

# CloneSelector 3.0 Users guide Sweetpotato

---

Nairobi, 16 July 2012, Luka M Wanjohi

# Table of Contents

<b>1. INTRODUCTION TO THE CLONESELECTOR USERS GUIDE .....</b>	<b>4</b>
<b>2. SWEETPOTATO BREEDING PROCEDURES .....</b>	<b>5</b>
2.1 INTRODUCTION .....	5
2.2 PROCEDURES .....	7
2.2.1. <i>Multiply clones for trial and maintain their identities</i> .....	7
2.2.2. <i>Trial types and selection schemes</i> .....	8
2.2.2.1 Farmer Participatory Variety Selection (FPVS) .....	12
2.2.3. <i>Data analysis, selection of clones and reporting of results</i> .....	13
<b>3. HOW TO USE CLONESELECTOR IN A SWEETPOTATO TRIAL .....</b>	<b>14</b>
3.1 HOW TO START CLONESELECTOR? .....	15
3.2 ELABORATION OF LIST OF GERMPLASM TO TEST .....	16
3.2.1. <i>Open List of Germplasm</i> .....	18
3.3 DESIGN OF FIELD TRIAL WITH ONE OR MORE EXPERIMENTS .....	19
3.3.1. <i>Make labels for planting material</i> .....	20
3.4 GENERATION OF A FIELD BOOK FOR EACH EXPERIMENT .....	22
3.4.1. <i>How to open a field book</i> .....	23
3.5 REGISTER METADATA FOR EACH EXPERIMENT .....	25
3.5.1. <i>General information</i> .....	25
3.5.2. <i>Installation information</i> .....	27
3.5.3. <i>Field information</i> .....	28
3.5.4. <i>Weather data</i> .....	28
3.5.5. <i>Crop management</i> .....	28
3.6 DATA COLLECTION IN FIELD AND POST HARVEST FOR EACH ENTRY .....	30
3.6.1. <i>Introduction to sweetpotato traits</i> .....	30
3.6.1.1 Pre-harvest data .....	31
3.6.1.2 Sweetpotato harvest data .....	33
3.6.1.3 Sweetpotato quality data .....	34
3.6.1.4 Piecemeal harvest data .....	35
3.6.1.5 Near-infrared reflectance spectroscopy quality data .....	37
3.6.2. <i>Printing the fieldbook sheet</i> .....	37
3.6.2.1 Adding fieldbook headers .....	38
3.6.2.2 Hide traits that will not be measured and set print area .....	38
3.7 ENTER COLLECTED DATA INTO ELECTRONIC FIELD BOOK .....	39
3.7.1.1 Checking data quality of the fieldbook data .....	39
3.8 CALCULATION OF DERIVED VARIABLES .....	41
3.8.1.1 Calculating variable in CloneSelector .....	41
3.9 STATISTICAL ANALYSIS OF TRIAL .....	43
3.10 ELABORATION OF CONSOLIDATED RESULT TABLE FOR EACH TRIAL .....	44
<b>4. OTHER FUNCTIONS IN CLONESELECTOR .....</b>	<b>46</b>
4.1 TOOLS FOR SORTING DATA AND RESULTS .....	46
4.1.1. <i>Sorting Fieldbook according to entry number</i> .....	46
4.1.2. <i>Sorting Fieldbook according to plot number</i> .....	46

4.1.3.	<i>Sorting Result table according to user selected column</i> .....	46
4.2	COPY EXISTING DATA SET INTO A CLONESELECTOR FIELDBOOK.....	47
4.2.1.	<i>How to create a CloneSelector Fieldbook based on an existing design</i> .....	47
4.2.2.	<i>Enter or copy your existing data into the CloneSelector Fieldbook</i> .....	48
4.3	SAVING FILES IN CLONESELECTOR .....	49
4.4	CONVERT FIELDBOOK TO CSV.....	50
4.5	RECREATE MASTER FORMULAS.....	51
4.6	FIX HIDE COLUMNS PROBLEM.....	52
4.7	SOLVE AUTOMATIC CALCULATION PROBLEM.....	52
<b>5.</b>	<b>CUSTOMIZING CLONESELECTOR TO YOUR NEEDS.....</b>	<b>54</b>
5.1	CLONESELECTOR FILE STRUCTURE .....	54
5.2	SETTING UP THE LOCATION OF THE CLONESELECTOR FILES AND TEMPLATES.....	55
5.3	CUSTOMIZING THE TRIAL TEMPLATE TO YOUR OWN NEEDS.....	56
5.3.2.	<i>Steps to changing the Trial Template</i> .....	56
5.3.2.1	Changing the Fieldbook sheet.....	57
5.3.2.2	Changing the Result and Master sheet .....	58
5.3.2.3	Changing the General, Installation, Field, Weather_data, and Crop_Management sheets.....	59
5.3.3.	<i>Steps to saving changes to trial template</i> .....	59
<b>6.</b>	<b>INSTALLATION OF CLONESELECTOR.....</b>	<b>61</b>
6.1	INSTALLING REXCEL .....	61
6.2	INSTALLING CLONESELECTOR.....	61
6.2.1.	<i>Solving Reference Error problem in Excel 2007 or 2003</i> .....	61

# 1. Introduction to the CloneSelector Users Guide

Jens Riis-Jacobsen

CloneSelector is a tool developed to **help plant breeder carry out field trials, analyze the results and make selection decisions**. In version 2.0 the functionality is focused on the process of carrying out a trial to test a list of germplasm, and it may help you in the following task:

- 1) Elaborate list of germplasm to test
- 2) Design of field trial with one or more experiments
- 3) Generation of a field book for each experiment
- 4) Register metadata for each experiment
- 5) Data collection in field and post harvest for each entry
- 6) Enter collected data into electronic field book
- 7) Calculation of derived variables (yield/ha etc)
- 8) Statistical analysis of trial
- 9) Elaboration of consolidated result table for each trial
- 10) For multi environment trial you may in addition:
  - a. Create a consolidated data set across multiple environments
  - b. Analyze each trait with advanced statistical methods
  - c. Create a consolidated result sheet
  - d. Generate selection indexes and rank in accordance with breeding objectives

The **CloneSelector is based on the Maize Fieldbook** developed by CIMMYT, and this tool has many additional functions especially related to seed management. For clonally propagated crops this functionality is not adequate and it has been eliminated in CloneSelector, however, for seed based crops it might be better to use the original Maize Fieldbook, which is available at CIMMYT. The current version of CloneSelector is targeted at sweetpotato and potato, but it can be used for any crop as long as the trial template is adapted to include the relevant traits.

The **CloneSelector User Guide is based the *Procedures for the evaluation and analysis of sweetpotato trials***, Gruneberg et al (2010), and as such **it is a practical tool that aims to help the sweetpotato breeders put these standardized methods into practice, and by doing that it aims to improve the efficiency and effectiveness of the breeding effort**. The Procedures have been incorporated into the user's guide in order to facilitate the learning and adoption of both tools. Furthermore, the process of operationalizing the procedures in a specific tool has lead to some minor adjustments on how the information is structured or presented, and the User Guide presents these changed data formats.

For the **installation** of CloneSelector see Chapter 6

## 2. Sweetpotato breeding procedures<sup>1</sup>

Gruneberg et al.

### 2.1 Introduction

Breeding programs involve large investments of time and money, but can pay very large returns on investment in the form of improved varieties which benefit farmers, societies and the environment. International breeding efforts involving multiple partners and targeting regionally important constraints have great potential for efficiently and rapidly achieving impact. Standardized information on the performance of progenies and selected clones across environments is necessary in order to assist breeders to efficiently make decisions about selection and variety release. Standardized methods also facilitate the reporting of breeding program results to the agencies that support us. **For sweetpotato these standardized methods have been documented** in *Procedures for the evaluation and analysis of sweetpotato trials*, Gruneberg et al (2010) which provides standard methods for partners in CIP's global breeding efforts, and the following is an edited version of chapter 1 and 2 in that document.

**Breeding is a process for adapting a crop to human needs.** An important component of breeding is the selection of new varieties. The selection of better varieties requires a good understanding of what is needed by farmers and societies, and it requires good biological and statistical knowledge.

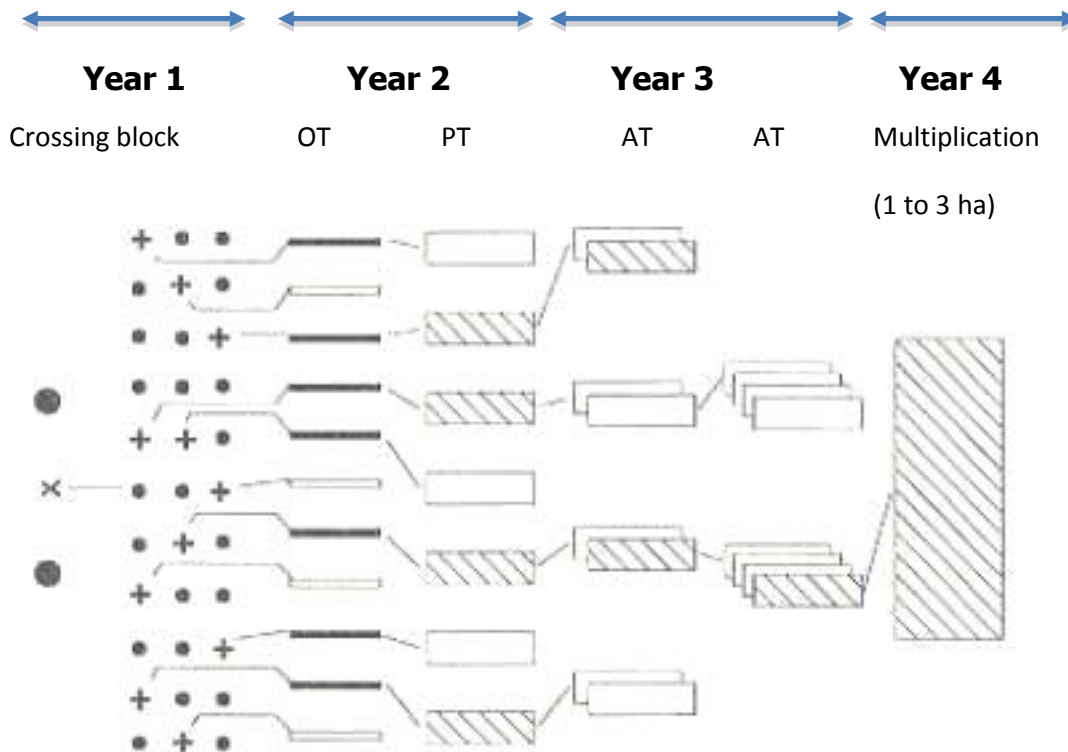
A variety is always characterized by several traits. **A better variety must have good performance over all traits and at least in one important trait it must be clearly superior to all other varieties, which are so far available in a region.** It is not possible to compare a set of new genotypes with all existing varieties across all target environments. Therefore we evaluate and compare new genotypes with important standard varieties (check varieties) in important environments (check environments). The check environments should be representative for the region we are aiming at.

**Selection of new varieties for a region requires comparing new genotypes with check varieties.** This is a very complex task, and may involve many partners working in different places throughout the region. Only what is comparable can be compared! If we evaluate different traits or if we use different procedures to evaluate these traits in different trials, we cannot compare the performance of genotypes across trials. The same is true if we use different sets of new genotypes and different check varieties. **Only commonly measured genotypes and check varieties can be compared.** We distinguish two designs which allow us to compare results of trials: (i) The complete design in which all genotypes are commonly tested across all check environments and (ii) the incomplete design in which only a fraction of all genotypes (at least 6) are tested across all check environments. If there are no or only very few genotypes (1 to 4 genotypes) commonly tested across check environments, it is not possible to make meaningful comparisons among trials. Hence there **needs to be agreement among breeders about:** (i) the **most important traits** to be evaluated, (ii) **standardized procedures** to record these traits, and (iii) **commonly used genotypes and checks varieties** (> 5 genotypes).

---

<sup>1</sup> This section is a slightly edited version of chapter 1 and 2 in *Procedures for the evaluation and analysis of sweetpotato trials*, Gruneberg et al (2010).

**Accelerated sweetpotato breeding (ASPB).** This method, which can be used for both pre-breeding (population improvement) and for varietal selection, is illustrated in figure 1, below. It can allow for completion of a selection and recombination cycle (in the case of prebreeding) in 1 to 2 years, or the selection and release of a new variety in a period of 3 to 4 years. The main features of the approach are to **use multiple selection sites from the initial stage of selection**, and to **minimize replication** (a maximum of two reps per trial are used) at each site so as to conserve resources while obtaining information on stability of genotypes being tested. The trial stages for which we present forms in this manual are the observational yield trial (OT), the preliminary yield trial (PT) and the advanced yield trial (AT). An important feature of our breeding and cultivar selection method is the use of farmer participation, which provides essential input to the breeding and selection process. Here, we detail the use of farmer input at the AT stage.



- Rules:**
- 1) early breeding stages: 1m row plots (thousands of genotypes) everything what can be made simultaneously is made simultaneously at 3 (4) locations no replications
  - 2) later breeding stages: 4-5 row plots (1st selection step: 300 clones, 3 locations, two replications; 2nd selection step: 40 clones, 6-12 locations, two replications)

Figure 1. Accelerated Sweetpotato Breeding Scheme

## 2.2 Procedures

### 2.2.1. Multiply clones for trial and maintain their identities

**Sweetpotato clones for trialing may be newly derived from seedling populations or may be important varieties or promising selections from other breeding programs which have been introduced as pathogen-tested in vitro plantlets.** Within regions sweetpotato clones may be moved as cuttings, following approved quarantine procedures. Quarantine procedures may slow down the exchange of breeding material, but are important. Who wants to become famous by introducing new pests or pathogens into environments where these have so far not been present?

**Clones have to be multiplied to produce planting materials for initial trials.** Locally important as well as standard check varieties should be included in multiplication plots to provide uniform planting materials for trials. Planting material of all the genotypes for any trial should come from a single source and the **health status of genotypes in the trial should be similar.** Often a common health status among clones is difficult to achieve. For example, local clones might not be virus free, whereas introduced clones are obtained pathogen free. In such a case the effect of the genotype might be confounded with its health status. Similar problems might occur although you are not working with introduced clones. Standard check varieties might have been used for a long time, without renewing them from a source of pathogen-free planting material. In contrast newly developed clones are young. It should be noted that health status does not only affect yield related traits. The same clone can appear morphologically different if it is pathogen-free or infected. **Therefore the multiplication of check varieties should, to the extent possible, routinely make use of pathogen-free mother plants.** These mother plants are maintained in greenhouses and are routinely checked to be pathogen free. A mother plant in which a pathogen has been detected must be immediately removed from the greenhouse. Mother clones with pathogens have to be cleaned up or replaced from pathogen-free in vitro source.

**Verify and maintain the identity of clones during the process of multiplication and evaluation.** The most frequently asked question by a professor in plant breeding to his PhD students was: Is genotype number 1 still number 1 or is it perhaps already number 2? Many plant breeders have encountered surprising results, only to realize that a mistake in the labeling of the genotypes must have happened, leading to a mix-up in identities. Such mistakes can be drastically reduced by giving clones both numbers and names. Mistakes in numbers occur more frequently than mistakes in names. If a genotype has no name, give the clone a 'code name'. A code name can easily be formed by the family the clone is tracing back [six digits; three digits for the father and three digits for the mother – (in case the clone derived from polycrosses put 000 for the father)] and then – separated by a decimal point – the clone number within the family. It is very unlikely that a mistake in the clone number and in the code name would happen at the same time, so you can usually quickly identify mistakes. It should be noted that mistakes cannot completely avoided, however, the problem starts when you cannot identify mistakes or when you cannot identify them rapidly.

Finally, we want to mention that the identity of clones can be confirmed by morphological characteristics and descriptor lists. However, these are usually only available for clones maintained in genebanks. Anyway, **the finally selected clones must be morphologically described and distinguishable**

**from other clones to allow a registration as a new variety.** If published and observed descriptors (including pigmentation of foliage and roots, and especially leaf shape) do not coincide (in other words, if you observe for the same clone two different types) both clones can still be entered in trials. In this case it is required to rename the clones to distinguish them. You should use the original name with an extension where the morphologically different type has been observed for the first time e.g. Jonathan-L for the original clone described and maintained at the sweetpotato genebank in Lima and a Jonathan-M for a clearly distinguishable Jonathan first observed in Maputo. Since sweetpotato has a tendency to mutate, it may well be that the new type is not the result of a mix-up, but could be a new Jonathan superior to the original Jonathan.

### 2.2.2. Trial types and selection schemes

Examples of **sweetpotato selection schemes** are provided by Hahn (1982), Martin (1983), Jones et al. (1986), Wilson et al. (1989), Kukimura et al. (1990), Saladaga et al. (1991), Tan et al. (2007) and Grüneberg et al. 2009a). Sweetpotato breeding as been reviewed by Martin and Jones (1986), Laurie and van den Berg (2002), Tan et al. (2007) and Grüneberg et al. (2009b). However, each sweetpotato variety is a highly heterozygous hybrid and we think that **the use of more breeding methods from hybrid breeding applied to sweetpotato breeding has merit** (for a discussion of this topic see Grüneberg et al. (2009b). This might lead to changes for sweetpotato population improvement in the future, but not to changes in variety development and selection.

In formal plant breeding **we distinguish between observation yield trials (OTs), preliminary yield trials (PTs) and advanced yield trials (ATs).** Note: ATs are occasionally designated as uniform as well as national trials, but the procedures used in both are usually the same. Therefore, we refer to both of these as ATs. Formal plant breeding (on station breeding) has been criticized for being slow to develop better varieties for resource poor farmers. For this reason CIP supports formal plant breeding programs which **involve farmers by (i) farmer participatory variety selection and (ii) on-farm trials.**

**Early breeding stages:** In the early breeding stages plants are raised from true seeds. Selection of single true seedling plants may not be advisable, because measurements on single plants have an extremely high error and plants grown from seeds are very different from those grown from cuttings with respect storage root formation. For this reason, evaluations of true seed plants are **limited to a few highly heritable traits such as susceptibility to pathogens or storage root flesh color.** Genotypes selected among true seed plants enter observation trials (OTs).

The **observation yield trial (OT)** belongs to the early breeding stages. The breeder has to evaluate many genotypes (several thousands) grown from cuttings of true seed plants and most of the genotypes grown in OTs clearly do not meet a lowest acceptable value in at least one character. OTs are also recommended for clones which are introduced from other regions of the world (i) to obtain a better understand how to handle foreign materials and (ii) to discard clones which clearly do not meet minimum acceptable values in the new region. In the sweetpotato breeding program at CIP we plant **OTs without replication**, in a single-row plot comprising 3 plants. However, **each clone is planted at 4 locations** (each location is treated as a replication of a randomized complete block design). Each OT is bordered by planting guard rows on all four sides of the trial to provide competition to all entries. The



name OT might seem to imply that only visual observations are made at this stage of selection. However, an OT grown across several locations merits to record and analyze highly heritable traits. Formerly CIP planted the OT in a single-row comprising 10 plants at only one location. Other breeders conduct the OT so that each genotype is evaluated on single plant basis (for example SPVD resistance or storage root flesh color). In the new OT design at CIP (OTs in one meter row plots across several locations) **it is possible to observe the stability of genotypes across locations and to separate the genotype by environment (GxE) effect from the genotype effect.** Heritability ( $h^2$ ) estimates show that the harvest index and **several storage root quality characteristics** (i. e. dry matter, protein, starch, sugar(s) and pro-vitamin A concentrations) can be evaluated in such an OT design with sufficient precision for selection ( $h^2 > 0.6$ ). To our surprise we observed **significant low to medium heritabilities ( $h^2$  about 0.4 to 0.5) for storage root yield.** However, visual selection for storage root size, shape and form - recorded using a single rating scale from 1 to 5 - showed a significant correlation with yield measurements in kg per plot in the new OT design. This can be explained by the extremely large genetic variation for storage root yield in early breeding stages due to very high segregation in hexaploid sweetpotato. However, this **allows us only to discard genotypes with poor yield performance since discrimination among clones with medium to high yield performance is only possible in larger plots** (see PTs and ATs).

