Animal Models of Diabetic Complications Consortium
(U24 DK076169-01)

Annual Report
(2009)

"Coordinating and Bioinformatics Unit for the AMDCC/MMPC"
Medical College of Georgia

Principal Investigator
Richard A. McIndoe, Ph.D.

Richard A. McIndoe, Ph.D.
Medical College of Georgia
Center of Biotechnology and Genomic Medicine
Augusta, GA 30912
Phone: (706) 721-3542 Fax: (706) 721-3688
Email: rmcindoe@mail.mcg.edu

Table of Contents

Animal Models of Diabetic Complications Consortium
(U24 DK076169-01)

Part A:

Principal Investigator's Summary

1. Program Accomplishments:

AMDCC Infrastructure Re-Design

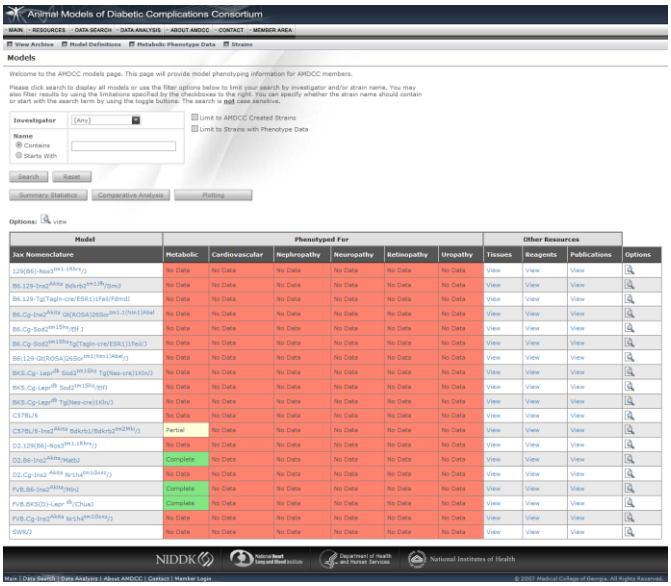
During this last funding period, we were focused on the re-design of the AMDCC web portal and adding functionality to the MMPC web portal. This effort required us to re-write/create key pages of the AMDCC, develop a better overall theme of data flow through the AMDCC website, add object models enhancements to both and redesign the database schemas for both. These design changes are technological advancements for the websites as a whole while others are more functional additions and enhancements. All the modifications to the AMDCC website were done in consultation with the NIDDK program directors by having bi-weekly conference calls. The following sections will describe these changes in more detail with example figures presented when appropriate.

AMDCC Portal re-design In order to better serve the public, we have re-designed the AMDCC web portal to make it easier to find data, protocols and model information for each of the covered diabetic complications. In addition to the redesign, we also implemented an automated system for extracting data from each of the models phenotyped and determining if they meet specific threshold criteria. During this second round of the AMDCC, the focus of the models/strains is much more restricted with their phenotyping being highly regimented and coordinated via the Mouse Generation and Husbandry Core (MGHC). We re-focused the web portal to be more model/strain centric relative to the animals being phenotyped/developed during this second round of the AMDCC. We also made it easier to find complication specific information as it relates to the models/strains, protocols, reagents, investigators, etc. In order to make these changes as flexible as possible it required that we develop a software engine to automate the complication validation and experiment criteria. This software engine provides the AMDCC website with the ability to create validation criteria for each of the diabetic complications. The validation criteria software is extremely flexible and allows the AMDCC to precisely define the criteria in terms of not only phenotype assay thresholds (e.g. blood glucose > 200) but also more time and independent variable oriented criteria. For example, the nephrophathy criteria looks for GFR inulin clearance with a 50% decrease over time and the presence of Arteriolar Hyalinosis. The GFR inulin clearance looks for a trend over some time while the Arteriolar Hyalinosis is categorical evaluation as to whether it is present or not.

