

Development of the *BoviAnalyser* cDNA Bovine Microarray

A.S. Leaflet R1877

Artur J.M. Rosa, Gang Ren, James M. Reecy
Department of Animal Science

Summary

DNA microarrays are powerful tools to analyze the expression pattern of thousands of genes simultaneously in response to treatments, during tissue differentiation or related to differences in traits such as marbling or disease resistance. We developed a bovine cDNA microarray, *BoviAnalyser*, containing 11,088 genes spotted in duplicate. A set of 10,608 ESTs was selected from MARC 1-4 libraries to include: 1) all the known genes, 2) unknown, putative or hypothetical genes expressed in muscle, liver or adipose. Functional assignments for these clones were compared to the human by BLASTx to add putative IDs to unknown genes. The selected set included 4484 known genes, 1170 hypothetical proteins, 497 unknown and 4364 not assigned genes. A database of human genes expressed in muscle, liver and adipose was used to evaluate the percentage of human genes are present on the microarray. The *BoviAnalyser* contains 66.37%, 71.02% and 67.34% of the human genes expressed in muscle, liver and adipose. The 12,000 cDNA clones were rearranged from the MARC libraries 1- 4 into 125 96- well plates. Two clones per plate were randomly sequenced to evaluate the consistency of the data. A total of 12 clones (5.5%) had sequence mismatches been excluded from the set. This new library was used to isolate plasmids containing the cDNA, amplify, purify and prepare the cDNAs to be spotted. The *boviAnalyser* also contain a set of 18 controls including negatives (blank, spotting solution and spiking), 10 spiking (*A. thaliana*), B-actin, GAPDH, combined 15 low, 15 average, 15 high expressed and the 45 combined spotted on every subgrid. A small array containing 192 triplicate genes was spotted to standardize the protocols for post processing the slides, probe synthesis, labeling, hybridization and scanning of the slides. Based on our evaluation, we would conclude that MARC libraries are a good source of ESTs for microarray production. The selected set allows the evaluation of changes in gene expression in different bovine tissues including muscle, liver and adipose. The *BoviAnalyser* is a bovine microarray suitable to be used to analyze molecular mechanisms controlling cell proliferation, differentiation in response to different treatments or variation in economic traits such as marbling and disease resistance.

Introduction

Large-scale sequencing projects have been conducted for several different organisms exponentially increasing the amount of information available for bacterial, yeast, plant, and animal genomes, including livestock species. A high throughput bovine EST–sequencing project was conducted by Smith et al. (2001) using four pooled-tissue normalized libraries in order to generate the greatest number of unique expressed sequence tags with the least amount of sequencing. Those sequences were deposited into GenBank and subsequently used in an interactive website, which is publicly available, constructed by The Institute for Genomic Research (TIGR), which facilitate the access to annotated data.

The availability of DNA sequence from organism genomes has laid the foundation for a post-genomic era from which new technologies are being developed, such as functional genomics, proteomics, metabolomics, for example, and will lead to a better understanding of the genetic control of the phenotypic variation observed between cells, animals and species. Animal scientists have long recognized the importance of understanding the gene's regulatory program to better understand it's functional role in the control of the biochemistry and physiology related to animal production. This new knowledge should allow for sustained increases in the efficiency of production system by the utilization of genetic selection, exogenous substances and nutritional management, for example (Young, 1987). The effect of different treatments or genetic variability on gene expression can be evaluated with microarray technology, in which the RNA is extracted, reverse transcribed into cDNA, labeled and hybridized to oligos or cDNA targets immobilized in a solid substrate, like a glass slide in a cDNA based microarray. The selection of a gene set representative of the tissue to be studied is an important and laborious task that can be performed more efficiently with the use of bioinformatic tools. Information of genes expression in different tissues in human, mouse or other organism can be used to help select cDNA clones in order to have a set of genes which are highly representative to be spotted in a glass slide, what would allow to increase the probability to develop a success experiment.

Materials and Methods

Data on the MARC libraries is available to researchers at the TIGR website (<http://www.tigr.org/tdb/tgi/btgi/>)

(Table 1, Table 2). A set of 10,608 genes was selected to include: 1) all the known genes, 2) unknown, 3) putative or 4) hypothetical genes expressed in muscle, liver or

adipose. In order to achieve this goal a list of human cDNA library representing muscle, adipose and liver tissues, available at MuscleNet website (Bertoluzzi et al., 1998) (<http://telethon.bio.unipd.it/GETProfiles/Index.html>) were compared to the bovine database by BLASTx utilizing the NCBI website (<http://www.ncbi.nlm.nih.gov>). Functional assignments for these clones were compared to the human by BLASTx to add putative IDs to unknown genes. The 12,000 ESTs clones were rearranged to generate a new library containing 125 96- well plates. Two clones per plate were sequenced and compared to the bovine database to evaluate the consistency of the data. The cDNAs were amplified, purified and prepared for spotting. A small microarray containing 192 triplicate genes was prepared to standardize the protocols for RNA extraction, cDNA synthesis, probe labeling, hybridization, slide washing and scanning.