Note: The OTs in the sweetpotato breeding program of CIP are directly used to select parents for crosses in hybridization blocks. Hence the breeding program operates with very short recurrent selection cycles to improve breeding populations and allow for variety development. Each year about 250 genotypes are selected from OTs: (i) to be used for crosses for population development and (ii) to enter PTs for variety development and selection, respectively (see below). All crosses are carried out by controlled crossings. In contrast to polycross nurseries, this results in more balanced seed production per parent, since both parents can be controlled. Selection theory tells us that controlled crosses are more efficient than polycrosses, however, considerably fewer seeds can be produced when using controlled crosses. [Current research aims to compare progress using both approaches.] Our crosses are carried out in a factorial design (the best with the rest), in which about 6 x (250-6) crosses are carried out. About 1/3 of all cross combinations results in no or low seed set so that each year about 900 families with about 20 to 40 seeds per family are developed for the OTs of the next recurrent selection cycle.

**To summarize:** OTs are characterized by a very large number of genotypes evaluated in very small plots without replications. The OT can be carried out at one location or at several locations and environments, respectively. The design of the OTs depends on the priorities of specific regions. For example in high virus pressure regions the breeder has first to eliminate all genotypes which show insufficient virus resistance, whereas in a drought prone region the breeder first has to eliminate all genotypes which show insufficient vine survival.

Here we do not prescribe traits to record in OTs, because **the traits to be recorded depend on the region, the country and the major breeding objectives, respectively.** However, we are aiming to obtain general OT trial information as well as information for the recorded traits in selected clones, which will allow us to group breeding programs into clusters to support appropriate true seed exchange among breeding programs. Moreover, we request to **record the parents and clone numbers of selected clones.**

The reason for this is that it is common knowledge among breeders that often good clones trace back to a few cross combinations. The record of parents and clone numbers of selected clones will allow us to determine the frequency of selected clones among parental combinations and to determine which promising crosses should be repeated on a larger scale. For details on record keeping for OTs, see section 3. Data collection forms for all trials are provided in Appendix 1.

**Later breeding stages:** Clones selected in OTs enter into later breeding stages or into variety selection and development. These **later breeding stages are comprised of preliminary yield trials (PTs) and advanced yield trials (ATs)**. In PTs and ATs the same characters are recorded, but on different plot size basis. Note: the ATs should be recorded on a plot size basis, which is identical with those required for official variety release - this is country specific. Moreover, the names uniform trials (UTs), national trials (NTs) and elite trials (ETs) are used for ATs. However, usually these have the same plot size and the same traits have to be recorded as for ATs. It should be noted that NTs and ETs are used for those finally selected clones to be tested against a group of check clones for variety release. In this case we distinguish between ATs and NTs or ETs, but all three have the same plot size and the same traits have to be recorded. They differ in the number of clones to be tested – usually in NTs or ETs a smaller number of clones are tested compared to ATs, so that the test precision / power is larger compared to ATs (see statistical analysis and multiple comparison procedures).

**The preliminary yield trial (PT)** is normally carried out in two row plots, 30-plants per plot (15 plants per row), and two plot replications. The PTs are planted in a randomized complete block design [RCBD (replications of clones are planted in blocks and in each block all genotypes are randomized)]. Single row plots should not be used because of the inter-plot competition (border effects due to neighbor plots within a block). Border effects are assumed to be large in sweetpotato, due to the large genetic variation for the upper biomass production among sweetpotato clones. The coefficient of variation (CV%) for the storage root yield error term in a single PT (two rows, 15 plants per row) is typically very large and ranges in our breeding programs from 28% and 52%. It would probably be larger if we used single rows.

**The PTs must be carried out in at least 3 locations and with plot replication for each genotype**, or it should not be considered as a preliminary yield trial. The clear advantage to conducting the PTs at three or more locations is that this saves time (years), because in sweetpotato trials, temporal variation of test environments (years) can be replaced by spatial variation of test environments (locations). Conducting PTs across locations, with two or more replications, allows us to separate the effects due to genotype, due to genotype by environment interactions, and due the plot error for each trait. Furthermore, with 3 locations it is possible to determine for each genotype, stability parameters, which must be considered as an additional character associated with yield.

**To summarize**, since 2008 CIP has aimed **to conduct PTs** as follows: (i) at least **two row plots** with (ii) at least **30 plants per plot**, (iii) **two plot replications** per genotype and (iv) in at least **three locations and environments**, respectively (you can generate environments in the same location by treatments i.e. irrigation or fertilization). The set of attributes and traits, respectively, to be recorded in PTs is fixed. For

details of data to record for PTs, see section 3<sup>2</sup>. The recording of additional traits is optional, and should be done if the breeder considers one or more traits ought to be recorded for an appropriate selection in the target environment.

Note: The main question in selection is **how many genotypes should be selected?** If almost none of the clones in a trial meet the lowest acceptable value in each trait there is not much choice. However, after good OTs most genotypes should meet the lowest acceptable value across all character. Variety selection is a multi-stage process, and for fixed entries (all genotypes of clonally propagated crops are fixed entries) this multi-stage selection problem has been well solved by selection theory. The results of selection theory show for very different selection scenarios [different ratios of variance components for genotypes, GxE and plot error, and different numbers of test capacities (total number of possible field plots to be allocated to genotypes, locations and replications)] that **at each selection stage 5 to 15% of the total number of genotypes should be selected**. Moreover, they show that **more than 3 to 4 selection stages do not result in a significant increase of genetic gains**. Hence a 3-step selection (one in OTs, one in PTs, and one in ATs) is sufficient to identify the most appropriate clones for variety registration in a breeding population. However, if the breeding population is still at an unsatisfactory level, there might be no or only very few clones which can be recommend for variety registration. In such a case the breeder must allocate more resources to population improvement (to increase the variety-generating ability of the breeding population) by (i) conducting more crosses (controlled crosses), (ii) using more parents, and (iii) shorter recurrent selection cycles (1 year to recombine parents and 1 year to select parents, which are more close to his breeding targets).

**The advanced trial (AT)** is the next selection stage of variety selection. It is usually planted as a RCB design, but in larger plots than in the PT. Our ATs in the breeding program of CIP have **5 row plots (15 plants per row) with 75-plant per plot, and two replications per location**. The ATs are carried out at **4 or more locations**. The coefficient of variation (CV%) for the storage root yield error term in a single AT (five rows, 15 plants per row) at CIP typically ranges between 25% and 32%, which is still large compared to ATs for grain crops. This shows that there is still room for improvement in trial designs for sweetpotato, and better experimental designs might be developed for sweetpotato in the future. The result of the selection process in ATs should be 5 to 8 clones with good performance over all traits and which have, in at least one trait, a clear advantage compared to all sweetpotato varieties available in the region. As mentioned above formal plant breeding has been criticized for not being successful at developing better varieties for resource poor farmers. For this reason, at least at one location, **the AT has to be carried out with farmer participation** (details see below). This eliminates the possibility of putting forward genotypes for official variety release that are not accepted by farmers. **The selected clones (5 to 8) are re-evaluated in the next growing season in a similar design at the same locations and additionally in more than 10 on-farm trials (OFT)**, which should be linked with the process of official variety release (for details of OFTs, see below).

---

<sup>2</sup> In CloneSelector the data collection formats provided in the annexes of the original Protocol have been converted into tables, which facilitates the analysis of the data, and some additional traits have also been added on request from the sweetpotato breeders.

**To summarize:** Since 2008 CIP proposes that **ATs be carried out:** (i) in **three and more row plots** with (ii) at least **75 plants per plot**, (iii) **2 plot replications** per genotype, and (iv) at least **four locations**. **Farmer participatory variety selection** is required in at least one location, and at least 10 OFTs in the final selection stage should be linked to official variety release. Instructions for data recording for ATs are given in section 3<sup>3</sup>. Note, the traits are used for ATs are the same as for PTs. Space is also provided for the collection of additional attributes.

**Note: Extended PTs and ATs to evaluate vine survival and piecemeal harvest are of special importance in sweetpotato.** Therefore we provide a brief discussion of these traits and a recommended method to evaluate these traits (this procedure might change in the future).

Sweetpotato has the highest food production per unit area per unit time. However, in drought-prone environments, a critical character of sweetpotato is vine survival from harvest to the next planting season. Additionally, an important trait of sweetpotato is the ability for use in piecemeal harvest, especially when it is used in home gardens. Neither trait has been addressed in PTs nor ATs to date, because typically the complete plot was harvested and no plants remained to determine vine survival and piecemeal harvest quality. We propose a design which allows observing these traits, by using extended plots in PTs and ATs. **These trials are carried out as described above, but rows are longer (about 5 planting positions per row).** The first and larger part of the plot is used to record characters as usual. **The second and smaller part of the plot is used later (2 to 3 months after the first harvest) to determine vine survival and piecemeal harvest traits.** We recommend that each partner should carry out extended PTs and ATs at 2 locations with two plot replications. It should be noted that in drought prone regions vine survival and sprouting potential of small roots determines the acceptance of a variety, since varieties that fail in these attributes will have no planting material available when the rain comes. Moreover, the ability to use a variety for piecemeal harvest is one of the most important characteristics in Sub-Saharan Africa; varieties that develop undesirable fiber or taste at later growing stages (5 months and more) are nearly always rejected by farmers. CIP hopes to come to an agreement with partners that these characters are determined in extended PTs and ATs in the future, and currently CloneSelector include a draft version of how to collect this data.

#### **2.2.2.1 Farmer Participatory Variety Selection (FPVS)**

This is an important part of the evaluation of AT clone. It should be carried out at a minimum of one AT location per country. **Farmers are invited to give their evaluations and comments on clones in ATs.** The evaluation is carried out on the basis of frequencies for the assessment of a limited number of traits (6 traits and one overall assessment) for each variety. The assessment is recorded using color cards to score each entry in the trial. (red = not acceptable; yellow = more or less acceptable; and green = clearly acceptable) and the overall performance of each variety. To assess the genotypes each farmer obtains 21 color cards for each genotype to be evaluated (one card of each color for each attribute to be evaluated). As each attribute is discussed, the farmer values each genotype by placing one color card

---

<sup>3</sup> In CloneSelector the data collection formats provided in the annexes of the original Protocol have been converted into tables, which facilitates the analysis of the data, and some additional traits have also been added on request from the sweetpotato breeders.

into a bag (based on the degree of acceptance for that attribute). There should be separate bags for women's and men's votes, or cards may be marked with M or F to indicated gender. Results are tallied using the data collection forms provided in the Appendix.

**ATs at the 2nd stage, which are carried out at several locations are used as mother trials for OFTs.**

Each farmer obtains 4 varieties from the AT and has to assess these varieties relative to his currently used variety. As with FPVS, the evaluation is carried out on the basis of frequencies for a limited number of traits (6 traits and one overall assessment for each variety). The farmer ranks the new genotypes relative to the performance of his currently preferred variety for each trait as well as for the overall performance across traits. We recommend conducting this on the basis of color cards as for FPVS: Green = improved or better than the local check, Yellow = equal or nearly equal to the local check, and Red = inferior compared to the local check. These data should be collected from farmers during visits to farms.

**2.2.3. Data analysis, selection of clones and reporting of results**

To facilitate analysis and decision-marking, the raw data collected in trials using standardized methods described in section 3 should be transformed into reference units. For instance, the number of harvested plants divided by the number of cuttings planted would be survival; yield measured in kg/plot must be converted to t/ha, etc. In CloneSelector this is done using a Master sheet as described in Section 3. After this processing step sort the data to be analyzed by location, genotype and replication. The ANOVA (Analysis of Variance), and mean comparisons (e.g., LSD or Tukey test) become useful tools for clone selection. In section 4 we give recommendations for breeders to analyze data from their trials, select clones and report results. On the other hand we all must follow the description of data collection in selection 3 (Appendix 1, 2 and 3) in order to report the results of our sweetpotato variety selection programs to our donors. The raw data recorded in the forms and sheets in appendix 1, 2 and 3 must be made available to CIP from 2008 on. CIP regional breeders will work with national partners on a continuing basis to achieve uniformity, quality, consistency and relevance of data from breeding trials.

### 3. How to use CloneSelector in a sweetpotato trial

Jens Riis-Jacobsen

This section describes how to use CloneSelector as an integrated tool in the process of carrying out a sweetpotato trial in accordance with the general protocol described in the previous chapter. The information is organized in the sequence of the general steps for carrying out a trial, and the user should consult the relevant section as the trial is carried out.

When CloneSelector is first installed the templates are in a generic format. In order to work more efficiently and not have to tweak for example the fieldbooks every time, the user may modify the templates to be more suitable for the local conditions, however, for the first trials it is recommended to use the standard templates in order to gain a better understanding before the templates are modified. The process for modifying the templates is described in a subsequent section.

Some general advice on using CloneSelector for trials

1. **Save your data frequently** as errors may happen and you may lose the work you have done
2. **Backup you data files frequently.**
3. **For the first trials follow the instructions in this guide as closely as possible.** Once you know the general working of CloneSelector you can start adapting it to your more specific needs.
4. **Experiment with CloneSelector before you using it in a real trial.** You can for example try to do all the steps in the guide with a trial from last season.
5. The fieldbook in CloneSelector as you will see below is made up of several sheets and these are linked in different manners. If you want **to make changes outside what is described in the following steps only do so on a copy, not in the original.** Changes may prevent CloneSelector from working, and you will probably not be able to undo the changes.
6. Remember, in the worst case your data is in Excel, and if CloneSelector fails to do exactly what you need, **you can always do your calculations or analysis as a normal Excel sheet or exporting the data to a statistical package of your choice.**

If you invest a little time on learning CloneSelector this will make you a more efficient breeder, and you will spend more time drawing the right conclusions from your trial, and less time on repetitive steps.

### 3.1 How to start CloneSelector?

CloneSelector is an add-in to Microsoft Excel, and to use it you must open it as any other Excel file<sup>4</sup>. After opening, Excel 2007 will have an additional menu in the Add-in section and in Excel 2003 the menu is added directly in the primary menu bar of the application.

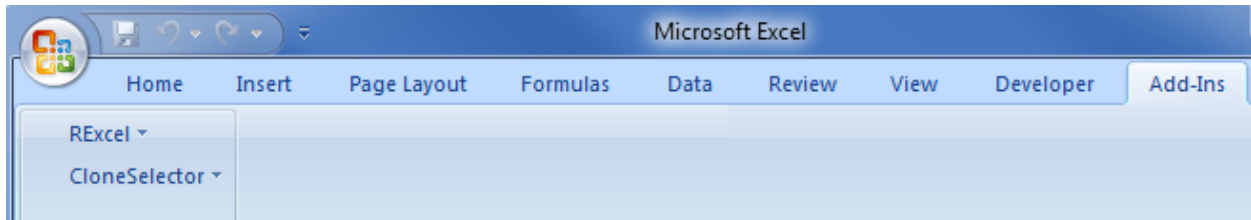


Figure 1 CloneSelector Add-in menu in Excel 2007

When you click on the CloneSelector menu it opens up. Initially, only the List of Germplasm, Trial and Options menus are available. **The rest of the menus will become available once you open or create a List of Germplasm or a Trial.** If you cannot access a menu it is because you are not in the right type of workbook, or you are not on the correct sheet.



**Warning:** CloneSelector use the name of the sheet to activate the relevant menus. **Do not change the names of the Fieldbook, Results and Master sheets** as the menus will not work properly.

CloneSelector is based on Excel, and apart from the tools in the system, you also have access to all the normal Excel functions. You must however, **be careful not to change the structure of the CloneSelector sheets as these will most likely not work correctly afterwards.** The safest solution is to insert a new sheet or use the space below the tables generated by CloneSelector. Similarly, it is preferable to **hide columns or rows, rather than deleting them** as this may cause problems.

If you have changed the structure of a worksheet and this is not working properly or if you have data in a normal Excel sheet that you want to analyze in CloneSelector the best solution is to create a new List of Germplasm or Trial sheet in CloneSelector and **copy the information over using Paste Special and only paste Values.** Doing this will not change your data and will maintain all the functionality and formatting in CloneSelector.

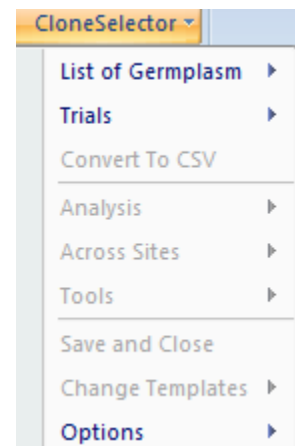


Figure 2 CloneSelector menus

<sup>4</sup> On Windows XP it is possible to add the CloneSelector to the XLSTART folder and it will start automatically when Excel is initiated, however, the security of Windows Vista and Windows 7 does not allow this approach.

### 3.2 Elaboration of list of germplasm to test

A trial in CloneSelector is conceptually a method for evaluating a list of germplasm with regard to certain traits. To some extent all the tools in CloneSelector are focused on manipulating the List of Germplasm and the first step in a trial is to prepare a List of Germplasm:

CloneSelector -> List of Germplasm -> New

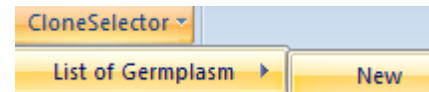


Figure 3 Making a new List of Germplasm

In the dialog that opens **enter or select the Season**. The recommendation is to indicate **Year** either as e.g. **2010** or **10** and **A, B, C** for the **Season**. The Season you enter will be the name of the folder where the file is stored. You must also enter a **List ID**, which will

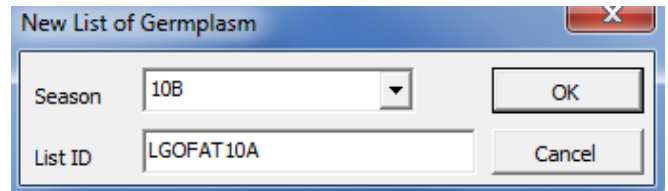


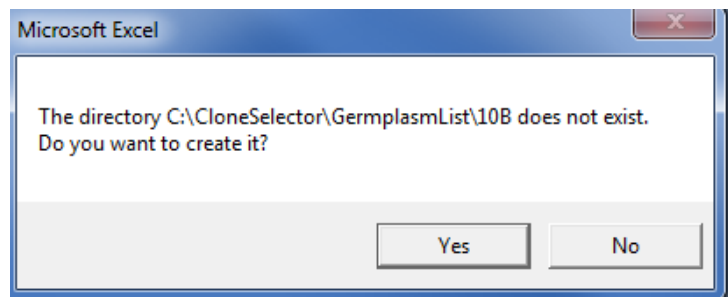
Figure 4 List of Germplasm dialog

become the name of the file. It is recommended that you use a code such as **LG** for **List of Germplasm**, **OF** for **color of material** and **AT** for **type of trial**, and then the **season** e.g. LGOFAT10A.



