

Analysis Protocol for ABI GA 310 AFLP-runs

Table of Contents

GenScan® 3.1 steps	2
1) Primer peak.....	2
2) Offscale peaks.....	2
3) Repeated runs polymer peaks.....	2
4) Folders with good runs.....	2
Remarks:	2
GelComparII 3.5 steps	3
Installation:	3
Concept:	3
Step 1: Import and normalization of gels	3
Step 2: Band assignment	3
Step 3: The database.....	3
Step 4: Band matching.....	3
How to start:.....	3
Processing data files.....	4
1) IMPORTING ABI files.....	4
2) Creating a new fingerprint type (experiment):	5
3) Alignment of the internal size standard (e.g. Gene Scan-500 ROX) and defining the reference gel (NORMALIZATION 1):	5
4) Linking the reference pattern to the STANDARD (NORMALIZATION 2).....	5
5) Linking data gels to the standard gel (NORMALIZATION 3):.....	6
6) Processing of data gels I: BAND ASSIGNMENT	6
7) Processing of data gels II: Adding the gel lanes to the database	9
8) Processing of data gels III: Defining ‘active zones’ and processing other data gels	10
9) Processing of data gels IV: BAND MATCHING	11
10) Analysis of data files	13
11) Final remarks	13

310 GenScan® 3.1 steps

- 1) Re-analyse your samples in order to remove the primer peak.
 - a) Create new sub-folders into the run-folder labelled “1-50”, “51-100”, “101-150”, ... in order to split the run files into groups of 50.
 - b) Open the project file, click on “*Sample File*” (all sample files should now be marked) | **Project | Remove Sample Files | Remove** |
 - c) **Project | Add Sample Files...** | select one of the new folders, e.g. “1-50” **Add All | Done**
 - d) Click on the button ▶ in the “*Parameters*” field and choose “*AFLP analysis range 3580-8250*” (for all samples the “*GS-500 ROX UT 3 (50)*” should be selected, if not change it in “*Size Standard field*” as described above for “*Parameters*”) | **Analyze** (if the button is not active click on the field/button to the left of the blue “*B*” button above the “*I*” button; all your samples will now be selected for analysis) |
 - e) The “*Analysis Log*” window should now pop-up (if not **Windows | Analysis Log** |). Most of the samples will match the size range 50-500 bps some even only the range 75-500 bps and by chance some will match the range **35-500 bps**. If so, remove this files from the project by clicking on the sample name | **Project | Remove Sample Files | Remove** |)
 - f) Repeat point d) - e) but select in the “*Parameters*” field the next lower analysis range, e.g. 3560-8225 ▶ 3540-8210 ▶ 3520-8205 ▶ 3500-8100 (if samples match the 35 bp size standard but **not** the 500 bp size standard, choose 3500-8150) ▶ 3480-8100 ▶... until all of the samples have matched the sizing range 35-500 bps (none of your samples will remain in the project).
 - g) If still some of the size standards are not correctly recognised, you have to change the size standard definitions by selecting one out of the list (number in brackets indicate size standard which did not match) and proceed these samples individually as described above.
 - h) As a final control add the just analyzed Samples again to the Project, and click on **Analyze** |. Control the “*Analysis Log*” for samples, which did not mach the sizing range, re-analyze them. Make back-ups of the analyzed run-files.
- 2) If there are offscale regions (listed in the analysis log) look at some samples (**Windows | Results Control** |) and write down the positions and the primer-colour (at the same positions you may find pull-up peaks within the other primers-colours, which you have to skip in your 0/1 matrix later on).
- 3) If you have repeated runs, overlay the different runs **Windows | Results Control** | (each primer-colour within one window) and select the good runs. Note peak shifts and other special things like polymer peaks and other artefacts.
- 4) Create a folder with all the good runs (only one run/sample) and rename them to PC compatibility (no special characters). Indicate the *MsEI*-primer and the plate on which the sample was run on the GA.

Remarks: When you are analysing your AFLP PCR products on the ABI 310 Genetic Analyzer, control the quality of the analysis regularly. If you encounter a strange AFLP pattern, an elevated baseline, unusual narrow peaks (polymer peaks), or an uncorrect sizing, immediately re-run this sample. In most of the cases the problem will be solved this way. Write down which runs are good or delete the bad runs immediately. If problems remain, write them down in detail. Some, like for example polymer peaks, can also be solved while transferring the patterns into a binary matrix.

GelComparII 3.5 steps

Installation: Install first Version 3.0 from the additional CD (the files will first be unzipped into the “c:\GelCompar II Setup\GelCompar II v.3.00” folder, click on the “setup” file within this folder). This turn around is only necessary to get the ABIconv filter installed {see point 1)} which is no longer provided by version 3.5. Installation of the sample database is not necessary, do **not** install the protection key drivers, **Cancel** !! Now install the Version 3.5 from the additional CD (again unzip the files into a folder you have to create first “c:\GelCompar II Setup\GelCompar II v.3.50”, click on the “setup” file within this folder). Now you have to install the protection key drivers, **OK** !. Restart your computer. You can delete the “GelCompar II Setup” folder.

