Users Guide to the Soybean Physical Map at the Website http://hbz.tamu.edu/.

Jeffry Shultz 1,
David A. Lightfoot 1

1 Department of Plant Soil and General Agriculture, Southern Illinois University, Carbondale, IL 62901-4415. E-mail: ga4082@siu.edu http://www.coalab.siu.edu/genome/

To develop the physical map of soybean, we have developed three large-insert soybean DNA libraries from 'Forrest' (Meksem et al., 2000). Texas A&M has generated restriction enzyme fragment fingerprints for 90,000 of 112,600 clones. About 50,000 clone fingerprints have been posted at at Physical Mapping sections of http://hbz.tamu.edu/. Each clone was tested for the content of known genetic markers and some ESTs (Wu et al., 2001) and SNPs (Zobrist et al., 2000; Meksem et al., 2001; Shultz et al., 2001). We have shown that fingerprints of BACs from other cultivars can be integrated with this data set (Shultz et al., 2001). The integrated genetic and physical map will inexpensively and rapidly enable identification of a number of genes underlying QTL of economic importance. Here we report methods of use for the genetic and physical map. This guide should allow soybean researchers to use the web-site at which data can be retrieved, clones submitted for fingerprinting and clones requested.

Methods

FPC Hitting Tool

There are three issues concerned with obtaining useful data from this source:
1. Improper interface to the actual FPC database
2. Only FPC hitting tool available (Individual clone matches)
3. Fundamental issues with FPC program and number of bands generated while fingerprinting

Improper interface:

This problem became visible when the data was first posted on the internet. At that time, clone "IDs" were used to report matching clones when using the FPC hitting tool. This meant that researchers would enter a plate location into the search engine, which would then give a list of positives using their clone IDs. This system seemed to work intermittently, giving expected matches at times, but most of the time it failed to produce the expected matching clones. This problem was identified, and decoded by SIU. Both the interface failures and the ambiguity of the clone ID reporting were discussed with Texas A&M personnel. The clone ID reporting was eventually replaced by plate location reporting. Unfortunately, the interface problem is still being rectified.

The previous decoding of the interface has allowed SIU to continue to use the database with good success. We have devised a method for interrogating the database tht seems to provide accurate and reliable results.

The first step is to convert the Web address to the actual plate address. This is done manually with table 1. All of the plate address stays the same, except that the numbers are converted. For example, Plate address H63A4 would be entered as H63A14 at the web browser. To reverse the procedure, if the Web browser lists H63A14, the user must convert this to H63A4 when attempting to isolate and perform experiments with this clone

Table 1. Web address conversion for Soybean FPC Database

Web Address (Based on incorrect conversion by database interface)

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
1	3	5	7	9	11	13	15	17	19	21	23	2	4	6	8	10	12	14	16	18	20	22	24

Plate Address

How to Contig two nearby (<1 cM apart) clones:

Enter each clones' converted ID into the FPC hitting Tool.

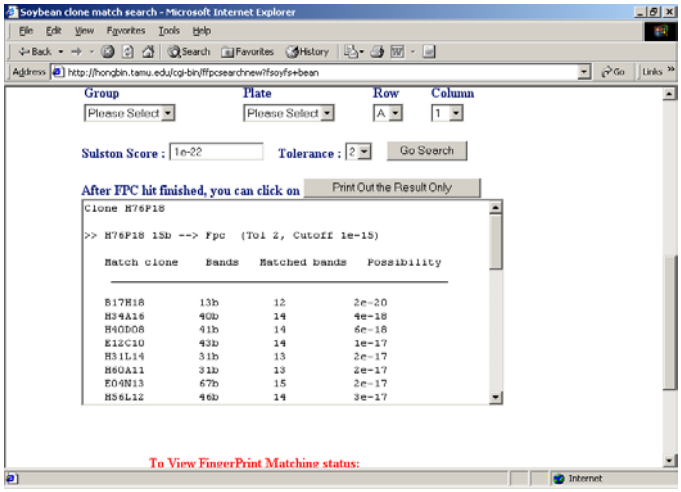

Figure 1. Texas A&M FPC hitting tool for Soybean.

