Form Approved Through 02/28/01
OMB No. 0925-0001

Department of Health and Human Services
Public Health Service

Grant Application

Follow instructions carefully.
Do not exceed character length restrictions indicated on sample.

LEAVE BLANK-FOR PHS USE ONLY.

Type	Activity	Number
Review Group	Formerly	
Council/Board *(Month, Year)*		Date Received

1. TITLE OF PROJECT *(Do not exceed 56 characters, including spaces and punctuation.)*

Informatics Center for Mouse Neurogenetics

2. RESPONSE TO SPECIFIC REQUEST FOR APPLICATIONS OR PROGRAM ANNOUNCEMENT ☐ NO ☒ YES *(If "Yes," state number and title)*

Number: PAR-99-138 Title: The Human Brain Project (Neuroinformatics): Phase I & Phase 2

3. PRINCIPAL INVESTIGATOR/PROGRAM DIRECTOR New Investigator ☐ YES

3a. NAME *(Last, first, middle)*: WILLIAMS, Robert W.

3b. DEGREE(S): Ph.D.

3c. SOCIAL SECURITY NO. ***Provide on Form Page KK.***

3d. POSITION TITLE: Professor

3e. MAILING ADDRESS *(Street, city, state, zip code)*

University of Tennessee, Memphis
Department of Anatomy and Neurobiology
855 Monroe Avenue
Memphis TN 38163

3f. DEPARTMENT, SERVICE, LABORATORY, OR EQUIVALENT: Department of Anatomy and Neurobiology

3g. MAJOR SUBDIVISION: College of Medicine

3h. TELEPHONE AND FAX *(Area code, number and extension)*

TEL: (901) 448-7018
FAX: (901) 448-7193

E-MAIL ADDRESS: rwilliam@nb.utmem.edu

4. HUMAN SUBJECTS ☒ No ☐ Yes
4a. If "Yes," Exemption no. **or** IRB approval date { ☐ Full IRB **or** ☐ Expedited Review
4b. Assurance of compliance No.

5. VERTEBRATE ANIMALS ☐ No ☒ Yes
5a. If "Yes," IACUC approval date: 07/09/99
5b. Animal welfare assurance no.: A-3325-01

6. DATES OF PROPOSED PERIOD OF SUPPORT *(month, day, year-MM/DD/YY)*		7. COSTS REQUESTED FOR INITIAL BUDGET PERIOD		8. COSTS REQUESTED FOR PROPOSED PERIOD OF SUPPORT	
From	Through	7a. Direct Costs ($)	7b. Total Costs ($)	8a. Direct Costs ($)	8b. Total Costs ($)
07/01/00	06/30/05	1,134,107	1,259,729	5,438,761	5,969,973

9. APPLICANT ORGANIZATION
Name: The University of Tennessee, Memphis
Address: 800 Madison Avenue
Memphis, TN 38163

10. TYPE OF ORGANIZATION
Public: → ☐ Federal ☒ State ☐ Local
Private: → ☐ Private Nonprofit
Forprofit: → ☐ General ☐ Small Business

11. ORGANIZATIONAL COMPONENT CODE: 01

12. ENTITY IDENTIFICATION NUMBER: 1-626001636B3
DUNS NO. *(if available)*

Congressional District: 9th

13. ADMINISTRATIVE OFFICIAL TO BE NOTIFIED IF AWARD IS MADE
Name: Robert L. Blackwell
Title: Vice Chancellor for Business and Finance
Address: The University of Tennessee, Memphis
62 Dunlap St., Suite 300
Memphis, TN 38163

Telephone: (901) 448-5523
FAX: (901) 448-6517
E-Mail: gbussell@utmem.edu

14. OFFICIAL SIGNING FOR APPLICANT ORGANIZATION
Name: Michael Dockter
Title: Associate Dean for Research
Address: The University of Tennessee, Memphis
8 South Dunlap Street, C109
Memphis, TN 38163

Phone: (901) 448-5587
FAX: (901) 448-7600
E-Mail: dsmith@utmem.edu

15. PRINCIPAL INVESTIGATOR/PROGRAM DIRECTOR ASSURANCE:
I certify that the statements herein are true, complete and accurate to the best of my knowledge. I am aware that any false, fictitious, or fraudulent statements or claims may subject me to criminal, civil, or administrative penalties. I agree to accept responsibility for the scientific conduct of the project and to provide the required progress reports if a grant is awarded as a result of this application.

SIGNATURE OF PI/PD NAMED IN 3a. *(In ink. "Per" signature not acceptable.)*

DATE: 02/07/01

16. APPLICANT ORGANIZATION CERTIFICATION AND ACCEPTANCE:
I certify that the statements herein are true, complete and accurate to the best of my knowledge, and accept the obligation to comply with Public Health Service terms and conditions if a grant is awarded as a result of this application. I am aware that any false, fictitious, or fraudulent statements or claims may subject me to criminal, civil, or administrative penalties.

SIGNATURE OF OFFICIAL NAMED IN 14. *(In ink. "Per" signature not acceptable.)*

DATE: 02/07/01

Introduction to the Application

The scientific motivation for this program project is to investigate the complex network of molecules and mechanisms that modulate the structure of different parts of the brain—from major subdivisions such as the hippocampus, thalamus, and cerebellum, down to the level of discrete subpopulations of neurons and glial cells in single nuclei. We plan to begin this work by producing the analytic tools, tissue resources, and genotypes that are essential for the systematic exploration of the complex genetics of mammalian brain architecture.

The bioinformatic tools and resources that we have been creating have very wide applicability, but our program project is specifically directed at exploiting a powerful new reductionist approach to explore the genetic basis of the very significant structural differences between the brains of different strains of mice. This approach, called either complex trait analysis or QTL analysis, developed rapidly in the late 1990s as a result of the hybridization of quantitative genetics and molecular genetics (Lander and Schork 1994). The suite of techniques associated with this approach has greatly extended the variety of CNS phenotypes that can be subjected to systematic molecular analysis. In neuroscience, complex trait analysis has been embraced by behavioral geneticists and neuropharmacologists (Plomin et al. 1991; Johnson et al. 1992; Crabbe et al. 1994; Takahashi et al. 1994; Kanes et al. 1996; Roubertoux et al. 1998). The time is ripe to apply these techniques to the genetic sources of structural differences in the brain.

Bioinformatic Tools and Resources

Taking advantage of rapid advances in both computer technology and molecular genetics, we will assemble an innovative multidisciplinary system that will greatly facilitate experimental analysis of the CNS. We will use this system to investigate the genetic basis of normal variation in mouse CNS. Other neuroscientists and geneticists will also benefit from this project, which will make a huge collection of histological slides available for data collection and stereological analysis over the Internet.

Project 1: The Mouse Brain Library. This integrated, multi-stain, multi-plane digital image resource for neuroscientists is housed at <nervenet.org/mbl/mbl.html>. The MBL now contains sections from over 600 brains imaged at 25 μm/pixel resolution and a subset of 120 brains imaged at 4.5 μm/pixel resolution. We intend to add nearly 2000 brains over the duration of the grant. Dr. Glenn Rosen, the PI of this project, will create relational databases to catalog slides, sections, images, and digital video clips. These databases will become crucial as the MBL collection grows more than 100-fold in volume—from 2 gigabytes (GB) to well over 200 GB. The databases will employ a coordinate reference system that will increase the utility of the collection for morphometric analysis in connection with Projects 2 and 3. The MBL is in a real sense the headwater of this project; it provides the major phenotypic resource on which all the other projects rely heavily.

Project 2: Internet Microscopy Systems for high-resolution imaging. We have developed an Internet-controlled microscope called the iScope to be used in conjunction with the MBL. This system acquires color images at a resolution of 0.1 μm/pixel. Our goal in this Project is to provide fast access to the entire slide collection using streaming video technology, enabling neuroscientists to acquire high-magnification images of any CNS region for any of the 2000 genetically defined mice in the MBL. Collaborative stereology across the Internet will become feasible.

Project 3: NeuroCartographer for segmentation of the MBL. Projects 1 and 2 will generate massive collections of images and through-focus series of the brains of several thousand mice. The main objective of Project 3 is to develop extremely efficient ways to extract numbers from a vast collection of digital micrographs and video clips. This ambitious project has begun with the construction of 3D atlases of the mouse brain that will be segmented into hundreds of nuclei and fiber tracts. The project will produce as its output hundreds of quantitative morphometric parameters that will then be analyzed as part of Project 4. Another major function of the NeuroCartographer project is to provide sophisticated navigational tools for use with the MBL and Internet microscopes.

Project 4: The Mouse Neurogenetics Tool Box. We will develop unique World Wide Web services that make it possible for neuroscientist to rapidly identify and map genes and quantitative trait loci, particularly those related to brain structure and, ultimately, to behavior. The Neurogenetics Tool Box will be tightly integrated with the MBL and NeuroCartographer Projects. Synergy between these

components of the program project will allow neuroscientists and geneticists to explore the complex genetics of CNS architecture.

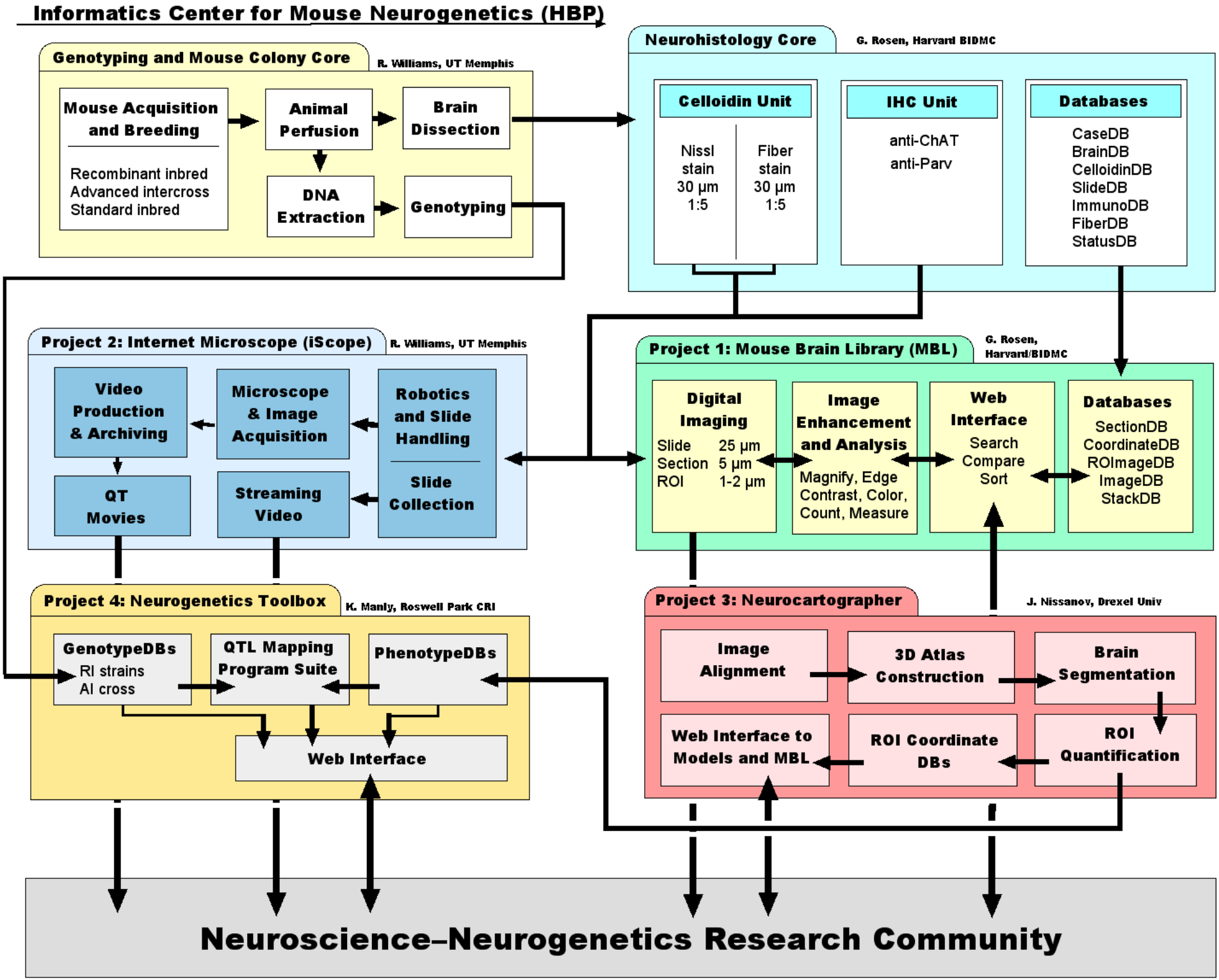
Our project will accomplish three important goals that will encourage further work on the genetics of mouse and human CNS. We expect these projects to illuminate how even seemingly small quantitative variation can catalyze research on the complex network of molecules that control brain size and cell composition.

1. **Universal Access to Neuroanatomical Phenotypes.** By imaging brain sections of several thousand mice at low, intermediate, and high magnification and then putting these images on the web, we will provide universal access to CNS neuroanatomical phenotypes that can be used for systematic quantitative analysis. As part of a pilot project, we have already put images of many brains into the MBL. Even neglecting for a moment our genetic focus, this is an amazing resource with which to explore sex differences, effects of age on the brain (age range is from 25 to 650 days), and differences between strains. We intend to triple the size of this collection over the next five years. Besides adding more genetically defined mice, we will add more depth to the collection. Our image database will span the entire optical range: from a resolution of 25 μm/pixel for entire slides, to 4.5 μm/pixel for individual sections, to 1 μm for specific regions of interest such as the hypothalamus and amygdala, down to 0.1 μm/pixel for high-magnification through-focus series (50 x 50 μm fields) used for unbiased stereological analysis of cell number and cell size. We will also add fiber-stained sections and anti-ChAT and anti-parvalbumin-labeled sections to begin the inevitable process of incorporating new histological and cytochemical dimensions in the MBL collection.
2. **Digital Dissection.** In the NeuroCartographer project, we will develop efficient and accurate methods to digitally dissect the massive MBL database of images and through-focus series. We will initially generate well-aligned images of sections from each brain in the MBL. The aligned set of images will then be semi-automatically segmented into several hundred anatomically defined compartments. Algorithms for segmentation have been developed at a rapid pace over the past decade, and the enormous advances in computer hardware have been a boon to this research. Until we initiated our large-scale bioinformatics project, however, there was no compelling reason to perfect these methods so that hundreds or thousands of brains could be digitally dissected. Now we have the motivation to push segmentation into high gear. Each region of interest (ROI) that we can define using segmentation programs will be associated with a full set of quantitative measures (volume, surface areas, shaped indices, position, and orientation).
3. **Comprehensive Genotype Data.** We will supply comprehensive genotype data for an unusually large mapping cross consisting of more than 1400 young and old mice. We will also generate and update databases on genotypes and phenotypes of a large number of recombinant inbred (RI) strains of mice. Our choice of strains and crosses has been thought out carefully with the aim of maximizing the precision and statistical power of QTL analysis. The large BXD set (35 extant strains) will be complemented by a tenth-generation advanced intercross between C57BL/6J and DBA/2J (the G_{10}). Using powerful statistical programs integrated in the Neurogenetics Tool Box (NTB), neuroscientists will be able to exploit these morphometric resources without processing either tissue or DNA. The extensive image resources in the MBL and the matched morphometric databases associated with the NeuroCartographer project will make it comparatively easy to map hundreds of QTLs that specifically modulate the size and cellular architecture of the mouse CNS. Each of these QTLs will be mapped with unprecedented precision because our major cross is unusually large (1400 animals) and because the multigeneration design provides approximately five times the resolution of a corresponding two-generation cross.

Personnel

This is a unique multidisciplinary collaboration between neuroscientists and geneticists, all of whom have strong backgrounds in bioinformatics and computer applications development. There is tremendous leverage in this collaboration, which has already produced impressive results.

Robert W. Williams, UT Memphis, will serve as Program Director, PI of Project 2 (the iScope), and director of the Administrative Core and the Genotyping and Mouse Colony Core. Dr. Williams, a developmental neurobiologist and mouse geneticist, has developed and maintained extensive Internet resources both for gene mapping and for the analysis of the mouse CNS and eye.


Informatics Center for Mouse Neurogenetics (HBP)
Genotyping and Mouse Colony Core
R. Williams, UT Memphis
Mouse Acquisition and Breeding
Recombinant inbred
Advanced intercross
Standard inbred
Animal Perfusion
Brain Dissection
DNA Extraction
Genotyping
Neurohistology Core
G. Rosen, Harvard BIDMC
Celloidin Unit
Nissl stain 30 µm 1:5
Fiber stain 30 µm 1:5
IHC Unit
anti-ChAT
anti-Parv
Databases
CaseDB
BrainDB
CelloidinDB
SlideDB
ImmunoDB
FiberDB
StatusDB
Project 2: Internet Microscope (iScope)
R. Williams, UT Memphis
Video Production & Archiving
Microscope & Image Acquisition
Robotics and Slide Handling
QT Movies
Streaming Video
Slide Collection
Project 1: Mouse Brain Library (MBL)
G. Rosen, Harvard/BIDMC
Digital Imaging
Slide 25 µm
Section 5 µm
ROI 1-2 µm
Image Enhancement and Analysis
Magnify, Edge
Contrast, Color,
Count, Measure
Web Interface
Search
Compare
Sort
Databases
SectionDB
CoordinateDB
ROImageDB
ImageDB
StackDB
Project 4: Neurogenetics Toolbox
K. Manly, Roswell Park CRI
GenotypeDBs
RI strains
AI cross
QTL Mapping Program Suite
PhenotypeDBs
Web Interface
Project 3: Neurocartographer
J. Nissanov, Drexel Univ
Image Alignment
3D Atlas Construction
Brain Segmentation
Web Interface to Models and MBL
ROI Coordinate DBs
ROI Quantification
Neuroscience–Neurogenetics Research Community

Glenn D. Rosen, Associate Professor of Neurology (Neuroscience) at Beth Israel Deaconess Medical Center and Harvard Medical School, will serve as PI of Project 1 (MBL) and director of the Neurohistology Core. Dr. Rosen is an expert stereologist and neuroanatomist, and his strong background in statistics and morphometry makes him particularly competent to develop the MBL.

Jonathan Nissanov, co-director of the Drexel University Center for Vertebrate Brain Mapping, will serve as PI of Project 3 (NeuroCartographer) and will also contribute to Project 2 (the iScope). Dr. Nissanov has a strong background in neuroscience, engineering, and bioinformatics. He has expertise in developing image analysis tools and helped to design, assemble, and test a robotic slide-handling system.

Kenneth Manly at the Roswell Park Cancer Institute will serve as PI of Project 4 (the Neurogenetics Tool Box). Dr. Manly is a well-known geneticist and the author of a set of powerful software programs for gene mapping (Map Manager, Map Manager QT, and Map Manager QTX).

Dan Goldowitz, Professor of Neurobiology at the University of Tennessee, Memphis, is a coinvestigator on Projects 1 and 2. Dr. Goldowitz, an expert neurogeneticist, is closely involved in the NIH mouse mutagenesis initiative and is the principal investigator of the Tennessee Mouse Genomics Consortium mutagenesis RFA application.

Melburn Park, Associate Professor of Neurobiology at UT Memphis, will be a coinvestigator for Project 2 (the iScope). Dr. Park is both a neuroanatomist and an accomplished programmer. He has written numerous large software programs for real-time data acquisition, iterative voltage clamp analysis, three-dimensional neuron reconstruction, text processing, mapping, and image processing using powerful deconvolution methods.

Oleh J. Tretiak has been Director of the Computer Vision Center for Vertebrate Brain Mapping at Drexel University, an NIH Biomedical Technology Resource, for the past 15 years. He has a strong background in engineering with a specific focus in computer vision and its application to aligning and segmenting histological material. Much of his research effort has dealt directly with the brain segmentation problem as it applies to rodent brains.

Investigating genetic variation

Normal variation can be impressive. Numbers of neurons in the human neocortex vary from 15 billion to 32 billion (Pakkenberg and Gundersen 1997). The volume and cell number of the human visual cortex vary three-fold, as does the density of cones in the foveal pit (Gilissen and Zilles 1996; Curcio et al. 1990). Numbers of ocular dominance columns within the primary visual cortex of rhesus monkeys vary more than 50% (Horton and Hocking 1996). These robust differences are caused not by mutations but by the cumulative action of many normally variable genes and by the action of numerous developmental and environmental factors. In the long run, normal genetic polymorphisms are the most critical source of variance: they are the substrate for evolutionary and developmental modification of brain size and cellular architecture (Williams and Herrup 1988; Lipp 1989; Williams et al. 1993).

In the late 1960s, Thomas Roderick, John Fuller, Douglas Wahlsten, and Richard and Cynthia Wimer began an ambitious program to manipulate neuroanatomical traits in mice by selective breeding (Roderick 1979). Their aim was to explore correlated changes in behavior. They gave the rapidly expanding field of behavioral neurogenetics a rigorous foundation in quantitative and statistical neuroanatomy (Wimer et al. 1969; Fuller and Geils 1972; Wahlsten 1975; Roderick et al. 1976; Fuller 1979; Wimer 1979; Wimer and Wimer 1985). Rather than relying on mutants, they exploited the substantial variation among standard inbred strains of mice. This work led to some important breakthroughs and some brick walls. One of the breakthroughs was successfully selecting for substantial differences in brain weight over less than 20 generations (Fuller 1979). An obvious limitation, highlighted by Roderick (1976), was that it was not possible to map gene loci responsible for the remarkable quantitative variation in CNS size, regional architecture, or behavior.

The research tools available to neuroscientists have changed radically in the past decade (Lander and Botstein 1989; Plomin et al. 1991; Belknap et al. 1992; Johnson et al. 1992; Tanksley 1993; Frankel 1995). Computational methods and molecular reagents—particularly the polymerase chain reaction (PCR) method—have become so powerful and economical that it is now practical to systematically dissect complex polygenic traits such as brain weight into sets of single well-defined QTLs. Virtually any heritable trait in mice,

whether structural, physiological, pharmacological, or behavioral, can be targeted for analysis. Recent examples in mice include epilepsy (Rise et al. 1991), effects of ethanol and haloperidol (Belknap et al. 1993; Plomin and McClearn 1993; Hitzemann et al. 1994; Kanes et al. 1996; Buck et al. 1997), patterns of sleep and activity (Toth and Williams 1998), and the mouse equivalent of anxiety (Flint et al. 1995). As illustrated in the work of Belknap and colleagues (1992), it is now feasible to continue the systematic genetic dissection of the mouse CNS begun in the late 1960s and to start identifying genes that underlie heritable variation in CNS size and structure.

QTLs are conventional genes that have two or more alleles that contribute to quantitative variation of specific traits (Roff 1997; Lynch and Walsh 1998). A trait may be a concentration or number; a size, weight, or density; an activity or behavior; a severity index; or an age of onset. QTLs are often contrasted with Mendelian loci that have discontinuous effects on phenotypes and predictable segregation patterns. Individual QTLs usually have more modest effects on a particular phenotype and are associated with phenotypes in a probabilistic way. A QTL might account for as little as 2% or as much as 50% of the total variance of a phenotype.

The clinical potential of QTL mapping in mice. The analysis of complex polygenic traits is an important facet of human genetics. Any recent issue of the *Journal of the American Society of Genetics* will include 5 to 10 reports on gene loci affecting disease susceptibility, onset, or severity in different human populations. These “susceptibility” genes are in fact QTLs. The genetics of alcoholism, schizophrenia, depression, heart disease, glaucoma, myopia, several dementias, breast cancer, colon cancer, and autoimmune disease all are benefiting enormously from the application of these methods.

Systematic histology of the huge human brain in a genetic context is daunting, to say the least, but in mice, the brain is the size of a plump pistachio. A 1-in-10 series of sections through the entire brain fits on a single 50 x 75 mm slide. Because so many diseases of the CNS have complex genetic and environmental etiologies (Parkinson's, schizophrenia, depression, Alzheimer’s, macular degeneration, etc.) it is vital to build up resources for a systematic genetic assault on complex CNS traits. We can investigate the genetic basis of many traits in the mouse, then look for corresponding features in humans.

We already know that morphometric CNS traits are often highly variable and highly heritable. For example, brain weight in the DBA/2J strain of mice is about 420 mg, and in C57BL/6J it is about 500 mg. Morphometric traits tend to have relatively high heritabilities compared to behavioral and fitness traits (Roff, 1997). Most heritabilities that we have computed (either broad or narrow sense) have been well above 0.3, with an average of about 0.5.

QTLs can be mapped routinely. Information on our set of over 5000 mice and over 100 strains is available online at <nervenet.org>. We have used this large data set to test the feasibility of mapping brain weight QTLs. The results have been extremely successful, and we have now mapped several QTLs that specifically modulate brain weight (see Appendix material). We have recently gone one step further to show that it is also entirely practical to map QTLs that affect specific subdivisions of the CNS. We have mapped QTLs that modulate the weight of the cerebellum (Airey et al. 1998, 1999) and one QTL that has an especially marked effect on the size of the hippocampus (Lu et al. 1999).

Probability of success in QTL mapping studies. What is the probability of successfully mapping one or more QTLs? For CNS traits that have heritabilities above 50%, it will usually be possible to map several QTLs in even a small cross consisting of 200 to 400 individuals. Behavioral traits such as open-field activity tend to have relatively low heritabilities (<30%), yet they have been successfully dissected into sets of QTLs (Flint et al. 1995). For example, Le Roy, Roubertoux, and colleagues (Le Roy et al. 1999) have successfully mapped more than a dozen QTLs that modulate the development of several behavioral traits in preweanling mice.

The main constraint in this work is the number of animals that can be phenotyped and genotyped. We have mapped four QTLs affecting retinal ganglion cell number (Williams et al. 1998a; Strom 1999), eye weight (Zhou and Williams 1999b; Williams and Zhou 1999), brain weight (Strom and Williams 1997; Strom 1999), and cerebellar weight (Gilissen and Williams 1997; Airey et al. 1998). In each case, approximately 150 animals were phenotyped per QTL. If multiple polymorphic loci are clustered near one another, then a small intercross may detect a single poorly localized QTL, in essence a polygenic QTL. In contrast, a large high-resolution cross, especially an advanced intercross of the type that will be a major resource in this program project, will separate the individual loci (Darvasi 1998; Williams 1998).

Although it is satisfying to decompose variation in brain weight into sets of individual QTLs, we still need to determine what parts of the CNS and what cell populations are most and least affected. For example, is there a subset of QTLs that specifically modulates the size of the cerebellum or the proliferation of granule cells? As part of this program project, our whole group is assembling the resources needed to carry out systematic stereological studies of individual nuclei and cell types. The main impediment for an individual investigator is the massive effort needed to section, stain, and count particular nuclei or regions in hundreds of cases. If a single cross is shared among many neuroscientists and then used to analyze many different CNS structures, then the effort required to map each QTL is reduced substantially. This, of course, is an important benefit of the program project: other investigators will be able to use the resources we develop to map multiple QTLs, yet these other groups will not have to process any tissue or genotype any animals.

Cloning QTLs. Mapping QTLs is the reconnaissance stage in a systematic effort to explore mechanisms that modulate the development of the CNS. The next step is matching each QTL with a single gene and its alternative alleles. QTLs will generally need to be mapped with a precision of 1 to 2 cM, a chromosomal interval that typically harbors 50–100 genes. Achieving this level of accuracy is not impractical, although it will often require an analysis of 1000 or more animals (Darvasi 1997, 1998). A small subset of positional candidate genes can then be chosen for further analysis on the basis of expression patterns, known function, and differences in DNA sequence among strains. The efficiency of the candidate gene approach will improve greatly in the next decade. The genome of C57BL/6J will have been sequenced within several years, and it is also likely that the utility of this code will be enhanced with sequence data from other major inbred strains such as 129, A, BALB/c, C3H, DBA/2, CAST/Ei, and SPRET/Ei. Once sequence data have been combined with expression maps for different parts of the mouse brain, candidate genes can be winnowed to a very short list.

Even before it has been cloned, a QTL can be used to study mechanisms of brain development or function. For example, we wanted to determine whether *Nnc1* modulates neurogenesis or cell death. To answer this question, we counted the cell population after neurogenesis but before the onset of cell death. We showed convincingly that the bimodality is produced by a fundamental difference in the total production of retinal ganglion cells (Strom and Williams 1998). *Nnc1* must influence the proliferation of retinal ganglion cells, for example, through variation in the progenitor pool size, pathways of cell differentiation, or cell cycle parameters. The effect is robust and different strains carry alternative phenotypes and alleles, so it should be possible to explore the relative importance of these processes and define more precisely how *Nnc1* modulates neurogenesis.

Generic resources versus specific goals

The resources generated by this program project and the two scientific cores will be open to the entire research community for collaborative and independent analysis. The new bioinformatic tools that we are now in the process of testing will ultimately catalyze far more extensive research on the structure of the mouse CNS. Each of the first three projects, when viewed separately, is a major resource for virtually any type of neuroscience. Furthermore, the methods are generic in the sense that they can be applied to a wide range of systems and problems, both in research and in teaching.

Our Neuroinformatics Program Project will allow us to combine brain maps with gene maps and thereby to explore the genetic basis of natural variation in CNS architecture—a topic that is now terra incognita. The mouse is the ideal mammal to exploit in developing these resources. We will address and answer a wide variety of questions: What genes control the size of the caudate nucleus? Why does one mouse have a large hippocampus relative to total brain size, whereas another has an unusually small hippocampus? What genes modulate the relative proportions of neurons and glial cells in neocortex, or the ratio of projection neurons and GABAergic interneurons in cerebellum or thalamus?

The most innovative component of our system is that it will have a completely open structure offering unrestricted Internet access to all key raw data and even to the neurohistological collection itself. As a result, other neuroscientists will be able to independently use both the tissue and the genotypes. The web site and the Mouse Brain Library (MBL) already illustrate the main features of an open research workbench. Our distributed team of investigators has already demonstrated the ability and drive to collaborate in assembling the key resources required to implement the system.

Only by combining efforts can scientists efficiently tackle the complex genetics of CNS structure. The amygdala, for example, has incredible structural complexity, and its architecture is undoubtedly variable among strains of mice. This structure may differ dramatically in mice such as *Mus spicelegus*, a colonial mound-building species from the Ukraine, and the standard commensal species *Mus musculus domesticus*. In assembling the MBL, we will produce highly detailed images of the amygdala for several thousand genetically defined mice, but we do not have the expertise or time to explore the genetic basis of anatomical and behavioral variation associated with the amygdala and many hundreds of other structures. This project will make the genotypes available along with powerful mapping tools in a "no strings attached" environment. We will not monitor or track who is using the collection, and we do not request coauthorship on papers that result from use of the resources.

In the next decade, this program project should revolutionize quantitative neuroanatomy and stereology and give this field a broad quantitative genetic foundation. The scientists involved in this project are fully aware of the bioinformatics challenge of integrating four disparate, if complementary, groups, but we are committed to overcoming obstacles in delivering bioinformatics tools and resources to the neuroscience community. We hope that we succeed in conveying the excitement all of us feel about the whole program and about our individual projects. Although this is an ambitious project, we are confident that we can assemble these tools and resources in five years.