Open the GelComparII program and type in the license string for Version 3.5. Now create the “GelComparII home directory” within your own directory where your personal files are located (e.g. “c:\Documents and Settings\your login name\My Documents\GelComparII home directory”) In the GelComparII main menu click on **Homedir** ! and select the home directory.

Make regularly back-ups from this home directory, all the GelComparII work you are doing will be saved within this folder.

Concept: The GelComparII Program combines “gel” (fingerprint) data of different individuals into a database. The different entries of the database can be used for different comparisons using band classes or only individual bands of the entries. Therefore processing data files (gels) in GelComparII follows 4 steps. In the first two steps the gel (experiment) will be imported, converted and normalised. The resulting data will be linked to database entries. The entries can then be combined to make comparisons, e.g. band matching.

Step 1: **Import and normalization** of gels

- a) Import and conversion of data (run) files. The conversion program converts densitometric multichannel sequencer files into “gels”.
- b) Fingerprint types are created which define the parameters of a specific fingerprint method (e.g. AFLP and its primer combination)
- c) A reference system (gel) has to be created and is used to standardize (normalize) the data gels.

Step 2: Data gel processing I: **Band assignment**

- a) Different parameters can be adjusted (filtering) which will improve the automatic band assignment.
- b) Correction of the assignment process can be done manually.

Step 3: **The database**

- a) Database entries will be created
- b) Gel lanes will be linked to this database entries.

Step 4: Data gel processing II: **Band matching**

- c) Band classes will be defined using different filtering methods.
- d) Corrections of the automatic band matching have to be done manually.
- e) Export of binary band matching character tables.

How to start: Go through all the steps described below while processing a first “gel”. This first “small gel” should be composed of the positive controls c2 (identical selective PCR-product run on each plate). This gel will be used to re-adjust the most important tools for automatic band assignment (“Richard-Lucy deconvolution” filtering parameters, {see point 6)b)iv}) and “Band search parameters”, {see point

6e)). In an ideal case they will be the same for all the different fingerprint types. The better parameters you can find (assigning the bands automatically in a consistent way) the less manual corrections have to be done later on. After that this parameters should not be changed anymore since no objective criteria will be available processing the “data gels” later on.

Use the positive controls c1 (the same template run on each plate for the whole AFLP procedure) to re-adjust the parameters (it may be necessary to increase the *Kernel size* in the *Richard-Lucy deconvolution* filter for some primer combinations {see point 6b)iv)(2) and 9b)iv)(2)(c)}

Processing data files

1) IMPORTING ABI files

- a) Create a new folder into the “GelComparII home directory” assigned to the species you are currently working with and the method you are using (e.g. *Veratrum*-AFLP). Then create a sub-folder within this folder which will later serve as your database directory containing all the automatically created files/folders of the analysis you are going to do. Name the folder as according to this analysis and give the database (see below) the same name: “species abbreviation”-“ study abbreviation”_”specification of the analysis you are doing (if necessary)” (e.g. vrab-s_test1).
- b) Create a new database: **GelCompar Startup program** | **New** (enter the same name as for the sub-folder above) | **Next** | **Brows** (select the sub-folder in the “GelComparII home directory” | **Next** | **Finish** | select “Use the local database” **OK** |
- c) Convert ABI files:
 - i) GelComparII can only process about 200-250 run-files at once (in Step 1 and 2). Therefore the run-files have to be split up into “small gels” of preferably each plate one “gel”. Thus, create subfolders of your run files, e.g. runs plate1, runs plate2, ...
 - ii) The ABICONV program:
 - (1) **GelCompar Startup program** | **Filters** | **ABICONV** (drag your ABI run-files from the Windows explorer window of your run folder onto the panel “ABI conversion program”
 - (2) Control the order of the files since the individual file names will no longer be available (should be the same as in the run folder). Number 1 will be lane 1 in the “gel”, number 2 lane 2 and so on.
 - (3) Change the order of the lanes or remove lanes using the **Lane** menu.
 - (4) Control if the 35 and 500 bp size standards are visible, otherwise reanalyse the samples with the 310 GenScan® 3.1 program {see point 1)}
 - (5) **File** | **Save file as** (give file name: use the same name as for the sub-folder above but add “_subset of samples”_”primer” at the end of the name, e.g. “vrab-s_test1_plate1_CAG”; the software appends to each primer-colour channel _01 [=JOE], _02 [=FAM], _03 [=TAMRA], and _04 [=ROX] (together with the primer specified in the name this defines the primer combination) | close window.
 - iii) The CrvConv program: This more general conversion program comes along with newer GelComparII versions. It converts different kinds of multichannel sequencer chromatograms.
 - (1) **GelCompar Startup program** | **Filters** | **CrvConv** | **File** | **Import sequences from file...** | (select the ABI run-files from your run folder) **Open** | **File** | **Export curves...** | (give file name) | **Don’t reverse** | close the window.
 - (2) ABICONV and CrvConv do not differ in how ABI files are converted (quality) but they proceed the four channels different: _01 [=JOE], _02 [=ROX], _03 [=FAM], and _04[=TAMRA]! This protocol uses the numbering of ABOCONV!

