High-throughput BAC Fingerprinting For Constructing Physical Maps

Ming-Cheng Luo

Department of Plant Sciences
University of California, Davis

Why still think about fingerprinting when sequencing is easily accessible?

- **Vast majority of species will not be sequenced in the near future.**

  For those species, contigs of large-insert genomic clones anchored to genetic maps represent a low-cost alternative to genome sequencing that would greatly enhance the accessibility of their genomes for biological research.

- For large genomes, a physical map is prerequisite for genome-wide sequencing.

Development of fingerprinting methods

By population of restriction fragments:

- One 6-cutter (Olson et al. 1986)
- One 6-cutter + one 4-cutter (Coulson et al. 1986; Klein et al. 2000)
- Multiplexing one 6-cutter + one 4-cutter (Ding et al. 1999)
- One type IIS restriction followed by determine of the nucleotide sequence at the cleavage site (Brenner and Livak 1989; Ding et al. 2000)
- Four 6-cutter + one 4-cutter (Luo et al. 2003)

Development of fingerprinting methods

By detection method:

- Staining (Olson et al. 1986)
- Radioisotope labeling (Coulson et al. 1986)
- Fluorescence dye labeling (Ding et al. 1999)

Development of fingerprinting methods

By electrophoresis type

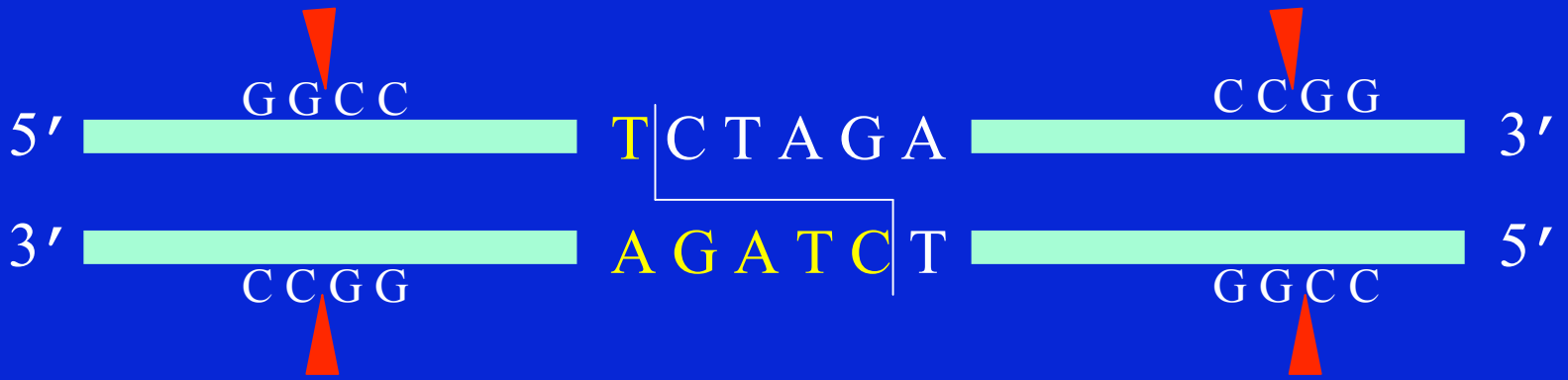
- Agarose gel (Olson et al. 1986)
- Polyacrylamide gels (Coulson et al. 1986)
- Polyacrylamide gel based sequencer (Gregory et al. 1997)
- Capillary sequencer (Luo et al. 2003)

Development of fingerprinting methods

- Multiple digestion
- Multi-color fluorescence labeling
- Capillary electrophoresis