Community-wide collaboration is familiar to geneticists who have used recombinant inbred (RI) strains in their research. A single genotype is replicated as an RI strain, and multiple individuals are phenotyped by many investigators. Data are accumulated on both phenotypes and genotypes. For example, John Belknap and colleagues have a database consisting of 400 behavior and neuropharmacological traits studied in the BXD/Ty set. Benjamin Taylor, Rosemary Elliott, and our own group maintain large databases on genotypes in RI strains. Mapping efforts in RI strains, however, suffer from some serious disadvantages. Even the largest RI panel, the BXD/Ty set mentioned above, consists of only 35 strains. RI strains currently have low statistical power and can be used only to map a major QTLs. This is one reason that we have pushed ahead with our advanced intercross. Not only will it complement the BXD set of strains, but it will significantly increase the statistical power to find and resolve QTLs.

Rationale

Why have we begun this ambitious project? Our primary purpose is to provide resources that are needed to efficiently and precisely map QTLs that influence CNS structure. Our first three projects generate the phenotypes and quantitative traits, and we also need to analyze the corresponding genotypes. Until now, mapping quantitative traits has been a cottage industry. A research group working within the constraints of a typical R01 grant budget can barely afford the 50,000 genotypes required to analyze 200–400 F_2 animals. Each group must also generate its own set of animals, genotypes, and phenotypes—a massive undertaking when quantitative neurohistology is involved. This extremely inefficient method only hinders QTL mapping. Furthermore, QTLs mapped with this small-scale approach always have poor positional precision (see explanation below). An unavoidable consequence is that the prospects of cloning QTLs by a positional candidate approach are usually very dim. The Genotyping and Mouse Colony Core is extremely valuable because we will do this onerous work for a very large sample of animals, so other researchers will not waste time and effort they would otherwise have to expend in breeding mice and processing tissue, at least in the context of neuroanatomical analysis of the CNS. Instead, they can focus on the quantitative analysis of the vast quantity of images and data that will be available in the MBL and the Neurogenetics Tool Box. By genotyping such a large genetic cross and then placing the results on the web, together with large collections of "virtual" slides, we will free other scientists from the budget and time constraints that would otherwise continue to hamper quantitative neuroanatomical studies.

We hope that scientists who use these resources will share their results so that the appropriate data on phenotype and genotypes can be added to the expanding database on variation of the mouse CNS housed at <nervenet.org>. This will be only a request, not a stipulation for using the MBL. Scientists have an incentive to share their results by publishing them in an archive that has already been set up (see, for example, <nervenet.org/netpapers/Rosen/RosenPapers.html>).

The External Advisory Board

Our group will consult with and be advised by a distinguished panel of senior scientists and database and web development experts on an annual basis. The following members of the External Advisory Board were selected on the basis of their stature in major areas of bioinformatics and neuroscience. (See the Administrative Core description for more detail.)

Adrianne Noe, Director of the National Museum of Health and Medicine in Washington, DC

Arthur Toga, Associate Director of the Neuroimaging Center at UCLA

John Belknap, Professor of Neuroscience, Oregon Health Science Center

Joel Richardson, Senior Programmer, co-PI of the Mouse Genome Database, Jackson Laboratory

Institutional support for this Program Project

The University of Tennessee, Memphis, has made a strong commitment to mouse neuroscience in general, and to this research project in particular. Institutional funds have provided direct financial support for work that has led up to this submission, including over $20,000 spent on the iScope prototype. The Associate Dean of Research, Dr. Michael Dockter, is pushing our efforts forward with enthusiasm, and he has made the following commitments on behalf of the Dean of the School of Medicine.

1. A new genotyping center headed by R. Williams will be established on campus. This genotyping center will have the capacity to generate over 200,000 genotypes per year and will be ready in the middle of Year 01. Over $150,000 will be expended by UT Memphis for equipment to supply the center. Our program project will have a guaranteed access to no less than 100,000 genotypes per year. This strong institutional support means that we do not need to request any funds for equipment for genotyping the extremely large intercross that will be used in all four projects.
2. UT Memphis has designated approximately 500 square feet of newly renovated laboratory and office space for use by the Internet microscope project (Project 2) starting in Year 02, a time at which we expect to have four microscopes online full time.
3. UT Memphis has agreed to purchase up to $40,000 worth of equipment in Year 01 that is required by all four projects and to continue investing up to $30,000/year for the duration of the grant.
4. UT Memphis has agreed to provide state-of-the-art high-bandwidth Internet connections for all web and video servers. At a minimum, all machines will have 100BaseT connections, but if demand continues to rise, our servers will be connected directly to the gigabit Ethernet backbone.

Conclusion

The complexity of CNS development is staggering. In mice approximately 75 million neurons and 25 million glial cells are generated, moved, killed or connected, and integrated into hundreds of different circuits over a period of 1 month (Williams 2000). The process is coordinated by the expression of a large fraction of the genome—as many as 40,000 genes are involved (Sutcliffe 1988; Adams et al. 1993). These same genes coordinate the development of the human brain, but more than a thousand times as many neurons are generated (Williams and Herrup 1988) and their integration and training take more than a decade. Five thousand genes probably have common roles in cellular metabolism, but another huge complement have selective, transient, and partially redundant roles in the development of different parts of the brain (Usui et al. 1994; Gautvik et al. 1996). Reductionist approaches that focus on isolated genes, molecules, and processes may seem hopelessly inadequate for disentangling the complex genetics of normal brain development, but they are essential at this early stage of analysis and understanding. To begin to understand the design of the brain, whether that of a mouse or a human, we need to extract the key genes that control and modulate cell proliferation, differentiation, and survival. A significant fraction of genes are polymorphic—that is, they exist in multiple forms or alleles—and differences in alleles generate variation in CNS structure that may be subtle or quite extreme. Normal variation is a significant source of the fascinating and sometimes disturbing behavioral differences among humans.

Operational Plan

Administrative aspects of the operational plan for this program project are discussed in the Administrative Core on page 120 of the application. The quality of network support will play an important role in the success of this project. We have appended letters from network administrators at UT Memphis (pp. 147–148), Beth Israel Deaconess Medical Center (p. 149), and from administrators at Roswell Park Cancer Institute and Drexel University (p. 154) affirming their willingness to provide a high level of support. Resources at UT Memphis in support of this bioinformatics project are extraordinary and are listed in letters from Dr. M. Dockter (Associate Dean for Research) and Dr. S.T. Kitai (Chair of the Department of Anatomy and Neurobiology) on pages 144–146 of the application.

The multidisciplinary and multi-institutional structure of this program project will require an engaged program director. Dr. R. Williams is actively involved in virtually all aspects of the work proposed here are will be able to devote a great deal of time and energy to this project and to ensure that communication among the four groups is constant and positive. Dr. Williams has worked for several years with both Drs. Rosen and Manly. Given our progress assembling the MBL, the distance between research centers has not been even the slightest impediment.

Permanence and Maintenance Plan: Phase 2

We recognize that maintenance is a serious problem, particularly with volatile web-based resources. The issue of permanence is taken up particularly in Project 2. This is one of the main reasons that we have asked Dr. A. Noe, Director of the National Museum of Health and Medicine, to join our External Advisory Board (p. 150). We have also gone to the effort to obtain a domain name for our bioinformatics effort that is independent of any particular research institution (<nervenet.org>). Thus, our resources will be permanently available at this URL regardless of long-term institutional affiliations of our members. This domain can accommodate and a tremendous amount of growth and will be able to host other neuroinformatics sites.

Internet publishing and archiving is now advancing rapidly (e.g., BioMed Central), and it seems probable that large and well-endowed organizations other than NIH will soon be equipped to handle, archive, and mirror sites such as <nervenet.org>. We also hope to collaborate with other neuroinformatics groups interested in the mouse CNS. This is one reason we are so pleased that Dr. Toga has been able to join our External Advisory Board, since he is the principal investigator of a newly funded Mouse Brain Mapping Project (letter p. 152) that will complement our own work by studying developmental changes in the CNS of several strains of mice.

Evaluation Plan

Evaluation of the progress will be assessed twice a year—once with our external advisory group and once without. Since all of our projects are in Phase II, one extremely effective way to evaluate progress is to track usage of web resources. We currently monitor usage of the <nervenet.org> site using the program Analog. We will also work closely with HBP staff to assess progress and to make changes in our program to improve quality and quantity of our research. Evaluation of progress of individual projects is taken up in the following four project descriptions.

Research Plan

Pages

RESEARCH PLAN

PROJECT 1: THE MOUSE BRAIN LIBRARY

PROBLEMS AND ISSUES ADDRESSED

In recent years there has been an explosion in the number of tools and techniques available to researchers interested in exploring the genetic basis of all aspects of CNS development and function. This program project focuses on new quantitative trait locus (QTL) methods that enable researchers to study any heritable trait—be it behavioral, pharmacological, or, as in our work, morphometric. Over the past two years, we have created the Mouse Brain Library (MBL) at < nervenet.org/mbl/mbl.html> in an ambitious attempt to provide high-quality histological material over the web for the purpose of mapping and ultimately cloning QTLs that modulate the architecture of the CNS. The MBL contains high-quality, uniformly processed sections from a variety of recombinant inbred, standard inbred, mutant, and outbred mouse strains, as well as from several F_2 intercrosses and one very large advanced intercross. The MBL currently contains images of more than 600 brains, and we will expand this number at least three-fold over the course of the project (see Core A). Because the MBL serves as the linchpin for the other three parts of this program project, we aim to greatly expand its capabilities as described below.

Aim 1. Extend the MBL.

The current MBL contains images of celloidin-embedded, Nissl-stained brains representing over 100 different strains. The brains were imaged at both 25 µm/pixel and 4.5 µm/pixel resolution. During the course of this proposal, we will extend and diversify this collection, adding recombinant inbred (RI) strains to complete our collection of BXD, AXB, and BXA sets. We will include over 600 mice from the tenth-generation advanced intercross (the G_{10} set—see Core B) and will complete the collection of standard inbred strains. We will also increase the depth of this resource by adding more images (1 µm/pixel) to the database, representing 40-50 anatomical structures of interest to the neuroscience community. We will begin to include immunohistochemically stained tissue and will stain adjacent sections from the celloidin collection for complementary information.

Aim 2. Generate a Relational Database for the MBL.

Our current image file archive is functional, but much more needs to be done to improve the underlying database architecture of the MBL collection. The MBL structure is not suitable for the various types of image comparison, manipulation, and analysis we intend to add over the next five years. We therefore aim to revamp the entire structure of the MBL to allow us to incorporate improvements, and to better integrate the MBL with the three other projects in this application. Substantial architectural changes in the MBL are essential to allow the collection to expand from the current 2 gigabytes of images up to a terabyte of images and through-focus series that we plan to acquire.

Aim 3. Enhance the Interface.

Changes in database and file structure outlined in Aims 1 and 2 above will increase the depth and power of the MBL. It is therefore imperative that this overhaul of the underlying core of the MBL be accompanied by a significant improvement in its ease of use. While some of the interface enhancements will no doubt be suggested by users of the MBL (and we plan to implement them during the course of the proposal), we currently have in mind a number of interface enhancements:

- Improve navigation from slide, to images, to specific coordinates
- Create JavaScript and Java applets for faster and more transparent navigation through the MBL collection
- Integrate the iScope through-focus series (short video clips) with static images in the MBL
- Combine segmented vector graphics from the NeuroCartographer Project with MBL images
- Allow searches by regions of interest (ROIs), anatomical structures, and standard stereotaxic coordinates
- Allow custom downloads of 12-bit and color images
- Allow creation of customized CD-ROM (or DVD-RAM) of images

Aim 4. Create Image Analysis Tools.

The usefulness of the MBL as a tool for the investigation of neuroanatomical QTLs lies not only in the ability to search and manipulate images, but in the ease of analysis. Toward that end, we will develop systems for direct web-based analysis of images using modifications and extensions of Image/J, a free Java image analysis program. We will allow fast morphometry and image enhancement of the MBL collection. For off-line analysis, we will provide hooks and tutorials that explain how to use other popular image analysis packages, including

NeuroZoom and NIH Image. We will also integrate data sets from the NeuroCartographer Project (Project 3) to aid in screening and analysis of the MBL. For example, once the MBL collection has been fully segmented it will be possible to call up specific sections simply by defining an ROI. This hybridization between Projects 1 and 3 will require the construction of special relational databases.

STATUS OF CURRENT RESEARCH EFFORTS

Our web site will allow "one-stop shopping" for scientists interested in mapping neuroanatomical quantitative trait loci (QTLs). Of course, the MBL is also extremely useful in studying nongenetic sources of CNS differences, for example, right-left asymmetries, sex differences, effects of age, and correlations between structures within and across genotypes. As discussed in the Introduction to this application, we will integrate quantitative information (e.g., brain weight, body weight, age, sex, etc.), high-quality histology, through-focus high-magnification movies, and the tools for segmenting ROIs, with detailed genetic information and the tools to map QTLs. In this section of the proposal we first briefly consider the complementary roles of complex trait analysis and stereological principles in biology and in neuroscience. This is followed by a discussion of our progress in mapping QTLs for different types of CNS traits. We conclude with an explanation of the Mouse Brain Library and our progress to date in creating this resource.

Mapping Quantitative Trait Loci

QTLs in Biology: What they are and why they are worth mapping

Quantitative trait loci are normal genes in every sense of the word. The somewhat unwieldy term *quantitative trait loci* highlights the fact that variant forms—or alleles—of QTLs have relatively subtle quantitative effects on phenotypes (Lynch and Walsh 1998). QTLs identified in human populations are often referred to as susceptibility genes because humans carrying certain alleles are at greater risk of developing particular diseases. QTLs are often contrasted with Mendelian loci that have pronounced and usually dichotomous effects on phenotypes, but the demarcation between quantitative and qualitative traits is often blurred. QTLs that have particularly large effects—often verging on producing Mendelian segregation patterns (e.g., 1:2:1)—are referred to as major-factor or major-effect QTLs. Their large effects make them the easiest QTLs to map.

If large numbers of genes collectively control variation in a CNS phenotype (neuron number, cell ratios, volume of a region, etc.), then allelic differences at any one locus will be associated with small, possibly undetectable differences in phenotype. This phenomenon is referred to as the *infinitesimal model* of polygenic action (Lai et al., 1994). In contrast, if only a small number of genes control much of the variation and if their alleles exert relatively large effects on the phenotypes, then it will often be possible to map single QTLs. Quantitative genetic analysis of variation in bristle number in *Drosophila* provides a superb example of what can be achieved by "disassembling" a polygene into its constituent QTLs. Normal allelic variants at the *achaete-scute* complex and at the *scabrous* locus each account for 5–10% of the variance in numbers of these sensory organs (Lai et al., 1994; Mackay, 1996). In this system, the QTLs controlling natural variation are also known to be vitally important developmental genes that had been identified previously based on mutant phenotypes associated with null alleles.

The ability to map QTLs in the CNS depends critically on the amount of variance explained by allelic differences at single QTLs and the technical precision with which traits can be measured (a signal-to-noise problem). The more accurate and reliable the method of phenotyping, the higher the heritability, and the greater the number of resolvable QTLs (Lander and Schork 1994).

QTLs in Neuroscience

The introduction of quantitative genetics and QTL mapping in neuroscience is a relatively new phenomenon. Among the traits that have been studied are EEG and evoked response potentials in humans (Begleiter et al. 1998), sensitivity to ischemia in rats (Rubattu et al. 1996; Jeffs et al. 1997), alcohol consumption, tolerance, and withdrawal (Berry et al. 1995; Gallaher et al. 1996; Gill et al. 1996; Buck et al. 1997; Erwin et al. 1997a, 1997b; Gehle and Erwin 1998), behavioral effects of stress (Cabib et al. 1997), tolerance to various drugs of abuse (Tolliver et al. 1994; Grisel et al. 1997), and retinal ganglion cell number (Williams et al. 1996b, 1998) and brain weight (Belknap et al. 1992; Williams et al. 1998) in mice. While the number of neuroscientists examining QTLs is relatively small compared to those mapping QTLs in biology as whole, the recent rapid growth of this field is unmistakable. Virtually every report mentioned above was published within the last four years, and the majority within the last two. In large measure, the reason for this recent surge in neurogenetics is that QTL methods have improved (Tanksley 1993; Lander and Schork 1994; Kearsey and Pooni 1996). These methods depend on the polymerase chain reaction (PCR), high-density genetic maps, and sophisticated statistics programs. The application of these methods in the next decade will revolutionize our understanding of normal genetic mechanisms controlling CNS development, susceptibility to disease, and even CNS evolution.

Stereology and Neuroscience: An Efficient Method for Acquiring Quantitative Data

Unbiased techniques to study cell populations and to estimate volumes from cross-sectional areas have been available for many years. But the introduction of highly efficient methods to count and measure cells in thick and thin sections, particularly direct three-dimensional counting (Williams and Rakic 1988) and several adaptations of the disector (Gundersen 1980), has made it easier to obtain accurate sets of neuroanatomical data. As a practical matter, however, the most important advance in quantitative techniques has been the rapid hybridization of microscopes, video systems, fast microcomputers, and powerful application programs such as NIH Image, Stereologer, and NeuroZoom. Investigators can now use more trustworthy quantitative techniques to study large numbers of cases. These technical advances make it possible to undertake a QTL analysis of CNS structure. Studies that previously would have taken months or years can now be done in weeks or months. The growth of stereology as a tool for neuroscientists is perhaps best illustrated by an examination of the frequency of published papers employing stereological techniques. In the past 20 years, more than 2000 published papers have relied on stereology—and most of these were published within the last decade. Of these, approximately 600 have been in neuroscience-related fields. It is clear, therefore, that the use of stereology in neuroscience has gained momentum in recent years as a method for accurate and efficient quantification of neuroanatomical traits. Using stereological principles, mapping QTLs for these neuroanatomical traits becomes quite feasible.

QTLs for CNS Morphometric Traits

Brain weight is a classic polygenic trait and one that is likely to be influenced during development by the activity of hundreds of genes. Brain weight is also affected by maternal factors and by myriad environmental factors (e.g., Collins 1970; Katz and Davies 1983; Wahlsten 1983). Many factors that affect body size have important pleiotropic or correlated effects on brain size, making the selectivity of action a critical problem (Lande 1979). From the point of view of genetic complexity, it is hard to imagine another morphometric trait that would be more difficult to resolve into individual QTLs. In spite of the difficulties, my collaborator, Dr. Williams has already succeeded in mapping QTLs that selectively affect the weight of the brain (Williams et al. 1996a, 1997).

Measuring brain weight. A simple analysis of differences in brain weight among numerous strains of mice has been carried out in part as a prelude to this program project. For the great majority of animals, we now have information on sex, body weight, age, and type and quality of fixation. Both sexes and a wide range of ages have been studied and incorporated into the MBL. Perhaps the most remarkable aspect of the data is that they reveal large differences in brain weight between several substrains of mice that carry only modest genetic differences. For instance, brain weights of BALB/cByJ and BALB/cJ mice differ by 76 mg. C57L/J and C57BL/6J differ by 88 mg. C3H/HeJ and C3H/HeSnJ also differ by 88 mg. The closely matched differences in these three pairs are tantalizing. Presumably, the differences are generated by a very small number of polymorphic genes—probably one or two genes. Since some of the parental strains were already fully inbred prior to being split into substrains, it is unlikely that the variation is due to the fixation of alternative alleles. Mutations or reversions are a more likely cause of this remarkable substrain variation. Unfortunately these large differences cannot easily be mapped because suitable polymorphic marker loci for this type of analysis have not yet been developed. However, comparative gene expression analysis (e.g., differential display), subtractive hybridization, and candidate gene approaches might be used to explore factors that contribute to these 80-mg differences.

In a standard mouse colony, variation in brain weight has a heritability that ranges from 0.35 to 0.7 (Roderick et al. 1973, 1976; Seyfried and Daniel 1977; Fuller 1979; Henderson 1979; Atchley et al. 1984; Williams et al. 1996a; Strom and Williams 1997). The correlation between values is a direct estimate of heritability in a narrow sense. Parents and offspring in this case are seven sequential generations (F_2 to G_{10}) of an advanced intercross between C57BL/6J and DBA/2J (G. Zhou and R.W. Williams, in progress). This estimate of heritability is somewhat lower than that found in most previous work, averaging 0.35 to 0.45 for different datasets and crosses.

In comparison to brain weight, variation in neuron number has a heritability of approximately 0.8 for granule cells in the dentate gyrus (Wimer and Wimer 1989) and between 0.7 and 0.9 for retinal ganglion cells (Williams et al. 1996a, 1998). These values are sufficiently high to motivate a QTL analysis.

Selectivity of QTLs. Before mapping QTLs, we need to consider the issue of specificity of gene action. The brain weight data in the MBL were not corrected for the significant differences in the mean body weights of these strains. Because brain weight may simply be a function of body size, there is a risk of mapping body weight QTLs instead of QTLs that have specific effects on brain weight or cell number (Hahn and Haber 1978; Lande 1979). To ensure that brain weight and not some other factor was mapped, variation in brain weight predictable from variation in body weight must be factored out. A crude way of doing this is to take the ratio of brain to body weight as the variable, but a computationally and conceptually better approach uses regression analysis to remove predictable variance associated with body size (Williams et al. 1997). Williams and colleagues used multiple regression to remove variance in cell number that was actually associated with total brain weight. This same logic applies in mapping QTLs that affect particular CNS cell populations (Williams et al. 1998).

The lesson to be taken from this discussion is that whatever types of QTLs we are trying to map, we need to carefully consider the higher-order structures and make sure that we have taken variation in these structures into account.

Mapping Brain Weight QTLs with RI Strains and F_2 Crosses. Several of the largest RI strain sets were generated by crossing strains with large and small brains. The BXD RI strains originated from a cross between C57BL/6J and DBA/2J (Taylor 1978; Taylor et al. 1999). The AXB and BXA strains originated from reciprocal crosses between C57BL/6J and A/J. In both cases brain weights of the parental stains differ by 80 mg. In an analysis of 20 BXD strains, John Belknap and colleagues (1992) identified three intervals that were likely to harbor QTLs: chromosome (Chr) 7 near *D7Rp2*, Chr 11 near *Hba*, and Chr 17 near *D17Tu7*. Williams and colleagues have extended this analysis to 26 BXD strains. Rather than mapping the data on brain weight itself, they mapped the difference between the actual brain weight and the value predicted on the basis of body size, age, and sex. They did not detect QTLs on Chr 7 or Chr 17. However, the proximal part of Chr 11 definitely does harbor a QTL—brain size control 1 (*Bsc1*)—that is specifically responsible for variation in brain weight (Williams et al. 1996a). This analysis is now being extended to the full set of 35 BXD strains.

Over the past two years, Williams and colleagues have used three F_2 intercrosses and additional RI strains to define several additional brain weight QTLs (Strom and Williams 1997). The advantage of the F_2 crosses is that QTLs with relatively small effects on brain weight can be mapped. One of these F_2 crosses is between BALB/cJ, a strain with a very large brain, and CAST/Ei, a wild Asian mouse with a small brain. In this cross, brain weights of the F_1 generation overlap those of the BALB/cJ parental strain. Although it appears that a heavy brain weight is inherited as a dominant trait, all of these F_1 progeny were born to BALB/cJ mothers, and maternal nongenetic factors are probably important. The variance among F_2 individuals is higher than that of the parental strains. This increase in variance is due to the segregation and assortment of QTLs that affect brain weight among F_2 offspring.

It is satisfying to move away from the phenomenology of brain weight and to describe variation in terms of individual gene loci. However, we still do not know what parts of the brain or what cell populations are most affected. We are taking two approaches to these questions. The first approach is to literally disassemble each brain into component parts and determine how the weights of these parts co-vary with the weight of the whole minus the weight of the part. Using this approach, we have succeeded in mapping four QTLs that have selective effects of the size of the cerebellum (Gilissen and Williams 1997; Airey et al. 1998). The second approach is to proceed directly to stereological investigations of CNS nuclei and cell populations. This, of course, is the main goal of this program project. A library of brain tissue from particular crosses would be especially useful since investigators could map multiple QTLs and study the genetic basis of correlations among different CNS regions. This strategy would also greatly reduce the amount of effort put into genotyping. This is precisely why we have spent the past two years beginning the slow process of building a large library of sectioned mouse brains suitable for quantitative genetic analysis of CNS architecture. This is also the reason why we have initially opted to concentrate our efforts on well-characterized lines of RI strains. While RI strains do suffer from some inherent problems for QTL analysis, they have the very significant advantage of making studies of genetic correlation between interconnected parts of the CNS possible. Investigators can also then use the data we generate for RI sets to explore behavioral, developmental, and neuropharmacological correlates of differences in CNS structure and cellular demographics.

QTLs for Retinal Ganglion Cells

Mapping Cell–Specific QTLs. It is practical to map QTLs that affect individual neuronal populations. Williams and colleagues (1998) have done this type of fine-grained analysis for one of the more accessible populations of neurons in the CNS—the projection neurons of the retina, also known as retinal ganglion cells. One reason they chose this population is that it is possible to count retinal ganglion cells easily and precisely. Each cell has one and only one axon in the optic nerve, and a quantitative electron microscopic census of axons in a single cross-section of the nerve provides a reliable and unbiased estimate of total neuron number (Rice et al. 1995; Williams et al. 1996b).

They began by estimating the size of this cell population in 5–10 individuals from each of 5–10 different inbred strains and extended the analysis to 20 common inbred strains (Williams et al. 1996b; Zhou and Williams 1999). Variation among strains was substantial, and they decided to count ganglion cells in ~8 individuals from each of the 26 BXD strains and 12 BXH strains. Strain averages tended to fall into well-defined modes that corresponded to the parental strain averages of 55,000 (C57BL/6J) and 63,000 (DBA/2J). This striking non-normal distribution of phenotypes across the set of BXD RI strains suggested that two or three QTLs were modulating neuron number.

Using 26 strains of BXD mice that were then available, they were able to map the QTL that is primarily responsible for the bimodality of strain averages (Williams et al. 1998). They found an excellent correspondence between phenotypes and genotypes on chromosome 11 near the *Tstap91a* gene. The correlation between strains with high and low neuron number and alleles at *Tstap91a* was 0.69 (see Table 3 in Williams et al. 1998). The genome-wide probability of getting a correlation this high by chance alone is less than 0.01.

These two examples of QTL mapping of quantitative neuroanatomical traits clearly demonstrate the utility of this approach in tackling issues of the genetic modulation of the CNS.

Gross morphologic measures as dependent measures

As part of ongoing research in our laboratory, we have consistently used various gross morphological measures as dependent measures in our studies. For example, we have been interested for a number of years in the biological substrates of anatomical asymmetry. We were initially struck by the simple notion that with respect to asymmetric brain regions, symmetry can result developmentally from either an increase in the normally smaller side, a decrease in the usually larger side, or a combination of the two processes. In the first case, the measure of total brain area of symmetric brain regions would be larger than their asymmetric counterparts whereas the opposite would be true in the second case. Brain areas would be similar if the third scenario were true. To test these alternatives, we examined photographs of the planum temporale (Geschwind and Levitsky 1968) to measure total planum area in the left and right sides (Galaburda et al. 1987). Similar to previous findings, we found a leftward asymmetry in the majority of cases. When we plotted the total planum area (right + left) against a measure of directionless asymmetry (magnitude), we found a significant negative correlation, indicating that as asymmetry increased, the total planum area decreased. These results demonstrated that symmetric brains were larger than their asymmetric counterparts and, further, support the hypothesis that asymmetry was the result of the production of a small side rather than the production of a large side. We have since replicated these findings in both rats and mice (Galaburda et al. 1986; Rosen et al. 1989a, 1989b, 1991, 1993).

We have also been aware for some time of the complications involved in estimating volume from two-dimensional structures. We were initially troubled by the potentially different methods of computation involved in estimating volume from cross-sectional areas. We compared the efficacy of each of the common methods for estimating volume by using mathematical simulations as well as actual morphometric measures from brain regions. As it turned out, there were no meaningful differences between the various estimates when many sections are used; with fewer sections, however, Cavalieri's estimator was most accurate. Yet while the Cavalieri approach provides a better approximation of volume under some circumstances, it requires equally spaced sections—a criterion sometimes difficult to meet in the real world (e.g., missing sections affect accuracy). We therefore devised methods for the estimation of brain volume from unequally spaced sections; these methods are quite accurate when large numbers of sections are used (Rosen and Harry 1990).

These studies demonstrate that in our experience, useful information can be obtained from studies involving gross morphometric measures. Moreover, we have been concerned about the computational problems involved in stereological measures and have devised methods to supplement these efficient and accurate methodologies.

The Mouse Brain Library

In the summer of 1997, Drs. Rosen and Williams began a collaboration to produce a large collection of sectioned brains specifically designed to enable us to begin quantitative genetic studies of neuroanatomical phenotypes. The Mouse Brain Library (MBL) now includes a large collection of standard inbred strains that provide a strong platform for building the MBL to the level described below. In addition to these inbred strains, we have a large collection of mice from 4 RI strain sets (BXD, AXB, BXA, and BXH) as well as some F_1 and F_2 intercrosses. During the course of this proposal, we anticipate adding approximately 300–400 mice each year to the library in order to enhance the representation of RI strains, standard inbred strains, F_1 and F_2 intercrosses, and the advanced G_{10} intercross (see Core B).

Whenever possible, strains have been obtained from the Jackson Laboratory (Bar Harbor, ME). Intercross progeny (F_1 and F_2) presently in the collection were generated at UT Memphis using Jackson parental strains, and those listed in the current proposal will also be generated at UT Memphis. Age, sex, and body and brain weights are known for almost all mice. In some cases, litter size and mother's parity is also known. See < nervnet.org> for more data. The roster of current processed brains is summarized in Table 1.

Bioinformatics components of the current Mouse Brain Library

One of our major goals in creating the Mouse Brain Library was to provide a central repository of high-quality histological material for the neuroscience community at large. Although originally thought of as a physical library from which researchers could "check out" slides for their experiments, the concept has evolved to that of a "virtual library." Over the past year in particular, we have concentrated significant effort and resources toward making the data contained in this library available over the web. Importantly, we have laid the groundwork for a future in which it will be possible to conduct a variety of neuroanatomical experiments over the Internet using readily available, platform-independent software.

Our web site at <nervenet.org> contains a wealth of information for neurogeneticists, including online versions of published and unpublished papers, links to important web resources, and a variety of databases. These databases include information on brain weight, sex, parity, body weight, etc., for over 5000 mice. In addition, the Mouse Brain Library database is available online and is available for downloading or can be searched using a web-based front end.