Find any clones that match between the two clones of interest:

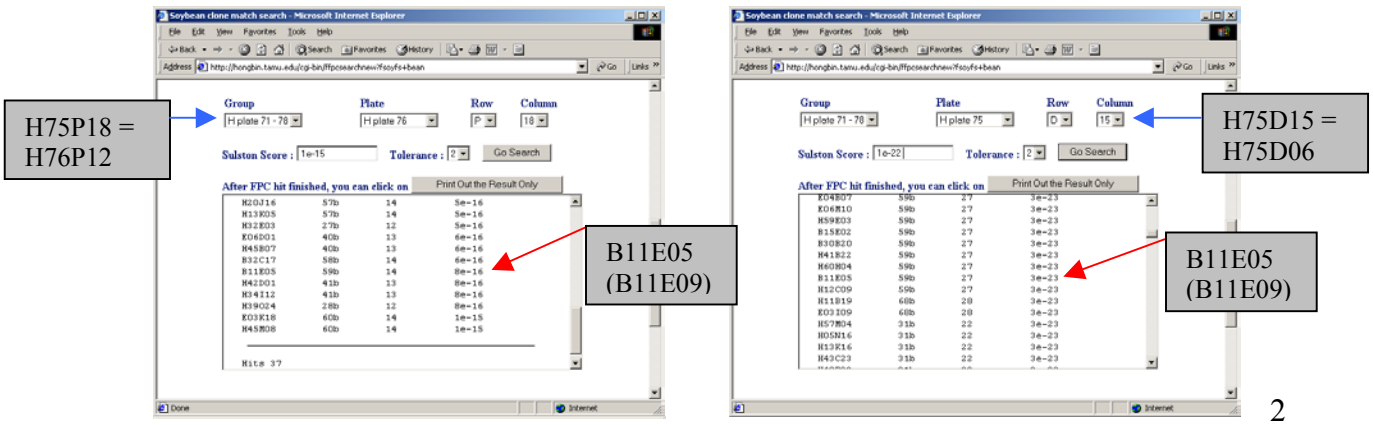

Figure 2. Clone plate entry and conversion in FPC, with linking clone indicated

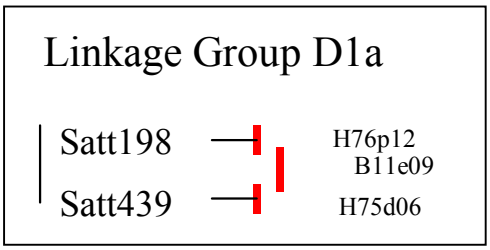

Figure 3, Contig formed on Soybean Linkage Group D1A from the Essex x Forrest cross.

How to find clones that are reasonably linked to your clone of interest using the FPC hitting tool, Microsoft Excel and Microsoft Wordpad.

First, start Microsoft Excel spreadsheet “Clonemacrosheet”, then open and point your web browser to the FPC web viewer. "Clonemacrosheet.xls" available at www.___.___

Second, convert the clone location to match the web location. Using table 1, H11B09 converts to H11B05. Start the FPC hit utility and enter the clones' plate location (Fig. 4).

