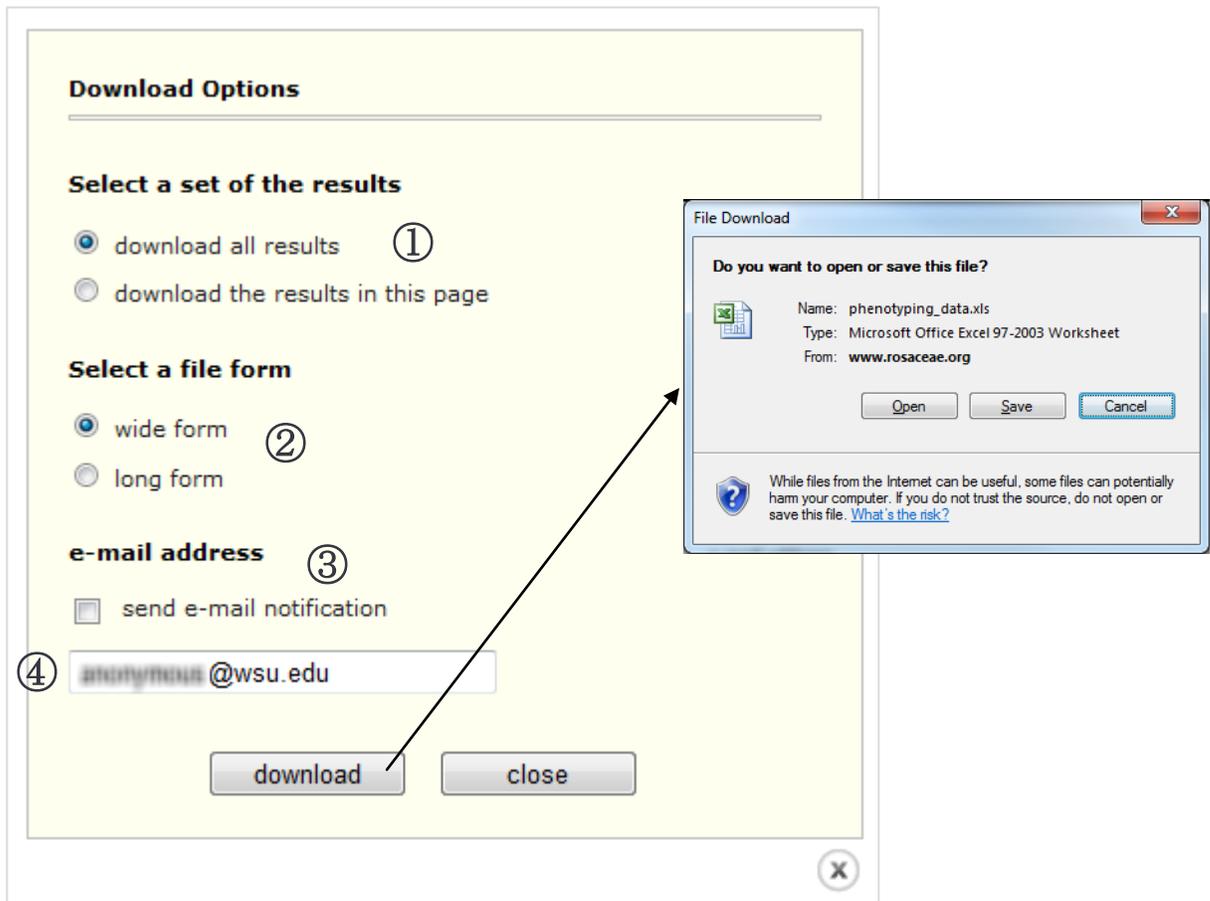


Fig. 4.2 Example of downloadable excel file

1. Datasets

Select "download all results" if you want to download evaluation data of all results of the search. It is common that a search returns more results than can fit in a single page. Therefore, results are viewed in several pages. The option "download the results in this page" is for a user to download evaluation data of the results only shown in the current page.

2. File form

Select a format ("wide" / "long") of a downloadable Excel file. Examples are shown in Fig. 4.4 and 4.5.

3. Email notification

Check the box and type your email address if you want to receive an email notification when the Excel file is generated.

4. Download excel file

Click "download" to download excel file.

4.3 Format of phenotyping data in an Excel file

An Excel file contains several sheets of evaluation data and one sheet for descriptions of phenotype descriptor (Fig. 4.3).

1	Category	Phenotype	Description	Code
2	Appearance	BITTERPIT	bitterpit presense	1:none;2:slight one fruit;3:medium 2or3 fruit;4:severe 4or5 fruit
3	Appearance	BLUSTRICOL	Type of red color	1: bhush; 2:stripe-blushed;3:blushed-striped;4:blushed
27	Covariate	BLOCK	Planting block or field	
28	Covariate	DATECHEM	n/a	
29	Covariate	DATEDIGI	n/a	
30	Covariate	HARVDATE	n/a	
37	Flavor	ACIDITY	Taste (acid)	1: no acid; 2: slightly acid; 3: medium; 4: acid; 5: very acid
38	Flavor	AROMA	Flavor (fruit aroma)	1: none; 2: slight aroma; 3: medium; 4: fruity; 5: very fruity
39	Flavor	MATURITY	sensory perception of ripeness	1:immature (starchy,granular);2:mature(no detectable startch);3:ov
40	Flavor	PH	pH	
41	Flavor	SS	Soluble Solids (refractometer reading from juice of 5 fruits) in Brix	
42	Flavor	STARCHRXXN	Cornell Starch Index for Ripeness	
43	Flavor	SWEET	Taste (sweetness)	1: no sweetness; 2: slightly sweet; 3: medium; 4: sweet; 5: very sw
44	Flavor	TA	% or milligrams per ml (malic acid equivalent), titrated to pH of 8.2	
45	Pests	FIREBLIGHT	disease severity	0: none; 1: slight; 2: medium; 3: severe
46	Pests	MILDEW	disease severity	0: none; 1: slight; 2: medium; 3: severe
47	Pests	SCAB	disease severity	0: none; 1: slight; 2: medium; 3: severe
48	Production	CROPLOAD	Crop load relative to size of tree	1: very light; 2: light; 3: medium; 4: heavy; 5: v.heavy
49	Production	DROPPING	amount of preharvest dropping of fruit relative to crop on tree	1: very light; 2: light; 3: medium; 4: heavy; 5: v.heavy
50	Texture	A1	Average Hardness of Region 1 (surface with no skin to 8.128mm)	
51	Texture	A2	Average Hardness of Region 2 (8.128mm to core)	
52	Texture	A3	Average hardness in core	
53	Texture	C0	10 pound creep force for 0.5 seconds at 0.32 inches 8.128mm	
54	Texture	C1	10 pound creep force for 0.5 seconds at boundary between Region 2 and core	
55	Texture	CN	Fourier transformation of force measurements in Region 2	
56	Texture	CRISP	Acoustic sensation that is detected by the ear during the fracturing	1: not crisp; 2: slightly crisp; 3: medium; 4: fairly crisp; 5: crisp

Fig. 4.3 Format of a downloadable Excel file

1	Dataset	Variety ID	Clone ID	Sample ID	Site	ISI	EP	Row	Position	R	PI	Acidity	Aroma	Maturity	pH	SS	StarchRxn	Sweet	TA
2	Apple_CRS_phenotyping_2010	Akane	Akane	Akane_00	CRS_APPLE			4	6-8			2.25	3.75	2.5	3.44	11.65	7.8	3.5	0.58705
3	Apple_CRS_phenotyping_2010	Akane	Akane	Akane_10	CRS_APPLE			4	6-8			3.25	3.5	2.375	3.53	12.5		3	0.3669

Fig. 4.4 "wide form" of phenotyping data in an Excel file

Q10 Acidity

phenotyping_data[1].xls [Compatibility Mode]

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
1	Dataset	Variety	Clone ID	Sample ID	Site	Pi	St	Ev	St	C	E	Pl	Ro	Position	Ro	Pl	phenotype	value
2	Apple_CRS_phenotyping_2010	Akane	Akane	Akane_00	CRS_APPLE								4	6-8			Acidity	2.25
3	Apple_CRS_phenotyping_2010	Akane	Akane	Akane_00	CRS_APPLE								4	6-8			Aroma	3.75
4	Apple_CRS_phenotyping_2010	Akane	Akane	Akane_00	CRS_APPLE								4	6-8			Maturity	2.5
5	Apple_CRS_phenotyping_2010	Akane	Akane	Akane_00	CRS_APPLE								4	6-8			pH	3.44
6	Apple_CRS_phenotyping_2010	Akane	Akane	Akane_00	CRS_APPLE								4	6-8			SS	11.65
7	Apple_CRS_phenotyping_2010	Akane	Akane	Akane_00	CRS_APPLE								4	6-8			StarchRxn	7.8
8	Apple_CRS_phenotyping_2010	Akane	Akane	Akane_00	CRS_APPLE								4	6-8			Sweet	3.5
9	Apple_CRS_phenotyping_2010	Akane	Akane	Akane_00	CRS_APPLE								4	6-8			TA	0.58705
10	Apple_CRS_phenotyping_2010	Akane	Akane	Akane_10	CRS_APPLE								4	6-8			Acidity	3.25
11	Apple_CRS_phenotyping_2010	Akane	Akane	Akane_10	CRS_APPLE								4	6-8			Aroma	3.5
12	Apple_CRS_phenotyping_2010	Akane	Akane	Akane_10	CRS_APPLE								4	6-8			Maturity	2.375
13	Apple_CRS_phenotyping_2010	Akane	Akane	Akane_10	CRS_APPLE								4	6-8			pH	3.53
14	Apple_CRS_phenotyping_2010	Akane	Akane	Akane_10	CRS_APPLE								4	6-8			SS	12.5
15	Apple_CRS_phenotyping_2010	Akane	Akane	Akane_10	CRS_APPLE								4	6-8			StarchRxn	
16	Apple_CRS_phenotyping_2010	Akane	Akane	Akane_10	CRS_APPLE								4	6-8			Sweet	3
17	Apple_CRS_phenotyping_2010	Akane	Akane	Akane_10	CRS_APPLE								4	6-8			TA	0.3669
18																		
19																		
20																		
21																		
22																		

1st row

2nd row

phenotypic descriptors values

Fig. 4.5 "long form" of phenotyping data in an Excel file

5. Search Phenotyping Data

The Breeders Toolbox allows user to search breeding data by variety names, traits, or by parentage. From the search results, you can access the record of the individual variety or download the evaluation data of all the varieties you have searched for.