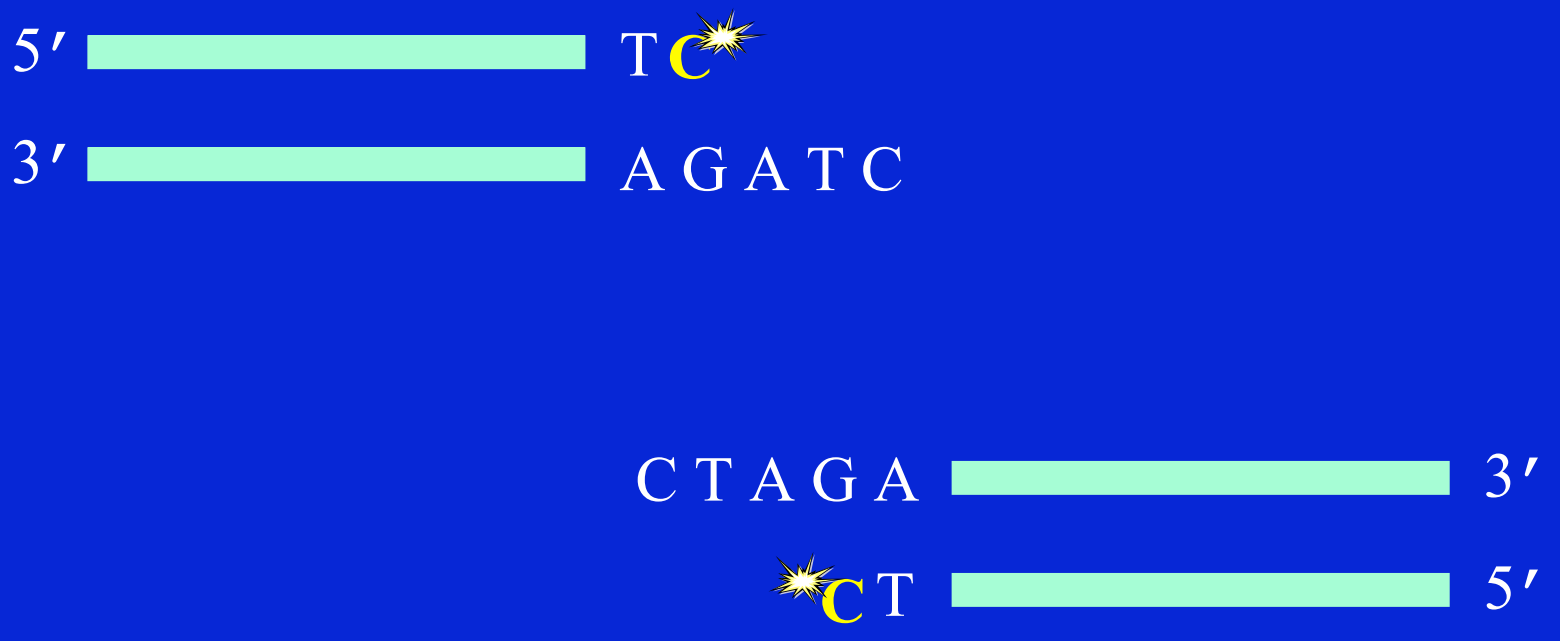
SNaPshot BAC Fingerprinting

Restriction cleavage



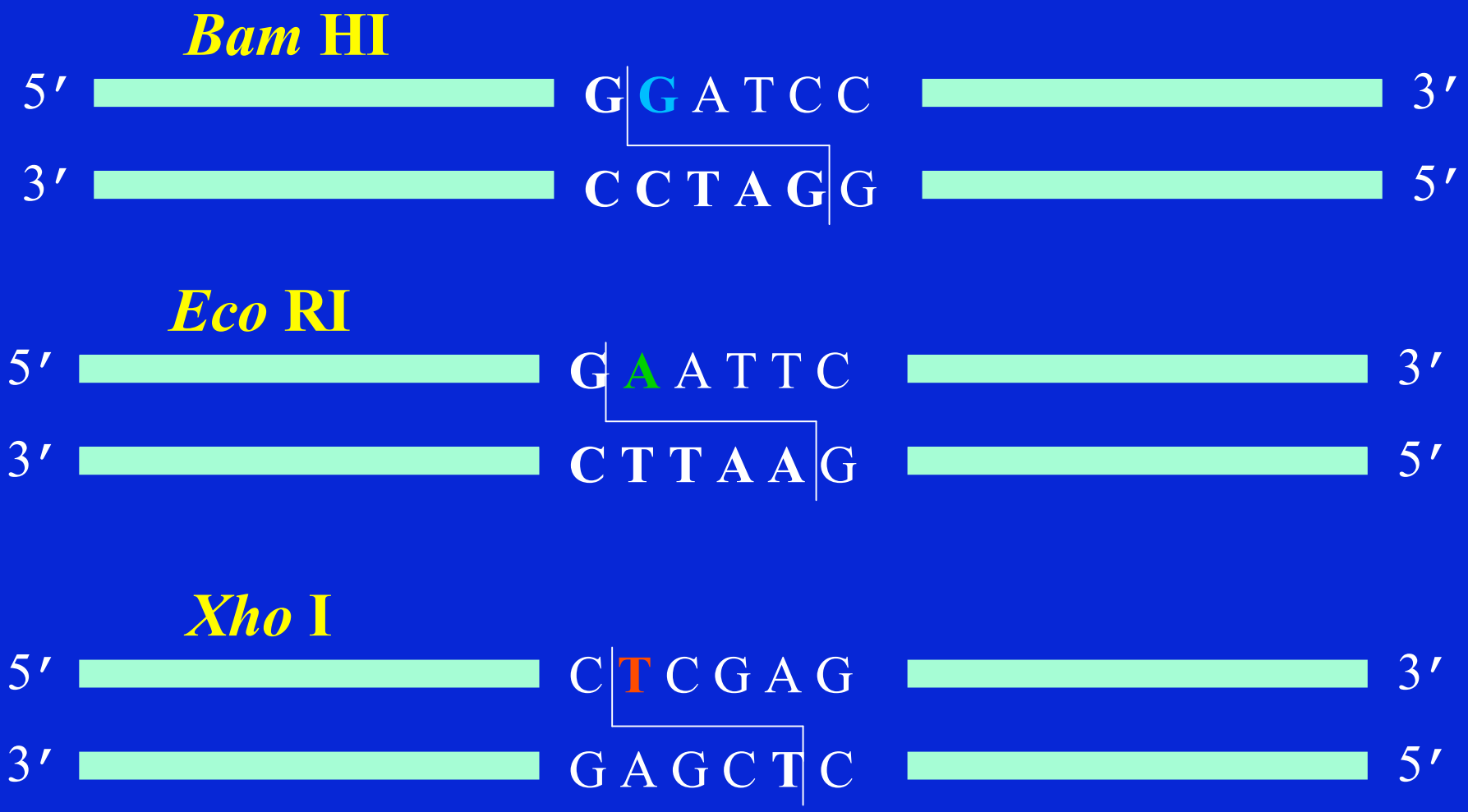
***Xba* I**

Fluorescent labeling



***Xba* I**

Restriction cleavage and fluorescent labeling



Characteristics of restriction sites and labeling of fragments

Restriction endonuclease	Restriction site	ddNTP	Fluorescent dye	Color of fragment
*Eco*RI	G^AATTC	A	dR6G	Green
*Bam*HI	G^GATTC	G	dR110	Blue
*Xba*I	T^CTAGA	C	dTAMRA	Yellow
*Xho*I	C^TCGAG	T	dROX	Red
*Hae*III	GG^CC	none		

Other enzyme combination?

Considerations:

- Generate 5' overhang
- AGCT only
- 6-bp cut ?
- Buffer compatibility

Cheap: 7 More expensive: 40 Don't care: 55

A	G	C	T
*Eco*RI ($212) * *Hin*dIII ($212)	*Bam*HI ($212) *Bgl*II ($1060)	*Xba*I ($840)	*Sal*I ($1060) *Xho*I ($504)

* Prices are based on NEB's list price of 50,000 units at available largest package

