IMPORTING HISTORICAL DATA INTO BMS

DATA IMPORT TOOL allows users to upload historical datasets generated from other software into Breeding Management System (BMS). This can be Nursery or Trial field data.

Contents

Obtaining GIDs to insert into the Data File 1

Import Germplasm List 1

Export germplasm list to retrieve GIDs 5

Obtaining Location_IDs to insert into the Data File 5

Obtaining Person_IDs to insert into the Data File 6

Format the Historical Data 7

Data import from an excel file using the excel wizard 7

Structure of an Excel File for data import 7

Add the data for factors and variates 8

Run the Data Import Application 8

Map data columns to ontology variables 9

Import the data values 11

Data import from a Fieldbook formatted file 12

Import the Germplasm List 12

Obtain a Fieldbook Template 12

Format the file for a single location 12

Import the File using the Fieldbook Wizard 14

Loading data for multiple locations in the same fieldbook file 16

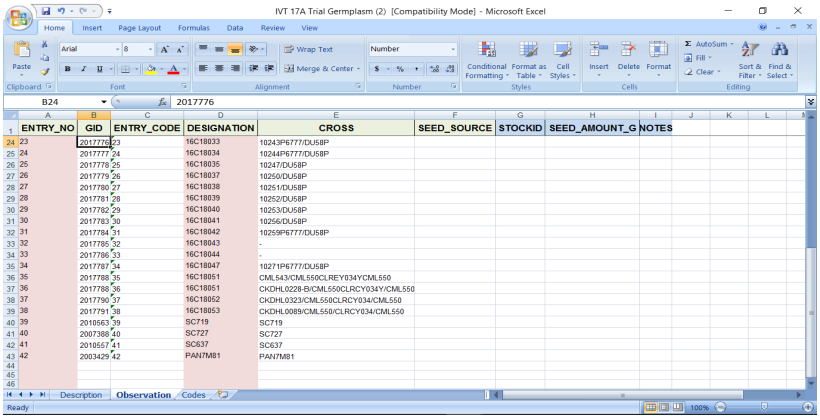
Using the VLOOKUP function to copy GIDs into an excel file 17

Obtaining GIDs to insert into the Data File

Import Germplasm List

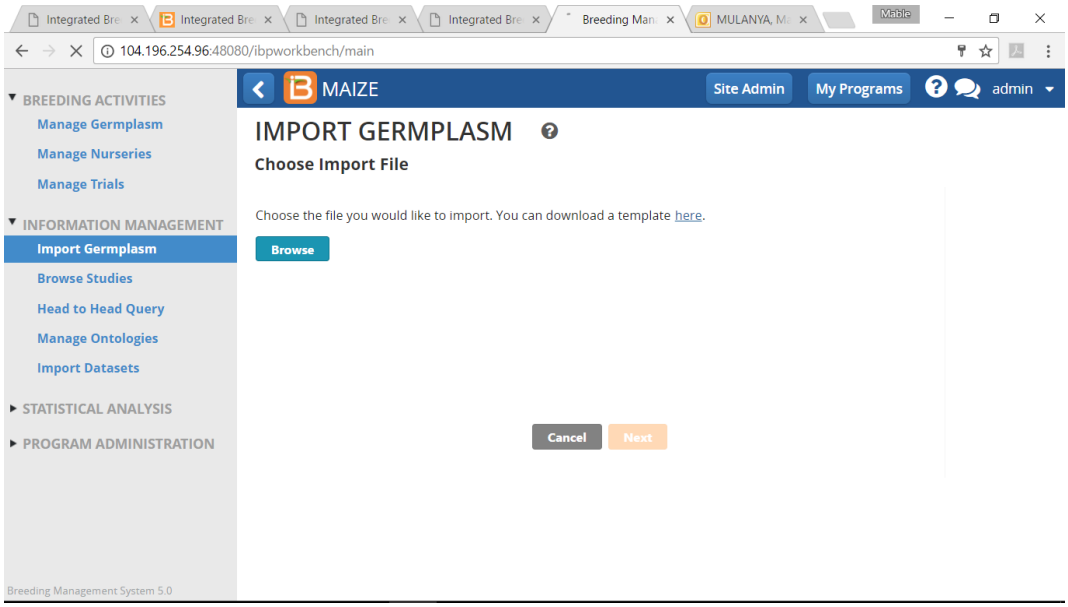
- Before importing any field data it's important to note that the **germplasm used in the trial or nursery must be allocated with GIDs** (Germplasm Identifiers) and these must appear in the data file. If the germplasm does not have GIDs, full BMS database integration requires that you first import the germplasm list into the database and assign GIDs to all germplasm and then enter those GIDs in the data file before importing data.

- Sort historical data by ENTRY NO per replication to identify germplasm used in the trial or nursery. **NOTE:** You only need germplasm entries of one replication to make a germplasm list to avoid duplicated GIDs. In this example we have 42 germplasm entries which have been replicated twice, therefore the total number of entries in the germplasm list will be 42 **(Refer to IVT 17A Trial Germplasm list.xls).**



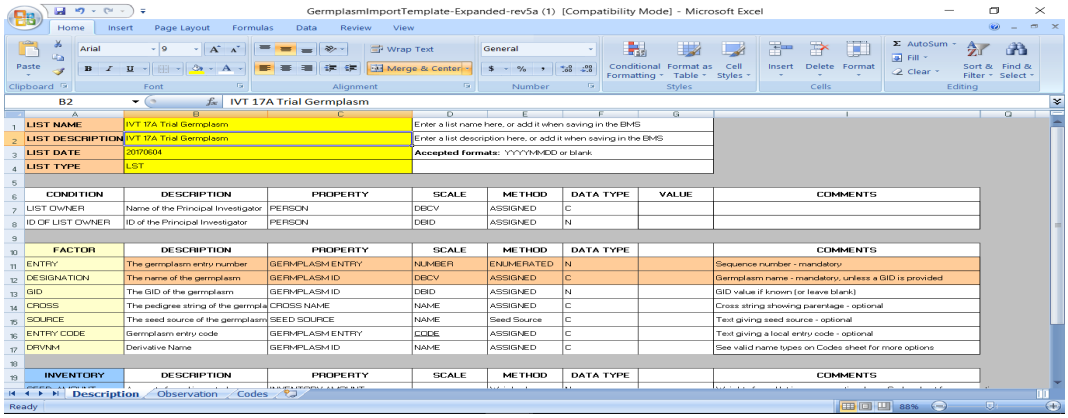
ENTRY_NO	GID	ENTRY_CODE	DESIGNATION	CROSS	SEED_SOURCE	STOCKID	SEED_AMOUNT_G	NOTES
23	2017776	23	16C18033	10243P6777/DU58P				
24	2017777	24	16C18034	10244P6777/DU58P				
25	2017778	25	16C18035	10247/DU58P				
26	2017779	26	16C18037	10250/DU58P				
27	2017780	27	16C18038	10251/DU58P				
28	2017781	28	16C18039	10252/DU58P				
29	2017782	29	16C18040	10253/DU58P				
30	2017783	30	16C18041	10256/DU58P				
31	2017784	31	16C18042	10259P6777/DU58P				
32	2017785	32	16C18043	-				
33	2017786	33	16C18044	-				
34	2017787	34	16C18047	10271P6777/DU58P				
35	2017788	35	16C18051	CML543/CML550CLREY034YCML550				
36	2017789	36	16C18051	CKDHL0228-B/CML550CLRCY034YCML550				
37	2017790	37	16C18052	CKDHL0323/CML550CLRCY034/CML550				
38	2017791	38	16C18053	CKDHL0089/CML550/CLRCY034/CML550				
39	2010563	39	SC719	SC719				
40	2007388	40	SC727	SC727				
41	2010557	41	SC637	SC637				
42	2003429	42	PAN7M81	PAN7M81				

- You need a BMS germplasm template to import any germplasm. Download BMS template from **IMPORT GERMPLASM** under **INFORMATION MANAGEMENT** section.

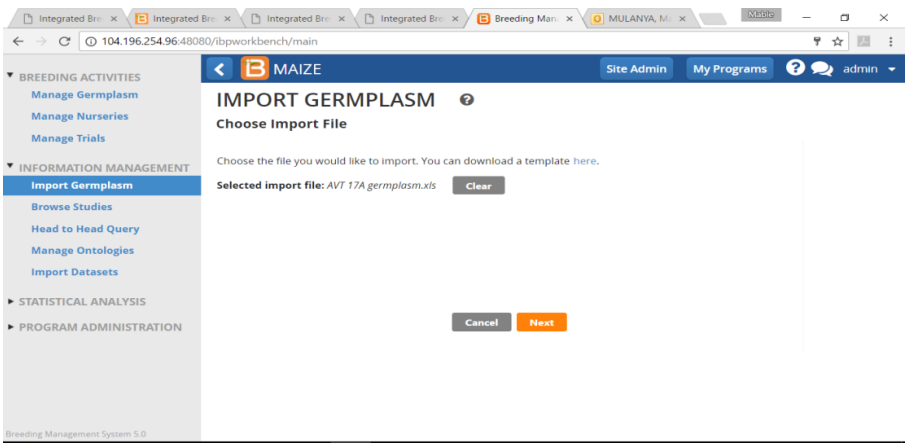


- The germplasm template (.xls) has three sheets:
- **Description**-Has meta data about the list
- **Observation**- Has germplasm details where ENTRY NO and DESISGNATION (germplasm name) are mandatory
- **Codes**- Provides optional column headers to customize germplasm details e.g the DVRNM appearing in the column header of observation sheet refer to derivative name in codes sheet.

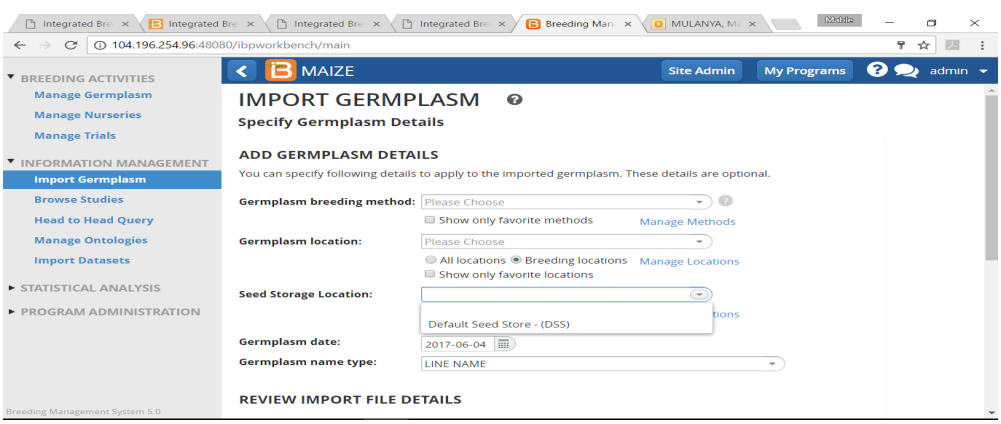
- In Description sheet, fill the details of germplasm name, description and the date the list was created.



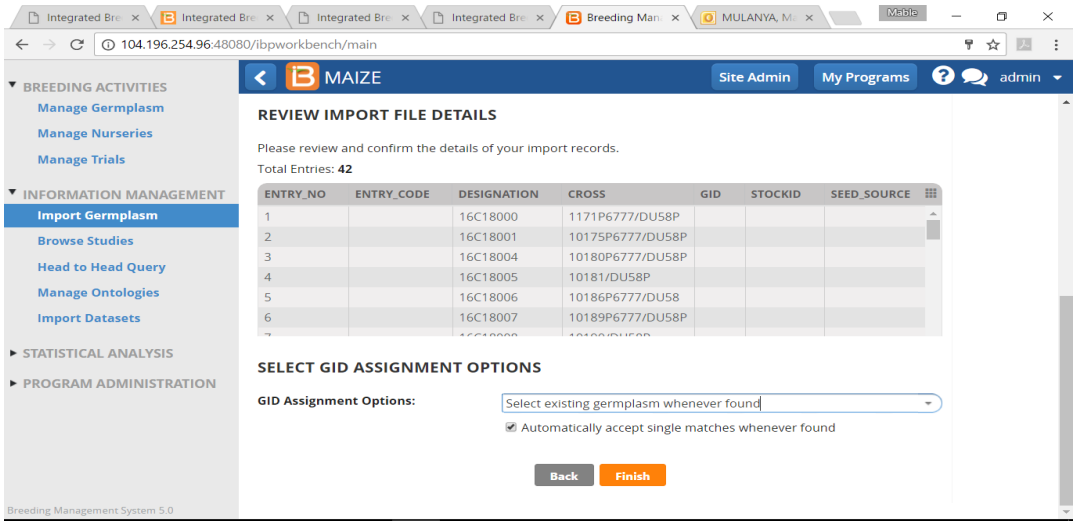
- When you have completed filling in details in the germplasm list, **SAVE** and return to BMS.
- Back to the BMS**>Information management>Import germplasm**. **Browse** to germplasm import file(.xls), upload the **GERMPLASM LIST (**Example file:*AVT 17A Germplasm.xls***)** and select **NEXT**.



- Fill in the appropriate details concerning the germplasm list using the dropdown arrows.