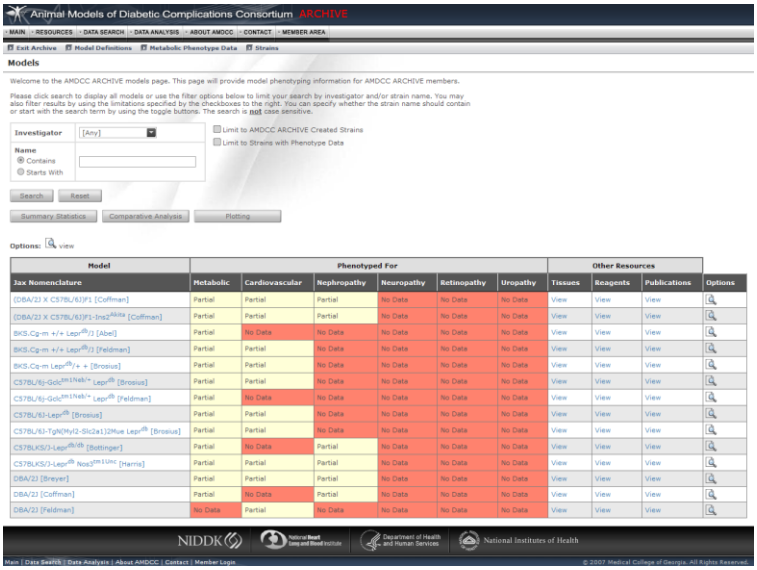
Model Centric Web Portal Because the strains in the second round of the AMDCC are very specific, we focused the new website on these strains/models. The MGHC provides the baseline phenotyping and data for all the animal models developed during this second round. The implications of these two concepts are that the data from the first round was “archived” with the default view of the web pages being the new data generated during the second round of the AMDCC. We archived the first round data because it was difficult to organize the strain data because each investigator provided differing levels of phenotyping. Although we have archived the old first round data, we provide an easy way for the public to view it if they request it (by simply clicking a hyperlink that toggles between the two data sets). To solve the differing degrees of phenotyping for the various strains, we provide the archived data by separating the data based on which lab generated the information. Therefore, if we have the same strain phenotyped in different laboratories, the data will be segregated based on each laboratory. New pages on the site default to the strains/models phenotyped in the second round of the

AMDCC. The ability to change between the new and old data required a re-write of web pages, not only to default to the new data, but also re-focus the context to be strain/model centric.