Table 1.
Current Contents of the Mouse Brain Library

Inbred		Recombinant Inbred							Intercrosses	
Strain	#	Strain		#	Strain		#		Strain	#
129/SvJ	8	AXB	1	8	BXD	1	11		ABXD5F2	44
A/J	5		2	8		2	6		C3HAF2	1
AKR/J	4		4	8		5	6		B6D2F1	4
BALB/cByJ	8		5	8		8	12		D2B6F1	4
BALB/cHeA	2		6	5		9	8		B6EiC3H	4
BALB/cJ	5		8	8		11	6			
C3H/HeJ	3		10	8		12	9			
C3H/HeSnJ	8		12	8		13	8		Outbred	#
C57BL/6J	11		13	5		14	6		CD-1	8
C57BL/10J	7		15	8		15	5			
C57L/J	6		18	8		18	7		Wild	#
C58/J	8		19	8		19	8		CASA/Rk	4
CBA/CaJ	6		20	9		20	5		CAST/Ei	7
CBA/J	2		24	7		22	8		MOLC/Rk	4
CE/J	4	BXA	1	8		23	8		MOLF/Ei	6
DBA/1J	7		2	8		24	8		PANCEVO/Ei	2
DBA/2J	10		4	8		25	8		SPRET/Ei	2
FVB/NJ	4		7	8		27	8		WSB/Ei	2
LG/J	7		8	8		28	8			
LP/J	6		11	8		29	8			
NOD/LtJ	9		12	8		30	8			
NZW/LacJ	5		13	8		31	8			
PL/J	5		14	8		32	13			
SJL/J	6		16	8		33	8			
SM/J	8		17	12		34	8			
SWR/J	6		24	6		35	8			
			25	11		36	8			
			26	8		38	8			
						39	8			
						40	8			
						42	8			
		BXH	3	2	BXH	10	5			
			4	6		11	4			
			6	5		12	4			
			7	5		14	5			
			8	1		19	3			
			9	6						

Limitations of Existing Approaches

Our goal in creating the MBL is to provide a high-quality repository of information useful for researchers wishing to examine the genetic basis of normal variation of mouse brain structure. As it is currently constructed, the MBL can be used for gross neuroanatomical measures (e.g., total forebrain volume) and for gross comparison between strains of mice. We currently have higher-resolution images (4.5 µm/pixel) available for approximately 20% of the cases, which can be used for a finer level of analysis. That being said, there are

several obvious ways to extend the MBL and to improve the ease of access to this collection. We need to expand the numbers of animals in the MBL to increase the power of the statistical analysis, and we need to provide images with higher resolution in order to widen the range of questions that can be addressed. The underlying database structure needs to be revamped to better reflect the type of information being stored. Finally, we are adding a wealth of new features to the MBL (see Projects 2, 3, and 4) that must be completely integrated and easily accessible.

EXPERIMENTAL PLAN

During the course of this proposal, we will extend the breadth and depth of the MBL to reflect the increase in subjects, types and numbers of images and through-focus series (Project 2), and information concerning regional segmentation of the brain (Project 3). We will provide the tools with which individuals can measure and analyze images, either working over the Internet or by downloading customized datasets onto their own computers. In addition, the MBL will be the key source for CNS phenotypes that are such a critical part of the gene mapping analysis system described in Project 4.

1. Extend the MBL

We will extend the depth and breadth of the MBL on three fronts. First, we will first increase the numbers of mice in the MBL by supplementing the current collections and by adding inbred strains, RI strains, and advanced intercrosses. Second, we will introduce brains that have been processed with stains other than what is currently in the MBL (i.e., Nissl stains). Finally, we will increase the range of images that will be stored in the MBL.

Additional Brains

Every brain that is added to the MBL originates in the Genotyping and Mouse Colony Core (Core B). The animals are perfused and the brains are shipped to the Neurohistology Core (Core A), where the brains are processed. For detailed methodologies, please see the respective Core descriptions.

Completing the RI strains. The current roster of RI strains in the MBL (BXD, AXB, BXA, and BXH) can be seen in Table 1. In the first year of the project, we will increase the number of animals representing each strain to at least 12 in order to improve the statistical power of subsequent analysis.

Advanced Intercrosses. As detailed in the description of the Genotyping and Mouse Colony Core (Core B), Dr. Williams and colleagues have been undertaking a large project to generate an advanced intercross between C57BL/6 and DBA/2 (the parental strains for the BXD RI strain). They currently have over 1400 tenth-generation (G_{10}) mice from this B6D2 intercross. During years 1 and 2 of the project, we will add at least 600 of these mice to the MBL. These mice range in age from P46 to P76. In year 3 we will add a minimum of 200 additional aged animals (2 years old and up). All G_{10} mice will be fully genotyped at 350 markers.

An additional 600 G_{10} animals will be held for celloidin or immunohistochemical processing (see below) in years 3 and 4 or will be used for whole brain dissection in years 4 and 5.

Complete Standard Inbred Strains. As shown in Table 1, the MBL already includes 20 inbred strains. As with the RI strains above, we will increase the number of animals representing each strain to at least 12. In addition to our standard horizontal and coronal planes of section, we will include some sagittally processed sections as well.

Additional RI sets. We will add more RI sets to the MBL. We are currently planning on adding the AKXD series (AKR/J x DBA/2J) and CXB. The Genotyping and Animal Facilities Core (Core B) has already produced the complete CXB set.

Potential Problems and Pitfalls. We are aiming to process a large number of animals over the course of the proposal, and a question can be fairly raised as to whether this is overly ambitious. In the 2 years from the inception of the MBL, we have processed over 600 brains with no funds directly earmarked for this purpose. The histological processing was performed by the equivalent of a half-time technician. Moreover, for each brain, 2 one-in-five series were generated (one for Rosen and one for Williams) whereas only one series will be used by all members of the program project. Given the level of staffing in the Neurohistology Core and in this project, we have every reason to expect that we can comfortably process the numbers of mice proposed.

Questions can also be raised as to our choice of subjects in the MBL. For example, we are not planning to include mutant or knockout mice. Instead, we are heavily weighted toward RI strains and the advanced intercross G_{10} progeny. This choice was made because our focus is on the source of normal variation in the CNS. This important question has never been systematically addressed using quantitative methods. That being said, the infrastructure would be in place for the addition of knockout and mutant mice should funding become available.

Additional Stains

The purpose of adding more stains to the MBL is to improve our ability to determine borders of neuroanatomical regions in the CNS. This is especially important for Project 3 (NeuroCartographer), in which we will use additional stains to enhance our ability to properly segment the tissue.

Celloidin-embedded tissue. Every brain currently in the MBL is cut in celloidin and stained for Nissl substance—a stain particularly well suited for determining neuroanatomical architecture and for performing cell counts. As detailed in the description of the Neurohistology Core, a number of other histological stains can be applied to tissue processed in this matter. During the course of the present proposal, we will begin using alternative stains on sections adjacent to those stained for Nissl substance in order to investigate their utility as additions to the MBL. Of the choices available, myelin staining (using the Loyez method) is likely to improve the ability to distinguish a variety of subcortical structures, including the striatum and thalamus. We therefore view the staining of adjacent series of sections to be an aid for the neuroanatomical analysis conducted on the Nissl-stained tissue.

Immunohistochemistry. While staining for Nissl substance provides an excellent overall picture of brain anatomy, other, more specialized stains can improve the delineation of certain neuroanatomical borders. Choline acetyltransferase (ChAT), for example, stains the ascending cholinergic system of the forebrain, and the distribution of ChAT-positive fibers and cells is regionally distinct. We will use this stain on 4 brains from each of the RI strains currently in the MBL (BXD, AXB, BXA) as well as 12 mice from the tenth-generation advanced intercross. We will also investigate the possibility of using other immunohistochemical stains (parvalbumin) and histochemical stains (AChE) as aids.

Potential Problems and Pitfalls. The potential technical problems associated with processing celloidin-embedded tissue and Immunohistochemistry are discussed in the description of the Neurohistology Core.

Imaging

In addition to the 25 µm/pixel and 4.5 µm/pixel resolution images, we will also acquire images at 1 µm/pixel resolution for the purpose of better visualizing individual neuroanatomical ROIs. Of greater importance to the design of the MBL in the future is the integration of these static images with the through-focus QuickTime movies being generated by the iScope (Project 2) as well as with the segmentation vectors provided by NeuroCartographer (Project 3). Because this integration requires precise alignment of images, global fiducial coordinates must be established which will require unique methodologies. In the sections below, we detail the procedure by which images will be acquired for the MBL.

A microscope stage fitted with LED sensors as described in Project 3 will be used to manipulate the slides for imaging. The sensors will determine an absolute point of origin for the slide by computing the intersection of the left and bottom sides of the slide. All other points on the slide will be described on an X,Y coordinate system with this point as the origin. As the slide is moved into position for photography, its travel along this X,Y coordinate space is recorded by electronic digital length gauges. As each image is captured, the coordinates of the bottom left and top right of the image are recorded. Each image therefore has two fiducial marks embedded with it, which, along with the point of origin, will allow proper alignment of images by triangulation. By establishing these fiducial marks, we can safely rotate and align individual images to allow for ease of measurement without losing the ability to link that image to through-focus series (Project 2) or information concerning segmentation (Project 3).

Several types of images will be generated. Most will be contrast-optimized 8-bit gray-scale images that have been compressed and saved as high-quality JPEGs (quality level 8 or 9). At the highest resolution, each compressed image requires about 0.6 to 1 megabytes of storage space. In each case, we will image the entire slide (25 µm/pixel), each section individually on the slide (4.5 µm/pixel), and then 40-50 ROIs at 1 µm/pixel. These regions will include, but will not be limited to, olfactory bulb, caudate/putamen complex, nucleus accumbens, basal forebrain, hypothalamus, septum, globus pallidus, amygdala, lateral geniculate nucleus, medial geniculate nucleus, ventrobasal and ventrolateral nucleus, anterior thalamic nuclei, lateral dorsal nucleus, posterior nuclei, hippocampal subregions, red nucleus, deep mesencephalic nuclei, inferior and superior colliculi, pontine nuclei, periacqueductal gray, and various cerebellar regions.

All images in the MBL have been and will continue to be imaged with a Micro Nikkor 60 mm camera lens with the exception of the 1 µm/pixel images, which will be captured using a Nikon microscope with a 0.5× objective. Images are digitized using a Kodak DCS560 digital camera in 16-bit gray scale mode. The CCD has a pixel count of 3060 x 2036. These digital images are subsequently transferred to Adobe PhotoShop for sharpening and rotation. During the first year of the proposal, we will be considering the implementation other imaging technologies, such as FlashPix <www.flashpix.com>, which may allow greater image depth and quality (see below).

The current collection of MBL gray-scale images has not been acquired with optical density standards. We have strived to optimize the range of contrast in every section of the collection. Thus faintly stained sections have been contrast enhanced to more closely match well-stained sections. In general, optical density values for Nissl-stained specimens are not of much analytic use (but see work of Zilles et al. 1980, 1985). However, calibrating the collection is relatively simple and will be of use to some image analysts, ourselves included. During slide photography sessions we will photograph a linear gradient neutral density filter (1 OD/cm from Edmund Scientific) in both axes. During batch processing of images, but prior to any image manipulation, we will superimpose 5-pixel-wide bands taken from the photograph of the linear gradient neutral density filter along the lower and right edges of the digital image. Color image calibration can be carried out using the simple LED device described by Beach and Duling (1993). Video camera calibration can also be carried out using a straightforward method that employs a single optical density filter described by Baldock and Poole (1993).

We will also adjust for shading corrections for the light source using a relatively simple technique. We will remove the slide or section and photograph the "uniform" light source, then divide each pixel value in the image by that in the empty field and multiply the product by ~250. This can be done automatically in batch mode by any of a number of programs, including PhotoShop 5.5, NIH Image, Image-Pro Plus (v 3.0 from Media Cybernetics), IP Spectrum, and IDL.

Potential Problems and Pitfalls. There are several potential difficulties in acquiring these images for the database. It is essential that static images located in the MBL be of exceptional quality, and we believe that the equipment being used will provide sharp images with good tonal range. Further, the ability to sharpen, change contrast, and otherwise adjust the image digitally in Adobe PhotoShop will ensure high quality.

In addition to image quality, the inclusion of fiducial points with each image is essential, as discussed above. It is important that these fiducial marks be transferable between the projects where slides will be manipulated. All projects will have identical stages made, and we will each have identical calibration slides. Before any imaging is to occur on a given day, the stage will be calibrated using this slide.

The use and construction of this specialized stage is an innovative aspect of the current proposal. Dr. Nissanov has a working prototype of this stage already operational. While we are confident that we will be able generate fiducial coordinates for all slides and sections, we are considering several rapid methods to create global series of coordinates. One possibility is to create a grid overlay to fix to the bottom of each slide before imaging. Thus, each image would be taken with grid marks visible within the plane of focus.

Finally, our ability to process the amount of images can be called into question. On the average, it takes approximately 1 minute to acquire each digital image (although this may be improved with the faster Macintosh G4 machine). It will therefore take approximately 20–30 minutes to produce the 25 µm/pixel and 4.5 µm/pixel image for a brain cut in the horizontal plane. A coronally sliced brain will take approximately 45 minutes. To acquire the 40–50 1 µm images will take another 75 minutes. We therefore conservatively estimate that a half-time imaging technician could process 3 slides/day for a total of approximately 600 slides/year. If that throughput is not enough, we will increase that technician's time on the imaging workstation.

2. Changes to the underlying data structure

Creation of Image-Specific Database Structure

The current MBL is a series of simple FileMaker Pro databases. There is a slide database (DB) that is linked to a database on tissue processing called CelloidinDB. This slide database is linked with a common gateway interface program that is able to locate images of three types (low power images of whole slides, higher power images of whole slides, and high power images of single sections). These DBs are linked to another DB in which brain weight, age, sex, parity, etc. are found. We do not have an image DB in which information concerning image type, pixel size, bit depth, acquisition parameters, contrast enhancement, digital enhancement applied, compression level, etc., are stored. In addition, Project 2 will be generating large numbers of QuickTime (QT) movies that that need to be organized and linked to the proper case in the MBL.

We will therefore design and implement a custom image DB and integrate it into the MBL. The question as to the best overall architecture for an image database that will eventually hold about 100 GB of 2D images and 500+ GB of QT movies is still open. Technologies for database design and management are constantly evolving, as are specific tools for image database management. As an example, if we choose to employ FlashPix technology <www.flashpix.com/> we gain the ability to zoom in and out of images with ease and without loss of resolution. There are a number of commercial database products designed to handle FlashPix technology, and implementing one of these or creation of our custom database would be feasible. That being said, there is no certainty that this technology will be the best available at the time that the DB is being

designed. During the first six months of the proposal, we will evaluate a range of software tools and image technologies before deciding on the platform for development of the image DB.

Archiving Images with Greater Depth

Although images in the current MBL are acquired at 12-bit depth, for reasons of storage space and speed of access they are saved as 8-bit images. We recognize the fact that users of the library may require higher-resolution images, perhaps even color, in order to address their research questions. One method of handling this issue is to archive high-resolution images off-line, and set up a system for users to request these images. We are uncertain whether these images can be eventually furnished online as the amount of data may preclude easy download. At the very least, we will be able to provide the ability to automatically order these images on CD-ROM or DVD-RAM online (see below).

As mentioned above, FlashPix technology may well solve this issue in a more elegant way. FlashPix Format and imaging architecture is a relatively new way of efficiently representing pixel-based images. The technology behind FlashPix more efficiently displays images by breaking pixel-based images into more manageable pieces. In the case of the MBL, we would create an "original" image that could be viewed at different resolutions and magnifications. Thus, rather than storing three images to represent the view of a slide, we could store one image and then call up the different views on the fly. Additionally, changes to a FlashPix image are stored as "edit files" (text based) rather than as digitized images. This has enormous potential utility with regard to implementing displays of segmentation (Project 3) and image manipulation and analysis (see below). Finally, images saved in FlashPix format can be displayed and saved in virtually any common image format.

Relating Other Databases to Image Database

In addition to the creation of a true image DB, we need to modify the existing DB to embrace the enhancements to the current MBL. Specifically, we need to create a DB of QT movies and directly link that to the individual cases and to their specific location on the sections from which they are taken. Information concerning the segmentation of regions of interest on the brains sections must also be incorporated into the database structure. Finally, we will provide hooks in the database to allow for the addition of other data as it comes online. For example, as individuals begin to measure CNS traits and map QTLs, that data will be stored in the database as well. In all cases, we will optimize the overall data structure for flexibility and speed.

Potential Pitfalls and Problems

The server running the MBL is currently a Macintosh G3. For the first two years of the project, we anticipate that the Macintosh platform will be sufficiently powerful to allow us achieve our aims. It is a distinct possibility that the MBL may require more processing power than the Macintosh operating system can handle. In that eventuality, we will switch to a Linux (or other open-source system, such as Mac OS X) operating system and transfer the database to open-source software (e.g., My SQL).

As with any database, care must be taken that the design is sufficiently flexible to allow for growth and modifications as new features are brought online. While we have done our best to carefully design the eventual architecture of the MBL, it is most certainly the case that we will be forced to deal with unanticipated additions and enhancements. Moreover, the database must be designed in such a way that migration to a different platform (e.g., Macintosh to Linux) is seamless. By using standard database programs that support SQL, Oracle, and ODBC standards, we can ensure future growth and compatibility.

3. Changes to the interface

Tour of the Current MBL

To better understand some of the changes necessary to the interface, it will perhaps be useful to discuss the current interface of the MBL and then point out how we plan to change it. Screen shots will be included, but the reader is encouraged to log into <nervenet.org/mbl/mbl.html> to view the current interface in detail.

Upon entering the MBL home page (screen shot, right), one can choose to browse the collection of images from the brains in the collection, linked to the Brain Atlas <nervenet.org/mbl/mbl_main/atlas170.html>, the brain databases <nervenet.org/mbl/mbl_main/mbl_databases.html>, a table of strains <nervenet.org/mbl/mbl_main/mbl_straintable.html>, the iScope <http://nervenet.org/mbl/iscope/aboutiscope.html>, or other links in the library

<nervenet.org/mbl/mbl_main/mbl_links.html>. Because most of the redesign of the interface of the MBL will involve the ways to view the photographic images taken from sections in the MBL, we will concentrate here on exploring this aspect of the current interface.

Clicking on "Browse the Library" takes the user to the interface for the FileMaker Pro image database (screen shot, right). There, users can browse or search on a variety of criteria, including strain, ID#, sex, age, brain weight, and body weight. All the cases matching the search criteria appear in the rightmost frame. Clicking on one of these thumbnails takes the user to a larger thumbnail of case (383 × 255 pixels) with its strain, cutting plane, sex, age, weight, brain weight, and ID# listed (screen shot below, left).

Clicking on this thumbnail then takes the user to the full 3060 × 2036 pixel image with a resolution of 25 µm/pixel. An example of one section of this image is shown in the screen shot on the right.

For 20% of the images currently in the database (mostly BXD animals), we have high-resolution images of individual sections captured at 4.5 µm/pixel. For those cases in which these section images are available, there is a slight modification to the screen shot seen above left. In this case (see right), links to individually numbered sections are present. Clicking on these links brings up a window with the high-resolution image (below left).

One can also access any brain in the database by initially linking to the table of strains on the MBL home page <nervenet.org/mbl/mbl_main/ mbl_straintable.html>. On this table, any subject with an asterisk next to it has high-resolution images available.

Changes that we will make to the interface will involve both improving the existing interface and adapting it for the addition of new components. The proposed changes are outlined below.

Improve Navigation from Slide to Images

Our current methods for moving from the thumbnail of a slide to an image of an individual section are rather clumsy. We propose to implement the ability to click on any section in the thumbnail to immediately call up the higher-resolution image of that section. One way to think of this would be as if we were using the thumbnail as an image map on a web page. While the idea behind this change in the interface is the same, our methods of implementation will be different, as we will have sophisticated links in place between all the images in a given brain.

Create JavaScript and Java Applets Tools and Techniques that Allow for Faster and More Transparent Navigation Through the MBL Collection.

The current methods of navigation through the MBL are serviceable but not yet optimized. During the course of the proposal, we will create a variety of tools that will expedite the most common requests of users of the MBL. For example, we will implement the ability to link directly from an image to any of the other data on the same mouse. Currently, we link to a few of the characteristics of the individual mouse, but as more and more phenotypes become measured, it is potentially important for researchers to be able to see what other traits are associated with an individual mouse brain.

Conversely, when one is analyzing a brain, it is important that all data that would prevent blind analysis be hidden from the user. We will set up a set of simple controls so that a user can select a series of brains for analysis and the images will appear with only an ID# present. When, at some point following the analysis, the user wishes to see the data associated with that brain, it could be called up.

Comparison of brains is likely to be of interest to the user of this resource. We will enable the ability to compare up to four individual sections on one screen. These sections could be from the same brain or from different brains. The criteria by which one chooses the comparison could be any of the data (i.e., strain, brain weight, gender, etc) or by region of interest (see below).

We are certain that there will be other interface enhancements that will become obvious only when the MBL comes into general use. We will closely monitor the ways that people use the library and will directly poll users for "wish lists" of functions and enhancements.

Integrate the iScope (Project 2) Through-focus Series with Static Images in the MBL

Integration of through-focus series will primarily involve incorporating navigation tools into the MBL that map and access these series onto the medium- and low-resolution images of slides and sections. For example, we would allow the user to put a translucent 20 x 20 pixel box over a section image (4.5 µm/pixel), and it would be linked to the QT movie closest to those coordinates and with a field of view of 100 x 100 µm. As detailed in Project 2, two types of iScope QT stacks will be acquired for each brain: (1) a set of approximately 200 systematic-random fields for unbiased stereological analysis of the whole brain (100 per slide) and (2) 100 obligatory fields representing specific parts of the CNS. Automatic and semi-automatic methods of acquiring these QT images are taken up in Projects 2 and 3.

Interface Segmented Vector Graphics (Project 3) onto MBL Images

Segmentation of images will occur primarily on the high-resolution single section images. Upon completion of segmentation, we will generate overlays onto section images. This type of overlay has been done for the C57BL/6J atlas <nervenet.org/mbl/mbl_main/atlas170.html>. Overlays will include the ability to choose whether to have text appear as well as the boundaries of the ROI. We will provide the user with the ability to call up individual sections from different brains based on the ROI. For example, one could click on the rostral tip of the lateral geniculate nucleus and then download section images from 4 cases at a time allowing comparison among them. In addition, we will allow for searches based on segmented ROIs.

Interface the Neurogenetics Tool Box

We will provide a direct interface to the Neurogenetics Toolbox. We envision that for the early portions of the grant, data manipulation will take place off-line, but we will make it easy for users to share their data with the neurogenetics community and for others to access this data.

Allow Custom Downloads of 12-Bit and Color Images

As mentions above, images in the MBL are now acquired at 12-bit depth, but are converted and saved as 8-bit images. Some clients may need greater bit depth, perhaps even color, in order to address research questions. We will implement archiving of 12-bit gray and 24-bit color images. MBL clients will be able to request these images. These images could be provided on an FTP site and, if web bandwidth permits, eventually provided online like our current 8-bit images. At the very least, we will be able to provide the ability to order these images on CD-ROM or DVD-RAM on-line.

Allow Creation of Customized CD-ROM (or DVD-RAM) of Images

The MBL is designed as an on-line resource, and we anticipate that for most researchers it will be used as such. For some potential users, however, there may be technical (low bandwidth on their network) or other reasons for wanting to work off-line. We will provide the ability to custom-order images, QT movies, segmentation information, and Map Manager software. Ordering of the images and burning of the CD-ROMs will be completely automated.

Potential Pitfalls and Problems

As mentioned above in the context of the redesign of the database, attempting to see into the future of the web is often a risky business. We have outlined here a number of ideas for how to improve and extend the current interface of the MBL, all of which are eminently achievable with technologies currently on hand. We are prepared, however, to modify these plans to take into account advances in web design and programming tools that are likely over the course of the proposal.

Integration of new information into the MBL will require close collaboration with the projects providing the new information. This is especially true for the integration of the QT movies from the iScope and the segmentation information from NeuroCartographer. It is for that reason that Williams (Project 2) and Nissanov (Project 3) are explicitly included in the professional staff of this project. The MBL itself demonstrates that Williams and Rosen have the ability to work in a close collaborative relationship. Drs. Rosen, Nissanov, and Williams have recently collaborated to create the 3-D mouse brain atlas detailed in Project 3.

4. Image analysis

The previous three goals of the experimental plan involved either enhancing or redesigning aspects of the MBL that were currently in place (Phase II). This goal, in contrast, focuses on the creation of a collection of novel functions to allow researchers to manipulate and analyze images online (Phase I).

The scientific community has been blessed with a pair of open-source image manipulation and analysis programs written by Wayne Rasband—NIH Image and Image/J. The latter is written in Java and it, along with its source code, will provide the foundation for all of the image manipulation and analysis tools that we will create. We have discussed our project with Dr. Rasband, and he has agreed to informally aid us by modifying Image/J to make it work well with the images on the MBL Below are some of the functions that we will implement

Image Manipulation

Among the common manipulations of images are contrast enhancement, sharpening, and brightness. Depending on the size of the image, these types of manipulations can be quite computationally intensive. One of the advantages of FlashPix technology is that manipulations such as these require less processing power and are therefore easier to imagine being performed over the Internet. Should we decide not to use FlashPix technology, we will consider the cost in terms of bandwidth and computing power to provide these types of manipulations and perhaps provide the tools to perform these manipulations locally.

Segmentation of Regions of Interest

As discussed in project 3, NeuroCartographer will do automatic segmenting of a variety of CNS structures. Dr. Nissanov will also provide tools for users to modify these automatically segmented structures, and these changes will propagate throughout the MBL. It is likely, however, that there will be some structures that are not easily segmented by the algorithms designed in Project 3, and we will therefore provide users of the MBL the ability to manually segment regions of interest.

Having these regions segmented, the next logical step will be to measure these segmented regions. Some of this will be done automatically during the process of segmentation, but we will also provide the tools to estimate volume of CNS regions using Cavalieri's rule. There are a number of macro routines that run in NIH Image that do this, and we will adapt these for Image/J.

Neuronal Counts

The iScope will provide high-quality through-focus QT movies of 100–200 fields in each brain. To facilitate the counting of these cells, we will adapt Gary Leydon's "Count Box" program for online usage. As detailed in Project 2, Dr. Leydon will be a consultant on this program project and will work with us to provide this important analysis tool.

Densitometry

Densitometry of Nissl-stained sections has been shown to be useful in architectonic parcellation of the cortex (Zilles et al. 1980; Zilles 1985). We anticipate that future additions to the MBL may well include autoradiographic or other images where densitometry is necessary for conducting analysis. We will allow densitometric measures within automatically designated segmented regions as well as those that are user-defined.

Hooks to other image analysis packages

We recognize that not all users will prefer to perform their analysis online. In addition, they may have established protocols for image analysis locally that are essential for their own research designs. We will therefore create hooks to allow users to directly access images from the MBL for analysis by their preferred software package (i.e., NIH Image, NeuroZoom).

Potential Pitfalls and Problems

The benefits of Java technology on the web lie mainly in its "machine agnostic" nature. By providing tools for image analysis that are platform independent, we increase the potential utility of the MBL enormously. The promise of Java, however, sometimes has exceeded its abilities. Performance issues are real concern, and badly written programs can execute slowly, even over a high-speed network. Although in an ideal world, Java programs should operate identically on all platforms, this is not always the case.

We will extensively test our software on all major platforms (PC, UNIX, Macintosh) to ensure its compatibility. During the development process, we will institute rigorous testing procedures and will provide simple and direct methods for users to report bugs and incompatibilities. By making all programs open source, we increase our chances of optimizing code; talented individuals in the community at large will be able to make enhancements and additions.

Time Line

Year 1

- Begin processing additional RI strains. Complete set of standard inbred strains.
- Begin redesign of the database structures using FileMaker Pro. This should be completed in 6–9 months.
- After deciding on the structure of several of the databases associated with the MBL, begin to acquire (and reacquire if necessary) images and other data categories (slide coordinates of ROIs in particular). This process will be ongoing for the duration of the project.
- Explore options for high-level image warehouse systems (e.g., Oracle 8i).
- Plan suite of image analysis and image enhancement tools.

Year 2

- Begin processing G_{10} brains. This should be completed by the end of Year 3.
- Revamp the interface, mostly to incorporate the data from Projects 2 and 3.
- Continue development of image database system.
- Test one or more high-level image warehouse systems.
- Begin implementing analysis tools.
- Acquire 1 µm/pixel images of ROIs.

Year 3

- Finish processing of G_{10} brains. This should be completed by the end of Year 3.
- Add through-focus series to the MBL.
- Continued development of image database system.
- Implement new high-level image warehouse systems in parallel to existing system (test only).
- Overlay coordinates of through-focus series on section and ROI images.
- Add segmentation interface between MBL and NeuroCartographer.
- Web alpha testing of MBL image analysis Java applets.
- Acquire 1 µm/pixel images of ROIs.

Year 4

- Add brains of 100–200 aged G_{10} brains to the MBL.
- Possible conversion of image database system to robust commercial DB such as Oracle 8i or other image warehouse DB.
- Web beta testing of MBL image analysis Java applets.
- Complete all standard and RI strains in the MBL.
- Acquire 1 µm/pixel images of ROIs.

Year 5

- Finish development on final MBL image database system.
- Near final testing of web MBL image analysis Java applets.
- Consider additional to MBL of a new set of immunohistochemical strains.
- Finish acquisition of 1 µm/pixel images of ROIs.

RESEARCH PLAN

PROJECT 2: INTERNET MICROSCOPY SYSTEMS (iSCOPE)

PROBLEMS AND ISSUES ADDRESSED

In this project we request funds to continue to develop resources and techniques for high-magnification quantitative analysis of the Mouse Brain Library collection. All four aims of this project are being actively pursued, and our strong Phase I progress and preliminary results justify long-term support for this Neuroinformatics Phase II project. Accomplishing the aims listed in this Project will make a collection of several thousand sectioned brains from genetically defined mice available to the entire biomedical research community for collaborative and independent research.

The microscope systems we are developing will have extremely widespread applicability in neuroscience and biology. We believe this project presents a new paradigm for generating and accessing biological reference collections. The bioinformatics challenges of designing and assembling this complex system are substantial, but the team of investigators (Williams, Park, and Nissanov) includes the right combination of expertise in video microscopy, database design, stereology, and image analysis to make this work a practical five-year objective.

Aim 1: iScope Microscope Design and Construction

We will assemble four new video microscopes over this five-year grant. Funds for equipment amount to less than the cost of a single commercial confocal system, and like our Phase I prototype, all of the resources we produce will be open without restriction or delay to the neuroscience community for beta testing and active research. We will be constructing microscopes for two different experimental purposes. First, two microscopes will be optimized for uninterrupted and automatic acquisition of through-focus image stacks (optical sections). These stacks will be saved as short video clips (40 to 60 frames) that will become a permanent part of the MBL collection (see *Aim 4*). Second, we will construct two microscopes based on inverted microscope stands that will be ideal for real-time streaming video. These microscopes will be interfaced with QuickTime4 streaming video servers as well as to the robotic slide handlers described in *Aim 2*. Using this system of hardware and software, investigators will be able to roam through the collection, perform real-time analysis and comparison of regions that have not been prerecorded, and verify the position and the source of images in the MBL At least one of these two microscopes will always be available for online streaming video delivery (*Aim 4*).