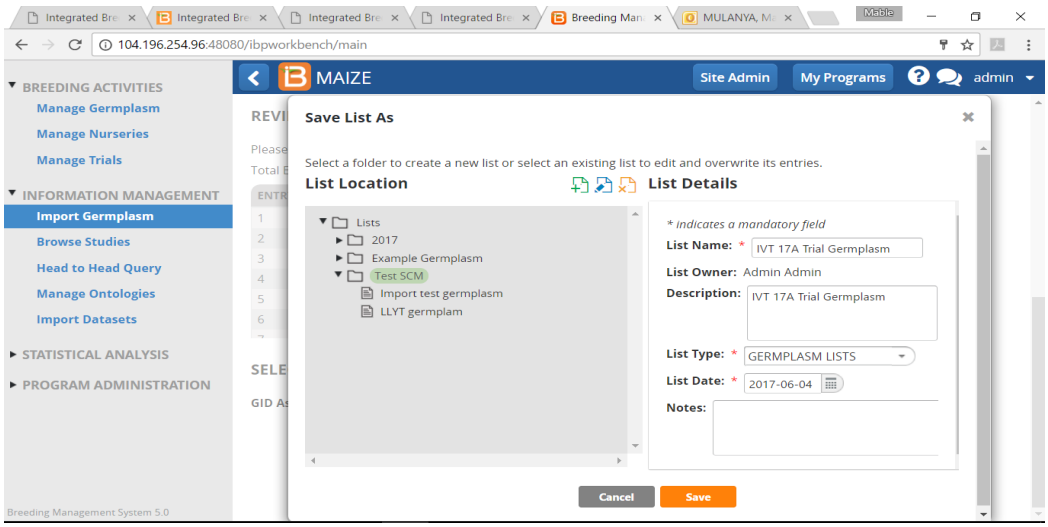


- There are two options when assigning GIDs:
- Assigning new GIDS-adding all entries to new records and assigning new GIDs
- Matching existing GIDs- Selecting existing germplasm whenever found to match Germplasm names in the import list to the ones existing in the database. It's advised to use this option to avoid duplication of GIDs.
- Select existing germplasm whenever found option and **FINISH**



- In case the germplasm in the import list matches with more than one in the database then you will have to select for the appropriate name from the existing list and select for the match to be used for other instances found in the import list.

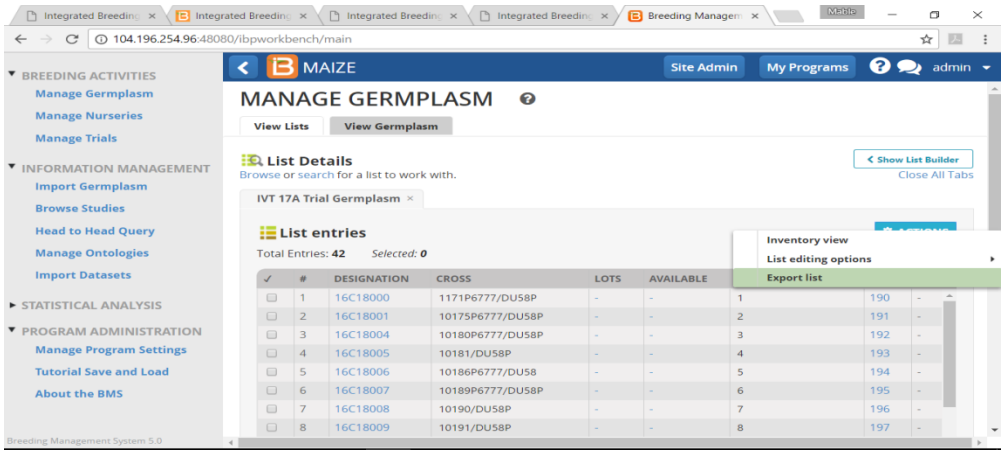
- Select continue, highlight the folder you will use and **SAVE** the list.



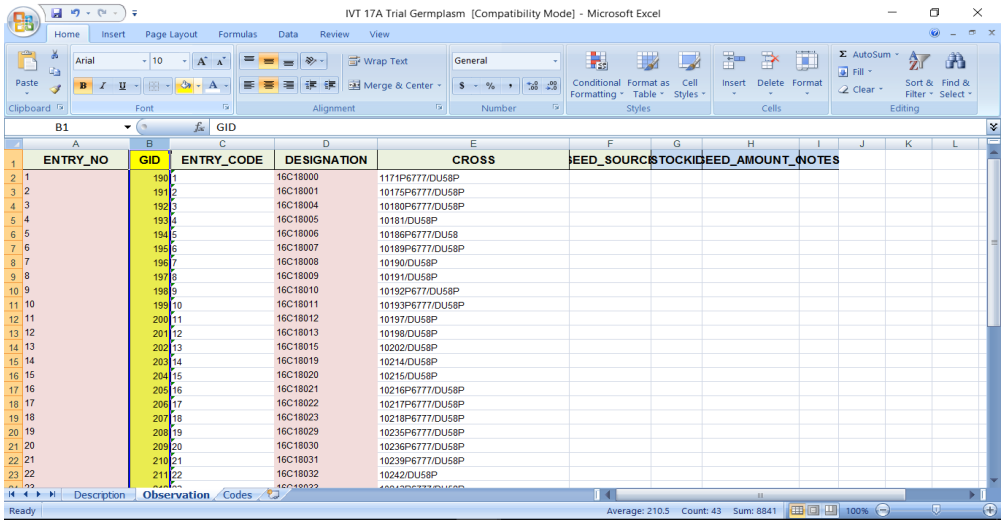
Export germplasm list to retrieve GIDs

The imported germplasm list must be exported from BMS to retrieve the GIDS.

- Select Manage lists from the menu. Browse to the imported germplasm list and open it. Close the selector window to view the list.
- From the actions menu select export list and You **MUST LOCK THE LIST** with the padlock icon appearing under the **ACTIONS** before you can export it.



- From the exported germplasm list, Select the observation sheet to view the GIDs.



Obtaining Location_IDs to insert into the Data File

It is important to have LOCATION_NAMEs and IDs in the data import files if the data is to be correctly assigned to environments. However the current Data Import tool does no look up location IDs so you need to find the IDs by looking directly in the Locations table of the database or by setting up a template trial with all the locations you intend to use (you may have to add some new ones) and then looking in the ENVIRONMENTS dataset of the template using the Browse Studies feature from the Information Management menu:

Browse or search for a study to work with.

AVT200 16-17 ×

Study Details

Factors

Variates

Datasets

Dataset of AVT200 16-17-ENVIRONMENT

TRIAL_INSTANCE	LOCATION_NAME	LOCATION_ID
1	RARS EARLY	20001
4	KRC EARLY	20005
8	SAVE VALLEY	11067
10	CHIREDZI	11045
11	MATOPOS	20013
12	CHISUMBANJE	20015
28	BALAKA (RS)	20065
29	SAVE VALLEY IRRIGATION	20088
30	CHIREDZI IRRIGATION	20012

For frequent use it is desirable to maintain a file of Location Names and IDs.

Obtaining Person_IDs to insert into the Data File

You may want to specify the PI_NAME (if using Fieldbook Wizard) or the COOPERATOR for each location if using either import method. To load this data you need to enter the persons name exactly as in the BMS and also the persons ID (in variables PI_NAME_ID or COOPERATOR_ID). To find these IDs you need to create a template study (not using the name of the study you intend to import) and add the person you want to find the ID of as PI or COOPERATOR and use Browse Studies to look in the FACTORS section for the PI name and ID:

Factors

NAME	DESCRIPTION	PROPERTY	SCALE	METHOD	DATATYPE	VALUE
STUDY_TYPE	Study type	Study	Type of STUDY_TYPE	Assigned	Categorical	Trial
TRIAL_INSTANCE	Trial instance - enumerated (number)	Trial instance	Number	Enumerated	Numeric	
START_DATE	Start date	Start date	Date (yyyymmdd)	Assigned	Date	20170602
LOCATION_NAME	Location - selected (DBCV)	Location	Location name	Assigned	Location	
GID	Germplasm identifier - assigned (DBID)	Germplasm id	Germplasm id	Assigned	Germplasm List	
END_DATE	End date	End date	Date (yyyymmdd)	Assigned	Date	20171231
LOCATION_ID	Location - selected (DBID)	Location	Location id	Assigned	Location	
DESIGNATION	Germplasm identifier - assigned (DBCV)	Germplasm id	Germplasm name	Assigned	Germplasm List	
PI_NAME	Principal investigator - assigned (DBCV)	Person	Person name	Assigned	Person	Christopher McLaren
COOPERATOR	COOPERATOR NAME	Person	Person name	Conducted	Person	
ENTRY_NO	Germplasm entry - enumerated (numb	Germplasm ent	Number	Enumerated	Numeric	
PI_NAME_ID	Principal investigator - assigned (DBID)	Person	Person id	Assigned	Person	6
COOPERATOR_ID	COOPERATOR ID -Assigned (DBID)	Person	Person id	Conducted	Person	
CROSS	The pedigree string of the germplasm	Cross history	Text	Assigned	Character	

Or in the ENVIRONMENT dataset to fine the name and ID of COOPERATORs

Dataset of AVTYY600-Meaning-ENVIRONMENT

TRIAL_INSTANCE	LOCATION_NAME	LOCATION_ID	COOPERATOR	COOPERATOR_ID	CROP_SEAS
1	RARS EARLY	20001	Christopher McLaren	6	2017 B sea
3	RARS LATE	20003	Christopher McLaren	6	2017 B sea
4	KRC EARLY	20005	Christopher McLaren	6	2018 A sea
8	ART MEDIUM	20010	Christopher McLaren	6	2017 B sea
15	LUSAKA WEST	11530	Christopher McLaren	6	2017 B sea
22	GART	20082	Christopher McLaren	6	2017 B sea
23	KAFUE	20078	Christopher McLaren	6	2018 A sea

It may be useful to retain a file of the persons and their IDs for easy reference.

Format the Historical Data

There are **TWO** options of file formats which can be used for historical data import into BMS:

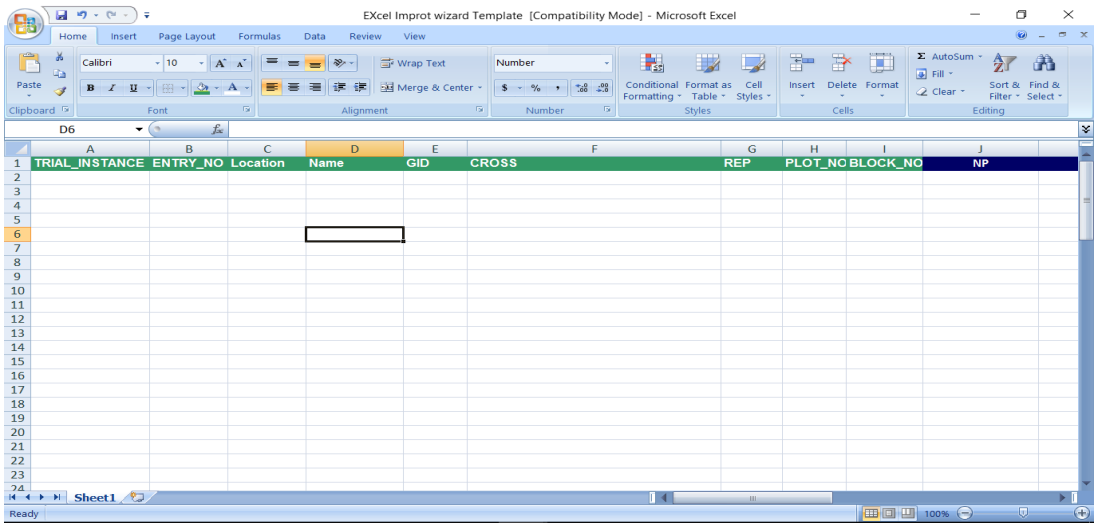
1. Use of Excel import wizard format is the most flexible format for importing historical data. It requires one sheet containing factors and variates for the dataset to be imported. There are **FOUR** mandatory columns: **TRIAL_INSTANCE, ENTRY_NO, PLOT_NO** and **GID** to import phenotypic observations. However, it is desirable to import **LOCATION_NAME** and **LOCATION_ID** as well. You can also include additional descriptive data such as environment codes, season codes, entry names, and design factors. While this is the most flexible format it requires manual mapping of factors and traits.

2. Use of excel Field book format (**Fieldbook format file.xls**). This format applies to files which have been previously generated from BMS e.g. trials or nurseries and exported out. Such files are perfectly formatted for historical data import into BMS. This format does not require manual mapping and so if you have many similar trials it is a very efficient way to import the data.

Data import from an excel file using the excel wizard.

Structure of an Excel File for data import

- **The Excel import wizard format** for data import reads one file at a time. The column headers highlighted in green are **FACTORS** and those in blue are **VARIATES**. There are **FOUR** mandatory columns: **TRIAL_INSTANCE, ENTRY_NO, PLOT_NO** and **GID.** The colors are just for illustration, and any other column headers can be used for other labels and traits.