- 2) Creating a new fingerprint type (experiment):
 - a) **GelCompar Startup program | Analyze | Experiments | Create new fingerprint type... |**
 - b) enter “GS-500 ROX”
 - c) **Next |** select “*Densitometric curves*” and “*12-bit*” **Next | Next | Finish |**
- 3) Alignment of the internal size standard (e.g. Gene Scan-500 ROX) and defining the reference gel (NORMALIZATION 1):
 - a) Activate (▶) file “..._04” in the “Files” window | **File | Open fingerprint file (data)... |** select “GS-500 ROX”, the fingerprint type you just defined **OK |**
 - b) **Edit | Change brightness & contrast... |** check “*Dynamic preview*” and adjust contrast (background noise disappears) and brightness (darkest bands should be almost black) **OK |**
 - c) **Lanes | Auto search lanes... |** click twice on the button ▶ (Next step) || click on the button  (Show normalized view) |
 - d) This point has only to be done once to define the external reference positions using the STANDARD line (thereafter you will see the STANDARD line left from the gel panel in the normalization step {see points f)ii) and 4}):
 - i) Scroll through the gel and choose a representative line having the bands at an average position compared to the other lines. This line will later be the STANDARD, the lane with the reference pattern, which is essential for normalization and alignment of the different “small gels” of the same primer combination. Write down which one it was!
 - ii) Click on the 35 bp band of this line | **References | Add external reference position** (enter the fragment size “35”) **OK |** repeat for all of the 16 size standards (the fragment size of number 9 is *not defined): 1.35, 2.50, 3.75, 4.100, 5.139, 6.150, 7.160, 8.200, 9.*, 10.300, 11.340, 12.350, 13.400, 14.450, 15.490, 16.500.
 - e) **References | Use all lanes as reference lanes | Yes |**
 - f) **Normalization |**
 - i) **Auto assign (bands)** (if the STANDARD line is not yet defined) |
 - ii) **Auto assign using standard (bands)** (if the STANDARD line is already defined {see point 4})) |
 - g) click on the button  (Update normalization) |
 - h) To correct wrong assignment click on the appropriate label at the reference position, hold the CTRL key and click on the reference band you want to assign to the same reference position,  |
 - i) **File | Save | Yes |** close the window
 - j) The next steps have only to be done the first time when the reference fingerprint type is created or {point i)} if several reference systems are available:
 - i) Double-click on “**GS-500 Rox**” under “*Fingerprint types*” in the “Experiments” panel || **Settings | Set as active reference system** (only necessary if several reference systems are available) |
 - ii) **Settings | Edit reference system** (new window pops up) | **Metrics | Copy markers from reference system... | Yes | Metrics | Assign units... |** type “bp” **OK |** close the two windows.
- 4) Linking the reference pattern to the STANDARD (NORMALIZATION 2). This step creates the standard pattern which will be displayed left from the gel panel in the normalization step for all further gels which will be normalized using the given internal size standard here defined as the GS-500 ROX fingerprint type:
 - a) **Database | Add new entries... |** in the field “*Database key: (will be assigned automatically if left empty)*” write “STANDARD@01” **OK |**

