

A Microsoft Access Based Database System Applied to the Management of Cotton Physical Mapping Data

James C. Estill¹, Barry Marler¹, Dawn Williams-Coplin¹, Richard D. Noyes²,
Mark Osterlund¹, Scott Lee¹, Katy M. Rainey³, Mila Pearce¹, Gary Pierce¹, Andrew H. Paterson¹

1. 158 Riverbend Research Lab, University of Georgia, Athens, GA 30602
2. Campus Box 334, EPO Biology, University of Colorado, Boulder, CO 80309
3. 320 Bradford Hall, Dept. Plant Breeding, Cornell University, Ithaca, NY 14853

Data Entry

The BACMan (BAC Management) application data entry user interface is a series of user friendly forms.

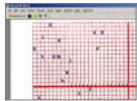


Scoring Films

Films with few hits are hand-scored and the data are manually entered. While films with large number of hits (>25) are manually scored on a transparency overlay. These data are then entered into the database using a manual data input method or text recognition software. All data are tractable by the films individual Bar-Code ID.

Manual Bar-Code Data Entry

An MS Access interface where a Bar-Code scanner is used to jump to the data entry page of a particular film which has been hand scored.

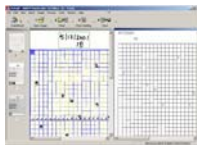


Manual BACGrid Data Entry

The MSWord VBA application BACGrid is used to manually digitize data from hand-scored overlays.

Text Recognition Data Entry

Hand scored overlays are scanned and read by the ABBY FineReader ICR software then sent to the MSEXcel VBA application ABBYEater.



Data Analysis

Routine analysis of BAC hybridization data can be quickly done in the BACMan forms user interface.

Deconvolution

BACMan can deconvolute multiplexed data in a number of different formats including the "grid" and "triangle" formats.

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46
41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66
61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86
81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100	101	102	103	104	105	106

Grid Format

12 Filters: A-L
36 Probes: 1-36

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46
41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66
61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86
81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100	101	102	103	104	105	106

Triangle Format

9 Filters: A-I
36 Probes: 1-36

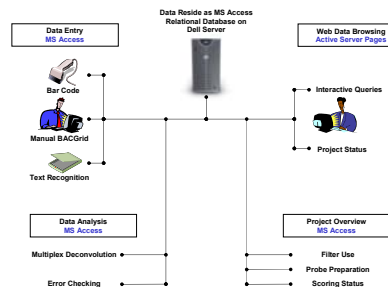
Error Checking

BACMan facilitates error checking in multiplexed data sets by identifying outliers in the deconvoluted data including the number of hits per filter and the number of hits per probe.

Abstract

A Microsoft Access database has been developed for the management of BAC hybridization data related to our ongoing physical mapping effort in three species of cotton, including the two cultivated 'AADD' tetraploid species *Gossypium barbadense* (Pima S6) and *G. hirsutum* (Acala Maxxa and Tamcot GCNH), and the wild DD genome species *G. raimondii*. BAC libraries for all three species are currently being assayed using genomic and cDNA clones derived from linkage maps, and also dispersed repetitive DNA clones. The MS Access BAC data management application BACMan allows for the flexible management of hybridization data, and facilitates data tracking by indexing autoradiographs using a barcode system. Our application also provides for the deconvolution of multiplexed BAC hybridization data, and is directly linked to a visual basic application that allows for the automation of film data entry. The physical mapping data set is available online using Active Server Page technology which directly queries the MS Access database. This database system can be easily scaled to large or small physical mapping projects, and thus provides a low-cost alternative to more high-end Unix based database systems.

BAC-MAN BAC Data Management



Database Relationships



Web Browsing

The BAC data are made available online as it is processed. The web interface was developed using MS InterDev 6 with MS FrontPage 2002. This interface directly queries the MS Access database providing availability of a constantly updated dataset.

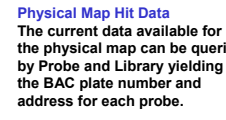
Cotton Genome Database

The web site for the cotton genome database is served on the same Dell server as the MS Access database.



Data Available

Data for the Genetic Map and Physical Map for cotton will both be available for browsing and queries.



Physical Map Hit Data

The current data available for the physical map can be queried by Probe and Library yielding the BAC plate number and address for each probe.



Physical Map Data Overview

Overview tables of project status are available online including the Number of recorded hits per library.

Project Status

Project status can be easily monitored in a series of forms providing data on filter use, probe preparation, and the scoring status of films.

Filter Use

BAC filter use and hit number data are available via forms based queries to help determine when to retire BAC filters.

Probe Preparation

Probe preparation, probe sequence, and probe type are linked to experimental data. The status of probe preparation and use are easily followed and thus provide a high degree of quality control.

Scoring Status

The scoring status of experiments is also available via a forms based query. This helps student workers determine which films need to be scored and helps track the number of hits per film and experiment.

Funding for this project has been provided by the National Science Foundation.

See a demo of this database Wednesday at 3:00 in the California Room

<http://www.plantgenome.uga.edu/cotton>

Software

ABBY Fine Reader 5.0 – <http://www.abby.com>
BACMan – <http://www.plantgenome.uga.edu/jestill/BACMan>
BACGrid – <http://www.plantgenome.uga.edu/jestill/BACGrid>
MS FrontPage 2002 – <http://www.microsoft.com/FrontPage>
MS Office 2000 – <http://www.microsoft.com/Office>
MS Access 2000 – <http://www.microsoft.com/Access>
MS InterDev 6 – <http://www.microsoft.com/vinterdev>