5.1a. Search by Variety Name

Search by Varieties | Search by Traits | Search by Parentage

Search breeding data by varieties, traits, or parentage. From the search results, you can access the record of the individual variety or download the evaluation data of all the varieties you have searched for.

Choose a destination...

Search by Varieties (see help) ← ①

Choose Crop Apple ②

Choose Breeding Group APPLE CRS/BPS MN ③

Names of Varieties exactly ④

(e.g.)
669NJ5, Akane
Fuji ⑤

Example 1:
select 'exactly' and type WSU11

Example 2:
select 'start with' and type
MN
WSU

includes aliases ⑥

Upload List of Varieties Browse... ⑦

Restricted by Datasets

```
Apple_CRS_phenotyping_2010
--- RosBREED Apple Crop Reference Set
Apple_BPS_MN_phenotyping_2010
--- RosBREED Apple Minnesota Breeding Pe
```

 ⑧

Search Reset

Fig. 5.1 Search by variety name form

1. Search help

Click "*see help*" to view help for the search form

2. Crop selection

Select a crop from the dropdown. The dropdown lists all currently available crops.

3. Breeding group selection

Select a breeding group from the dropdown. The dropdown lists all breeding groups that the current user belongs to. The dropdown appears only if the current user belongs to more than one breeding group.

4. Select a condition

Select search condition of query keywords. The conditions, "*exactly*", "*start with*", "*contains*" and "*end with*" are available.

5. Type variety names

Type complete or partial names of varieties in the box. Names can be listed on different rows and/or names can be separated by commas in the same row (Fig. 5.2).

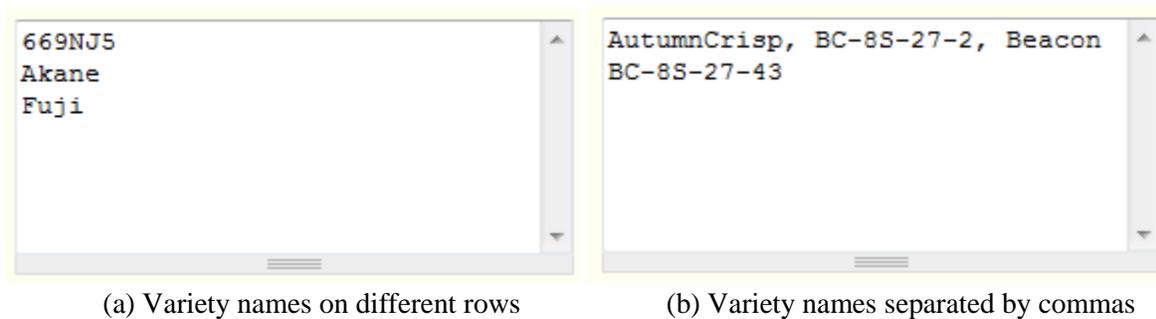


Fig. 5.2 Example of searching multiple variety names

6. Alias checkbox

By default, it searches for matches in variety names as well as their alias(es). If users don't want to search on alias, they can uncheck the alias checkbox.

7. Upload a file

A user can upload a file with multiple variety names.

8. Restrict the search by datasets.

Narrow down your search result by selecting one or more datasets. To select multiple datasets, keep pressing 'Ctrl' button while clicking datasets in the list (Fig. 5.3).

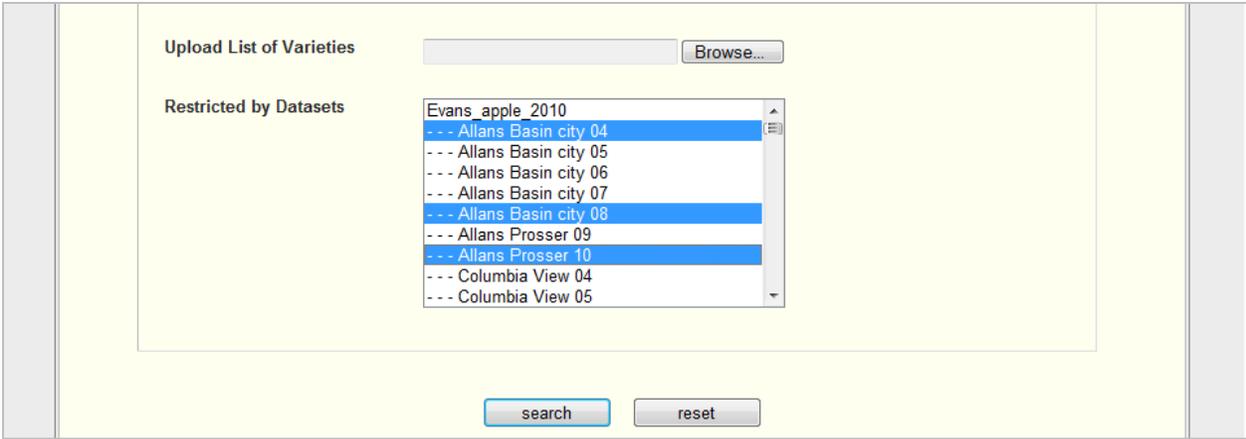


Fig. 5.3 Selecting multiple datasets

5.1b. Results of Search by Variety Name

Search by Varieties Search by Traits Search by Parentages

Your Search

varieties that exactly matched with '669NJ5 ', 'Akane ', 'Fuji'

Your search for variety returned 3 record(s)

Click "Select phenotypic characters" to bring up a screen from which to select the traits or variables which you would like to have downloaded.

Click on "Download search results" to start downloading a file containing the results of your variety and trait selection request.

After downloading one set of search results, you can return to this page to select different sets of phenotypic characters to download.

Select phenotypic characters

Download search results

Download help

669NJ5	Akane	Fuji
--------	-------	------

Fig. 5.4 Example result of search by variety name

1. Summary of the search

"*Your Search*" section summarizes your search criteria and its result.

2. Result table

The table lists all varieties returned by the search. The variety names are linked to the variety page.

3. Select phenotypic characteristics

See section 4.1 "*Select phenotype characteristics*"

4. Download the results in an Excel file

See section 4.2 "*Download phenotypic evaluation data*"

5. Download help

Click "*Download help*" button to view download help

5.2a. Search by Traits

Search by Varieties | Search by Traits | Search by Parentages

[Go Genotyping Data Search](#)

Search breeding data by varieties, traits, or parentages. From the search results, you can access the record of the individual variety or download the evaluation data of all the varieties you have searched for.

Search by Traits ①

Search for varieties with a certain set of phenotypic values. You can set a range for numeric phenotypic values and choose multiple coded values by holding down the control key ([see help](#)).

Selected Breeder Group: ②

③ Appearance

BITTERPI.. 1: none 2: slight one fruit 3: medium 2or3 fruit	BLUSTRIC.. 1: blush 2: stripe-blushed 3: blushed-striped
CALYXO 1: all closed 2: mixed 3: all open	COREO 0: open 1: closed
CRACK 1: None 2: Slight – one fruit 3: Medium – 2 or 3 fruit	DIAM <input type="text"/> ≤ DIAM ≤ <input type="text"/>
GREASE 1: none 2: slight 3: medium	GRNDCOL 1: green 2: pale green 3: green-yellow

④

M2 <input type="text"/> ≤ M2 ≤ <input type="text"/>	M3 <input type="text"/> ≤ M3 ≤ <input type="text"/>
MAG-TAYL.. <input type="text"/> ≤ MAG-TAYL.. ≤ <input type="text"/>	MATURITY 1: immature (starchy, granular) 2: mature(no detectable starch) 3: over mature (soft, lacks acid an
OAH <input type="text"/> ≤ OAH ≤ <input type="text"/>	OMH <input type="text"/> ≤ OMH ≤ <input type="text"/>
QF <input type="text"/> ≤ QF ≤ <input type="text"/>	STARCHRX.. <input type="text"/> ≤ STARCHRX.. ≤ <input type="text"/>
WEIGHT <input type="text"/> ≤ WEIGHT ≤ <input type="text"/>	

Phenotype category

Fig. 5.5 Search by traits form

1. Search help

Click "see help" to view help for searching form

2. Breeding group selection

Select a breeding group from the dropdown. The dropdown lists all breeding groups that the current user belongs to. The dropdown appears only if the current user belongs to more than one breeding group.

3. Phenotypic descriptors

Click the information icon  to view descriptions of phenotypic descriptors.

4. Search criteria on evaluation data

Search for varieties with a certain set of phenotypic values. You can set a range for numeric phenotypic values and choose multiple coded values by holding down the "Ctrl" key.

(a) Set a range of coded values

Multiple values can be selected. Click target values while pressing "Ctrl" key.



Fig. 5.6 Set ranges of coded values (SHAPE and SCALD)

(b) Set a range of quantitative values

Set the minimum and maximum values on the phenotypic data. To get an exact value, input the same numerical value in both maximum and minimum boxes.

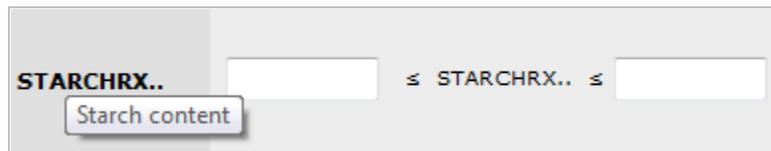


Fig. 5.7 Set ranges on "starch content"

5.2b. Results of Search by Traits

Search by Varieties Search by Traits Search by Parentages

Your Search

BitterPit	: 3 : medium 2or3 fruit
BluStriCol	: 1 : blush, 3 : blushed-striped
CoreO	: 1 : closed
Diam	: [min] 3
Diam	: [max] 6

i descriptor information

Your search for variety returned 1 record(s)

Click "Select phenotypic characters" to bring up a screen from which to select the traits or variables which you would like to have downloaded.

Click on "Download search results" to start downloading a file containing the results of your variety and trait selection request.

After downloading one set of search results, you can return to this page to select different sets of phenotypic characters to download.

0410-0015

Select phenotypic characters

Download search results

Download help

Fig. 5.8 Example result of search by traits

1. Summary of the search

"*Your Search*" section summarizes your search criteria and the result.

2. Result table

The table lists all the varieties returned by the search. The variety names are linked to the variety page.

3. Phenotypic descriptors

Click the information icon *i* to view definitions of phenotypic descriptors.