- b) **Database** | **Add new information field...** | enter field name “label” **OK** |
 - c) Double-click on the new database entry || enter “GS-500 ROX” in the label field | **OK** |
 - d) Activate (▶) the file “..._04” in the “Files” window | **File** | **Open fingerprint file (entries)...** |
 - e) Drag the link arrow  of the line with the reference pattern {see point 3)d)} to the new database entry “STANDARD” and close the fingerprint file (entry) window
 - f) Double-click on “GS-500 ROX” in the “fingerprint types” window || drag the  button to the database entry “STANDARD”.
 - g) Close the “fingerprint type” window
- 5) Linking data gels to the standard gel (NORMALIZATION 3):
- a) Creating the fingerprint type (experiment) for a given primer combination:
 - i) Activate (▶) the file “..._01” in the “Files” window | **File** | **Open fingerprint file (data)...** | (since there is not yet a fingerprint type defined for this primer-combination you have to create a new one, otherwise proceed with {v}))
 - ii) **Create new...**(this is basically the same as you did for the standard “gel” in {point 2}) |
 - iii) Enter a name e.g. “vrab-AFLP_CAG_01”; the name defines the species, the method (AFLP), and the primer combination (CAG-01). The method and the primer combination together are the fingerprint type! **Next** |
 - iv) Select “*Densitometric curves*” and “12-bit” (4095 grey levels) **Next** | **Next** | **Finish** |
 - v) select the fingerprint just created (if the appropriate fingerprint type is already defined proceed with {b}) **OK** |
 - b) Activate (▶) the file “..._04” in the “Files” window | **File** | **Open fingerprint file (data)...** | minimise the window of this reference gel!
 - c) Link the data gel to its reference gel by dragging the greyed button  of the data gel and dropping it on the minimised window of the reference gel (“..._04”) **Yes** | the linkage button becomes un-greyed.
 - d) Close the data gel window | **Yes** |
 - e) Close the minimised window of the standard gel
 - f) Now only the gel-files of the active fingerprint type are visible in the files panel
 - g) Editing of the data fingerprint type, which has to be done only for a new created fingerprint type:
 - i) Double-click on the new fingerprint type, e.g. “vrab-AFLP_CAG_01”, under “*Fingerprint types*” in the “Experiments” panel ||
 - ii) **Settings** | **Edit reference system** (new window pops up) | **Metrics** | **Copy markers from reference system...** | **Yes** | **Metrics** | **Assign units...** | type “bp” **OK** | close the window.
- 6) Processing of data gels I: BAND ASSIGNMENT
- a) Activate (▶) the file “..._01” in the “Files” window | **File** | **Open fingerprint file (data)** |
 - b) Click once on the button ▶ | **curves** | **spectral analysis** | (“spectral analysis” window pops up, move it to the left-top corner | If a Error message appears you have to unselect first the filters described in points i)-iv).
 - i) **Edit** (fingerprint data window) | **Edit settings...** | select “*Densitometric curves*”:
 - ii) “*Background subtraction*” check “*Apply*” and set the “*disk size*” to the value which was estimated in the “*Spectral analysis*” as the “*Background scale*” (round the percentage value to the next integer)
 - iii) “*Filtering*” check “*Arithmetic average*”, “*Apply least square filtering*” and set the “*cut of below*” value to the value which was estimated in the “*Spectral analysis*” as

- the “*Wiener cutoff scale*” (here the exact value can be entered, leave the default value of 2.0 for “*Power*”)
- iv) These parameters have to be re-adjusted only once {see [How to start](#)}: “*Apply*” “*Richard-Lucy deconvolution*”
 - (1) set the “*iteration*” to an appropriate value (decreasing this value reduces the noise level in the gel): A value of 30 worked well for all primer combinations.
 - (2) set the “*Kernel size*” (increasing this value reduces the split up of shoulders into separate peaks): A value of 9.5 pixels worked well for most primer combinations. If there are problems with split up’s, especially in large peaks, increase this value (e.g. 15 pixels worked well for joe-CAC). But see first {[How to start](#) and point 9)b)iv)(2)(c)}
 - v) **OK** ! close the “spectral analysis” window.
 - vi) **Edit** (fingerprint data window) ! Check **Rescale curves** ! a line can be selected by clicking on it (blue lines get red) and in the panel “*Densitometric curves*” the effects of filtering (grey peaks, now used by the program to define bands) can be compared with the original curve (black line). They may appear quite different and some peaks will even no longer be recognized by the program (especially for short fragments). This effect is mainly due to the *Richard-Lucy deconvolution* which deblurs the gel. The main (positive) effect is, that more peaks will be recognized (especially larger fragments) and the consistency of band recognition is improved {see [e](#)}.
- c) Click on the button  !  !  !
- d) **Edit** ! **Edit tone curve...** !
- i) Scroll through the preview images and select a part which contains weak and dark bands (search at the end of the gel)
 - ii) click the “*Linear*” button
 - iii) Play around with the settings to get a clear, sharp and contrastfull picture (try “*Increase zero level*” 1×, “*Enhance weak bands*” 10×, and “*Reduce contrast*” 2×)
 - iv) **OK** !
- e) **Bands** ! **Auto search bands** or  !
- i) “band search” window, set “*Min. profiling*” to 1.1%, “*Grey zone*” to 1.2%, “*Min. area*” to 0.15%, “*Shoulder sens.*” to 1, and check “*Relative to max. value of lane*”
 - (1) High-density bands or offscale peaks {see report file, [f\)ii](#)} have a height of 4095 pts (the 12-bit OD range set in the beginning). A minimum of 45 pts (which is not the same as a peak height of 45 in the GenScan® 3.1 Software) would be 1.1%, the “*Min. profiling*”.
 - (2) The “*Grey zone*” defines a range where bands will be marked as uncertain. Setting 1.2% defines this zone from 45 to 94 pts (similar to the mostly used lower threshold level of 50 in GenScan® 3.1).
 - (3) In summary: Bands higher than ca 94 pts (2.3% of the highest band) will be marked as certain, bands between ca 45 and 94 (1.1-2.3% of the highest band) will be marked as uncertain, and bands smaller than 45 (1.1%) will not be marked.
 - (4) Shoulders or band doublets are split-up in two or more bands quite conservatively (shoulder sens.), see also {point [b\)iv\)\(2\)](#)}.
 - ii) optionally look at the **Preview** ! the pink mask shows the threshold level, the blue mask the grey zone
 - iii) **Search on all lanes** ! certain bands are marked with a green horizontal line, whereas uncertain bands are marked with a small green ellipse.
 - iv) Now the “*Richard-Lucy deconvolution*” filtering parameters {see point [b\)iv](#)} as well as the “*Band search parameters*” {see point [e](#)} can be re-adjusted to get a consistent band recognition. However, the program will never find the same bands