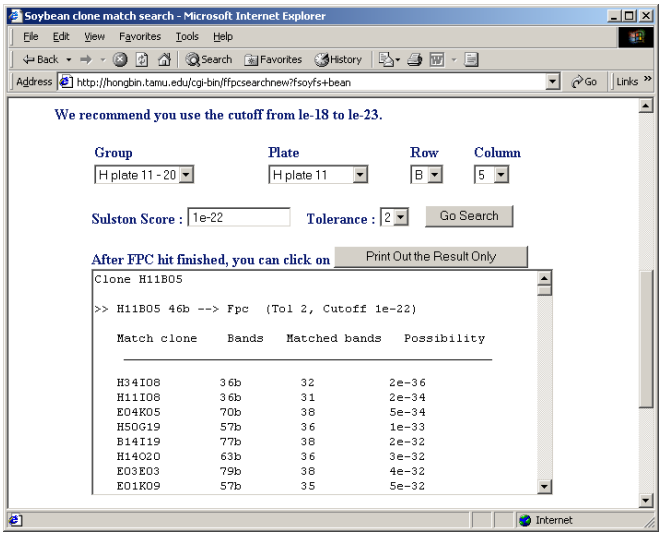

Figure 4. Clone plate entry and conversion in FPC

Now, highlight the output by right-clicking your mouse and highlighting “Select All”

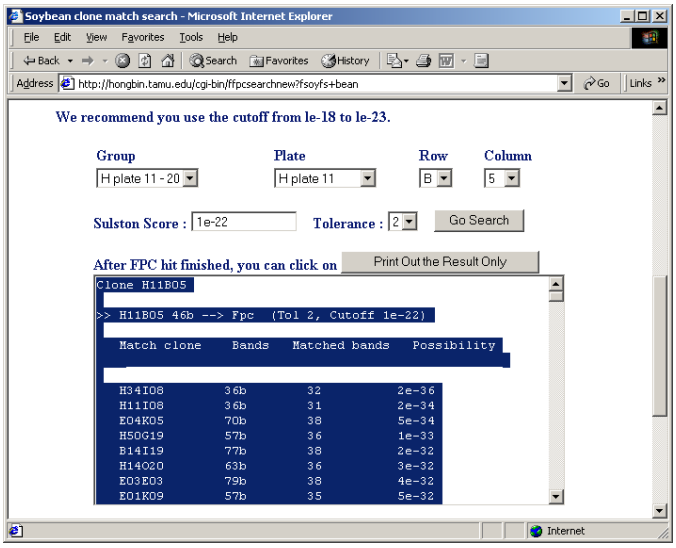

Figure 5. Select All, highlight and copy from FPC window

Right-click again and select “Copy”

Open the Microsoft Wordpad document “Clonepastewordpad” and “Select All” from the edit menu. Now select Paste. These steps will overwrite the previous output in this file and replace it with the current listings. Once the FPC output is pasted, Wordpad goes to the end of the file, where you will be able to see the number of positive clones (in this case there are 1369 for H11B09 at Tolerance 2, Cutoff 1e-22.

It is important that you use the same file for this step, since the excel macro will only open “Clonepastewordpad”.

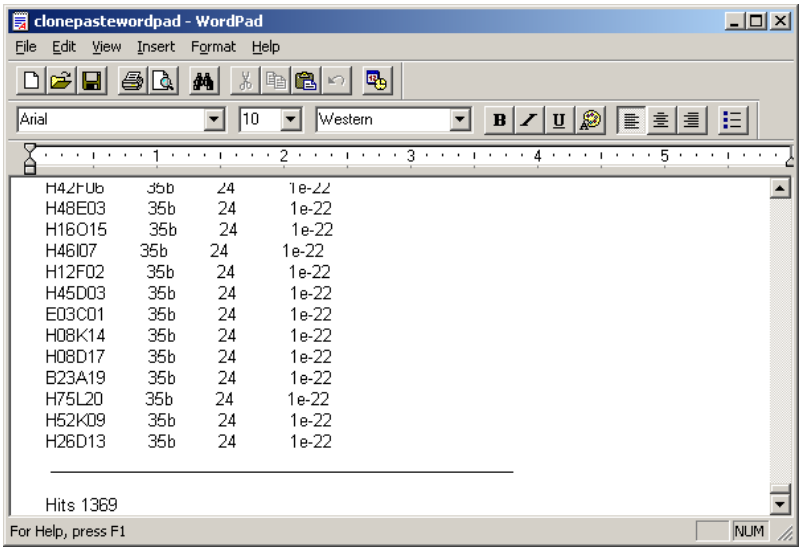

Figure 6. Wordpad file "Clonepastewordpad" after pasting FPC output into the file

Click on the “X” to close the file and select “Yes” when asked to save changes. (Wordpad must be closed for the excel macro to work).

