AMDCC

THE ANIMAL MODELS OF DIABETIC
COMPLICATIONS CONSORTIUM

AMDCC Coordinating and Bioinformatics Unit

1) Web portal overview

2) Website IT Structure Changes

3) Future Work and Enhancements

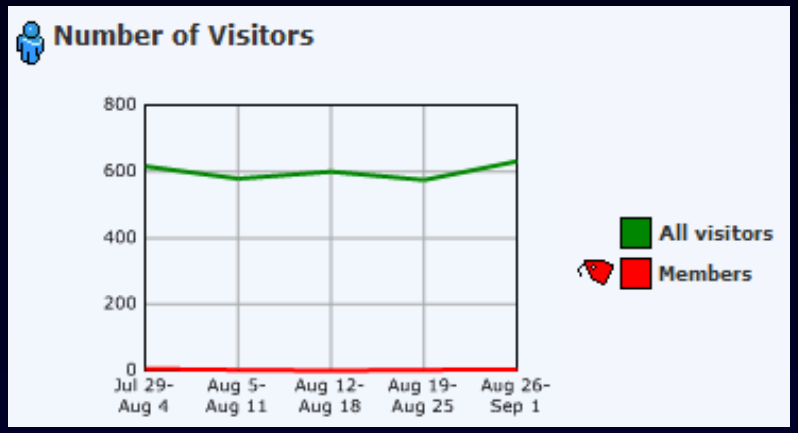
Web Portal Overview

Average number of visitors per month: 3149

Average number of visits per week: 601

Average number of visitors per day: 98

Sample Month



Last six months

All numbers reflect non-AMDCC members

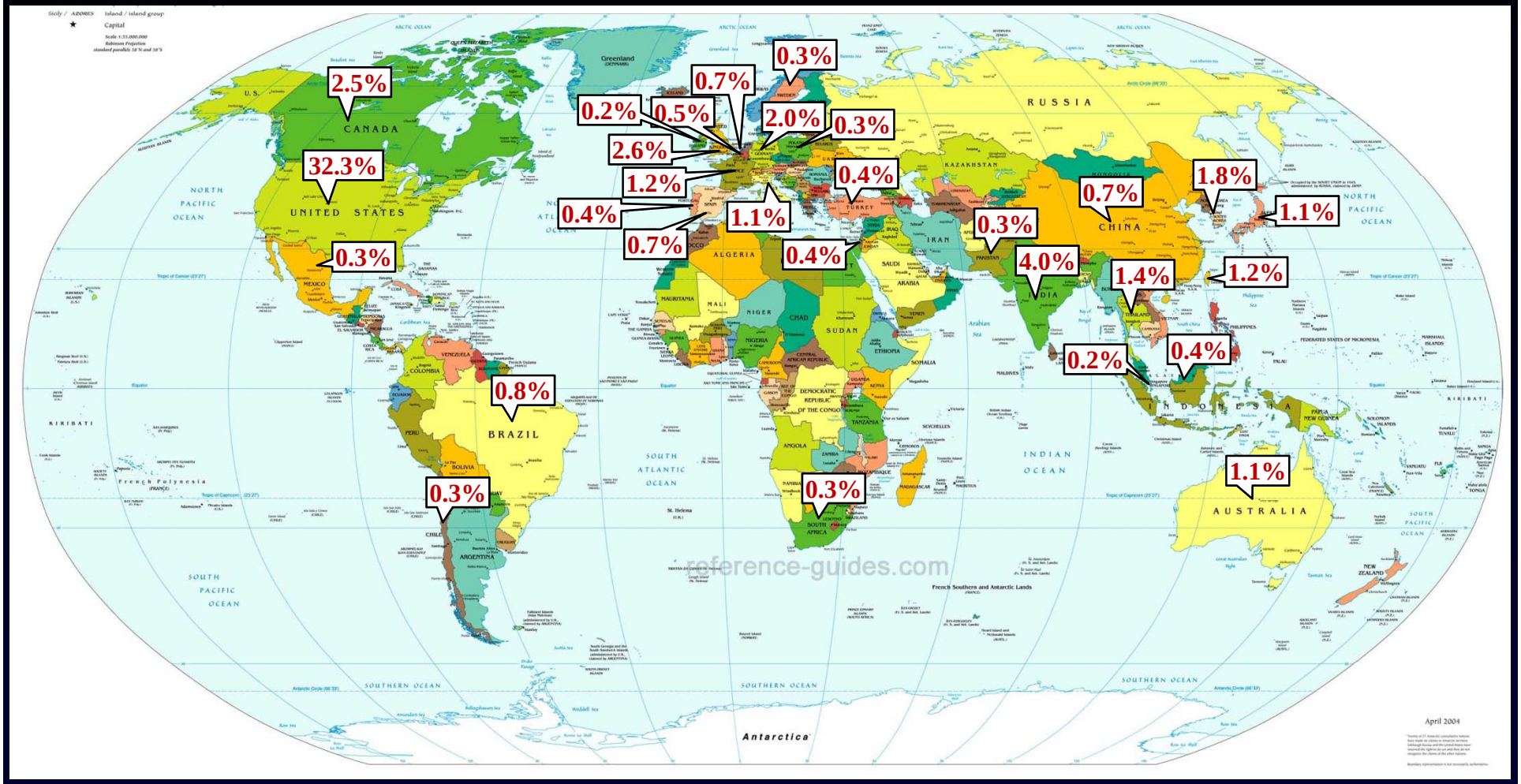
Top Search Engine Referrers

Google	63.4%
http://en.wikipedia.org	2.8%
Yahoo!	2.6%
MSN	0.9%
http://www.niddk.nih.gov	0.5%
Windows Live	0.5%
http://www.med.umich.edu	0.4%
http://www.answers.com	0.3%
http://dels.nas.edu	0.3%
AOL	0.3%
http://www.mmpc.org	0.2%
http://www.cbgm.mcg.edu	0.2%
http://grants.nih.gov	0.2%
http://www.mcg.edu	0.1%
http://www.dxy.cn	0.1%
Other	2.9%
None	24.2%

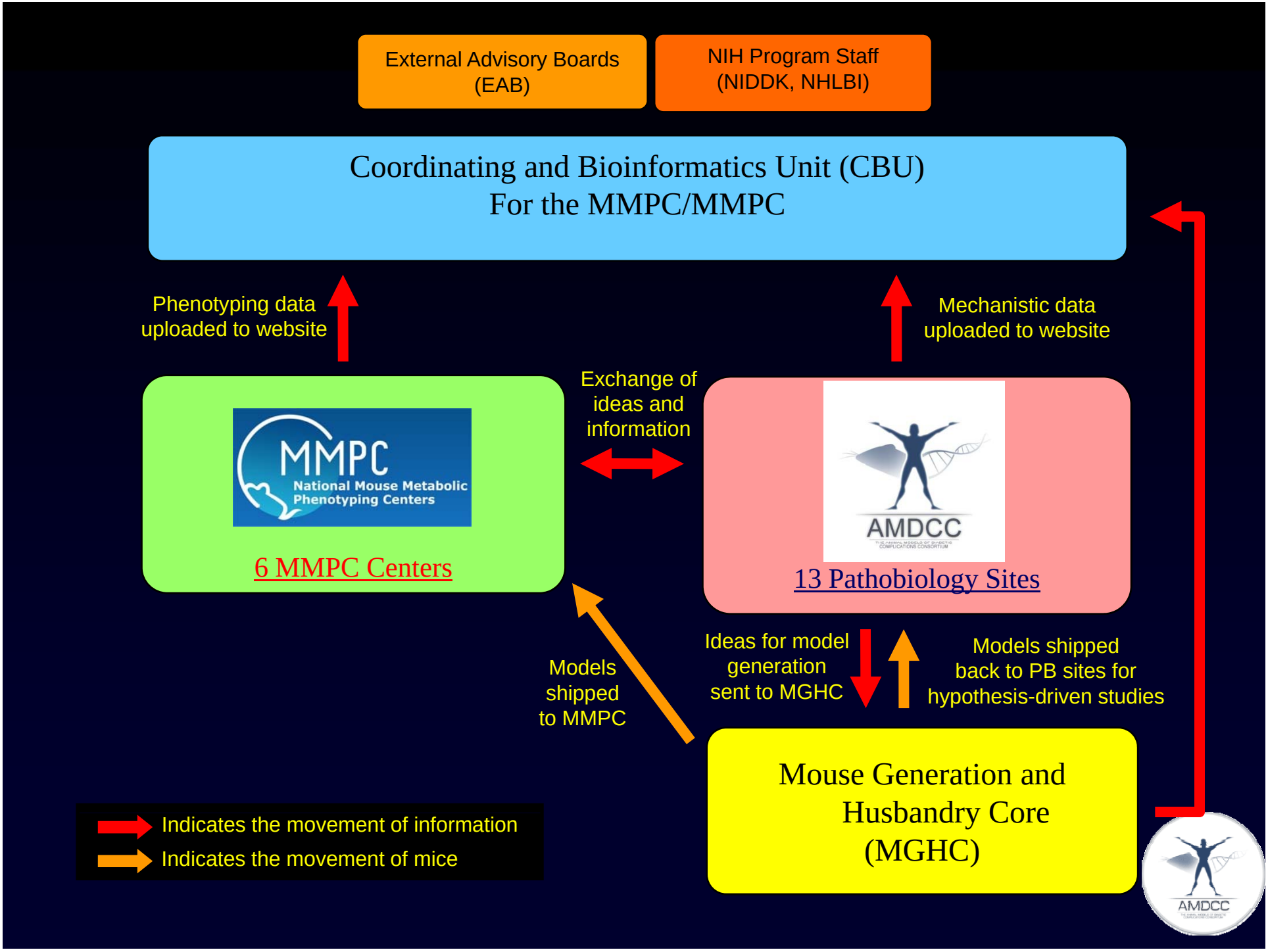
Top Search Terms

tbars	6.3%
thiobarbituric acid	2.2%
catalase assay	2.0%
amdcc	1.6%
ipgtt	1.2%
tbars assay	1.1%
diabets	0.7%
mouse dissection	0.7%
thiobarbituric acid reactive substances	0.7%
tail flick	0.6%
nadh oxidase	0.5%
glucose clamp	0.5%
ampk assay	0.4%
aortic banding	0.4%
thiobarbituric acid assay	0.4%
amdcc protocols	0.4%
mouse perfusion	0.3%
animal models of diabetic complications consortium	0.3%
tbars protocol	0.3%
thiobarbituric	0.3%
insulin elisa	0.3%
streptozotocin	0.2%
catalase protocol	0.2%
amplex red	0.2%
intraperitoneal glucose tolerance test	0.2%
Other	45.2%
None	32.4%