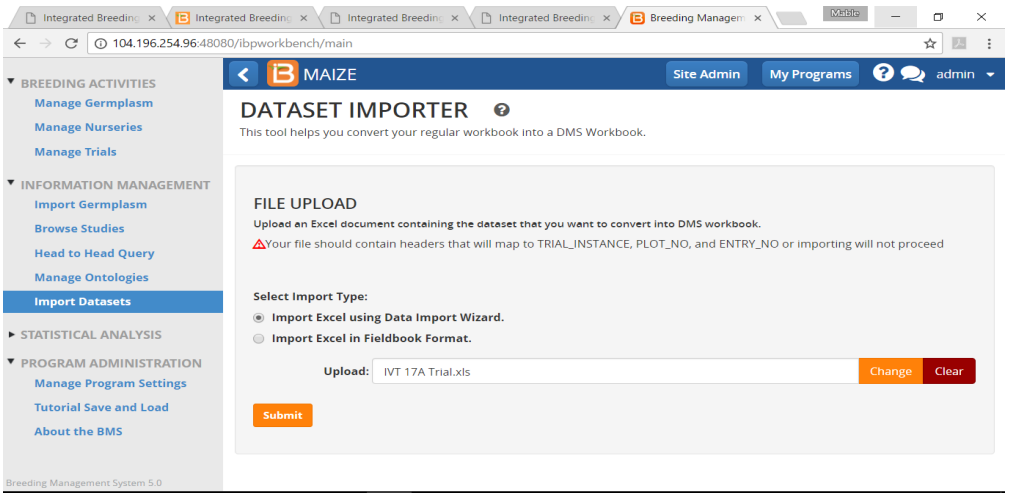


Add the data for factors and variates

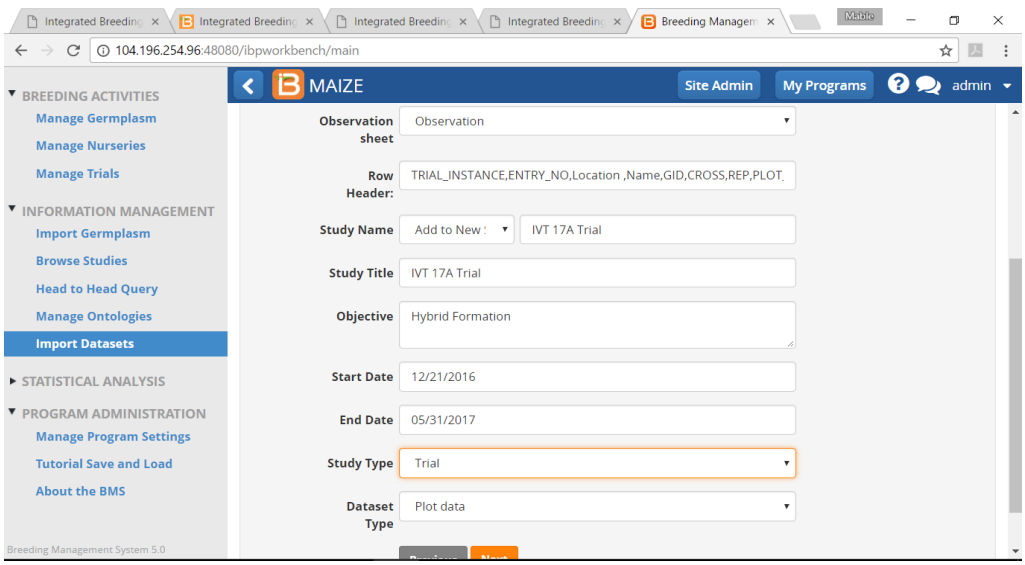
- Copy and paste the information in your trial into the **EXCEL IMPORT WIZARD TEMPLATE** as shown below. The TRIAL_INSTANCE column should contain distinct integers identifying distinct trial sites or environments. The optional column Location could contain location names (repeated for each plot at that location).
- The Name column in the example above could contain the germplasm name or DESIGNATION and mapping the columns will be easier if the column is in fact headed DESIGNATION. The CROSS column could contain a cross string for each entry. Name, GID and CROSS can all be filled from the germplasm list file using the VLOOKUP function as described in the last section of this file. IF the CROSS column is present, no cells can be empty – just add "-", this is a known bug.
- In the column header named **GID**, copy the GIDs from the germplasm list that was exported from BMS. You can use the Excel VLOOKUP function as described in the last section of this file. Ensure that the **ENTRY NO** or **DESIGNATION** (name of germplasm) that appear in the Germplasm list match exactly those of the trial.
- PLOT_NO must just contain integers 1,2 …n where n is the number of plots in each trial instance, for each trial instance. If the Trial design is different from the template format then you will need to edit the factors in the headers. If you have an incomplete block design be sure to use the header BLOCK_NO for the sub-block factor (and not BLK or BLK_NO) – this is a known bug.
- If you intend to add **LOCATION_NAME** and **LOCATION_ID** (which you should), you must add those columns to the data sheet and fill them with the values from your locations file or from the template view (as shown above). Each plot for each location must have the same value of name and ID.

Run the Data Import Application

- Launch the **INFORMATION MANAGEMENT** from BMS then select **DATA IMPORT TOOL**. Select the import **EXCEL USING DATA IMPORT WIZARD** and select an excel file to import and **SUBMIT**.



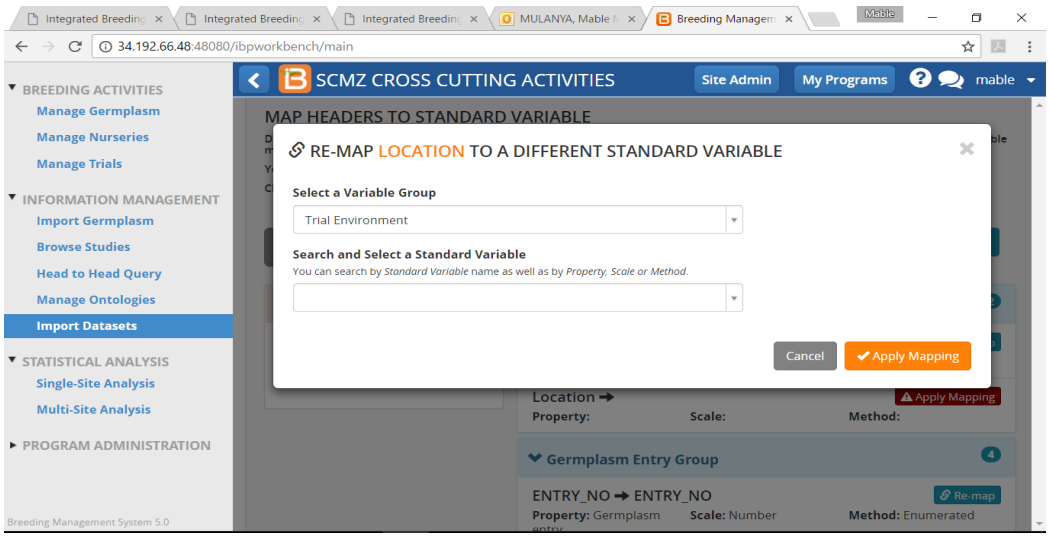
- Specify the observation sheet containing phenotypic data within the excel file. Select the appropriate column header and close. Enter study name, title and type.



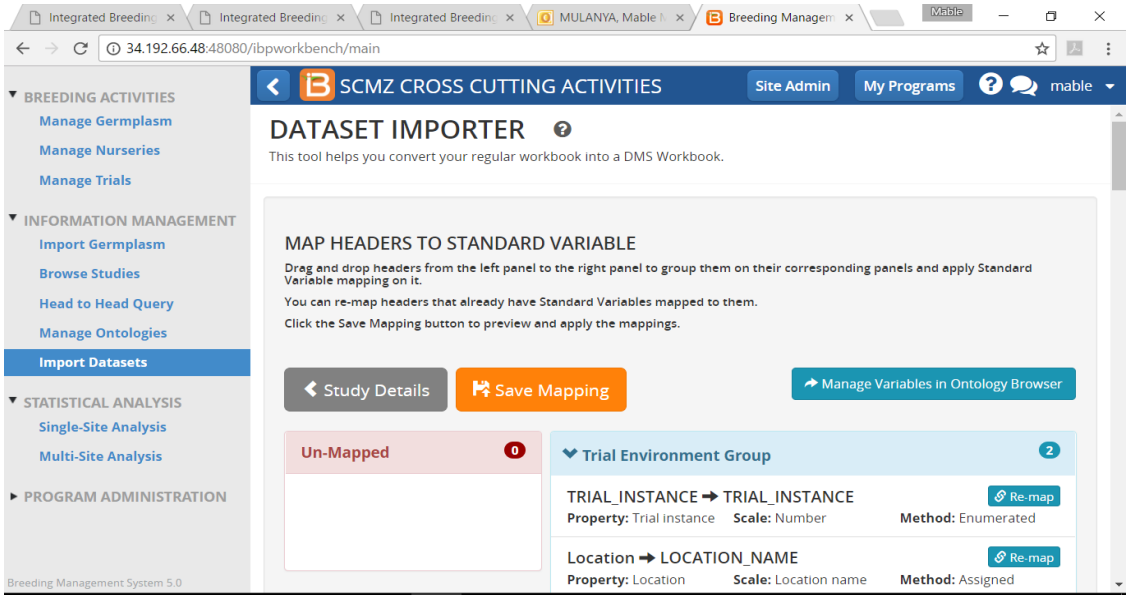
Map data columns to ontology variables

- The BMS will attempt to map the excel sheet column headers to corresponding database terms. Carefully review the mapped terms to ensure that the BMS correctly matched the truly meaning and measurement of the imported data.
- Review the details of the suggested matches by selecting re-map icon and Search for alternative matches by typing in the appropriate field (either **TRIAL ENVIRONMENT, GERMPLASM,TRIAL DESIGN** or **VARIATE GROUP**)

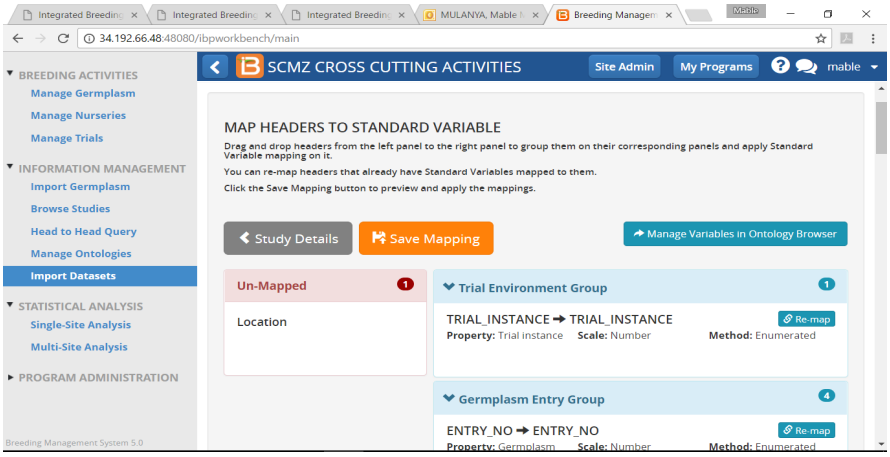
- In this example the term **LOCATION** is mapped under the **TRIAL ENVIRONMENT GROUP**. IF you have added LOCATION_NAME and LOCATION_ID instead they should go to the ENVIRONEMNT group. Similarly Season_Code or management codes should go into this group.



- For the **UN-MAPPED TERMS**, drag and drop them to their appropriate grouping either (**TRIAL ENVIRONMENT, GERMPLASM,TRIAL DESIGN** or **VARIATE GROUP)**.If the alternative term is an actual match ,select apply mapping. If a match is not found, you will have to go back to ontology manager to add the new variable.

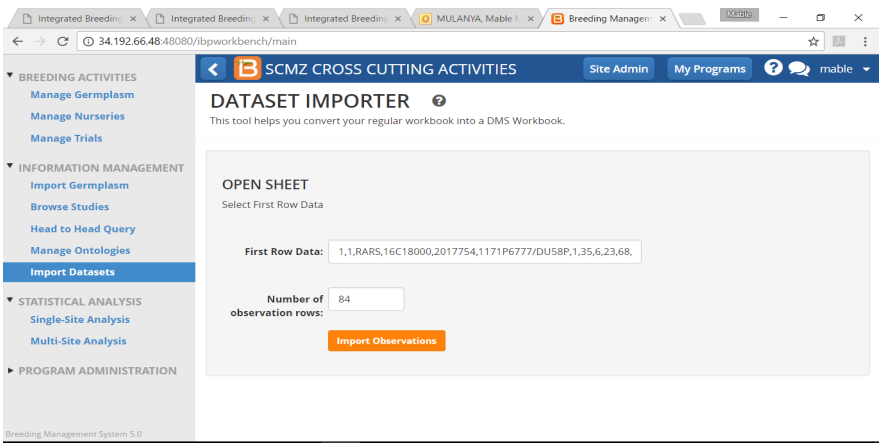


- When mapping is complete, save it and select confirm header mapping.

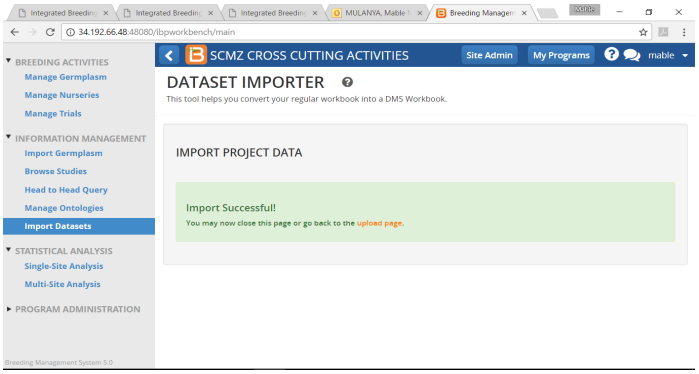


Import the data values

- Select the first row of the data and import the observations. The phenotypic observations for this trial will be saved in the database and you will receive a message for successful importation.



- You will receive a notification if import is successful.



- You can retrieve the trial by selecting Manage Trials from the MENU and BROWSE for the Trial!

Data import from a Fieldbook formatted file

Import the Germplasm List

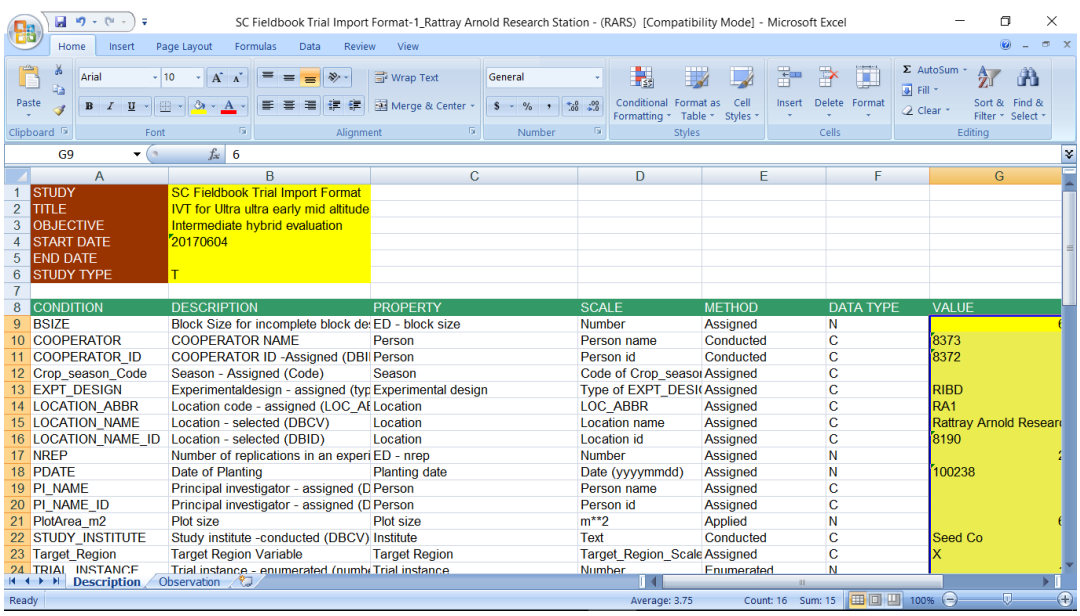
- You have to import a germplasm list into BMS to generate GIDs as indicated above.
- Export the list to get the GIDs to insert into the Fieldbook

Obtain a Fieldbook Template

- The fieldbook format option expects the data to be delivered in a file formatted according to the BMS fieldbook format like files exported from trials already in BMS. The easiest way to obtain such a file is to create a template trial which matches the trial you wish to import in terms of variables – factors and traits. This template should not have the same name as the trial you want to import. Then export the fieldbook from the template as though you were going to collect the data. The fieldbook format has two sheets: Description and Observation sheet.