CURRENT



ARCHIVE

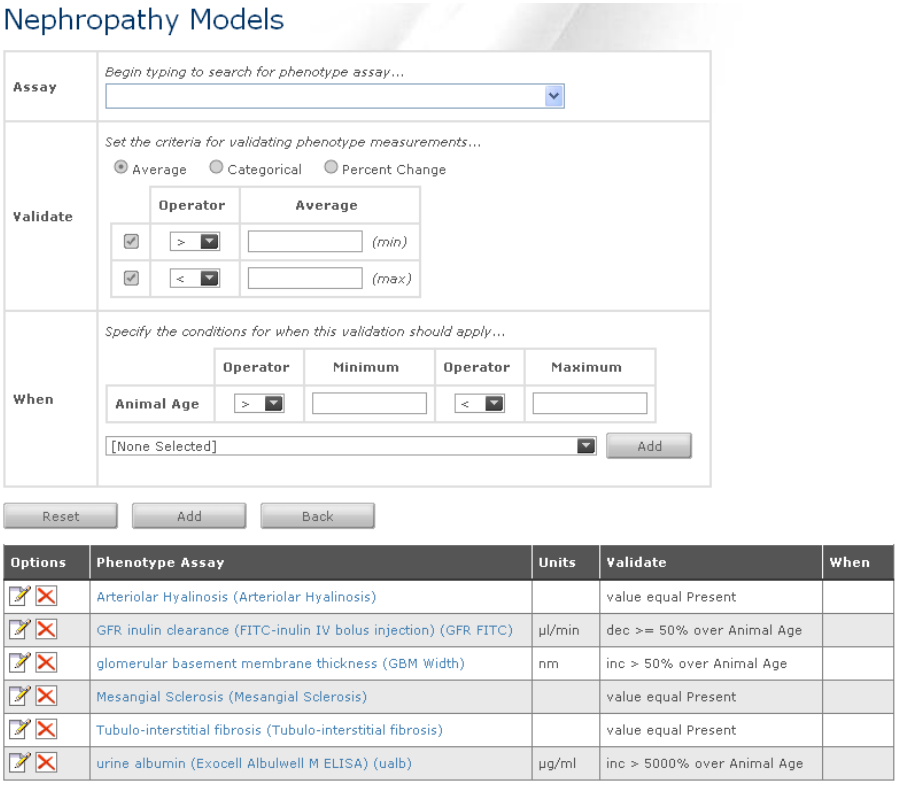


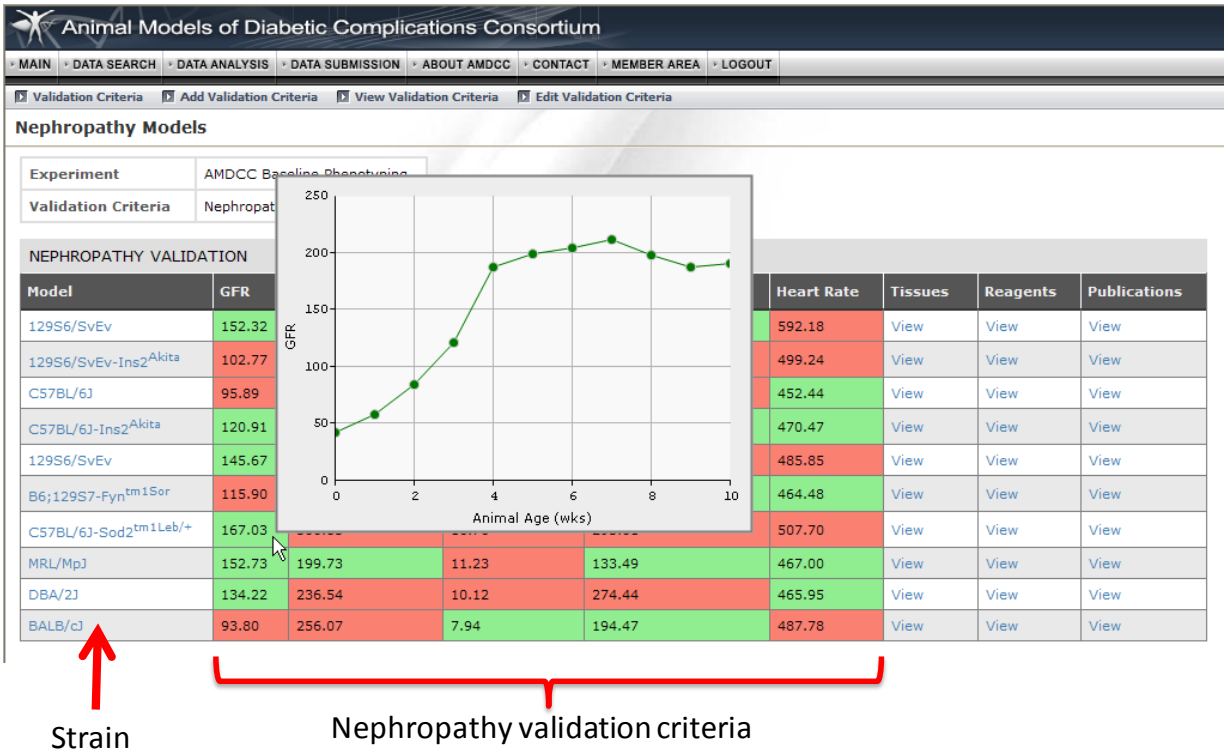
Models/Strains phenotyping information for both the basic metabolic screening and any complication specific phenotyping. Also provides easy links to tissues and reagents for each of the strains.

Complication Specific Information. As stated previously, to make the portal easier to find complication specific resources and information we re-designed the web site. We have significantly reduced the amount of "clutter" on the home page and focused the portal to be complication specific. For example, clicking on the Model resource tab on the left menu on the home page provides the user with access to each of the complication specific models/strains for both archived and current models.

Complication Validation Criteria. One of the main goals of this new redesign is to provide a bird's eye view of each mouse model and how well the strain phenotype fits specific complication phenotypes. Each of the complication committees are responsible for determining what experimental outcomes define a valid model for a particular complication. These validation criteria can be quite complicated and include criteria that are not simple mathematical inequalities (e.g. blood glucose >= 200 mg/dL). They can include quantitative ranges, change of values over time or categorical conclusions (e.g. presence of fibrosis). In order to provide a flexible system, we have developed an algorithm that allows us to define specific rules for the validation criteria and use those rules to determine whether or not a strain passes the criteria. To accomplish this goal, we created the database schema to store the rules as well as the object model to interpret and implement the rules. Because we had previously developed the concept of datasets, the validation criteria API uses a dataset object as the input for the data to validated against. We have stored predefined datasets for nephropathy and cardiovascular complications as these committees have "finalized" the criteria. The figure to the right shows the interface used to define the validation criteria. We are working with the neuropathy and uropathy committees to determine the specifics for their respective validation criteria. We have completed the full implementation of the automated validation criteria system as well as updated

the web site pages to use this system. An example Nephropathy page is given below to illustrate the concept. As can be seen, all the nephropathy models are listed with the various complication specific phenotyping criteria. For each of the assays that define the validation criteria, the system automatically color codes the cells based on whether or not that model passes those specific criteria. In addition, we developed AJAX based data retrieval that occurs during mouse over events to present more data (this is illustrated by the graph generated while hovering over the GFR cell).