4. Select phenotypic characteristics

See section 4.1 "*Select phenotype characteristics*"

5. Download the result in an Excel file

See section 4.2 "*Download phenotypic evaluation data*"

6. Download help

Click "*Download help*" button to view download help

5.3a. Search by Parentage

Search by Varieties Search by Traits Search by Parentages

[Go Genotyping Data Search](#)

Search breeding data by varieties, traits, or parentages. From the search results, you can access the record of the individual variety or download the evaluation data of all the varieties you have searched for.

Search Varieties by Parentages

Selected Breeder Group ①

Maternal Parents

- 74-6-77
- Arlet
- AurGoldGa
- AutumnCrisp
- Blackjon
- Braeburn** ②
- Clivia
- Cortland
- Cox
- Crandall

or
or
and ④

Paternal Parents

- 669NJ5** ③
- Akane
- Antonovka
- Arlet
- AurGoldGa
- AutumnCrisp
- BC-8S-27-2
- BC-8S-27-43
- BeautBath
- BenDavis

Fig. 5.9 Search by parentage form

1. Breeding group selection

Select a breeding group from the dropdown. The dropdown lists all the breeding groups that the current user belongs to. The dropdown appears only if the current user belongs to more than one breeding group.

2. Maternal parent selection

Select one of the varieties from the list as a maternal parent.

3. Paternal parent selection

Select one of the varieties from the list as a paternal parent.

4. "OR" and "AND" logical operators

Select one of the logical operators to change the search condition.

(e.g.)

A user selects "*Braeburn*" as a maternal parent and "*669NJ5*" as a paternal parent in Fig. 5.8.

By selecting "*and*", the search will return all varieties whose maternal parent is "*Braeburn*" and paternal parent is "*669NJ5*".

By selecting "or", the search will return all varieties whose maternal parent is "Braeburn" or paternal parent is "669NJ5".

5.3b. Results of Search by Parentage

The screenshot shows a web interface for searching by parentage. At the top, there are three tabs: "Search by Varieties", "Search by Traits", and "Search by Parentages". The "Search by Parentages" tab is selected. Below the tabs is a search box containing the text "You've selected Braeburn (mother) or 669NJ5 (father)". Below the search box, it says "Your search for variety returned 21 record(s)".

There are three numbered callouts: 1 points to the search box, 2 points to the results table, and 3, 4, and 5 point to the buttons "Select phenotypic characters", "Download search results", and "Download help" respectively.

Below the search box, there are three instructions:

- Click "Select phenotypic characters" to bring up a screen from which to select the traits or variables which you would like to have downloaded.
- Click on "Download search results" to start downloading a file containing the results of your variety and trait selection request.
- After downloading one set of search results, you can return to this page to select different sets of phenotypic characters to download.

The results table is as follows:

Coop18	GR9P7#232	GR9P7#233	GR9P7#234
GR9P7#235	GR9P7#236	GR9P7#237	GR9P7#238
GR9P7#239	GR9P7#240	GR9P7#241	GR9P7#242
GR9P7#243	GR9P7#245	GR9P7#246	GR9P7#258
GR9P7#259	GR9P7#262	GR9P7#263	GR9P7#268

At the bottom of the table, there is a pagination link: "1 2 next > last »".

Fig. 5.10 Example of the results from search by parentages

1. Summary of the search

"Your Search" section summarizes your search criteria and the result.

2. Results table

The table lists all the varieties returned by the search. The variety names are linked to the variety page.

3. Select phenotypic characteristics

See section 4.1 "Select phenotype characteristics"

4. Download the result in an Excel file

See section 4.2 "Download phenotypic evaluation data"

5. Download help

Click "Download help" button to view download help

6. Genotyping Data

Users can download the genotyping evaluation data of varieties in an Excel file. However, evaluation data may be too large to be downloaded online. In such a case, an Excel file is generated offline and stored in our server. Upon creation of an Excel file, an email notification with the file location will be sent to the requested user.

6.1 Download genotypic evaluation data

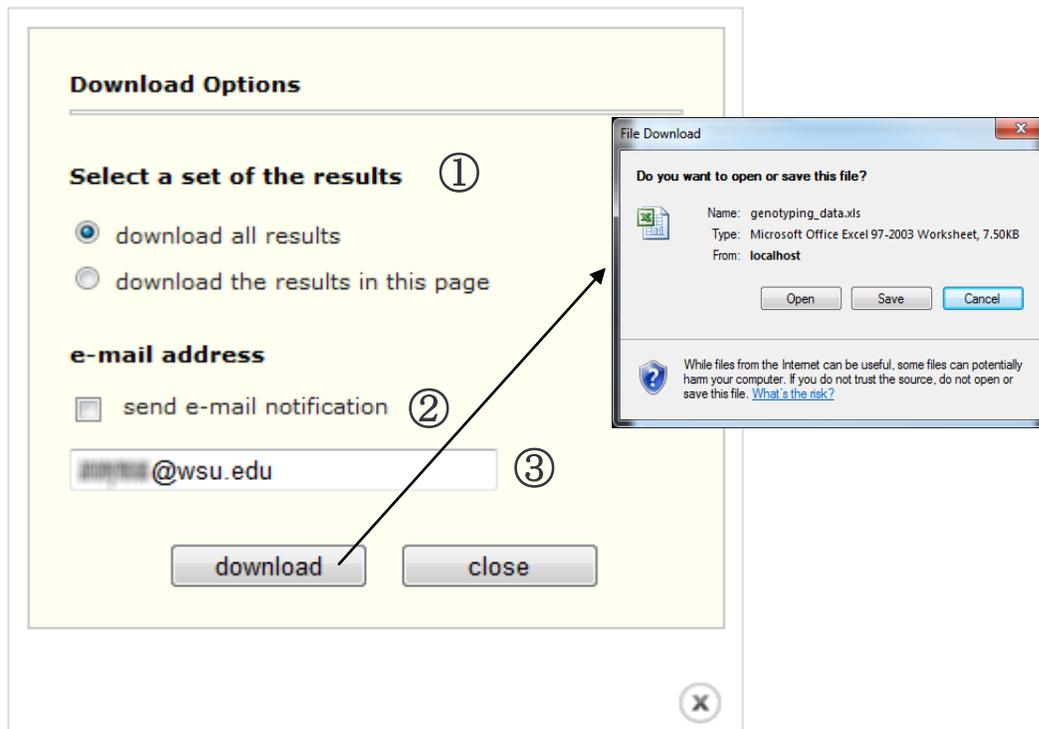


Fig. 6.2 Options for downloading genotyping data

1. Datasets

Select "*download all results*" if you want to download evaluation data of all the results of the search. It is common that a search returns results in multiple pages. The option "*download the results in this page*" is for a user to download evaluation data of the results only shown in the current page.

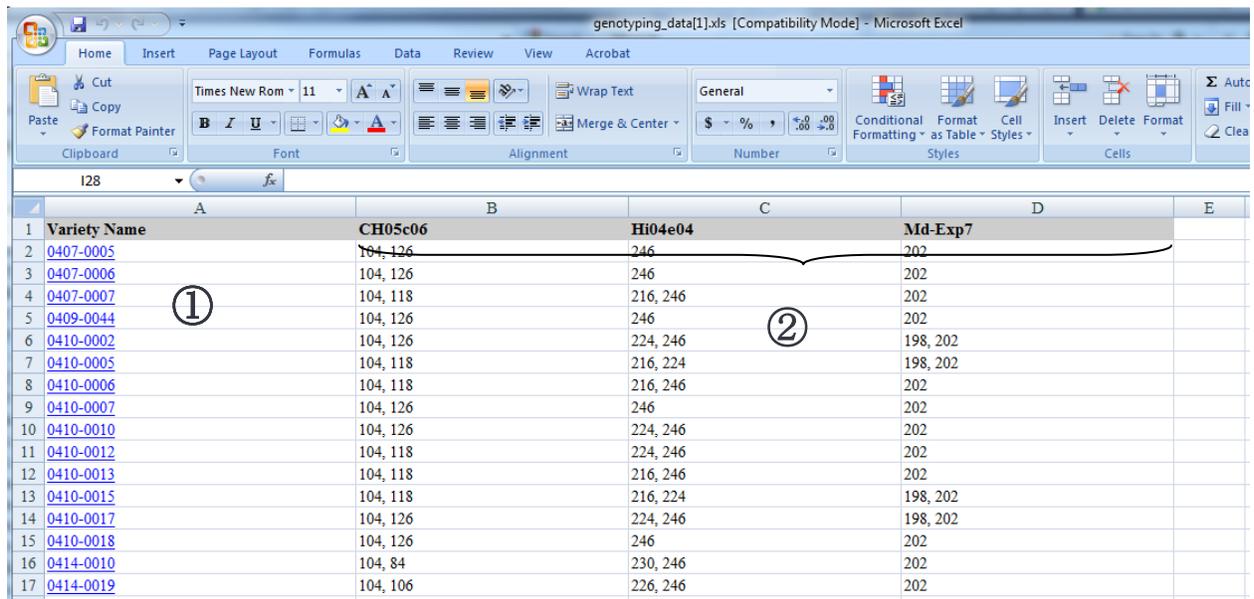
2. Email notification

Check the box and type your email address if you want to receive an email notification when an Excel file is generated.

3. Download excel file

Click "*download*" to download an excel file.

6.2 Format of genotyping data in Excel



	A	B	C	D	E
1	Variety Name	CH05c06	Hi04e04	Md-Exp7	
2	0407-0005	104, 126	246	202	
3	0407-0006	104, 126	246	202	
4	0407-0007	104, 118	216, 246	202	
5	0409-0044 ①	104, 126	246	202	
6	0410-0002	104, 126	224, 246	198, 202	
7	0410-0005	104, 118	216, 224	198, 202	
8	0410-0006	104, 118	216, 246	202	
9	0410-0007	104, 126	246	202	
10	0410-0010	104, 126	224, 246	202	
11	0410-0012	104, 118	224, 246	202	
12	0410-0013	104, 118	216, 246	202	
13	0410-0015	104, 118	216, 224	198, 202	
14	0410-0017	104, 126	224, 246	198, 202	
15	0410-0018	104, 126	246	202	
16	0414-0010	104, 84	230, 246	202	
17	0414-0019	104, 106	226, 246	202	

Fig. 6.2 Format of genotyping data in an Excel file

1. Variety Name column

Variety names are hyperlinked to their variety pages.