**Warning: Do not use spaces in the name**, as this will cause problems when trying to generate the field book.

**Click OK** once you have entered the Season and List ID. If it is the first List of Germplasm in a season, a dialog will appear asking if you want to create a new directory, **click Yes**<sup>5</sup>. After this a **List of Germplasm file will open with the name you indicated as List ID**.



Two fields must be filled in the List of Germplasm: **Entry number** and **Vines/plot**.

**Entry number must be a consecutive list of numbers starting with 1**, such as 1,2,3,4, you cannot leave gaps in numbers such as 1, 2, 5, 6 or start at a different number 100, 101 etc. Use Fill down in Excel to generate list.

**Vines/plot** is the number of vines to be planted in each plot and must be a positive integer number such as 5 or 10.

Once you have established the List of Germplasm click **Save** as you would any other Excel file.

	A	C	D	M
1	Entry	Name	BreedersPedigree	SeedsPer
18	Season: 10B			
19	Exp: LGOFAT10A			
20				
23	Entry	Name	Pedigree	Vines /plot
24				
25	1	W119 06-39		20
26	2	UXIPHONE 06-1		20
27	3	UW119 06-296		20
28	4	UW119 06-32		20
29	5	MUSG 0703-37		20
30	6	Libertada		20

Figure 5 Example of List of Germplasm

<sup>5</sup> If you click No CloneSelector cannot create the Germplasm list and you will get an error. In that case try again.



Note: **CloneSelector organizes the lists of germplasm in a separate folder and inside this it creates a folder for each season.** It is **important that you accept the default locations** assigned when saving files in CloneSelector. If you use Save As and saves a file outside the default structure CloneSelector cannot locate the file for example when it is need for generating a fieldbook for your trial. When CloneSelector creates a list of germplasm or fieldbook it is automatically stored in the correct folder, and as long as you do not change it, it will work fine.

The List of Germplasm workbook has three sheets:

- **GermplasmList** contains a list of the entries to be tested in the experiment and the quantities of vines or seed required for each plot
- **Design** is where the experimental design is stored once it has been generated as will be shown in the next section
- **SeedLabels** contains the labels for the planting material for each plot

### 3.2.1. Open List of Germplasm

CloneSelector automatically organizes the List of Germplasm in a separate folder for each year which helps keeping in order all the files related to a breeding program. You can open the files as normal Excel files, but this may implicate some clicking to arrive at the right folder, unless of course the file is still in the list of most recent files used in Excel in which case you just click on it to open it. In order to ensure that the file organization of CloneSelector works well it is, however, preferable to use the Open command:

**CloneSelector -> List of Germplas -> Open**

This will open the Open Germplasm List dialog. The drop down list for Season will have a list of all the seasons for which you have created a list of germplasm. **Select the season** in which you want to locate the list. After selecting the season the List ID drop down will now have a list of all the lists of germplasm generated in that season. **Select the List ID and click OK to open the List of Germplasm.**

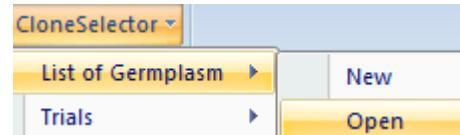


Figure 6 Open List of Germplasm menu

In Excel 2007 you may get a message from the Microsoft Office Excel – Compatibility Checker, which is a product of the changes made in the different versions of Excel. If this message appears you must **clear the check box for Check compatibility when saving this workbook** or this message will appear all the time.

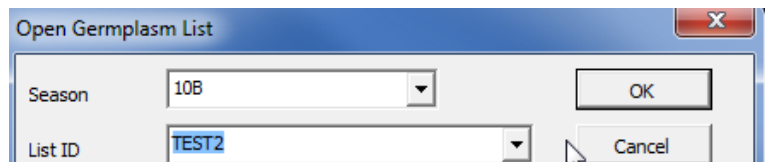


Figure 8 Open List of Germplasm dialog

Unfortunately in Excel 2007 it is not possible to completely disable the annoying feature.

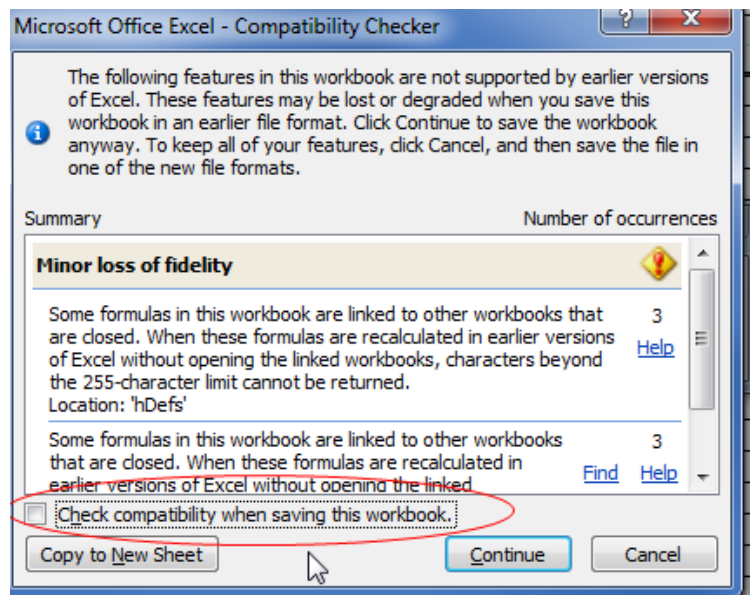


Figure 7 Excel 2007 Compatibility Checker problem

### 3.3 Design of field trial with one or more experiments

Next step is to create a design. To do this the List of germplasm must be open and if that is not the case, see the previous section on how to open it.

**CloneSelector -> List of Germplasm -> Create Design**

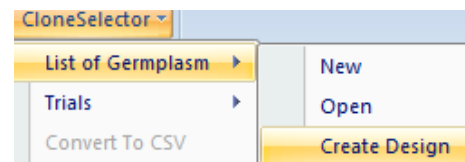


Figure 9 Create Design menu

Note: Your germplasm list must be open and the **GermplasmList sheet must be the active sheet**, else the Create Design menu is disabled.

**In the dialog Create Design:**

Note: The Create design will automatically detect the # entries, and you should not change this. If the number is incorrect, you may not have used a consecutive list of numbers starting with 1.

The design options are divided between **unreplicated** and replicated designs.

Among the former you may choose **No randomization**, which means the design is simply a copy of the list of germplasm

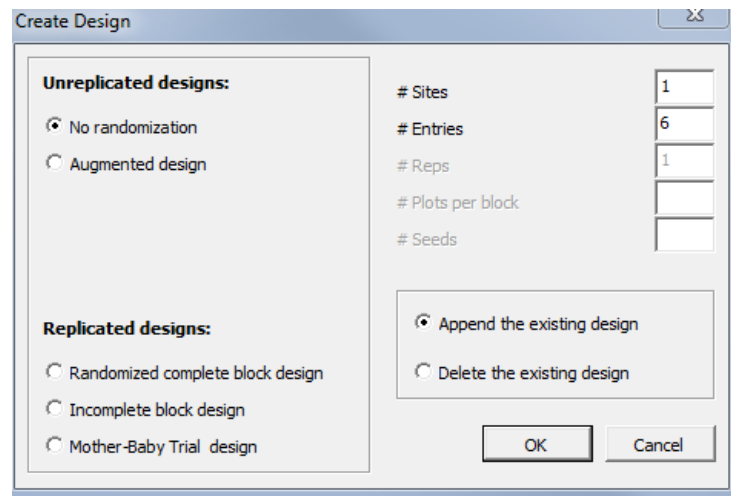


Figure 10 Create Design dialog

for each site. “When some treatments (checks) are replicated  $r$  times and other treatments (new treatments) are replicated less than  $r$  times, an **augmented design** may be used. These designs may be minimum variance designs for estimating contrasts of check effects, of new variety effects, of new variety versus checks, or of all check and new varieties simultaneously”.<sup>6</sup> For an Unreplicated design designs you must indicate the **# of sites** or if you have more than one experiment at a site indicate in the total number of experiments. The # of sites determines the number of field books that will be generated in a subsequent step.

For Replicated designs you can choose **Randomized Complete Block** design, Incomplete block design<sup>7</sup> or Mother-Baby trials<sup>8</sup>. For RCBD indicate the **# of sites** and **# replications per site**.

If you have already created a previous design and want to change it you may use the option Delete the existing design. Note if you have not made a design yet, or if you just opened the file this option may not

<sup>6</sup> On Augmented Designs, W. T. Federer and D. Raghavarao, Biometrics, Vol. 31, No. 1 (Mar., 1975), pp. 29-35, Published by: International Biometric Society, <http://www.jstor.org/stable/2529707?seq=1>

<sup>7</sup> See for example: <https://docs.google.com/viewer?url=http://www.designcomputing.net/pdf/federer.pdf>

<sup>8</sup> See for example: <http://www.cazs.bangor.ac.uk/ccstudio/Library/publications/ReviewPapers/mother.pdf>

be available. In the latter case you may have to create a design twice to be able to delete properly the previous version.

**Click OK to generate design.**

A Dialog will ask if you want to create labels, see below for that option.

Once the design has been generated you must **click on the Design tab**, and **verify that the design** that was generated has the correct number of sites, number of reps, etc. You must also manually define Blocks if that is relevant. It is important to ensure that the design is correct before you generate field books, as the design in the field book is not updated if you change it later on.

Exp	Rep	Block	Plot	Entry	SeedsPerF
1	1	1	1	1	3
1	1	1	1	2	6
1	1	1	1	3	5
1	1	1	1	4	2
1	1	1	1	5	1
1	1	1	1	6	4
1	2	1	7	4	12
1	2	1	8	3	12
1	2	1	9	1	12
1	2	1	10	2	12
1	2	1	11	6	12
1	2	1	12	5	12

Figure 11 Example of design of experiment

In the example you can see how the first column in the design indicates the Experiment number, the second column the Replication within the experiment, the third indicates the Block (must be entered by user as CloneSelector cannot know what blocks are relevant to your field), the fourth is the consecutive list of plot numbers, and the fifth is the Entry numbers from the list of germplasm, and these will be randomized if that is part of the chosen design.

**If you want to use CloneSelector to analyze an existing trial**, you must first enter the list of germplasm as described above and then generate the design. **If the trial was made with a randomized design you should copy the original trial design and overwrite the design that was generated.** You can for example do this by sorting your original design by Experiment, Replication and Plot and then copy the Entry numbers into the design or manually enter the Entry numbers. If you do this your fieldbooks will subsequently have the same layout as your trial, and it will be much easier to enter the data.

### 3.3.1. Make labels for planting material

**CloneSelector -> List of Germplasm -> Seed Labels**

In the Labels: **Choose Sites dialog**, you must **choose for which sites you want to create labels**, typically you want to choose **Select All**.

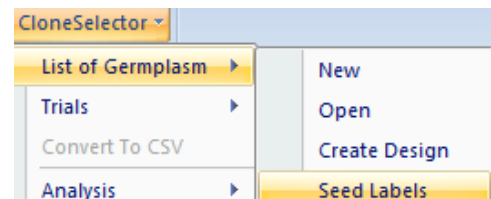


Figure 12 Seed Labels menu

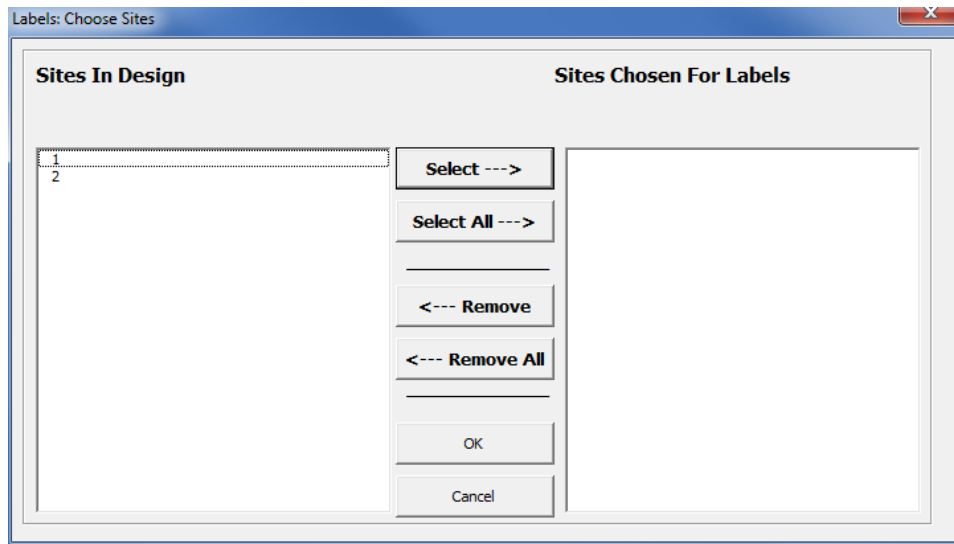


Figure 13 Choose Sites for generating labels dialog

**Click OK to define the sites.**

In **Make Seed Label** select the **# of labels in a row** on your paper. If for example you have labels on an A4 sized sheet you may have 3 labels in each row.

**Click OK to generate labels.**

The labels generated may not have the exact size of your labels, and you may have to adjust the size. **Print a small test sample before trying to print all labels**, in order to be able to adjust the print size to your labels.

**Save** the List of Germplasm and **Close** it.

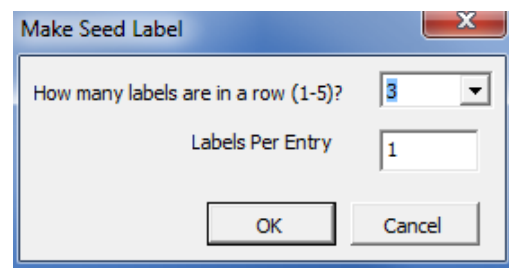


Figure 14 Layout of seed labels dialog

### 3.4 Generation of a field book for each experiment

Make sure **your List of Germplasm is closed** before you start making field books. If this is not the case you may get an error as CloneSelector will attempt to open the file again.

This step will generate one fieldbook in a separate Excel file for each experiment defined in the design of the germplasm list. For example if you entered 3 sites in the Create Design dialog, then CloneSelector will generate 3 fieldbooks. What happens behind the scene is that CloneSelector takes the information from the List of Germplasm and inserts it in all the right places in a fieldbook template. The first time you may find the process a bit confusing, but soon you will appreciate the work saved in this step.

**CloneSelector -> Trials -> New**

In **New Trial Series Dialog**.

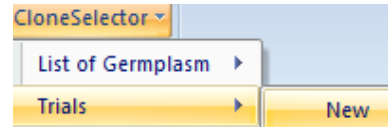


Figure 15 New trial menu

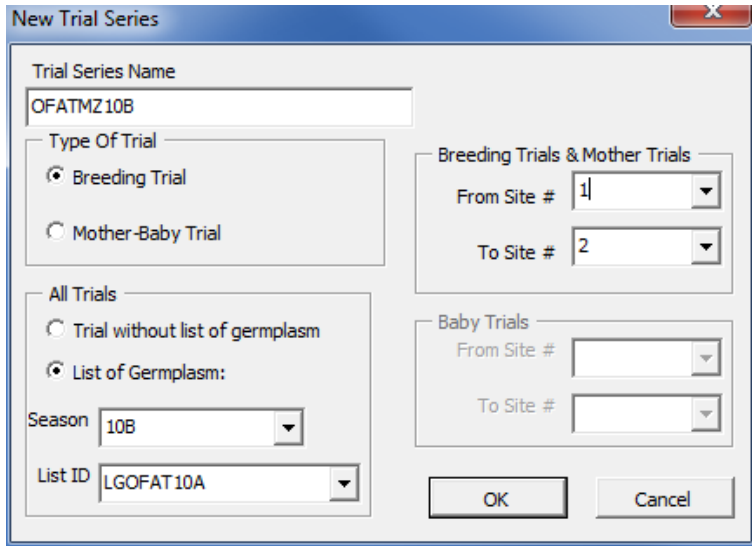
A screenshot of the 'New Trial Series' dialog box. The 'Trial Series Name' field contains 'OFATMZ10B'. Under 'Type Of Trial', 'Breeding Trial' is selected. Under 'All Trials', 'List of Germplasm:' is selected. The 'Season' dropdown is set to '10B' and the 'List ID' dropdown is set to 'LGOFAT10A'. Under 'Breeding Trials & Mother Trials', 'From Site #' is '1' and 'To Site #' is '2'. Under 'Baby Trials', both 'From Site #' and 'To Site #' are empty. 'OK' and 'Cancel' buttons are at the bottom.

Figure 16 New Trial Series dialog

1. **Enter a name for the trial series**, Suggested name is 2 letters for color of material, (OF, WH, YE etc), 2 letters for Type of Trial (OT, AT, etc), 2 letters for breeding program/country, and finally year and season, and example could be OFATMZ10B. **Names should not include spaces**. The name entered will be the name of the Fieldbook files and for each site a number will be added such as for example OFATMZ10B1, OFATMZ10B2 etc.
2. Leave **Type of Trial** as **Breeding Trial** (unless of course it is a Mother Baby trial)
3. In drop down list **Select the Season** in which you put the Germplasm List, typically this should be 10B or something similar.
4. **Select the List ID** that is the name you assigned to the List of Germplasm. Note: If you have not made a Design in the List of Germplasm, you will get a message, in that case click Cancel and open List of Germplasm to create design before proceeding.

- The **system will automatically detect the number of Sites in the From Site # to Site #**, and if that is not correct you have a problem with the design in your List of Germplasm. Cancel and fix before proceeding.

**Click OK to start process of fieldbook generation**

Note: If it is a large germplasm list this may take time, **wait**, do not click or close.

**If it is the first Fieldbook in the season Excel will ask if you want to create a new directory.** Click **Yes**. If you click No the process will be canceled, and you may get an error if you Click OK again. In that case close Click Cancel in New Trial Series and start again.

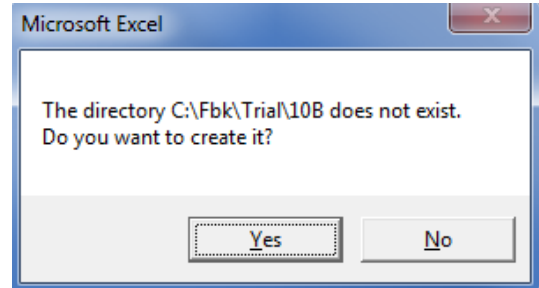


Figure 17 Create directory for trials in new season dialog

**Important in Excel 2007:** Excel will open a **Dialog Compatibility Checker**. Make sure to **Uncheck** the box Check Compatibility When Saving this book, before clicking **Continue**. If you do not uncheck the option you may have to wait for some time. Note: The **Dialog will appear once for each field book**, and if you have for example 4 sites it will appear 4 times.