Analysis of Google Search referrer: world wide hits



AMDCC IT Structure

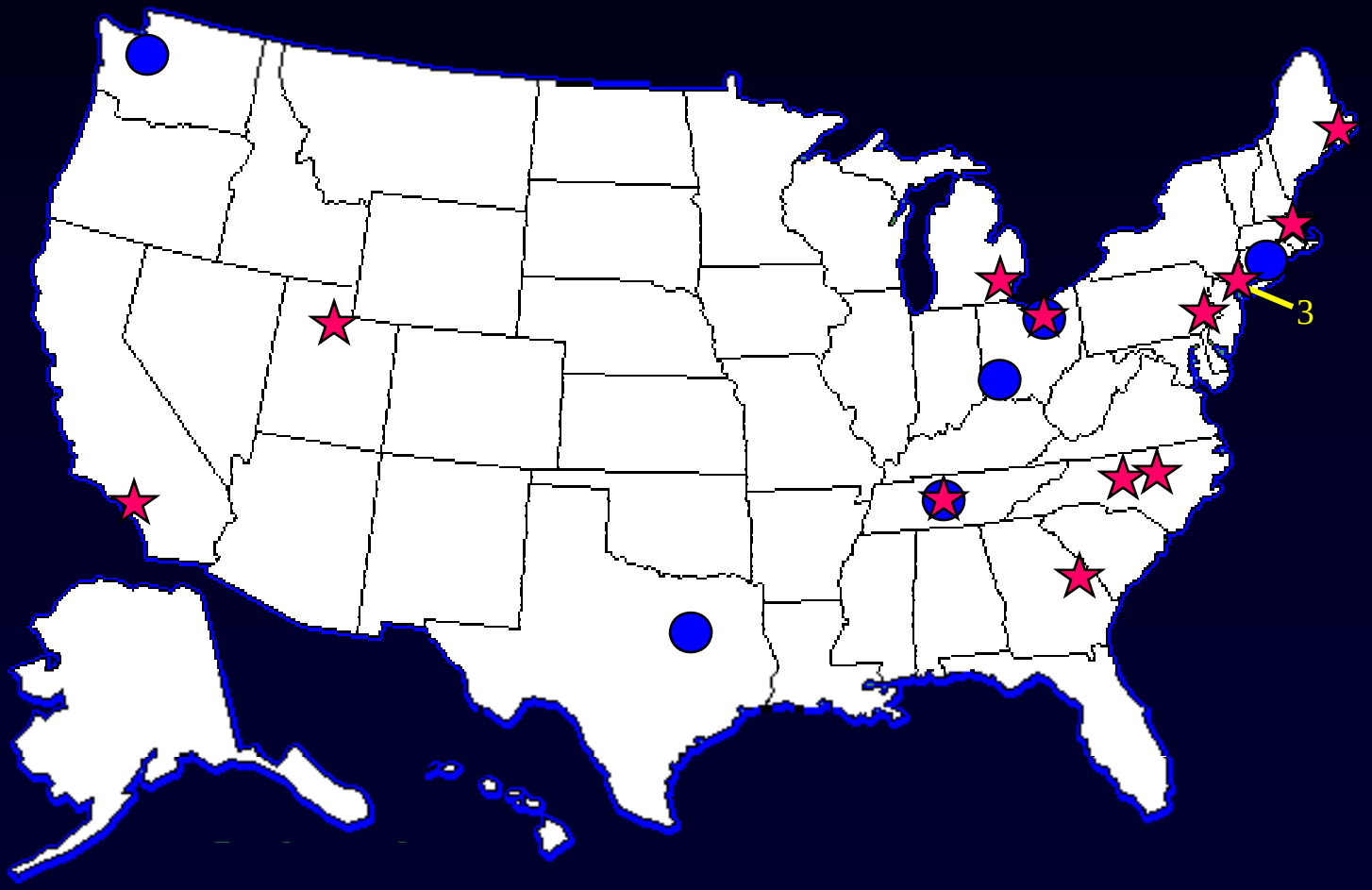

External Advisory Boards (EAB)
NIH Program Staff (NIDDK, NHLBI)
Coordinating and Bioinformatics Unit (CBU) For the MMPC/MMPC
Phenotyping data uploaded to website
Mechanistic data uploaded to website
Exchange of ideas and information
MMPC
National Mouse Metabolic Phenotyping Centers
6 MMPC Centers
AMDCC
13 Pathobiology Sites
Ideas for model generation sent to MGHC
Models shipped back to PB sites for hypothesis-driven studies
Models shipped to MMPC
Mouse Generation and Husbandry Core (MGHC)
Indicates the movement of information
Indicates the movement of mice
AMDCC

AMDCC
THE ANIMAL MODELS OF DIABETIC COMPLICATIONS CONSORTIUM

= ★

MMPC
National Mouse Metabolic Phenotyping Centers

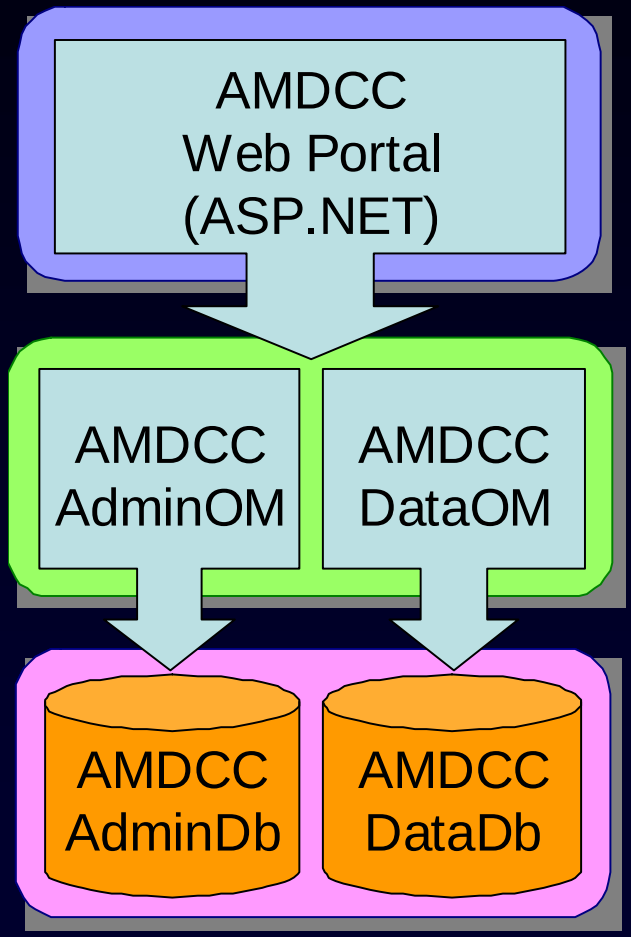
= ●



AMDCC

One of the missions of the CBU for the AMDCC/MMPC was to create an integrated data structure so information and data can be exchanged easily between the MMPC, AMDCC and MGHC.

Original AMDCC IT Structure
(3-Tier Architecture)

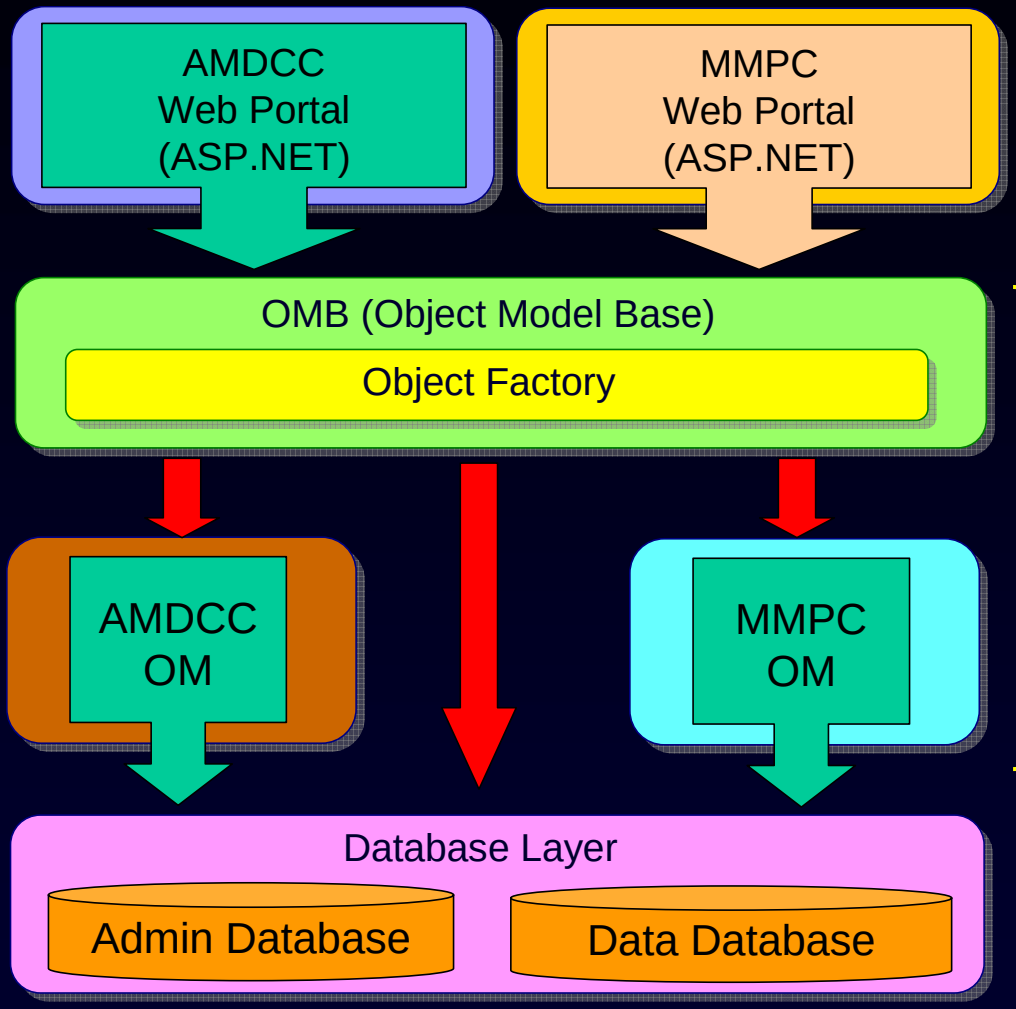


User Interface

Business Layer (OM-API)

Data Layer (Database)