Selection of enzyme combination

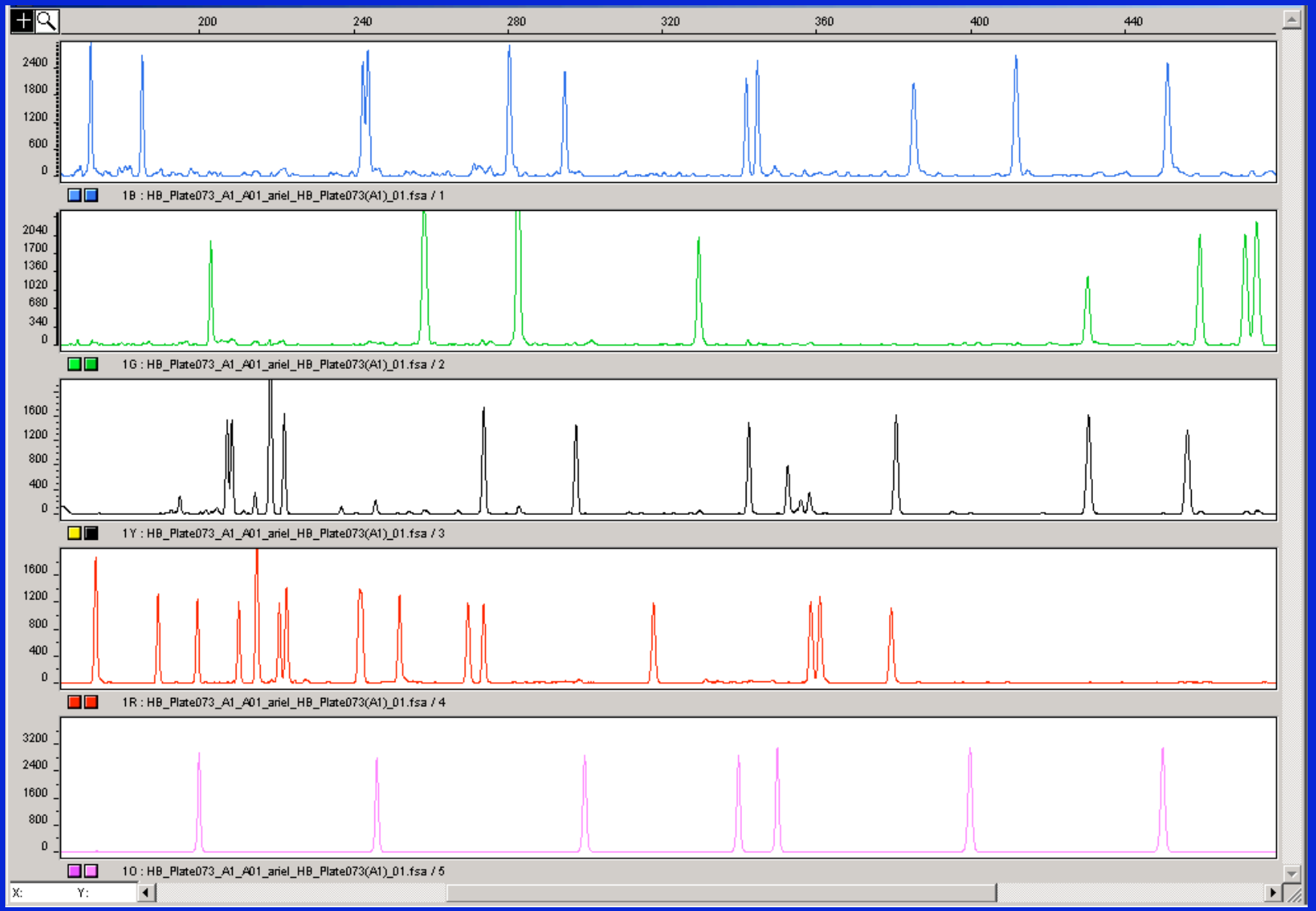
Predicted numbers of restriction fragments in SNaPshot fingerprints of two *Triticum monococcum* and two *T. turgidum* BACs in the range of 50-500 bp

Enzyme	116F2 (107.3 kb)	115G1 (128.6 kb)	BAC1 (173.4 kb)	BAC2 (147.6 kb)	Total (556.9 kb)
*Eco*RI	31	38	32	32	133
*Bam*HI	21	36	53	32	141
*Xba*I	31	47	38	41	157
*Xho*I	26	30	46	23	125
Total	108	151	168	128	--
*Hin*dIII	43	51	68	77	239

Fragment sizing w/ ABI3730

- Any5Dye
- Denatured condition (Hi-Di, 95 ºC, 5')
- 36 cm capillary array / 50 cm capillary array
- Liz-500 Size Standard / longer range