Maximize “Clonemacrosheet.XLS” and press ctrl+p. the program will automatically import the data from "Clonepastewordpad" into columns, calculate the percentage of positive bands relative to total bands in each clone, then filter out all clones that do not match your criteria (we typically use 70 to 75% band matching) This macro also recognizes and flags clones that have too many bands.

An essential part of this procedure is the ease with which the user can change the percent of matching bands. This is vital as the band count for the clone of interest becomes larger. For example, Table 2 shows the effect of percent matching bands on number of matching clones at Tolerance 2, Cutoff 1e-22 for H11B09 (46 bands).

Table 2. The effect of changing the percent of matching bands on number of matching clones at Tolerance 2, Cutoff 1e-22 for H11B09 (46 bands).

Percent of Clones' Bands that match H11B09 (%)	Number of Matching Clones (1369 total)
50	1237
60	669
70	149
80	20
90	1

Figure 7. Microsoft Excel spreadsheet output listing all clones with at least 80% of their bands matching the clone of interest (H11B09).

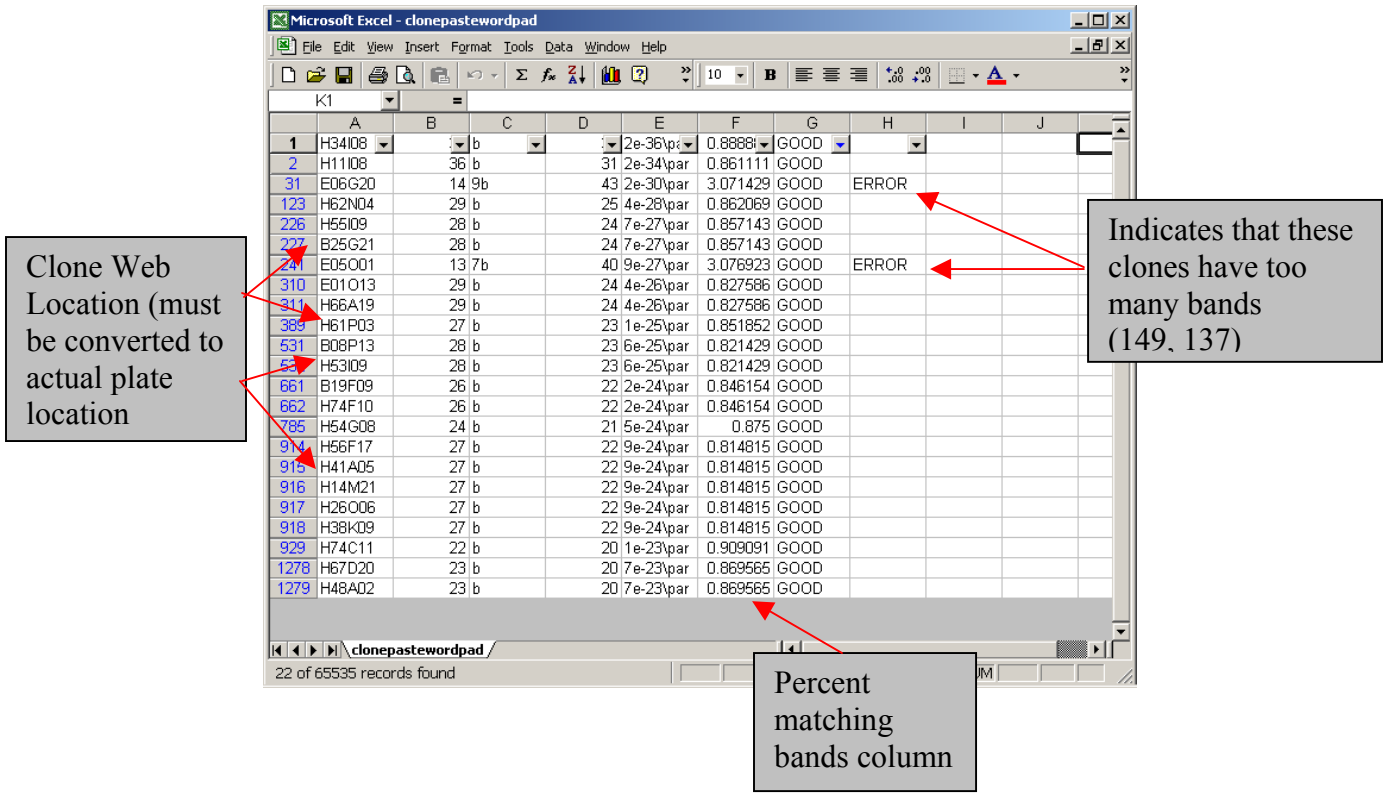

Table 3. The effect of increasing Cutoff on positive clone number using Clone H11B09 (46 bands) with a Tolerance of 2

Cutoff	Number of Matching Clones
1e-22	1369
1e-26	241
1e-30	24
1e-34	1 (H34I08)
1e-38	0

High cutoff can be used for high band number clones (>40), but is inappropriate for low band number clones, since this favors false linkages with high band number clones. Using a high cutoff also eliminates large numbers of clones that have low band numbers (<30) from matching to middle (30-40) to high (>40) band number clones. For instance, 22 of the 24 positive clones listed in table three for H11B09 at cutoff 1e-30 are between 57 and 89 bands with a range of 35 to 39 matches; only 2 of the 24 are in the middle (30s) band range and are highly probable as matches with 32 and 31 matching bands. Both of these clones are at the top of the list for the 80% matching band Excel spreadsheet output, which also includes 18 other clones that are in the 20 band range (Figure 7).