**CloneSelector generates automatically a new field book for each experiment, but they are closed.**

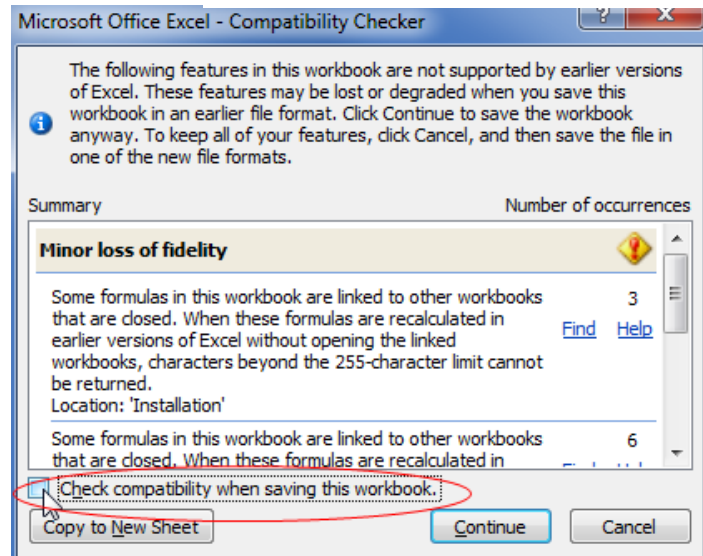


Figure 18 Excel 2007 Compatibility checker problem

### 3.4.1. How to open a field book

To verify the new field books click **Trial -> Open**.

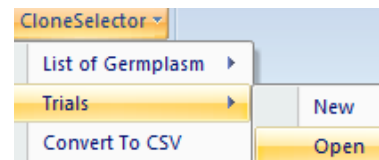


Figure 19 Trial Open menu

Select **Season** and **Trial name** in the dialog. The files will have the Name of the Trial series plus a consecutive number.

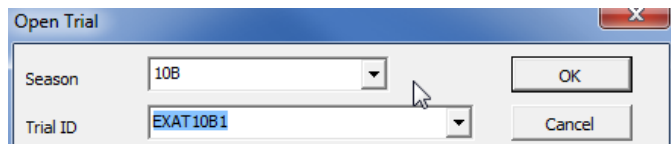


Figure 20 Open Trial dialog

Each Excel workbook has a number of sheets: **Fieldbook**, **Results**, and **Master** are all related to the data being recollected on each entry and the analysis of this information. **General**, **Installation**, **Field**, **Climate\_data**, and **Crop\_management** are all related to the general information on the setup of the experiment. The use of these will be explained in the subsequent steps.

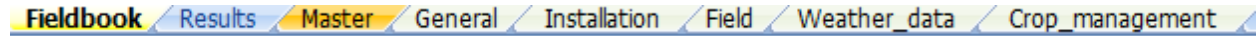


Figure 21 Sheets in a CloneSelector fieldbook

Open the **field books and verify that they are correct**. In particular verify the germplasm list and design of the trial on the Fieldbook sheet. If you identify a problem you must correct it in the List of Germplasm and generate the field books again.

The field books are stored in **C:\Fbk\Trial\Season**, where the season is for example 10B.



### 3.5 Register metadata for each experiment

The interpretation of an experiment is not only based on the specific observations, but also on the context under which they were observed such as for example where was the experiment undertaken, by whom, with what purpose, what was the management, the soil and the weather like? This information can be considered metadata i.e. data about data. In CloneSelector you can register a wide range of metadata and each Fieldbook has five sheets for documenting the experiment:

- **General** with general information on the trial
- **Installation** which is the specific setup. You must add the **Plot size**
- **Field** is soil type and previous management
- **Weather data**
- **Crop management**

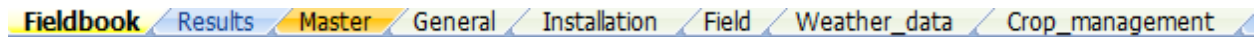


Figure 22 Meta data sheets in Fieldbook

You should not change or delete the standard fields, but if you do not need a field you can leave it blank and consider hiding the row. Below the table of standard fields you can add new fields (rows) if needed.

#### 3.5.1. General information

This form contains general background information on the trial, and includes information on who carried it out, where it was done, and what material was used as checks, and finally it contains the Dublin Core standard information.

Table 1 General information

Factor	Value
Short name or Title: *	
Project Name and Code:	
Name of Contact Scientist: *	
Institution:	
Phone Number Office:	
Phone Number Cell:	
Country: *	
Department	
Province	
District	
Site Name	
Agro-Ecology	
Altitude (masl) *	
Latitude (Degree/Minutes)*	
Longitude (Degree/Minutes)*	

Type of Trial : *	1-Observation Trial (OT), 2-Preliminary (PT), 3- Advanced (AT), 4- On-farm (OFT)
Type of Observational Trial:	Single plant basis
Type of Trial:	
Season	1- Wet, long rains; 2- Wet, short rains; 3- Dry
Check variety name :	
Trait 1 to evaluate during the observational trial:	
Trait 2 to evaluate during the observational trial:	
Trait 3 to evaluate during the observational trial:	
Trait 4 to evaluate during the observational trial:	
Trait 5 to evaluate during the observational trial:	
Number of checks varieties:	
Name of check variety #1:	
Type of check variety #1	
Name of check variety #2:	
Type of check variety #2	
Name of check variety #3:	
Type of check variety #3	
Name of check variety #4:	
Type of check variety #4	
Comments	
Description: *	
Objectives:	
Comments:	
Year	
Version	
Contributor	An entity responsible for making contributions to the resource, typically the name of the breeder
Owner	The owner of the data is typically the institution financing the work or employing the breeder
Subject	Typically, the subject will be represented using keywords, key phrases, or classification codes.
Publisher	An entity responsible for making the resource available.
Date	A point or period of time associated with an event in the lifecycle of the resource
Format	The file format, physical medium, or dimensions of the resource.
Identifier	An unambiguous reference to the resource within a given context.
Source	A related resource from which the described resource is derived.

Language	A language of the resource.
Relation	A related resource.
Rights	Typically, rights information includes a statement about various property rights associated with the resource, including intellectual property rights.
Rightsholder	A person or organization owning or managing rights over the resource.
License	A legal document giving official permission to do something with the resource.
Audience	A class of entity for whom the resource is intended or useful.
Provenance	A statement of any changes in ownership and custody of the resource since its creation that is significant for its authenticity, integrity, and interpretation.
Embargo till	
Status	
Data	

### 3.5.2. Installation information

The installation form contains the information related to establishing the trial in the field. In particular it is important to indicate the **Plot size (m<sup>2</sup>)** as this is used in many calculations such as the conversion of the yield into t/ha.

Table 2 Installation information

Factor	Value
Trial design *	
No of entries *	
No. of replications *	
No. plants per plot *	
No. plants per row	
Number of rows per plot	
Plot size (m <sup>2</sup> ) *	
Distance between plants (m) *	
Distance between rows (m) *	
Density (plants/ha)	
Planting date (mm/dd/yyyy):	
Verification of Establishment (mm/dd/yyyy):	
1st Virus Symptom Evaluation (mm/dd/yyyy):	
2nd Virus Symptom Evaluation (mm/dd/yyyy):	
Harvest date (mm/dd/yyyy):	
Date of vine cutting / killing (mm/dd/yyyy) *	
Crop duration (days):	
Number of plants per genotype:	
Plot type:	
Number of border rows or rows of mounds per plot:	

Number of plants intended for final harvest :	
Cuttings per plot actually used to archive target:	
Target plant spacing within rows (m):	

### 3.5.3. Field information

The field information is related to soil characteristics and prior use of the field.

Table 3 Field information

Factor	Value
Predominant Soil Texture *	
Soil group:	
Organic matter (%) *	
Soil pH (0 - 14) *	
Crop duration (days):	
Crop rotation - Previous crops:	
Crops in rotation # 1:	
Crops in rotation # 2:	
Crops in rotation # 3:	

### 3.5.4. Weather data

The weather data contains monthly average data for the weather in the trial periode.

Table 4 Weather data

	Month	Avg. precipitation (mm)	Average temperature (°C)	Avg. min. temperature (°C)	Avg. max. temperature (°C)	Avg. day/night temp. fluctuation	Relative humidity (%)	Radiation level
1	Jan-YY							
2	Feb-YY							
3	Mar-YY							
4	Apr-YY							
5	May-YY							
6	Jun-YY							
7	Jul-YY							
8	Aug-YY							
9	Sep-YY							
10	Oct-YY							
11	Nov-YY							
12	Dec-YY							

### 3.5.5. Crop management

Crop management indicates how the trial was managed.

Table 5 Crop management

Intervention category	Intervention type	Date of intervention	Name of product *	Product concentration	Dose of application	Observations

Fertilizing						
Weed control						
Pest control						
Disease control						
Irrigation						

## 3.6 Data collection in field and post harvest for each entry

### 3.6.1. Introduction to sweetpotato traits<sup>9</sup>

Gruneberg et al.

[Note: While under some circumstances, there are different meanings of the words “variable, trait and attribute,” for our purposes we use these words synonymously.] In experiments, we distinguish between **classification variables**, and **response variables which have to be analyzed**. Classification variables help us to identify plots and how experimental factors and factor levels are applied to plots (e.g., year, location, genotype, genotype name, and replication). The classification variables can be comparable to our home address by which we can identify who is living where (country, town, street, name). Classification variables allow us to identify a plot, its location, which genotype was planted in it, and how the plot was treated. Moreover, classification variables are needed to provide statistical program packages with information about how the data were organized (e.g. data records are sorted by locations, within locations by genotypes and within genotypes by replications) and they are used to inform statistical procedures about how to analyze the data records. In CloneSelector **the classification variables are mainly entered in the List of Germplasm** and automatically transferred when the fieldbooks are generated, **or the variables are entered as trial metadata** as described in the previous section. **This section is focused on how to collect data on the response variables.**

**Variables which have to be analyzed can be distinguished:** (i) **parametric variables** (real numbers following an approximately normal distribution (e.g., storage root yield, upper biomass yield, storage root dry matter, etc.)), (ii) **non parametric variables** (real numbers which show strong deviations from a normal distribution (e.g., disease damage)), (iii) **rank variables** (e.g., scores with a scale of 1 to 9 for vine strength or scores with a scale of 1 to 5 for storage root fiber content, or scores from 1 to 3 corresponding to the color cards used during farmer assessment of varieties). It should be noted that the analysis of variance (ANOVA) procedure is relatively robust to deviations from the normal distribution, so that even symmetrically distributed rank scores from 1 to 9 can be analyzed by an ANOVA. However, the ANOVA is very sensitive to deviations due to variance in homogeneity of the error [this is the case when a genotype obtains a common score value across replications (e.g., 1 for vine strength = no vine survival), while other genotypes obtain different scores (e.g. 6 to 9 for vine strength = vine survival)]. No variation among replications results in variance in homogeneity, and the requirements of an ANOVA are not fulfilled! Rank variables with scores from 1 to 3 should never be analyzed by an ANOVA; however, the frequency means provide useful information. Rank variables with scores from 1 to 3 have to be analyzed by rank statistics and for these, significance tests and multiple comparison procedures (procedures to compare each clone with a check or among all other clones in the trials) are available.

Note: in some places, established breeding programs may have already developed different rating scales (e.g. 1 to 3, or 1 to 5 or 1 to 9). It is our hope that all of our partners in this collaborative sweetpotato

---

<sup>9</sup> This section builds mainly on Chapter 3: Description of Data Collection Forms in Procedures for the Evaluation and Analysis of Sweetpotato Trials, Gruneberg et al, 2010.

breeding effort will be willing to adopt and use the scales given below, so that all may benefit from the power of the information provided from comparative analysis of our combined results.

**The following presents the standardized traits recommend for sweetpotato.** The traits are grouped according to the different stages in an experiment or the method used for measurement:

- Pre-harvest data
- Harvest data
- Quality data
- Piecemeal harvest data
- Near-infrared reflectance spectroscopy quality data

**In the fieldbook each group of traits have been coded with a color** in order to easily be able to identify them. **Below is a description of the traits in each group, and how they should be measured or assessed.** It is very important that all breeders use the same definition of traits and methods of measurement in order to be able to analyze data jointly from different breeding programs.

Initially you should use the fieldbooks without modifications, but **after a while you will have identified which traits are relevant for your work**, and at that point **you should modify the trial template in accordance with your own needs.** This way you will avoid having to add the same information every time or hide the same columns. However, in order to know what to modify it is preferable to acquire some experience before doing the modifications.

**In the fieldbook you should hide the traits you are not going to measure. Do not delete any columns as this may create errors** and you may not be able to analyze the data. If you need to measure traits that are not in the standard list each group have two or three traits available where breeding programs can define additional traits relevant for their local area.

### 3.6.1.1 Pre-harvest data

In the Fieldbook tab the traits with a light green header color are the ones collected before the harvest

Vines Planted	Plants Established	Virus symp 1-9 6-8 wks	Virus symp 1-9 1 M. bef harv	Vigor 1-9 1 M bef harv	Alternaria Symp1 1-9	Alternaria Symp2 1-9	Trail1	Trail2	Trail3
---------------	--------------------	------------------------	------------------------------	------------------------	----------------------	----------------------	--------	--------	--------

Figure 23 Pre-harvest traits

**Vines Planted**, NOPS = number of plants (cuttings) planted per plot.

**Plants Established**, NOPE= number of plants (cuttings) established per plot (to be determined 3 weeks after planting).

**Virus symp 1-9 6-8 wks**, VIR1 = virus symptoms 1st evaluation (to be determined 4 to 6 weeks after planting; recorded in scores from 1 to 9, where:

1 = No virus symptoms

- 2 = Unclear virus symptoms
- 3 = Clear virus symptoms < 5% of plants per plot
- 4 = clear virus symptoms at 6 to 15% of plants per plot
- 5 = clear virus symptoms at 16 to 33% of plants per plot
- 6 = Clear virus symptoms at 34 to 66% of plants per plot (more than 1/3, less than 2/3)
- 7 = Clear virus symptoms at 67 to 99 % of plants per plot (2/3 to almost all)
- 8 = Clear virus symptoms at all plants per plot (not stunted)
- 9 = Severe virus symptoms in all plants per plot (stunted)

**Virus symp 1-9 1 M. bef harv**, VIR2 = virus symptoms 2nd evaluation (to be determined 1 month before harvest; recorded in scores from 1 to 9 as described for VIR1).

**Vigor 1-9, 1M. pre-harv**, VV = Vine vigor 1st evaluation (to be determined 1 month before harvest; recorded in scores from 1 to 9, where

- 1 = nearly no vines,
- 2= weak vines, thin stems, very long internode distances,
- 3 = weak to medium strong vines, medium thick stems, and long internode distances,
- 4 = medium strong vines, medium thick stems, and medium internode distances,
- 5 = medium strong vines, thick vines, and long internode distances,
- 6 = medium strong vines, thick stems, and medium internode distances,
- 7 = strong vines, thick stems, short internode distances, and medium-long vines,
- 8 = strong vines, thick stems, short internode distances, and long vines,
- 9 = very strong vine strength, thick stems, short internode distances, and very long vines).

**Alternaria symp1 1-9**, 6-8 weeks after planting.

- 1- No symptoms
- 2- Unclear symptoms
- 3- Clear symptoms at <5% per plot
- 4- Clear symptoms at 6 to 15% of plants per plot
- 5- Clear symptoms at 16 to 33% of plants per plot (less than 1/3)
- 6- Clear symptoms at 34 to 66% of plants per plot (more than 1/3, less than 2/3)
- 7- Clear symptoms at 67 to 99 % of plants per plot (2/3 to almost all)



8- Clear symptoms at all plants (not fully defoliated)

9- Severe symptoms at all plants per plot (fully defoliated)

**Alternaria symp2 1-9**, 1 M. pre harvest. Recorded on scale of 1-9 as described in previous trait.

**Trait1, Trait2, Trait3...Trait 12** There are 12 columns available to take data on additional traits if required. Indicate the trait being evaluated in rows 23 and 24 on the Fieldbook sheet. Use a 1 to 9 scale if making a rating, and provide the description of the trait on the Installation sheet. The user can change for example the label Trait1 in row 23 and 24 on the Fieldbook sheet, and also in row 20 and 21 on the Results sheet. There are also 10 notes fields which may be used if desired. Note: For CloneSelector to work, no changes should be made in row 1 on any of the sheets as this will most likely cause problems. Nor should the order of columns be changed.

### 3.6.1.2 Sweetpotato harvest data

The following are the traits to be collected at the sweetpotato harvest

# plants Harvested	# plants w. Roots	# Roots Marketable	# Roots NonMarket	Weight Root Marketable	Weight Root NonMarket	Vine Weight kg	Skin Color 1-9	Secnd Skin Color 0-9	Flesh Color
Prim Flesh Color 1-9	Secnd Flesh Color 0-9	Root Size 1-5	Root Form 1-5	Any Damage 1-5	Weevil Damage 1-5	Trait4	Trait5	Trait6	

Figure 24 Harvest traits

**# plants Harvested**, NOPH = Number of plants harvested.

**# plants w. Roots**, NOPR = Number of plants with storage roots.

**# Roots Marketable**, NOCR = Number of commercial storage roots per NET plot

**# Roots NonMarket**, NONC = Number of non-commercial storage roots per NET plot

**Weight Root Marketable**, CRW = weight of commercial storage roots per NET plot in kg

**Weight Root NonMarket**, NCRW = weight of non- commercial storage roots per NET plot in kg

**Vine Weight kg**, VW = weight of vines per NET plot in kg

**Skin Color 1-9**,

**Flesh color**, FC = storage root flesh color, to be determined on 5 storage roots per plot using color charts provided by CIP, the page number is used as reference.

**Root Size 1-9**, RS = storage root size. Overall assessment of storage root size based on inspection of the harvested roots. Use a 1 to 9 scale, where 1 = excellent, 3 = good, 5 = fair, 7 = poor and 9 = terrible, with numbers in between representing intermediate ratings

**Root Form 1-9**, RF = storage root form. Overall assessment of storage root form based on inspection of the harvested roots. Use a 1 to 9 scale, where 1 = excellent, 3 = good, 5 = fair, 7 = poor and 9 = terrible, with numbers in between representing intermediate ratings.

**Root defects 1-9**, DAMR = defects to storage roots including cracks, veins, etc. [to be recorded on plot basis as scores from 1 to 5; 1 = Very Severe (>60% roots affected), 2 = Severe (30-60% roots affected), 3 = Moderate (10-30% damaged), 4 = Light (few roots affected), 5 = None].

**Weevil Damage 1-9**, WED1 = weevil damage 1st evaluation Overall assessment of weevil damage based on inspection of the harvested roots. Use a 1 to 9 scale, where 1 = no damage, 3 = minor, 5 = moderate, 7 = heavy and 9 = severe damage, with numbers in between representing intermediate ratings.

**Trait4, Trait5, Trait6**, See note with respect to Trait 1, Trait 2, Trait 3 above. If the number of traits recorded exceeds the 3 available here, additional Trait columns may be used from elsewhere in the fieldbook.

### 3.6.1.3 Sweetpotato quality data

The following traits are used to assess the quality of the sweetpotato storage roots. Note in PTs only the top fraction of clones (15 to 25% of all PT clones) needs to be determined, whereas in ATs all clones have to be determined. Note also, that an alternative is to use near-infrared reflectance spectroscopy methods to assess the quality, and the traits used for this are described under NIRS quality data.