New Combined IT Structure

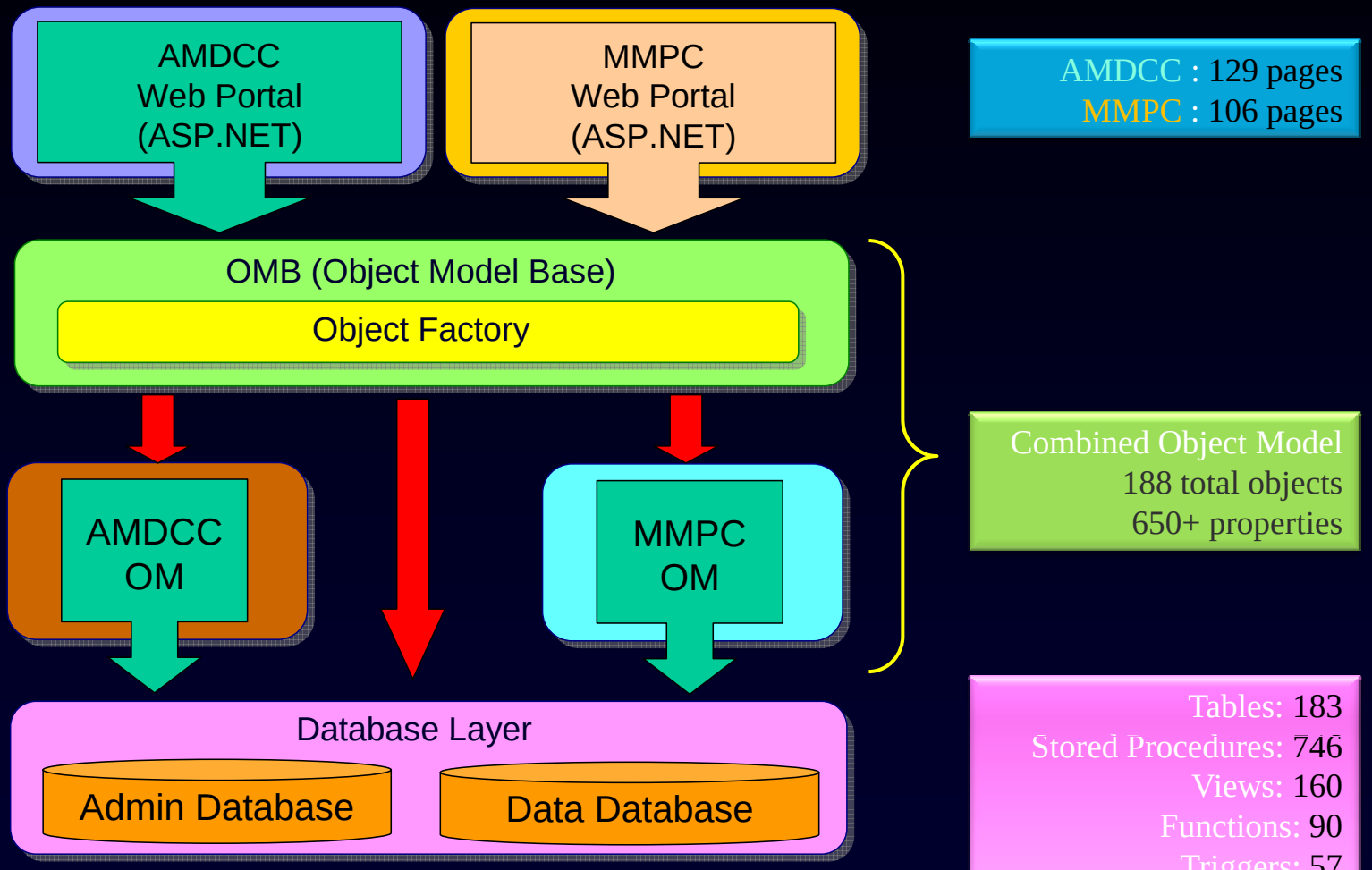


User Interfaces

Business Layer

Data Layer

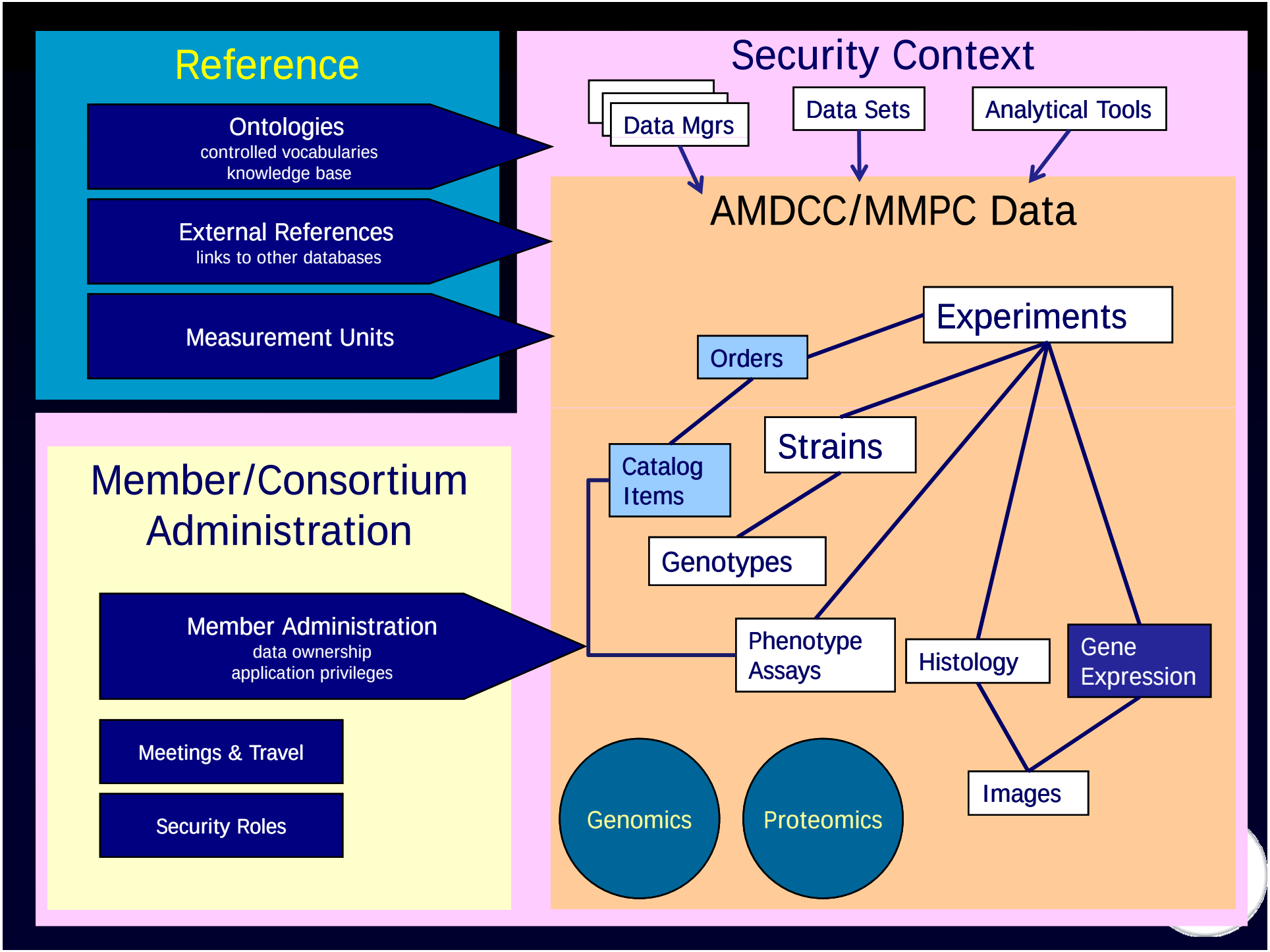
OMB provides the same interface and functionality to both portals. This allows for all common functionality to be implemented in the OMB code base and prevents code from having to be duplicated and separately maintained for each portal. OMB also allows its functionality to be overridden and / or customized through the concept of Object Factory.



AMDCC : 129 pages
MMPC : 106 pages

Combined Object Model
188 total objects
650+ properties

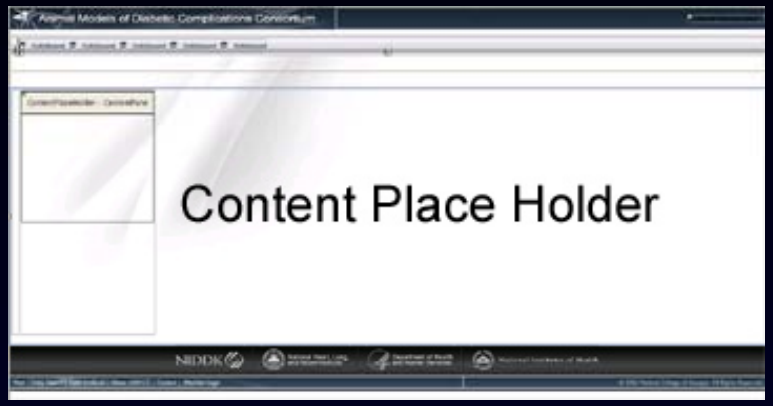
Tables: 183
Stored Procedures: 746
Views: 160
Functions: 90
Triggers: 57


Reference
Ontologies
controlled vocabularies
knowledge base
External References
links to other databases
Measurement Units
Security Context
Data Mgrs
Data Sets
Analytical Tools
AMDCC/MMPC Data
Experiments
Orders
Strains
Catalog Items
Genotypes
Phenotype Assays
Histology
Gene Expression
Images
Genomics
Proteomics
Member/Consortium Administration
Member Administration
data ownership
application privileges
Meetings & Travel
Security Roles