Aim 2: Robotic Slide-Handling System

Dr. Nissanov and colleagues at Drexel University have spent the past two years designing and building a unique and effective slide-handling system that works in conjunction with a motorized microscope stage. This slide-handling system consists a set of slide carousels, each of which can accommodate ~50 2 x 3 inch glass slides. Slides are delivered from the carousel to a transfer belt that reliably loads slides onto the microscope stage. An optical feedback system ensures that slides are mounted precisely. In Year 01 we will build a second, larger and more refined carousel system and motorized X-Y stages to work in conjunction with one of the inverted microscopes described in *Aim 1*. This system will allow us to open a large part of the MBL to online high-magnification analysis. Drs. Rosen, Nissanov, and Williams will collaborate closely in designing a *Slide-and-Coordinate* database (described in *Project 1*). Our design goal is to deliver web microscopists to any site on any slide in the MBL with a precision of ±200 µm. Our long-term goal is to achieve an accuracy of ±50 µm.

Late in Year 02 the MBL will have expanded to more than 5000 slides, and we will be ready to experiment with a higher-capacity and more flexible slide-handling system. For this reason we request funds to adapt a commercial robot arm used in DNA automation laboratories to deliver slides to any of four microscope stations and from an arbitrarily large collection of slides. As part of this project, we will develop a novel slide-mounting system. Two to four slides will be mounted in rigid plastic holders. The holders will improve robotic handling, slide delivery, and registration of slides on the stage; in addition, we can apply machine-readable bar codes to the holders to facilitate machine reading and sorting of the MBL slide collection. Our goal is to assemble a system that can be extended from the 1200 slides now in the MBL to a collection of more than 10,000 slides. The slide system will include a master rack that securely holds the majority of the collection and a cache rack for recently and frequently accessed slides. Our target is to load slides held in the cache in less than 1 minute and to load slides held in the main rack in less than 2 minutes.

Aim 3: A web archive of image stacks and software for stack analysis

We will develop an automated system to acquire large archives of through-focus image stacks. This system will ultimately free users from the microscopes, because most of the high-magnification images they want will already be archived. In

combination with software we are writing, these stacks will be ideal for quantitative stereological analysis of CNS cell populations. Online microscopists will be able to obtain high-magnification images directly from the MBL. Our plan is to create a system that will allow us and other research groups to generate virtual brain tissue that can be stored digitally, transferred without loss, and analyzed both on- and off-line. We have three immediate goals:

1. To develop batch processing methods to acquire and process through-focus series. Accomplishing this requires integration of databases, coordinate registration and verification, and the simultaneous computer control of a robot, a microscope, and a digital video camera. This is a challenging bioinformatic and hardware problem.
2. To exploit this batch image acquisition system in order to automatically acquire sets of as many as 400 stacks from each case in the MBL. We will design an interface that will allow neuroscientists to order the microscopes to image particular coordinates on particular slides at high magnification and automatically generate through-focus series. These stacks will then be added to the MBL and will be integrated into the NeuroCartographer Project. Our goal is to make it possible for a scientist to order sets of systematic high-magnification sample sites from any slide or structure in the MBL collection.
3. To adapt existing open-code software (including the Leydon-Williams Counting Box program, MacMeasure, Wayne Rasband's Image/J, and Melburn Park's deconvolution programs) so that clients can use modern stereological methods to study through-focus series. Our stack analysis programs will be provided as open-source Java applet and compiled C programs. They will permit the simultaneous viewing and analysis of multiple stacks.

Aim 4: Streaming video

Streaming video technology is developing extraordinarily rapidly. In Year 01 we will apply Apple Computer's QuickTime 4 (QT4) streaming server technology to the existing iScope. The microscope is already in place with all of the necessary computer hardware and software. The QT4 architecture has an impressive feature set, and we will be able to simultaneously deliver streams of color video at multiple resolutions. The QT4 streaming system will allow us to deliver at least 10 frames per second (fps) at a resolution of 160 x 120 pixels, even to clients on 56 Kbps connections. At the other end of the spectrum, the full digital video camera output amounts to 3.6 MB/sec, and this bandwidth will become widely available in the next decade. We expect to eventually serve the original digital video from our archive or directly from the cameras. The main advantage of streaming video is that Internet microscopists will be able to use the collection in a more natural way than would otherwise be possible. For example, they will be able to record their own series or montages of images from any brain in the MBL; they can search for rare features such as dying neurons or mitotic figures; or they can define the precise coordinates of sites they would like us to acquire on their behalf using our rapid through-focus image acquisition batch processing system.

In Aims 3 and 4 we will provide two types of video: live streaming video in color and prerecorded compressed progressive video in both color and 8-bit monochrome. The gamma of video images will be customized for both Windows and Macintosh operating systems. Two streaming servers will be put on the UT Memphis fiber optic backbone using 100BaseT or a direct fiber optic connection (see appendix letters). Progressive/prerecorded video will be integrated into the MBL library and handled by conventional Apache web servers running on Macintosh G4 computers. The slide coordinates of prerecorded high-magnification movies will be marked on the low-magnification slides in the MBL, and simple clicking on a marked coordinate spot in the MBL will open the corresponding QT4 through-focus series.

Letters in the appendix emphasize the enthusiasm of our research and network administrators at UT Memphis for the Internet microscopy lab. They view this project as a powerful teaching tool, and they guarantee their full cooperation in ensuring that our servers will have optimized connections to the Internet. They have also offered new space and substantial funds for the duration of the grant to purchase and upgrade computer equipment and software.

OPPORTUNITIES

In this section we highlight the promise and the problems associated with building a robust real-time Internet microscope system. The promise of Internet microscopy is in some respects obvious: such systems can make a large and unique collections of tissue available to investigators without the standard limits of time and space. A system of this type is needed for the following purposes.

Creating digital slide collections

Investigators accumulate vast and often irreplaceable histological collections (e.g., the Scheibels' collection of Golgi material, the Altmans' series of thymidine-labeled material, the primate brain collections of Rakic and Goldman-Rakic, Paul Yakovlev's extensive collection of adult and fetal human brains, the Brain Biodiversity Bank started by Woolsey and continued by Wally Welker). The list could go on for pages. Almost every one of us accumulates some tissue that is

unique. For example, I have spent over a decade studying a line of mutant Belgian sheepdogs that completely lack an optic chiasm. The anatomical repercussions of this mutation are without precedent (Williams et al. 1994; Hogan et al. 1999). What is the long-term fate of this sort of collection? Often a dumpster. The Yakovlev collection narrowly escaped such a fate and was fortunately placed in a permanent collection maintained by the Armed Forces Institute of Pathology.

Key parts of these collections can now be saved by being digitized. The problem of curating digital archives can ultimately be reduced to finding the right digital medium and the most effective database structure and means of distribution. Only a few years ago, digitizing entire collections would have been totally impractical, but in the past year the cost of digital storage has dropped precipitously. One terabyte—1000 gigabytes (GB) of hard drive space— is enough capacity to accommodate over 300,000 extremely high-resolution images that closely match the quality of 35-mm ASA 64 slide film (3000 x 2000 x 24-bit images). Today this amount of digital storage space can be purchased for about $8000, and the cost is likely to drop further. The storage cost of a single archived high-resolution image should drop to a quarter of a cent.

Even at current hardware costs, the important bottleneck is not storage but limitations in the technology used to acquire and process digital images. If a technician making $10–15 an hour takes about 2 minutes to capture and process an image, then labor costs more than 100 times as much as storage. We need to develop methods for rapidly and automatically acquiring batches of high-resolution images from existing collections. A major motivation of the iScope project has therefore been to develop methods to convert collections of glass slides to their digital counterparts quickly and with minimal human intervention (*Aim 4*).

We have made an excellent start in archiving, distributing and exploiting low-power images from the MBL. These images are extremely useful. For example, two colleagues of mine are now mapping QTLs that modulate the size of the hippocampus (Lu et al. 1999) and the cerebellum (Airey et al. 1999). Their computer workstations are located less than 20 feet from where the MBL slides are stored in a room full of microscopes. But rather than using the slides and microscopes, Drs. Lu and Airey find it far easier to download the 25- and 4.5-µm/pixel images from the web site. They load these images into NIH Image, outline boundaries, and transfer the data into Excel worksheets.

Low-power images, however, are not sufficient. We would like to be able to zoom into any part of a slide or section, from the equivalent of a 1x objective to the equivalent of a 100x objective. Ideally all the images, or virtual tissue, would be available immediately from a digital collection. This is feasible for low and intermediate scales, from 25 µm to 2.5 µm per pixel. But high-magnification images with pixel dimensions of under 0.5 µm are more challenging to deliver. Aim 4 of this project describes how we intend to use the iScopes to automatically and systematically acquire high-resolution image stacks for several hundred sites per brain.

Collaborative research

Email, the Internet, and courier services have transformed the way scientists collaborate. These innovations are fostering new trans-institutional collaborations. It is now possible to assemble the best team to carry out a complex multidisciplinary research project from among universities and corporations scattered widely across the globe. The iScope is just one more tool among many that will increase the ability of groups to collaborate effectively across institutional boundaries. This resource will be accessible to neuroscientists around the world without restriction or tracking. The only limitation will be that we may have more users than our microscopes or bandwidth will be able to accommodate. We have already set up a simple but effective arbitration scheme that schedules microscope users by category on a first come, first served basis.

Efficient use of animal resources

A collection such as the MBL, which already contains samples from ~100 genetically defined strains of mice, increases the efficiency with which neurogenetic problems can be addressed. For example, a scientist investigating the effects of age on the volume of the mouse neocortex can explore and exploit the MBL collection rather than ordering and then aging a set of mice and then processing all of their brains. In a matter of a few days, this investigator can mine the MBL and obtain excellent data for animals ranging in age from one month to two years. Information about many other factors, such as sex differences in the volume of the amygdala or hippocampus of mice, is likely to be embedded in the MBL's superb collection of celloidin sections. The answers to scientific questions can be obtained quickly without sacrificing a single additional mouse, and without the huge cost of processing hundreds of brains.

Teaching neuroscience and neurogenetics

Teaching neuroscience will be transformed over the next decade by Internet resources. Not only will publications be available on the Internet, but the datasets and analytic tools that are critical in carrying out research will become more accessible. For example, tutorials and resources already available on the <nervenet.org> site could be used as part of a course in quantitative neuroanatomy and neurogenetics. The iScope and MBL in particular would be useful adjuncts in many neuroanatomy courses.

CURRENT STATUS

At its inception the MBL was intended to be a standard library: investigators would request small shipments of slides, which they would return promptly after analysis and photography. Reality intervened: the first shipment sent out was damaged during transit; backup slides were sent, but after more than a year none of these slides has yet been returned. This may have been good fortune. We were motivated to look for a better way to organize and distribute the collection. A microscope rescued from surplus became the stand for our first Internet microscope (soon shortened to the iScope). Over the past year, we have made substantial progress in developing the iScope. By far the best way to assess our progress is to test-drive the system at <nervenet.org/mbl/mbl.html>; however, because the iScope is still in development, it is not online at all times.

The present iScope is a modified upright Zeiss Universal microscope equipped with a x40 achromatic DIC objective and a MiniDV digital camcorder (Canon XL-1). The iScope has been operating continuously without hardware failure (or even a bulb change) since August 20, 1999. Although the iScope has broad applicability in research and education, in the context of our research program it is intended primarily as a high-magnification extension of the MBL, and this is where the neuroinformatics challenge lies. When the project is complete, all sections in the collection will be available for online analysis. Neuroscientists will be able to obtain streaming video at sufficient resolution (320 x 240 pixels at 10 fps, or 640 x 480 pixels at 2.5 fps) to navigate effectively through the collection. This virtual microscopy lab will be tightly integrated with image databases that are part of the MBL and NeuroCartographer projects. In conjunction with the MBL, investigators can select areas they want to examine in greater detail, then zoom in from low-resolution images of entire slides, to medium-resolution images of individual sections, to high-resolution through-focus image stacks of neurons and glial cells.

Modern stereology and the iScope project

We intend to implement an automated process that will generate a large collection of through-focus Z-axis image stacks of each brain in the MBL at very high magnification (<0.8 µm resolution). Each prerecorded stack will require less than 1 MB of storage space. These images will be ideal for unbiased stereological analysis using either direct 3D counting or the optical disector. Two microscopes (one for development, one for production) will be used to generate the stacks. These video clips will be ideal for the high-magnification cell counts that have become a keystone of unbiased stereology (Williams and Rakic 1988a, b). All high-resolution prerecorded material will also be available on our web and ftp sites and on DV-RAM disks.

Novelty of the neuroinformatics component

What is the bioinformatics challenge of Internet microscopy that we will address in this project? There are a significant number of informatics, database, and connectivity bandwidth problems associated with effectively controlling a microscope over the Internet. We will come back to these shortly. One of the major challenge is in assembling an entire suite of applications and images that allows a neuroscientist to sweep through a large collection of slides from the lowest power to the highest power without actually having physical access to either the slides or a microscope. In our Phase I research, we have demonstrated competence to deal with the low- and high-power images, but the informatics challenge now is to bridge the entire range of magnifications using two complementary approaches—real-time imaging and analysis of prerecorded high-magnification virtual tissue. By virtual tissue we mean 3D stacks of images that when played back on a computer give neuroscientists the same feel for the tissue that they would have in turning the focus control at the microscope.

Limitations and problems

What about the problems associated with Internet microscopy? One problem at present is that the effectiveness of virtual microscopy depends on the bandwidth of connections between our servers and our clients. Aim 4 is therefore hostage to the quality of Internet connections. We have partial control of this factor, and we will need to ensure that our streaming servers can handle the load of requests. But our clients will probably be scattered all over the world, and their Internet connections will often be woefully slow. Thus, we have to design robust interfaces for the iScope and for all other parts of this program. Clients will be able to do useful work even over a slow connection—a modem creaking along at less than 30 Kbps. This will be a challenge for the image-rich data types that we are providing, though we have used the MBL effectively over a 56 Kbps modem connection. The process is slow, but the initial steps permit the investigator to sort and search for just the right strain, mouse, slide, or section before actually downloading a 500 KB file. Rapid advances in digital communication and the spread of high-speed Internet connections will gradually (or perhaps abruptly) increase the usefulness of the iScope.

Precedents

Telepresence microscopy has been used on several occasions to operate sophisticated high-voltage electron microscopes (see Zaluzec 1995, at Argonne National Laboratory; Parvin et al. 1995, at the Lawrence Berkeley Laboratory; Fan and

Ellisman 1993, 1995, at the University of California San Diego). These efforts have been motivated by the desire to make these extremely powerful microscopes more available to the research community.

One recent experiment in telepresence microscopy over the Internet involved teaching students at Lehigh University how to use a transmission electron microscope situated at the Oak Ridge National Laboratory (Voelkl et al. 1997). Although the microscopes involved are very different, the Oak Ridge project and our own share many common aspects. Parallel video and audio communication between the two sites was maintained using a Connectix digital camera and CU-SeeMe software. Two parallel video links established between the sites allowed an instructor to explain the use of the electron microscope while simultaneously showing images generated by the microscope. This is essentially the same system we have used to set up a laboratory camera that images our microscopy laboratory at UT Memphis (see <labcam.nervenet.org>)

As pointed out by Voelkl and colleagues, the main drawback in the Oak Ridge experiment was the limited bandwidth of the Internet. An uncompressed 256 x 256 pixel, 8-bit monochrome video running at 30 fps requires a bandwidth of more than 1 Mbps, or at least 20 times faster than a 56 Kbps modem. Fast JPEG compression can achieve at least a 5-fold reduction, and the use of more sophisticated compression methods that compress across video frames (temporal compression) can result in an even greater efficiency. For our purposes, we do not expect bandwidth to be a major problem. The point of doing of real-time streaming video is to navigate through the image and find the right spot for detailed analysis. Once the point of interest has been located, a slow but high-quality video feed can be ordered. In our opinion, the ideal system for real-time video microscopy will combine a clever interface with the ability to rapidly switch between resolution modes.

Light microscopes on the Internet

Over the past two years, a number of groups have begun to experiment with motorized light microscopes on the Internet for collaborative projects and to market microscopes. The current implementation of the iScope does not differ greatly from several of these efforts, and when possible we hope to collaborate and learn from these other projects.

1. Leica Microsystems Imaging Solutions Ltd. (Cambridge, England) has had a motorized Leica RXA transmitted light microscope online for at least 18 months at <www.leica.co.uk/microscope/>. This is a true Internet microscope available to anyone.
2. The Microsystems Technology Laboratory at MIT has assembled a system for collaborative inspection of semiconductor wafers over the Internet. Their microscope, which been in operation since 1996, is part of a collaboration between Stanford, MIT, and several other institutions. This system is available only to registered users.
3. The Corporation for National Research Initiatives (Reston, VA) has undertaken a project similar to our own. They give unrestricted access to a light microscope and have also made Java code available for use by other groups interested in Internet microscopy at <sal.kachinatech.com/z/2/REMOTE-MICROSCOPE.html>. This effort also focuses on collaborative wafer inspection.
4. A group in Germany has a particularly successful Internet microscope implementation at <amba.charite.de/telemic/intro1.html>. Most of the major features of a transmitted light microscope can be controlled. Their interface is similar to that of the iScope.
5. Manchester University Materials Science Center has assembled a series of digital micrographs similar to the MBL, but with emphasis on metallurgical specimens. The image database, with controls that allow a user to manipulate the images, is at <www.umist.ac.uk/~intmic/>.

Unique aspects of the iScope

The iScope differs from these earlier efforts in some important ways.

1. In implementation, the iScope microscopes will have a relatively simple configuration with one fixed objective and only two magnifications.
2. The iScopes will be designed specifically for high-magnification quantitative analysis. They will complement the low- and medium-power images in the MBL.
3. The iScopes will be interfaced with slide-handling equipment that will allow a large collection of brain tissue slides to be accessed over the Internet.

Compromises and cost

The design of the microscope and video equipment is inevitably a compromise between cost, ease of implementation, and image quality. For example, use of oil-immersion objectives with high numerical aperture would improve resolution from 0.8 to 0.25 µm, but for obvious reasons, oil immersion is not practical for an automated slide-handling system. We believe that for streaming video and even for acquiring stacks of images for 3D reconstruction, our best option is a dry objective with a moderately high numerical aperture (NA of 0.6–0.8) and a moderately long working distance. We now use an

objective that can tolerate some variation in cover glass thickness without suffering serious spherical aberration. An alternative would be to use microscopes with motorized nosepieces and multiple objectives, but besides being far more expensive, these options would complicate the implementation without greatly extending the utility of the MBL. To provide a broader range of magnification, we will use a beam splitter and two video cameras each equipped with a different intermediate lens on microscopes designed for streaming real-time video. We can readily achieve a 3:1 ratio of magnification from two cameras using a single 40x objective. With this equipment, microscopists will have access to both high- and low-magnification video streams simultaneously, and they will of course have access to the 25 and 4.5 µm images for all cases in the library.

Streaming video

There are numerous special requirements for effective streaming microscopy on the web. Not only must the video system provide a rapid update from the camera, but the web microscopist needs to be able to rapidly control the movement of the stage. Achieving fluid two-way communication is a challenge. For example, in our current implementation, the duty cycle of the system—that is, the time between a request to move the stage and the time at which the video is actually refreshed on the client's web browser—is approximately 10–20 seconds, even on an in-house connection. With streaming video, this lag will disappear. The current system is a standard web server that orders new video frames from the camera at 10-second intervals. For true streaming video, faster low-resolution video is preferable to slow high-resolution video during the navigational stage. A relatively large field of view is also preferable to a small high-magnification view. For this reason, each microscope used for streaming video will be equipped with two video ports and two digital cameras, each with its own magnification. One camera will cover a large-field view optimal for navigation. The other will have the narrow field of view of the present system, optimal for analysis and image capture.

Future directions

By the end of Year 03, the iScope will be a means for creating large digital reference collections of blocks of neuronal or other tissue captured at high magnification. We intend to test and refine image processing methods that will automatically perform optical segmentation of neuronal somata. Applied to the digitized reference collections, the segmentation process will produce what are essentially digital representations of cell packing and conformation. Applied to differentially stained material, we should be able to generate libraries representing the cellular architecture of different parts of the mouse CNS. This work has clear application to the human brain project. We recognize, as well, that a fully featured implementation of iScope could have value to laboratories wishing to obtain digital reference collections of their own material.

EXPERIMENTAL PLAN

Duration

The experiments, tests, and development proposed in this application will take five years. Because the iScope project is not hypothesis-driven, this section of the project appears to have an open-ended character. However, we have specific and well-defined aims, and given our accomplishments to date, we believe that within five years we can refine and apply the technology for highly effective Internet microscopy. Furthermore, this is enough time to gather a substantial collection of through-focus image stacks. These stacks will be a major component of the MBL and will make available an excellent collection of stereological material from hundreds of strains of mice. In addition, we will make our very large QTL mapping cross available to the genetics, neurogenetics, and neuroscience communities.

We will build four new iScope systems during the first four years of this grant. The first two microscopes will explore important different hardware options, foremost, the utility of inverted versus fixed-stage upright microscope stands.

Aim 1: iScope Microscope Design and Assembly

General comments on timing

In Year 01 of the grant, we will build two new but relatively inexpensive Internet microscope systems with ample free space above the specimen. Drs. Williams, Park, and Nissanov will work on integrating the carousel or jukebox slide-handling system (*Aim 2*) with the microscopes. One system will be based on an inverted microscope, the other either on another inverted stand or on a fixed-stage upright microscope. In Year 03 we will acquire two additional microscope stands with the aim of eventually putting five fully configured iScopes on the web. This is a realistic schedule—the first

iScope was assembled over a two-week period. The major delay was in having custom adapters made for the Canon XL-1 camera.

Cost

Most research microscopes are now, in our opinion, overpriced and over-equipped. We will continue to do our best to provide top-notch image quality on a relatively modest budget. We will design and construct custom microscope parts and video adapters. When practical, and when the price and quality are right, we will purchase used microscope stands, in particular older-generation Zeiss ICM and Universal stands. We have already purchased most of the optics used by Zeiss Axioplan microscopes, and we will therefore buy at least one of these microscope stands. They have an extraordinarily precise focusing mechanism based on a "harmonic" drive that has almost no mechanical backlash. Our major expenses will be motorized stages, optics (condenser, an objective, and Nomarski prisms), digital video cameras, and fast microcomputers. These "peripherals" are actually the key pieces of equipment for this project. Collectively these parts cost approximately $25,000 for our first iScope. We can purchase stable microscope stands with good focusing blocks for about $10,000. Thus, over the course of this grant we expect to spend approximately $140,000 on the microscopes—less than the cost of a single confocal microscope system. All four microscopes, along with our exiting iScope, will remain online after the end of funding.

iScope Feature Set

The following is a short list of key features of the iScopes.

1. For streaming video by web microscopists, we are leaning toward use of a single x40 dry objective (NA 0.65) with proven optical resolution of 0.8 µm and pixel resolution as high as 0.1 µm/pixel. We will experiment with higher-NA dry and water immersion objectives for microscopes intended specifically to acquire very high resolution Z-axis image stacks. The higher the NA, however, the more difficult the implementation and the greater the maintenance problems.
2. For streaming video, two magnifications will be provided simultaneously via dual 3-CCD DV cameras, one with a field of view of 60 x 80 µm, the second with a field of 200 x 300 µm.
3. DC servo motorized stage and motorized Z axis adapted for the carousel slide delivery system.
4. Differential interference contrast optics.
6. Autofocus.
5. Alternative light sources: a low-voltage tungsten bulb with a very extended life for streaming web microscopy and a high-pressure Xenon-Mercury system (1000 h bulb life) for acquiring Z-axis image stacks.

Justification for DIC optics of Nissl-stained tissue

Standard brightfield microscopy of Nissl-stained specimens primarily conveys information about the optical density of the strain. DIC optics reveals differences in the refractive properties of the tissue. DIC optics of stained tissue is thus a multimodal imaging method and as such provides more information about the tissue (Farkas et al. 1995; Glasbey and Martin 1996) than simple brightfield microscopy. A number of other advantages of DIC microscopy have been enumerated by Williams and Rakic (1988a, b). One of the most important advantages has to do with the unique modulation transfer function of DIC images (Inoué, 1986). In short, DIC optics enhances contrast of high-spatial-frequency image content and suppresses contrast of low-spatial-frequency components of the image; basically, DIC acts as a high-pass optical filter and edge enhancer. Furthermore, the depth of field of the high spatial frequencies is remarkably shallow, making it possible to optically section DIC through-focus series far more effectively than corresponding non-DIC images. DIC can be thought of as a pseudo-confocal system in which the out-of-focus blur from tissue above and below the focal plane is significantly attenuated (Oldenbourg et al. 1993; Hibbard et al. 1996).

Autofocusing the microscope

Developing efficient autofocusing methods for the iScope is critical both when the microscopes are being controlled during a streaming video session and during the capture of through-focus series (Aims 3 and 4 below). Focus would not be quite so critical if the slides were not so large and if the sections were pressed perfectly flat to the slide surface. Unfortunately, the difference in the z-axis coordinate of the top of different sections on a single slide can be up to 100 µm. Thus, without automatic z-axis focusing it will be very difficult for our clients to remain in focus when leaping from one section to the next. What are some possible solutions? Connecting the XL-1 camera's internal autofocusing system with our motorized fine-focus may be an inexpensive and elegant solution. This is particularly practical with the Canon XL-1 because all lens controls, including those that drive the automatic lens focus, are readily accessible (the autofocus lens dismounts on this camcorder). Focus in the XL-1 is driven by the image quality itself, not by an infrared system. John Zemek, president of Applied Scientific Imaging, has offered to help us test the feasibility of this approach. A standard fallback is to use image analysis utility programs to access image quality and independently dive the z-axis. For example, Wu et al. (1996) at Mt. Sinai School of Medicine have developed efficient methods to drive an autofocus mechanism. More than 10 autofocusing algorithms that we will consider are succinctly reviewed by Santos et al. (1997).

One of the advantages of DIC optics is that even in completely unstained parts of the section there will be ample phase contrast to generate the high-spatial-frequency signals of a well-focused image. When the image is not over tissue but instead is over a blank part of the slide, we do not want the focusing mechanism to hunt forever for focus or to focus on the top of the cover glass. One way to prevent these problems is to force the focus control through one cycle of the maximum z-axis travel permitted for the stage—say ± 200 µm—and if no z coordinate with adequate focus is detected, rest the microscope at the last focused z-axis coordinate. If the lag in control of the fine-focus can be reduced to a sufficiently short interval of < 1 second between command and response, then it may be easier simply to disengage the autofocus mechanisms entirely.

Z-axis calibration

The PI has spent a considerable amount of time working on this problem. There are several vertical calibration standards in the form of series of polymer films of very precisely known thickness. However, these films have refractive indices that differ from glass and oil (~ n = 1.52), adding complications to the vertical calibration process. Boddeke et al. (1997) have developed a simple and elegant method to calibrate the z-axis using a "tilted slide" preparation. The slide can be a standard calibration slide with a series of closely spaced parallel lines. Moving the tilted slide changes the position of the line that is in best focus as a function of the tangent of the tilt angle. Their protocols permit estimates of backlash in the focusing block, long-term stability, and absolute precision of focus.

Relocation accuracy

Several factors affect how accurately a given coordinate on any slide can be acquired and then reacquired. One factor is the adequacy of the database that describes the slide and the section coordinates. Assuming that the database is correct, then there are a series of physical factors that affect relocation accuracy: how precisely the slide is positioned on the stage, the consistency of encoding on different microscopes, etc. Tucker et al. (1994) achieved a remarkable relocation precision of less than 17 µm between two different microscopes and a precision of less than 7 µm on the same microscope. R. Williams has constructed an encoded stage using a trio of Heidenhain digital length gauges that can consistently reposition a manually placed slide to within a 5 µm radius. The ASI DC servomotor stage that is now fit to the iScope has a relocation accuracy of better than 40 µm even when the microscopist wanders all over the 2 x 3 inch slide prior to centering.

For the purposes of this research project, we will initially be satisfied if a client can select a point on one of our low-resolution slide images (25 µm/pixel) and then be delivered to that coordinate within 100 µm on any of the microscopes. This will require a database with an entry for each slide. This Slide-Coordinate database (described in Project 1) will be used to navigate from images in the MBL to points on particular slides. Slide registration on the microscope stage must be precise, and this registration will be improved by the use of the holders that will be designed with this problem in mind. We ultimately hope to attain relocation accuracy of better than 100 µm. The main purpose is to deliver users to single regions that were previously used to generate through-focus series. Relocation accuracy of ± 20 µm would be required to put the user in a field that overlaps the initial images. Having this level of precision will be useful to confirm and extend datasets generated using only MBL image resources. With the use of absolute slide coordinates and the ability to rezero each slide after it is loaded, we should be able to achieve good relocation precision.

Objective centering

A single objective will be used in each iScope, simplifying relocation accuracy because we will not have to worry about nosepiece position and centering multiple objectives. However, we will be using multiple iScopes, and it will be necessary to measure offsets of each objective on each microscope so that systematic bias in location among microscopes is minimized. Using a single objective per microscope obviously simplifies this procedure. Each microscope may require its own offset.

iScope use arbitration

The current implementation is a simple first come, first served system. Any user can work or play with the microscope for a 10-minute period. Registered scientists actively using the MBL and iScope will be able to sign in for periods of 2 hours or more, depending on demand. We do not expect access problems because two microscopes will be on reserve specifically to acquire high-power image stacks at locations requested by our clients. It should be possible to define these locations using the 4.5-µm pixel images. Thus, even heavy users should not need long periods of uninterrupted access to the microscopes.

Aim 2: Robotic slide-handling systems

Dr. Nissanov and colleagues at Drexel University have spent more than two years designing and building a unique and highly effective slide-handling system that works in conjunction with a motorized microscope stage. The prototype

robotic slide feeder operates with a very low rate of jamming (less than 1 jam/100 slides) and jams do not damage slides. A duplicate of this system, which we refer to as a jukebox, will be installed during year 01. It consists of a slide jukebox, a linear conveyor, and an automatic slide clamp. The system is designed to be compatible with a 3-axis-motorized stage on the microscope. Our goal is to deliver web microscopists to any site on any slide in the MBL with a precision of ±200 µm. The jukebox system will be used throughout the grant as part of our automatic z-axis image acquisition system.

The jukebox is reminiscent of carousels used in 35mm photographic slide projectors, except that the slides exit and reenter the carousel radially and the ejector mechanism exerts a force from the center outward. The carousel is mounted vertically on a central horizontal spindle. The stationary outside sleeve is fully enclosed except for a single slot while the internal sleeve has a central opening for the ejector arm. The spindle can rotate about its axis to position the desired slide behind the front slot so that an ejector arm from the spindle can push out the slide. Five carousels can be placed on the spindle; each accommodates 60 slides to give a total of 300 slide slots. Translation along the axis of the spindle selects the different carousels.