on all lanes. In the band matching (Step 3) good filters will be applied to get rid of this problem. Adjust this parameters only once {see [How to start](#)}.

- f) You can manually adjust band assignment (how it works practically {see point [vi](#)}) by comparing the automatically created pattern of certain/uncertain bands with the densitometric curve of the lines. If you add bands, only add them as uncertain. This will help to make the final decision in the band matching process later on! Thus better add a band (uncertain) when you see a peak in the densitometric curve than leave it unmarked. However you can also skip the manual corrections of band assignment and first see which improvement you can get with the very powerful filters in the band matching part {see point [9\)b\)iv](#)}. Later you can always go back to correct the band assignment. But you should only make corrections if your samples are in a random order and you do not know to which population the samples belong! Otherwise you will run into the danger to assign weak peaks as bands to samples which are more similar to each other (e.g. within populations) and doing the opposite in case of very different patterns. Therefore, try to apply a objective criterion and do the same for all samples, or let the software do it for you!
- i) **Edit** ! Check **Snap to peaks** (now the cursor automatically jumps to the closest peak) !
 - ii) When you click on a band its position (bp in brackets) is shown at the bottom of the window in the small panel to the right. There is a possibility to get a text file of the band **File** ! **Export report** !
 - iii) Do not spend time on band assignment of fragments outside the analysis ranges: all fragments larger than 500 bp and fragments smaller 50 bp. To avoid artefacts caused by primer multiplication and amplification, especially in multiplexing PCR (I found such peaks in my negative controls in the size range below 100 bp), you should increase the lower threshold level to 100 bp {see your notes of the 310 GenScan® 3.1 point [3](#)}. Later this “zones” will be excluded from the analysis {see point [8\)b\)i](#)}
 - iv) Control regions with off-scale peaks which you have written down in the {310 GenScan® 3.1 step [2](#)}. If there are off-scale peaks in a EcoR1 pattern, e.g. 01 (=JOE), you will often find pull-up peaks (bad band assignment is inconsistent peak is recognised as two bands) in the other patters, e.g. _02 (=FAM), _03 (=TAMRA) of the given MsE1-primer.
 - (1) One option is to delete such bands (pull-up peaks) or other know artefacts {see notes of the 310 GenScan® 3.1 step [3](#)}, but write down what you deleted, since you have to do the same thing in all the “small gels” of the same MsE1-primer.
 - (2) Another option is not to delete such bands (which is easier in large gels) but to skip this band classes {see point [9\)b\)v](#)}(1)}later on.
 - v) Correct shoulders which where wrongly marked as double peaks or the other way around (but also here, the correction can be made later on by deleting or matching whole band classes)!
 - vi) Manual band assignment:
 - (1) Select a group of bands with holding the SHIFT key and dragging the mouse pointer holding the left mouse button.
 - (2) Delete bands with the DEL key
 - (3) Use the  buttons to zoom in of zoom out
 - (4) press **F6** to mark uncertain bands as certain (or **F5** to mark bands as uncertain)
 - (5) add new bands with selecting a band with the curser (jumping of the cursor to the closest peak can be avoided by holding down the TAB key while clicking on a band) and pressing the ENTER key (and F5 to mark them as uncertain).
 - vii) **File** ! **Save** !