Format the file for a single location

- Set the Study Name, Title, Objective, Start Date, End Date. The Study Type MUST be set to T for Trial. These values remain the same for subsequent files for each location. Edit the value fields in the CONDITION and CONSTANT sections on the **Description sheet** to reflect the information for the environment you are loading. In particular you can enter the LOCATION_NAME and LOCATION_ID in the CONSTANT section.
- You must delete the row describing PLOT_ID from the description sheet, and the corresponding column from the observation sheet. You can similarly delete rows and corresponding columns for any variables for which you do not have data.

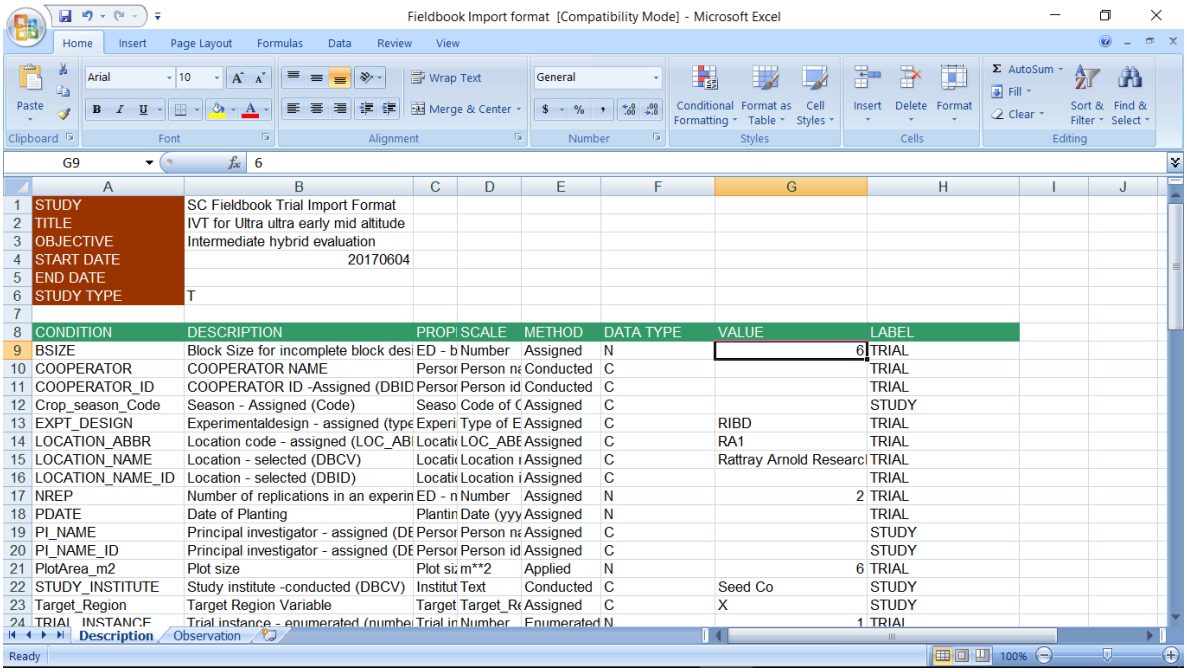


	A	B	C	D	E	F	G
1	STUDY	SC Fieldbook Trial Import Format					
2	TITLE	IVT for Ultra ultra early mid altitude					
3	OBJECTIVE	Intermediate hybrid evaluation					
4	START DATE	20170604					
5	END DATE						
6	STUDY TYPE	T					
7							
8	CONDITION	DESCRIPTION	PROPERTY	SCALE	METHOD	DATA TYPE	VALUE
9	BSIZE	Block Size for incomplete block de	ED - block size	Number	Assigned	N	
10	COOPERATOR	COOPERATOR NAME	Person	Person name	Conducted	C	8373
11	COOPERATOR_ID	COOPERATOR ID -Assigned (DBI	Person	Person id	Conducted	C	8372
12	Crop_season_Code	Season - Assigned (Code)	Season	Code of Crop_seaso	Assigned	C	
13	EXPT_DESIGN	Experimentaldesign - assigned (typ	Experimental design	Type of EXPT_DESI	Assigned	C	RIBD
14	LOCATION_ABBR	Location code - assigned (LOC_A	Location	LOC_ABBR	Assigned	C	RA1
15	LOCATION_NAME	Location - selected (DBCV)	Location	Location name	Assigned	C	Rattray Arnold Resear
16	LOCATION_NAME_ID	Location - selected (DBID)	Location	Location id	Assigned	C	8190
17	NREP	Number of replications in an experi	ED - nrep	Number	Assigned	N	
18	PDATE	Date of Planting	Planting date	Date (yyyymmdd)	Assigned	N	100238
19	PI_NAME	Principal investigator - assigned (D	Person	Person name	Assigned	C	
20	PI_NAME_ID	Principal investigator - assigned (D	Person	Person id	Assigned	C	
21	PlotArea_m2	Plot size	Plot size	m**2	Applied	N	
22	STUDY_INSTITUTE	Study institute -conducted (DBCV)	Institute	Text	Conducted	C	Seed Co
23	Target_Region	Target Region Variable	Target Region	Target_Region_Scale	Assigned	C	X
24	TRIAL_INSTANCE	Trial instance - enumerated (numb	Trial instance	Number	Enumerated	N	

- Copy the GIDS using **VLOOKUP FUNCTION** from the germplasm list into a column header **GID** in the trial observation sheet as described in the last section of this file. You can also copy DESIGNATION and CROSS in this way if required.

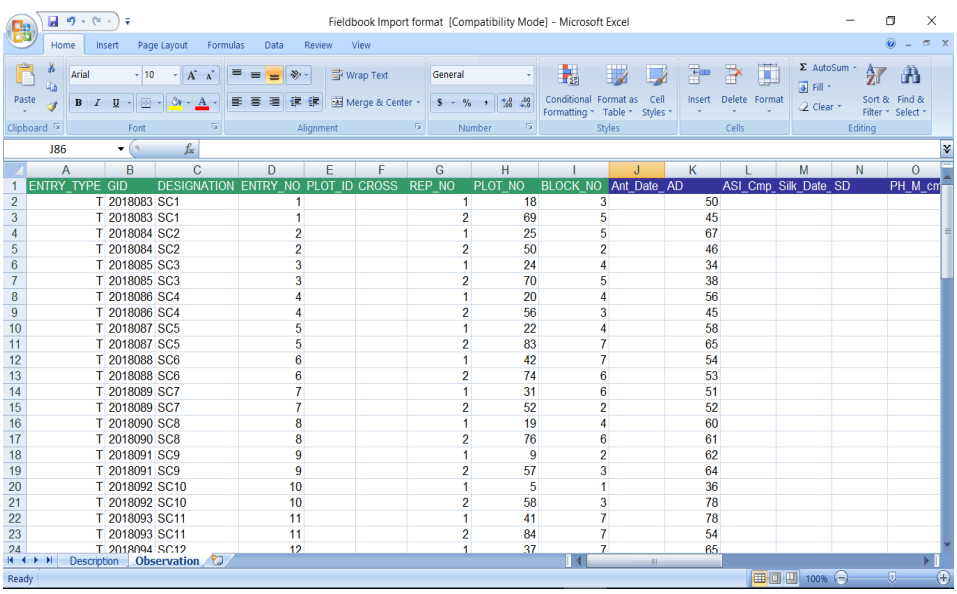
- Copy the rest of the data from the historical data file into the observation sheet of the fieldbook. Save the file and review the trial to confirm that the mandatory column headers (**TRIAL_INSTANCE, ENTRY_NO, PLOT_NO and GIDs**) are present. A complete fieldbook formatted trial with the two sheets is as below

Description sheet



1	STUDY	SC Fieldbook Trial Import Format						
2	TITLE	IVT for Ultra ultra early mid altitude						
3	OBJECTIVE	Intermediate hybrid evaluation						
4	START DATE	20170604						
5	END DATE							
6	STUDY TYPE	T						
7								
8	CONDITION	DESCRIPTION	PROP	SCALE	METHOD	DATA TYPE	VALUE	LABEL
9	BSIZE	Block Size for incomplete block des	ED - b	Number	Assigned	N	6	TRIAL
10	COOPERATOR	COOPERATOR NAME	Persor	Person na	Conducted	C		TRIAL
11	COOPERATOR_ID	COOPERATOR ID -Assigned (DBID	Persor	Person id	Conducted	C		TRIAL
12	Crop_season_Code	Season - Assigned (Code)	Seaso	Code of C	Assigned	C		STUDY
13	EXPT_DESIGN	Experimentaldesign - assigned (type	Experi	Type of E	Assigned	C	RIBD	TRIAL
14	LOCATION_ABBR	Location code - assigned (LOC_AB	Locatic	LOC_ABE	Assigned	C	RA1	TRIAL
15	LOCATION_NAME	Location - selected (DBCV)	Locatic	Location i	Assigned	C	Rattray Arnold Researc	TRIAL
16	LOCATION_NAME_ID	Location - selected (DBID)	Locatic	Location i	Assigned	C		TRIAL
17	NREP	Number of replications in an experin	ED - n	Number	Assigned	N	2	TRIAL
18	PDATE	Date of Planting	Plantin	Date (yyy	Assigned	N		TRIAL
19	PI_NAME	Principal investigator - assigned (DE	Persor	Person na	Assigned	C		STUDY
20	PI_NAME_ID	Principal investigator - assigned (DE	Persor	Person id	Assigned	C		STUDY
21	PlotArea_m2	Plot size	Plot si	m**2	Applied	N	6	TRIAL
22	STUDY_INSTITUTE	Study institute -conducted (DBCV)	Institut	Text	Conducted	C	Seed Co	STUDY
23	Target_Region	Target Region Variable	Target	Target_Re	Assigned	C	X	STUDY
24	TRIAL_INSTANCE	Trial instance - enumerated (numbe	Trial in	Number	Enumerated	N	1	TRIAL

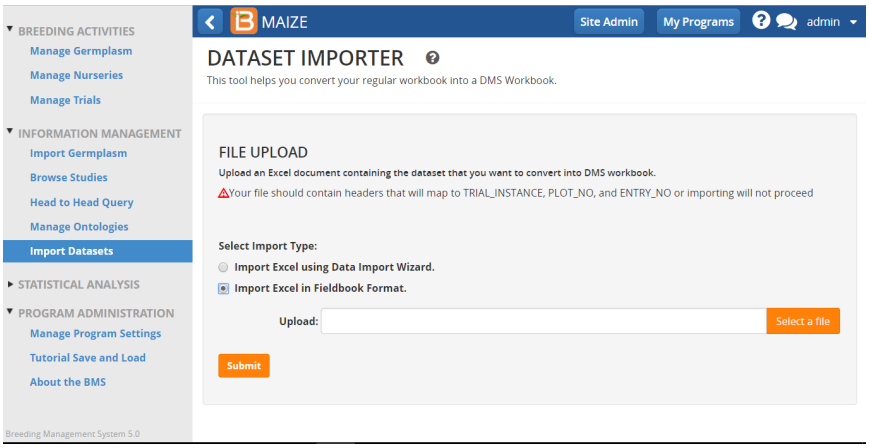
Observation Sheet



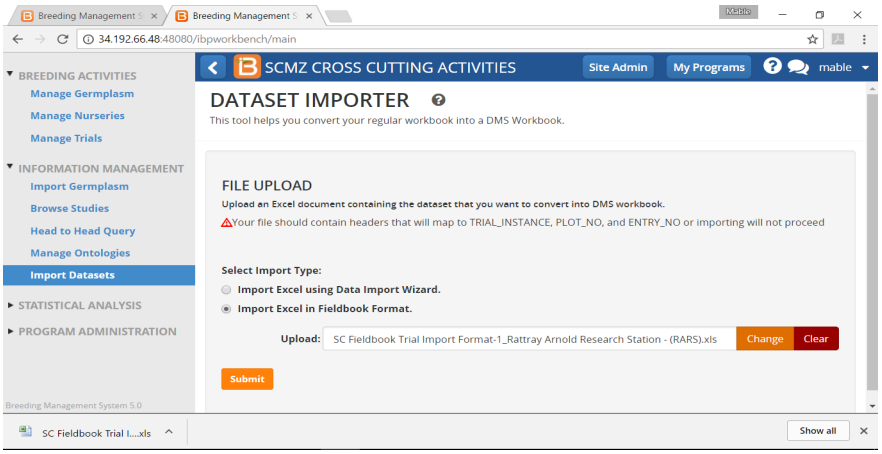
	ENTRY_TYPE	GID	DESIGNATION	ENTRY_NO	PLOT_ID	CROSS	REP_NO	PLOT_NO	BLOCK_NO	Ant_Date_AD	ASI_Cmp	Silk_Date	SD	PH_M_cm
2	T	2018083	SC1	1			1	18	3	50				
3	T	2018083	SC1	1			2	69	5	45				
4	T	2018084	SC2	2			1	25	5	67				
5	T	2018084	SC2	2			2	50	2	46				
6	T	2018085	SC3	3			1	24	4	34				
7	T	2018085	SC3	3			2	70	5	38				
8	T	2018086	SC4	4			1	20	4	56				
9	T	2018086	SC4	4			2	56	3	45				
10	T	2018087	SC5	5			1	22	4	58				
11	T	2018087	SC5	5			2	83	7	65				
12	T	2018088	SC6	6			1	42	7	54				
13	T	2018088	SC6	6			2	74	6	53				
14	T	2018089	SC7	7			1	31	6	51				
15	T	2018089	SC7	7			2	52	2	52				
16	T	2018090	SC8	8			1	19	4	60				
17	T	2018090	SC8	8			2	76	6	61				
18	T	2018091	SC9	9			1	9	2	62				
19	T	2018091	SC9	9			2	57	3	64				
20	T	2018092	SC10	10			1	5	1	36				
21	T	2018092	SC10	10			2	58	3	78				
22	T	2018093	SC11	11			1	41	7	78				
23	T	2018093	SC11	11			2	84	7	54				
24	T	2018094	SC12	12			1	37	7	65				