The ejector is a small lever, actuated by Nitinol (a NiTi shape memory alloy) that pushes the slide from the inner radius of the carousel outward. The linear conveyor (motorized rollers) draws the slide the rest of the way from the carousel and places it partially on the slide clamp, which is mounted on the imaging stage (microscope stage for microscopy). Another small Nitinol-actuated lever drives the slide the remaining distance, gently pressing it against the end stop pins of the slide clamp. A second lever than presses it into a fixed position. Three position sensors obtain accurate (resolution 5 µm) measures of the location of two of the slide edges from which the slide coordinate origin is computed.

After imaging, the slide is released. Another Nitinol lever mounted on the mounting frame ejects the slide, and the rollers draw it out of the clamp and drive it most of the way back into the carousel. A fourth Nitinol lever guarantees the full reinsertion of the slide into the carousel, and the cycle is complete.

A dedicated microcontroller board controls the jukebox, the conveyor, the clamp, and all the levers. A network of optical and other sensors will permit the microcontroller to monitor the progress of a slide through the system. A serial datalink allows commands and data to be exchanged with a desktop computer, allowing synchronization of the slide manipulation system with the image acquisition system.

By year 03 the MBL will have grown to more than 5000 slides. We will have gained a great deal of experience with the microscopes, the slide databases, and patterns of use. We will be ready to experiment with a higher-capacity and more flexible slide-handling system. Beginning in Year 02 a new robotic system will be designed and built that will accommodate as many as 20,000 slides. Instead of a carousel jukebox, the new system will consist of vertical racks of slides serviced by an ORCA robot arm designed for high-throughput DNA sequencing applications. Slides will be mounted in a plastic frame similar to a standard microtiter plate. The robot arm will ride on a 3-meter rail and retrieve the appropriate slide holder. Both slides in the holder will be locked into a fixed position on a large XYZ stage mounted under the microscope. Our target is to load slides in the cache in less than 1 minute and to load slides in the main rack in less than 2 minutes.

Aim 3: A web archive of image stacks and software for stereological analysis

We will develop a system to automatically acquire batches of through-focus series in the form of 1- to 2-second digital video clips. This process will involve systematically driving the stage to specific coordinates generated by the Slide-Coordinate database. Through-focus series with 1-µm z-axis step size will be acquired at each site in <10 seconds. Z-axis coordinates will be corrected for optical foreshortening (Williams and Rakic 1988). Using powerful commercial video processing programs and utilities (DVEdit, Media Cleaner Pro 4 Sorenson Video Pro) we will develop batch processing methods to collect and process several hundred z-axis stacks per day. Original digital video (DV) will consist of less than 60 frames at full-DV resolution and will be approximately 7.2 MB in size. These videos will be put on an ftp server. Two G4 computers running the asymmetric Sorenson Video codec (compression-decompression program) will compress clips to under 1 MB. These clips will be uploaded onto the MBL servers as QT4 movie files.

Once we have developed and extensively tested the system described above, we plan to acquire two complementary sets of z-axis stacks.

1. Starting in Year 02 we plan to acquire a systematic, random, and unbiased sample of through-focus series (~200 points/case) from every mouse brain in the MBL for advanced stereological analysis. This density of sampling is adequate for analysis of large structures such as the caudate, neocortex, and cerebellum. This online collection can be used to obtain precise estimates of total neuron and glial cell populations either by manual counting or by using automatic cell recognition programs. Higher-density sampling (~400 points/case) may be justified in some cases to allow even fairly small nuclei to be analyzed.
2. Starting in year 03, we plan to acquire stacks from a list of ~100 structures or regions defined semi-automatically by the NeuroCartographer project. As part of this project, slide coordinates of these regions

will be automatically generated during the segmentation process. Once the segmentation coordinates have been manually verified and adjusted, we will use these NeuroCartographer coordinates to generate a list of "greatest hits" for each brain.

Mechanics of acquiring through-focus series

Most microscope stages have appreciable mechanical backlash. For example, the Nikon Diaphot has a backlash of about 0.6 µm. The new Zeiss Harmonic drive and the older Leitz planatary gear focusing blocks have backlash that is almost negligible (<0.4 µm). The Zeiss Universal blocks are quite good and have a backlash estimated to be about 0.4 µm (Williams, unpublished). We plan to acquire all image stacks in the same sequence, focusing from the bottom to the top (driving the stage upward, or in a fixed-stage configuration, driving the objective down (see Boddeke et al. 1997). Long-term drift will not be a significant problem: each new slide will be automatically focused in the z axis. The client will be able to rezero the z as needed during a streaming video session.

Throughput

Once this system is in place it should be feasible to chose a specimen and a coordinate and then define the lower and upper focal planes of the region of interest. A click of the mouse should then initiate the capture of images and the construction of QT4 movies. We may need to run each frame through a series of Adobe PhotoShop 5 filters (unsharp mask, level adjust, discard color information, resize, etc.) prior to assembling the movie. These operations will be done automatically using the batch processing feature of PhotoShop 5. We expect to be able to generate several thousand QT movies per year.

Sampling considerations for iScope stacks

Systematic random sample is more appropriate and is unbiased and will generally produce a lower coefficient of error (SE/mean) than a random sample. A systematic sample of this type is referred to by Gundersen as a fractionator. Our system will make a two-stage systematic sample possible. In the first stage every nth section through the target is selected for stage 2 analysis. The second stage consists of sampling counting boxes that are systematically spaced in the target.

Sample size considerations

From the point of view of obtaining robust estimates that apply to an entire population or genotype of animals, the precision of individual estimates of cell number from a nucleus of a single animal should approximately match the variability of the genotype or strain of mice that are the subject of analysis (West and Gundersen 1990; Glaser and Wilson 1998). However, this latter value is generally not known in advance, nor is the magnitude of technical error generally well characterized. For this reason, the precision of individual estimates should by design be somewhat higher than the estimate of the standard deviation of the sample of animals. For example, if the number of cells in the brain of C57BL/6J animals has a standard deviation of ±10 million cells (a CV of 10%), then it would be reasonable to target our estimates to have a CE of ±5 million cells. Glaser and Wilson have demonstrated, not surprisingly, that the CE is an inverse function of the number of cells counted per case when the counting box or disector volume is held constant. For both practical and statistical reasons, at a fixed total count it is better to obtain a count from a higher density of smaller counting boxes than a lower density of large counting boxes (Williams and Rakic 1988). The Scheaffer-Mendenhall-Ott approach to computing the CE of a sample is summarized well by Glaser and Wilson (1998).

To obtain a high level of accuracy for whole brain stereology will require a minimum of 200 sample sites per brain. Furthermore, to obtain a robust estimate of the within-genotype variation we would like a reasonable number of samples. This is the reason our target number for the MBL is 12 animals per genotype or strain. For genetic crosses such as our tenth-generation advanced intercross, each individual mouse is genetically unique and a somewhat different sampling strategy needs to be employed (Williams et al. 1996). In essence all animals are initially analyzed using a low=density fractionator and outliers are then analyzed using higher-density fractionators.

Aim 4: Streaming video

Streaming video has become a standard technology used in the most mundane of Internet applications. We find that it answers specific needs in this project. It is essential for a truly interactive microscope implementation.

Introduction

Over the last three years computers and connections to the Internet have improved by at least an order of magnitude. A personal Road Runner cable modem connection <www.rr.com> of the type now available widely to home customers runs at ~3000 kbps (Hamilton, 1999)—approximately 54 times faster than a 56 Kbps modem. This is also twice the speed of a standard T1 connection that runs at 1500 Kbps. Peak download speeds with a cable modem can reach 7 Mbps. This is enough bandwidth for high-quality streaming video. In fact, cable modem speeds now often exceed the bandwidth of major university server connections, an odd and ironic reversal of fortune, and one that highlights the importance of

university network upgrades such as the Internet 2 initiative. Similar changes are occurring on the hardware and video fronts. Extremely fast computer and video accelerators are now routinely used to play complex and demanding interactive Internet video games. Many home computers are now more powerful than UNIX workstations that are only 1 or 2 years old. Computers equipped with 10–20 GB drives and CPUs capable of processing nearly a billion floating point operations per second are not on the horizon—they are already in the living room. A terabyte of hard disk capacity, enough to hold one million 1 MB images of the type we have put into the MBL, now costs just under $10,000 ($212 per 25 GB IBM DeskStar drive).

Even in a computer industry that is used to rapid advances, this is blistering progress. The opportunities in terms of video microscopy are enormous, and in our opinion the major challenge is not in developing new technology but in deciding on the best way to take advantage of what is already available. This is particularly the case when it comes to streaming video. In comparison to the equipment demands of a typical Internet game aficionado, what we propose to do in terms of streaming video and in terms of controlling three axes of movement of a single microscope is a simple task. The bottlenecks for the iScope project for the foreseeable future will not be disk space, computer hardware, or connectivity. The main problems will really center around our ability to process suitable tissue for imaging and then to photograph the tissue and put images and movies on a web server. We spent a great deal of effort assembling the images in the 1999 edition of the MBL. Compared to any other image databases of which we are aware, it is an extremely large and well structured collection.

In Year 01 we will move our current iScope onto a QuickTime 4 (QT4) streaming server. The present system is based around Reardon Technologies Java applet SiteCam software. This is not a true streaming technology: Camera images are captured from the Canon XL-1 videocamera and are immediately compressed using JPEG and then uploaded into the Java applets window. The applet itself is embedded into an html page that also includes rudimentary navigational controls (x,y,z movements). The response time of the system is slow. And in combination with our JavaScript code and a C++ program that controls stage position, it takes on the order of 10–20 seconds between initiation of a movement and receipt of the updated image.

All the necessary hardware and software to upgrade the iScope to a true streaming server is in place. We have a Macintosh G3 that runs the new Mac OSX, a UNIX operating system with a Macintosh-like shell. OS X comes with a custom QuickTime 4 streaming server. The QT4 streaming architecture has an impressive feature set, and we will be able to simultaneously deliver streams of video at multiple resolutions. The QT4 streaming system will allow us to deliver at least 10 fps at resolution of 160 x 120 pixels even to clients on 56 Kb connections. At the other end of the spectrum, the full MiniDV camera output amounts to 3.6 MB/sec, and it is likely that this bandwidth will soon be available to many of our clients, giving them the same video quality that we have at the camera. Letters in the appendix emphasize the enthusiasm of senior network and computer administrators of UT Memphis for this virtual microscopy lab. They view this project as an extremely powerful teaching tool, and they have guaranteed us their full cooperation is ensuring that our streaming servers will have the most suitable connections to the Internet.

Progressive video will be handled by a conventional Apache web server running on a Macintosh G3. Four streaming servers will be put on the UT Memphis fiber optic backbone using 100BaseT or higher bandwidth (see appendix letters). There will be two sources of video: html progressive video and live streaming video. HTTP streaming is also referred to a “progressive download” streaming or “on-demand video”; this method is much most appropriate for acquiring and playing QT movies in the MBL collection. Neuroscientists and stereologists are likely to run through the image stack both forward and backward, effectively changing the focal plane in the tissue. The time dimension has been converted to the third spatial dimension, and it is important to be able to move up and down with equal speed and equal image quality. In contrast, true streaming is essentially a live video feed from the iScopes to the client through a QT intermediary. This method requires the bandwidth of the media feed from the microscope to match that of the viewer’s connection. We will provide both types of video, but using different servers. Gamma will be customized for either Windows or Macintosh operating systems.

Video format and distortion

Aspect ratios of different video formats vary considerably (4:3 for NTSC, 16:9 for HDTV, 2.21:1 for Cinemascope). The majority of images will be acquired and processed at using the CCIR601 format: 720 x 486 pixels, 30 fps digital video format. This format has pixels with a ratio of 1:1.08. Aspect ratio is not distorted on analog playback. However, all standard computer monitors display square pixels and, as a consequence, if uncorrected, DV format appears wider than expected. Non-square pixel correction will be performed as a batch process using Media Cleaner Pro 4. Archival material will be retained in the original miniDV format, but through-focus series will be processed and uploaded to our servers as 640 x 480 pixel movies.

Navigational aids for streaming videomicroscopy

The current iScope provides xyz field coordinates relative to an arbitrarily define 0.0.0 point. We will need to integrate the Slide-Coordinate database with the iScope code in order to have a consistently defined 0.0.0 point of reference. The system of fiducial marks and coordinate scheme for each slide is one of the key aims of Project 1 (MBL) and is taken up elsewhere in more detail. Rather than displaying the xyz coordinates in a separate part of the web browser, we will merge

the coordinates and a calibration scale onto an alpha channel of the streaming video. A simple click will either add or subtract this alpha channel from the client's view of the section. Lumsdon et al. (1995) have described a simple analog video mixer that allowed them to align video data with electrophysiological recordings. Our solution will be along the same lines, but we will be able to exploit the alpha channel to do this interactively using a web interface. The QuickTime streaming server supports an 8-bit alpha channel.

Capture format

We plan to acquire essentially all movies in the miniDV format. The choice of this format is a crucial design decision and one that needs to be justified. Our first justification is that this format is of significantly higher image quality than conventional NTSC video currently used in most laboratories. It is also a higher quality than Hi8 or S-VHS standards. A major problem with NTSC and PAL video is that both are interlaced formats. The two fields that make up a frame can give rise to a comb artifact. This artifact is objectionable when NTSC signals are displayed on non-interlaced computer monitors. The raw NTSC format also takes up a huge amount of space: 27 Mb per second of video. In contrast, the miniDV format incorporates progressive scan, non-interlaced acquisition at 30 fps. The resolution is also more than adequate for the task at hand: at the video magnification at which the majority of image stacks will be acquired 720 x 486 pixels/frame will be able to capture even the finest details. For example, with the current iScope and a 40x objective, individual pixels have dimensions of about 0.15 µm, much finer than the resolving powers of the objective. This ensures a high modulation transfer from the image plane to the CCDs and ultimately to the display.

Like NTSC video, the miniDV format has an aspect ratio of 4:3, and it is therefore well matched to display on conventional televisions and most computer monitors. This aspect ratio is more suited for video microscopy (although 1:1 would be ideal) than the 16:9 ratio of digital TV.

Of greatest significance, miniDV is digital from its inception in the camera and can be copied from generation to generation without loss. The data stream output from the camera is fixed at a constant 3.6 MB/sec—a tractable rate that can be handled without interruption by modern computer hardware and disk drives. This 3.6 MB/sec is fed via the now ubiquitous Firewire/IEEE1394/I-Link directly to a computer. Firewire cables can be up to 10 meters in length, allowing us to separate microscopes from computers if necessary.

Cost is a major factor that has driven our decision to adopt the MiniDV format. Very high quality 3-CCD cameras can be purchased for approximately $4,500. The camera we are currently using is an extremely versatile unit with "progressive scan." Each frame is recorded as a single non-interlaced image. Other advantages include excellent sensitivity (allowing us to run the tungsten bulbs at a lower voltage), built in color-correction, control of functions from the computer, and a trouble-free constant operation in camera mode. One important practical advantage of the Canon XL-1 is that the lens can be dismounted, and once a custom adapter piece has been machined, the CCDs can be placed directly at the image plane of the objective. This simple arrangement can yield extraordinary high quality.

The particular advantage of the Canon XL-1 is that it has interchangeable lenses. We have machined a special adapter for the XL-1 that allows it to be mounted directly to our Zeiss Universal.

MiniDV compression

An uncompressed or source NTSC signal required a bandwidth of approximately 27 MB/sec. In contrast, the higher-quality miniDV format requires only 3.6 MB/sec. This trick is accomplished by an extremely efficient compression method carried out in the camera. The miniDV format uses 4:1:1 or 4:2:0 color subsampling compression. The luminance signal is not subsampled at all, but the chromanance channel has either one-quarter or one-half the original spatial resolution. Compression can result in subtle color aliasing artifacts. However, in the context of imaging either Nissl-stained or immunohistologically stained specimens that are (or will be) part of the MBL, which will be used for data acquisition and analysis, the subtle color aliasing artifact is not important when weighed against the numerous advantages of this digital format.

Use of an alpha channel

An alpha channel is often used in video applications to apply text and graphic overlays onto a video source. In the context of this project, we will be using an alpha channel to label each frame with a Z-axis position as well as with calibration. We have experimented with various ways to use alpha channels with QT4 movies (see Special Features section of the <nervenet.org/mbl/mbl.html> site).

Processing bottlenecks

The main bottleneck will involve those steps that require human intervention. The only significant computation bottleneck is the asymmetric compression of image stacks for QT4 web publication. The Sorenson Video codec takes up to 2 seconds to compress each frame. It will therefore take 1–2 minutes to compress each through-focus master clip to the version for web distribution.

Keyframes or not

In a through-focus series at 1-µm steps, there is a great deal of z-axis image redundancy that can be stripped away without significantly reducing image quality. This "temporal" or interframe compression can be carried out using several

compression-decompression utilities: Indeo (an Intel Co. method), the Sorenson Video codec (our current preferred codec for QuickTime 4), or MPEG-2 (used for full broadcast quality DVD-Video). To maintain high quality throughout the Z-axis set, we plan to set keyframes every 5 frames or every 5 µm. A keyframe is a frame that is spared from compression during the production of final output movie. Most compression methods assume that the movie will be played in a preferred direction, moving forward in time. But for a through-focus series, there is no preferred direction, and as mentioned above, it is critical that our video microscopists be able to focus up and down with equal speed and image quality. Finding the optimal compression method that meets this design criterion and that leaves us with z-stacks under 2 MB is our goal. Given that the original unprocessed stack is under 7.2 MB, we may be able to achieve compact files just using a mild JPEG compression on each frame. An example of this type of compression is shown on our web site at <nervenet.org/mbl/mbl.html>. This simple would be ideal because each frame would have precisely the same image quality. Whatever solution we strike upon, we are not making an irrevocable decision—all stacks will be archived in their original full miniDV resolution.

Construction of through-focus image stacks

Each through-focus movie will be 40–50 frames. Thirty to 40 frames (enough to extend through the entire section) will be devoted to the stack itself. The remaining frames will be appended JPEG frames. They appended frames will consist of the following:

1. A generic MBL title frame with support credit with web address
2. Synopsis of case data from our phenotype databases
3. Image of whole slide
4. Image of section with stack target marked and precise acquisition target coordinates,
5. Data on time and microscope/video set-up used to acquire QT movie
6. Full-frame grid-type calibration standard
7. Edge-only marginal calibration
8. Gray scale gradients
9. Standard color bar
10. Counting frame (50 x 50 µm)
11. Counting frame (25 x 25 µm)

A small calibration "watermark" will be placed on every frame. Each frame will be labeled with its z-axis position (corrected for Snell's Law).

Data rate of the DV format is 3.6 MB/sec. This is a real-time 1:8 compression relative to a corresponding NTSC signal that is implemented by the camera hardware. A two-sided DV-RAM disk with a capacity of 5.2 GB can hold up to 1400 seconds, or 24 minutes of material. A single DV-RAM could therefore comfortably fit well over 1200 original quality through-focus series. Access to individual files on such a disk is rapid. A single miniDV tape could store several hours' worth of original quality QT movies.

The QuickTime multimedia architecture is an effective vehicle for delivering through-focus image stacks. This architecture is currently the best choice for cross-platform compatibility, whether for streaming video, html progressive movie delivery, or the production of CD-ROM or DV-RAM disks.

The challenge of developing batch acquisition and processing procedures to automatically generate, archive, and compress QT4 movies

The *co*mpression-*dec*ompression codec that we will use in conjunction with QT4 movies is the Sorenson Video Pro edition. This is a highly asymmetric codec—that is, the compression is much more computationally intensive that the decompression. This is advantageous from the point of view of achieving fast playback of movies, but it puts the rate at which movies can be compressed as slow as 1 sec per frame even with extremely fast microcomputers (such as the Macintosh G3). It will therefore require approximately an order of magnitude more time to compress through-focus series than it does to acquire the stack. We do not yet know what the optimal solution and match will be between stack acquisition and stack compression. Even in a worst-case scenario, we do not expect it to take more than 5 minutes to compress the 30 to 40 frames that will be acquired at each site. We think that we should easily be able to acquire 20 to 30 movies per hour. The highest sampling density that we will use to "scan" a slide will be a 1 mm grid pattern. This relatively fine sampling grid will generate between 200 and 300 sample sites per slide, a number that we expect eventually to be able to acquire in well under an hour.

Attributes of the QuickTime architecture that we will exploit include the following.

1. *QuickTime poster frames.* Poster frames make it possible to embed a single frame in a page, which when selected is replaced by the movie. Large numbers of QT4 movies can be embedded in a single page.
2. *QuickTime chapter list.* This feature enables clients to click on a pop-up menu and then directly jump to a selected part of a video. Chapters are defined when the QT movies are assembled. Chapters might be particularly useful if we decide to experiment with providing multiple through-focus series as one continuous movie. For example, an

entire set of 240 through-focus series from one brain could be put together as a 4-minute clip. Chapters could then be used to define the structure or coordinates of each of the stacks.

3. *QuickTime fast start.* The fast start feature allows a movie to begin playing before it has been fully downloaded. This feature provides feedback faster, an important interface consideration.
4. *QuickTime text track.* It is possible in QT to add a text track with subtitles. This feature makes it easy to add z-axis coordinate to each frame in a stack.

RESEARCH PLAN

PROJECT 3: NEUROCARTOGRAPHER AND SEGMENTATION OF THE MBL (NISSANOV, PI)

In a twist of a well-worn aphorism, John Morrison and Patrick Hof (1998) note, "A number is worth a thousand pictures." Projects 1 and 2 will generate massive collections of images and through-focus series from brains of several thousand mice. The next task is to extract numbers from the MBL's vast collection of digital micrographs. Therefore, the main objective of Project 3 is to develop efficient tools to support transformation of huge image databases into compact quantitative representations of variation in the architecture of the mouse CNS. Using the MBL and iScope images, the NeuroCartographer Project will continue a productive line of research at Drexel University.

Aim 1: Construction of mouse 3D brain atlases and navigation software

Efficient segmentation of the MBL collection ultimately relies on fully delineated 3D models (atlases) of the mouse brain. (Segmentation refers to image analysis procedures that define spatial boundaries of neuroanatomical structures.) We will construct 10-µm isotropic 3D digital representations for both males and females of two strains of mice, C57BL/6J and DBA/2J, which are the parental strains of both our set of 1400 advanced intercross progeny and the 35 BXD recombinant inbred strains now in the MBL. Three-dimensional boundaries of approximately 1200 structures—homologs of the structures defined in our rat brain atlas (Nissanov and Bertrand 1998a)—will be manually delineated. The spatial coordinates of one of the atlases will serve as the standard coordinate system of the MBL. The other atlases will be mapped into this coordinate system with an accuracy of 20 µm. To support viewing of the atlases and their anatomical delineations over the Internet, navigation software (NeuroTerrain) will be developed.

Aim 2: Affine mapping into standard coordinates

Each brain in the MBL is stored as a series of ordered but unaligned images of cresyl violet–stained sections interlaced with Loyez fiber–stained sections. In-plane image resolution now ranges from 25 to 4.5 µm; off-plane sampling for a given stain is one section in five (150 µm). To exploit this vast image database efficiently, we will compute affine transformations that best aligns each section to the MBL coordinate system. This registration process will allow neuroscientists using NeuroTerrain to find homologous planes of view across the brains stored in the database at a resolution of approximately ±225 µm. NeuroTerrain will be enhanced to support viewing matched planes from multiple brains simultaneously. New alignment software will also be made available to users for registering a 2D section set into a 3D atlas.

Aim 3: Segmentation of the MBL and extraction of quantitative traits

The alignment performed in Aim 2 will be refined to permit accurate segmentation of sections. Using automatically segmented fiducial points for registration, a nonlinear transformation will be computed that maps each section into the standard coordinate system. (The application to be developed for nonlinear alignment will be released to the research community.) Our five-year objective is to segment each brain in the Mouse Brain Library into 1200 standard anatomical structures (at a resolution of 36 µm) from which quantitative values can be rapidly extracted. New features in NeuroTerrain will include support for user-defined volumes of interest (VOIs) and visualization tools for exploring anatomical differences between strains.

Software developed in all phases of the project will be available without charge to neuroscientists using the MBL. With the variation maps of the mouse brain generated as part of Aim 3, scientists, ourselves included, can rapidly compare brains from any two strains, or from groups of individuals within a strain, as a function of age, sex, body weight, or environmental factors. This extremely powerful extension of the MBL will highlight those parts of the brain most amenable to quantitative genetic analysis. Using the software utilities and web interfaces developed in Aim 3, online researchers will be able to capture, simplify, and display information about the variation among CNS compartments and then generate variation maps of the mouse brain. Data will be extracted and displayed in atlases generated in Aim 1. Neuroscientists will be able to assess, for example, how variation in volume of the striatum correlates to variation in the volume of the anterior thalamic nucleus, the substantia nigra, and the cerebellum.

PROBLEMS AND ISSUES ADDRESSED IN THIS PROJECT

A spatial database, like any other database, consists of content and infrastructure to support access and analysis. This project is largely concerned with infrastructure, focusing on the analytic tools needed as part of a neuroanatomical

database. To transform a collection of images such as the MBL into a powerful research database, we need a suite of special tools that have some capacity to detect, define, and extract numerous compartments, nuclei, ventricles, and fiber tracts that make up the mouse brain.

Typically, a user of the MBL will want to compare a region of interest (ROI) across strains. The investigator may wish to visually inspect histological sections, perhaps sorted over the range of strains according to the size of the ROI. Alternatively, the investigator can issue a command to the iScope system (Project 2) to automatically sample a small volume of tissue (typically 50 x 50 x 30 µm) within an ROI across a set of brains. If manual delineation of each brain were a prerequisite for data extraction, the utility of the NeuroCartographer project would be diminished for the simple reason that our throughput would be severely limited.

Automatic segmentation is crucial. In addition to saving time, automated segmentation avoids the subjective judgement involved in manual delineation that is the major source of error in data analysis (Eilbert et al. 1990). Our objective is to develop a reliable computerized system to expedite analysis. Algorithms will be designed to segment each brain in the MBL into a set of standard anatomical structures like those defined in the rat atlas produced by Nissanov and Bertrand (1998a). An investigator who prefers to define a nonstandard division will be able to modify neuroanatomical templates residing in the atlases and have these modifications automatically propagated throughout the database. Besides providing the means to navigate through the MBL, segmentation of the database content will have another benefit: positional variability and volume estimates of the anatomical structures will be analyzed in the Neurogenetics Tool Box (NTB; see Project 4), then used to map QTLs.

Automatic brain parcellation is a familiar theme in neuroinformatics (Toga and Thompson 1999). We believe that this project and this neurogenetics research setting lends itself particularly well to automation: the data set is highly homogeneous, consisting of cresyl violet– and Loyez fiber–stained tissue fixed in precisely the same manner, and embedded and cut in a single laboratory. These celloidin-embedded sections are of the highest quality. It is our hope that the technological innovations proposed here not only will serve well in the context of MBL but will provide insight to the general problem of brain segmentation.

Status of current research efforts

A prominent concept in neuroimaging is the use of anatomical templates positioned in a standard coordinate system. Experimental data are aligned into this coordinate system and consequently inherit labels for numerous neuroanatomical compartments. In addition to establishing homology at the level of anatomical structures, coordinates can be used directly to address equivalent points across brains. A number of coordinate systems have been developed for human neuroimaging. The most commonly used is defined by the Talairach atlas, which was constructed from histological sections derived from postmortem material (Talairach and Tournoux 1988). Another system was defined by 3D MRI (Evans et al. 1994). In nonhuman research, the closest equivalent to a standard system consists of stereotaxic coordinates defined by their positional relation to bony landmarks (Slotnick and Brown 1980; Paxinos et al. 1985), but these are ill suited for registration of brains to an anatomical template because the relation between the brain and the cranium is lost during histological procedures. Furthermore, even if the relation could be recovered, the variability of the skull is often substantial and does not necessarily support accurate template placement. Brain atlases are more suitable for this purpose.

There are a number of 2D digital rodent brain atlases (Toga et al. 1989; Bloom et al. 1990; Nissanov and Bertrand 1998a). To truly model the brain and to accommodate arbitrary planes of sectioning, however, 3D atlases are needed. Such atlases have been constructed using MRI or blockface imaging for a number of species (Black et al. 1997; Cannestra et al. 1997; Ghosh et al. 1994; Toga et al. 1994, 1995), but these are not ideal in the setting of the MBL. As the underpinning for a coordinate system, atlases serve two purposes. First, their demarcation into anatomical regions defines a standard template to be used in parcellation of experimental material. Second, features on the atlas, either geometric or gray value, drive the registration process. MRI and blockface imaging fail to provide the resolution and the contrast needed. Computerized 3D atlases from stained sections, a more suitable basis for an MBL atlas, have been generated previously for the rat (Nissanov and Bertrand 1998b; Nissanov et al., submitted) and the mouse (Celio et al. 1998; Davidson et al. 1997; Reed et al. 1999). During the past few years, we have developed progressively more effective tactics for atlas construction. We have used these tactics to construct a 2D rat brain atlas (Nissanov and Bertrand 1998a), a 42-µm isotropic 3D rat brain atlas (Nissanov et al., submitted; Nissanov and Bertrand 1998b), and a 20-µm isotropic 3D atlas of C57BL/6J (Reed et al. 1999). In ongoing studies we are developing methods to construct atlases from celloidin-embedded tissue. Once constructed, our new atlases will guide registration of MBL data into the standard coordinate system. Two important considerations are the types of information used for this alignment and the classes of transformations used to register sections. Information for alignment can be either geometrical (model) or image intensity–based. In geometrical alignment, points, lines, surfaces, or geometrical invariants are abstracted, automatically or manually, from experimental images. These are matched to homologous geometric objects in the atlas coordinate system. In image intensity–based methods, a global disparity between the experimental and atlas image intensities is computed, and alignment is adjusted to minimize this disparity. Classes of transformation for alignment range from geometrical "linear" mappings (e.g., rigid-body or affine transformations) to complex nonlinear mappings.

In intensity-based procedures, gray value disparity measures such as squared differences or correlation are useful in matching images with the same modality (Hibbard and Hawkins 1988; Collins et al. 1994; Thevenaz and Unser 1995). In the setting of histological material, a given stain constitutes a modality. More sophisticated and flexible disparity measures such as conditional variance (Woods et al. 1993) and mutual entropy (Wells et al. 1996; Thevenaz and Unser 1996; Maes et al. 1997) have been found effective for both intra- and intermodality registration. Many investigators are using the programs AIR (automatic image registration) from the University of California at Los Angeles (Woods et al. 1998 a, b) and ANIMAL (automatic nonlinear registration) from the Montreal Neurological Institute (Collins and Evans 1996). AIR can perform both linear (rigid body, affine) and nonlinear (polynomial) alignment with a choice of three gray value comparison functions, allowing inter- and intramodality alignment. ANIMAL is designed for intramodality intersubject alignment; it incorporates a powerful nonlinear spline alignment model.