Results and Discussion

- The MARC 1- 4 libraries were used to select 10,608 ESTs (Table 3). The selected set included: 1) 4484 known genes, 2) 1170 hypothetical proteins, 3) 494 unknown, 4) 4364 not assigned and 5) 18 different controls.
- The overlap analysis against a human database (MuscleNET website) showed that the MARC bovine 1-4 libraries contain 66.37%, 71.02% and 67.34% of the

human genes expressed in muscle, liver and adipose (Figure 1).

- A total of 272 cDNA clones were selected from all 126 96- well plates and evaluated via sequencing. Approximately 12 clones (5.5%) had sequence mismatches been excluded from the set.
- The scanning results for the 192 microarray can be visualized on Figure 2.

Based on our evaluation, we would conclude that MARC libraries are a good source of ESTs for microarray production. The selected set allows the evaluation of changes in gene expression in different bovine tissues including muscle, liver and adipose. The BoviAnalyser is a bovine microarray suitable to analyze the molecular mechanisms controlling cell proliferation, differentiation in response to different treatments or variation in economic traits such as marbling and disease resistance.

Acknowledgements

The authors thank Cargill and the Department of Animal Science- Iowa State University

Table 1: Total bovine sequences in the TIGR

Total sequences			
	In TCs	singletons	total
ESTs	183,797	44,144	227,941
ETs	2,798	577	3,375
Totals	186,595	44,721	231,316

Table 2: Breakdown of the Selected EST clones.

Known Genes	Hypot. Proteins	Unknown	Not Assigned	Controls	Total
4484	1170	494	4364	576	11088
0.404	0.106	0.045	0.394	0.052	1

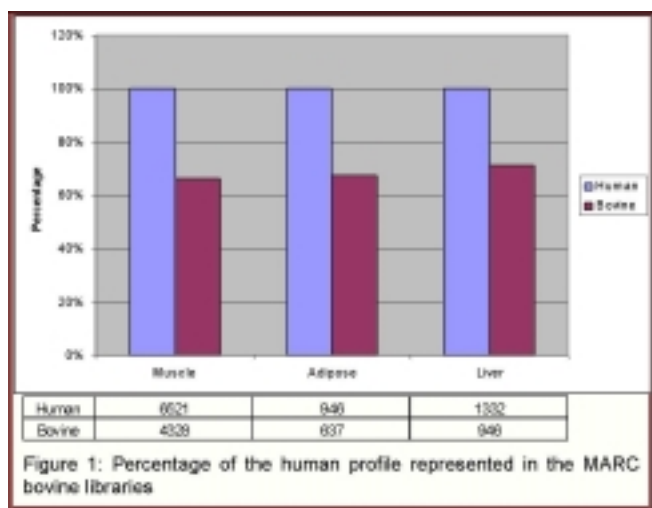


Figure 1: Percentage of the human profile represented in the MARC bovine libraries

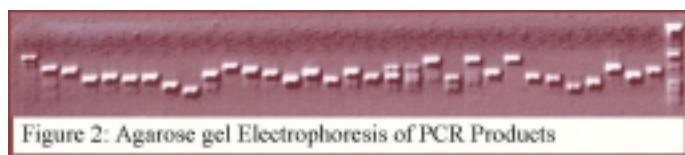


Figure 2: Agarose gel Electrophoresis of PCR Products

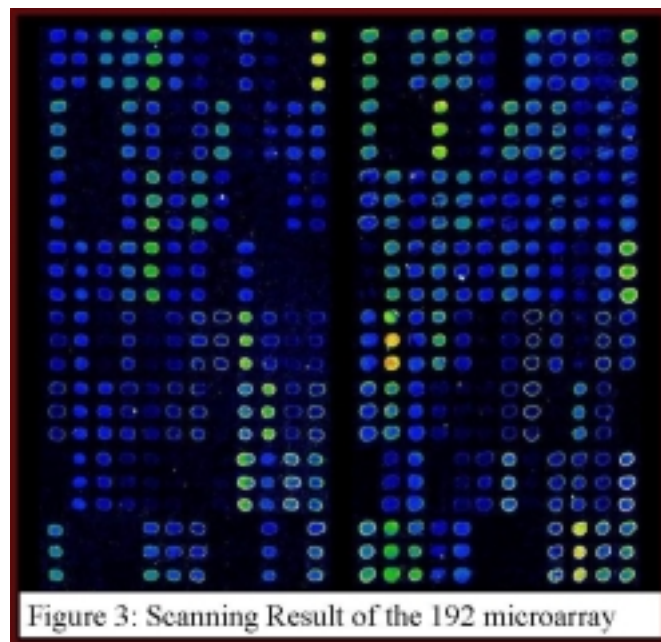


Figure 3: Scanning Result of the 192 microarray