Portion of multi-color fingerprinting profile of a BAC clone



BamHI

***Eco*RI**

***Xba*I**

***Xho*I**

Liz-500

Fingerprinting throughput per single sequencer

Instrument	Daily	Weekly (7 days)	Monthly (28 days)	Annually (330 days)
ABI 3100	480	3,360	13,440	158,400
ABI 3730	2,160	15,120	60,480	712,800
ABI 3730XL	4,320	30,240	120,960	1,424,000

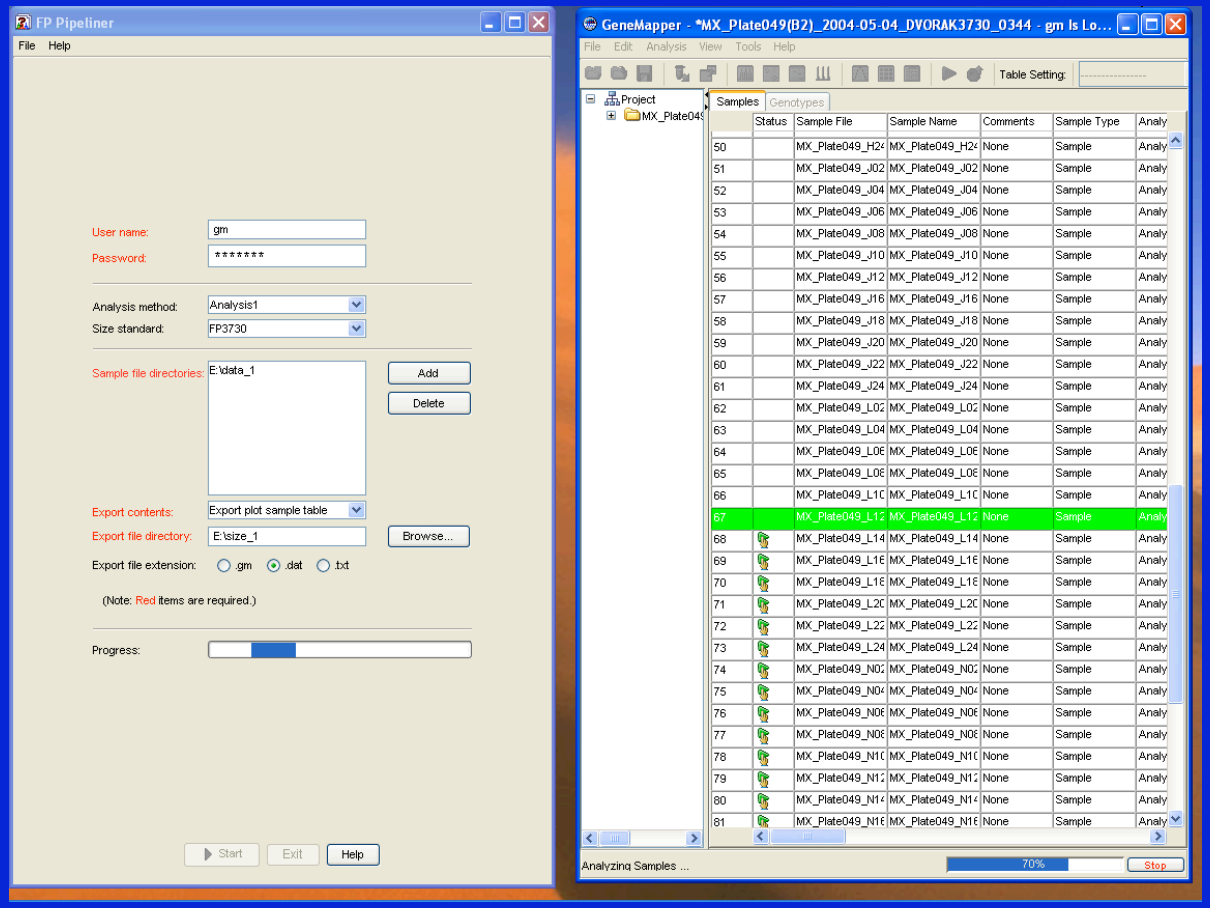
Success rate of the SNaPshot based fingerprinting procedure (ABI 3100)