Geometric-based methods rely on prior image segmentation of fiducials. Once obtained, the fiducial registration points can be used to warp brains into a standard coordinate system. The class of transformations should be matched to the character of the matching problems; for example, rigid-body transformations are suitable, up to a point, in 3-D reconstruction (Hibbard et al. 1987) and multimodality images from the same subject (Pellizari et al. 1989), and the affine transformation is the classical method for intersubject alignment (Collins et al. 1994; Fox et al. 1994). A greater range of intrasubject variability can be accommodated with nonlinear transformations, which include polynomials (Friston et al. 1995; Woods et al. 1998 a, b), product splines (Collins et al. 1995), and transformations based on continuum mechanics (Bajcsy et al. 1983; Christensen et al. 1996; Kim et al. 1997; Gabrani and Tretiak 1999).

For rodent brains, rigid-body alignment is surprisingly accurate even in the setting of interanimal alignment (Ozturk et al, submitted; see Appendix). We have employed 3D distance-based alignment to register brains using the outer surface as reference. Misregistration of the internal structures was found to be 193.5 μm (rms). Improvement will require nonlinear transformation guided by both the outer brain surface and internal fiducials. We propose to use spline interpolation for this purpose. Spline functions may be designed with the requisite degree of differentiability and are stable to local perturbation. The two general spline methods for multidimensional interpolation are (tensor) product splines (de Boor 1978) and radial splines (Duchon 1977; Meniguet 1979; Bookstein 1989; Gabrani and Tretiak 1999). Fiducial structures in brains occur at irregular locations, and radial spline interpolation is well adapted to such data. Product splines have certain computational advantages and have been used for brain registration by Collins et al. (1994). The key issue in constructing transformation functions is accuracy, which improves as more fiducial points are used and also depends on the smoothness of the function being interpolated. To reduce time and effort, it is desirable to use the fewest fiducial points possible. A commonsense resolution to this problem is to use more fiducial points in regions of high variability.

Although fiducial points are very easy to use, they are relatively hard to locate. Surfaces (or outlines in 2D images) are now commonly employed and have been the basis of many successful algorithms (Amit 1997; Borgefors 1988; Pellizari et al. 1989; Besl and McKay 1992; Jiang et al. 1992; Huang and Cohen 1996; Kozinska et al. 1997; Thompson et al. 1997; Cohen el al. 1998; Ibrahim and Cohen 1998; Gabrani and Tretiak 1999). The method we employ for computing transformations (described by Gabrani and Tretiak 1999) can use both types of data. Spline theory is based on defining an "energy" functional, typically the integral of the square of a derivative of the interpolating function. A spline interpolator is the function that minimizes the integral of the energy. For example, cubic splines minimize the integral of the square of the second derivative of the interpolation function, and Bookstein's (two-dimensional) thin-plate splines minimize the square of a Laplacian function. When we use surfaces for spline alignment, points from the surface of one object must be transformed to the surface of the second object. The location of the point on the second surface is not known but is found by minimizing the energy functional. The method is directly applicable to heterogeneous data consisting of fiducial surfaces, lines, and points.

No matter how the brains are aligned, the end result is immensely valuable. In addition to its use in segmentation, brain alignment has been employed to detect abnormalities (Thompson et al. 1997; Thompson and Toga 1997), for surgical planning (Levy et al. 1997; St.-Jean et al. 1998), and, most relevant to the present application, for database retrieval (Fox et al. 1994). For our purpose, alignment brings the brains into the coordinate system. Neuroscientists will then be able to access corresponding points across subjects by simply referring to a coordinate in the standardized system.

Limitations of existing approaches

Construction of high-fidelity atlases has relied on blockface imaging (Nissanov et al. submitted). The tissue block is imaged prior to collection of each section, and this image is used to guide 3D reconstruction. There are two major difficulties with this approach. First, blockface imaging demands specialized equipment that is difficult to adapt for use with celloidin-embedded material. Second, alignment of data sets collected using different techniques (intermodal) has limited reconstruction accuracy. We propose a section-to-section alignment strategy that should permit reconstruction accuracy approaching ±10 μm.

The difficulty with current alignment strategies is only partly due to the lack of appropriate algorithms. Although the task inherent to this setting is 2.5 dimensional (ordered, unaligned sections that do not necessarily constitute a complete series of sections) and most available alignment strategies are either 2D or 3D (exceptions do exist; see Kozinska et al. 1996; Kim et al. 1997;Schormann and Zilles 1998; Cohen et al. 1998; Ibrahim and Cohen 1998), we plan a

straightforward modification of existing approaches to tackle this problem. We propose to extend our 3D distance-based alignment method (Kozinska et al. 1997)) to solve the 2.5D registration problem at a coarse level of accuracy. Intersubject registration will then be improved using implicit fiducials and splines.

The most severe limitation of available technology, however, is the present interface to the MBL. Navigation within the current system is inefficient: an investigator wishing to look at homologous planes from multiple animals must inspect the data from each animal at a moderate level of resolution, establish homology visually, select the desired section from each brain separately, and then download the high-resolution views. To estimate volumes of structures, the investigator must delineate the ROIs. If unbiased cell density is to be determined from through-focus series, the user must navigate across the slide while it is mounted on the iScope and select the field desired. The technology to be developed in this project will facilitate access to the MBL, increasing its usefulness to collaborating scientists.

Relationship of project to objectives of the Human Brain Project

Substantial effort toward development of neuroinformatics tools has been expended since the Committee on a National Neural Circuitry Database was convened in October 1989. After the committee published its recommendation (Pechura and Martin 1991), the Human Brain Project was launched, and the new discipline of neuroinformatics blossomed (Koslow and Huerta 1997). The work described in this proposal will link neuronal phenotypes with corresponding genotypes. The influence of the MBL and the NeuroCartographer projects will extend far beyond the field of mouse neurogenetics. The MBL will become a well-traveled bridge connecting neuroanatomical and genome databases in other species. For example, if a QTL is identified on proximal mouse chromosome 7 that controls the volume of the caudate nucleus, then we can look for the homolog of this gene on human chromosome 19 or 11. Data sets generated by the NeuroCartographer project will make it feasible to traverse from mice to humans (and other species) by virtue of structural and genetic homologies. As a result, we will be in a far better position to explore the genetic basis of variation in human CNS structure.

Integrative relationship between brain research and informatics components of the project

Complex trait analysis and QTL mapping are data-intensive tasks. To establish the genetic and environmental basis of morphometric differences among genetically defined mice, a huge number of brains must be examined. Efficient exploitation of an image warehouse like the MBL cannot occur without efficient and effective informatics solutions. This project will extract quantitative CNS traits from brains in the MBL. These numbers will be a far more succinct and usable representation of the MBL. These quantitative traits will then be the "deliverables" sent from the NeuroCartographer Project to the Neurogenetics Tool Box. These numbers will also be fed back to the MBL to provide far more powerful navigational tools than would otherwise be available. It will be possible to issue a command in the MBL that is equivalent to "Display sections of the anterior thalamic nuclear complex of male mice between the ages of 200 and 300 days in these 10 strains." It will also be possible to issue a command to the iScope such as "Acquire through-focus series for stereological analysis (gray scale please) of the lateral habenula on both right and left sides of a sample of the first 100 advanced intercross progeny in the MBL." All of these powerful new tools rely on advanced segmentation protocols to digitally dissect the MBL collection.

Novelty of the informatics component

The bioinformatic novelty of this project is three-fold.

First, this project has an absolutely unique research setting in the context of the quantitative genetic dissection of the mouse CNS. Our atlases and segmented data sets exist in a very complex informatics environment that requires integration with the MBL, the iScope project, and the Neurogenetics Tool Box. This integration across levels of analysis and even across major research divisions (neuroscience and genetics) requires a new level of sophistication in database design.

Second, we intend to make our tools accessible and malleable using web-based interfaces. Our goal is to allow neurogeneticists to modify and add to our brain models. If the pattern of in situ hybridization of a particular cDNA or mRNA defines a novel CNS compartment, then we need to allow this compartment to be added to our 3D model. The same applies to antibodies that label specific compartments, such as the anti-ChAT and anti-parvalbumin that we will test as part of the MBL. One can imagine an almost infinite number of ways in which molecular tags (whether epitopes recognized by antibodies or mRNAs) might segment the brain. This level of plasticity in models is definitely not yet the norm, and devising ways to make our models plastic over the Internet will be a significant challenge.

Third, this project is associated with a classic Phase II problem: how does one move effectively from small demonstration projects (Phase I) to the massive high-throughput production mode that should ideally characterize more mature Phase II and III technology? In this project we need to segment brains and produce numbers at a pace that can keep up with a phenomenally productive Neurohistology Core that even without major funding has processed and digitized more than 600 brains. Gearing up to this challenge may seem like just a quantitative difference, but at this level of throughput, quantitative differences mandate qualitative changes in procedures.

EXPERIMENTAL PLAN

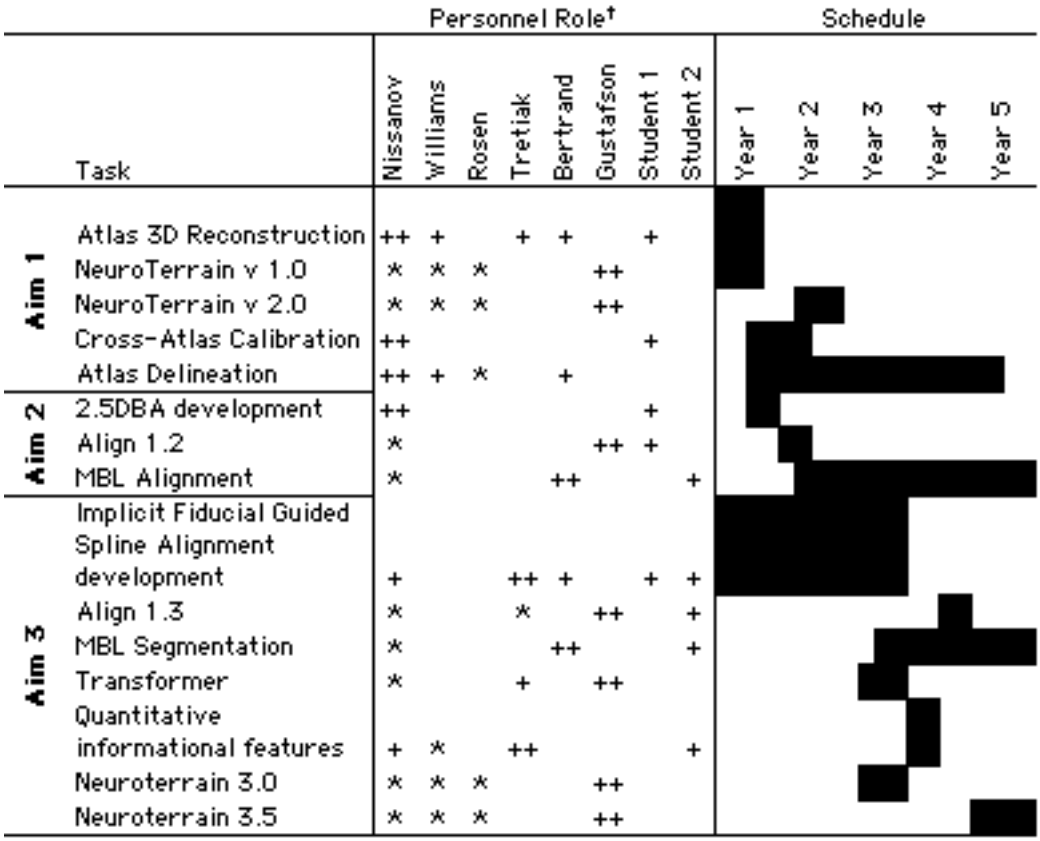


Figure 1. Timeline. Definition of personnel role notation: ++ team leader (manages day-to-day operation and reports to PI regularly); + team member (participates in data analysis or coding on a regular basis); * advisor (advises team regularly)

Figure 1 is a timeline for all proposed tasks as well as the role of personnel. In this plan we outline procedures associated with each task and describe our accuracy objectives, with details on how these objectives will be assessed. Finally, we describe alternative tactics we will use if initial methods do not meet our criteria.

NEW METHODS AND THEIR ADVANTAGES

Aim 1: Construction of mouse 3D brain atlases and navigation software

In Aim 1, we propose to construct eight 10-μm isotropic brain atlases: two stains (Nissl and fiber), two strains (C57BL/6J and DBA/2J), and both sexes. The brains will be manually delineated. A transformation function for mapping coordinates across the brains will be established. To view the atlases over the Internet, a JAVA version of MacOStat, our existing atlas viewer, will be developed and released.

3D atlas reconstruction

The sections to be aligned in the atlas construction phase will be imaged at a resolution of 4.5 μm. Recently, we have begun examining whether the volumes could be reconstructed without the aid of blockface imaging. Instead we have performed section-to-section alignment using AIR (Woods et al. 1998 a, b). This method yields surprisingly good global alignment (Fig. 2), and we have used it to generate a 30-μm isotropic reconstructed brain from celloidin-embedded sections (Fig. 3). Because rigid-body transformation was used for this alignment, cutting-induced distortion reduced reconstruction accuracy. To examine whether nonlinear registration will be required to reconstruct 10 μm isotropic

atlases, we aligned consecutive sections using punctate fiducials (blood vessels). The residual misregistration was found to be about 45–50 μm.

To overcome this reconstruction error, fiducials will be used to drive nonlinear transformation. After rigid-body inter-section alignment, distortion will be corrected by selecting ROIs in one section, cross-correlating these with nearby regions on adjacent sections, and obtaining displacement vectors. We hypothesize that overlap of cells that occur on both sections will cause a correlation peak. If the two sections contain information from linear structures (e.g., blood vessels) or planar structures (e g., cell layers or ventricle edges) that are oblique relative to the section plane, then the correlation peaks due to these features may be displaced. The possibility of such bias will be mitigated by (1) avoiding regions that contain such structures and (2) high-pass filtering the two images to suppress broad correlation peaks. The number of such alignment points needed depends on the spatial frequency content of the distortion field, which will be established empirically by constructing such a field from many windows on a sample of section pairs. Preliminary data suggest that about 25 (5 x 5) masks, each 200 x 200 μm, will suffice. A distortion field from section *i* to section *I*+1 will be constructed by assuming that the *I* th section is distortion-free. Spline interpolation will be used. After all such maps are collected, they will be globally "relaxed" by minimizing the total mechanical distortion (strain) energy over the brain. Our overall objective is to achieve an unprecedented alignment accuracy of better than 10 μm.

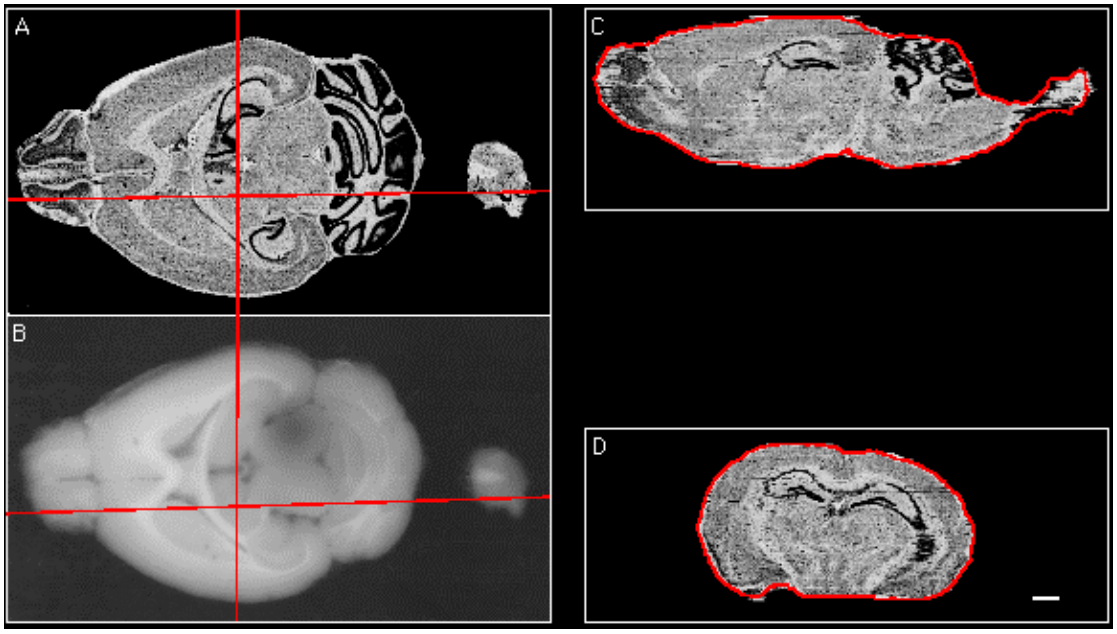


Figure 2. Section-to-section rigid-body alignment. Horizontal brain sections (C57BL/6J) such as the one shown in panel A were cut with a cryostat equipped with a blockface imaging system. The system yields aligned blockface images (Nissanov et al., submitted). Shown in panel B is the blockface image corresponding to the section shown in A. The stack of aligned and stained sections was resampled perpendicular to the cutting plane at the location indicated in *A* to yield sagittal and coronal views (*C* and *D*, respectively). The blockface stack was also resampled perpendicular to the same location as in A on the corresponding blockface image (*B*). The outer contour of the resampled blockface views is superimposed on the corresponding resampled sectional stack (C and *D*). Note that AIR preserved the global shape of the brain well. Scale bar = 1mm and applies to all panels.

Atlas delineation

Dr. Nissanov, Dr. Williams and Ms. Bertrand, all of whom have expertise in neuroanatomy, will delineate the atlases in three dimensions using available software. For this phase of atlas creation, the anisotropic volume will be employed (4.5 μm in plane, 10 μm in the orthogonal axis). To accomplish this task, we have built a software tool, BRAIN (available from the Computer Vision Center for Vertebrate Brain Mapping at Drexel), and have employed it in delineating our 2D-atlas fully (Nissanov and Bertrand 1998a) and the 3D-atlas partially (Nissanov et al., submitted). Delineation will be smoothed in 3D using morphological operators. In addition, other investigators will have the option of inserting their own delineations. Thus, users can select which anatomical template they wish to employ. Complete delineation of each atlas (approximately 600 VOIs per side, 1200 total) will take approximately 6 months. They will be made available as they are generated.

Cross atlas alignment

Our objective is to define a single standard coordinate system for the MBL. One atlas will be selected for that purpose and the others will be transformed into that system. Point-to-point mapping will first be determined by an affine transformation computed with our previously developed 3D distance-based alignment algorithm (3DBA), followed by a "small" radial spline transformation. The transformations will be computed with a mixed fiducial point and fiducial surface alignment algorithm based on the method in Gabrani and Tretiak (1999). The fiducials will consist of a set of

about 100 three-dimensional points delineated manually. These and the outer surface will be used to drive the algorithm. The required software is already available at Drexel.

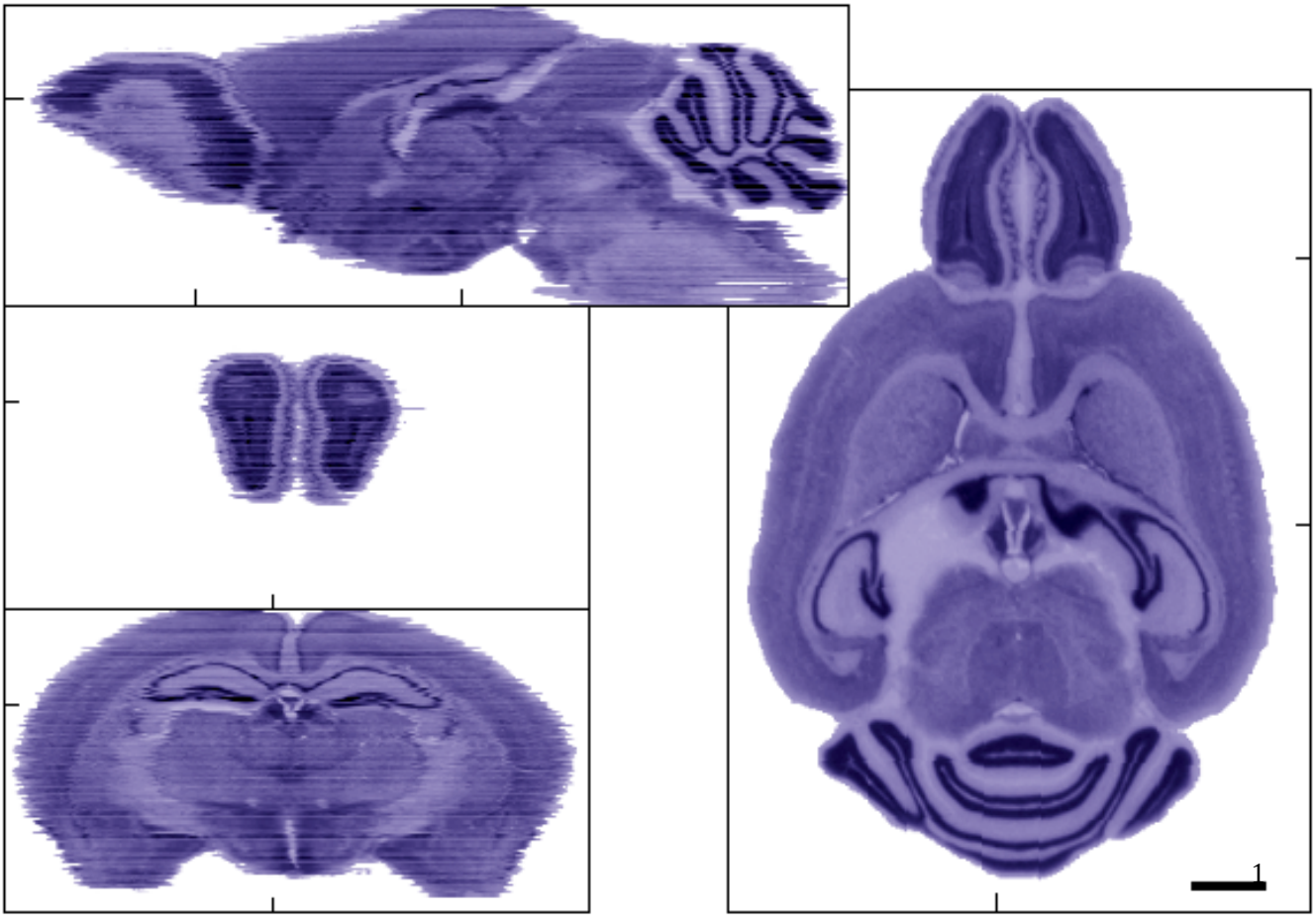


Figure 3. A 30 µm isotropic 3D reconstruction of C57BL/6J brain. Images celloidin sections were aligned using AIR (rigid-body). The original cutting plane was horizontal (far right, *D*). Sagittal (top, *A*) and two coronal planes (lower left, C and D) were obtained by resampling the volume. Location of views is cross-indicated by tick marks with adjacent notation referring to the panel of the given view. No gray value alignment (e.g., Nissanov et al., submitted) was performed. Scale bar applies to all panels.

To assist in matching fiducials in cresyl violet–stained atlases with Loyez fiber–stained atlases, 4 brains (male and female individuals of the strains C57BL/6J and DBA/2J) will be cut, and the two stains will be interlaced. We will seek to achieve an overall accuracy of 20 µm on this task.

NeuroTerrain versions 1.0 and 2.0

MacOStat, atlas navigation software for the Macintosh platform, is available from the Brain Mapping Center at Drexel. This application allows users to interactively define an arbitrary plane of view through a 3D data set. The view is dynamically updated as the user rotates or translates about x, y, or z axis. The atlas also supports VOI display. The outline generated by intersection of VOIs and the plane of view can be superimposed over the atlas.

MacOStat currently is written in C++ and runs on the Macintosh operating system. We will port it to JAVA to support navigation over the Internet (NeuroTerrain) and will work with Dr. Rosen to achieve a smooth interface with the MBL. A consideration is the speed of navigation over the Internet. A view through the 10-µm isotropic atlas planes along the greatest dimension requires 800 KB of computer storage space. In MacOStat we have achieved dramatic reduction (44%) in memory (both disk and RAM) footprint using a macrovoxel data representation (Nissanov et al., submitted). Voxel data are organized first into macro voxels, which taken together form the entire volume. An internal coordinate system relative to the macro voxels is used for slicing operations. The whole number portion of a coordinate triple specifies the macro voxel containing that point, and the fractional portion of the coordinate triple specifies the micro voxel within the macro voxel. Any macro voxel in which all micro voxels have the same value is eliminated. In practice, any macro voxel that is all white (the image background) is eliminated because the class representing the volume map will return a white voxel

for any missing macro voxel. Thus, in navigating through the atlas, the data transfer requirement is reduced to less than 450 KB per view. With current networks, this translates to about a view every 5.5 sec. To accelerate this process, we will employ a progressive encoding scheme that supports a 10-fold speed increase during coarse navigation. We estimate release of the first navigator version will be 9 months after commencement of the grant period. A manual and a tutorial will be distributed online with the applet at no charge.

MacOStat currently supports simultaneous viewing of homologous arbitrary planes through multiple volumes. This feature will be incorporated into the NeuroTerrain version (v 2.0) of the navigator by the end of Year 02.

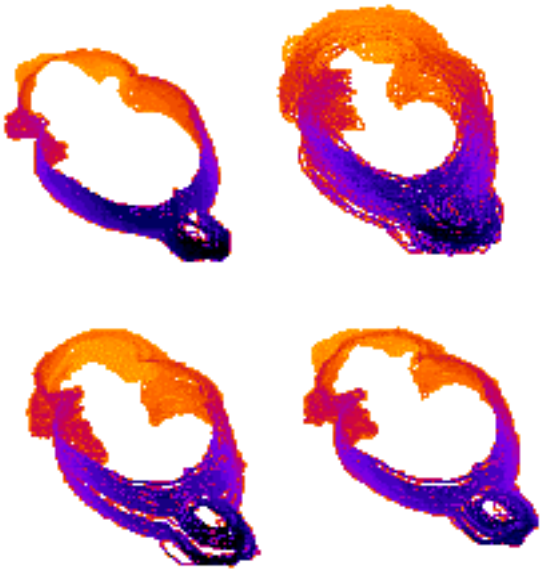


Figure 4. 2.5 distance-based alignment. Images of sections from a rat brain were aligned to the blockface data of the same animal. Shown in A is the stack of outer contour from the set of 100 sections. The contour from the set of contour is shown in B prior to alignment, in C after section-to-section alignment with moments and in D after 2.5DBA. Some distortion in shape along the axis perpendicular to the horizontal plane can be observed in panel C relative to A. 2.5DBA improves the alignment appreciably. Color codes for distance from observer.

AIM 2: Affine mapping into the MBL standard coordinate system

The main objective of this aim is to establish homology at a resolution of ±225 μm among the brains stored in the MBL so that scientists using the MBL can readily retrieve similar planes of view from multiple animals by specifying the desired plane on the atlas. As part of this aim, we will develop a new version of our alignment toolbox that incorporates 2.5D registration, and we will release a new version of NeuroTerrain that will support simultaneous viewing of multiple brains.

2.5-dimensional alignment

An obstacle in registration of histological material to a 3D atlas is that the sectional material is an ordered but unaligned 2D data set. When every section has been cut and stained, correlational rigid-body alignment (Toga 1993) using the section silhouette can be employed to first reconstruct the volume. 3DBA (Kozinska et al., 1997) can then be used to register the brain to the most suitable atlas. When gaps exist in the data set, as is the case in the MBL data, reconstruction using correlation is not possible. We will instead use a 2.5D extension of 3DBA (2.5DBA).

The distance-based alignment algorithm is available as a MatLab application <coe.drexel.edu/ICVC/Align/align11.html>. The latest version, Align 1.2 (not yet released), includes an implementation of 2.5D. Our tests of this approach have been confined to intra-animal alignment and phantoms where it does perform well (Fig. 4). For use in the setting of interanimal registration, the algorithm will be modified. The initial step in the registration process is section-to-section moments-based alignment. This achieves a fair degree of local registration accuracy. Next is an iterative alternation between 3D-3D and 2D-2D alignment. The 3D-3D alignment positions the volume as a whole relative to the atlas while the 2D-2D seeks the best in-plane rotation and translation of a given section to match the homologous atlas plane. In the present implementation, the 2D-2D alignment is unconstrained, and substantial inter-section shear may be introduced. To limit these problems and maintain a smooth section-to-section transition, we will impose inter-section coupling to penalize excursion from the initial position. The actual alignment is estimated to take less than 1 hour/brain and will require only 5–10 minutes of human inspection of automatic segmentation (using isodata thresholding) of the outer section contour.

Align 1.1 is now available on multiple platforms (IRIX, SOLARIS, Windows and MacOS). If proven successful, our extension of 2.5DBA will be released with an online manual and tutorial as we have done for Align 1.2.

MBL alignment.. After the 2.5DBA tests are complete, MBL sections will be aligned. After each data set is aligned, a resampled view orthogonal to the cutting plane and the homologous atlas plane will be generated and visually inspected. The reoriented images will replace the raw data on the database, and the transformation parameters will be stored. This aspect of the proposed work will be ongoing throughout the grant period..

Aim 3: Segmentation of the MBL and extraction of quantitative traits

The final objective of this project is segmentation of the MBL and measurement of quantitative traits. To that end, we propose to develop fiducial-based nonlinear alignment. The algorithms will be tested, implemented as part of the Align

software package, and released. Additionally, they will be used to segment the MBL into over 1200 VOIs (see Appendix for complete listing). We will also develop a program to translate coordinates on one brain to another and will add features to NeuroTerrain that will support delineation of VOIs.

Implicit fiducial–guided spline alignment.

Fiducial points and surfaces will be used for constructing the atlases (Aim 1) and fine alignment of the MBL. Current image-processing techniques can locate relatively few fiducial structures in the brain (Nissanov and McEacrhon 1991; Schliecher et al. 1999). Manual segmentation is possible, though time-consuming. We propose to identify fiducial location by using local comparison of atlas and MBL regions. A similar concept was used by Bajcsy and Kovacic (1989) and Amit (1997). The general notion is that for each fiducial, we will select a 3D region *R* in the atlas reference brain, then compare the gray value function in this region with the MBL brain gray values over *R* at various translations, rotations, and scalings. The comparison process produces a scalar that reaches an extremum (maximum or minimum) when the two gray distributions are most similar. Comparison may be made by cross-correlation (Hibbard and Hawkins 1987; Bajcsy and Kovacic 1989), squared difference (Thevenaz and Unser 1995), conditional variance (Woods et al. 1993), and mutual information (Thevenaz and Unser 1996; Wells et al. 1996; Maes et al. 1997).