Import the File using the Fieldbook Wizard

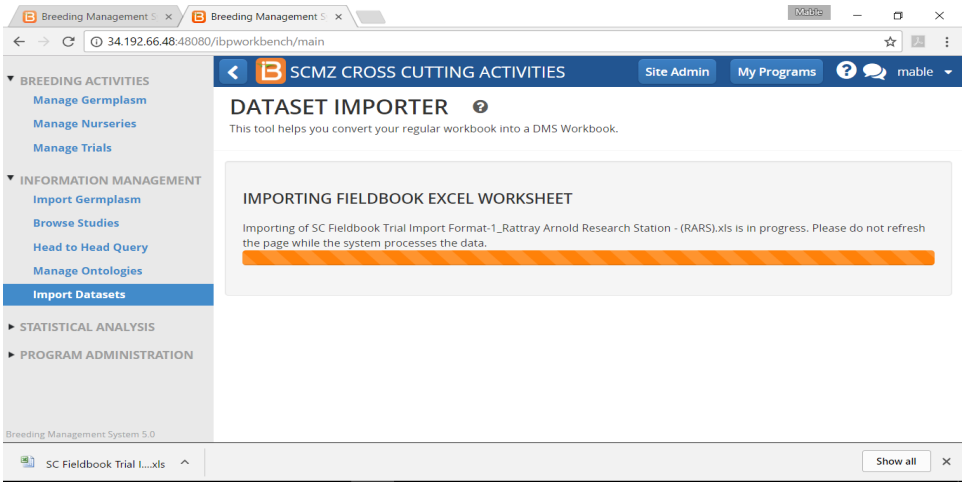
- In the BMS, select **IMPORT DATASETS** from the INFORMATION MANAGEMENT. Select, **IMPORT EXCEL IN FIELDBOOK FORMAT.**



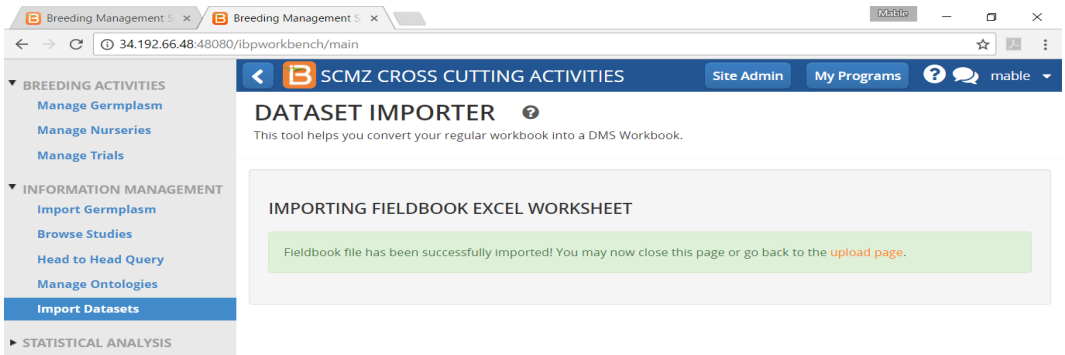
- Select the fieldbook formatted file and upload it.



- BMS will automatically map the terms in the fieldbook formatted trial with the database.

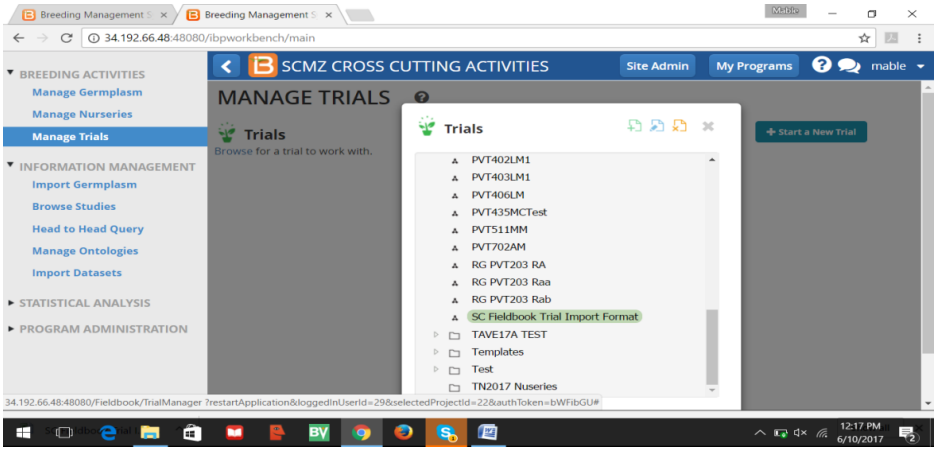


- When import is successful you will receive a notification.

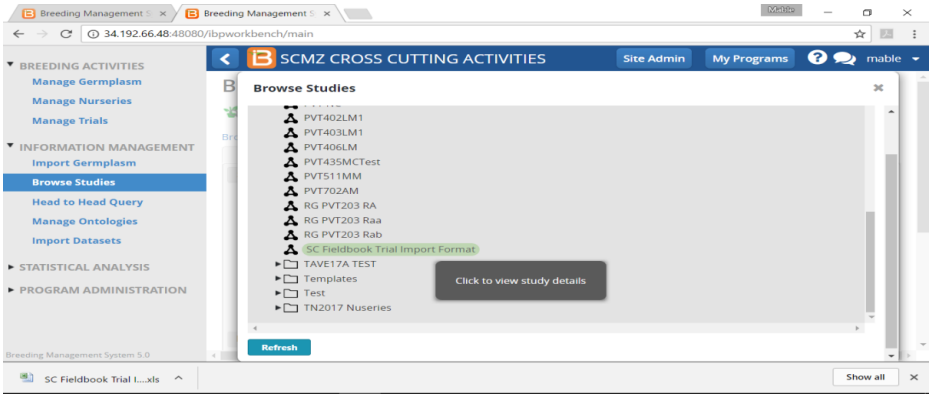


Retrieve dataset

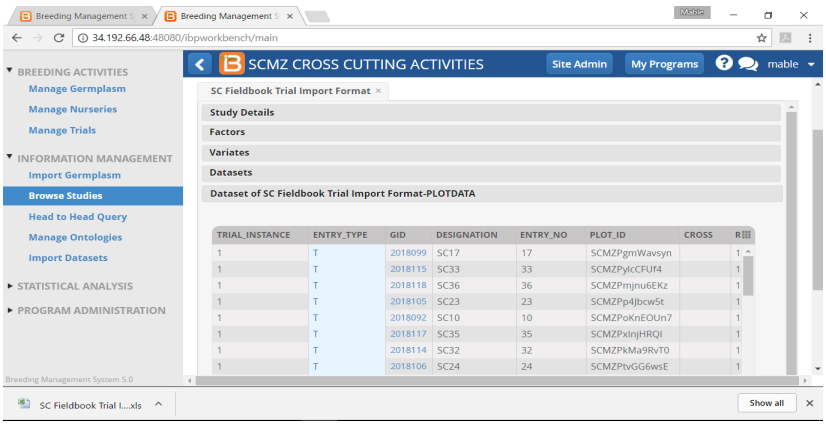
- You can add subsequent locations to the same trial by editing the CONDITION and CONSTANT sections of the Description sheet and pasting the GIDs and data for the next location in the observation sheet. Leave the Study name and other Study level variables unchanged.
- Retrieve the fieldbook formatted trial from the **MANAGE TRIALS** and **BROWSE** to select the file.



- You can also retrieve the trial from the **BROWSE STUDIES** option under the **INFORMATION MANAGEMENT** and select the trial**.**



- Close the Browse studies window and select **DATASETS** and then **PLOT DATA** to view the Trial.



Loading data for multiple locations in the same fieldbook file

You can load data for multiple locations in the same fieldbook file but you need to change the structure of the template.

- Move rows describing TRIAL labels (as indicated in column H) from the CONDITION section of the Description sheet to the FACTOR section. (Leave only rows describing STUDY conditions)

	A	B	C	D	E	F	G	H
8	CONDITION	DESCRIPTION	PROPERTY	SCALE	METHOD	DATA TYPE	VALUE	LABEL
9	PI_NAME	Principal investigator - assig	Person	Person name	Assigned	C	Tavengwa Ndowa	STUDY
10	PI_NAME_ID	Principal investigator - assig	Person	Person id	Assigned	C	39	STUDY
11	STUDY_INSTITUTE	Study institute -conducted (I	Institute	Text	Conducted	C	Seed Co	STUDY
12	Target_Region	Target Region Variable	Target Region	Target_Region_Scale	Assigned	C	N	STUDY
13								
14	FACTOR	DESCRIPTION	PROPERTY	SCALE	METHOD	DATA TYPE	VALUE	LABEL
15	TRIAL_INSTANCE	Trial instance - enumerated	Trial instance	Number	Enumerated	N		TRIAL
16	LOCATION_NAME	Location - selected (DBCV)	Location	Location name	Assigned	C		TRIAL
17	LOCATION_ID	Location - selected (DBID)	Location	Location id	Assigned	C		TRIAL
18	COOPERATOR	COOPERATOR NAME	Person	Person name	Conducted	C		TRIAL
19	COOPERATOR_ID	COOPERATOR ID -Assigne	Person	Person id	Conducted	C		TRIAL
20	Crop_season_Code	Season - Assigned (Code)	Season	Code of Crop_seaso	Assigned	C		TRIAL
21	EXPT_DESIGN	Experimentaldesign - assign	Experimental design	Type of EXPT_DESI	Assigned	C		TRIAL
22	NREP	Number of replications in an	ED - nrep	Number	Assigned	N		TRIAL
23	BSIZE	Block Size for incomplete bl	ED - block size	Number	Assigned	N		TRIAL
24	PlotArea_m2	Plot size	Plot size	m**2	Applied	N		TRIAL
25	SEEDING_DATE	Date Seeded - applied (yyy	Planting date	Date (yyyymmdd) of	Applied	N		TRIAL
26	GID	Germplasm identifier - assig	Germplasm id	Germplasm id	Assigned	C		ENTRY
27	DESIGNATION	Germplasm identifier - assig	Germplasm id	Germplasm name	Assigned	C		ENTRY
28	ENTRY_NO	Germplasm entry - enumerat	Germplasm entry	Number	Enumerated	N		ENTRY
29	CROSS	The pedigree string of the ge	Cross history	Text	Assigned	C		ENTRY
30	REP_NO	Replication - assigned (numt	Replication factor	Number	Enumerated	N		PLOT
31	PLOT_NO	Field plot - enumerated (num	Field plot	Number	Enumerated	N		PLOT
32	BLOCK_NO	Block - assigned (number)	Blocking factor	Number	Enumerated	N		PLOT

- Move rows describing TRIAL variables (as indicated in column H) from the CONSTANT section of the Description sheet to the VARIATE section.

	A	B	C	D	E	F	G	H
34	CONSTANT	DESCRIPTION	PROPERTY	SCALE	METHOD	DATA TYPE	VALUE	SAMPLE LEVEL
35								
36	VARIATE	DESCRIPTION	PROPERTY	SCALE	METHOD	DATA TYPE	VALUE	SAMPLE LEVEL
37	SITE_SOIL_PH	Soil acidity - ph meter (pH)	Soil acidity	pH	Measured	N		TRIAL
38	Rainfall	Total rainfall over the growin	Rainfall	mm	Measured	N		TRIAL
39	Irrigation	Total irrigation over the grow	Irrigation	mm	Measured	N		TRIAL
40	Fertilizer	Fertilizer applied during the	Fertilizer use	Text	Applied	C		TRIAL
41	AD	Also Ant_DT_day. Anthesis t	Anthesis time	Day	Days to anthesis - Co	N		PLOT
42	SD	Also Silk_DT_day. Silking tin	Silking time	Day	DTS - Computation	N		PLOT

- Insert columns in the observation sheet and enter the names of the new variables you have just added to the FACTOR and VARIATE sections into the first row. The variable names must be exactly as

they appear in the description sheet and in the same order they appear in the CONDITION and VARIATE sections of that sheet.