Fresh Weight 5 roots	Dry Weight	Method Drying	Fresh Weight Vines	Dry Weight Vines	Method Drying Vines 1-4	Cooked Fiber1 1-5	Cooked Sugars1 1-5	Cooked Starch1 1-5	Cooked Taste1 1-5	Trait7	Trait8
-------------------------	------------	------------------	-----------------------	---------------------	----------------------------	----------------------	-----------------------	-----------------------	----------------------	--------	--------

Figure 25 Quality traits

**Fresh Weight 5 roots**, DMF = fresh weight storage root samples

**Dry weight**, DMD = dry weight of storage root samples

**Method Drying**, DMM = dry matter assessment method (1- Sun-dried, 2- Laboratory oven dried, 3- freeze dried, 4-specific gravity)

**Fresh Weight Vines**, DMVF = fresh weight vine sample

**Dry Weight Vines**, DMVD = dry weight of vine sample

**Method Drying Vines 1-4**, DMMV = dry matter assessment method (1- Sun-dried, 2- Laboratory oven dried, 3- freeze dried, 4-specific gravity)

**Cooked Fiber1 1-5**, COOF1= Fiber in cooked samples assessed by tasting using a 1 to 9 scale where 1 = non-fibrous, 3 = slightly fibrous, 5 = moderately fibrous, 7 = fibrous and 9 = extremely fibrous, with numbers in between representing intermediate ratings.

**Cooked Sugars1 1-9**, COOSU= Sweetness in cooked samples assessed using a 1 to 9 scale where 1 = non-sweet, 3 = barely sweet, 5 = moderately sweet, 7 = sweet and 9 = extremely sweet, with numbers in between representing intermediate ratings.

**Cooked Starch1 1-9**, COOST= Starchiness in cooked samples assessed using a 1 to 9 scale where 1 = very low starch, 3 = slightly starchy, 5 = moderately starchy, 7 = starchy and 9 = very starchy, with numbers in between representing intermediate ratings.

**Cooked Taste1 1-9**, COOT1= Overall taste of cooked samples assessed using a 1 to 9 scale where 1 = excellent, 3 = good, 5 = fair, 7 = poor and 9= horrible, with numbers in between representing intermediate ratings.

**Cooked Appearance 1-9**. COOAP1 = Appearance of cooked samples assessed using a 1 to 9 scale where 1 = very attractive, 3 = attractive, 5 = somewhat attractive, 7 = unattractive and 9 = very unattractive, with numbers in between representing intermediate scores.

**Trait7 and Trait8**. See note with respect to Trait 1, Trait 2, Trait 3 above. If the number of traits recorded exceeds the available space here, additional Trait columns may be used from elsewhere in the fieldbook.

### 3.6.1.4 *Piecemeal harvest data*

The following traits are recommended for piecemeal harvest data

Plot size	Vine	Virus symp 1-9	Weevil	Cooked	Cooked	Cooked	Cooked	Root	Trait9	Trait10
piece harvest	Vigor2 1-9	3ra ev.	Damage2 1-5	Fiber2 1-5	Sugars2 1-5	Starch2 1-5	Taste2 1-5	Sprouting		

Figure 26 Piecemeal harvest traits

**Plot size piece harvest**, PS2 = Size of the plot harvested in piecemeal harvesting.

**Vine Vigor2, 1-9**, VV2 = vine vigor in piecemeal/extended harvest plot (to be determined at final harvest; recorded in scores from 1 to 9, where

- 1 = nearly no vines,
- 2= weak vines, thin stems, very long internode distances,
- 3 = weak to medium strong vines, medium thick stems, and long internode distances,
- 4 = medium strong vines, medium thick stems, and medium internode distances,
- 5 = medium strong vines, thick vines, and long internode distances,
- 6 = medium strong vines, thick stems, and medium internode distances,
- 7 = strong vines, thick stems, short internode distances, and medium-long vines,
- 8 = strong vines, thick stems, short internode distances, and long vines,
- 9 = very strong vine strength, thick stems, short internode distances, and very long vines).

**Virus symp 1-9, 3<sup>rd</sup> ev, VIR3** = Virus evaluation determined at final harvest in piecemeal/extended harvest plot; recorded in scores from 1 to 9, where:

1 = No virus symptoms

2 = Unclear virus symptoms

3 = Clear virus symptoms < 5% of plants per plot

4 = clear virus symptoms at 6 to 15% of plants per plot

5 = clear virus symptoms at 16 to 33% of plants per plot

6 = Clear virus symptoms at 34 to 66% of plants per plot (more than 1/3, less than 2/3)

7 = Clear virus symptoms at 67 to 99 % of plants per plot (2/3 to almost all)

8 = Clear virus symptoms at all plants per plot (not stunted)

9 = Severe virus symptoms in all plants per plot (stunted)

**Weevil Damage2 1-9, WED2** = Overall assessment of weevil damage in piecemeal/extended harvest plot (to be determined at final harvest) based on inspection of the harvested roots. Use a 1 to 9 scale, where 1 = no damage, 3 = minor, 5 = moderate, 7 = heavy and 9 = severe damage, with numbers in between representing intermediate ratings.

**Cooked Fiber2 1-9, COOF2**= Fiber in cooked samples from piecemeal/extended harvest plot (to be determined at final harvest) assessed by tasting using a 1 to 9 scale where 1 = non-fibrous, 3 = slightly fibrous, 5 = moderately fibrous, 7 = fibrous and 9 = extremely fibrous, with numbers in between representing intermediate ratings.

**Cooked Sugars2 1-9, COOSU2**= Sweetness in cooked samples assessed using a 1 to 9 scale where 1 = non-sweet, 3 = barely sweet, 5 = moderately sweet, 7 = sweet and 9 = extremely sweet, with numbers in between representing intermediate ratings.

**Cooked Starch2 1-9, COOST2**= Starchiness in cooked samples assessed using a 1 to 9 scale where 1 = very low starch, 3 = slightly starchy, 5 = moderately starchy, 7 = starchy and 9 = very starchy, with numbers in between representing intermediate ratings.

**Cooked Taste2 1-9, COOT2**= Overall taste of cooked samples assessed using a 1 to 9 scale where 1 = excellent, 3 = good, 5 = fair, 7 = poor and 9= horrible, with numbers in between representing intermediate ratings.

**Cooked Texture2 1-9. COOTX2** = Texture/mouth feel of cooked samples assessed using a 1 to 9 scale where 1 = very moist, 3 = moist, 5 = medium dry, 7 = dry and 9 = very dry, with numbers in between representing intermediate scores.

**Cooked Appearance2 1-9. COOAP2** = Appearance of cooked samples assessed using a 1 to 9 scale where 1 = very attractive, 3 = attractive, 5 = somewhat attractive, 7 = unattractive and 9 = very unattractive, with numbers in between representing intermediate scores.

**Root Sprouting**, RSPR = Root Sprouting noted at final harvest of piecemeal/extended harvest plots. Use a 1 to 9 scale where 1 = absent, 3 = some sprouting, 5 = moderate, 7 = abundant and 9 = profuse sprouting, with numbers in between representing intermediate ratings.

**Trait9 and Trait10**, See note with respect to Trait 1, Trait 2, Trait 3 above. If the number of traits recorded exceeds the available space here, additional Trait columns may be used from elsewhere in the fieldbook.

### 3.6.1.5 Near-infrared reflectance spectroscopy quality data

Near-infrared reflectance spectroscopy, NIRS, can be used for measuring various quality traits on sweetpotato:

Protein	Fe	Zn	Ca	Mg	Beta Carotene	Total Carotenoids	Starch	Fructose	Glucose	Sucrose	Maltose	Trait11	Trait12

Figure 27 NIRS quality data

Protein

Fe

Zn

Ca

Mg

Beta Carotene

Total Carotenoids

Starch

Fructose

Glucose

Sucrose

Maltose

### 3.6.2. Printing the fieldbook sheet

In order to collect data in the field you can print the Fieldbook sheet, however, before doing so you must do some preparations such as adding header information to the fieldbook sheet and hiding the traits that will not be scored in this particular experiment.

### 3.6.2.1 Adding fieldbook headers

When working in the field it is necessary to be able to identify well which fieldbook to use for which experiment and this can be done automatically if you have filled in the relevant metadata as described above.

The fieldbook header includes the following information:

- Name of experiment (the same as the name of the file)
- Site name or location
- Agroecology
- Planting date
- Harvest date
- Plot size
- Collaborator
- Institution

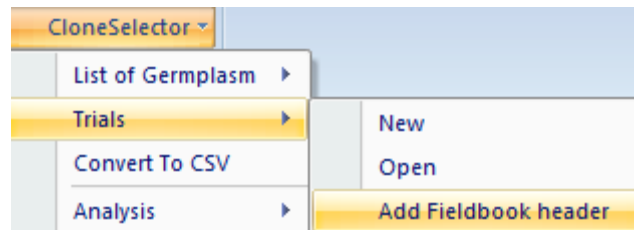


Figure 28 Add Fieldbook header menu

This information should be added in to the metadata (see above) before inserting the fieldbook header

#### **CloneSelector -> Trials -> Add Fieldbook header**

After inserting the header you may have to move it or adjust the size of the box or the rows for it to be properly positioned.

### 3.6.2.2 Hide traits that will not be measured and set print area

The fieldbook include many traits and you may not measure all of them in the particular trial, and before printing you should **hide the traits that will not be used**. Note: **Do not delete columns** as this may cause problems when trying to analyze the data.

#### **Before Printing fieldbooks**

1. **Hide the columns** you do not wish to collect. **Do not delete columns**, only hide.
2. **Define the Print Area** as in any Excel sheet
3. **Use Print Preview** and **Set Print Area** as you would in any other Excel sheet.

Once you have configured the print setup **print the fieldbook as any other Excel sheet**.

Note: If you have problems with printing a fieldbook, you may consider copying the relevant headers and germplasm list into a temporary Excel file, but do only use this for printing.

Note: In some cases you may have problems if you try to hide too many columns, and in that case use printing via a temporary sheet.

**Save the fieldbook once the printing has been set up correctly.**

### 3.7 Enter collected data into electronic field book

Once **data has been collected in the field it must be entered in the fieldbook sheet**. The data is entered as in any normal Excel sheet.

When data is entered **missing values should be entered as a point or dot i.e. "."**. This is very important as **Excel in some cases considers an empty cell as zero and this may lead to errors in your data analysis**.

#### 3.7.1.1 Checking data quality of the fieldbook data

CloneSelector include tools to **verify that the data are on a valid scale** and it will also **flag outliers in relation to a normal distribution**, however, these tools will only detect the obvious errors and **it is very important that you take precautions to ensure the data is transcribed without errors**.

After data Entry you must check the data quality of the fieldbook. To do this it is useful to have a paper copy of the data present so that the problems identified can be corrected immediately.

**CloneSelector -> Analysis -> 1. Check Field Data**

A **Dialog Correct Errors/Outliers** will appear in the upper left corner if problems are encountered.

If you are not able to enter the correct data **you must enter a point i.e. "." as a missing value**. If you leave the cell blank then this may cause problems when calculating the derived values in the Master sheet as **Excel considers an empty cell as zero, and instead of a missing value the calculation may return a zero, and you would end up with a difficult to detect error in the results sheet**.

In the Correct Errors Dialog the first line indicates the type of error or problem for example in this case a 10 is an invalid data as the scale is only 1-9, and CloneSelector propose to change this to a "." i.e. a missing value.

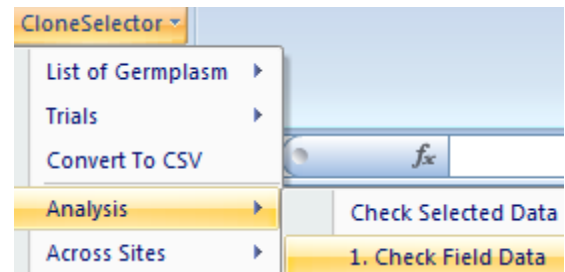


Figure 29 Check Field Data function

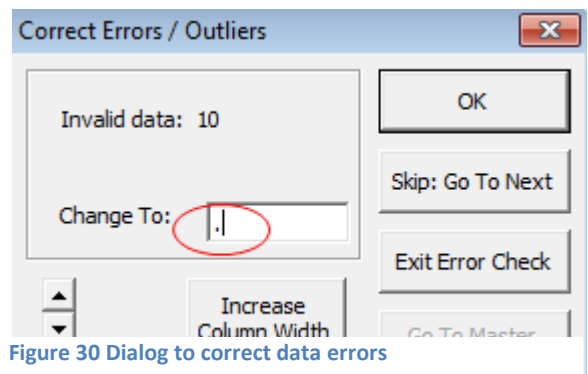


Figure 30 Dialog to correct data errors

The data check is both for false values such as letters in a numeric columns or outliers. The latter is values that are more than two standard deviations away from the normal distribution of the other values observed. In columns with very homogenous results, this may well flag correct values, and the check is only meant for you to ensure that the value is not an entry error.

In the Dialog **you may enter a new value** if you discover a mistake, or **accept the automatic proposal** by clicking **OK**. When clicking Ok the **value of the Cell is changed**, and it is **colored yellow** to flag that a change was made to the data.

The Option **Skip, Go to Next** will leave the value unchanged, but will color the cell red so that you can check it later.

The check is finished with a **Dialog Data Change Completed**.

**Check all the cells in red before proceeding with the data analysis.** If you subsequently change a data value you will have to redo the other steps, so **it is important to complete one step fully before moving on to the next.**

**Save the Fieldbook after data check of field data.**



### 3.8 Calculation of derived variables

Several of the traits observed in the field may need to be converted in order to be more easily interpreted or be comparable with results from other trials. For example the yield per plot should be converted to yield per hectare and the number of plants harvested should be compared to the number of vines planted in order to analyze the vine survival. The following table includes some of the most commonly derived variables used in sweetpotato breeding.

Table 6 Calculated fields for sweetpotato

Field to be calculated	Notes on how to calculate the field
Average commercial root size	Weight of commercial roots CRW/ # commercial roots NOCR
Biomass yield	Total root yield + foliage yield
Commercial yield t/ ha	Weight commercial root CRW* 10/plot area m2
Dry matter foliage yield	Dry matter foliage % * total foliage yield/100
Dry matter root yield	Dry matter % * total root yield/100
Foliage total yield per hectare	FYTha VW*10/plot area m2
Harvest index	Storage root yield (CRW + NCRW) / (foliage yield VW + storage yield (CRW + NCRW))
Harvest index rank	Ranking of entries according to harvest index rank
Harvest sowing index (survival)	Number of plants harvested/ # plants planted per net plot NOPH/NOPS
Non commercial yield t/ ha	Weight non commercial root NCRW* 10/plot area m2
Number of roots per plant	Total # roots NOCR+NONC / # plants harvested NOPH
Percent marketable roots (commercial index)	Weight marketable CRW/ total root yield (CRW + NCRW)
Root foliage ratio	Dry matter root yield/ dry matter foliage yield
Storage root dry matter content (%)	Dry matter weight (DMD)/ fresh weight (DMF) * 100
Yield of total roots per hectare t/ha	Commercial root yield + non commercial root yield 10*[(CRW + NCRW)/plot area in m2]
Yield per plant Kg	(CRW+NCRW)/NOPH
Yield rank in trial	Ranking of entries according to total yield

In CloneSelector this conversion or calculation is done in the **Master sheet where the first line contains the formulas for calculating the variable**. You can click at the formulas in the first data row and you will see that the formulas are simple Excel formulas that take the data from the Fieldbook sheet and calculate the relevant variables.

#### 3.8.1.1 Calculating variable in CloneSelector

The variables are calculated almost automatically, however, for the best result you must do some preparations:

1. **Finish the Check Fieldbook data quality** as described in the previous step and ensure that all data is valid or if it is missing a "." has been inserted.
2. **Make sure that the Plot size has been entered on the Installation sheet**. This value is used in all conversions to hectares.

3. Note: **Do not hide or unhide columns in the Master sheet** as this may cause problems when running the statistical analysis.
4. On the Master sheet check in the first second row the formulas that calculate the values. Especially the values #Div/0 . Delete the formula if the problem is that you are not collecting the data in this experiment.

Once you have done the preparations and checked the formulas you are ready to do the calculation:

**CloneSelector -> Analysis -> 2. Fill in Master sheet**

This should copy the formulas down to fill in all the rows in trial.

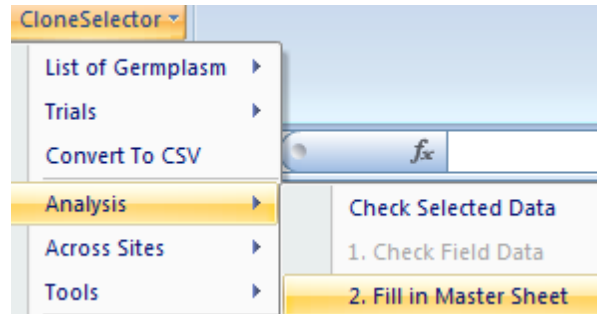


Figure 31 Fill in Master sheet

Next step is to check the Master data for errors and particularly to solve issues with missing data:

**CloneSelector -> Analysis -> 3. Check Master Data**

The two most common error in the Master data are generated by

1. An empty value in the fieldbook is indicated as "." but in the Master sheet it must be an empty cell for R to analyze the data correctly.
2. If the "." In Fieldbook is used in a formula in the Master it will generate an error indicated as #VALUE!

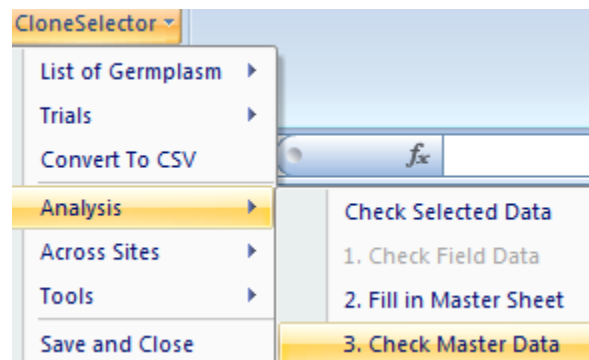


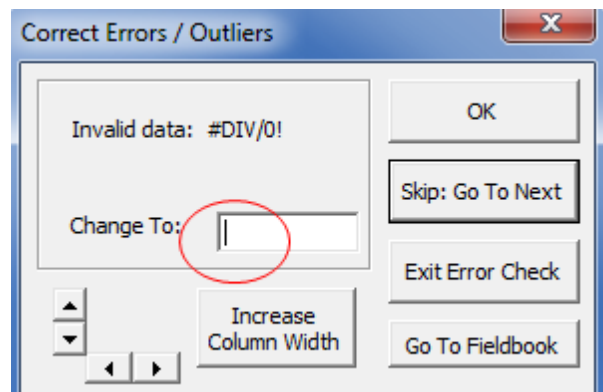
Figure 32 Check Master Data

In both cases **Check Master Data will automatically change the value in the cell to empty and the color to yellow to indicate the change.**

A **Dialog Correct Errors/Outliers** will appear in the upper left corner if other problems than the two previously mentioned are encountered.

If you are not able to enter the correct data **you must leave the cell blank for a missing value.** RExcel which is used for the analysis will give an error if you enter a "." Or anything else that a blank cell for missing values.

In the Correct Errors Dialog the first line indicates



the type of error or problem for example in this case a missing value that creates a “division by zero error”, and CloneSelector propose to change this to a blank cell i.e. a missing value.

In the Dialog **you may enter a new value** if you discover a mistake, or **accept the automatic proposal** by clicking **OK**. When clicking Ok the **value of the Cell is changed**, and it is **colored yellow** to flag that a change was made to the data.