Library	No. clones tried	No. clones succeeded	Success rate
*Bam*HI	58,261	56,187	96.44%
*Bam*HI BiBAC	19,730	19,455	98.61%
*Eco*RI	52,914	51,217	96.79%
*Hind*III	58,615	56,803	96.91%
*Hind*III BiBAC	26,125	25,907	99.17%
Total	215,645	209,569	97.18%

Fingerprints edit and management

- Fragment size-calling
- True fragments vs. background noises
- Batch fingerprint editing
- BAC cross-contamination check and removal
- Fragment frequency analysis
- Dataset management (clones and fragments)
- BAC-marker hybridization data conversion

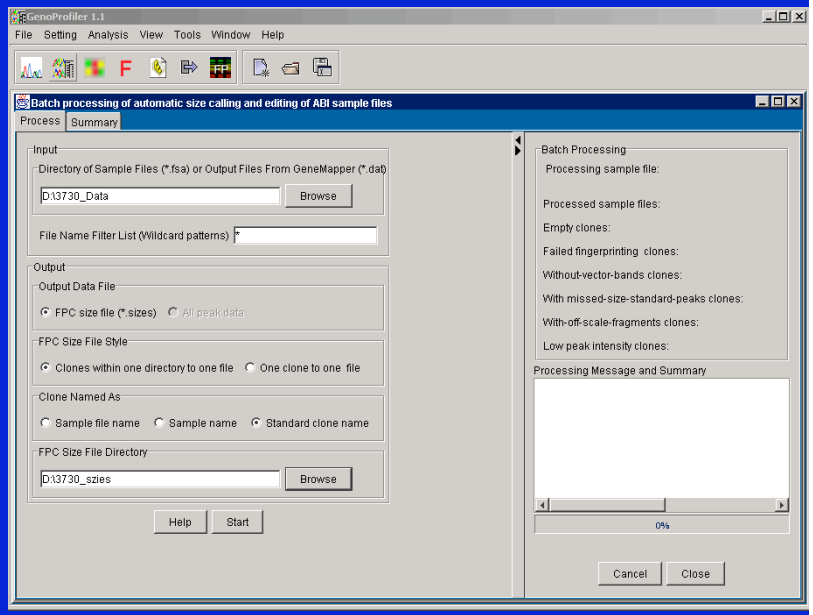
Fingerprints Edit

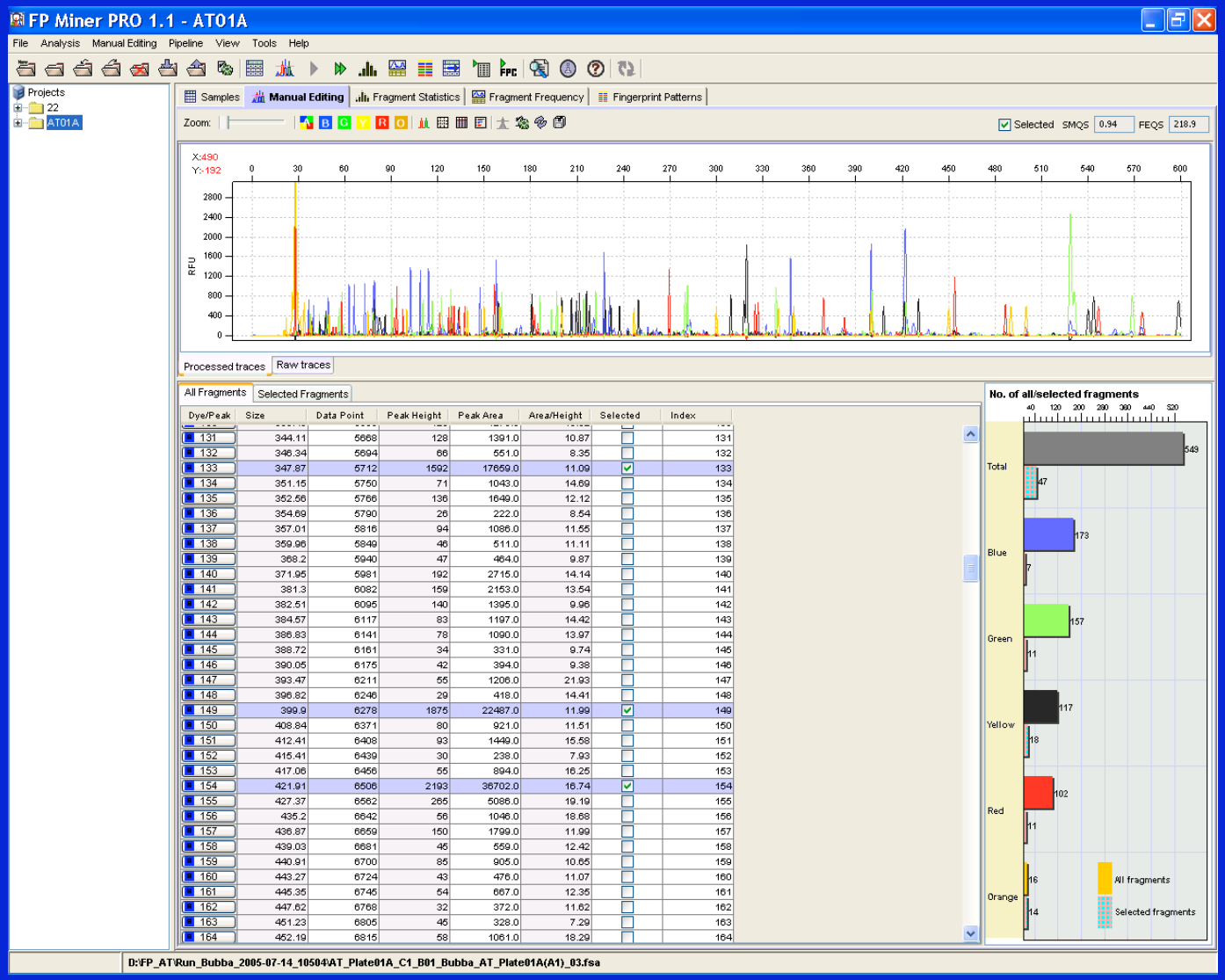


FP Pipeliner (www.bioinforsoft.com)

GeneMapper (www.appliedbiosystems.com)