We plan to introduce several refinements on the above method. The comparison technique intrinsically identifies a transformation between regions. We will identify a segmentable point in the atlas region and associate the best translation as the displacement of this point: this is what we mean by an implicit fiducial (IF). These data (fiducial point and displacement) are directly applicable for constructing a spline interpolant transformation. Segmenting the same point in the MBL data set allows us to validate the accuracy of the given implicit fiducial. Such validation, on a subset of the MBL brains, will allow us to select stable and reliable implicit fiducials. We will use a multiresolution strategy in the search (Bajcy and Kovacic 1989). An initial approximate alignment will be performed; subsequently, we will look for a few low-resolution IFs that are not sensitive to (moderate) rotations. These will be used to compute an approximate spline transformation. After this is applied, we will use higher-resolution IFs to refine the transformation. These can be applied over smaller search zones. Proper care will be taken to address aliasing problems caused by coarse z-axis sampling and imprecise alignment in the MBL. Some smoothing will be performed in the z direction on the atlas data, to remove possible sensitivity to angular misalignment, and a set of slices through the atlas will be compared with a group of regions in adjacent MBL sections. This set will be stepped along the z axis in the atlas in order to achieve accuracy finer than inter-section spacing. We propose to compare the summed squared difference between the atlas and the MBL, with adaptation for brightness and contrast differences (Thevenaz and Unser 1995). Two strategies will be explored to account for MBL x-y misalignment. The data may be smoothed in the x-y plane to mitigate effects of this jitter. Alternatively, the MBL sections can be translated by different amounts so that each produces a best match, and differences in displacement may be used to refine alignment.

MBL segmentation and quality control

Our main objective in Aim 3 is to use the computed transformation to map the 3D anatomical template onto at least 1200 brains to be stored in the MBL (Project 1). The 4.5-μm/pixel images scanned into the MBL collection will be aligned to standard coordinates. To match the resolution of the atlas, the images will be downsampled to 10 μm/pixel.

For quality control purposes, 5 ROIs will be delineated on each brain, with each region delineated on one section. These ROIs will be compared to the automatic segmentation results. To limit the error associated with manual delineation, we have chosen structures that can easily be segmented using the semiautomatic procedure of thresholding and editing. Our objective is to limit both false positive and false negative pixel assignment to 20% for a 225-μm-diameter circular ROI (this corresponds to roughly 3 times the error on repeated manual delineations of strong edge). In addition to this quantitative performance measure, the quality of the segmentation will also be evaluated by visual inspection. Sections with their anatomical overlap and the corresponding atlas planes will be displayed using NeuroTerrain.

We plan to complete software development, test the alignment algorithm, and begin segmentation of the MBL by the middle of Year 03. Complete delineation of the MBL will require a segmentation throughput of 4 brains per day. With automatic segmentation, we believe that the limiting factor will be availability of brains rather than the alignment algorithm.

Transformer

Transformer will be an object (program plus parameters) that evaluates coordinate transformations between data sets—either atlases or MBL brains. For the purpose of database access, the physical coordinates of one of the atlases will be selected as the coordinate system (see Aim 1). Using the Internet, an investigator can specify either a region (point, line segment, rectangular window, stack) in the native coordinates (i.e., voxels) of one brain. The transformer will give corresponding coordinates in the native coordinates of another brain. Since considerable care will be taken in establishing the physical location of sections on the slide (Project 1) it will be possible to translate to both image coordinates and to the physical slide coordinates. The coordinate transformations will be computed by composing transformations from one brain to the MBL standard coordinate system and from the coordinate system to another brain. This way, only about 2500

transformations need to be stored. Transformations will be implemented as product cubic splines. Assuming the knot set is on a 10 x 10 x 10 grid, each transformation will require about 12,000 floating point numbers.

Although the above transformation specifications are in the form of product splines, the alignment procedure will compute radial splines from fiducial point data. We use radial splines for alignment because fiducial points range over an irregular grid; product splines, however, are faster to compute. Product splines will be derived from radial splines by sampling these splines over a grid. The grid will be selected to provide adequately accurate approximation of the product spline.

Quantitative informational features

Once segmented, a variety of quantitative information can be extracted and fed to the NTB (Project 4). A straightforward measure is estimate of the volumes of VOIs. Another numerical feature is the magnitude of the nonlinear displacement of a given point in the MBL coordinate system. One can compute point configurations (Amit 1997) or morphological Euclidean invariants as developed by Bookstein (1991). A natural set of features can be derived using tools developed in continuum mechanics (Chandrasekariah and Debnath 1994). Let $f: R^3 \to R^3$ be a C^1 alignment transformation, and let $\mathbf{F} = \nabla f$ be the deformation gradient tensor.. Then det**F** is the Jacobian, or local volume ratio, and $\det \mathbf{F}(\mathbf{r}) / \det \mathbf{F}(\mathbf{r}_0)$ is an affine invariant, so it is insensitive to difference in (uniform) shrinkage of the brains. Thus one possible set of numerical indices is the values of the Jacobian, indexed by the MBL coordinate system. Mechanical or functional causes for differences between brains can be indicated by regions where the value of the Jacobian is different from unity. If we evaluate the polar decomposition $\mathbf{F} = \mathbf{VQ}$, where **V** is symmetric and **Q** is orthogonal, then the eigenvalues of **V** are Euclidean invariants of the transformation, and differences between eigenvalues and unity indicate non-isotropic growth. In the presence of an intrinsic direction, e.g., the normal to cortex, $\| \mathbf{Fn} \|$ measures size difference along that axis. If manually segmented information of structures is available, one can compare changes in volume of a given nucleus using either the Jacobian or the overall size difference at that location.

Early in this phase of the project, the facilities to support user-defined VOIs will be incorporated (NeuroTerrain version 3.0, Fig. 1). It will be possible to propagate these VOIs across the entire database. This version of NeuroTerrain will also support manual refinement of the placement of the anatomical template on any given brain, yielding substantially higher segmentation accuracy.

In a later version (NeuroTerrain 3.5), facilities to support visualization of interbrain differences will be added. As described above, we will compute a variety of measures (e.g., non-isotropic compression and expansion and magnitude of displacement relative to the MBL coordinate system). Users will be provided with a number of ways to access that data. One will consist of a point-and-click interface where the variation across strains for the selected coordinate will be retrieved. Another will allow investigators to request a voxel-by-voxel difference map between any two selected strains. Such a map could be displayed as an overlay on any of the eight atlases.

We believe that these measures will have considerable ramification for QTL mapping. Statistical techniques to compare these maps and methods to explore these in the context of developmental models will be a major research direction beyond the present grant period.

ASSESSMENT OF FEASIBILITY

The major tests of feasibility to be performed are the accuracy of the reconstructed atlases and their cross alignment (Aim 1), the performance of 2.5DBA (Aim 2), and the intrinsic fiducial–guided spline interpolation (Aim 3). In this section we explain how these test data will be evaluated.

Aim 1: Construction of mouse 3D brain atlases and navigation software

3D atlas reconstruction

To evaluate atlas reconstruction accuracy, the contiguous section faces on consecutive atlas sections will be examined. High-resolution images will be collected; bisected cells will be identified manually and their misregistration error measured. Figure 5 demonstrates convincingly that such a match between sections can be achieved. We estimate the accuracy of matching fiducial point sets—centroids of cell nuclei—to be better than 4 μm (including error in converting slide coordinates to aligned-image coordinates). With the aid of the iScope (Project 2), we will collect 25 well-dispersed fiducials in each of 15 contiguous section pairs from all the different atlas animals. For this purpose, iScope accuracy will be enhanced by calibrating the position using intrinsic fiducials. The three section pairs to be examined from each animal will be collected from the ventral, middle, and dorsal regions. Our objective is to reduce misalignment to below 10 μm as determined by the error of the 14th worst measure of the 75 measures on each animal (i.e., ~ median + 1 standard deviation).

Cross-atlas alignment

Inter-atlas transformations will be validated by transforming region delineated in one brain to another and then comparing the alignment. An additional quantitative validation will be performed by leaving out one knot, computing the spline on the remaining knots, evaluating the error on the remaining knot, and cycling over the points (leave-one-out paradigm). We will seek to achieve an accuracy of 20 µm.
AIM 2: Linear alignment of the MBL into the standard coordinate system

2.5-dimensional alignment

Performance of 2.5DBA will be assessed using two methods. In the first, we will employ 30µm-isotropic data sets (male C57BL/6J) similar to the one shown in Figure 3. These will be aligned to the male C57BL/6J atlas (cresyl violet–stained) using 3DBA and 2.5DBA. For the latter, we will use every fifth section (i.e., the same unimodality sampling as for the MBL data). Taking the result of the former alignment as our gold standard, we will assess misregistration of the 2.5DBA. The placement error can be reduced to the maximum distance (the maximum distance error on the outer section contour) between the position of a given section under the two alignment regimes. Our objective is to obtain a mean of the absolute value of that error across the entire brain of less than 200 µm (i.e., a mean rotational error of less than 1.5°, or a displacement of less than 7 pixels).

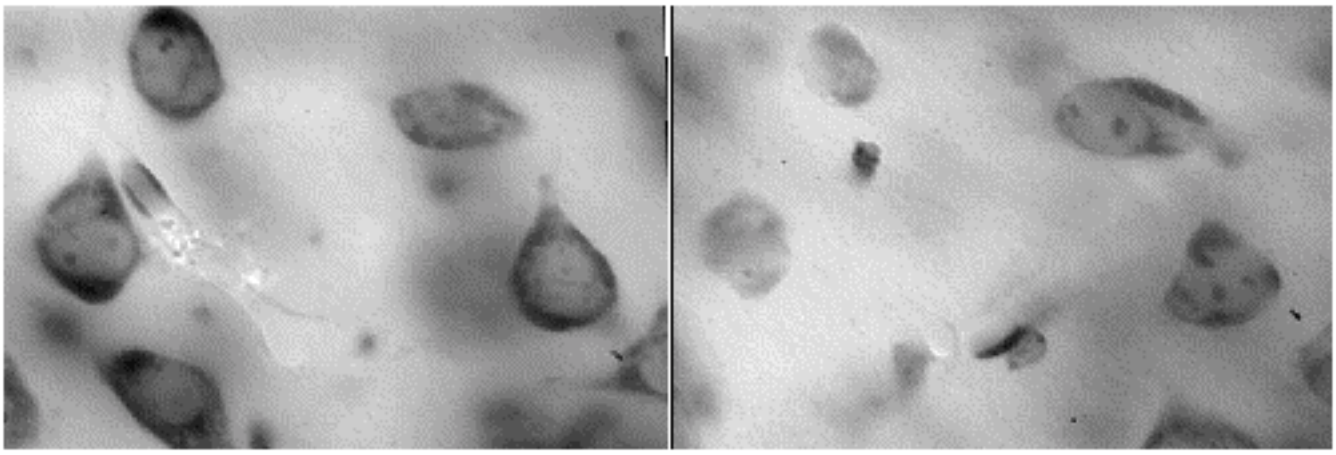


Figure 5. Cell level matching on consecutive sections. Continuity of blood vessel (*) as well as individual cells (numbered) can be followed across the cutting plane between consecutive sections (A & B). The sections are the top and bottom plane, respectively, of a high-magnification z-axis series of each section taken with a 1.25 NA 100x plan-achromatic objective, narrow-band green illumination (546 ±ß10 nm), and video-enhanced differential interference contrast (Nomarski) optics. For this test we used a Hamamatsu 1-inch ultraviolet-visible Ultricon (N983) camera. The full NA of the objective was used during imaging. In-plane optical resolution is ~0.22 µm. The field area is 84 x 67 µm. The actual pixel size of images is 633 x 473 pixels, giving a pixel area approximately 0.13 x 0.14 µm, a 6% X-Y distortion. The vertical video resolution is approximately 0.14 µm divided by the Kell correction of 0.7 (Inoué 1986, p. 171). This gives a final video resolution of 0.20 µm, closely matched to the optical resolution. The effective resolution of video images is lower than the modulation transfer function cutoff point. These images have an effective resolution is 0.3 to 0.4 µm.

In a second performance task, we will compare matching accuracy between 2.5DBA and experts. As described below, human matching accuracy in the framework of rigid body is about ±225 µm (using the same measure as described above). We are seeking to roughly match this level with the automatic system.
AIM 3: Segmentation of the MBL and extraction of quantitative traits
Implicit fiducial–guided spline alignment.

To test the virtual fiducial algorithm, we will use 45 MBL brains with 100 manually defined fiducials (due to the coarser z-plane of these relative to the atlas, delineation will take only about 1 month). Thirty-two of the brains will be aligned to the atlas using manual VOIs. Stable fiducials will then be automatically generated by examining the correlation at fixed standard coordinate positions across the data set. The stable set will than be used to align the remaining 13 brains, and the accuracy of the alignment will be assessed. Our objective is positional accuracy of 36 µm (the displacement of a 250-µm-diameter circular ROI leading to a 20% false positive and 20% false negative rate).

POTENTIAL PITFALLS AND POTENTIAL SOLUTIONS

Aim 1: Construction of mouse 3D brain atlases and navigation software

3D atlas reconstruction

Reconstruction using fiducials located by the correlation filter may fail. Should implicit fiducials defined using a matched filter prove insufficient for reconstruction of any given atlas, point fiducials will be obtained using bisected cells or blood vessels (the iScope will be used to acquire images of these). While this approach is straightforward and is likely to yield highly accurate reconstruction, it requires a high degree of operator interaction and thus is not as attractive as using matched filters. If this approach is required, it will add another 6 months to the atlas construction phase of the work. However, it will introduce only a moderate delay in the alignment phase of Specific Aim 2, which can commence when one atlas becomes available.

Cross-atlas alignment

Failure to achieve acceptable cross-atlas alignment as proposed will necessitate use of a larger number of fiducials. Since each brain will eventually be delineated, all VOIs can be used for this purpose.

AIM 2: Linear alignment of MBL into the standard coordinate system

2.5-dimensional alignment

The possibility exists that 2.5DBA might fail. A fallback approach to registration is the use of semiautomatic tools already in place in MacOStat. In this approach, an operator visually selects the homologous atlas plane for the first and last sections of a slab of tissue. The matching atlas planes for the intervening sections are defined by interpolation. Each experimental section is then automatically aligned to its homolog using an affine transformation.
To test the accuracy of this approach, we aligned 3 rat brains to our rat atlas. Each brain consisted of >90 coronal sections, and all the sections were matched independently to the atlas. Matching accuracy was assessed by comparing placement of sequential sections under the assumption that adjacent sections should match parallel atlas planes with a distance equal to the known intersection interval. The deviation from that expectation was defined to be the error and was measured as the maximum distance error on the outer section contour between the expected and the assigned position. That error was measured to be 212 μm. In this manual mode, alignment requires about 2 hours per brain.

AIM 3: Segmentation of the MBL and extraction of quantitative traits

Implicit fiducial–guided spline alignment

Automatic extraction of fiducials may fail. The fallback position is manual delineation. From the experiment described above (see the feasibility assesment section above), we will establish the number of fiducials required to achieve adequate alignment accuracy. Assuming all 100 are needed, it would take approximately 2.5 years to generate and align 1200 MBL brains. The desired positional accuracy of 36 μm (3.6 x the resolution of the atlas) might be unattainable even with all 100 fiducials. If that is the case, the MBL will be segmented as best as can be achieved, and finer accuracy will demand manual adjustment of the anatomical templates on a case-by-case basis.
While we are aware that the positional exactness sought is ambitious, we believe that given the homogeneity of the MBL, the algorithm will support this degree of accuracy on most brains. The quality control test embedded in the procedure will flag those brains in which performance is poorer. For those we will manually refine the segmentation using NeuroTerrain.

Means of assessing success with respect to Phase II objectives

One major objective of Phase II is to refine neuroinformatics tools. The central feature in this project is the use of a standard coordinate system to provide database access to anatomical images in the MBL. BrainMap is a well-developed database built around this idea. Our contribution will be a substantial refinement in accuracy, the development of new web-based software with a powerful user interface, and a new application setting.
As a neuroanatomical database supporting navigation and morphometric investigations at high resolution over the Internet, the MBL breaks new ground. Given the clearly defined need for image warehouses of histological material, we are optimistic that our approach and our software will soon be adopted by other neuroinformatics centers. To support the spread of this technology, all software along with source code will be made freely available. At the Computer Vision Center for Vertebrate Brain Mapping, we have had many years of experience in engineering software for the neuroscience community. The first personal computer–based autoradiographic software was developed at our center (McEachron et al. 1987), and we have been developing sophisticated software for neuroscientists ever since. Applications now run on MacOS, UNIX, and Windows in over 300 laboratories worldwide. We have extensive experience not only in software development but also in software testing, support, and documentation (Gustafson 1995, 1997, 1999; Nissanov 1998).

Legacy data and other collections

In the context of the present proposal, the MBL is exclusively produced in-house. However, our approach is generic, and in the future we plan to incorporate material from other sources. For example, the entire Yakovlev-Hakem collection of

human brains at the Armed Forces Institute of Pathology, which were also processed in celloidin and cut at 35 µm, is an obvious candidate. Relaxation of the strict protocols for tissue processing and data collection imposed in Project 1 will initially limit the scope of material that can be segmented with accuracy. Yet we believe that modification of the algorithms will eventually make it possible to incorporate a wider range of data sets without compromising the precision of segmentation. Certainly at this stage, accommodation of legacy data can made at the level of anatomical homology. Indeed, this mode of linkage may have the greatest repercussions in neuroscience and clinical research. At this level we may provide a gateway to genomics databases for other vertebrate brain phenotyping efforts. For example, consider the following scenario: An abnormality is discovered in a functional neuroimaging study (e.g., Gur et al. 1995) and indicates a serious structural abnormality (e.g., Arnold et al. 1995). If variations in the corresponding structures have already been identified in the MBL, then it may be possible to establish a link between relevant QTLs and genes in mouse that may underlie the abnormality in humans. Anatomical and genomic homologies are essentially cross-referenced.

Assessing success

How can one assess success of this project in the context of the Human Brain Project? Assessment should be simple: at one level we can simply look at patterns of usage of our web sites. Over the past six months use of <nervenet.org> resources has surged—almost entirely due to the addition of a highly detailed atlas of C57BL/6J and to the MBL's image collection. In the long run our success will be measured by the scientific breakthroughs this work will facilitate and engender. We should expect a dense map of QTLs that will provide the research community with a framework to explore the genetic basis of structural differences in CNS. For Project 3 in particular, one key metric to assess success is the accuracy of segmentation and the accuracy of quantitative trait values NeuroCartographer it will supply to the NTB (see Project 4). Finally, our success will also be measured by the degree to which the approaches and software we develop are adopted and modified by other neuroinformatics centers.

RELEVANCE OF PROJECT 3 TO THE WHOLE PROGRAM

This project is closely linked to all other projects of this program. It takes data generated in Projects 1 (MBL) and 2 (iScope) and returns tools and data for navigating through a sea of images and movies. It provides the iScope with standard coordinates to guide acquisition of high-resolution micrographs, as well as an indexing system for the resulting high-resolution image stacks. In turn, NeuroCartographer will rely on the iScope to test atlas reconstruction accuracy. NeuroCartographer will feed the NTB with the essential quantitative trait measurements required to perform complex trait analysis. NeuroCartographer will also provides the basis, in the form of standard anatomical divisions, for interconnecting to other neuroanatomical databases.

RESEARCH PLAN

THE NEUROGENETICS TOOL BOX

Our plan in this section of the application is to develop unique World Wide Web services that will make it possible for neurogeneticists to rapidly identify and map genes and quantitative trait loci, particularly those related to brain structure and behavior. The Neurogenetics Tool Box (NTB), which our group at the Roswell Park Cancer Institute will design, test, and distribute on the web, will be tightly integrated with the MBL and NeuroCartographer Projects. Synergy between these components of the program project will allow neuroscientists and geneticists to explore the complex genetics of CNS architecture and to map QTLs much more efficiently than is now possible. The NTB will consist of three software modules and complete online documentation and tutorials.

1. The NTB biostatistical evaluation module

An essential prelude to gene mapping is to analyze the statistical properties of the trait, in particular, its normality, its covariance with other phenotypes, and of course, its heritability. We will construct a web interface using Python scripts, CGIs, and custom server software that will allow neuroscientists to submit quantitative trait data from any sample population. In response, this NTB module will rapidly return (1) key descriptive statistics, (2) results of tests of normality and homoscedasticity, (3) correlation matrices between submitted traits and other selected traits (e.g., brain weight, body weight, age, and sex), (4) estimates of narrow and broad-sense heritability, and, when appropriate, (5) estimates of the minimum number of effective genes controlling CNS traits. This system will include modules which clients will use to perform transformations of trait values to improve normality, to control for nongenetic variance (e.g., age), to test for outliers, and to test for uniformity of variance across complex data sets. Clients will be able to rapidly evaluate the suitability of their phenotypes for gene mapping without becoming expert statisticians.

2. The NTB gene and QTL mapping module

The web interface of this system will allow rapid QTL mapping of CNS traits for major recombinant inbred (RI) sets, for shared intercrosses and backcrosses, and particularly for a very large advanced intercross population (the G_{10}) described in the Genotyping and Mouse Colony Core. The NTB will include curated data for marker genotypes and map distances for all crosses and RI strains in the Mouse Brain Library. Using the resources developed by this program project, neurogeneticists will be able to generate phenotypes, then feed the data directly into the NTB's biostatistics and gene mapping modules. Furthermore, the NTB will include an extensive legacy database on behavior and neuropharmacology of RI strains. (The core of this database will be provided by John Belknap, a member of the External Advisory Committee.) It will therefore be feasible to explore the genetic basis of differences in a broad range of CNS traits.

The NTB gene mapping module will initially be based on algorithms already successfully implemented in the programs Map Manager QT and Map Manager QTX (Manly and Olson 1999). The NTB will perform single-locus association tests, simple interval mapping, and composite interval mapping and will also search for epistatic interactions between QTLs. The mapping module will be implemented in Python and C++, with Python used for interaction with web forms and data formatting and C++ for the analysis routines.

For functions in which processing speed is not limiting, results will be returned to clients immediately in the form of web pages; for lengthy calculations, data sets and figures will be returned by email. This system will include empirical significance thresholds established by permutation tests or by a new method that promises to be much faster (H-P. Piepho, personal communication). Our NTB mapping program will also allow for weighted linear regression for those data sets in which the variance of phenotype values is available, and it will provide nonparametric mapping methods for traits with distributions far from normal.

The web interface of the mapping module will also allow clients to submit private data sets (trait values, marker genotypes, and map distances) for QTL analysis in crosses that are not curated by the NTB. Clients will have access to the same range of analysis options as described above except that correlation with other traits will necessarily be limited to those traits submitted with the population marker data. A high level of confidentiality will be guaranteed to clients using this automatic gene mapping system. As part of this system, we will assemble and maintain an extensive database on phenotypes and genotypes for the major recombinant inbred strains (BXD, AXB-BXA, CXB, BXH, AKXD, etc).

A powerful and unique feature of the gene mapping module is that it will allow investigators to search for genetic correlations among large numbers of traits and QTLs in our NTB archive. The NTB databases will become progressively more powerful as more traits are mapped on the same crosses. We anticipate that networks of gene interactions and gene pleiotropy will be exposed.

3. NTB data archiving and export modules

This subsystem will allow clients to permanently archive published and unpublished results for our NTB advanced intercross and major RI strains on the NTB server. All data generated by Drs. Rosen, Williams, and Nissanov as part of the program projects will be archived, but we anticipate that many other neuroscientists and geneticists will want to build

onto the data sets in the NTB. Several options for security level and release conditions will be provided to our clients. We will also make it possible for clients to export trait data combined with the NTB's curated marker genotypes and map distances in formats suitable for analysis using Map Manager, QTL Cartographer, or conventional text editors. These data will be sent to the requesting user within their browser, by email, or by FTP.

4. NTB documentation and tutorials

The functions of the NTB will be explained in a user manual that will be available as an electronic hypertext document. The manual for Map Manager QT, included as appendix item, provides an example of the proposed documentation. For those clients with high-speed connections, streaming video tutorials (QuickTime format) will demonstrate the process of mapping QTLs using the NTB.

PROBLEMS AND ISSUES ADDRESSED

Many genetic traits of medical and economic importance are complex quantitative traits. Such traits are usually affected by numerous genetic loci and also have significant environmental or nongenetic components. A complex trait is typically measured as a numerical value rather than a categorical phenotype. In recent years, with the development of numerous easily typed genetic markers based on DNA sequence polymorphisms, it has finally become practical to identify many of the individual loci that contribute to quantitative traits in mammalian populations (Lander and Botstein 1989). In the introduction to this application, we reviewed the impact that complex trait analysis (also known as QTL analysis) is now having on structural, functional, and behavioral analysis of the CNS. In fact, in mice, a strong case can now be made that complex trait analysis is proving to be as important in the rapid advances in neurogenetics as is the analysis of transgenic, knockout, and mutant lines. In this section of the application, we provide a much more comprehensive bioinformatic and technical review of QTL mapping and the programs and utilities now used in this rapidly expanding field. We demonstrate that the work we have carried out up to the present meets and exceeds the standards of a Phase I Neuroinformatics Project. We explain how, with a relatively modest investment of funds, this project will significantly enhance the analysis of complex traits of the mouse CNS. The PPG should put QTL analysis into a new bioinformatic sphere in which collaboration becomes the new norm.

QTL analysis is a highly specialized area of research that may be unfamiliar even to readers with a strong genetics background. Recent papers and a tutorial by the PI (KM) and the coPI (RW) are included in the appendix and are also available online at <mcbio.med.buffalo.edu/mmQT.html> and <nervenet.org/papers/papers.html>.

The Human Genome Project has provided extensive DNA sequence information for genomes of humans and model organisms. In the near future, this information is expected to be far more comprehensive and nearly complete. However, connecting this sequence information with traits of medical and economic importance will remain a challenge for many years. In many cases, the first step for connecting traits with sequence information will be QTL mapping, both in humans and in mice. QTL mapping can enumerate genes contributing to a trait, indicate an approximate chromosomal location for each gene, and estimate their relative importance. In this sense, QTL mapping is a classic forward genetic approach that starts with phenotypes and particular biological problems and then searches for genes that contribute to or modulate variance in those phenotypes. This approach is a compelling advantage for neuroscientists, who almost invariably become interested in the genetics of a specific phenomenon or mechanism because they have spent years studying it—the migration of neurons, modulation of cell proliferation and cell death, the control of receptor density, or even the source of emotionality. Traits like these can be systematically explored or genetically dissected using QTL methods (Williams 1998, 2000).

The feasibility of QTL mapping and its potential importance have stimulated theoretical work on mapping methods and software development to implement those methods. This work has included methods designed for human populations, where pedigree methods are necessary, and those designed for experimental organisms, where inbred lines are available. Although QTL mapping is now feasible, it remains beyond the reach of many neuroscientists because it requires specialized statistical methods and data sets consisting of tens or hundreds of thousands of genotypes derived from large crosses. Current mapping software makes many statistical methods available, but often in forms difficult for biologists to use. The required genotyping, which is expensive and time-consuming, is an even more serious deterrent. For example, as part of the PPG, our Genotyping and Mouse Colony Core will generate nearly 400,000 genotypes, a number far beyond the reach of most R01-supported laboratories. The Neurogenetics Tool Box, in combination with mapping resources and the huge collection of images and phenotypes generated in the other three projects, will substantially reduce these barriers and allow neuroscientists to exploit the power of QTL detection and mapping. (See the Introduction for specific examples.)

STATUS OF CURRENT RESEARCH

Background

QTL mapping includes three levels of increasing complexity. First is single-locus mapping: statistical tests of association between trait values and the genotypes of marker loci throughout the genome. A significant association is interpreted as indicating the presence of a QTL linked to the marker that shows the association. Since these tests consider only one locus at a time, they do not require any information about adjacent loci; that is, they do not require that marker loci be mapped.

The second method, simple interval mapping (SIM), requires construction of a marker genetic map. This method evaluates the association between the trait values and the expected genotype of a hypothetical QTL (the target QTL) at multiple analysis points between each pair of adjacent marker loci (the target interval). The expected QTL genotype is estimated from the genotypes of the flanking marker loci and their distance from the QTL. Since there is usually uncertainty in the QTL genotype, the analysis is based on a sum of terms (one for each possible genotype, weighted by the probability of that genotype). The analysis point that yields the most significant association may be taken as the location of a putative QTL.

The third method is called composite interval mapping (CIM: Zeng 1993; Jiang and Zeng 1995) or multiple QTL mapping (MQM: Jansen 1993). Like simple interval mapping, this method evaluates the possibility of a target QTL at multiple analysis points across each interlocus interval. However, at each point it also includes in the analysis the effect of one or more markers elsewhere in the genome. These markers, sometimes called background markers, have previously been shown to be associated with the trait and therefore are each presumably close to another QTL (a background QTL) affecting the trait. The background marker genotypes are used in place of the (unknown) background QTL genotypes even though there will usually be differences due to recombination between these markers and the QTL. The inclusion of a background marker in the analysis helps in one of two ways, depending on whether the background marker and the target interval are linked. If they are not linked, inclusion of the background marker makes the analysis more sensitive to the presence of a QTL in the target interval. If they are linked, inclusion of the background marker may help separate the target QTL from linked QTLs on the far side of the background marker (Zeng 1993, 1994).

Each of the above methods differs in the form of the equation which is assumed to relate a trait value to the genotypes of marker loci. For single-locus mapping in a backcross population, the equation has the form

$$Y_i = a + bx_i + e_i$$

where Y_i is the trait value for the ith individual, x_i is the coded genotype of that individual, e_i is the environmental effect in that individual, and a and b are parameters to be estimated. For SIM, the second term (bx_i) is replaced by a sum of terms for each possible QTL genotype weighted by its probability. For CIM, the equation has this same sum and has, in addition, a bx_i term for each background locus. For an intercross population, each bx_i term may be replaced by a sum of two terms, one representing an additive component and one representing a dominance component. For example, the equation for CIM in an intercross population might be

$$Y_i = a + b_a x_i + b_d x_i + \sum_j b_j x_{ij} + e_i$$

where b_a and b_d represent the additive and dominance contribution of the target QTL and each b_j represents the contribution of a background marker (assumed to be additive).

For all methods, each mouse or RI strain provides the data for one equation, and the statistical problem is to estimate a and b (or each of several bs). Two methods have been used for this problem: (1) least-squares regression and (2) maximum-likelihood estimation using the estimation-maximization algorithm. Regression is computationally simpler, but it makes assumptions about the distribution of trait values that are only approximately true. Nevertheless, in practice, regression seems effective (Haley and Knott 1992; Martinez and Curnow 1992; Jansen 1993). To avoid one of the assumptions about trait value distributions, it is possible to use a nonparametric statistic to evaluate the significance of an association, a measure based on the rank of each trait value rather than the trait value itself (Kruglyak and Lander 1995).

Recently, Whittaker, Thompson, and Visscher (1996) published a regression method which, for additive QTLs, is equivalent to the methods of Haley and Knott (1992) and Martinez and Curnow (1992) but which is computationally much simpler. This method allows the estimation of QTL position and effect from flanking-marker regression coefficients without fitting a regression equation at multiple points between the flanking markers. This increased speed will be especially welcome in the calculation of significance thresholds by permutation tests, described below. We do not want NTB clients to have to wait too long for results to be displayed.