The Option **Skip, Go to Next will leave the value unchanged**, but will **color the cell red** so that you can check it later.

The check is finished with a **Dialog Data Change Completed**.

**Check all the cells in red before proceeding with the data analysis.** If you subsequently change a data value you will have to redo the other steps, so it is important to complete one step fully before moving on to the next.

**Save the Fieldbook after the master data has been checked**

### 3.9 Statistical analysis of trial

CloneSelector include an option for analyzing a trial in a randomized complete block design with a standardized method that generates 2 outputs

1. A table with the mean value across the replications of each trait for each entry
2. A table with statistical information for each trait
  - a. Mean value of the observations of the trait
  - b. Variance estimation for genotypes
  - c. Variance estimation for blocks
  - d. Residual variance estimation
  - e. Coefficient of variation
  - f. Minimum value
  - g. Maximum value

The statistical variables can only be calculated if the trait has some variation within, and are not calculated for values that are the same in all plots such as for example Number of vines planted.

**Before doing the Statistical analysis you must Check the Master data** as described in the previous step.

**Make sure your data is Saved before attempting the analysis.**

Note: The statistical analysis **requires RExcel to be installed**, and CloneSelector must have the correct Reference to this. RExcel must also be correctly Referenced which is done in the VBA Editor under Tools -> Reference, where there should be a reference to RExcelVBA Lib. This is a different library in Excel 2003 and 2007. In some cases it may have a reference such as for example Missing RExcelVBA LIB 2007 in Office 2003, in that case just remove the reference, and make the correct one. **If this has not been**

**done, you probably need help from an IT person to do this.** We will be working on a more user-friendly installer, but for now you may need expert help.

To do the analysis:

#### **CloneSelector -> Analysis -> 4. Analyze Master Data**

This takes time so **WAIT**, it may take several minutes, depending on the size of your trial and your computer. **When it is finished it will automatically show the Results sheet.**

Note: This step implies sending data to a statistical program, and many things can go wrong, so error may occur, and it is important to save before running the analysis.

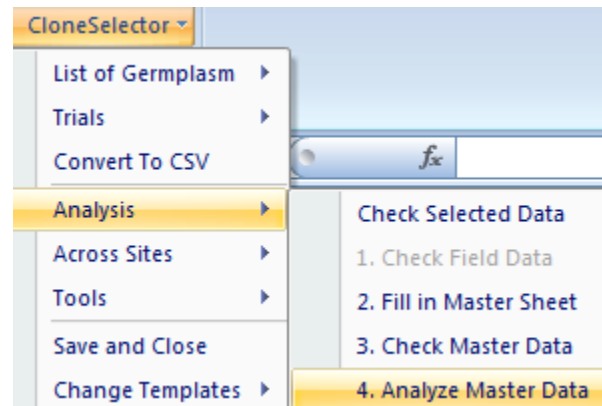


Figure 33 Analyze master data

Check that the Result columns are correctly aligned with the labels on the Result table. The Row just above the first Entry has the name of the column in the Master Sheet, and this should correspond to the column in the Result table. If they do not coincide you have probably hidden/unhidden columns in the Master sheet. To correct it you can either try to revert the Master sheet back to the original state, or manually move the content of the result table. We are working on making this more robust.

### **3.10 Elaboration of consolidated result table for each trial**

After running Analyze Master Data in the previous step the Result sheet should have two tables:

1. A table with the mean value across the replications of each trait for each entry
2. A table with statistical information for each trait
  - a. Mean value of the observations of the trait
  - b. Variance estimation for genotypes
  - c. Variance estimation for blocks
  - d. Residual variance estimation
  - e. Coefficient of variation
  - f. Minimum value
  - g. Maximum value

The final step in the analysis is to check that the Result data are within the range defined for each trait.

On the **Result Sheet**, **below the Statistical data** you will see two rows, **Lower and Upper Limit**, and in these you can define **the range that is valid for each trait**. In the initial use of CloneSelector you may change the values directly in the Result Sheet, but in order not to do this every time you may also consider changing it in the Trial Template as described in another section.

Lower Limit				0.00	1	0.0		1.0	1.0	
Upper Limit				50.00		100.0		5.0	9.0	

Figure 34 Lower and Upper limits for results

Once you have defined the valid ranges or simply accepting the default values, you should check the result data:

**CloneSelector -> Analysis -> Check Result Data**

**CloneSelector will now apply a yellow color to all results that are outside the defined range for the trait.** It will also check the Min and Max value for each trait, so that if you have many entries you only have to revise the Min and Max to see if there is a problem in a trait.

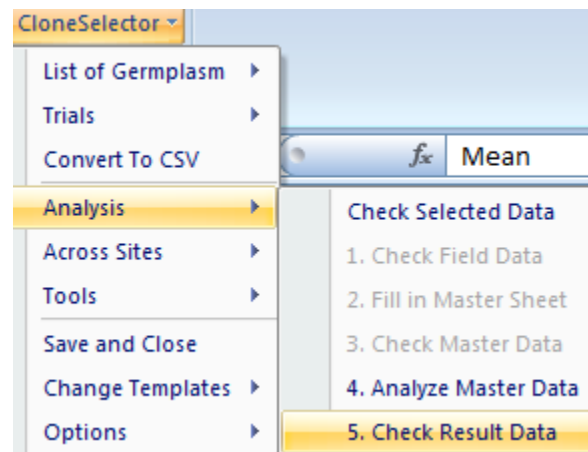


Figure 35 Check result menu

That a result is **out of bounds** may have different **explanations**:

- It could be a **problem with the data on the Fieldbook sheet**, but this is not so likely if you checked the Fieldbook data and corrected any problem found there. If you find that you have a problem on the Fieldbook sheet, you should verify that the Upper and Lower limits are the same for the trait on the Fieldbook and Results sheets. If you correct data in the Fieldbook sheet you must rerun the whole analytical process, from Check Fieldbook data and all subsequent steps.
- You may also have a **problem with formula on the Master sheet**, particularly, a problem of the data not being converted to the correct scale, and for example your data range may be defined considering results in gram, but the actual data has been calculated in kg. In this case you can
  - Change the Upper and Lower limits to reflect the used unit
  - Or change the formula used in the first line on the Master Sheet, and you must then rerun the analysis from Calculate derived fields above i.e. initiate with Fill in Master Sheet, and then redo the subsequent steps.
- Or you may find that your **Upper or Lower Limits are not correctly defined on Result sheet**, in which case you should correct them and rerun Check Result data. Or you may accept that in this particular case the result is acceptable.

What is important to understand is that **when you change data in the Fieldbook, formulas on the Master sheet or limits on the result sheet you must rerun the analytical process from the relevant step.**

## 4. Other functions in CloneSelector

The previous section of the CloneSelector outlines the uses related to carrying out a trial which is the main workflow at this point in time. This section explains some additional functionality which may more occasionally be useful such as for example sorting the fieldbook observations or the result sheet.

### 4.1 Tools for sorting data and results

The Fieldbook and result sheets in CloneSelector have quite a complicated structure and if you need to sort them according to a certain column you have to be careful not to accidentally mix up your data. To avoid problems it is therefore recommend using the sorting tools presented in this section.

#### 4.1.1. Sorting Fieldbook according to entry number

Sorting the fieldbook according to entry number will order the fieldbook so that all the plots for one entry can be observed easily. To sort by entry the Fieldbook must be the active sheet or the menu will not be available.

**CloneSelector -> Tools ->Fieldbook -> Sort by Entry**

After this you can observe in the Entry column how the fieldbook has been ordered i.e. it starts with all the replicates of entry number one, followed by entry number two etc.

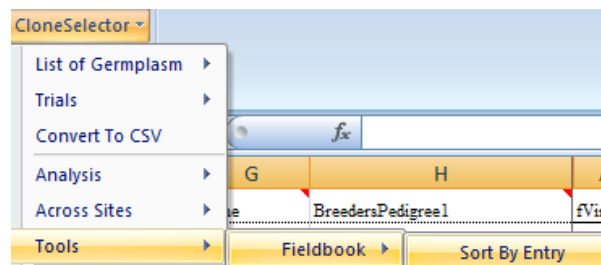


Figure 36 Sort fieldbook by entry menu

#### 4.1.2. Sorting Fieldbook according to plot number

Sorting the fieldbook according to plot number is the default situation and it will bring the fieldbook back to its original order, which means for example the entries will be show the way they were randomized. To sort by plot the Fieldbook must be the active sheet or the menu will not be available.

**CloneSelector -> Tools ->Fieldbook -> Sort by Plot**

After this you can observe in the Plot column how the fieldbook has been ordered i.e. it starts with plot one, followed by plot two etc.

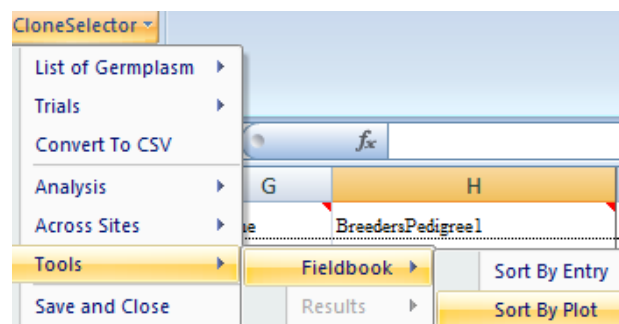


Figure 37 Sort fieldbook by plot menu

#### 4.1.3. Sorting Result table according to user selected column

As part of the analysis of the results it can be useful to sort the Results table according to a column such as for example Total Root Yield. To do this the Results sheet must be active.

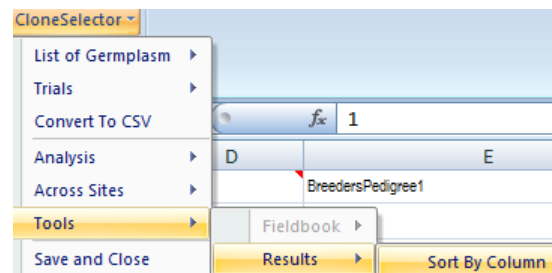


Figure 38 Sort results by column menu

## CloneSelector -> Tools -> Results -> Sort by Column

After clicking the menu a Sort dialog appears where you must **enter the column which will determine the sort order** and you can also **choose if it should be ascending or descending**. Once you have entered the information **click OK** to do the sorting. After sorting by yield you can for example easily fill in the Yield Rank column.

The default sort order in results is according to Entry, and to return to the original table, you can enter B as the Sort by Column and Sort Ascending which will then return the results to the original order.

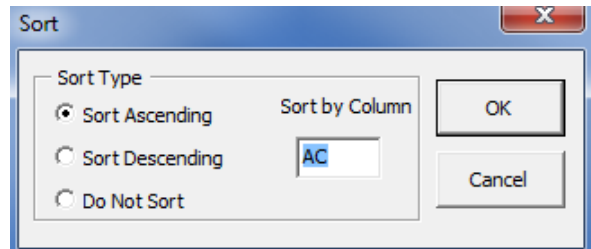


Figure 39 Sort results dialog

## 4.2 Copy existing data set into a CloneSelector Fieldbook

You may have a trial you already carried out or are in the process of carrying out, and you would like to enter the data in CloneSelector, so that you can verify the data and analyze the data. To do this you must first create a Fieldbook with the same design as your original design, and subsequently copy over the data.

### 4.2.1. How to create a CloneSelector Fieldbook based on an existing design

It is assumed that you already know how to create a germplasm list with a design, and how to generate a Fieldbook from this. If that is not the case then consult the relevant sections in chapter 2. **The critical issue is in Step 3 where you must ensure you use the same randomization as in your original trial**, as that is the only way to copy over the data with a limited effort.

**Step 1:** Generate a Germplasm List and enter the list of germplasm from your original trial (see also 2.2 Elaboration of list of germplasm to test).

Note: CloneSelector requires Entry numbers to be a consecutive list of numbers starting with 1, and if that is not the case in your original experiment, you must assign new Entry numbers.

**Step 2:** Generate a Design of a field trial with the same number of sites and repetitions as you have in your historical data for example 2 sites with 3 reps (see 2.3 Design of Field Trial with one or more experiments)

**Step 3:** From your original trial you must enter the randomization of the entries into the Entry column on the Design sheet in

Exp	Rep	Block	Plot	Entry	SeedsPerPac
1	1	1	1	1	4
1	1	1	1	2	5
1	1	1	1	3	1
1	1	1	1	4	3
1	1	1	1	5	2
1	1	1	1	6	6
1	2	1	1	7	5
1	2	1	1	8	6
1	2	1	1	9	1
1	2	1	1	10	4
1	2	1	1	11	3
1	2	1	1	12	2
1	3	1	1	13	2
1	3	1	1	14	1
1	3	1	1	15	3
1	3	1	1	16	5
1	3	1	1	17	4
1	3	1	1	18	6

the Germplasm list. This is very important as this will help you copy/paste or enter your data into the Fieldbook. (see 2.3 Design of Field Trial with one or more experiments)

If your data is in Excel, you should make sure the data is sorted according to Rep before you copy the Randomization over.

After you have modified the randomization to match the original trial **Save the Germplasm list.**

**Step 4:** Generate the CloneSelector Fieldbooks using the Germplasm list you created in the previous steps. Note: It is critical you have entered the original randomization in the Design before you generate the fieldbooks. (See 2.4 Generation of a fieldbook for each experiment)

#### 4.2.2. Enter or copy your existing data into the CloneSelector Fieldbook

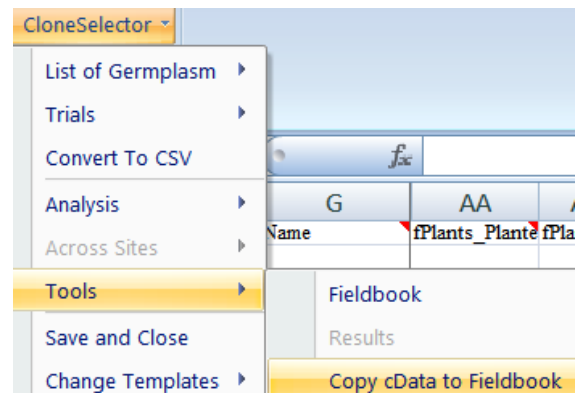
Before you do this you must as described in the previous section Generate a Fieldbook in CloneSelector with the same design as your original data.

**Step 1:** If your data is on paper you will have to transcribe them into Excel. If you already have the data in Excel you can obviously omit this step.

	A	B	C	D	E	F	G
1	32	32	5	2	8	1	
2	32	30	3	1	6	3	
3	32	32	4	3	5	2	
4	32	32	7	2	4	2	
5	32	31	8	1	5	1	
6	32	32	5	3	3	3	
7	32	30	8	2	5	4	
8	32	30	6	1	8	4	
9	32	30	5	3	6	2	
10	32	31	4	2	3	1	
11	32	30	5	2	4	1	
12	32	27	3	1	5	2	
13	32	30	5	3	5	1	
14	32	31	8	4	3	3	
15	32	30	6	4	4	2	
16	32	31	3	2	7	1	
17	32	32	4	1	8	3	
18	32	32	5	1	5	2	
19	NOPS	NOPE	PUNI	PGH	PVIG	FLOD	

Installation Field Weather\_data Crop\_management cData

**Step2:** Copy your data into the cData sheet and insert **below** each column the official abbreviation for each trait. The abbreviations are listed on the Fieldbook Sheet in the second row. Note the abbreviations must be below and not above the columns with data. The order of the columns on the cData is not important as long as you have the correct abbreviation in each one. When the data is copied to the Fieldbook sheet, they will be reordered according by matching the abbreviations.

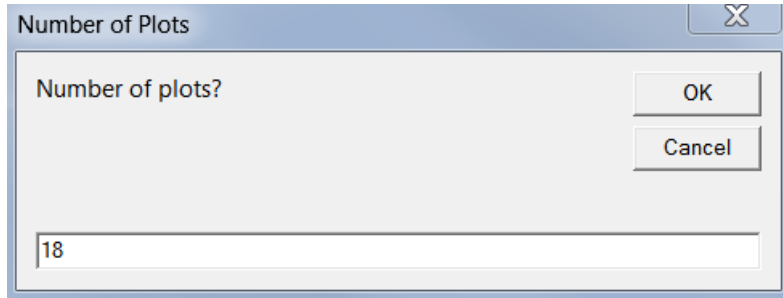


**Step 3:** On the Fieldbook or cData sheet select:

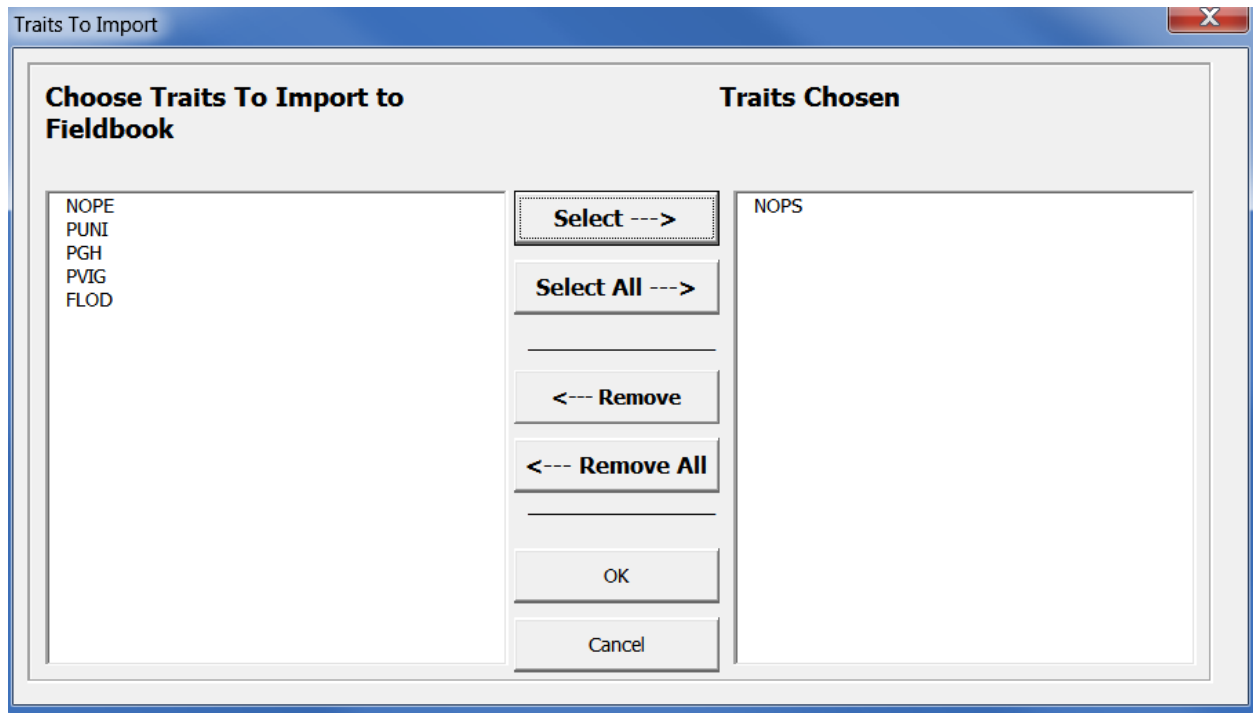
**CloneSelector -> Tools -> Copy cData to Fieldbook**



**Step 4:** Confirm the number of plots by **clicking OK in the Number of plots dialog**. If the number of plots in your data is not equal to the number of plots in your Fieldbook, the automatic import may not work. Though, in some cases you may have data with fewer repetitions.