2. Marker names columns

The associated markers are listed in horizontal direction. The associated alleles are listed in the marker column.

7. Search Genotyping Data

Search genotyping data by variety name, combination of variety name and marker, or combination of marker and allele. From the search results, you can access the record of the individual variety or download all the genotyping data you have searched for.

7.1a Search by Variety Name

The screenshot shows a web interface for searching genotyping data. At the top, there are three tabs: "Search by Varieties", "Search by Variety / Marker", and "Search by Marker / Allele". Below the tabs is a descriptive paragraph and a "Choose a destination..." dropdown menu. The main form area is titled "Search by Varieties (see help)" with a circled "1" and an arrow pointing to the "see help" link. The form contains several fields: "Choose Crop" with a dropdown menu showing "Apple" (circled "2"); "Choose Breeding Group" with a dropdown menu showing "APPLE CRS" (circled "3"); "Names of Varieties" with a dropdown menu showing "exactly" (circled "4"); a text input field containing "Akane, Fuji" (circled "5") with a small "(e.g.) Akane, Fuji" example below it; a checkbox labeled "includes aliases" which is checked (circled "6"); and an "Upload List of Varieties" section with a file input field and a "Browse..." button (circled "7"). To the right of the text input field, there are two examples: "Example 1: Select exactly and type 'COOP15'" and "Example 2: Select start with and type 'COOP' 'GMAL'". At the bottom of the form are "Search" and "Reset" buttons.

Fig. 7.1 Search by variety name form

1. Search help

Click "*see help*" to view help for searching form

2. Crop selection

Select a crop from the dropdown. The dropdown lists all currently available crops.

3. Breeding group selection

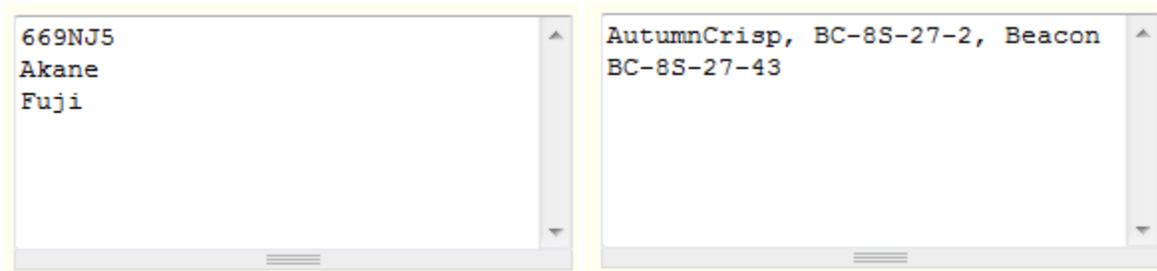
Select a breeding group from the dropdown. The dropdown lists all breeding groups that the current user belongs to. The dropdown appears only if the current user belongs to more than one breeding group.

4. Select a condition

Select search condition of query keywords. The conditions, "exactly", "start with", "contains" and "end with" are available.

5. Type variety names

Users can type complete or partial names of varieties in the box. Names can be listed on different rows and/or names can be separated by commas in the same row (Fig. 7.2).



(a) variety names on different rows

(b) variety names separated by commas

Fig. 7.2 Example of searching multiple variety names

6. Alias checkbox

By defaults, variety names along with their alias(es) will be searched. Uncheck the checkbox If you want to exclude aliases from your search.

7. Upload a file

A user can upload a file with multiple variety names.

7.1b Results of Search by Variety Name

Search by Varieties | Search by Variety / Marker | Search by Marker / Allele

① Your Search
varieties that start with 'COOP'

Your search for variety returned 3 record(s)

Click on "Download search results" to download a file with the genotyping data. [Download search results](#) ③

② [Coop15 \(view data\)](#) [Coop17 \(view data\)](#) [Coop18 \(view data\)](#)

Fig. 7.3 Example of the results from search by variety name

1. Summary of the search

"Your Search" section summarizes your search criteria and the result.

2. Results table

The table lists all varieties that are returned by the search. The variety names are linked to the variety page. "view data" next to variety name is linked to "Variety / Marker" page (see Section 7.2b).

3. Download genotyping data

See section 6.1 "Download genotypic evaluation data".

7.2a Search by Variety / Marker

Search by Varieties | Search by Variety / Marker | Search by Marker / Allele

Search genotyping data by variety name, combination of variety name and marker, or combination of marker and allele. From the search results, you can access the record of the individual variety or download all the genotyping data you have searched for. Choose a destination...

Search by Variety / Marker

Choose Crop Apple ①

Choose Breeding Group APPLE CRS ②

Variety Name

- GR14P9#60
- GR14P9#61
- GR14P9#64 ③
- GR14P9#65**
- GR14P9#66
- GR14P9#67
- GR14P9#68
- GR14P9#69

Name Filter gr14p9#6 ④ (to display only the variety names that starts with)

Associated Marker Names

GR14P9#65 has genotyping results from 3 markers

- CH05c06
- Hi04e04
- Md-Exp7 ⑤

Search

Fig. 7.4 Search by variety / marker form

1. Crop selection

Select a crop from the dropdown. The dropdown lists all currently available crops.

2. Breeding group selection

Select a breeding group from the dropdown. The dropdown lists all breeding groups that the current user belongs to. The dropdown appears only if the current user belongs to more than one breeding group.

3. Variety name list

All the names of varieties are listed for users to select. When a variety name in the list is selected, it shows all the associated markers in the "*Marker table*" at bottom of page. A list of markers is dynamically changed depending on the which variety in the list is selected.

4. Name filter

Users can type some letters in the filter to list only the variety names that starts with those letters.

5. Marker table

All the markers associated with the selected variety are listed in the marker table. Marker names are linked to the marker pages.

(e.g.) Example in Fig. 7.4

(a) A user has typed "*GR14P9#6*" in the name filter box. Therefore, only the varieties whose names start with

"*GR14P9#6*" are shown in the list.

(b) "*GR14P9#65*" has been selected in the list. Its associated markers "*CH05c06*", "*Hi04e04*" and "*Md-Exp7*" are listed in the marker table.

7.2b Results of Search by Variety / Marker

The screenshot shows a web interface with three tabs: "Search by Varieties", "Search by Variety / Marker", and "Search by Marker / Allele". The "Search by Variety / Marker" tab is active. A summary box (1) displays "You've selected 0409-0043" and "Your search returned 3 markers". Below this, a button (5) labeled "Download search results" is visible. A table (2) lists three markers: CH05c06, Hi04e04, and Md-Exp7. The table has three columns: "Marker" (2), "Allele (bold: associated alleles)" (3), and "View all genotyping data" (4). The alleles for CH05c06 are 104, **106**, 108, 110, 114, 116, 118, 120, 122, **126**, 128, 132, 84, 90. The alleles for Hi04e04 are 204, 216, 220, 222, 224, 225, **226**, 228, 230, **246**. The alleles for Md-Exp7 are 198, 199, 200, **202**, 204, 206, **210**, 214.

Marker ^②	Allele (bold: associated alleles) ^③	View all genotyping data ^④
CH05c06	104, 106 , 108, 110, 114, 116, 118, 120, 122, 126 , 128, 132, 84, 90	view all genotyping data for CH05c06
Hi04e04	204, 216, 220, 222, 224, 225, 226 , 228, 230, 246	view all genotyping data for Hi04e04
Md-Exp7	198, 199, 200, 202 , 204, 206, 210 , 214	view all genotyping data for Md-Exp7

Fig. 7.5 Example of the result of search by variety / marker

1. Summary of the search

It shows name of the selected variety and the number of markers that are associated with the variety.

2. Marker name column

The associated markers are listed. Each marker is linked to the GDR marker page.

3. Allele column

All the alleles associated with the marker on the left-most column are listed. Alleles in bold are alleles that have been detected in the selected variety.

4. View all genotyping data column

It is linked to "Marker / Allele" page (see Section 7.4).

5. Download genotyping data

See section 6.1 "Download genotypic evaluation data"

7.3 Search by Marker / Allele

Search by Varieties | Search by Variety / Marker | Search by Marker / Allele

Search genotyping data by variety name, combination of variety name and marker, or combination of marker and allele. From the search results, you can access the record of the individual variety or download all the genotyping data you have searched for.

Choose a destination...

Search by Marker / Allele

Choose Crop Apple ①

Choose Breeding Group APPLE CRS ②

Marker Name Md-Exp7 ③
CH05c06
Hi04e04
Md-Exp7

Associated Alleles

Md-Exp7 has 8 alleles ④

198	199	200	202	204
206	210	214		

Fig. 7.6 Search by marker / allele form

1. Crop selection

Select a crop from the dropdown. The dropdown lists all currently available crops.

2. Breeding group selection

Select a breeding group from the dropdown. The dropdown lists all the breeding groups that the current user belongs to. The dropdown appears only if the current user belongs to more than one breeding group.

2. Marker name list

All markers are listed in the dropdown. When a marker name in the list is selected, it shows all associated alleles in the allele table at the bottom of the page. A list of alleles is dynamically changed depending on which marker in the list is selected.

3. Allele table

The allele table lists all the alleles associated with the selected marker from the marker list. Alleles are linked to "*Marker / Allele*" page (see Section 7.4).

7.4 Marker / Allele page

Search by Varieties Search by Variety / Marker Search by Marker / Allele

List of varieties associated with the following marker and alleles

# of varieties	18
marker name	Md-Exp7
selected allele	200

Click on "Download search results" to download a file with the genotyping data.

Download search results

AE0210-149	AE0210-151	AE0210-152	AE0210-153
AE0210-158	AE0210-159	AE0210-160	AE0210-163
AE0210-164	AE0210-165	AE0210-168	AE0210-171
AE0210-172	AE0210-174	AE0210-178	AE0251-325
GMAL4329	RussianSeedling		

Fig. 7.7 Marker / allele page

1. Marker /Allele information table

It shows information about the selected marker/allele and number of varieties in the dataset that has the selected marker/allele. Marker name is linked to GDR marker page.

2. Download genotyping data

See section 6.1 "*Download genotypic evaluation data*"

3. Results table

All the varieties associated with the selected marker and allele(s) are listed in the variety table. Variety names are linked to variety page.

Marker / Allele page can be reached from (a) "*Search marker/allele*" page and (b) "*Result of search by variety/marker*" page (Fig. 7.8a and 7.8b).