	A	B	C	D	E	F	G	H	I	J	K	L
1	TRIAL_INSTANCE	LOCATION_NAME	LOCATION_ID	COOPERATOR	COOPERATOR_ID	Crop_season_Code	EXPT_DESIGN	NREP	BSIZE	PlotArea_m2	SEEDING_DATE	GID
2	8	ART MEDIUM	20010	Tavengwa Ndow	39	2017 B season	RIBD	3	5	5	20170602	2009949
3	8	ART MEDIUM	20010	Tavengwa Ndow	39	2017 B season	RIBD	3	5	5	20170602	2007777

and

R	S	T	U	V	W
BLOCK_N	SITE_SOIL_PH	Rainfall	Irrigation	Fertilizer	AD
1	2.66	108	100	40	69

- Append the data for the concatenated set of locations to the Observation sheet as for the Excel import mode described above. Take care that the location CONDITIONS and CONSTANTS contain the same value for every plot at that location.
- Note I had some trouble with Crop_Season_code where I had to enter the meaning of the code (2017 B season) rather than the code itself (17B). The same is true for Target Region which is a STUDY condition on the description sheet. You can look up the meanings of the valid values in the scales section of the Ontology Browser.

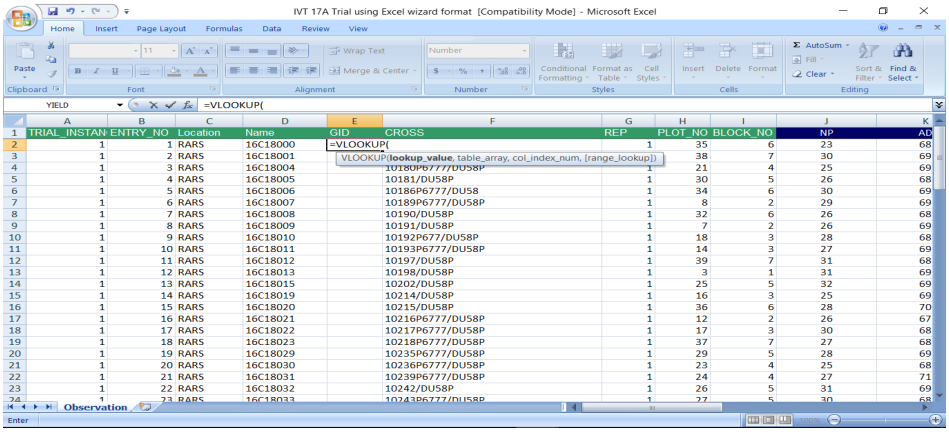
	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	TRIAL_INSTANCE	LOCATION_NAME	LOCATION_ID	COOPERATOR	COOPERATOR_ID	Crop_season_Code	EXPT_DESIGN	NREP	BSIZE	PlotArea_m2	SEEDING_DATE	GID	DESIGNAT	ENTRY_N
2	8	ART MEDIUM	20010	Tavengwa Ndow	39	2017 B season	RIBD	3	5	5	20170602	2009949	12C28564	17
3	8	ART MEDIUM	20010	Tavengwa Ndow	39	2017 B season	RIBD	3	5	5	20170602	2007772	15C3052	7
4	8	ART MEDIUM	20010	Tavengwa Ndow	39	2017 B season	RIBD	3	5	5	20170602	2009950	12C28570	15
5	8	ART MEDIUM	20010	Tavengwa Ndow	39	2017 B season	RIBD	3	5	5	20170602	2007766	15C3062	13
6	8	ART MEDIUM	20010	Tavengwa Ndow	39	2017 B season	RIBD	3	5	5	20170602	2007214	STD3	25
7	8	ART MEDIUM	20010	Tavengwa Ndow	39	2017 B season	RIBD	3	5	5	20170602	2010337	12C28577	20
8	8	ART MEDIUM	20010	Tavengwa Ndow	39	2017 B season	RIBD	3	5	5	20170602	2010332	12C28582	6
9	8	ART MEDIUM	20010	Tavengwa Ndow	39	2017 B season	RIBD	3	5	5	20170602	2007755	15C3051	5
10	8	ART MEDIUM	20010	Tavengwa Ndow	39	2017 B season	RIBD	3	5	5	20170602	2010338	12C28581	21
11	8	ART MEDIUM	20010	Tavengwa Ndow	39	2017 B season	RIBD	3	5	5	20170602	2009946	12C28575	22
12	8	ART MEDIUM	20010	Tavengwa Ndow	39	2017 B season	RIBD	3	5	5	20170602	2007754	15C3050	8
13	8	ART MEDIUM	20010	Tavengwa Ndow	39	2017 B season	RIBD	3	5	5	20170602	2007533	SC402	3
14	8	ART MEDIUM	20010	Tavengwa Ndow	39	2017 B season	RIBD	3	5	5	20170602	2007161	STD2	24
15	8	ART MEDIUM	20010	Tavengwa Ndow	39	2017 B season	RIBD	3	5	5	20170602	2010334	12C28567	16
16	8	ART MEDIUM	20010	Tavengwa Ndow	39	2017 B season	RIBD	3	5	5	20170602	2007160	STD1	23
17	8	ART MEDIUM	20010	Tavengwa Ndow	39	2017 B season	RIBD	3	5	5	20170602	2010411	SC608	1
18	8	ART MEDIUM	20010	Tavengwa Ndow	39	2017 B season	RIBD	3	5	5	20170602	2007775	10C8427	11

and

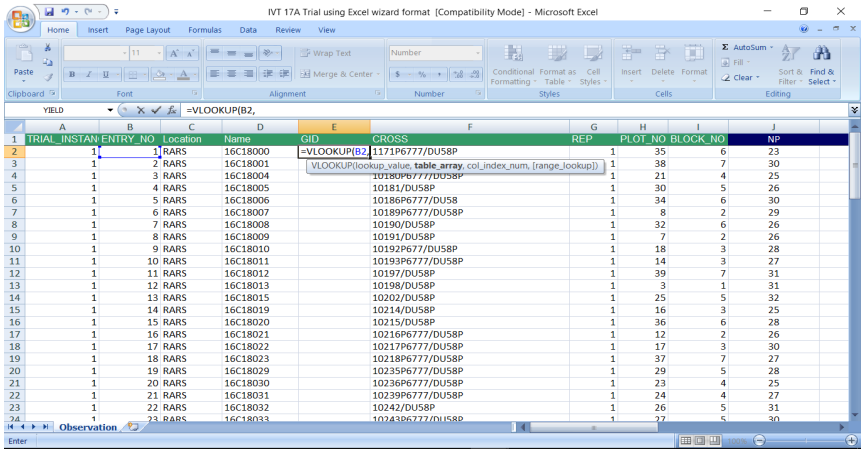
R	S	T	U	V	W	X	Y	Z	AA	AB	AC
BLOCK_N	SITE_SOIL_PH	Rainfall	Irrigation	Fertilizer	AD	SD	ASI	NP	PH	EH	EPR
1	2.66	108	100	40	69	70	1	47	2.35	1.35	0.574468
1	2.66	108	100	40	70	69	-1	46	2.25	1.1	0.488889
1	2.66	108	100	40	73	74	1	48	2.05	1.2	0.585366
1	2.66	108	100	40	68	69	1	48	2.45	1.35	0.55102
1	2.66	108	100	40	70	71	1	47	2.5	1.54	0.616
2	2.66	108	100	40	70	71	1	48	2.71	1.45	0.535055
2	2.66	108	100	40	71	73	2	46	2.6	1.4	0.538462
2	2.66	108	100	40	67	68	1	48	2.5	1.15	0.46
2	2.66	108	100	40	74	76	2	46	2.6	1.4	0.538462
2	2.66	108	100	40	69	70	1	48	2.4	1.3	0.541667
3	2.66	108	100	40	59	61	2	48	2.38	0.95	0.39916
3	2.66	108	100	40	57	58	1	46	2.3	1.1	0.478261
3	2.66	108	100	40	68	69	1	48	2.48	1.3	0.524194
3	2.66	108	100	40	73	72	-1	47	2.3	1.21	0.526087
3	2.66	108	100	40	70	71	1	46	2.54	1.32	0.519685
4	2.66	108	100	40	67	68	1	48	2.46	1.2	0.487805
4	2.66	108	100	40	70	71	1	48	2.63	1.3	0.494297
4	2.66	108	100	40	69	70	1	48	2.36	1.16	0.491525
4	2.66	108	100	40	69	70	1	48	2.48	1.29	0.520161
4	2.66	108	100	40	73	74	1	48	2.65	1.5	0.566038

Using the VLOOKUP function to copy GIDs into an excel file.

- Open the formatted Trial (either the simple excel file or a Fieldbook file) and insert the VLOOKUP formula into the first cell just below the GID column header on the sheet containing the observations. In our case this will be cell E2. To insert the formula, type =VLOOKUP(in cell GID column. In this example, the VLOOKUP formula will be in will be cell E2.

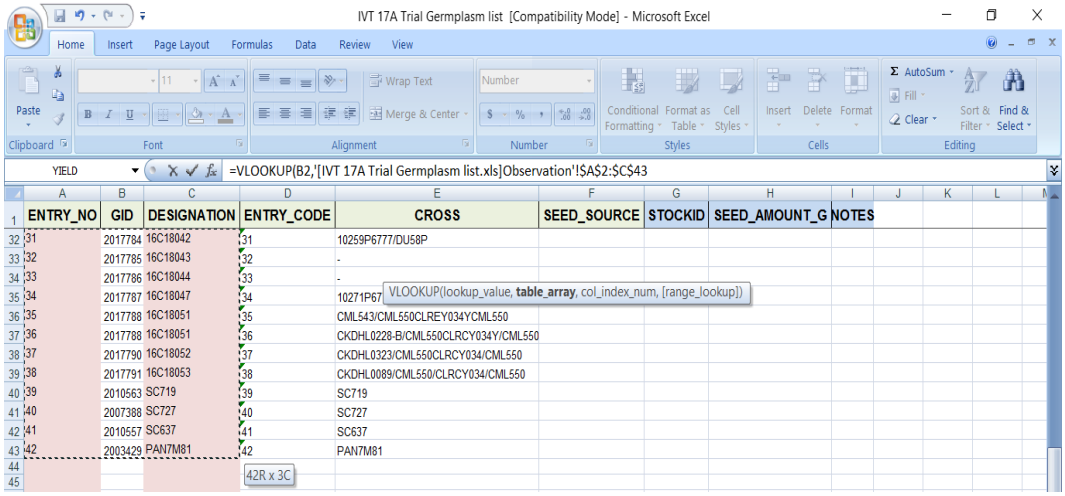


- The VLOOKUP formula has 4 elements;
    a. **Lookup value**- the value you want to look for
    b. **Table Array**- range where your lookup value is located (the lookup value must be in the first column of this range)
    c. **Col Index Num**- The column number in the range containing the return value (relative to the first or look-up column)
    d. **Range Lookup**- matching choice you want (can be either TRUE meaning it's an approximate match or FALSE meaning it's an exact match)
- In this example, the lookup value is Entry NO.1 which is in cell B2. We will use Entry No. because is the common column header in the two sheets.

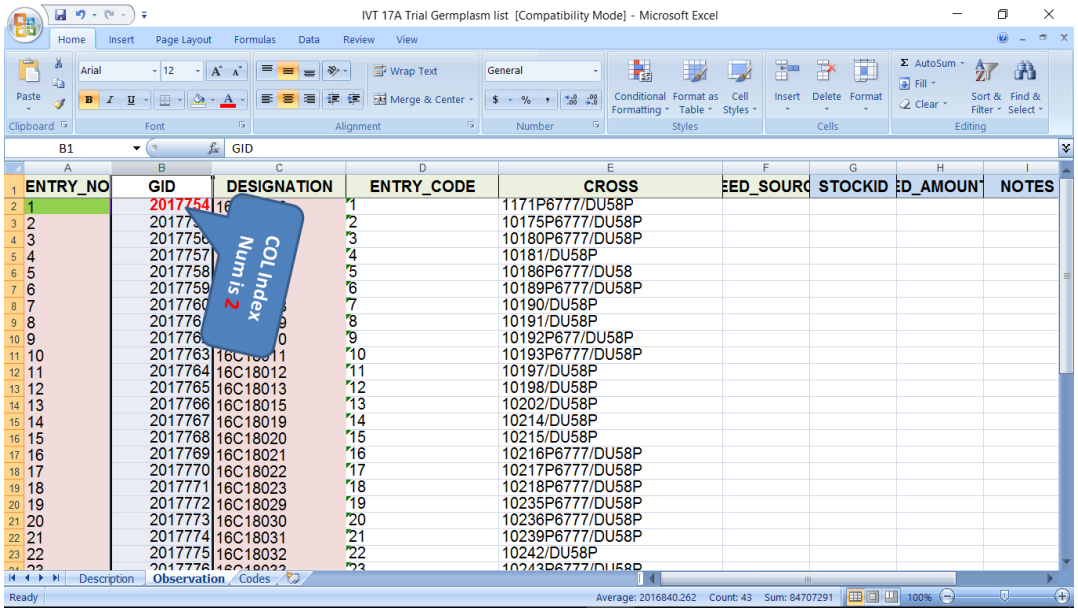


- Enter the Table array [IVT 17A Trial Germplasm list.xls] Observation'!$A$2:$C$43] which is the range including our LOOKUP VALUE and the GID column we need to copy. In this case the Table Array is to come from the Germplams List. (You can also copy the lookup array from the other file into another sheet or even the same sheet of the destination file before you set this range if you prefer.)
- If your excel Version does not automatically generate dollar signs in the formula as ($A$2:$C$43) then you will have to add dollar signs ($) to the table array address so that A2:C43 becomes $A$2:$C$43. Putting the dollar signs in the table array makes the cell references absolute so

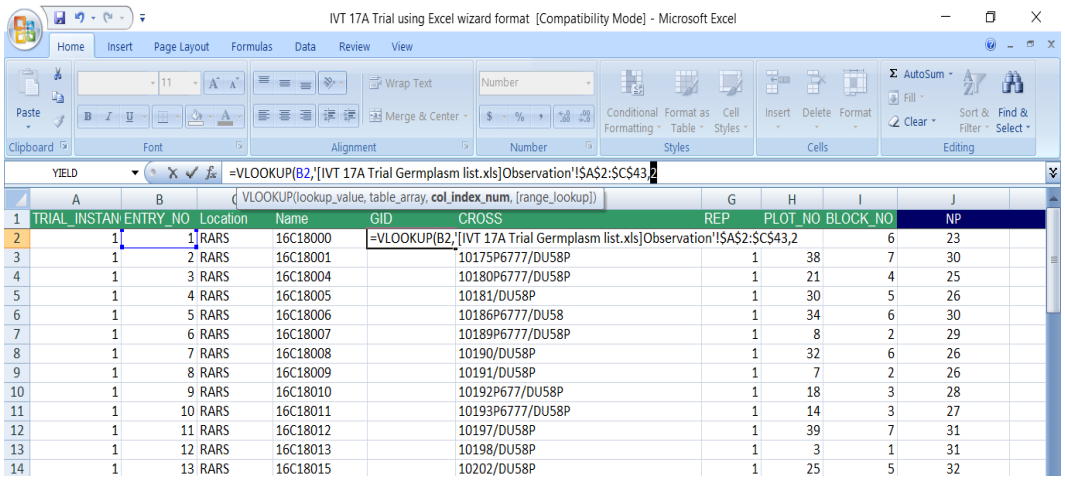
that when you copy the formula down the GID column, the VLOOKUP function will still be looking at the same table (Germplasm list) you referred to.