**Step 5:** Choose the traits to import to the Fieldbook from cData by selecting traits and **clicking Select or clicking Select All**. The traits in the list Traits chosen will be imported. Remember you must put the trait abbreviations below the data in each column in the cData sheet. When you have selected the traits you want to import **click OK**.



**Step 6:** On the Fieldbook **verify** that your data was imported correctly, and **Save** the work sheet.

### 4.3 Saving files in CloneSelector

CloneSelector files are generally just normal Excel spreadsheets and you can save and close them as any other spreadsheet. However, CloneSelector also

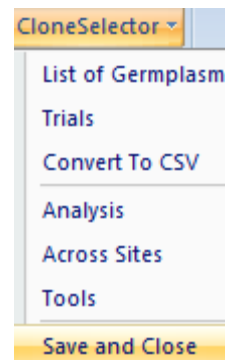


Figure 40 Save and Close menu.

include a menu that will save and close the current sheet with just one click.

### CloneSelector -> Save and Close

The outcome of doing it this way is no different from just using Excel menus.

CloneSelector automatically stores the Germplasm lists and trials in a folder structure order by seasons, and assigning standardized names to the files. This will help you organize your work, and at the same time ensure the good functionality of CloneSelector, so it is recommend to use the default options with regard to naming and storage locations.

## 4.4 Convert Fieldbook to CSV

CSV is an abbreviation for comma separated value file and is a standard format for data exchange of information presented in tables. Converting fieldbooks to CSV is relevant for using CloneSelector with FieldLog, which is an application for logging trial data on a hand held computer, a PocketPC. At this stage the functionality has not been tested, however, **it is expected that over the next year we will do some testing of the use of hand held computers for data logging.** More complete information will be included in this section once the testing has been done.

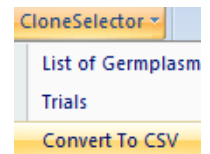


Figure 41 Convert to CSV menu

### CloneSelector -> Convert To CSV

In practice you will want to for example convert all your fieldbooks from one season and store them on the hand held computer and the conversion is therefore set up for bulk conversion so that you select all the files you want to convert and it is all done in one operation.

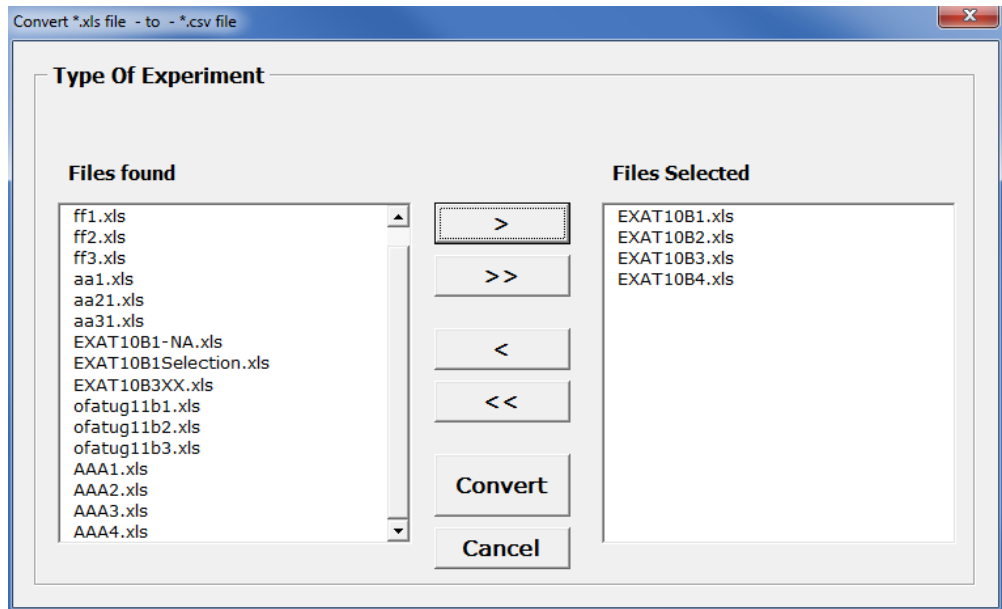


Figure 42 Convert to CSV dialog

To select one or more files click on them and then click > to move over to selected or use < to move a file back if a mistake was made.

Once you have selected the files to be converted click **Convert** to do the conversion.

The CSV files will be saved in the Trial and season folder with the name CSV + current time. Inside the folder you will find the fieldbooks of all the selected files converted to CSV format.

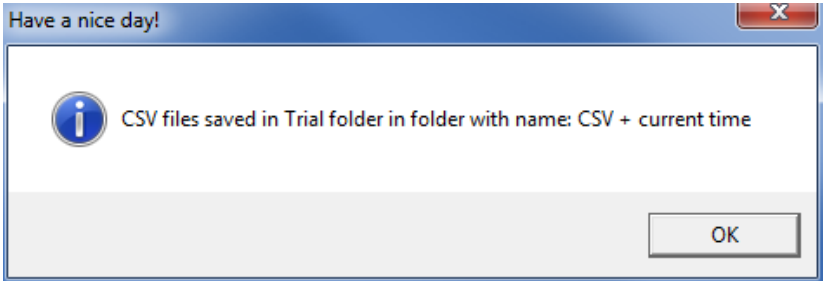


Figure 43 Dialog advising that CSV conversion has finished.

Computer ▶ OS (C:) ▶ CloneSelector ▶ Trial ▶ 10B ▶ CSV\_20100901\_14.9

Figure 44 Example of name and location of CSV file.

## 4.5 Recreate Master Formulas

You may occasionally have deleted or changed the formulas on the Master sheet, and to recreate them manually is quite a lot of work and prone to errors. CloneSelector has therefore a function that recreates the formulas as in the original template. Note that if you have changed the default formulas, you will have to enter the changes again.

One common way to accidentally delete the master formulas is when there is a missing value in the first row of the Fieldbook sheet, which in turn may lead to missing values in the first row of the Master Sheet. If the Fill in Master sheet is done followed by the Check master sheet, then cells in the first row on the Master sheet will be left empty as the formulas are deleted. If the Fill in Master sheet is clicked again, this will delete the formulas in the whole column. If you do the analysis of a trial repeatedly, this may happen, but if you only analyze it once you will not have the problem.

To recreate the formulas in the Master sheet:

**CloneSelector -> Tools -> Recreate Master formulae**

This will insert the formulae in the first row with data on the Master Sheet. You should now check the formulas and include any changes to the default formulas you had made. Once the formulas are ok, and then click

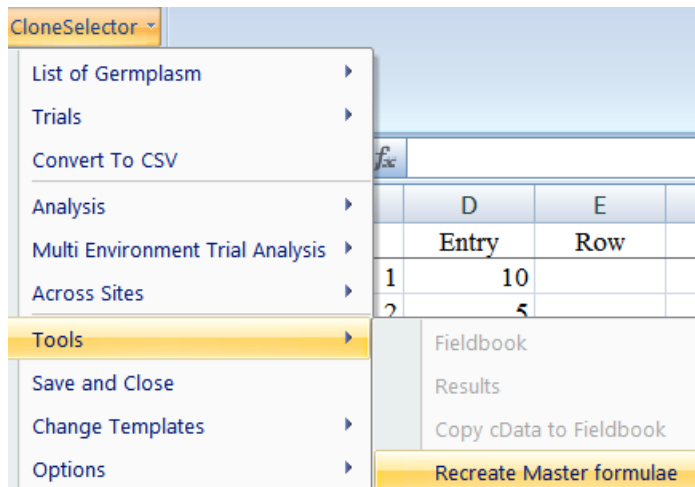


Figure 45 Recreate Master formulae

**CloneSelector -> Analysis -> Fill in Master Sheet**

This will copy the formulas from the first data row to the rest of the Master Sheet.

#### 4.6 Fix hide columns problem

Especially in the Fieldbook sheet you may want to hide many columns, but occasionally when trying to hide a large number of columns you will get the following message:

And you are not able to hide any more columns.

If this happens you simply click:

**CloneSelector -> Tools -> Fix hide columns problem**

This will solve the problem, and afterwards you can hide as many columns as you want.

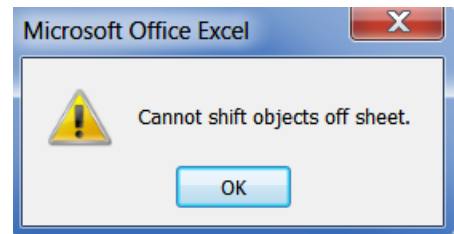


Figure 46 Problem when trying to hide columns

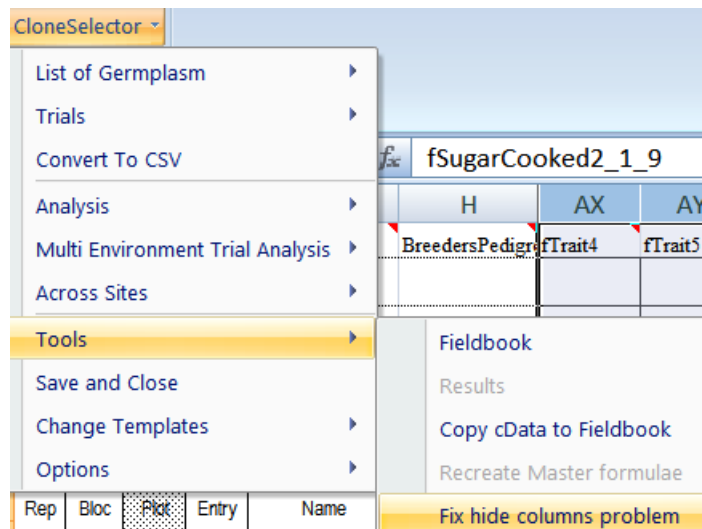


Figure 47 Fixing hide column problem

#### 4.7 Solve Automatic Calculation problem

Occasionally on the Master sheet you will see that all the numbers in a column are the same, even when the data on the Fieldbook varies for the trait. This indicates that Excel is not recalculating automatically the workbook, and that it has been set to manual recalculation. You can solve this in Excel by going into Excel Options and under Formulas set the Calculation options to Automatic, however, in CloneSelector you can also

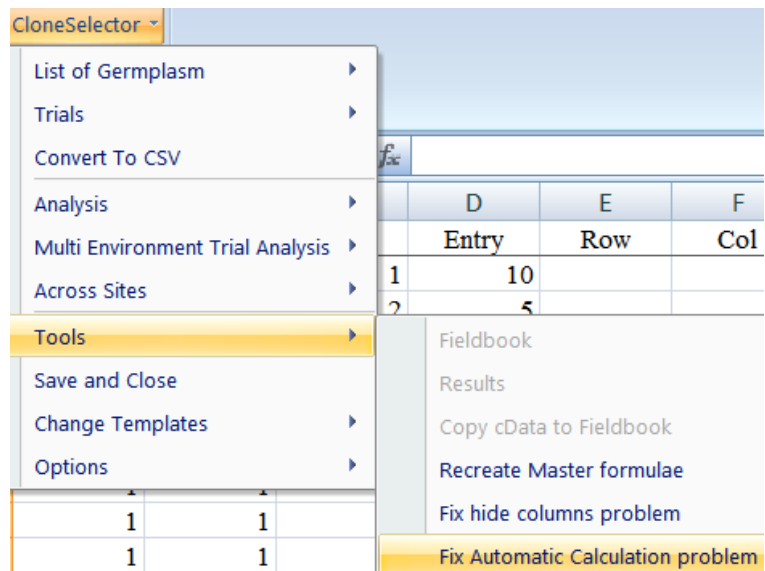


Figure 48 Fix automatic calculation problem

simply click:

**CloneSelector -> Tools -> Fix Automatic Calculation problem**

This will set Excel to automatically recalculate after each change as is the default function.

## 5. Customizing CloneSelector to your needs

The potato version of CloneSelector is developed in accordance with the established protocol for potato breeding and as such it should be generally applicable for most potato breeders, however, you may still have to do some modification to work as efficiently as possible.

1. If you have not placed the CloneSelector folder in the default location directly on the C disk, then you will have to adjust the file locations in the options menu as explained below
2. Once you have acquired some experience with CloneSelector you may decide to modify the trial template in order not to have to repeat the same modifications for each trial. This does however, require some general understanding of the different elements of CloneSelector and it is recommended to carry out a few trials before this is attempted. For the specific instructions see below.

However, before you start out modifying CloneSelector you should understand the basic structure of the files and folders used as is described in the following section.

### 5.1 CloneSelector file structure

The CloneSelector folder contains a number of folders and files.

**The Excel file CloneSelector** is the file you must open to start CloneSelector.

Note: with time the CloneSelector file may get corrupted and will not work properly. It is therefore recommend that you maintain a backup of CloneSelector, and if you experience problems, you can simply copy the backup version into the CloneSelector folder. To do so make sure that Excel is closed.

The **fieldbook.R** document is an R script used for analyzing the CloneSelector data with R. With time we hope to include more analytical options.

The **tpFbk** folder is where the templates are stored. Everytime CloneSelector generates a new Germplasm List or Trial it is based on a template, and these are stored in **tpFbk/Potato**. The modification of the templates is described below.

The **GermplasmList** and **Trial** folders are for storing user data. A new Germplasm List is sorted in GermplasmList/Season and a trial in Trial/Season where Season is folders named for example 10A and 10B for the two seasons in 2010. CloneSelector automatically stores the information in these folders, however, for this to work well the user should adopt naming conventions of both seasons and trial series.

**BackupCopy** contains a backup of the CloneSelector application and the templates in case any of them get corrupted. Should that happen simply copy the relevant files into the original place and restart CloneSelector by closing down Excel and reopening CloneSelector.

**Summary** is currently not used, but the intention is to be able to generate season summaries in this folder.

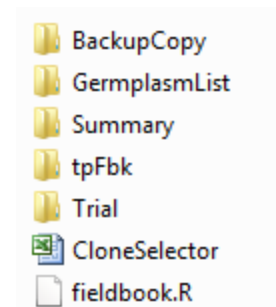
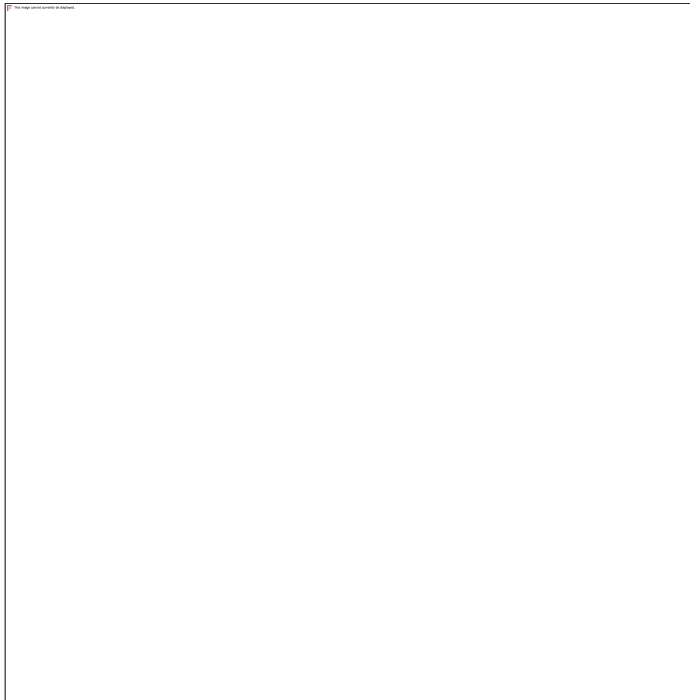


Figure 49 CloneSelector file structure

## 5.2 Setting up the location of the CloneSelector files and templates

CloneSelector aims to help automate and organize certain parts of the trial process. To do this it needs to be able to find the templates used for example for generating a trial or for storing the fieldbooks. CloneSelector is made to be installed by default directly in on the C drive i.e. in C:\\CloneSelector. If you install it in another place you will need to adjust the File Locations.

### CloneSelector -> Options -> File Locations



This will open the File Locations dialog. To change the locations you must click **Browse for CloneSelector Folder** and **Browse For tpFbk Folder**.

If you click one of the **Browse For** buttons that will open the **Browse for Folder dialog**, and you must here locate the relevant folder. For **CloneSelector folder** you must locate a folder named **CloneSelector** and for the templates the folder must be named **tpFbk** or they will not be accepted. When you have selected the relevant folder click OK to confirm the selection. To close the File Locations dialog and save your selection click OK again.

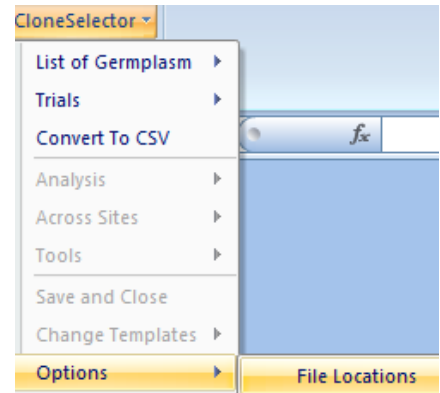


Figure 50 File Locations menu

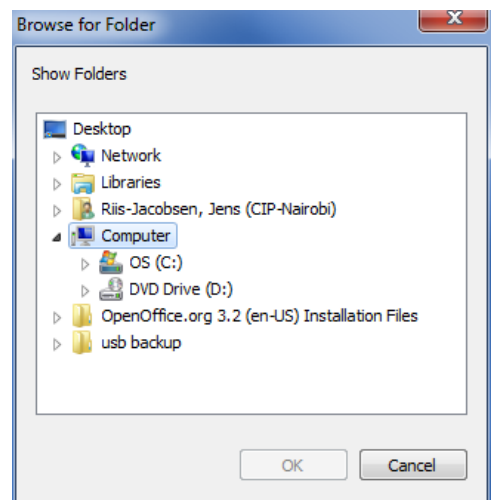


Figure 51 File Locations dialog

Note: At this point in time the Language option is not activated and CloneSelector only works in English, though it is hoped at a later stage that we will at least also have templates in other languages.

### 5.3 Customizing the trial template to your own needs

The Trial template is what CloneSelector uses to generate fieldbooks for your trials. When you customize the template, the changes will be reflected in all the new fieldbooks you generate, but existing fieldbooks will be left unchanged. The benefits of changes may include:

1. You may **hide the traits that are not relevant for your work**, and that way your future fieldbooks will only show what is relevant for you.
2. You can **add meta data** such as Country, project leader and other **information that is the same in all or most trials** you generate, so that you only enter it once.
3. You can use the place holders Trait1, Trait2 etc to **insert new traits** that are not included as standard traits.
4. You can **modify the formula for calculating derived variables** e.g. if you want to use plot size instead of plant density to calculate yield per hectare, then you will need to change the formula in the master sheet.

#### 5.3.2. Steps to changing the Trial Template

Changing the template is a delicate procedure and it is recommended you make a backup of the template folder before you do any modifications. You should also close all open trial files or similar to avoid problems.

Note: **You should not delete Rows or Columns in the Trial Template**, and also be very careful with moving information around. If you do this CloneSelector may not work properly, or you may have made your data incompatible with other breeders data. **You should use “Hide” instead of “Delete”**, if you do not want information to be shown.

**Step 1:** Make a backup of the folder C:\CloneSelector\tpFbk , which is the folder with all the templates. If you have put the CloneSelector in another location the route is different, but should include CloneSelector\tpFbk.