Complication Experiment Criteria. In order to keep track of the phenotyping aspect of the animal models, we also developed a software system to determine whether or not the phenotyping has been completed. These 'Experiment Criteria' are flexible and can define the number of phenotype assays measurements expected during a specific time period. For example, the Baseline Metabolic Phenotyping has expectations of at least 3 spot urine albumin measurements and 5 plasma glucose measurements when the animal is between 4 and 26

weeks of age. The portal changes the cell color of the grid on the website based on how much of the Experiment Criteria are completed (ie. Green when complete, Yellow when in progress).

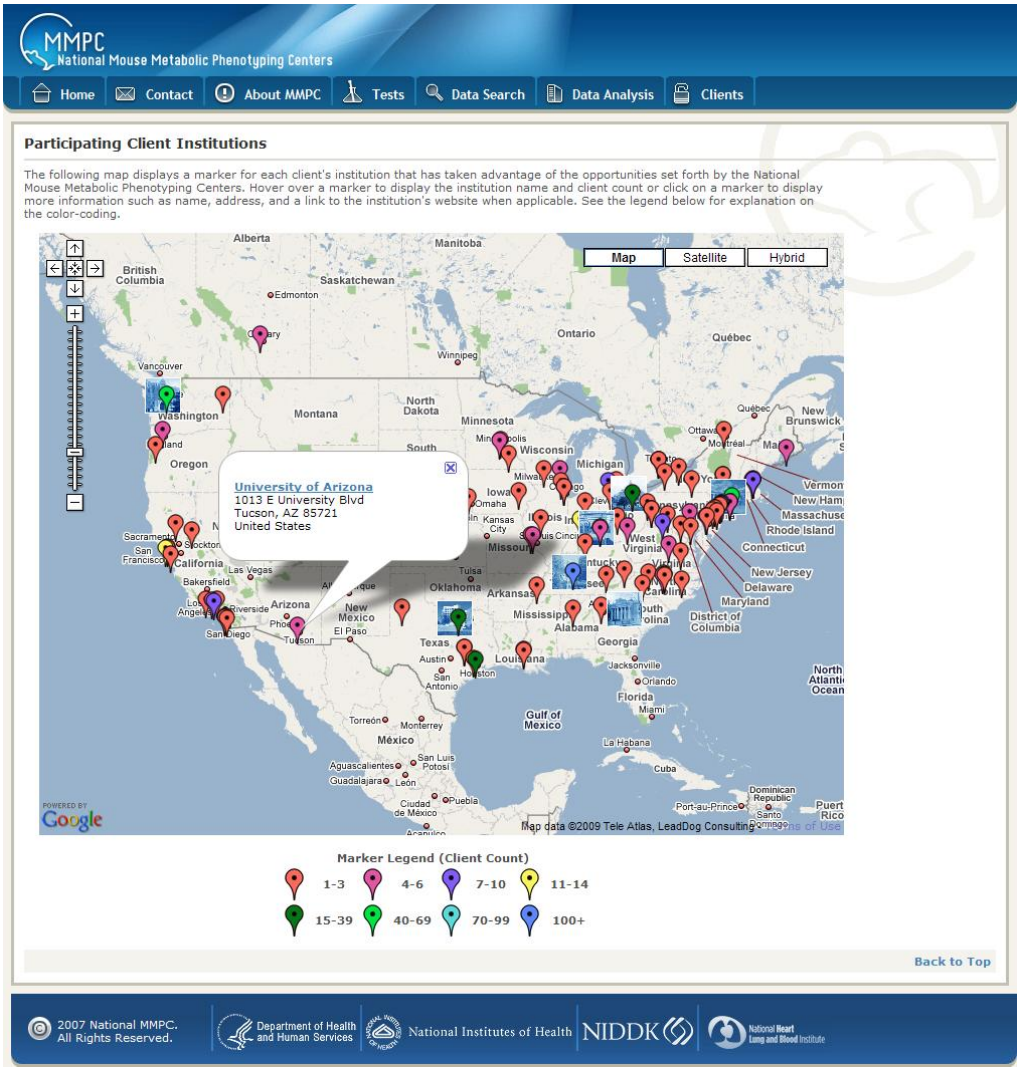
AMDCC Pilot and Feasibility Program. Because the AMDCC has increased its activities regarding funding Pilot and Feasibility projects, we took advantage of the infrastructure we built for the MMPC funding program management. Specifically, we activated the same structure in the AMDCC web portal. The funding program system provides a complete workflow for both the review and management of P&F applications. One of the differences between the AMDCC and MMPC is that AMDCC accounts are restricted to specific funded investigators. However, the P&F program is available to the public and not to AMDCC investigators which required us to update the AMDCC portal to allow new accounts for individuals wanting to submit a P&F application without requiring an account review. Because these are highly restricted accounts, we had to update our privilege and group security roles and the web portal pages to deal with these new restricted user accounts. We have completed all the necessary updates and have fully implemented the funding program system in the AMDCC.