- g) Close window (the following question may appear: “*Settings have been changed. Do you want to use the current settings as new defaults?*”
- i) Answer **Yes** only the first time. The settings will be saved along with the fingerprint type, thus the new defaults setting for this specific fingerprint type.
 - ii) Answer **No** for all the small data gels you are analysing later on, the settings will only be saved along with the gel, while leaving the fingerprint type unchanged.
- 7) Processing of data gels II: Adding the gel lanes to the database
- a) Activate (▶) the file “..._01” in the “Files” window | **File | Open fingerprint file (entries)... |**
 - i) If no database entries are defined for the current gel lanes the program can create new entries and directly link the gel lanes to the entries. This option can be used for the gels of the first fingerprint type **Database | Add all lanes to database |** but {see point iii)}
 - ii) Link the lanes of the other fingerprint types (other primer combinations) individually to the already existing entry of the same sample, thus drag the link arrow  of each line to the belonging database entry. If you try to link a lane to an entry which already has a lane of the same experiment type linked to it, the program will ask whether you want to create a duplicate key for this entry. The program assigns an extension /#x to such duplicates and automatically updates fields from the “master” entry.
 - iii) An other option to create database entries is to import the field data from an Excel worksheet. This needs some configurations and adding of columns in the Excel worksheet:
 - (1) 1. Column, the GelComparII database “Key”. This is a unique key assigned to each entry of the database. This key is equivalent to a unique label, which identifies a sample. This can be something like: ‘*species*’@’*region code*’@’*population numbering inside the region*’@’*subpopulation numbering inside the population*’@’*plant numbering inside the subpopulation*’@’*additional number for duplicated runs*’, e.g. vrab@JUR@002@01@010@1.
 2. Column, the “label” of the samples (e.g. the number you gave the samples during the AFLP procedure)
 3. Column, the “small gel code” {see point 1)c)i}
 4. Column, numbering of the samples within the “small gel”. The numbering must be the same as the samples sequence in the “run folder” of the “small gel” since only then the numbering will correspond to the lane numbering in the “small gel”.
 - (2) Now you have to define a name for the data you want to make accessible for another program (e.g. through the Open Database Connectivity (ODBC) protocol).
 - (a) Define the range you want to make accessible. You have 2 options:
 - (I) Mark only cells you want to make accessible
 - (II) Mark whole columns
 - (III) Later on you can always change the selection
 - (b) **Insert | Name ▶ Define...** | give a name {e.g. see 1)a)}
 - (c) The name appears now in the name field (left top corner). If you have defined different names within your worksheet, you can easily jump between them by selecting the appropriate name:
 - (d) Save the workbook.
 - (3) Now you have to set up the ODBC link in the GelComparII program
 - (a) **Database | ODBC link | Configure external database link... |**

- (b) “*ODBC data source:*” click **Select...** ! select “*Excel-files ...*” ! select the Excel Workbook you saved before.
 - (c) “*Database table or query:*” here you have to fill in the name you specified before within the Excel worksheet {see point (2)(b)}
 - (d) **OK** ! now the ODBC database import window pops up. Link the fields from the local GelComparII database (left column) with the fields from the external database (right column) **Link** !
 - (e) If all links are established press **OK** ! A link to the GelComparII key field is obligatory, since here the data base entries are compared. Never use a key twice (GelComparII takes only one entry per key), add an additional number at the end of the key {see point (1)}
 - (f) Update all GelComparII database entries from the external data source **Database ! ODBC link ! Copy from external database** (existing entries are updated and new entries are automatically added) !
 - (g) Now you can add your data gel lanes to the entries as described above {point ii)}
- b) At the end, each entry will have as many linked patterns (green dots) as primer combinations are available (e.g. MsE1^{primer 1}-EcoR1^{fam}_(_02), MsE1^{primer 1}-EcoR1^{joe}_(_01), MsE1^{primer 1}-EcoR1^{tamra}_(_03), MsE1^{primer 2}-EcoR1^{fam}_(_02), MsE1^{primer 2}-EcoR1^{joe}_(_01), MsE1^{primer 2}-EcoR1^{tamra}_(_03))
- 8) Processing of data gels III: Defining ‘active zones’ and processing the other data gels:
- a) The fingerprint types (experiments) for the other primer combinations can be created in the standard way {see point 5)} or, in order to reduce to work of parameter setting, as follows:
 - b) First you have to finish the editing of the data fingerprint type available until now:
 - i) Defining the ‘active zone’ which will be used for the analysis (band matching):
 - (1) Click on an entry in the database window
 - (2) Activate (▶) the data experiment “vrab-AFLP_CAG_01” in the “*Fingerprint types*” window ! **Experiments ! Edit experiment type...** !
 - (3) At the bottom of the window you see the gel line of the selected database entry
 - (4) To exclude the lower threshold region hold the left mouse button and the SHIFT key at the same time while dragging the mouse pointer over the fingerprint.
 - (5) First drag it only a little bit from the left to the right into the gel strip. In the top panel the left number in the squared bracket shows to which percentage the range is excluded (thus the lower limit of the active zone)
 - (6) Try to get this value equal to the value in the “*Reference systems*” for 100 bp - 0.2% to take the variation in migration of fragments into account (or set lower values to a minimum of 50 bp {see point 6)f)iii}) by repeating point {(4)}. You have to start with the mouse pointer within the crosshatched red region (the excluded region).
 - (7) To re-include a region, hold the left mouse button (without holding the SHIFT key), while dragging the mouse pointer over the fingerprint.
 - (8) Repeat for the higher (500 bp) threshold level. The right number in the squared bracket shows the higher limit of the active zone. Give 0.2% more to include a 1 bp variation at the 500 bp position.
 - (9) With this method multiple you can exclude multiple regions (regions of known artefacts or not analysable patterns).
 - ii) Defining the position tolerance for band matching **Settings ! Position tolerance settings...** ! {explanations see point 9)b)iii}) !
 - (1) “*Optimization*” (shift between two patterns), set 0.2.
 - (2) “*Position tolerance*” (maximal shift between two bands), set 0.2