GenoProfiler (wheatdb.ucdavis.edu:8080/wheatdb





Fingerprints Edit

FP Miner

http://www.bioinforsoft.com/

GenoProfiler

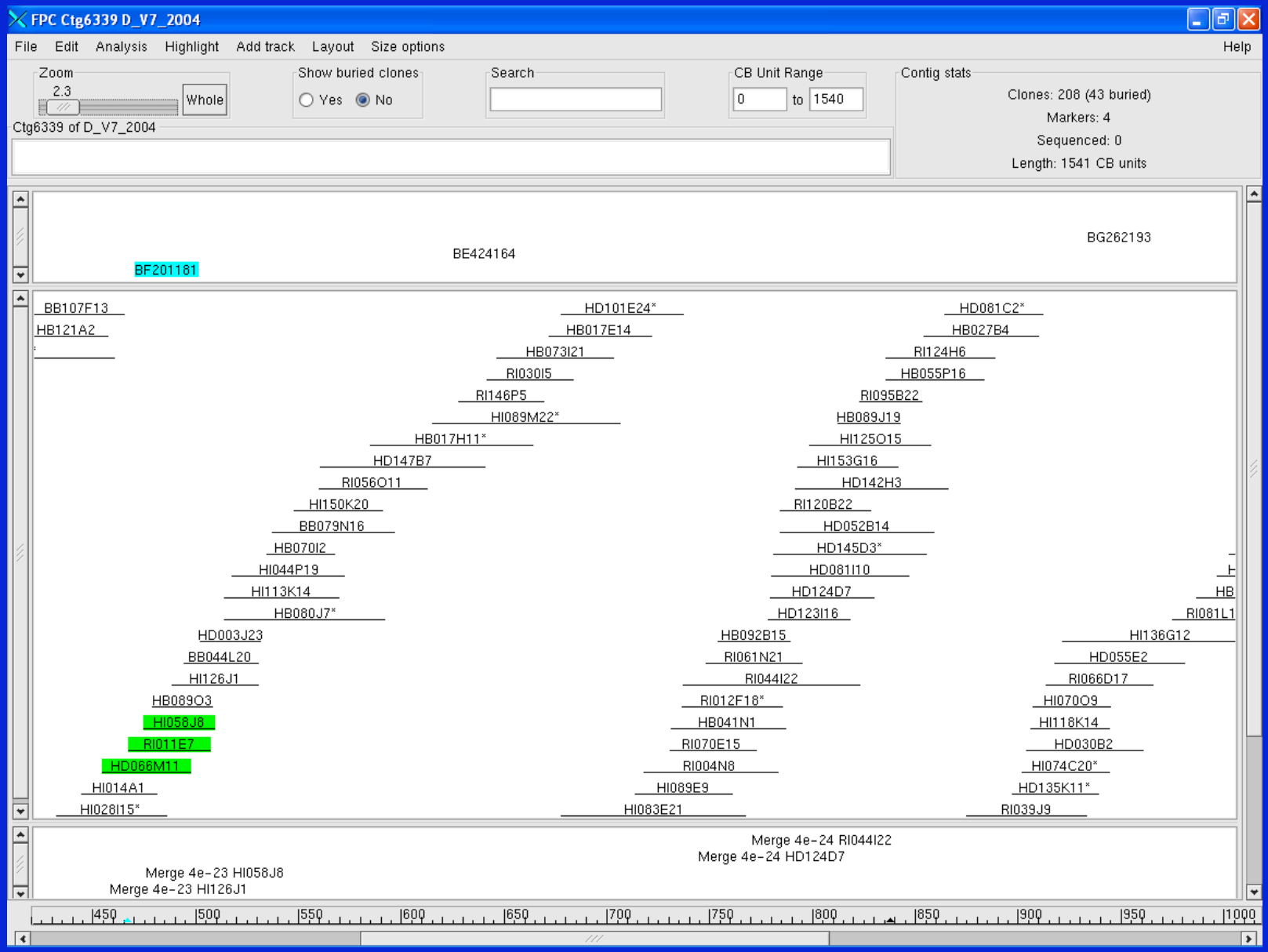
To edit 100,000 BAC fingerprints:

- **input -- GeneMapper exported text file**
  ≈ 5 min.
- **input -- 3100/3700 samples files (.fsa)**
  ≈ 4.5 hr.

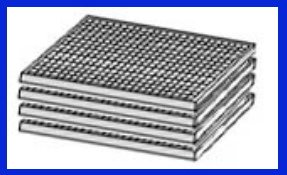
Computer: 3.2 GHz CPU, 1.0 GB RAM

http://wheatdb.ucdavis.edu:8080/wheatdb/

A sample contig (≈ 3 Mb)

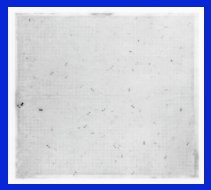


BACs



DNA isolation

Reactions



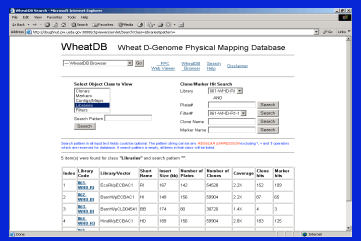
BAC-marker integration

ABI 3730XL

Fragment sizing

Data processing

Contig assembling



Integrated databases

http://wheatdb.ucdavis.edu: 8080/wheatdb

Luo *et al. Genomics* 82:378-389, 2003

Current Applications

Wheat

Rice

Barley

Soybean

Citrus

Sorghum

Brassica

Grape

Tomato

…

Catfish

Rainbow Trout

…

Summary

- High information contents
- Accurate fragment sizing
- Simple procedure, high reproducibility
- Relatively inexpensive
- Automated pipeline, high-throughput

Acknowledgements

Jan Dvorak
Carolyn Thomas
Patrick E. McGuire

Olin D. Anderson
Frank M. You

C. Robin Buell
Shu Ouyang
Jo Hsiao

DBI-0077766

NSF/Plant Genome Research Program

“Mapping and sequencing large genome: Let’s get physical!”

_ Meyers *et al.* (2004) Nat. Rev. Genet. 5: 578