- Enter the Col Index Num. This is the column number in the range (Table array of the germplasm list) containing the actual value you want to copy in relation to the Lookup value. In this case, ENTRY NO 1 is our lookup value and we want to copy GID for ENTRY NO 1. We check in our germplasm list and the GID of ENTRY NO.1 is in column number 2.

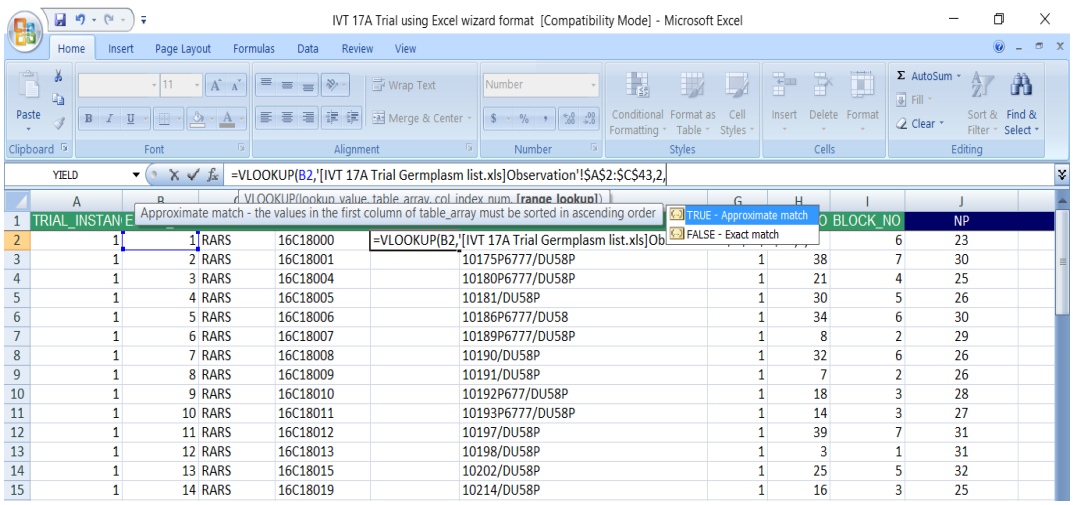


- We then type 2 as our Col Index Num value.



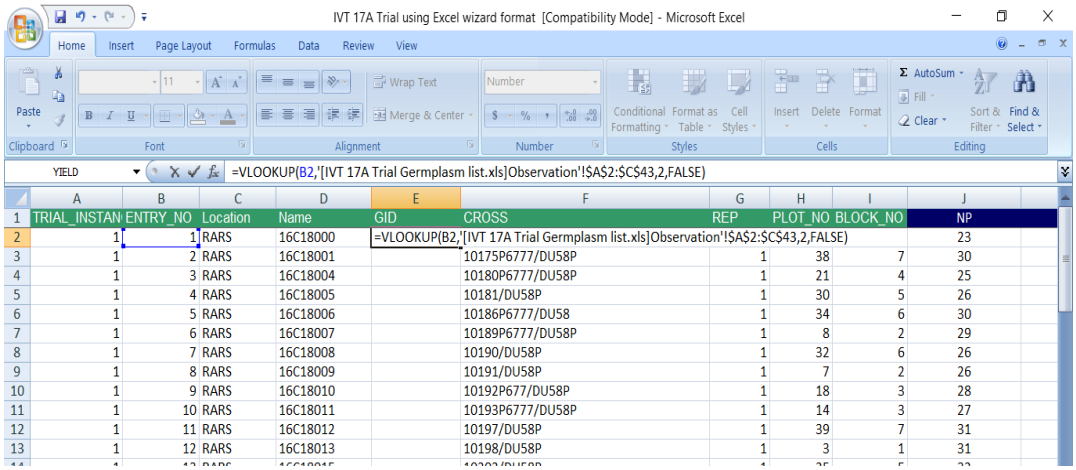
TRIAL_INSTAN	ENTRY_NO	Location	Name	GID	CROSS	REP	PLOT_NO	BLOCK_NO	NP
1	1	RARS	16C18000	=VLOOKUP(B2,'[IVT 17A Trial Germplasm list.xls]Observation'!$A$2:$C$43,2				6	23
1	2	RARS	16C18001		10175P6777/DU58P	1	38	7	30
1	3	RARS	16C18004		10180P6777/DU58P	1	21	4	25
1	4	RARS	16C18005		10181/DU58P	1	30	5	26
1	5	RARS	16C18006		10186P6777/DU58	1	34	6	30
1	6	RARS	16C18007		10189P6777/DU58P	1	8	2	29
1	7	RARS	16C18008		10190/DU58P	1	32	6	26
1	8	RARS	16C18009		10191/DU58P	1	7	2	26
1	9	RARS	16C18010		10192P677/DU58P	1	18	3	28
1	10	RARS	16C18011		10193P6777/DU58P	1	14	3	27
1	11	RARS	16C18012		10197/DU58P	1	39	7	31
1	12	RARS	16C18013		10198/DU58P	1	3	1	31
1	13	RARS	16C18015		10202/DU58P	1	25	5	32

- Next, enter the RANGE LOOKUP which is the matching option. An approximate match is TRUE and Exact match is FALSE. We want an exact match so we type FALSE in the formula.



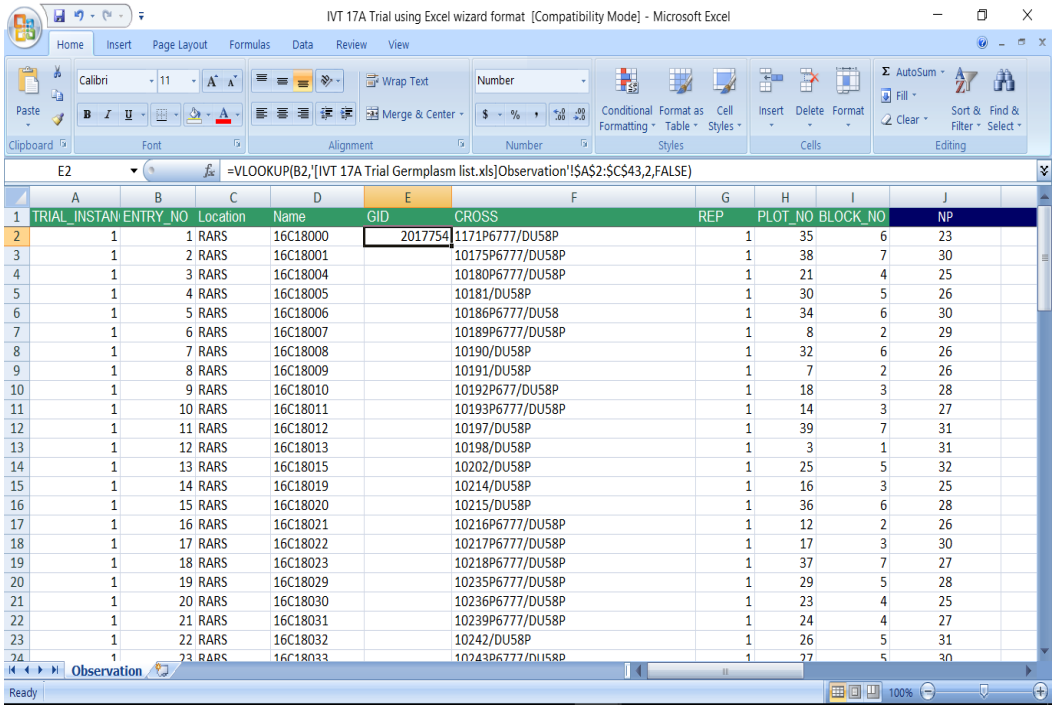
TRIAL_INSTAN	ENTRY_NO	Location	Name	GID	CROSS	REP	PLOT_NO	BLOCK_NO	NP
1	1	RARS	16C18000	=VLOOKUP(B2,'[IVT 17A Trial Germplasm list.xls]Ob				6	23
1	2	RARS	16C18001		10175P6777/DU58P	1	38	7	30
1	3	RARS	16C18004		10180P6777/DU58P	1	21	4	25
1	4	RARS	16C18005		10181/DU58P	1	30	5	26
1	5	RARS	16C18006		10186P6777/DU58	1	34	6	30
1	6	RARS	16C18007		10189P6777/DU58P	1	8	2	29
1	7	RARS	16C18008		10190/DU58P	1	32	6	26
1	8	RARS	16C18009		10191/DU58P	1	7	2	26
1	9	RARS	16C18010		10192P677/DU58P	1	18	3	28
1	10	RARS	16C18011		10193P6777/DU58P	1	14	3	27
1	11	RARS	16C18012		10197/DU58P	1	39	7	31
1	12	RARS	16C18013		10198/DU58P	1	3	1	31
1	13	RARS	16C18015		10202/DU58P	1	25	5	32
1	14	RARS	16C18019		10214/DU58P	1	16	3	25

- Finally type the CLOSING BRACKET and ENTER.



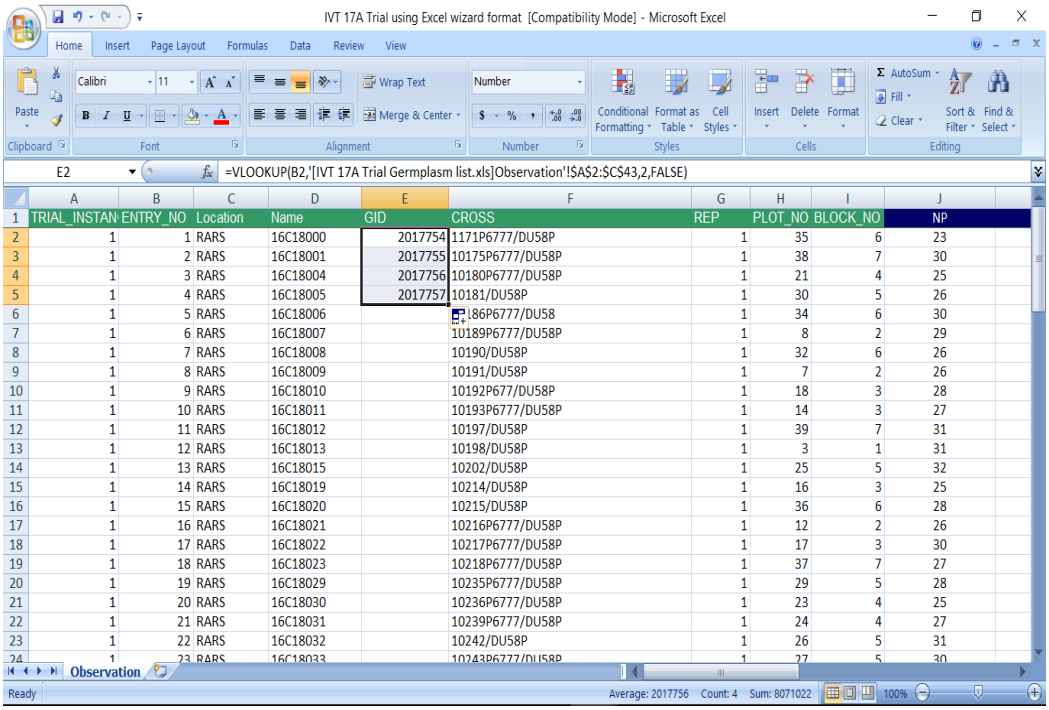
TRIAL_INSTAN	ENTRY_NO	Location	Name	GID	CROSS	REP	PLOT_NO	BLOCK_NO	NP
1	1	RARS	16C18000	=VLOOKUP(B2,'[IVT 17A Trial Germplasm list.xls]Observation'!$A$2:$C$43,2,FALSE)					23
1	2	RARS	16C18001		10175P6777/DU58P	1	38	7	30
1	3	RARS	16C18004		10180P6777/DU58P	1	21	4	25
1	4	RARS	16C18005		10181/DU58P	1	30	5	26
1	5	RARS	16C18006		10186P6777/DU58	1	34	6	30
1	6	RARS	16C18007		10189P6777/DU58P	1	8	2	29
1	7	RARS	16C18008		10190/DU58P	1	32	6	26
1	8	RARS	16C18009		10191/DU58P	1	7	2	26
1	9	RARS	16C18010		10192P677/DU58P	1	18	3	28
1	10	RARS	16C18011		10193P6777/DU58P	1	14	3	27
1	11	RARS	16C18012		10197/DU58P	1	39	7	31
1	12	RARS	16C18013		10198/DU58P	1	3	1	31

- We notice that the GID for ENTRY NO.1 has been copied in the Trial.