**Step 2:** Select **CloneSelector -> Change Templates -> Change**

**Step 3:** Select **MSTRIAL** as **Template to modify** and

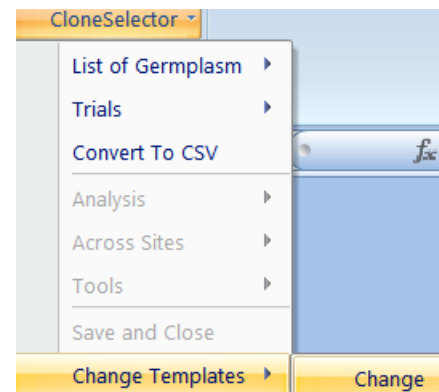


Figure 52 Change Template menu

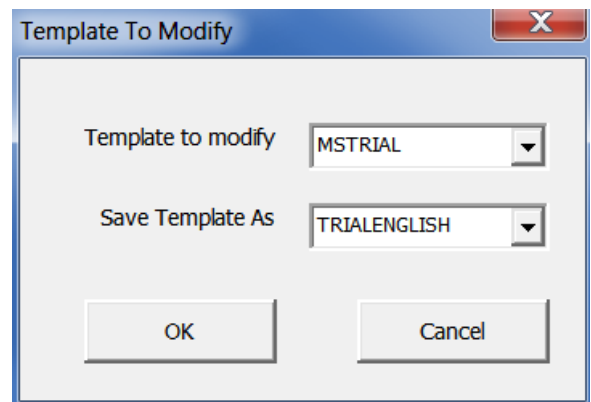


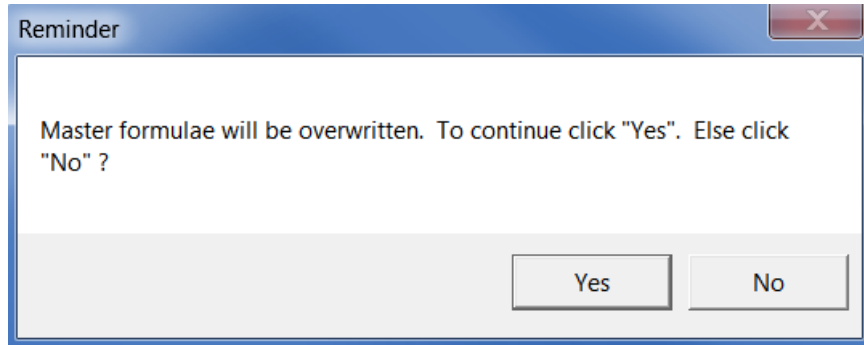
Figure 53 Select template to modify



**TRIALENGLISH as Save Template As.**

Note: We hope to be able to work with templates in different languages, but in this version you have to use the English one.

**Step 4:** You will be asked if you want to overwrite the Master formulae. If you click Yes, your formulas will be set to the default



CloneSelector calculations. If you answer No, the formulas will remain unchanged. The latter is useful if you have entered your own formulas.

**5.3.2.1 Changing the Fieldbook sheet**

Typical changes to the Fieldbook sheet include: Setting upper and lower limits to control range of valid values, modify labels of traits or inserting new traits on placeholders (Trait1, Trait2 etc).

To modify the Fieldbook sheet **click on the tab tpFieldbook.**

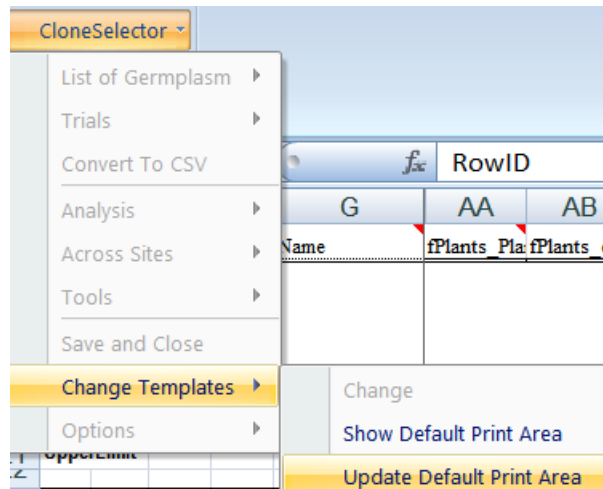
**Hide traits that are not measured or observed**

You are unlikely to routinely use all traits in the Fieldbook, and you should hide the traits you never or rarely use, so that you only see what you need.

**Step 1:** Hide all the columns you do not want to use in your Fieldbook

**Step 2:** Select **CloneSelector -> Change Templates -> Update Default Print Area**

Note: If you have already modified the template before you can click **Show Default Print Area**, and this will hide/show all the columns from the previous modifications.



**Modify range of valid observations**

The lower and upper limits of each trait in the Fieldbook are defined in rows 20 and 21, and to modify the range you simply change the existing values. If you leave a cell blank, not check will be undertaken.

20	LowerLimit					0	0	1
21	UpperLimit					40	40	9
23	Rep	Bloc	Plor	Entry	Name	Plants	Plants	Plant
24						Planted	Emerged	Uniformity

Figure 54 Setting lower and upper limits in Fieldbook

**Insert a new trait to be measured or observed**

To insert a new Trait you simply modify the text in row 23 and 24 where it says Trait1, Trait2,.....Trait14. If you want the trait included in the Result sheet either directly or in a calculated form you should also change the label on the tpResult, and insert the formula on the tpMaster sheet. For example if we want to include a trait: Insect Damage on a scale of 1-9

**Step one:** On **tpFieldbook** in row **23** and **25**, modify the Trait2 to Insect Damage and

in row **20** and **21** set the lower and upper limits to 1 and 9 respectively.

20	LowerLimit						1	
21	UpperLimit						9	
22								
23	Rep	Bloc	Plot	Entry	Name	Trait1	Insect Damage	Trait3
24								

Figure 55 Modify a placeholder on tpFieldbook e.g. Trait2

**Step two:** On **tpResults** in row **20** and **21**, modify Trait 2 to Insect Damage

**Step three:** On **tpMaster**, enter the formula for the trait under **mTrait2**. In our case it is simply =tpFieldbook!A125 , which means the value is simply copied over. If you want to

make more complicated calculations look at the other formulas on the Master sheet, and

you can probably find one that shows you how to do it. Note, if you use a value such as Plant Density or Plot size from the Installation sheet, you typically want to make it as a fixed reference (by adding \$ in front of the row and column reference e.g. \$A\$1), so that when it is copied down in the Master sheet you maintain the correct reference. Again, you may need help from an Excel expert. If you have problems with the formula get help from someone with good knowledge of Excel formulas.

=tpFieldbook!A125				
D	E	F	R	S
Entry	Row	Col	mTrait1	mTrait2
				0.00

Figure 56 Insert formula in Master sheet

Note: It is important that you insert the formula in exactly the same placeholder trait on the Master and the Result sheets, e.g. use mTrait2 and Trait2, as the analysis automatically put the Results in the same order as on the Master sheet. But the relation to the Fieldbook sheet is only via the formula on the master sheet.

**Insert a new trait to be calculated without new observations**

To insert a new calculated trait you simply insert the trait on the tpResult and tpMaster as described in step 2 and 3 above. You do not need to insert it on the tpFieldbook unless you also have something to observe or measure in each plot.

**5.3.2.2 Changing the Result and Master sheet**

The Result and Master sheets are linked by the analysis, and it is crucial that the order of the columns is exactly the same on both sheets, which means if you change something on the Result sheet, you will probably also have to change the same column in the Master sheet.

**Step One:** On **tpResults** in row **20** and **21** modify the placeholder e.g. Trait 2 is modified to Insect Damage

**Step two:** On **tpMaster** the formula is entered for the trait in the same placeholder, e.g. **mTrait2**.

See also above example on modifying Fieldbook for full example.

### 5.3.2.3 Changing the General, Installation, Field, Weather\_data, and Crop\_Management sheets

On these sheets you should enter any variable that does only change infrequently within your breeding program. Especially on the General and Installation sheets you may identify variable that are the same in all your trials such as: Institution, Leader, and Country.

If you do not need some of the variables you can hide the rows, but should not delete them. If you need to add new variables you can do so below the existing ones.

### 5.3.3. Steps to saving changes to trial template

Once you have done the desired modifications to the Trial Template you must save the changes, and to do this **you must use Update Template Settings**. If you simply click Save, this will not work.

The update process involves a serious of steps and allows you to do additional changes. In the beginning it is probably a bit confusing, but you can **simply click OK or Yes** to go to the next step, and not do any additional changes.

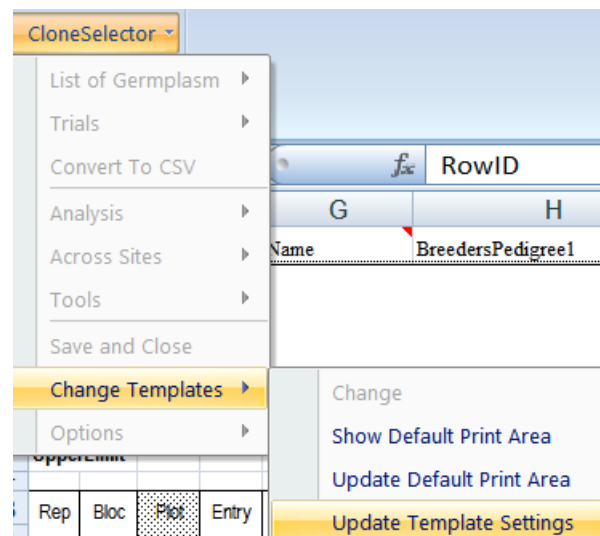
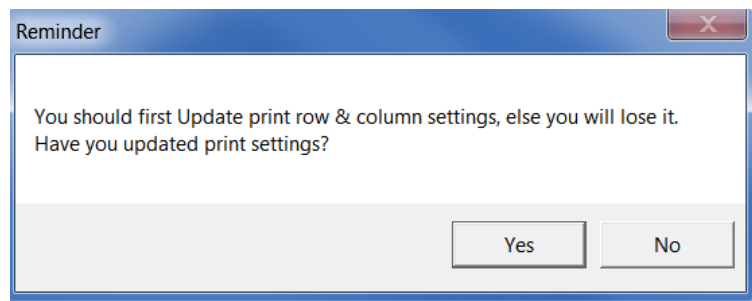


Figure 57 Use Update Template Settings to save changes

First you get a reminder that you must Click Update Default Print Area if you have changed which columns are visible. **Click Yes**, and you can move on.



Next dialog is **Trial Options**. You can here change the position of the data on the Fieldbook and Results sheets. **You should NOT make any changes on this form. Click OK to move on.**

In the **Trait Limits dialog** you can set the **Lower** and **Upper** values that are acceptable on the **Result sheet**. You can scroll down to see more traits. When you are satisfied with the ranges **click OK** and **wait**.

Your trial template is now being updated and it will close once the update is finished. **Wait until the template closes by itself.**

Test the modification of the trial template by creating a new trial.

	Lower	Upper
SelectionIndex1		
SelectionIndex2		
PlantsEmer_Per	0	100
Plant_uniformity	1	9
Plant_vigor	1	9
Flowering_degree	0	5
Senescence	0	5
Trait1		
Trait2		
Trait3		

## 6. Installation of CloneSelector

CloneSelector requires RExcel and some additional R components to be installed in order to work. An automated install packaged is available on the Sweetpotato Knowledge portal. After installing RExcel you run the CloneSelector installer.

Note: **RExcel requires Excel 2003, 2007, or 2010** to be present on the computer. **It will not work with Excel 2000 or older.** As CloneSelector requires RExcel to be present it only works with the same versions of Excel.

### 6.1 Installing RExcel

Follow the instruction in the document for RExcel installation available at on the Sweetpotato portal:

[http://sweetpotatoknowledge.org/germplasm/research-methods/cloneselector-1/R/installer\\_final\\_0.doc/view](http://sweetpotatoknowledge.org/germplasm/research-methods/cloneselector-1/R/installer_final_0.doc/view)

You can download the RExcel installer from:

<http://sweetpotatoknowledge.org/germplasm/research-methods/cloneselector-1/R-with-rcom-rscproxy-rexcel-installer.zip/view>

If you already have RExcel installed you may just add the package agricolae, which can be done with R standard Add Package procedure.

### 6.2 Installing CloneSelector

Follow the instruction in the document for CloneSelector installation available at on the Sweetpotato portal:

<http://sweetpotatoknowledge.org/germplasm/research-methods/cloneselector-1>

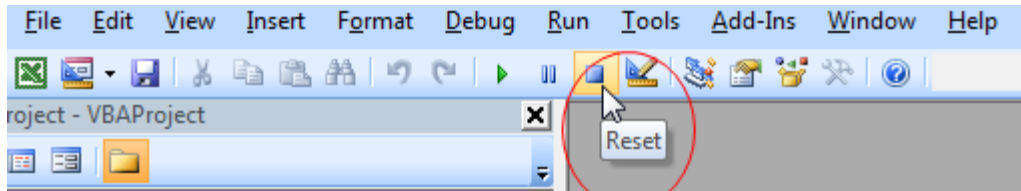
You can download the CloneSelector installer from the same place.

**If you already have CloneSelector installed should back it up before running the CloneSelector installer.** Remember CloneSelector stores data inside the CloneSelector folder, and you must therefore make sure not to lose your data.

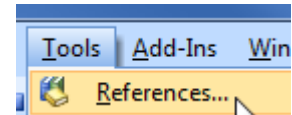
#### 6.2.1. Solving Reference Error problem in Excel 2007 or 2003

RExcel for Office 2003 is different from RExcel in Office 2007, and this may cause a Reference problem when trying to start CloneSelector. The solution is to remove the RExcel 2003 reference and put the correct office 2007 or vice versa:

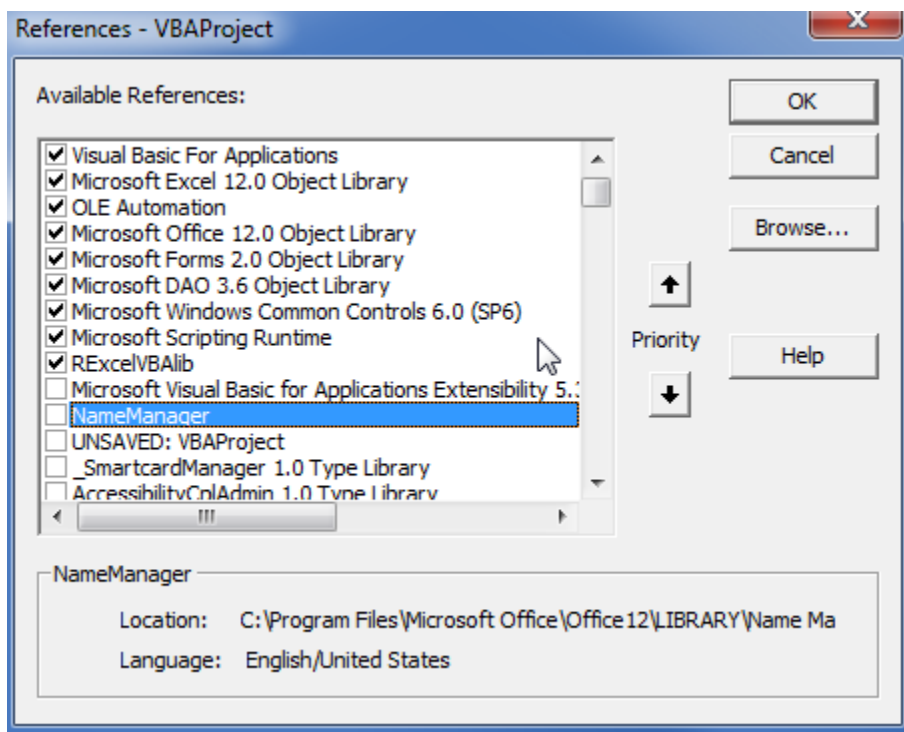
1. When you receive a **message that some library or reference** is missing, say **Yes** to start **debugger**.
2. You must then stop the debugger by **clicking on the Reset square in the toolbar**.



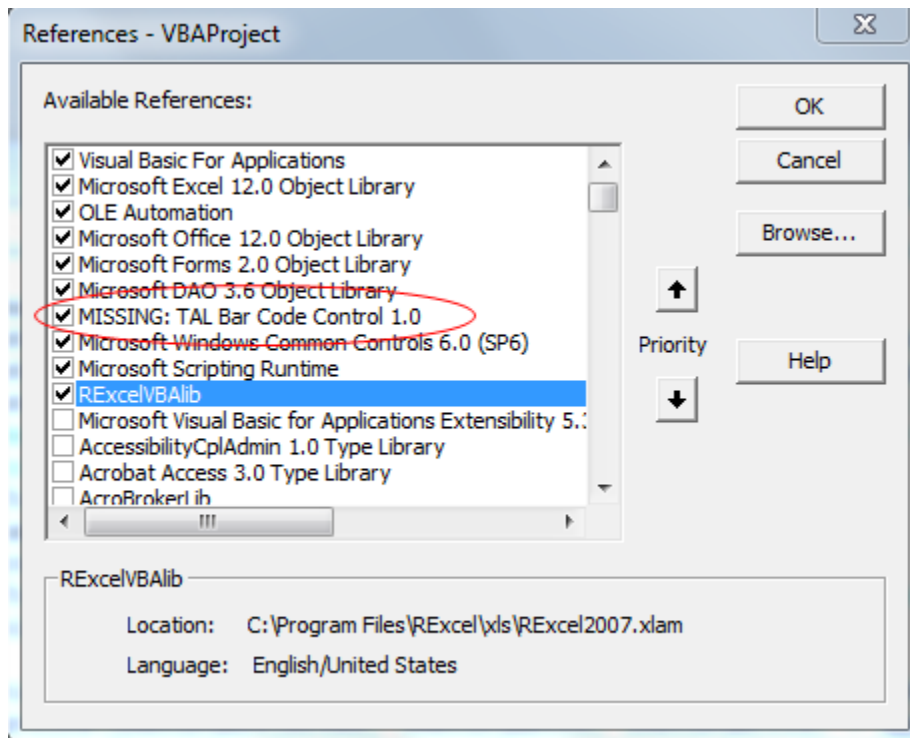
3. Then click **Tools -> References**



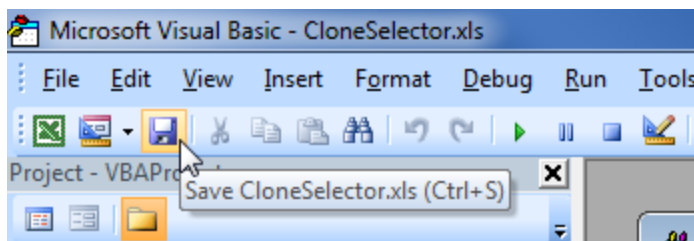
4. Your References on Office 2007 should look like:



5. **RExcelVBAlib** is called something else on Office 2003, and you may have to **'untick'** the RExcel reference that is **MISSING** and **tick the correct RExcelVBAlib**. Or as the Example below you have a missing TAL Barc Code Control, which you should 'untick'.



6. Once done click **OK** to close Reference dialog
7. **Save CloneSelector** in the VBA window



8. **Close Excel**
9. **Open Excel**
10. **Open CloneSelector**, as indicated in the section “How to start CloneSelector” and now you should not get the reference error.