(a) Linked from "*Search marker / allele*" page (Fig. 7.6)

Search by Varieties Search by Variety / Marker Search by Marker / Allele

You've selected 0409-0043
Your search returned 3 markers

Click on "Download search results" to download a file with the genotyping data.

Marker ▲	Allele (bold: associated alleles)	View all genotyping data
CH05c06	104, 106 , 108, 110, 114, 116, 118, 120, 122, 126 , 128, 132, 84, 90	view all genotyping data for CH05c06
Hi04e04	204, 216, 220, 222, 224, 225, 226 , 228, 230, 246	view all genotyping data for Hi04e04
Md-Exp7	198, 199, 200, 202 , 204, 206, 210 , 214	view all genotyping data for Md-Exp7

# of varieties	409
marker name	CH05c06
all alleles	<div style="border: 1px solid red; padding: 2px;"> 104 106 108 110 114 </div>

Fig. 7.8a Marker/Allele information table

When "view all genotyping data" column is clicked, all the associated alleles are listed in Marker /Allele information table. The result table lists all the varieties that have been tested with the selected marker.

(b) Linked from "Result of search by variety / marker" page (Fig. 7.5).

Search by Varieties | Search by Variety / Marker | Search by Marker / Allele

Search genotyping data by variety name, combination of variety name and marker, or combination of marker and allele. From the search results, you can access the record of the individual variety or download all the genotyping data you have searched for. Choose a destination...

Search by Marker / Allele

Choose Crop: Apple

Choose Breeding Group: APPLE CRS

Marker Name: Md-Exp7

Associated Alleles: Md-Exp7

Md-Exp7 has 8 alleles

198	199	200	202	204
206	210	214		

# of varieties	18
marker name	Md-Exp7
selected allele	200

Fig. 7.8b Marker/Allele information table

When one of the alleles listed is clicked, only the selected allele is listed in Marker /Allele information table. The table lists all the varieties that have the selected allele for the marker.

8. Pedimap

Pedimap is a tool for exploring and visualizing the flow of phenotypes and alleles (observed or based on Identity-by-Descent calculations) through pedigrees. The Breeders Toolbox allows users to generate an input file for the Pedimap software. Users may extract evaluation data from database and embed those values in an input file for use with Pedimap.

8.1 Generation of Pedimap input file

The screenshot shows a web form titled "Generate Input File for Pedimap". At the top right, there is a dropdown menu labeled "Choose a destination...". The main form area is divided into two columns. The left column contains a heading "Generate Input File for Pedimap", a paragraph of introductory text, a link to the Pedimap site (annotated with a circled 1), and two dropdown menus: "Choose Crop" (annotated with a circled 2) set to "Apple" and "Choose Breeding Group" (annotated with a circled 3) set to "APPLE CRS". The right column contains a text input field for "Name of Variety *" (annotated with a circled 4) set to "Arlet", a dropdown for "Ancestor Generations" (annotated with a circled 5) set to "3", and a dropdown for "Progeny Generations" (annotated with a circled 5) set to "2". Below these are radio buttons for "Trace Pedigree Through": "Female line" (selected, annotated with a circled 6), "Male line", and "Both". There are checkboxes for "Include progeny parents" (annotated with a circled 7) and "Include sibs" (checked, annotated with a circled 8). At the bottom of the right column is a text input field for "File name" (annotated with a circled 9) set to "pedimap.input.txt", with a note below it: "(Enter the desired name for the Pedimap input file that will be downloaded)". A "Search Pedigree" button is located at the bottom center of the form.

Fig. 8.1 Form of generating Pedimap input file

1. Pedimap home page

It is linked to Pedimap home page (http://www.plantbreeding.wur.nl/UK/software_pedimap.html).

2. Crop selection

3. Breeding group selection

Select a breeding group from the dropdown. The dropdown lists all breeding groups that the current user belongs to. The dropdown appears only if the current user belongs to more than one breeding group.

4. Variety name

Users can type the name of the target variety.

5. Ancestor / progeny generations

Select how many generations of ancestor and progeny to be included.

6. Trace pedigree

Select which line to be traced.

7. Progeny parents

Check if you want to include progeny parents.

8. Sibs

Check if you want to include siblings.

9. File name

Users need to type a name for the input file

8.2 Generating the Pedimap input file

Go back to Pedimap selection page

① There are 53 varieties in the pedigree

✓ **select phenotypes** to be included in Pedimap input file
(phenotypes should be selected to see default values)

NAME
<input type="radio"/> Arlet
<input type="radio"/> AurGoldGa
<input type="radio"/> CrippsPink
<input type="radio"/> Enterprise
<input type="radio"/> EsopSpitz
<input type="radio"/> GoldenDel
<input type="radio"/> Honeycrisp
<input type="radio"/> Idared

②

include evaluation data ③

④

Click a radio button in the list above to change value ⑤

Initially, no phenotypes are selected. So there is no phenotype column. Need to add phenotype columns (see Section 8.2)

Fig. 8.2 Example of generating a Pedimap input file

1. Select Phenotypes

Click "*select phenotypes*" to bring up the "*Phenotype selector*" to select phenotypes to be included in Pedimap input file (Fig. 8.3). To select phenotype, see Section 8.2.

2. Variety table

All the varieties in the pedigree, selected from the form in Fig. 8.1, are listed at the left-most column of the table. To add/edit phenotype columns, click link "*select phenotypes*" above. The phenotype columns display the average of all the recorded values for the specific variety and the trait.

3. Evaluation data

Check "*include evaluation data*" if you want to include evaluation data in a Pedimap input file.

4. Generate a Pedimap input file

Click "*generate Pedimap input file*" to save a Pedimap file.

5. Evaluation data table

Evaluation table will be shown when one of radio buttons in the "*Variety table*" is clicked (Fig. 8.4). See Section 8.2b to see how to use "*Evaluation data table*" to modify the trait values to be included in the Pedimap input file.

8.2a Select phenotypes

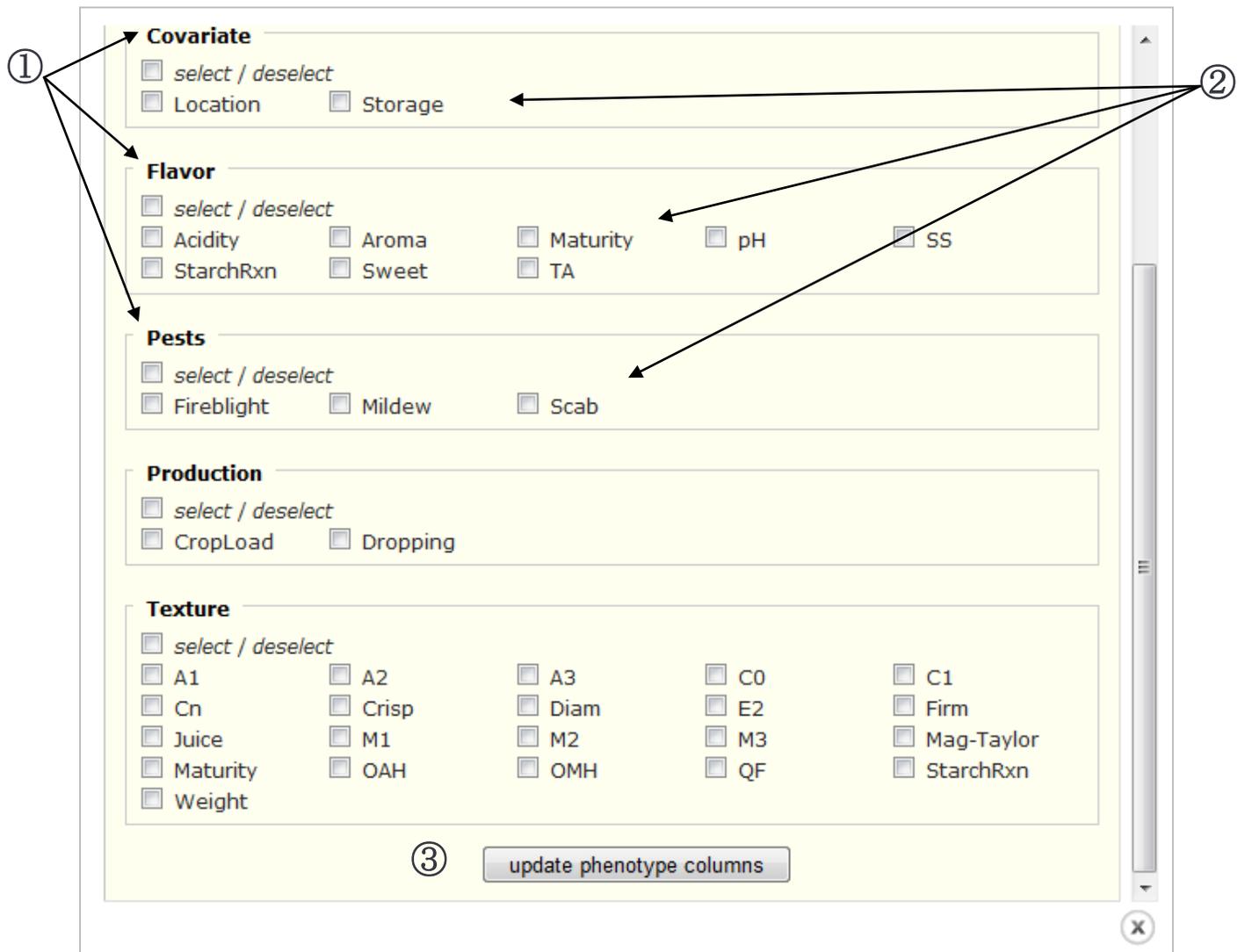


Fig. 8.3 Phenotype selector

1. Phenotypic categories

Phenotypic descriptors are organized by the categories that a breeder has provided.

2. Descriptor checkboxes

Check all descriptor checkboxes to be included in a Pedimap input file.