- (3) “*Change towards the end of fingerprint*” (of position tolerance), set 0.2.
 - (4) “Uncertain bands” check “include”
 - (5) The blue arrows below the gel line visualise the position tolerance settings.
 - (6) **OK** !
 - iii) Close fingerprint type window
 - c) Create copies of fingerprints (to do that you have to close the *GelComparII* program):
 - i) In Windows Explorer go to the “*GelComparII* home directory” and select the sub-folder you created in the very beginning {see point 1)a)}. Open the folder “Fprint”.
 - ii) Create as many copies of the data fingerprint type (e.g. ) as there are primer-combinations {see point 7)b)} and give them the appropriate name.
 - d) Open the *GelComparII* program again, **Analyze** ! and repeat {point 6)} (band assignment) and {point 7)} (adding the gel lanes to the database) for the other data gels.
- 9) Processing of data gels IV: BAND MATCHING
- a) Composite data set
 - i) **Experiments** ! **Create new composite data set** !
 - ii) Enter the same name as for the sub-folder in the very beginning (“species abbreviation”-“ study abbreviation”_”specification of the analysis you are doing”, if necessary) **OK** !
 - iii) Select the fingerprint type(s) **Experiment** ! **Use in composite data set** !
 - iv) Close the “Composite data set window” window.
 - b) Creating a band matching
 - i) Select the database entries in the main window on which you want to create the band matching:
 - (1) Manually select entries with CTRL + mouse click (repeating this will unselect the entry)
 - (2) Automatic search and select entries **Edit** ! **Search entries** ! (* replaces any range of characters whereas ? replaces only a single character)
 - (3) A selection of entries can be saved as a **Subset** ! **Create new** (with the copy, cut, and past function you can add or remove entries in subsets)
 - ii) **Comparison** ! **Create new comparison** ! select the appropriate fingerprint type in the experiments selection bar at the bottom of the window and press the show image button  (or **Layout** ! **Show image**),
 - (1) Press the button  (show metrics scale)
 - (2) **Layout** ! **Stretch (X dir)** ! (enlarges the image in the horizontal direction only)
 - iii) **Bandmatching** ! **Perform band matching** !
 - (1) the “Position tolerance settings” dialog box for the fingerprint type pops up for which the appropriate values are already set {see point 8)b)ii)}. 465 bps (35 bp - 500 bp) are about 94% of a gel line, 0.20% equals 1bp. 0.2% allows a shift of 1 bp between two bands (tolerance) or two patterns (optimisation).
 - (2) The “*Change to the end of fingerprint*” value is additive to the “*Position tolerance*” value, thus 0.2% allows a continuous increase of the tolerance to 2 bp (0.2% + 0.2%) in tolerance towards the end of the gel line.
 - (3) **OK** ! **Yes** (since you want to perform the band matching only inside the selected entries)
 - iv) **Bandmatching** ! **Band class filters...**
 - (1) “*Remove all bands below*”: Leave both values at zero since with this filter differences between the samples are created. This would need a hypothesis about the quantity of fragments at a given loci which is not possible without additional experiments with known crosses! PCR reactions with a lot of fragments are very complex, especially multiplexing PCR’s. The presence of some fragments

(which depends on the individual (sample) under investigation) may influence the amplification of other fragments, i.e. there may be a lot of products of a given fragment in one sample but much less in another, which does not say anything about the presence or absence of a fragment. AFLP markers are qualitative markers!