AMDCC IT Changes/Updates

New Software Infrastructure was re-written/written in .NET 2.0

Master Page



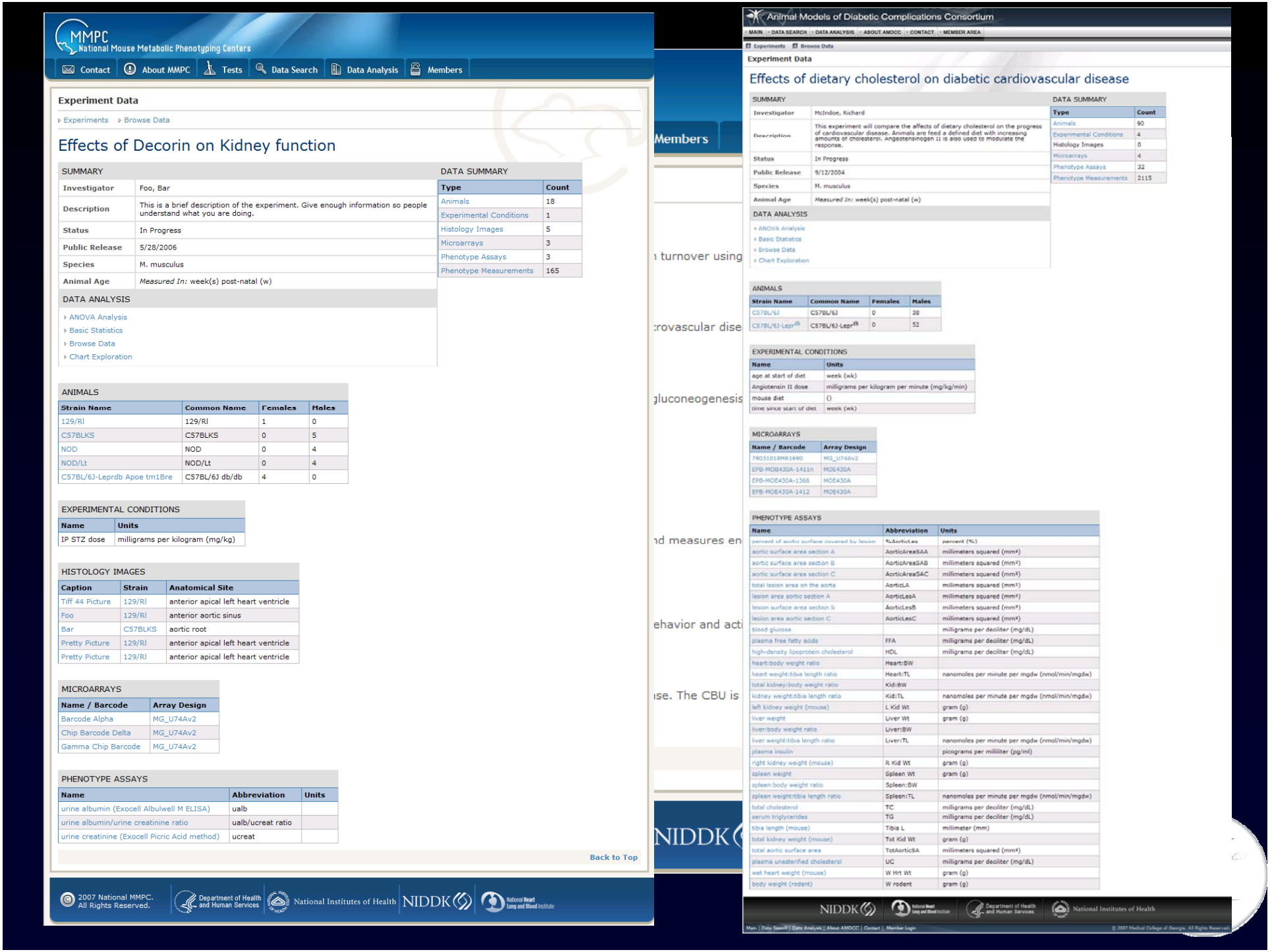
+

Content Page

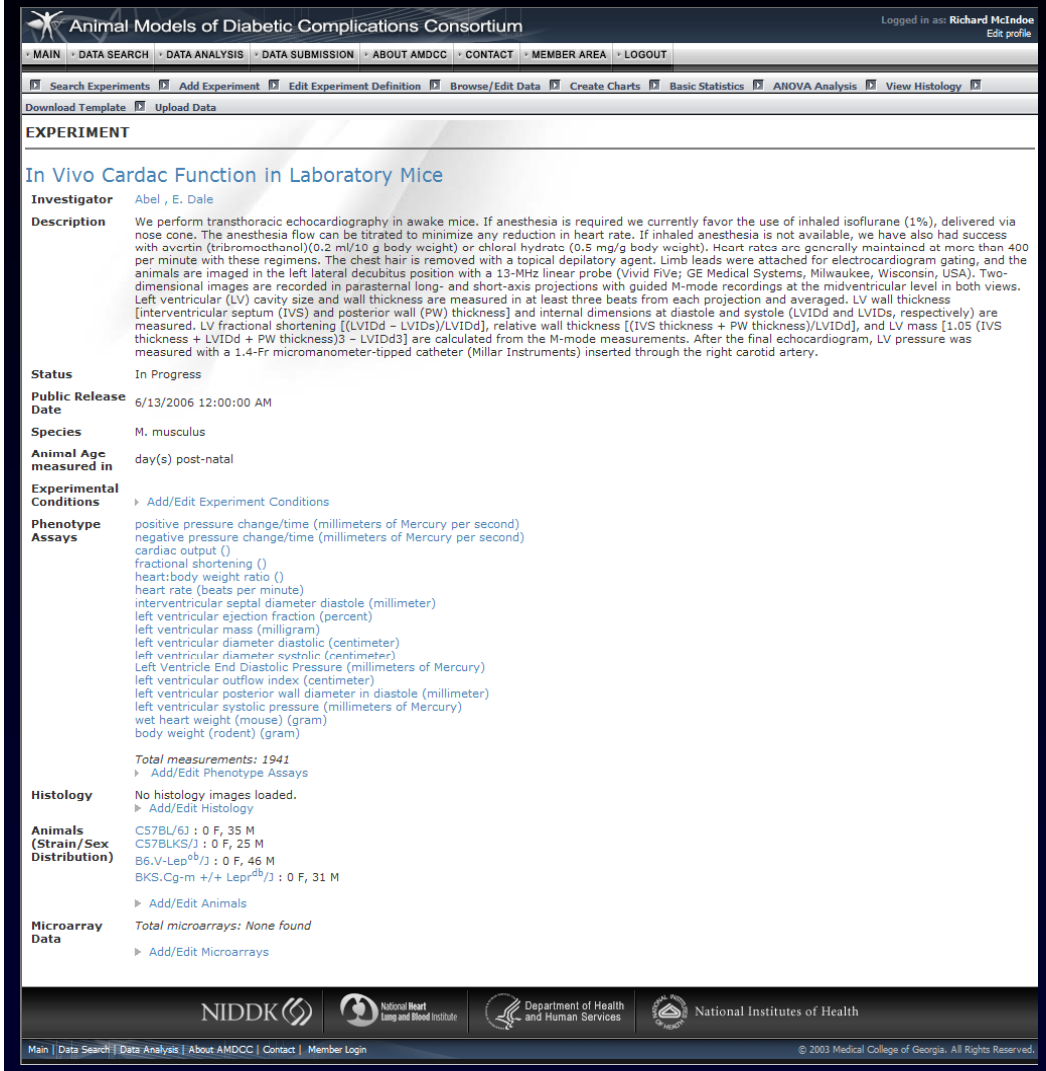
=

Final Rendered Page

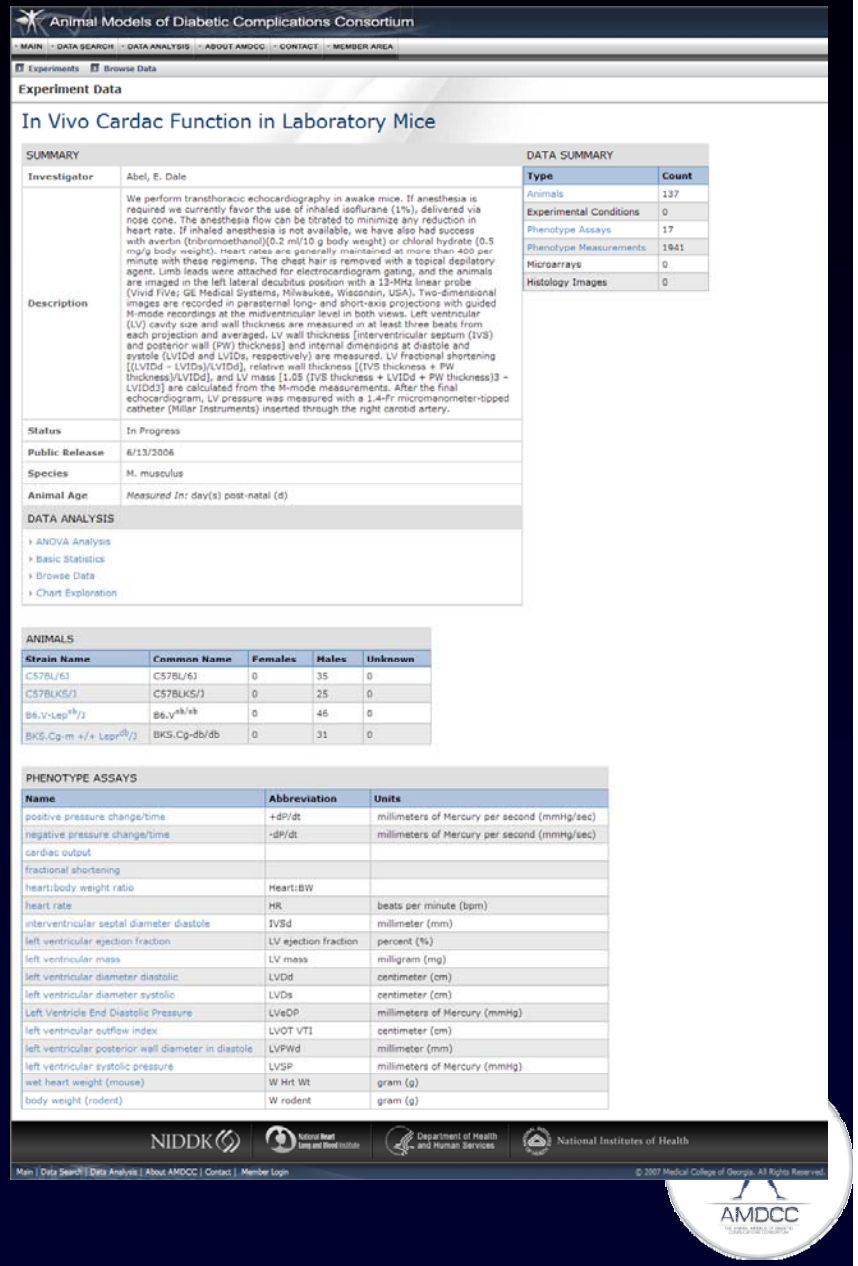
AMDCC


MMPC
National Mouse Metabolic Phenotyping Centers
Contact
About MMPC
Tests
Data Search
Data Analysis
Members
Experiment Data
Experiments
Browse Data
Effects of Decorin on Kidney function
SUMMARY
Investigator
Foo, Bar
Description
This is a brief description of the experiment. Give enough information so people understand what you are doing.
Status
In Progress
Public Release
5/28/2006
Species
M. musculus
Animal Age
Measured In: week(s) post-natal (w)
DATA SUMMARY
Type
Count
Animals
18
Experimental Conditions
1
Histology Images
5
Microarrays
3
Phenotype Assays
3
Phenotype Measurements
165
DATA ANALYSIS
ANOVA Analysis
Basic Statistics
Browse Data
Chart Exploration
ANIMALS
Strain Name
Common Name
Females
Males
129/Rl
129/Rl
1
0
C57BLKS
C57BLKS
0
5
NOD
NOD
0
4
NOD/Lt
NOD/Lt
0
4
C57BL/6J-Leprdb Apoe tm1Bre
C57BL/6J db/db
4
0
EXPERIMENTAL CONDITIONS
Name
Units
IP STZ dose
milligrams per kilogram (mg/kg)
HISTOLOGY IMAGES
Caption
Strain
Anatomical Site
Tiff 44 Picture
129/Rl
anterior apical left heart ventricle
Foo
129/Rl
anterior aortic sinus
Bar
C57BLKS
aortic root
Pretty Picture
129/Rl
anterior apical left heart ventricle
Pretty Picture
129/Rl
anterior apical left heart ventricle
MICROARRAYS
Name / Barcode
Array Design
Barcode Alpha
MG_U74Av2
Chip Barcode Delta
MG_U74Av2
Gamma Chip Barcode
MG_U74Av2
PHENOTYPE ASSAYS
Name
Abbreviation
Units
urine albumin (Exocell Albulwell M ELISA)
ualb
urine albumin/urine creatinine ratio
ualb/ucreat ratio
urine creatinine (Exocell Picric Acid method)
ucreat
Back to Top
2007 National MMPC. All Rights Reserved.
Department of Health and Human Services
National Institutes of Health
NIDDK
National Heart Lung and Blood Institute
Members
turnover using
rovascular dise
gluconeogenesis
nd measures en
ehavior and acti
se. The CBU is
NIDDK
Animal Models of Diabetic Complications Consortium
MAIN
DATA SEARCH
DATA ANALYSIS
ABOUT AMDCC
CONTACT
MEMBER AREA
Experiments
Browse Data
Experiment Data
Effects of dietary cholesterol on diabetic cardiovascular disease
SUMMARY
Investigator
McIndoe, Richard
Description
This experiment will compare the affects of dietary cholesterol on the progress of cardiovascular disease. Animals are feed a defined diet with increasing amounts of cholesterol. Angiotensinogen II is also used to modulate the response.
Status
In Progress
Public Release
9/12/2004
Species
M. musculus
Animal Age
Measured In: week(s) post-natal (w)
DATA ANALYSIS
ANOVA Analysis
Basic Statistics
Browse Data
Chart Exploration
DATA SUMMARY
Type
Count
Animals
90
Experimental Conditions
4
Histology Images
0
Microarrays
4
Phenotype Assays
32
Phenotype Measurements
2115
ANIMALS
Strain Name
Common Name
Females
Males
C57BL/6J
C57BL/6J
0
38
C57BL/6J-Leprdb
C57BL/6J-Leprdb
0
52
EXPERIMENTAL CONDITIONS
Name
Units
age at start of diet
week (wk)
Angiotensin II dose
milligrams per kilogram per minute (mg/kg/min)
mouse diet
()
time since start of diet
week (wk)
MICROARRAYS
Name / Barcode
Array Design
79031018MR1690
MG_U74Av2
EPB-MOE430A-1411n
MOE430A
EPB-MOE430A-1566
MOE430A
EPB-MOE430A-1412
MOE430A
PHENOTYPE ASSAYS
Name
Abbreviation
Units
percent of aortic surface covered by lesion
%AorticLes
percent (%)
aortic surface area section A
AorticAreaSAA
millimeters squared (mm²)
aortic surface area section B
AorticAreaSAB
millimeters squared (mm²)
aortic surface area section C
AorticAreaSAC
millimeters squared (mm²)
total lesion area on the aorta
AorticLA
millimeters squared (mm²)
lesion area aortic section A
AorticLesA
millimeters squared (mm²)
lesion surface area section b
AorticLesB
millimeters squared (mm²)
lesion area aortic section C
AorticLesC
millimeters squared (mm²)
blood glucose
milligrams per deciliter (mg/dL)
plasma free fatty acids
FFA
milligrams per deciliter (mg/dL)
high-density lipoprotein cholesterol
HDL
milligrams per deciliter (mg/dL)
heart:body weight ratio
Heart:BW
heart weight:tibia length ratio
Heart:TL
nanomoles per minute per mgdw (nmol/min/mgdw)
total kidney:body weight ratio
Kid:BW
kidney weight:tibia length ratio
Kid:TL
nanomoles per minute per mgdw (nmol/min/mgdw)
left kidney weight (mouse)
L Kid Wt
gram (g)
liver weight
Liver Wt
gram (g)
liver:body weight ratio
Liver:BW
liver weight:tibia length ratio
Liver:TL
nanomoles per minute per mgdw (nmol/min/mgdw)
plasma insulin
picograms per milliliter (pg/ml)
right kidney weight (mouse)
R Kid Wt
gram (g)
spleen weight
Spleen Wt
gram (g)
spleen:body weight ratio
Spleen:BW
spleen weight:tibia length ratio
Spleen:TL
nanomoles per minute per mgdw (nmol/min/mgdw)
total cholesterol
TC
milligrams per deciliter (mg/dL)
serum triglycerides
TG
milligrams per deciliter (mg/dL)
tibia length (mouse)
Tibia L
millimeter (mm)
total kidney weight (mouse)
Tot Kid Wt
gram (g)
total aortic surface area
TotAorticSA
millimeters squared (mm²)
plasma unesterified cholesterol
UC
milligrams per deciliter (mg/dL)
wet heart weight (mouse)
W Hrt Wt
gram (g)
body weight (rodent)
W rodent
gram (g)
NIDDK
National Heart Lung and Blood Institute
Department of Health and Human Services
National Institutes of Health
Main | Data Search | Data Analysis | About AMDCC | Contact | Member Login