3. Update phenotypes

Click "update phenotype columns" button after selecting phenotypes to be included in a Pedimap input file. See an example in Section 8.2b8.2b Example of generating a Pedimap input file

[Go back to Pedimap selection page](#)

There are 53 varieties in the pedigree

✓ **select phenotypes** to be included in Pedmap input file
(phenotypes should be selected to see default values)

Scroll down for more varieties

NAME	BitterPit	Diam	OvrCol	Location	Acidity
<input type="radio"/> Arlet	<input type="text" value="3"/>	<input type="text" value="6.246622"/>	<input type="text" value="3"/>	<input type="text" value="1"/>	<input type="text" value="2.25"/>
<input checked="" type="radio"/> AurGoldGa	<input type="text" value="1"/>	<input type="text" value="7.36346"/>	<input type="text" value="4"/>	<input type="text" value="3"/>	<input type="text" value="3"/>
<input type="radio"/> CrippsPink	<input type="text" value="1"/>	<input type="text" value="7.63397"/>	<input type="text" value="4"/>	<input type="text" value="3"/>	<input type="text" value="3"/>
<input type="radio"/> Enterprise	<input type="text" value="1"/>	<input type="text" value="8.44296"/>	<input type="text" value="4"/>	<input type="text" value="3"/>	<input type="text" value="3"/>
<input type="radio"/> EsopSpitz	<input type="text" value="*"/>	<input type="text" value="*"/>	<input type="text" value="*"/>	<input type="text" value="*"/>	<input type="text" value="*"/>
<input type="radio"/> GoldenDel	<input type="text" value="*"/>	<input type="text" value="*"/>	<input type="text" value="*"/>	<input type="text" value="*"/>	<input type="text" value="*"/>
<input type="radio"/> Honeycrisp	<input type="text" value="2"/>	<input type="text" value="7.100824"/>	<input type="text" value="3"/>	<input type="text" value="1"/>	<input type="text" value="2.625"/>
<input type="radio"/> Idared	<input type="text" value="*"/>	<input type="text" value="*"/>	<input type="text" value="*"/>	<input type="text" value="*"/>	<input type="text" value="*"/>

include evaluation data

sample	BitterPit	Diam	OvrCol	Location	Acidity	A1	A:
default value	<input checked="" type="radio"/> 1	<input type="text" value="7.24408"/>	<input checked="" type="radio"/> 4	<input checked="" type="radio"/> 3	<input type="text" value="2.75"/>		
AurGoldGa	<input type="radio"/> 1	<input checked="" type="radio"/> 7.36346	<input type="radio"/> 4	<input type="radio"/> 3	<input type="text" value="2.5"/>	51.36361	1
AurGoldGa	<input type="radio"/> 1	<input type="radio"/> 7.1247	<input type="radio"/> *	<input type="radio"/> 3	<input checked="" type="radio"/> 3	47.16004	8

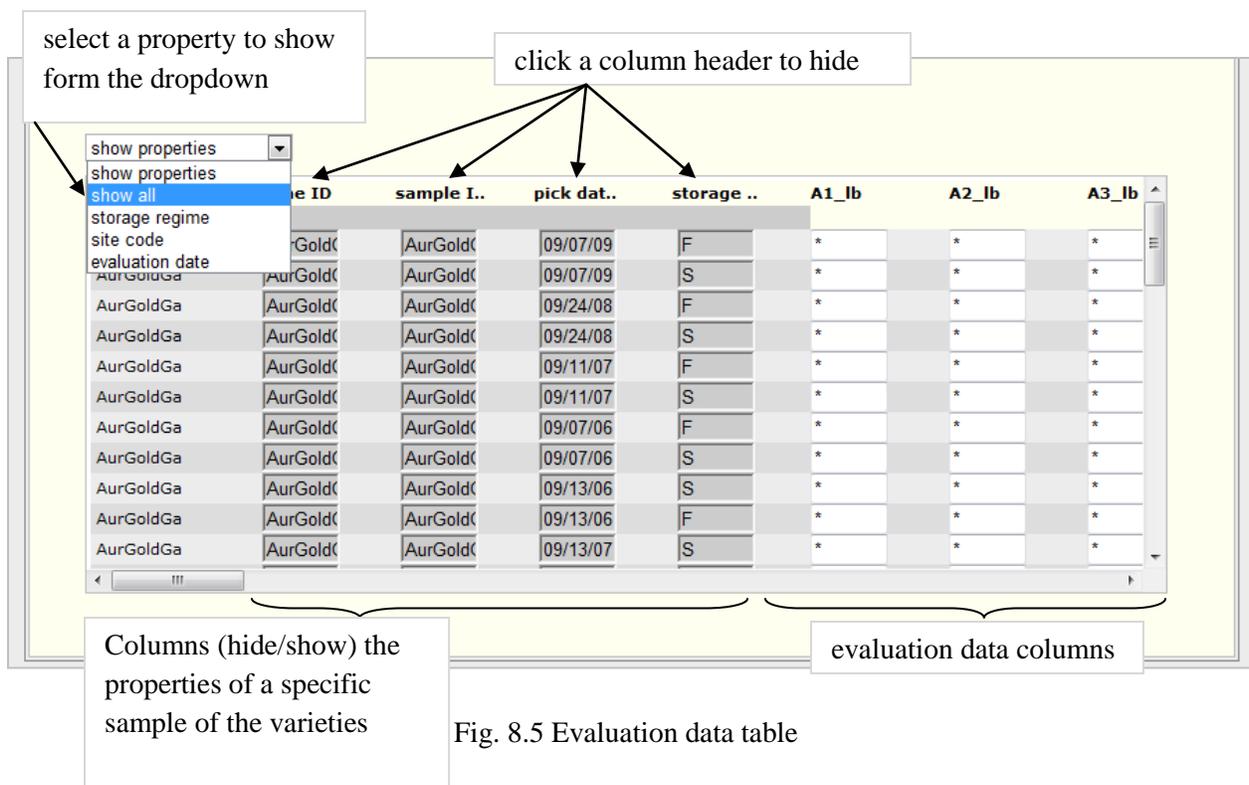
Select values on the selected phenotypes

Scroll left to see other evaluation data

Fig. 8.4 Example of generating a Pedimap input file

1. Variety table

In the example, five phenotypes has been selected and updated in the "*Phenotype selector*". The table has variety names and evaluation data of the selected phenotypes. Initially, default values of these phenotypic evaluation data are assigned. The default values are the average of all the recorded values for the specific variety and the trait.



2. Evaluation data table

The radio button of "AurGoldGa" in the Variety table has been selected. The evaluation data of "AurGoldGa" are listed in the Evaluation data table. The left-most column of the Evaluation data table is variety name, followed by sample properties and evaluation data. (Fig. 8.5) The columns for sample properties can be hidden by clicking the headers. The hidden columns are listed in the dropdown and can be displayed again by selecting property in the dropdown.

The Evaluation data table is used to change the values in the variety table. The first row of the table shows default values. The next rows show the phenotypic values of each specific sample of the variety. A user can click ratio button to select one of the values in each phenotype column. All the phenotypic values recorded for the specific sample of the variety are also shown to help users to decide which trait value to select.

* Users are allowed to change values of the selected phenotypes only.

3. Generating a Pedimap Input file

select phenotypes to be included in Pedimap input file
(phenotypes should be selected to see default values)

NAME	BitterPit	Diam	OvrCol	Location	Acidity
<input type="radio"/> Arlet	3	6.246622	3	1	2.25
<input type="radio"/> AurGoldGa	1	7.24408	4	3	2.75
<input type="radio"/> CrippsPink	1	7.63397	4	3	3
<input type="radio"/> Enterprise	1	8.44296	4	3	3
<input type="radio"/> EsopSpitz	*	*	*	*	*
<input type="radio"/> GoldenDel	*	*	*	*	*
<input type="radio"/> Honeycrisp	2	7.100824	3	1	2.625
<input type="radio"/> Idared	*	*	*	*	*

include evaluation data

generate PEDIMAP input file

File Download

Do you want to open or save this file?

Name: pedimap_input.txt
Type: Text Document, 6.86KB
From: localhost

Open Save Cancel

While files from the Internet can be useful, some files can potentially harm your computer. If you do not trust the source, do not open or save this file. [What's the risk?](#)

```

pedimap_input[1].txt x
; Example of a Pedimap data file. All keywords are shown in UPPERCASE↓
↓
POPULATION = pedimap_input.txt↓
↓
UNKNOWN = *↓
↓
NULLHOMOZ = $ ; this is the default↓
PLOIDY = 2 ; this is the default↓
↓
PEDIGREE↓
NAME FEMALE MALE BitterPit Diam OvrCol Location Acidity↓
"Arlet" "GoldenDel" "Idared" 3.000000 6.246622 3.000000 1.000000 2.250000↓
"AurGoldGa" * * 1.000000 7.244080 4.000000 3.000000 2.750000↓
"CrippsPink" * * 1.000000 7.633970 4.000000 3.000000 3.000000↓
"Enterprise" * * 1.000000 8.442960 4.000000 3.000000 3.000000↓
"GoldenDel" * * * * * * *↓
"Honeycrisp" * * 2.000000 7.100824 3.000000 1.000000 2.625000↓
"Idared" * * "Jonathan" "Wagener" * * * * *↓
    
```

Fig. 8.6 Generating Pedimap input file

After modifying data in the variety table, click "*generate PEDIMAP input file*" to generate an input file for Pedimap. Data in the variety table is embedded in the input file.

If you have any questions or feedback please don't hesitate to contact us via the GDR contact form using the breeders toolbox category. <http://www.rosaceae.org/contact>.

9. Cross Assist

Cross Assist is a decision-support tool for breeders to plan crosses. Cross Assist will determine efficient parent combinations (“crosses”) from among a breeder’s available parents. This efficiency is determined by the number of seedlings that would need to be created to result in a target number of seedlings predicted to each perform within specific trait thresholds. An efficient cross would be one that achieves the trait thresholds with relatively few seedlings.

Cross Assist uses available information on the breeding value of each candidate parent to predict how well resulting seedlings would perform. Available information can be phenotype only, knowledge of breeding values based on experience either subjectively or objectively via calculation of quantitative genetics parameters, or genotype-based breeding values that involves information on functional haplotypes for each parent for known QTLs. To aid the user in comparing between the value of these three information types, three methods of determination of efficient crosses are offered: “*Phenotype*” (uses only phenotypic information of individuals in the dataset), “*+Pedigree*” (adds information provided by pedigree; specifically, estimated breeding value), and “*+Ped+DNA*” (adds information from pedigree and any further information provided by DNA-based functional genotypes). Where such additional information is not available, calculations behind the second and third revert to use of phenotypic information only.

After choosing the calculation method, the user provides two specifications as input: target number of seedlings and target trait thresholds. Cross Assist’s output is then a list of crosses, sorted initially by the estimated number of seedlings that would need to be created to meet the trait thresholds. The preceding description covers the current features of Cross Assist;.

9.1 Front Page

The Front Page enables the user to choose their crop of interest, available dataset of interest, and calculation method.