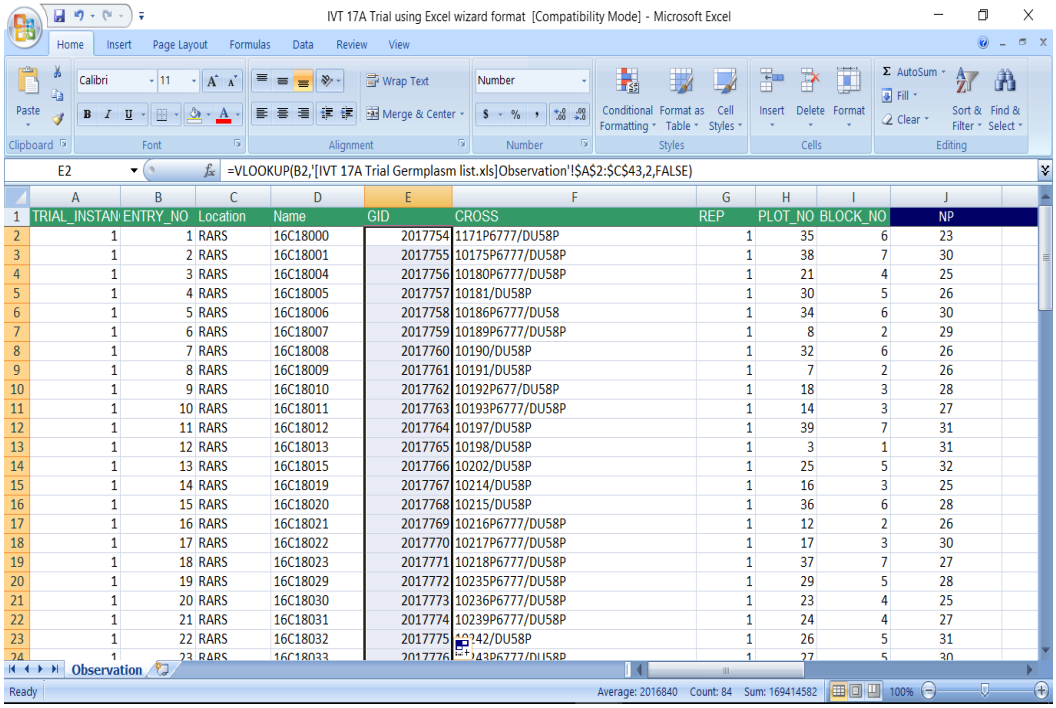
TRIAL_INSTAN	ENTRY_NO	Location	Name	GID	CROSS	REP	PLOT_NO	BLOCK_NO	NP
1	1	RARS	16C18000	2017754	1171P6777/DU58P	1	35	6	23
1	2	RARS	16C18001		10175P6777/DU58P	1	38	7	30
1	3	RARS	16C18004		10180P6777/DU58P	1	21	4	25
1	4	RARS	16C18005		10181/DU58P	1	30	5	26
1	5	RARS	16C18006		10186P6777/DU58	1	34	6	30
1	6	RARS	16C18007		10189P6777/DU58P	1	8	2	29
1	7	RARS	16C18008		10190/DU58P	1	32	6	26
1	8	RARS	16C18009		10191/DU58P	1	7	2	26
1	9	RARS	16C18010		10192P677/DU58P	1	18	3	28
1	10	RARS	16C18011		10193P6777/DU58P	1	14	3	27
1	11	RARS	16C18012		10197/DU58P	1	39	7	31
1	12	RARS	16C18013		10198/DU58P	1	3	1	31
1	13	RARS	16C18015		10202/DU58P	1	25	5	32
1	14	RARS	16C18019		10214/DU58P	1	16	3	25
1	15	RARS	16C18020		10215/DU58P	1	36	6	28
1	16	RARS	16C18021		10216P6777/DU58P	1	12	2	26
1	17	RARS	16C18022		10217P6777/DU58P	1	17	3	30
1	18	RARS	16C18023		10218P6777/DU58P	1	37	7	27
1	19	RARS	16C18029		10235P6777/DU58P	1	29	5	28
1	20	RARS	16C18030		10236P6777/DU58P	1	23	4	25
1	21	RARS	16C18031		10239P6777/DU58P	1	24	4	27
1	22	RARS	16C18032		10242/DU58P	1	26	5	31
1	23	RARS	16C18033		10243P6777/DU58P	1	27	5	30

- To copy GIDs in the whole column, use the drag down tool in the bottom right corner of the cell E (GID column).



TRIAL_INSTAN	ENTRY_NO	Location	Name	GID	CROSS	REP	PLOT_NO	BLOCK_NO	NP
1	1	RARS	16C18000	2017754	1171P6777/DU58P	1	35	6	23
1	2	RARS	16C18001	2017755	10175P6777/DU58P	1	38	7	30
1	3	RARS	16C18004	2017756	10180P6777/DU58P	1	21	4	25
1	4	RARS	16C18005	2017757	10181/DU58P	1	30	5	26
1	5	RARS	16C18006		186P6777/DU58	1	34	6	30
1	6	RARS	16C18007		10189P6777/DU58P	1	8	2	29
1	7	RARS	16C18008		10190/DU58P	1	32	6	26
1	8	RARS	16C18009		10191/DU58P	1	7	2	26
1	9	RARS	16C18010		10192P677/DU58P	1	18	3	28
1	10	RARS	16C18011		10193P6777/DU58P	1	14	3	27
1	11	RARS	16C18012		10197/DU58P	1	39	7	31
1	12	RARS	16C18013		10198/DU58P	1	3	1	31
1	13	RARS	16C18015		10202/DU58P	1	25	5	32
1	14	RARS	16C18019		10214/DU58P	1	16	3	25
1	15	RARS	16C18020		10215/DU58P	1	36	6	28
1	16	RARS	16C18021		10216P6777/DU58P	1	12	2	26
1	17	RARS	16C18022		10217P6777/DU58P	1	17	3	30
1	18	RARS	16C18023		10218P6777/DU58P	1	37	7	27
1	19	RARS	16C18029		10235P6777/DU58P	1	29	5	28
1	20	RARS	16C18030		10236P6777/DU58P	1	23	4	25
1	21	RARS	16C18031		10239P6777/DU58P	1	24	4	27
1	22	RARS	16C18032		10242/DU58P	1	26	5	31
1	23	RARS	16C18033		10243P6777/DU58P	1	27	5	30

- The GIDs have now been copied and you need to cross check the two lists to confirm if the copying was correctly done.



TRIAL_INSTAN	ENTRY_NO	Location	Name	GID	CROSS	REP	PLOT_NO	BLOCK_NO	NP
1	1	RARS	16C18000	2017754	1171P6777/DU58P	1	35	6	23
1	2	RARS	16C18001	2017755	10175P6777/DU58P	1	38	7	30
1	3	RARS	16C18004	2017756	10180P6777/DU58P	1	21	4	25
1	4	RARS	16C18005	2017757	10181/DU58P	1	30	5	26
1	5	RARS	16C18006	2017758	10186P6777/DU58	1	34	6	30
1	6	RARS	16C18007	2017759	10189P6777/DU58P	1	8	2	29
1	7	RARS	16C18008	2017760	10190/DU58P	1	32	6	26
1	8	RARS	16C18009	2017761	10191/DU58P	1	7	2	26
1	9	RARS	16C18010	2017762	10192P677/DU58P	1	18	3	28
1	10	RARS	16C18011	2017763	10193P6777/DU58P	1	14	3	27
1	11	RARS	16C18012	2017764	10197/DU58P	1	39	7	31
1	12	RARS	16C18013	2017765	10198/DU58P	1	3	1	31
1	13	RARS	16C18015	2017766	10202/DU58P	1	25	5	32
1	14	RARS	16C18019	2017767	10214/DU58P	1	16	3	25
1	15	RARS	16C18020	2017768	10215/DU58P	1	36	6	28
1	16	RARS	16C18021	2017769	10216P6777/DU58P	1	12	2	26
1	17	RARS	16C18022	2017770	10217P6777/DU58P	1	17	3	30
1	18	RARS	16C18023	2017771	10218P6777/DU58P	1	37	7	27
1	19	RARS	16C18029	2017772	10235P6777/DU58P	1	29	5	28
1	20	RARS	16C18030	2017773	10236P6777/DU58P	1	23	4	25
1	21	RARS	16C18031	2017774	10239P6777/DU58P	1	24	4	27
1	22	RARS	16C18032	2017775	10242/DU58P	1	26	5	31
1	23	RARS	16C18033	2017776	10243P6777/DU58P	1	27	5	30

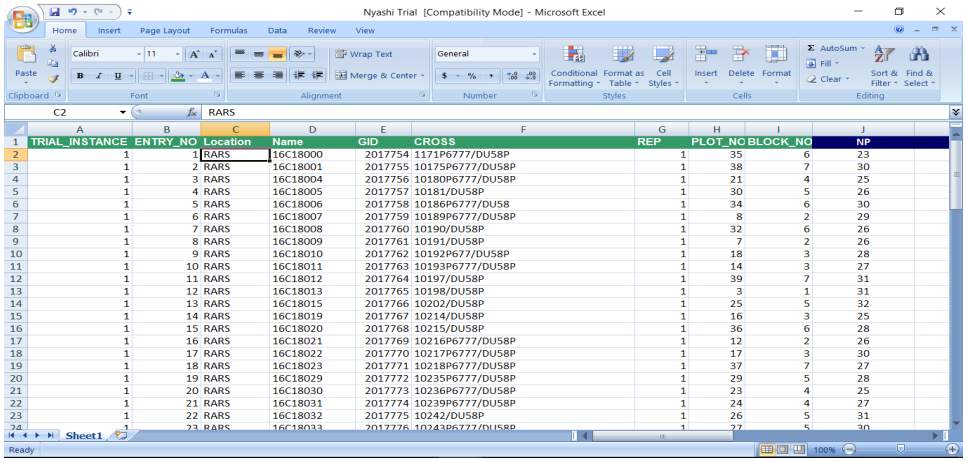
Germplasm list

ENTRY_NO	GID	DESIGNATION	ENTRY_CODE
1	2017754	16C18000	1
2	2017755	16C18001	2
3	2017756	16C18004	3
4	2017757	16C18005	4
5	2017758	16C18006	5
6	2017759	16C18007	6
7	2017760	16C18008	7
8	2017761	16C18009	8
9	2017762	16C18010	9
10	2017763	16C18011	10
11	2017764	16C18012	11
12	2017765	16C18013	12
13	2017766	16C18015	13
14	2017767	16C18019	14
15	2017768	16C18020	15
16	2017769	16C18021	16
17	2017770	16C18022	17
18	2017771	16C18023	18
19	2017772	16C18029	19
20	2017773	16C18030	20
21	2017774	16C18031	21
22	2017775	16C18032	22

Trial

TRIAL_INSTAN	ENTRY_NO	Location	Name	GID
1	1	RARS	16C18000	2017754
1	2	RARS	16C18001	2017755
1	3	RARS	16C18004	2017756
1	4	RARS	16C18005	2017757
1	5	RARS	16C18006	2017758
1	6	RARS	16C18007	2017759
1	7	RARS	16C18008	2017760
1	8	RARS	16C18009	2017761
1	9	RARS	16C18010	2017762
1	10	RARS	16C18011	2017763
1	11	RARS	16C18012	2017764
1	12	RARS	16C18013	2017765
1	13	RARS	16C18015	2017766
1	14	RARS	16C18019	2017767
1	15	RARS	16C18020	2017768
1	16	RARS	16C18021	2017769
1	17	RARS	16C18022	2017770
1	18	RARS	16C18023	2017771
1	19	RARS	16C18029	2017772
1	20	RARS	16C18030	2017773
1	21	RARS	16C18031	2017774
1	22	RARS	16C18032	2017775
1	23	RARS	16C18033	2017776

- Once you have checked that the correct GIDs are going to the correct entry numbers you must copy the whole GID column in the template and paste it as values – you cannot import a file with formulas in the cells.
- **This is a Complete formatted Trial** ready for import into BMS.



Nyashi Trial [Compatibility Mode] - Microsoft Excel

TRIAL_INSTANCE	ENTRY_NO	Location	Name	GID	CROSS	REP	PLOT_NO	BLOCK_NO	NP
1	1	RARS	16C18000	2017754	1171P6777/DU58P	1	35	6	23
1	2	RARS	16C18001	2017755	10175P6777/DU58P	1	38	7	30
1	3	RARS	16C18004	2017756	10180P6777/DU58P	1	21	4	25
1	4	RARS	16C18005	2017757	10181/DU58P	1	30	5	26
1	5	RARS	16C18006	2017758	10186P6777/DU58	1	34	6	30
1	6	RARS	16C18007	2017759	10189P6777/DU58P	1	8	2	29
1	7	RARS	16C18008	2017760	10190/DU58P	1	32	6	26
1	8	RARS	16C18009	2017761	10191/DU58P	1	7	2	26
1	9	RARS	16C18010	2017762	10192P677/DU58P	1	18	3	28
1	10	RARS	16C18011	2017763	10193P6777/DU58P	1	14	3	27
1	11	RARS	16C18012	2017764	10197/DU58P	1	39	7	31
1	12	RARS	16C18013	2017765	10198/DU58P	1	3	1	31
1	13	RARS	16C18015	2017766	10202/DU58P	1	25	5	32
1	14	RARS	16C18019	2017767	10214/DU58P	1	16	3	25
1	15	RARS	16C18020	2017768	10215/DU58P	1	36	6	28
1	16	RARS	16C18021	2017769	10216P6777/DU58P	1	12	2	26
1	17	RARS	16C18022	2017770	10217P6777/DU58P	1	17	3	30
1	18	RARS	16C18023	2017771	10218P6777/DU58P	1	37	7	27
1	19	RARS	16C18029	2017772	10235P6777/DU58P	1	29	5	28
1	20	RARS	16C18030	2017773	10236P6777/DU58P	1	23	4	25
1	21	RARS	16C18031	2017774	10239P6777/DU58P	1	24	4	27
1	22	RARS	16C18032	2017775	10242/DU58P	1	26	5	31
1	23	RARS	16C18033	2017776	10243P6777/DU58P	1	27	5	30

Learn more about BMS from the Tutorials at **https://www.integratedbreeding.net**