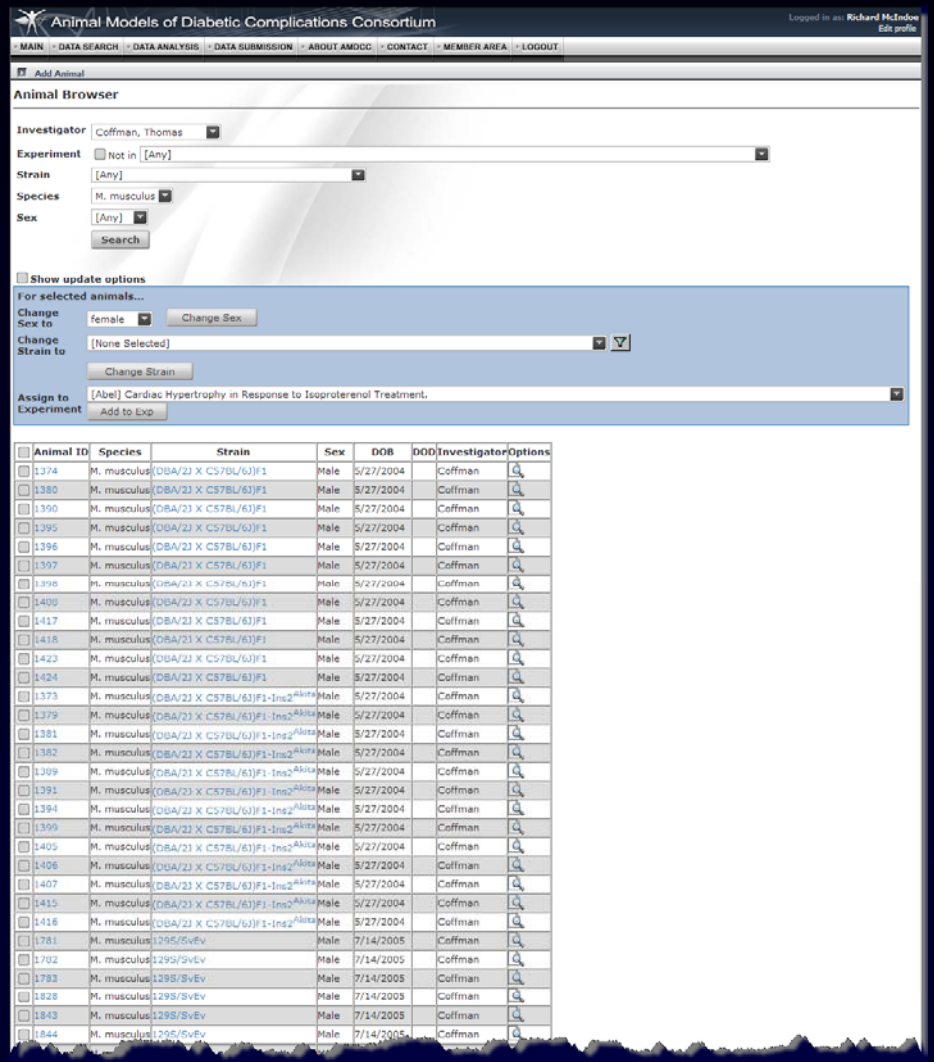
Current Experiment Page



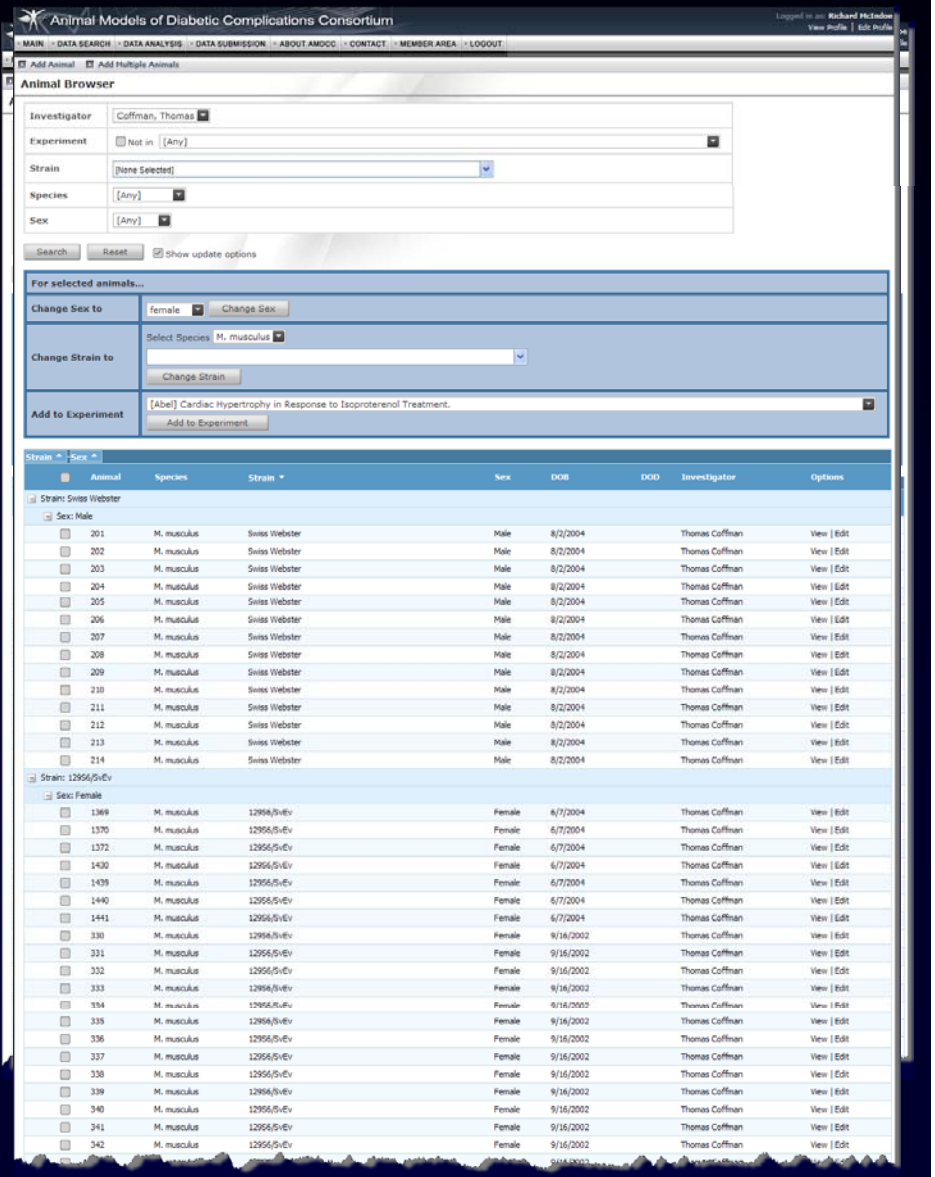
New Experiment Page



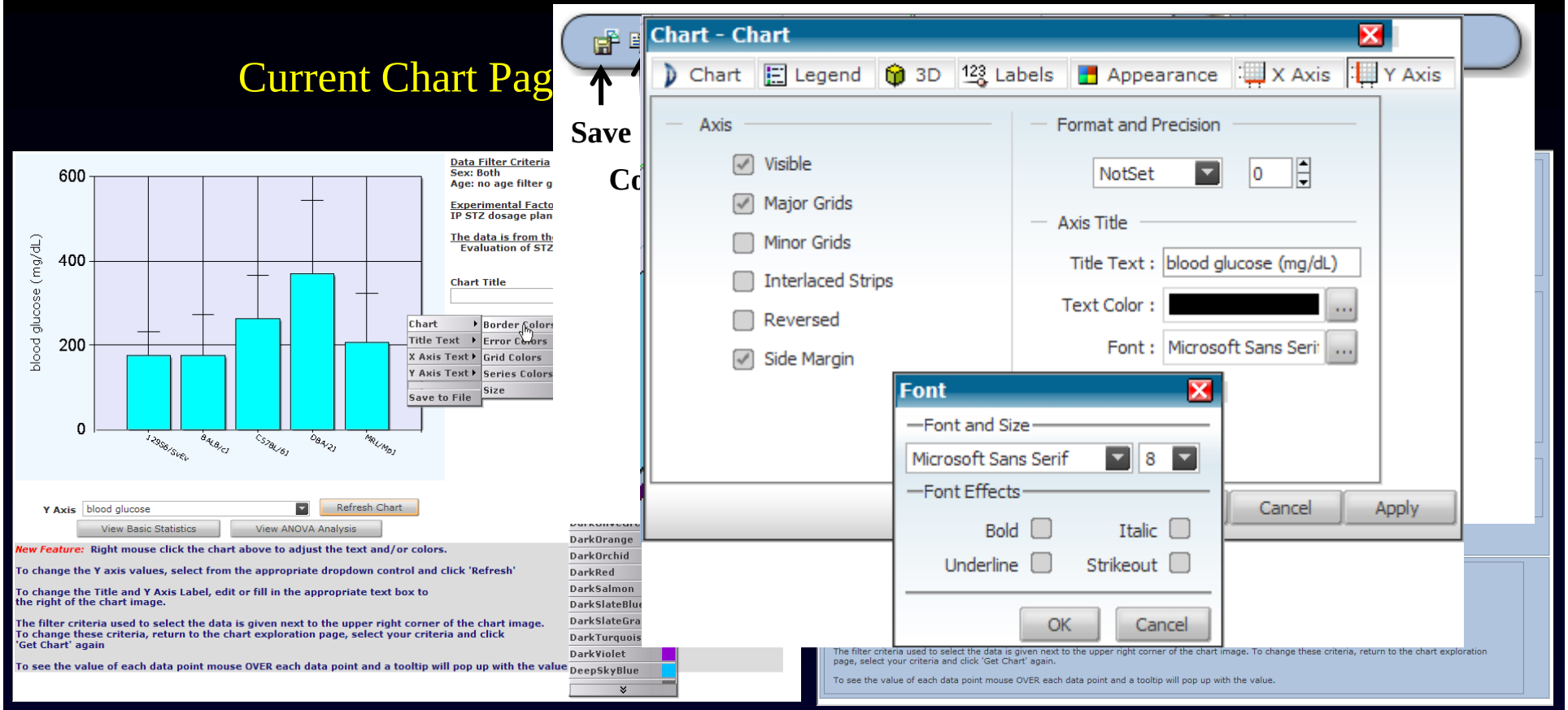
Current Animal Page



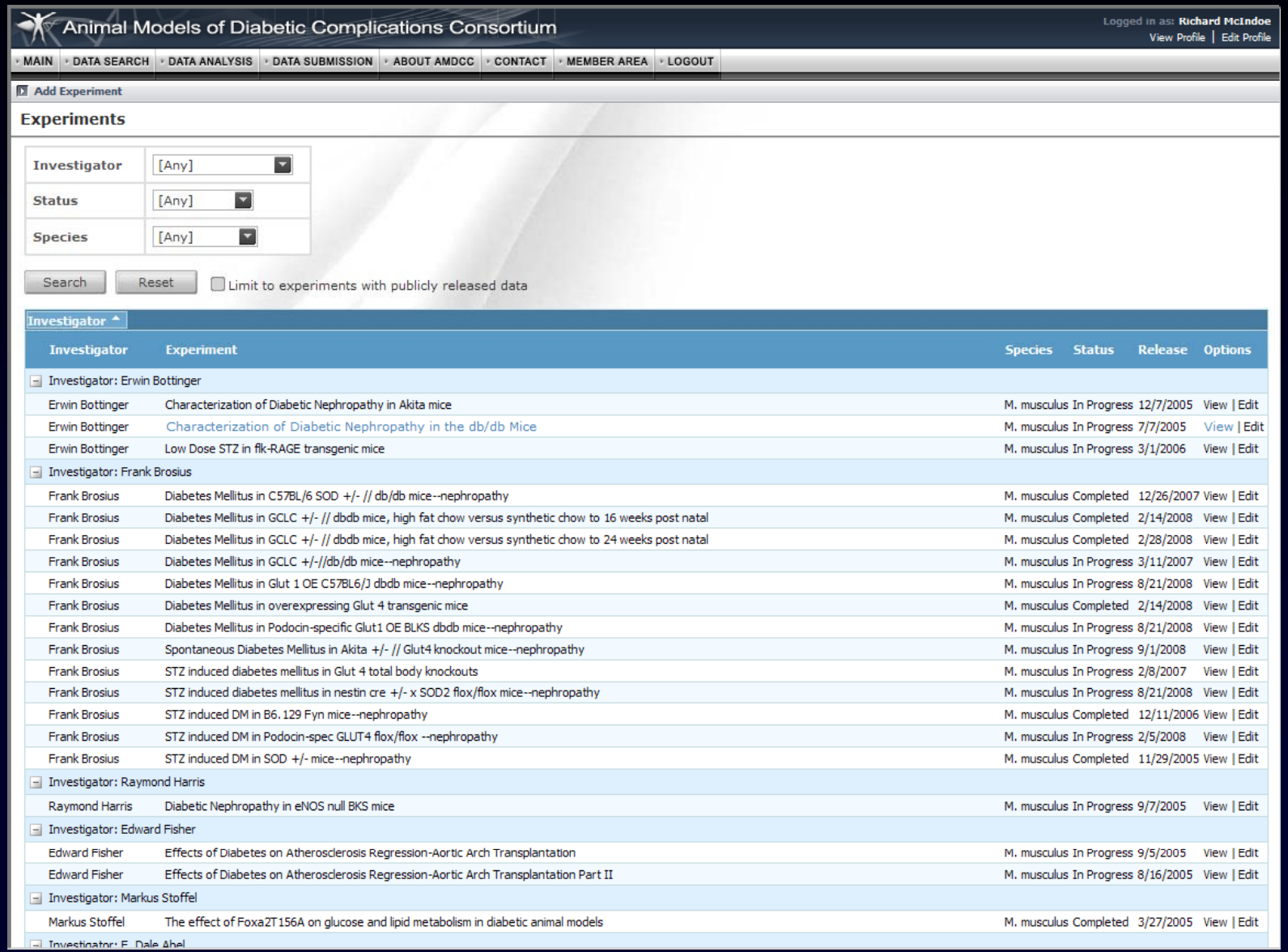
New Animal Page



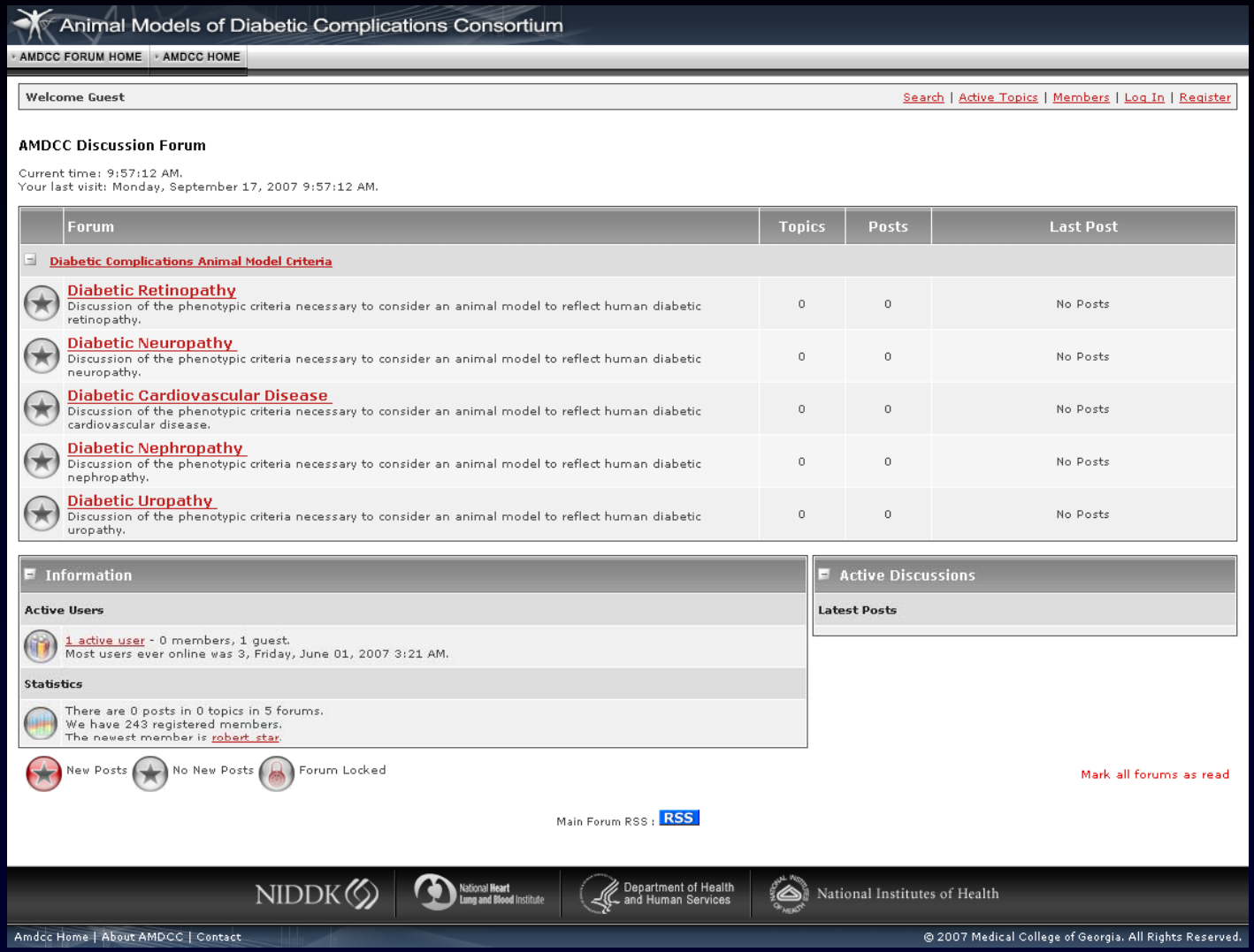
Current Chart Pag



AJAX enable the website:



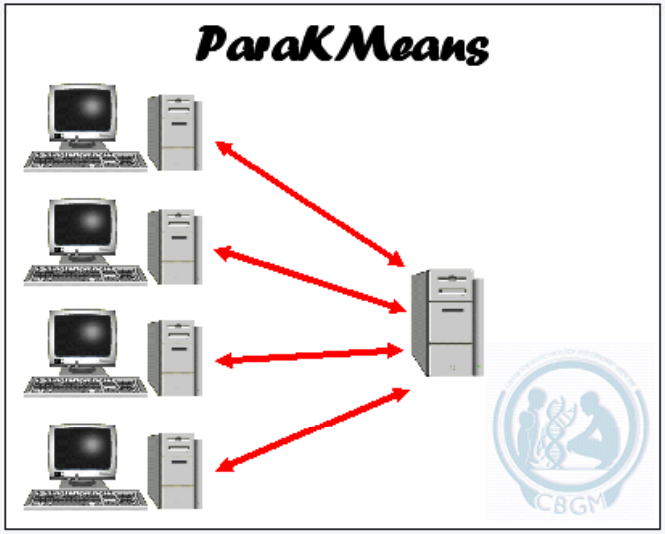
AMDCC Forums section.



ParaKMeans Web Application - A Parallel K-means Clustering Algorithm for Microarray Data

Center for Biotechnology and Genomic Medicine at the Medical College of Georgia

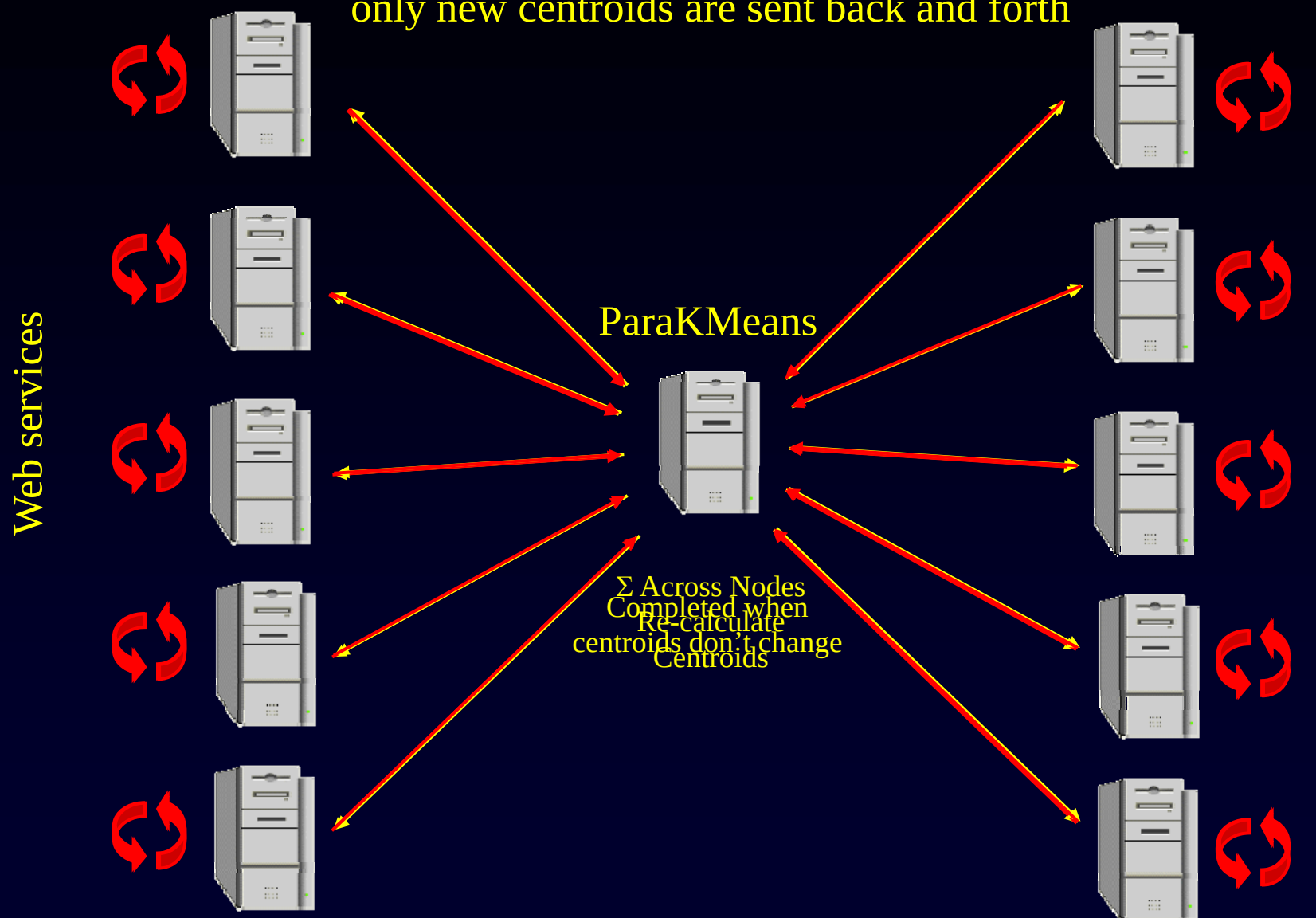
Main | Options | Manage Nodes | Output | Help