9.1.1 Page layout

Cross Assist v1 is an online tool to support your crossing decisions. You specify how many seedlings you'd like to create that are predicted to be above/within your specified trait thresholds. Then from among your available parent pool and according to databased knowledge of performance, breeding values, and/or DNA-based functional genotypes, Cross Assist v1 will generate a list of the most efficient parental combinations to achieve your targets.

Cross Assist v1 is the primary data-interfacing module for breeders developed by RosBREED's Breeding Information Management System Team to make routine marker-assisted breeding a reality for U.S. rosaceous crops.

Choose Crop Apple

Parental Dataset APPLE CRS

Choose Method

Phenotype

+Pedigree

+Ped+DNA

Select Parents

Fig. 9.1 Layout of Front Page of Cross Assist

9.1.2 Structure (Fig. 9.1):

- ① Choose Crop: for users to choose a crop.
- ② Choose Dataset: for users to choose a dataset of all candidate parents and example seedling populations, containing phenotypes at a minimum and containing breeding values and functional haplotypes for use of advanced calculations. All datasets to which the user has access are listed.
- ③ Choose Calculation Method: for users to choose one of the three calculation methods
- ④ Select Parents: for users to submit choices and proceed to Targets Page

9.1.3 Steps to use Front Page

1. Choose the crop of interest from the dropdown list, e.g. Apple, Tart Cherry. Only a single crop can be specified at a time.
2. Choose the dataset containing parents of interest from among those available to the users permissions. For example, the Apple CRS is available for anyone to select, but private data can only be selected by that dataset owner(s). The dataset contains potential parents as well as a population of individuals for which the phenotypic distribution is representative of that expected from random crossing among parents. For example, a dataset composed only of cultivars and advanced selections may contain all potential parents a user is interested in but is unlikely to represent accurate phenotypic distributions of resulting seedlings. A dataset composed of cultivars and numerous seedling populations from multiple populations (i.e. a typical breeding dataset) is well suited to Cross Assist analysis. The dataset providers have specified which individuals are possible parents (*Parent pool*) and which are to be used for variation estimates (*Seedling pool*).
3. Choose the calculation method: *Phenotype*, *+Pedigree*, or *+Ped+DNA*. The method estimates two parameters from the dataset for each user-specified trait: a mean (μ) and variance (σ^2) of the phenotypic distribution for a hypothetical seedling population resulting from each possible parental combination ($P_i \times P_j$). For each specified trait, μ and σ^2 are then used to project a Normal distribution (current default), and the probability that a data point (individual seedling) is within/above/below a specified trait threshold is calculated. Across multiple traits, the joint probability of meeting specified thresholds is the product of probabilities for individual traits.

$$\text{Number of Seedlings Required} = \text{Target Number of Seedlings} / \text{Prob(J)}$$

$$\text{Prob(J)} = \prod_{i=1}^{i=n} P_i$$

where...

Prob(J): joint probability of a seedling meeting the thresholds for multiple chosen traits,

P_i : probability of trait i ,

n : number of traits,

Phenotype: Uses only phenotypic information of individuals in the dataset. For each trait of interest, the parameter μ is calculated as the phenotypic average of the two parents of any considered cross. The parameter σ^2 is the same for all parental combinations and is calculated as the variation (standard deviation) of the phenotypes of the specified *Seedling pool*.

$$\text{Prob(J)} = \prod_{i=1}^{i=n} P_i = \prod_{i=1}^{i=n} (f(x_2; \mu, \sigma^2) - f(x_1; \mu, \sigma^2))$$

where...

Prob(J): joint probability of a seedling meeting the thresholds for multiple chosen traits,

$$f(x; \mu, \sigma^2) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \text{ (for Normal distribution, the default distribution),}$$

x_2 : upper threshold,

x_1 : lower threshold,

if only lower threshold is defined, then upper threshold is $+\infty$ by default; if only upper threshold is defined, lower threshold is $-\infty$. Note that $f(+\infty; \mu, \sigma^2) = 1$, $f(-\infty; \mu, \sigma^2) = 0$.

+Pedigree: Uses breeding value estimates based on pedigree information, where available. For each trait of interest, μ is calculated as the average of the mean-adjusted breeding values of the two parents of any considered cross. The “mean” to which breeding value estimates are anchored is the mean of the specified *Seedling pool*. The parameter σ^2 is the same as for the **Phenotype** method above. If pedigree information is not available for a specified trait, parameter estimates revert to the **Phenotype** method for that trait.

+Ped+DNA: Uses DNA-based functional genotypic information, in addition to breeding value estimates based on pedigree information, where available. Considering DNA information available for a single QTL influencing a trait of interest, four values of μ and σ^2 are determined as the mean-adjusted effects of the four possible genotypes resulting from a cross. The “mean” to which functional genotype effects are anchored is the mean of the specified *Seedling pool*. For multiple QTLs, μ and σ^2 are determined from joint functional genotype effects for as many functional genotypic categories as the DNA information discriminates and characterizes. The probability that a data point meets a specified trait threshold is then determined for each of the projected distributions of functional genotypes, and each such probability multiplied by their expected frequencies and summed across all genotypes. If DNA information is not available for a specified trait, parameter estimates for that trait revert to the **+Pedigree** method, or revert to the **Phenotype** method if no pedigree information is available.

For a biparental cross ($H_1H_2 \times H_3H_4$) of four haplotypes (H_1H_3 , H_1H_4 , H_2H_3 , and H_2H_4):

$$\begin{aligned} \text{Prob(J)} &= \prod_{i=1}^{t=n} P_i = \prod_{i=1}^{t=n} \sum_{j=1}^4 \text{freq}_j P_{ij} \\ &= \prod_{i=1}^{t=n} \sum_{j=1}^4 (\text{freq}_j (f(x_2; \mu_{ij}, \sigma_{ij}^2) - f(x_1; \mu_{ij}, \sigma_{ij}^2))) \end{aligned}$$

where...

Prob(J): joint probability of a seedling meeting the thresholds for multiple chosen traits,
j: haplotypes number. e.g. H_1H_3 (*j* = 1), H_1H_4 (*j* = 2), H_2H_3 (*j* = 3), H_2H_4 (*j* = 4).

freq_j: frequency of haplotype *j*,

P_{ij}: probability of haplotype *j* having trait *i*,

x_2 : upper threshold,

x_1 : lower threshold,

μ_{ij} : mean value for haplotype *j*,

σ_{ij}^2 : variance of haplotype *j*.

4. Click the "Select Parents" button to proceed to the "Targets Page".

9.2 Targets Page - Input

The Targets Page is where the user specifies their parameters of interest: the target number of seedlings and traits of interest with their target levels.

9.2.1 Page layout

The screenshot displays the 'Targets Page' interface with the following components:

- 1 Target Number of Seedlings:** Input field containing '100'.
- 2 Method:** Set to 'Phenotype'. Options include:
 - Phenotyping Data
 - +Pedigree
 - +Ped+DNA
- 3 Description of Phenotypes:** Information icon.
- 4 Choose Traits (Continuous):**
 - Button: 'Choose Traits' | Label: 'Select Quantitative (Continuous) Phenotypes'
 - Panel for 'Weight' trait:

SS	mean	min	max	std	#dp
13.04	7.75	19.90	1.56	347	

Weight	mean	min	max	std	#dp
182.88	17.40	574.25	68.24	374	

 Units: ° (degree) | grams
 Range: >= 12 | <=> 180-300
 - Panel for trait selection:
 - BitterPit
 - CalyxO
 - Crack
 - GrndCol
 - MoldyCore
 - PercRuss
 - RussLoc2
 - RussLoc4
 - Shape
 - SunBurn
 - BluStriCol
 - CoreO
 - Grease
 - IntBrown
 - OvrCol
 - RussLoc1
 - RussLoc3
 - Scald
 - Shrivel
 - WaterCore
 - Covariate:**
 - Block
 - Location
 - Row
 - StoReqim
 - update button
- 5 Choose Traits (Discontinuous):**
 - Button: 'Choose Traits' | Label: 'Select Quantitative (Discontinuous) Phenotypes'
 - Panel for 'BitterPit' trait:

mean	min	max	std	#dp
1.43	1.00	4.00	0.92	370

 Range: <= 1
 - Panel for 'Crisp' trait:

mean	min	max	std	#dp
2.67	1.00	4.25	0.73	374

 Range: >= 3
 - Codes:
 - BitterPit: 1 : none, 2 : slight one fruit, 3 : medium 2 or 3 fruit, 4 : severe 4 or 5 fruit
 - Crisp: 1 : no noise, 2 : low noise, 3 : medium, 4 : noisy, 5 : very noisy
- 6 Histogram of OAH:**
 - Number of Data Points: 293
 - Number of Intervals: 50
 - Interval width: 1.979
 - Max Frequency: 22

Buttons at the bottom: 'Generate Cross Efficiency List' and 'Reset'.

Fig.9.2 Layout of Targets Page of Cross Assist

9.2.2 Structure (Fig. 9.2)

- ① Target Number of Seedlings: for users to specify their target number of seedlings.
- ② Method: indicates the calculation method chosen by the user on the Front Page. The icons shown here also appear next to trait names in the trait input forms where such information is available for that trait – e.g., in Fig. 9.2, the traits SS and TA have underlying phenotypic and pedigree information in the dataset, while BitterPit has only phenotypic information, and Crisp has phenotypic, pedigree, and DNA information.
- ③ Description of Phenotypes: user can click the icon () to generate a pop-up window with a short text description and units of each trait in the dataset.
- ④ ⑤ Choose Traits: for users to choose traits of interest (via a pop-up window) and specify thresholds for each trait within the trait input forms. Traits are placed in two categories: either Quantitative (Continuous) or Quantitative (Discontinuous). Each trait input form includes a short statistical description about trait, with mean, minimum (*min*), maximum (*max*), standard deviation (*std*), the number of data points (*#dp*), and an icon to view the histogram and several further data descriptions. Units in which a trait is recorded are displayed directly for continuous traits or codes are displayed for discontinuous traits. The Range section provides a dropdown menu of comparison operators and a text box for users to set trait thresholds.
- ⑥ Generate Cross Efficiency List: for users to submit choices and proceed to the Cross Efficiency Results Page.
Reset: for users to reset all choices on the Targets Page.