- (2) “*Remove all band classes that have no bands exceeding*”: Here a quite conservative approach should be applied. We do not know which part of the genome yields the fragments and which fragments are used for the analysis (depends on the primer combination, the range we decide to use for the band matching, ...). Therefore it is better to skip a whole band class which consists just of some weak bands, polymer peaks or artefacts. Keeping these band classes would bear the risk to include polymorphic loci into the analysis which are just caused by weak amplifications or artefacts.
 - (a) Set 2.3% for “*minimum profiling*”. The value is similar to the values you set in the band matching step {see point 6e)i)} meaning that a band class needs at least one band having a height of 82 pts to be considered as relevant. Do not increase this value since large fragments give rather flat but wide peaks and you would lose those. To deal with this problem the area filter can be applied.
 - (b) Set 0.5% for “*minimum area*”. This filter treats shorter fragments (high and narrow, dark patterns) not different from larger fragments (flat and wide, weaker patterns). It removes quite some band classes from the band matching but leaves the loci with at least one clear and relevant band in the band matching table.
 - (c) If still a lot of polymorphisms are detected using the positive controls c2 or c1 {see [How to start](#)} increase the value for “minimum area” or readjust the *Kernel size* in the *Richard-Lucy deconvolution* filter for some primer combinations {see point 6b)iv)(2)}.
- v) Manual editing of a band matching (you can easily see which bands belong to a given band class by double clicking the vertical blue line that represents the class: all bands that belong to the class are selected with a green flag):
 - (1) Press the button  (show curves)
 - (2) Delete a whole band class as follows: Click on a band belonging to the band class, right-click on the band, and select **Band classes Remove band class**!. Delete all band classes larger than 500.25 bp and of known artefacts. Especially bands larger than 400 often give wrong shoulders which have to be merged to one class {see also point (5)} or simply one band class has to be deleted..
 - (3) Assign a band to a band class: Select a band class to which the band ought to be assigned (the selector becomes blue), right-click on the band, and select **Band classes Assign band to class** !.
 - (4) Create a new band class: right-click on the band which should belong to a new class and select **Band classes Add new band class**!, right-click again on the band and select **Band classes Auto assign bands class** !
 - (5) Merge band classes: Choose a band which occurs quite in the middle of the two classes, add a new class, remove the other two classes and right click again on the band on the new class **Band classes Auto assign bands class** !
 - (6) Remove a single band from a class (which should only be done for known artefacts, since otherwise polymorphisms are created): Select a band and press DEL
 - (7) After reassigning bands, the band class position may not be the center anymore: right click on the band class **Band classes Center class position** !
- vi) Save the band matching **File** ! **Save**

- vii) Adding entries to a band matching while preserving the existing band classes and introducing new classes where the new entries have bands that do not fit in any of the existing classes:
- (1) Select the new entries in the database and select **Edit ! Copy selection**
 - (2) **Edit ! Past selection** into the comparison on which the band matching was created.
 - (3) **Bandmatching ! Search band classes !** choose the same tolerance settings as before
 - (4) To the question “*Remove existing band classes?*” answer **No** (in order to preserve the existing band matching) !
 - (5) To the question “*Search only inside the selection?*” answer **Yes** (in order to search only for new band classes) !
 - (6) Apply the band class filters again {see point iv)}
- c) Creating a band matching table for polymorphism analysis:
- i) All band classes
 - (1) Click on the composite data set button  in the experiment type selection bar in the bottom of the comparison window !
 - (2) Once the character table is present, it is still possible to edit the band class assignments on the patterns. The character table is updated automatically. Band classes that have been created by the user are marked with an asterisk.
 - (3) **Composite ! Export character table !** “Use tab-delineated fields” **Yes !**
 - (4) A tab-delineated text file appears (RESULT.TXT) which shows on the first lines the band classes (columns) and thereafter the presence/absence table starting with the sample names (lines).
 - (5) To correct a mismatch in the column label add a TAB in the very beginning of the document
 - (6) **File ! Save as... !**
 - ii) Only polymorphic band classes:
 - (1) This tool is only working for selected entries, thus if you want to use the whole dataset, first select all entries {see point b)i}
 - (2) **Bandmatching ! Polymorphic classes only (for selection list)**
 - (3) continue with steps {i)(3)-(6)}
- 10) Analysis of data files (samples):
- a) Make a back-up from your “GelComparII home directory”
 - b) Create a new sub-folder and database as described in {points 1)a) & 1)b)}
 - c) Close the GelComparII program
 - d) Copy all files from the previous sub-folder (the test analyses you just did on the control) into the new sub-folder!
 - e) Open the GelComparII program again **Analyze !**
 - i) Remove all entries except the STANDARD
 - ii) Remove all files (test gels) except the one on which the STANDARD is defined!
Right-click on the file ! **Delete experiment file... !**
 - iii) Close the GelComparII program again
 - f) Now your new database is ready (only containing the STANDARD on which all your small gels will be normalised {see point 4})) to start the analysis of all your data sample files.
- 11) Final remarks: I used the GelCompar software with the aim to transfer the AFLP raw data into a binary matrix as automatic as possible. The most difficult part is to find good filtering parameters, in order for densitometric curves of all the control samples to be fitted in such a

way, that the same bands will be maintained in the matrix. The parameters mentioned in this protocol are the ones I used. Thereafter, I let the program do the work. The only manual editing I did where the following:

- a) After bandmatching, I checked the resulting banding pattern visually for tight band classes (< 0.5 bp up to fragments of 200 bp then < 0.8 bp).
- b) They were corrected into one class only if less than 1.5% of the samples had both fragments (shoulder) or if wide peaks were inconsistently split up into two peaks.
- c) I did those corrections in the band assignment. You have just to open the small gels on which you have to correct something as described in the Band Assignment {see point 6)}. Any time you want, here you can change manually the band assignment and save the file, but not as default settings {see point 6)g)} and never do a new band search otherwise you will lose all the corrections. The advantage to do the corrections in the band assignment is, that you can do later on new bandmatchings with a subset of your samples and the corrected band assignments will always be used while corrections in the bandmatching have to be done again.