A Parallel Version of the K-Means Clustering Algorithm

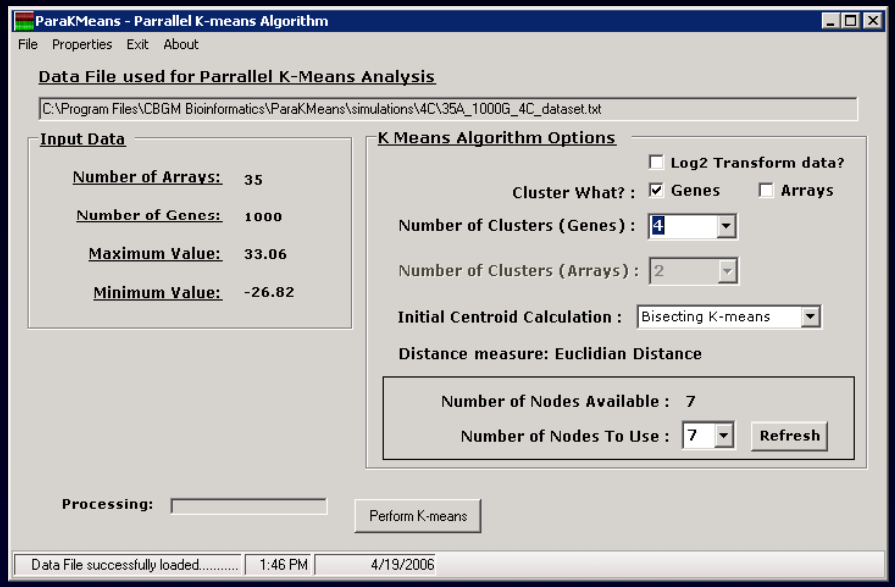
Piotr Kraj, Ph.D. and Richard A. McIndoe, Ph.D.

Data sent to each node and remains there, only new centroids are sent back and forth

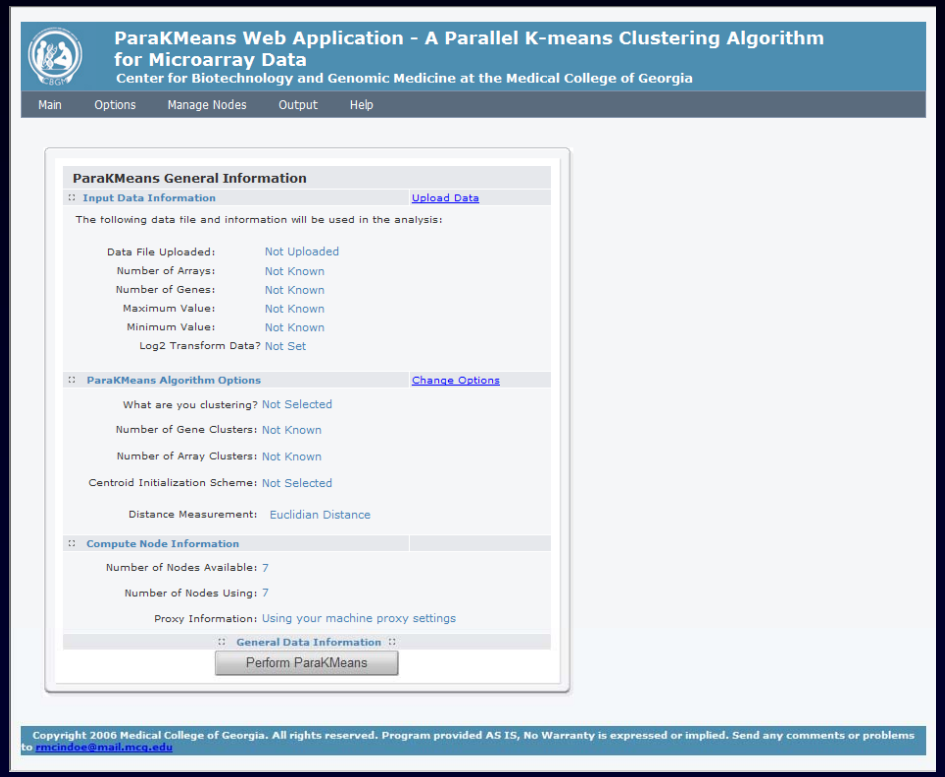


Have two user interfaces.

Windows Version

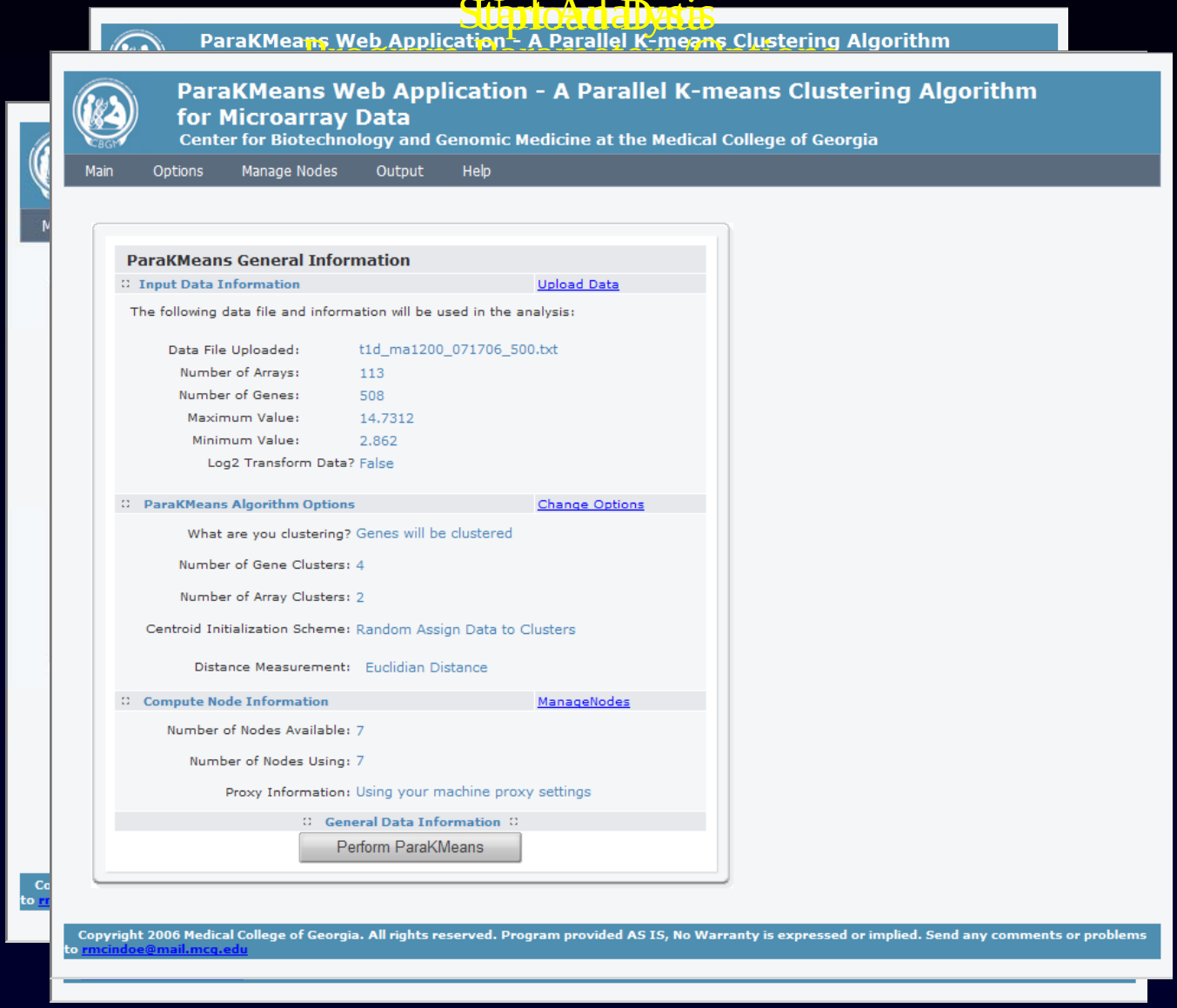


Web Version



AMDCC

Manage Nodes



Results Page

ParaKMeans Web Application - A Parallel K-means Clustering Algorithm for Microarray Data

Center for Biotechnology and Genomic Medicine at the Medical College of Georgia

Main | Options | Manage Nodes | Output | Help

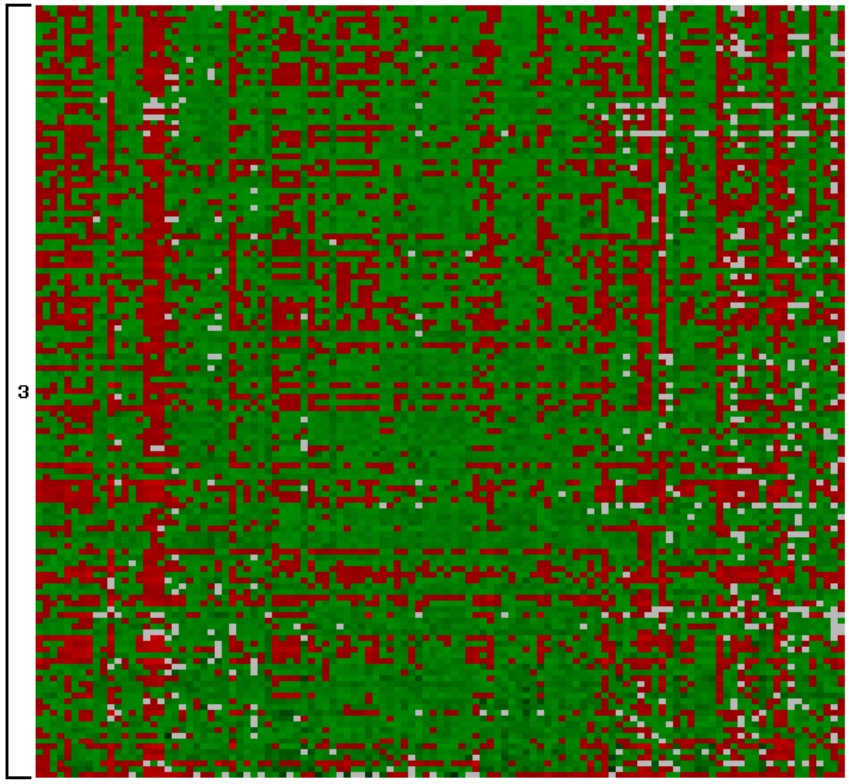
Color Scheme: Red-Green | Adjust Gain: 1 | Download **Complete** Data Result: Click Here

Gene Cluster to Show: 3 | Download Only **Cluster Assignments**: Click Here

Refresh

Execution Time: 00:00:00.9375000

2.9 — 14.7



Assess the speed, accuracy and stability of ParaKMeans.