9.2.3 Steps to use Targets Page

1. Enter the target number of seedlings meeting trait thresholds specified at step 3 below that a breeder would want to have to then be able to select for other traits of interest. The default number is 100.
2. Click "Choose Traits" button to proceed to a pop-up window that lists available traits, and click the boxes next to each trait to be considered for in efficiency comparisons among possible parental combinations. Click "update" after all choices are made. To subsequently adjust choices, simply return to this pop-up window by again clicking "Choose Traits".
3. Choose the desired seedling performance range for each trait. Clicking on the histogram icon () generates a pop-up window displaying a histogram for the trait. In the Range section of each trait input form, click on the dropdown menu to choose a comparison operator other than the default of ">=" (greater than or equal to). The operators "<>" and "<=>" are for specifying a range that has a minimum and a maximum. Type into the textbox the desired threshold(s) for each trait. The text format for a range is two numerical values separated by an en dash ("–"; a hyphen or em dash, -, will also work) between two spaces. For example, to specify a range of between 1.25 and 5.7, type "1.25 - 5.7".
4. Click the "Generate Cross Efficiency List" button for Cross Assist to perform the necessary calculations and to proceed to the Cross Efficiency Results Page.

9.2.4 Example

Leave "Target Number of Seedlings" on default value of 100. Choose traits of SS (soluble solids), Weight, BitterPit, and Crisp. Set thresholds of SS ≥ 12 °Brix, Weight 180 to 300 g, BitterPit ≤ 1 (i.e., "none" only), and Crisp ≥ 3 (i.e., "medium", "noisy", and "very noisy") (as in Fig. 9.3). Click "Generate Cross Efficiency List". (This example is continued in 9.3.4.)

The screenshot displays two panels for setting trait targets. The top panel, titled "Quantitative (Continuous)", shows settings for "SS" and "Weight". The bottom panel, titled "Quantitative (Discontinuous)", shows settings for "BitterPit" and "Crisp".

Traits	SS	Weight
Stats	mean: 13.04, min: 7.75, max: 19.90	mean: 182.88, min: 17.40, max: 574.25
	std: 1.56, #dp: 347	std: 68.24, #dp: 374
Units	° (degree)	grams
Range	≥ 12	$\in [180, 300]$

Traits	BitterPit	Crisp
Stats	mean: 1.43, min: 1.00, max: 4.00	mean: 2.67, min: 1.00, max: 4.25
	std: 0.92, #dp: 370	std: 0.73, #dp: 374
Range	≤ 1	≥ 3
Codes	1 : none 2 : slight one fruit 3 : medium 2 or 3 fruit 4 : severe 4 or 5 fruit	1 : no noise 2 : low noise 3 : medium 4 : noisy 5 : very noisy

Fig. 9.3 An example of using the Targets Page, choosing thresholds for traits of "SS", "Weight", "BitterPit", and "Crisp"

9.3 Cross Efficiency Results Page - Output

The Cross Efficiency Results Page displays the results of Cross Assist's calculations of cross efficiency.

9.3.1 Page layout

[Go Back to Targets Page](#)

Choices Made

Parental Dataset	APPLE_CRS	Quantitative (continuous)	SS	: >= 12
Target # of Seedlings	100		Weight	: <=> 180-300
# of Parents	47	Quantitative (discontinuous)	BitterPit	: <= 1
# of Pairs of Parents	1081		Crisp	: >= 3
Method	Phenotype			

Progress Bar 100%

[Estimate # Seedlings](#)

Cross List

# of Seedlings Required	Mother	Father	Formula
481	Cox	MN1914	f_x
481	Delicious	MN1914	f_x
	Empire	MN1914	f_x
	Fiesta		f_x
	Fuji		f_x
	Gala		f_x
	GoldenD		f_x
481	Goodlan		f_x
481	Jonafree		f_x
481	Jonathar		f_x
481	Keepsal		f_x
481	MN1914		f_x
481	MN1914		f_x
481	McIntosl		f_x
504	Ambrosia		f_x
521	Enterpri		f_x
543	Nicola		f_x
559	MN1914		f_x
598	MN1914		f_x
635	Crimscrip	MN1914	f_x
640	MN1702	MN1914	f_x
668	Ambrosia	Nicola	f_x
675	Hatsuaki	MN1914	f_x
691	Ambrosia	Coop15	f_x
691	Ambrosia	Cox	f_x

Sorted # of Seedlings Required by reverse order

Number of Seedlings Required = Target Number of Seedlings * 1 / Prob(J)

481 = 100 x 1 / (

SS : 0.99
= [1.0 - NORMDIST(12.00, 15.93, 1.56)]
*mean : 15.93 = (ND / 15.93) / 2

Weight : 0.49
= [NORMDIST(300.00, 186.88, 68.24) - NORMDIST(180.00, 186.88, 68.24)]
*mean : 186.88 = (ND / 186.88) / 2

BitterPit : 0.50
= [NORMDIST(1.00, 1.00, 0.92)]
*mean : 1.00 = (ND / 1.00) / 2

Crisp : 0.85
= [1.0 - NORMDIST(3.00, 3.75, 0.73)]
*mean : 3.75 = (ND / 3.75) / 2

4
[First](#)
[Prev](#)
1 2 3 4 5 6 7 8 9 10
[Next](#)
[Last](#)

Fig. 9.4 Cross Efficiency List

9.3.2 Structure (Fig 9.4)

- ① Choices Made: for users to view a summary of choices made on the Front Page and Targets Page.
- ② Cross List table: lists all possible parent combinations in default order of the estimated number of seedlings required to achieve the target number of seedlings that meet the specified trait target thresholds. The table can be sorted by any column, in ascending or descending order for “# of Seedlings Required” and alphabetical or reverse alphabetical order for “Mother” and “Father”.
- ③ The formula and data used in the calculation for each cross can be reached as a pop-up window via the “ f_x ” symbol in the last column of the Cross List table. The probability of achieving the target threshold(s) for each trait is shown in bold; e.g., in Fig. 9.4 for SS the probability is 0.99 in the cross between ‘Cox’ and ‘MN1914’. The display includes the mean of the two parents of each cross that is used in the calculation; e.g., in Fig. 9.4 for Weight the mean is 186.88, which in this case was calculated from only one of the two parents (‘MN1914’) as indicated by the “ND” notation for one of the parents. When no data is available for both parents of a cross, the resulting # of Seedlings Required is returned as “n/a” and in the calculation window there is a red dot at the start of the formula. If there are some cases of “ND”, a yellow dot is displayed (as in the Fig. 9.4 inset). If there is no missing data for both parents, the dot is green.
- ④ Each page shows 25 entries at a time; further entries can be accessed via the page navigation list.

9.3.3 Steps to view Cross Assist results

1. Identify the most efficient cross(es) by observing Mother-Father combination with the smallest number for “# of Seedlings Required” – which will be the first cross listed. The number listed represents the estimated number of seedlings required to achieve the target number of seedlings that meet the specified trait target thresholds; e.g., in Fig 9.4 the first cross of ‘Cox’ × ‘MN1914’ is estimated to require 481 seedlings to achieve 100 seedlings meeting the specified trait thresholds.
2. To adjust the displayed order of crosses, click on the column headings. Clicking twice will reverse the order (ascending/descending, alphabetical/reverse alphabetical).
3. Interrogate the calculation behind the # of Seedlings Required for a particular cross by clicking on the “ f_x ” symbol in the last column to reveal a pop-up window.

9.3.4 Example

Following the example introduced in 9.2.4, after clicking "Generate Cross Efficiency List", the Cross Efficiency List Page will appear (Fig. 9.5). Click "Estimate # Seedling" and wait until the Progress Bar reaches 100%, when a “computation complete” box will appear (click “OK”). The Cross Efficiency Table will then be populated. In this example, the first 14 crosses all have an estimated required seedling number of 481, and all have ‘MN1914’ as a parent. View the underlying calculation by clicking on f_x in the "Formula" column for the first cross, ‘Cox’ x ‘MN1914’. The calculation for all of these 14 crosses is the same as in Fig. 9.4.



Fig. 9.5 An example Cross Efficiency Results

When multiple traits are chosen and two or more crosses have the same outcome and share a parent, it is likely that there is missing data for one of the parents (not the common parent) for at least one of the traits. In such cases, the probability for a trait is only based on the parent that has underlying data – effectively, the result mimics a self of that parent. In the phenotypic data underlying the example here, only ‘MN1914’ has data for all four chosen traits; the other 14 parents do not have records in the database for all of these traits. (Next versions of Cross Assist will indicate such crosses and allow sorting according to complete or levels of missing data or for crosses with missing data to be

discarded from consideration.) ‘MN1914’ clearly has desirable phenotypes for the chosen traits because not only does it feature effectively as a self in the first 14 crosses, but it is also a parent of the next seven crosses.

View the underlying calculation for the cross ‘Ambrosia’ × ‘MN1914’, which has a result of 504 seedlings, by clicking on f_x in the "Formula" column. The resulting pop-up window (Fig. 9.6) indicates that both parents have phenotypic data in the dataset for all four chosen traits (because there is a green dot at the start of the formula and “ND” does not appear for any of the individual trait probability calculations). The probability of this cross providing seedlings that meet the threshold of ≥ 12 °Brix is estimated to be 0.95, from the calculation of $1.0 - \text{NORMDIST}(12.00, 14.51, 1.56)$, where 12.00 is the threshold, 14.51 is the mean of the two parents (‘Ambrosia’ = 13.10, ‘MN1914’ = 15.93), and 1.56 is the variance of all individuals in the underlying dataset’s designated seedling pool. Following the same logic, the probability that this cross provides seedlings meeting specified thresholds for Weight, BitterPit, and Crisp are 0.55, 0.5, and 0.75, respectively. The joint probability that seedlings of the cross meet thresholds for all four traits is the product of the individual probabilities, i.e., $0.95 * 0.55 * 0.5 * 0.75$, which is approximately 0.196 – about one in five. Therefore, to obtain 100 seedlings that meet all thresholds, a breeder would need to create about 500 seedlings from this cross (504 exactly).

Incidentally, for the next best cross, ‘Enterprise’ × ‘MN1914’, with a result of a slightly less efficient 521 seedlings than for the previous cross, the trait levels of the other parent ‘Enterprise’ are different to ‘Ambrosia’ for three of the four traits. ‘Enterprise’ is slightly lower for SS, slightly greater for Weight, the same for BitterPit, and slightly lower for Crisp, which provides the slightly increase in the estimated required seedling number when crossed with ‘MN1914’.



Fig. 9.6 An example Cross Assist calculation