Using this information, efforts are underway to create groups of highly probable clones using the clones that are already integrated into the soybean map. These groups can then be used to create chromosome specific sub-libraries and to create limited contigs. We believe that these steps alleviate some of the fundamental issues with FPC program and high number of bands generated while fingerprinting.

References

Meksem K. Zhang H.B. and D.A. Lightfoot. 2000. Two transformation ready large insert clone libraries for soybean: Physical mapping of resistance to Soybean Cyst Nematode and Sudden Death Syndrome. Theor. Appl. Genet. 100:747-755.

Meksem K., Ruben E., Hyten D., Triwitayakorn K., Lightfoot DA. 2001. Conversion of AFLP bands to high-throughput DNA markers. Mol Gen Genomics 265:207-214.

Meksem K., Ruben E, Hyten D Schmidt M and Lightfoot DA. 2001a. High-throughput detection of polymorphism physically linked soybean cyst nematode resistance gene *Rhg4* using Taqman Probes. Molecular Breeding 7: 19-28.

Wu C, FA Santos, P Nimmakayala, R Springman K Meksem, DA Lightfoot† and H-B Zhang 2001. The NBS-LRR gene family in the soybean genome is organized and evolves by a process of birth-and-death, supercluster formation and adaptive selection**.** Proc. Natl Acad Sci USA (submitted).

Zobrist K, Meksem K, Wu C, Tao Q, Zhang H, and Lightfoot DA (2000) Integrated physical mapping of the soybean genome: A tool for rapid identification of economically important genes. Soybean Genetics Newsletter 27: 10-15.

Zobrist K, 2000a. Integrative physical mapping of the soybean genome. MS thesis SIUC Carbondale pp213.