Used simulated datasets to evaluate the program:

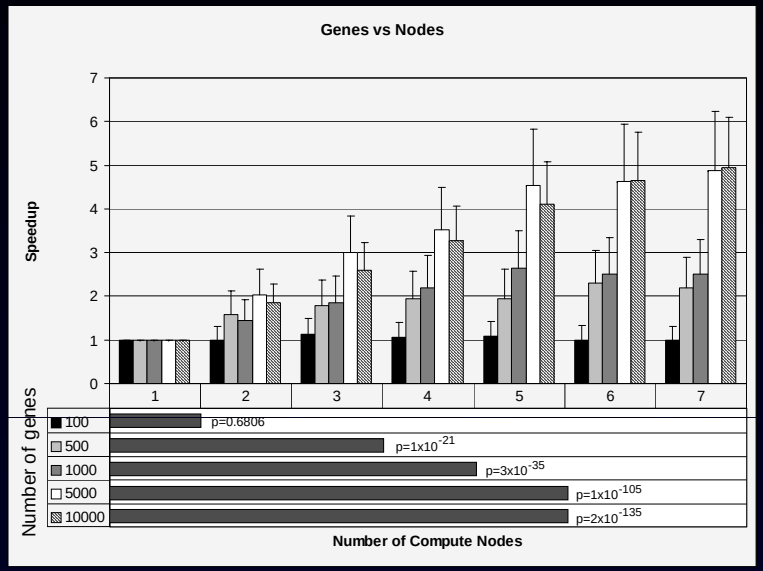
Number of arrays: 35, 100, 200
Number of genes: 100, 500, 1000, 5000, 10000
Number of cluster: 4, 10 , 20

Created data files for all combinations: 45 files total
Ran each file 12 times

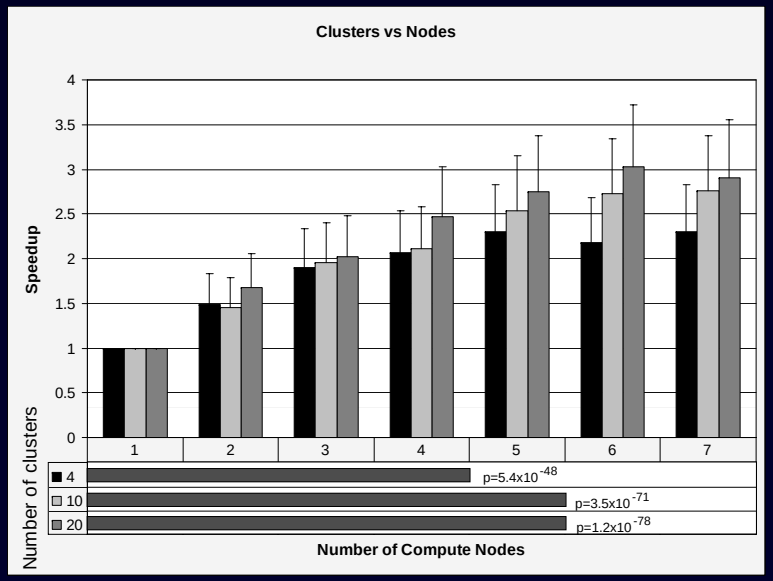
Number of nodes used: 1-7

Affect of nodes, genes, clusters on speedup

A

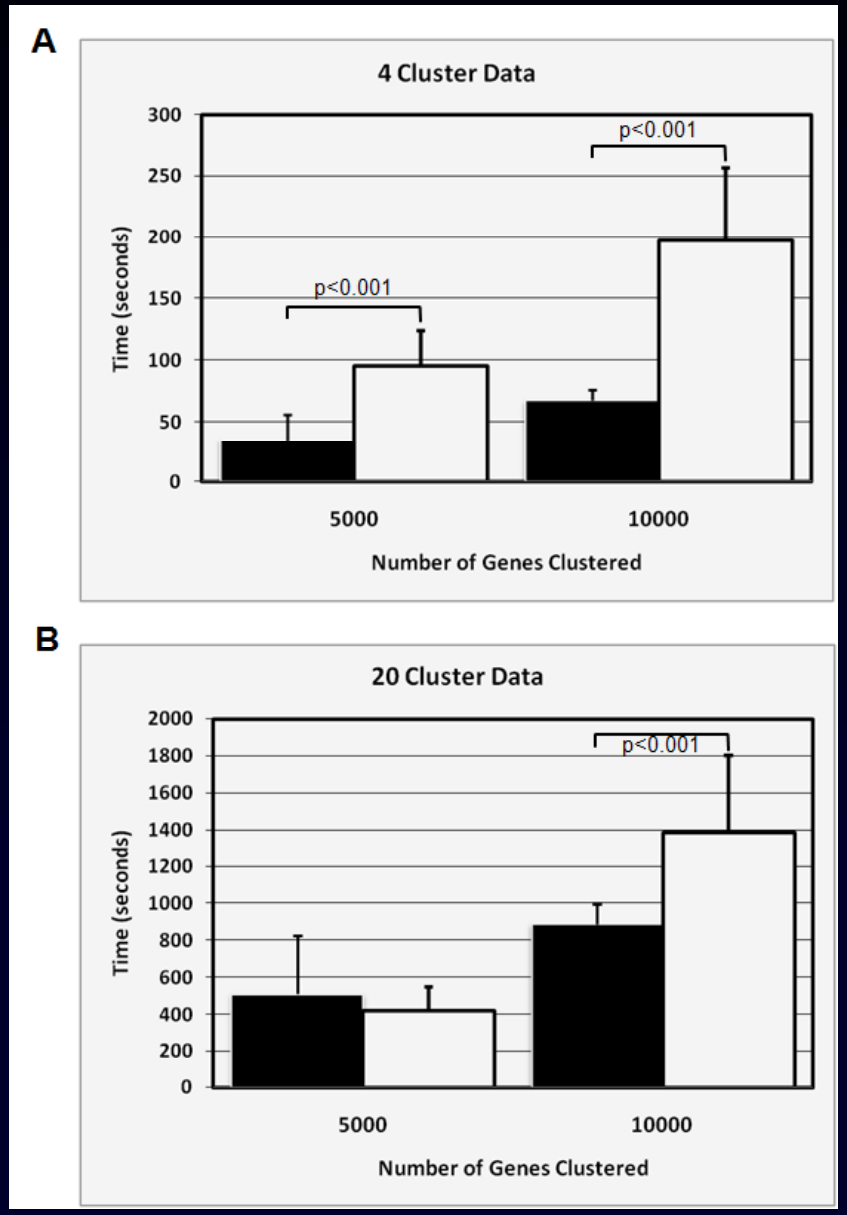


B



Comparison to Eisen's *Cluster*

1 node/RIA initialization



Accuracy and Stability

Adjusted Rand Index: statistic indicates the fraction of agreement between two cluster partitions

0.0 – 1.0 (perfect agreement)

N = 12

Table 1. Accuracy and stability results using different clustering programs, initialization schemes and number of genes/clusters.

	Clusters	5000 genes	10000 genes	Combined
Accuracy				
Cluster*	**4**	0.405 (0.404-0.405)	0.594 (0.569-0.604)	0.487 (0.404-0.604)
PKM-RIA*	**4**	**0.519 (0.453-0.597)**	**0.896 (0.896-0.896)**	**0.747 (0.453-0.896)**
PKM-RFD*	**4**	0.519 (0.322-0.519)	0.770 (0.586-0.896)	0.553 (0.322-0.896)
PKM-BKM*	**4**	0.489 (0.489-0.489)	0.770 (0.770-0.770)	0.629 (0.629-0.629)
Cluster*	**20**	0.163 (0.124-0.183)	0.256 (0.196-0.297)	0.190 (0.124-0.297)
PKM-RIA*	**20**	0.231 (0.211-0.461)	0.216 (0.208-0.233)	0.227 (0.208-0.461)
PKM-RFD*	**20**	0.189 (0.178-0.226)	0.210 (0.202-0.252)	0.202 (0.178-0.252)
PKM-BKM*	**20**	**0.400 (0.400-0.400)**	**0.691 (0.691-0.691)**	**0.545 (0.400-0.691)**
Stability				
Cluster*	**4**	0.439 (0.436-0.442)	0.797 (0.783-0.812)	0.618 (0.436-0.812)
PKM-RIA*	**4**	0.514 (0.435-1.00)	**1.00 (1.00-1.00)**	**1.00 (0.435-1.00)**
PKM-RFD*	**4**	0.569 (0.483-1.00)	0.769 (0.586-0.770)	0.711 (0.483-1.00)
PKM-BKM*	**4**	**1.00 (1.00-1.00)**	**1.00 (1.00-1.00)**	**1.00 (1.00-1.00)**
Cluster*	**20**	0.347 (0.321-0.393)	0.492 (0.418-0.907)	0.405 (0.321-0.907)
PKM-RIA*	**20**	0.738 (0.444-0.888)	0.904 (0.682-0.994)	0.788 (0.444-0.994)
PKM-RFD*	**20**	0.594 (0.548-0.643)	0.652 (0.573-0.668)	0.634 (0.548-0.668)
PKM-BKM*	**20**	**1.00 (1.00-1.00)**	**1.00 (1.00-1.00)**	**1.00 (1.00-1.00)**

* - Cluster: Single Node (N=12), All PKM: 1 and 7 node data (N=12)

All values are the median adjusted Rand Index with the range of values in parentheses.

Accuracy = cluster results vs. known assignments, Stability = over all aggreement between cluster results

PKM=ParaKMeans, RFD=Random From Data, RIA=Random Initial Assignment,

BKM=Bissecting K Means, Cluster=Eisen Cluster program.

Accuracy:

4 Cluster data: PKM-RIA > PKM-BKM > PKM-RFD > Cluster

20 Cluster data: PKM-BKM > PKM-RIA > PKM-RFD > Cluster

Stability:

4 Cluster data: PKM-BKM > PKM-RIA > PKM-RFD > Cluster

20 Cluster data: PKM-BKM > PKM-RIA > PKM-RFD > Cluster

ParaKMeans is freeware and available from the AMDCC website.
(Source code available upon request)
http://www.amdcc.org/bioinformatics/bioinformatics.aspx

We deployed the web application for the public to access our compute cluster.
http://bioanalysis.genomics.mcg.edu/parakmeans

Future Directions

1. **Work with MGHC to do a better job of integration**

2. **Provide a better assessment of animal model progress and validation criteria outcomes.**

3. **Continue to develop and update training modules for portal.**

The End

If you have any problems, questions or concerns don't hesitate to email us.
Use any of the following addresses:

General web:
amdcc_web-l@lists.amdcc.org (Listserv)

Data, assays, experiment help:
spatel@mail.mcg.edu (Sachen Patel)

Website
vdoshi@mail.mcg.edu (Vishal Doshi)