During this last funding cycle, the AMDCC solicited applications for the 2009 Pilot and Feasibility program. The solicitations were open to the scientific community with applications restricted to investigators that are not funded AMDCC investigators and scientific projects that have relevance to the AMDCC mission. All the information regarding the scope of the applications, application submission instructions and the review/funding process are posted on the AMDCC website and available for download. For this cycle, we received 31 applications and funded 12 of them. Each application was reviewed and scored by two external reviewers and funding decisions were made by the NIDDK/NHLBI program officers and the AMDCC external advisory committee.

AMDCC Webservices. Because of the extensive changes in the AMDCC object model, we needed to update the AMDCC webservices to use this new model. During this last year, we completed re-writing and updating the AMDCC web services to use the new unified object model. These web services provide external systems to access the publicly released data without using the AMDCC web portal (ie electronic transfer of data on request). The web services cover all the models/strains, experiments, assays, animals, etc. that are stored in the system. We are currently working on creating a validation criteria web service where other investigators could supply their animal model data and check it against our criteria. This is not implemented yet.

MMPC Portal Updates

During the last year, the MMPC has increased the number of courses and programs available to the public. To make this a more dynamic process, we have updated the site to handle course information as well as new announcements for upcoming events related to the MMPC's activities. To provide a better accounting of the both the AMDCC and MMPC activities, we also developed a number of scripts that automatically run once a month to search PubMed for manuscripts that cite the AMDCC/MMPC grants and insert these into their respective publication database. For the MMPC, the institutions page was also re-designed to use a Google Maps application to illustrate the number and origin of clients that use the MMPC. This page queries the database for the client institutions with completed orders and inserts a pin at that location on the US/World map. The pins are color coded based on the number of clients from the institution.



MMPC Catalog Webservices In addition to the web portal updates and management we have also created a series of web services that provide access to the MMPC catalog information. These web services provide the individual MMPC centers a mechanism for electronically retrieving catalog data for their Center's and displaying it on their center web sites. Meaning, the centers can maintain their catalog information on the National MMPC site, but still have it displayed on their custom center sites look and feel. This will allow both the National and individual MMPC sites to remain in sync with respect to their catalog information. To demonstrate the principal, we downloaded the Case Western MMPC web site web pages and CSS files and created a catalog page using their look and feel. The figure below is that page. The web services are extremely fast and do not show any signs of delay when rendering the page.

MMPC Catalog Web Services

Case Western Catalog at www.mmpc.org

Case Western Catalog using CWRU styles

MMPC Funding Programs

During this last cycle, we will have conducted 8 funding program solicitations. These were the yearly MMPC Pilot and Feasibility program followed by six quarters for the MMPC MICROMouse program. The 2008 MMPC P&F program had 7 applications with 3 being funded through the MMPC. These awarded applications will be completed in November of 2009. The 2009 MMPC P&F had 3 applications with the funding decisions not yet determined. The MMPC MICROMouse program accepts applications throughout the year with reviews occurring quarterly. This last year we reviewed 8 applications over four quarters (2008) and funded 5 applications. During the first 2 quarters of 2009, we reviewed 5 applications. Two of these were funded and three are still under review.

2. Address previous EAC comments:

We thank the EAC for the comments regarding the direction of the AMDCC web portal. Regarding one specific comment:
"Once these modules are completed, it may be time for a refresher course for AMDCC investigators to encourage data entry."

We agree with this assessment and have provided web conferences using GotToMeeting for investigators needing to upload data. Based on the response from the AMDCC PIs and available funds, we may need to go to specific sites for those investigators that are having conceptual or technical problems that cannot be resolved via a web conference.

3. Publications:

A modified hyper plane clustering algorithm allows for efficient and accurate clustering of extremely large datasets. Sharma A, Podolsky R, Zhao J, McIndoe RA. 2009 Bioinformatics 25:1152-1157.

Brosius III FC, Alpers CE, Bottinger EP, Breyer MD, Coffman TM, Harris RC, Kakoki M, Kretzler M, Leiter EH, Levi M, McIndoe RA, Sharma K, Smithies O, Susztak K, Takahashi N, Takahashi T for the AMDCC. Mouse Models of Diabetic Nephropathy. JASN 2009 (In Press)

A parallelized version of the Significance Analysis of Microarrays algorithm provides a significant increase in speed and dataset size for gene expression analysis. Ashok Sharma, Robert Podolsky and Richard A. McIndoe. *Manuscript in preparation*