A constant problem in QTL mapping is establishing appropriate significance thresholds. First, although a likelihood ratio statistic can be calculated and interpreted as a χ^2 statistic, this ratio is only approximately χ^2 because the number of progeny is finite (and sometimes small). Second, because mapping involves multiple tests across the genome, significance thresholds must be correspondingly more stringent (Lander and Kruglyak 1995). Permutation tests have been described which address both problems by calculating empirical thresholds specific for a data set and an appropriate range of the genome (Churchill and Doerge 1994; Doerge and Churchill 1996). These empirical thresholds are essential, and the NTB will support their calculation. However, the calculations are time-consuming, and a recently proposed method offers a comparable empirical threshold with greatly reduced calculation time (Piepho, personal communication; Davies 1977, 1987), so the NTB will offer this method as well.

Community mapping resources

The Genotyping and Mouse Colony Core describes a large advanced intercross that is already beginning to provide data to map and identify large numbers of loci that control variation in the architecture of the CNS. A major goal of the NTB is to provide Web-based software that will allow neuroscientists and geneticists, including those not currently familiar with QTL mapping, to collaborate with one another to exploit the resources of the MBL and NeuroCartography projects. The large size of this tenth-generation advanced intercross and its derivation by multiple generations of recombination will allow QTLs to be mapped with a precision of less than 4 cM. We expect that the discovery of genetic sources of structural variation will rapidly catalyze correlative analysis of functional and behavioral traits. The recombinant inbred strains mentioned in the next paragraph (Belknap 1998, Williams 1998) provide a superb resource for extensive brain-behavioral correlative genetic analysis.

Sets of recombinant inbred (RI) lines will provide another major shared resource to be exploited with the aid of the NTB. The Mouse Brain Library already has a very extensive collection of brains from RI strains—an average of more than 6 animals from each of 35 BXD strains and ~24 AXB/BXA strains. In the first few years of the grant we will add at least 30 RI strains (for example, the BXH and CXB strains are already in progress), for a cumulative set of well over 80 RI strains in the collection. RI strains provide multiple individuals for estimating trait values (Belknap 1998). Using average trait value for 10 or 20 individuals greatly reduces the nongenetic variance, in effect, giving the trait a higher heritability. This, in turn, increases the likelihood of success in mapping QTLs (Darvasi and Soller 1997). Using this approach, Toth and Williams (1999 a, b, c) recorded EEGs from an average of 12 animals from each CXB RI strain, and they succeeded in mapping at least one QTL controlling the sleep cycle (REM vs. slow-wave sleep) in mice—a trait that normally has a very low heritability ($h^2 < 0.3$). Another key advantage of RI lines is that hundreds of different types of traits (behavioral, neuropharmacological, neuroanatomical) can be studied by many investigators over a period of years using one fixed set of genotypes. All data, whether genotypes or phenotypes, are cumulative. Using these far more extensive sets of data on entire strains of mice, investigators can detect genes and QTLs that have diverse or pleiotropic effects. The main limitation of RI strains is that their statistical power tends to be low due to the relatively small number of strains in each set. But this is not a major problem at this early stage of analysis, before any QTLs affecting CNS structure have been mapped. In fact, we would like to start with QTLs that have major effects because they may have more prominent behavioral consequences.

Current software

Nearly a dozen software packages are currently available for mapping QTLs (Manly and Olson 1999). Despite this apparent variety, only Map Manager QT is a noncommercial, general-purpose mapping program designed from the beginning with an effective graphic user interface that can be mastered within a week or two. QTL mapping is a somewhat esoteric field. Despite Map Manager QT's accessible interface, many users, particularly those who have not been immersed in mouse genetics for several years, find mapping QTLs to be challenging. They need an experienced collaborator. Most of the neuroscientists and even many of the neurogeneticists using the MBL and NeuroCartographer will be expert neuroanatomists, developmental or molecular biologists, or stereologists, but they will not be expert statisticians or quantitative geneticists. In order to effectively exploit the raw mapping resources we will offer, clients will need accessible and well-documented software tools to act as a remote collaborator. With this resource, scientists will be able to do QTL mapping without getting totally immersed in quantitative genetics. This is the major interface challenge in the NTB project—devising a web interface that allows any neuroscientist to access powerful gene mapping utilities in a way that makes their use as simple as possible.

As discussed below, the QTL detection and mapping functions of Map Manager QT and Map Manager QTX are prototypes for the corresponding functions proposed for the NTB. We will be able to adapt and extend the code from Map Manager QTX for use in the NTB. The code will initially be adapted for the NTB by simplifying it—removing some features that are not required by the uncluttered user interface of the NTB. Once this "first-pass" interface has been successfully implemented, and after the NTB has been tightly linked with our extensive phenotype databases, we will progressively add more sophisticated mapping options and novel features for mapping epistatic interactions.

Existing Map Manager software

The Map Manager family of genetic mapping software consists of four programs, the latest of which will be released before the start of this program project. These programs have been developed with support from the National Human Genome Research Institute and the Rockefeller Foundation. The programs and user manuals have been distributed without charge.

The first two existing programs are described fully in the user manuals for Map Manager QT included in the appendix. Map Manager Classic maps Mendelian loci in recombinant inbred lines, backcrosses, and intercrosses with codominant markers. It was started in 1988 as a program to manage and use the rapidly growing volume of data from mouse recombinant inbred lines (Manly and Elliott 1991), and additional functions were added later (Manly 1993). Map Manager Classic runs only under Mac OS and is written in Pascal. Map Manager QT, an augmented version of Map Manager Classic, includes functions for the detection and mapping of QTLs by regression-based single-locus association, simple interval mapping, and composite interval mapping. It also provides permutation tests for establishing empirical significance levels for putative QTLs. Map Manager QT has been used extensively in the preliminary work leading to this proposal.

Map Manager XP is a mapping program for Mendelian traits, but unlike Map Manager Classic, it accepts dominant alleles and populations in which adjacent loci show different segregation patterns. It is written in C++ and, more important, uses a commercial portable library called XVT. This allows the code, with minor modifications, to be compiled for a variety of machines and/or operating systems. Map Manager QTX, an augmented version of Map Manager XP with the same functions for detection and mapping of QTLs as Map Manager QT, will handle additional types of mapping populations, including advanced backcrosses and advanced intercrosses such as the G_{10} intercross described in the Genotyping and Mouse Colony Core. The analysis routines in Map Manager QTX will be adapted for use in the NTB. The first version of Map Manager QTX will be released early in 2000. In addition to providing code for the Web-based mapping functions proposed here, Map Manager QTX will itself be used by members of this program project.

Limitations of existing approaches

All QTL mapping programs require that the user provide a large data set of marker genotypes (typically tens of thousands of genotypes) and, in most cases, a corresponding set of map positions for the markers. For the nonspecialist, obtaining these data in a suitable format can be a significant barrier to QTL mapping, even when the data are publicly available. We discuss this problem at greater length in the Introduction. The NTB will be a major resource in mouse neurogenetics, and one that we think will rival ongoing mutagenesis programs in impact. The NTB will include all of the necessary markers, genotypes, and map positions in a data structure that requires no attention from neuroscientists interested in tracking down genes that control any region or nucleus in the CNS. We will also have a very extensive preloaded list of phenotypes in the NTB, initially key parameters for multiple regression analysis such as sex, body weight, brain weight, age, etc. Eventually, when the NeuroCartographer Project has segmented the majority of brains in the MBL collection, the NTB will contain several hundred phenotypes for each animal and RI strain. Once the NTB system is in place, it will be relatively simple to extend it to other test crosses. We expect to get many requests to curate new QTL data sets. For example, an advanced intercross between C57BL/6J and A/J to complement our AXB-BXA RI set would be a very welcome addition.

NOVELTY OF THE NEUROGENETICS TOOL BOX

The novelty and utility of this project lie in its integration with the image libraries of the Mouse Brain Library, the data of the NeuroCartographer Project, and the genotypes provided by the Genotyping Core. Six innovative features will be integrated into the NTB over this 5-year grant cycle.

1. Analysis methods not previously implemented. The fast interval mapping method (Whittaker, Thompson et al. 1996) has only recently been implemented in Map Manager QTX; it is not otherwise available in mapping software. The fast method for empirical significance thresholds (Peipho, personal communication) is not currently available in any mapping software.
2. The integration of mapping software and genotype data. All current mapping programs come “empty.” A research group has to generate both genotypes and phenotypes. In contrast, the NTB will come loaded with a very large set of genotypes as well as with all the phenotypes automatically generated by the NeuroCartographer Project, which are simply a part of the MBL database (age, sex, body weight, brain weight, litter size, etc). We will have done much more than half of the data generation neuroscientists need for mapping QTLs. The NeuroCartographer Project will produce several hundred quantitative traits for each RI strain and for each animal in the advanced intercross. These values will be available to QTL mappers.
3. Direct web access to QTL mapping tools. We intend to start with a simple and robust web interface in which investigators paste their phenotypes into a form with tab-delimited lists of values and then select or define trait names. For CNS phenotypes already in the NTB, users will be able to remap quickly to compare one trait with another. While web access to QTL mapping has been tried before (most notably in the Kearsey and Seatom QTL Café Project, 1997–1998, at the University of Birmingham), there is a major difference in our

project. In our case, most of the data used for mapping will already be resident in the NTB. Neuroscientists will primarily be adding small numbers of traits. This fact gives us more control over the design of the interface. For example, we know exactly how many cases there are, and we will know how to display their genotypes most effectively over the web. A prototype form for this purpose is illustrated and explained below.

4. Correlative analysis of multiple traits. This is crucial for analysis of the CNS, in which different structures are often interconnected. For example, a group studying cell populations in the septal region will of course be very interested in knowing what genes control the size of the hippocampus. These types of correlated genetic studies have never been possible before because phenotype data have never been curated. This is even true for the RI strains. There is now no accessible database of the phenotypes of RI strains. In contrast, there are several sources for the genotypes of RI strains (B. Taylor, R. Elliott, R. Williams, and the Jackson Laboratory all maintain RI genotype databases). As part of the NTB, we will assemble phenotype databases for RI strains as well as for all of the advanced intercross animals and standard inbred strains in the MBL.
5. Help in the early stages of QTL mapping. The neuroscientists who generate the trait values will often not have a strong statistics background. They will need help to evaluate the suitability of traits for QTL mapping. For example, we expect many neuroscientists to submit percentage and ratio data to the NTB: the ratio of the volume of the caudate nucleus to the substantia nigra, or the percentage of glial cells in the neocortex. These types of data will often need to be transformed and normalized prior to mapping. The NTB will include a set of web utilities that will screen and suggest how best to modify a data set prior to mapping.
6. Extensive interactive online tutorials on QTL mapping. The statistical and bioinformatic aspects of QTL mapping need to be made more tractable for neuroscientists. Complex trait analysis should make a significant contribution to neuroscience. Successful penetration of this powerful method requires the right learning environment. A tutorial specifically tuned for neuroscientists is online at <nervenet.org/papers/shortcourse98.html>, and an extensive general tutorial is online at <mcbio.med.buffalo.edu/mmQT.html>. These tutorials will be expanded and made significantly more attractive. We will even add short QuickTime movies using the expertise of the iScope project in video editing and QT movie production. We will step neuroscientists through the process of mapping QTLs using specific data sets from the MBL and the <nervenet.org> databases.

TIME COURSE OF DEVELOPMENT

Experience with development of the Map Manager programs has taught some important lessons about development of this type of software. The most important is that development of an easy-to-use interface requires multiple cycles of development, user feedback, and modification. This process takes longer than it seems it should. With that in mind, we plan on the following time course for developing and testing services that will be made available to the scientific community.

Year 01: Basic descriptive statistics and heritability measures for traits. Single-locus association and fast interval mapping for curated data sets from recombinant inbred lines. Fast empirical significance thresholds and/or permutation tests for single-locus association and fast interval mapping. Documentation and tutorials for these functions.

Year 02: Single locus association and standard interval mapping for the G_{10} advanced intercross. Graphic display of interval mapping results. Fast empirical significance thresholds and/or permutation tests for standard interval mapping. Data export in formats for other mapping software. Documentation and tutorials for these functions.

Year 03: Composite interval mapping for all curated data sets and empirical significance thresholds for these. Correlation of submitted traits among each other and with traits stored in the NTB. Documentation and tutorials for these functions.

Year 04: Detection and analysis of epistatic interactions between QTLs. Analysis of crosses in which both trait values and marker genotypes are submitted by the user. Documentation and tutorials for these functions.

Year 05: Bootstrap methods for confidence intervals on QTL location. Weighted regression and nonparametric regression methods. Documentation and tutorials for these functions.

EXPERIMENTAL PLAN

Staffing and oversight

The programmer requested for this project will be responsible for coding the functions of the NTB under Dr. Manly's direction. This programmer will also perform some initial testing of the system but will not be responsible for exhaustive testing. The intern to be hired in Year 02 will be responsible for exhaustive testing, under Dr. Manly's direction, using simulated and real data. The Genotyping and Mouse Colony Core, under the direction of Drs Williams and Gu, will generate the marker genotypes, in particular the genotypes of the G_{10} advanced intercross, needed for operation of the NTB. Genotypes needed for QTL mapping with recombinant inbred lines will be selected from public data by Dr. Williams or Dr. Manly. Documentation and tutorials will be written by Dr. Manly with advice from Dr. Williams.

Once the NTB is operational, the production servers will be maintained at the University of Tennessee, Memphis, by Dr. Williams. Both Drs. Manly and Williams will have access to these servers, and they will be jointly responsible for maintaining the software and associated data files.

Communication

Development of the NTB will not require large-scale or high-speed data transfer between Buffalo and Memphis. The code developed in Buffalo will be tested locally with local copies of the genotype data files, then transferred periodically to the production server in Memphis. The marker genotype data needed by the NTB, although extensive, will be relatively static, and copies in Buffalo will need updating only infrequently. Although we may establish a shared file system between Buffalo and Memphis, ftp would be sufficient.

Interval mapping

The heart of the interval mapping method is a set of expressions that give the expected effect for a QTL in the target interval, dependent on the analysis point (the hypothetical position of the QTL in the interval), the genotype of the flanking markers, the population type, and the dominance and interference models. The options for dominance models and interference models described above are implemented by choosing the appropriate set of expressions. Sets of these expressions for different situations have been published by many authors (Lander and Botstein 1989; Knapp, Bridges et al. 1990; Carbonell, Gerig et al. 1992; Haley and Knott 1992; Martinez and Curnow 1992; Moreno-Gonzalez 1992). Similar expressions for dominant markers have apparently not been published, but, assuming normal segregation ratios, these can be created as appropriately weighted averages of the terms for codominant markers.

Establishing QTL significance

After estimating regression coefficients under the assumption of a normal error distribution, the NTB, like Map Manager QT, will calculate a likelihood ratio statistic or LRS (Haley and Knott 1992) as a measure of the significance of a hypothetical QTL. For interval mapping, it will calculate this statistic for every analysis point in the intervals considered. The point at which the LRS is a maximum is interpreted as a possible location of a QTL, and the value of the LRS at that point is interpreted as a measure of the statistical significance of that QTL. The LRS can be interpreted as a χ^2 statistic. Strictly speaking, it is an approximate χ^2 statistic; but, by comparing this LRS with the likelihood ratio calculated by the maximum-likelihood method in Mapmaker, Haley and Knott (1992) showed that the approximation can be quite good. In cases for which the LRS is based on one degree of freedom, the LRS can also be converted to a traditional base-10 LOD score. The significance of a QTL can be estimated by comparing the maximum LRS value with significance thresholds for χ^2 statistics or LODs. However, permutation tests provide significance thresholds that are not affected by the "mixture problem" or by the distribution of the environmental effect. In addition, a new empirical method has been proposed that promises to provide empirical thresholds with greatly reduced computational time (Piepho, personal communication). This new method is a specific application, validated by simulation, of an approximate threshold formula proposed by Davies (1977, 1987). The NTB will provide one or both of these functions for generating empirical significance thresholds, and neuroscientists will be encouraged to use these methods routinely.

Graphical display of interval mapping analysis

Like Map Manager QT, the NTB will use graphical display of analytical results. An example of an interval mapping of simulated F_2 data by Map Manager QT appears in Figure 1. The NTB will display these figures in color in an HTML page; such figures can be copied and pasted into a graphics program or saved as a file.

Although these figures have become a standard method for presenting QTL mapping results, they are largely window-dressing. The only useful information they convey is the maxima for the LRS, points which indicate possible QTL positions. Although the shape of the LRS curve has been interpreted as giving a confidence interval for QTL position (Lander and Botstein 1989), this method is not reliable (Ooijen 1992; Mangin, Goffinet et al. 1994; Darvasi and Soller 1997).

For some types of crosses, the graphical display is unnecessary. As explained below under Fast Interval Mapping, Whittaker, Thompson, and coworkers (1996) have published an elegant method for additive QTLs which determines LRS maxima directly, This method is implemented only in Map Manager QTX. The NTB will use this method for mapping in recombinant inbred lines, which detect only additive QTL components, reserving the graphical display for intercrosses, which can detect dominance components.

Precision of QTL location

The NTB will provide two estimates of confidence intervals for QTL position, both empirically derived. Lander and Botstein (1989) proposed a simple theoretical rule for constructing confidence intervals for QTL position: the position of the maximum LOD score is taken as the position of the QTL, and the region in which the LOD score is within one LOD unit of the maximum is taken as a 96.8% confidence interval. However, this method is not accurate except for strong QTLs. Better methods for confidence intervals have been derived by theoretical analysis (Mangin, Goffinet et al. 1994) and simulation (Ooijen 1992; Darvasi and Soller 1997). For moderate QTLs, Darvasi provides the expression 530/Nv for a 95% confidence interval, where N is the population size and v is the proportion of variance explained (Darvasi and Soller 1997; Darvasi 1998). However, a bootstrap method seems to be the most accurate current method for estimating a confidence interval (Visscher, Thompson et al. 1996). The NTB will provide both of these estimates.

Software design

The NTB will be written largely in Python and C++. Python <www.python.org> will be used for interface routines that present online forms and/or interpret the information returned in those forms. We anticipate that these interface routines will change more frequently, in response to user suggestions, than the analysis routines, and the flexibility of an interpreted language will be useful in implementing those changes. Python was recommended for this purpose by K. Jacobs, one of the designers of the graphic user interface for Statistical Analysis for Genetic Epidemiology <darwin.cwru.edu/sagegui/intro.html>. Speed will be important for the analysis routines, so these will be written in C++. Most of these will be simplified versions of code written for Map Manager QTX. The analysis routines can be integrated with the Python code by compiling them as new Python modules. We may also use Java applets for some functions in the data entry forms, but in the interest of making the forms as responsive as possible, these will be kept to a minimum and/or made optional.

Software produced by this project will be copyrighted to control modification and distribution, but it will be made available without charge, either as executables or as source code, to nonprofit organizations.

Data entry, editing, importing, and transformation

The NTB will provide at least two methods for submission of trait information. The more general method will be a form with an entry field into which a list of strain names (or progeny numbers) and trait values can be entered. This method will require the user to follow some simple formatting rules in entering the data. The second method, for shared data sets such as RI lines, will provide separate forms for each shared data set, in which each strain or individual will be represented by a separate entry field for the trait value. This method will eliminate all data formatting problems. Figure 2 shows a prototype of such a submission form. The user need enter only a trait name and trait values for each inbred line and then

click the Submit button. If we can do so without sacrificing responsiveness, we will use a Java applet to perform sanity-checking on the trait values (detection of extreme outliers) to alert the user to possible entry errors before the data is submitted. The same form will be used to request descriptive statistics, single-locus association test, and simple interval mapping. A form for composite interval mapping would be somewhat more complicated, because the user would be given the option of choosing among loci previously shown to be associated with the trait.

It is often useful to be able to rescale trait data or to transform it with simple functions to make the distribution of the data more nearly normal or to equalize the variance of the different trait phenotypes. Such a transformation does not affect the problem described above as the "mixture problem," but it may reduce non-normality derived from the distribution of the environmental effect and thus improve small sample accuracy of the χ^2 approximation. The NTB will offer a number of transformation functions and will offer a probit plot of the trait distribution, a graphical indication of the distribution's normality.

The early implementations of the NTB will focus on quantitative traits expressed in metric data. This is the most common type of trait data. Traits expressed in ordinal data (relative rank) will be addressed in later years when nonparametric methods are implemented.

Marker redundancy in data sets

QTL mapping of marker loci in recombinant inbred mice uses data sets that were not developed solely for this purpose. Such data sets have many markers that are, for QTL mapping purposes, redundant. Markers that are not separated by recombinations, and those that are spaced more closely than 10 cM add little or no information for conventional backcross and intercross populations (Lander and Botstein 1989; Rebai, Goffinet et al. 1995). The markers chosen for the recombinant inbred data sets supported by the NTB will be optimized for QTL mapping, with appropriate marker density.

Advanced intercross populations, on the other hand, benefit from having more closely spaced markers (Darvasi and Soller 1995). The choice of markers will be especially important for the large G_{10} advanced intercross described elsewhere in this application. We have created a complete database of all MIT CA-repeat microsatellite loci, which can be downloaded at < www.nervenet.org/main/dictionary.html>. Of a total of 6310 microsatellites in the Whitehead Institute database and our own FileMaker Pro version, 2957 are polymorphic between C57BL/6J and DBA/2J. Of these, 1269 have polymorphisms in the range of 8 to 30 bp. Given the very large number of polymorphic loci, it has been easy for us to select a subset of 350 evenly spaced markers that will cover >95% of the mouse genome, the Y chromosome excepted.

Data and map export

The NTB is intended to supplement rather than replace existing QTL mapping programs. In particular, it will make some existing programs more accessible to neuroscientists. Some existing programs, such as Mapmaker/QTL (Lincoln, Daly et al. 1992) and QTL Cartographer (Basten, Weir et al. 1994), provide sophisticated methods, including maximum-likelihood evaluation and automated model-building. However, they require the user to create specially formatted data files and marker maps. These requirements can be an insurmountable barrier for the nonspecialist. To enable neuroscientists to use these programs, the NTB will allow export of marker data, trait data, and marker maps in the formats required by these programs.

We will make public the genotype data generated by our Genotyping and Mouse Colony Core as soon as possible, in formats suitable for Map Manager QT and QTX, Mapmaker/QTL, and QTL Cartographer. Even before the proposed Web interface is in place, we will provide Map Manager QT data files for downloading from the <nervenet.org> site. Map Manager data files have been available for downloading from this site since early in 1994. Data sets include complete RI genotype files, complete MIT SSLP map files, and several other major mapping panels. A complete list of Map Manager files currently available online is at <www.nervenet.org/MMfiles/MMlist.html>.

Dominance models

Dominance refers to the fact that the trait value of a heterozygote may not be the mean of the values of the two homozygotes. This phenomenon can be analyzed in intercrosses, in which separate contributions from additive and dominance effects can be estimated. These effects are expressed in the trait equation as coefficients in two terms, one in which the three genotypes (maternal, heterozygous, and paternal) are coded as –1, 0, and 1, and one in which the genotypes are coded as 0, 1, and 0. For interval mapping, the expected genotypes are typically nonintegral and are derived from the genotypes and distances of flanking markers.

For the free dominance model, the two coefficients are allowed to assume any value. For a constrained additive model, the dominance coefficient is set to 0; for constrained dominant (or recessive) models, the dominance coefficient is set to the value of (or the negative of) the additive coefficient. The free model uses two degrees of freedom, the others only one. The NTB will allow a choice of these models for the target QTL in simple or composite interval mapping. For the background loci in composite interval mapping, it will allow any of the constrained models, that is, those that use one degree of freedom.

Interference models

Genetic interference refers to the observation that, at least in some organisms, two recombination events close to each other are much less common than would be expected by chance. This fact affects the estimation of a QTL genotype from the genotypes of flanking markers. No completely satisfactory model of interference has been described, and even if one were available, its implementation would probably be computationally demanding. However, we can easily implement two extreme models, one in which interference is absent and one in which it is complete (multiple recombinants are absent, at least within an interval).

In recombinant inbred lines and advanced intercrosses, recombination events are derived from multiple rounds of mating, and since adjacent recombinations may be derived from different meioses, interference is assumed to be absent. In addition, the probability of recombination is greater than the nominal map distance between markers (Jiang and Zeng 1997). QTL mapping in recombinant inbred lines and in the G_{10} advanced intercross, therefore, will be analyzed by a no-interference model that accounts for the map expansion from multiple rounds of mating.

Fast interval mapping

Whittaker and coworkers (1996) have derived a relationship for estimating the position and effect of an additive QTL directly from the regression coefficients derived from regression of the trait on the two flanking marker genotypes. Background markers can be included in this analysis; that is, this method can be analogous to either SIM or CIM. This method avoids regression at multiple analysis points between the flanking markers, and it should be several-fold faster than SIM or CIM as described above. This elegant method, which has been implemented for the first time in Map Manager QTX, will be implemented in the NTB. It will be the standard method for mapping with RI strains, since there is no question of distinguishing dominance effects with these. This method will also be an option for mapping with intercross and advanced intercross populations.

Weighted regression

In some experimental situations, particularly those that involve recombinant inbred lines, an estimate of the measurement error may be available for each trait value. In cases where the measurement error is not uniform across trait values, the estimate of QTL effect may be improved by an inverse-variance weighting, that is, by weighting the contribution of each trait value with the inverse of the variance for that value. The user interface for this option can be quite simple. The variance of the trait will be entered as another trait, and an option will allow one trait to be designated as the weight for another.

Missing marker data

Missing data constitute an important practical problem in genetic mapping, including QTL mapping. If a trait value is missing for an individual, that individual must be omitted from the regression. If marker data are missing or ambiguous, however, it is usually possible to calculate an expected value based on the genotypes of flanking markers (Martinez and Curnow 1994; Jiang and Zeng 1997). The NTB will use the method of Jiang and Zeng. For sets of marker data curated by the NTB (recombinant inbred strains and shared intercrosses), the expected values for missing data (which should be rare) can be calculated once and stored with the marker data.

Epistasis testing and search

There is growing awareness of the importance of epistasis (interactions between nonlinked loci) in complex traits and a plea for software which will detect and analyze such effects (Frankel and Schork 1997). Cheverud and Rotman (1995) recently published a method for analysis of epistatic effects, but it has not yet been implemented in mapping software. The NTB will implement that method in two ways. The first will produce a simple report on epistatic effects between any two marker loci with respect to a chosen trait. The second will allow a search for marker loci that show significant epistatic effects with a given locus. These methods will not be interval methods; they will be analogous to the single-locus QTL mapping method described above, in which single marker loci are taken as indicators for nearby quantitative trait loci. In the case of epistasis, however, two marker loci are involved, each an indicator for a different QTL. Normally, single-locus QTL mapping will detect epistatic effects between QTLs only if there is also a significant additive or dominance effect. The Cheverud and Rotman method, in contrast, will allow the detection of pairs of loci whose effect on a trait is purely epistatic, that is, pairs of loci for which there is no significant additive or dominance effect.

Empirical significance thresholds

As mentioned above, one of the critical problems in QTL mapping is the establishment of appropriate significance thresholds. The NTB will implement more than one method for establishing significance thresholds, including a new one. There are several aspects to this problem, the most important of which is that mapping involves testing multiple hypotheses for one data set. The importance of sufficiently stringent significance thresholds has been amply demonstrated, and *a priori* thresholds have been established (Lander and Kruglyak 1995). More important, methods have been described to calculate empirical thresholds, tailored to the idiosyncrasies of the data set (Piepho, personal communication; Churchill and Doerge 1994; Doerge and Churchill 1996). Three of these are permutation methods—one to establish a threshold for SIM and two to do the same for CIM. The fourth is an approximate method that promises to be much faster than permutation. Map Manager QT currently implements the first method; the NTB will implement all of them.

Nonparametric statistics

For traits expressed on an ordinal scale or whose distribution is far from normal, the NTB will provide the option of evaluation with a generalized Wilcoxon rank-sum test (Kruglyak and Lander 1995). The behavior of this statistic has not been described for the case of composite interval mapping, but in any case users should rely on the permutation tests provided by the NTB to calculate significance thresholds.

Other covariates

The NTB will provide an option by which a trait can be designated as a covariate and included in the analysis as a nongenetic determinant of the trait being analyzed. This option may be useful when environmental or nongenetic conditions (such as age) are known to affect the trait.

Testing

The mapping functions of the NTB will be tested with simulated data generated by a set of functions written for the Mathematica mathematics software. These functions can generate marker data sets with markers at prescribed intervals; in addition, they can generate traits based on any number of quantitative trait loci, using any of the standard dominance models and using any statistical distribution for the environmental effect. Other routines can perform simple and composite interval mapping, generating the same type of figures as will be generated by the NTB. A person not part of the programming team will use these Mathematica functions to generate data sets and compare a Mathematica analysis and an analysis by Map Manager QTX with an analysis by the NTB. The budget for Project 4 includes, in Years 02 to 05, a stipend for a student intern to perform this work.

User manual and tutorials

The user manual will be produced and made available online as hypertext and PDF files. The design will resemble the user manual for Map Manager QT (see appendix). However, the online manual for NTB will include interactive tutorials with real data sets. This material will be modeled on the superb ActiveStats statistics tutorial (DataDescription Inc). Feedback from users of Map Manager QT since its release include many questions on the rationale and strategy of QTL mapping rather than the operation of the software. Our online documentation for the NTB will need to teach the critical statistical concepts needed for successful QTL mapping:

- Differences between regression methods and maximum-likelihood methods and other software that makes maximum-likelihood methods available
- Interpretation of the Haley and Knott likelihood ratio statistic and its relationship to the χ^2 statistic and to LRS and LOD scores
- Multiple hypothesis testing and its effect on the probability of inference errors
- Alternative dominance models, alternative epistasis models, choice of background loci, and mapping itself as examples of multiple hypothesis testing
- Advantages and limitations of the nonparametric (rank-sum) method
- Different strategies of composite interval mapping for linked and unlinked QTLs
- Limitations on the ability of any QTL mapping method to resolve closely linked QTLs unless the effect of the QTLs is in the opposite direction
- The importance of empirical significance thresholds established by permutation tests and the rationale for different permutation tests

Core Descriptions

Core Descriptions

Introduction

The Research Core units will serve different projects of the Program. These units will be a resource available for grant-supported research projects and protocols related to the overall goals of the participating laboratories. A Core Committee made up of Drs. Williams and Rosen (Core Directors) will consult weekly to review Core progress and procedures. The reviews will address experimental design, data management, and budgetary effects on the Core units. The use of the Core budgetary resources will be at the discretion of the Core director. Any redistribution of funds will require the approval of the Program director. The use of the Cores is summarized in Table 1.

The Neurohistology Core (Core A) will process mouse brains for inclusion in the Mouse Brain Library (MBL). The Genotyping and Mouse Colony Core will provide genotypes of G_{10} animals that will be added to the MBL. In addition, all brains to be processed by Core A will be perfused in this Core.

Table 1

QUANTITATIVE USE (%) OF CORE UNITS BY COMPONENT RESEARCH PROJECT (YEAR 1)

		CORE UNIT	
INVESTIGATOR	PROJECT	NEUROHISTOLOGY	GENOTYPING AND MOUSE COLONY
ROSEN	1	40%	40%
WILLIAMS	2	40%	0%
NISSANOV	3	20%	0%
MANLY	4	0%	60%
		100%	100%

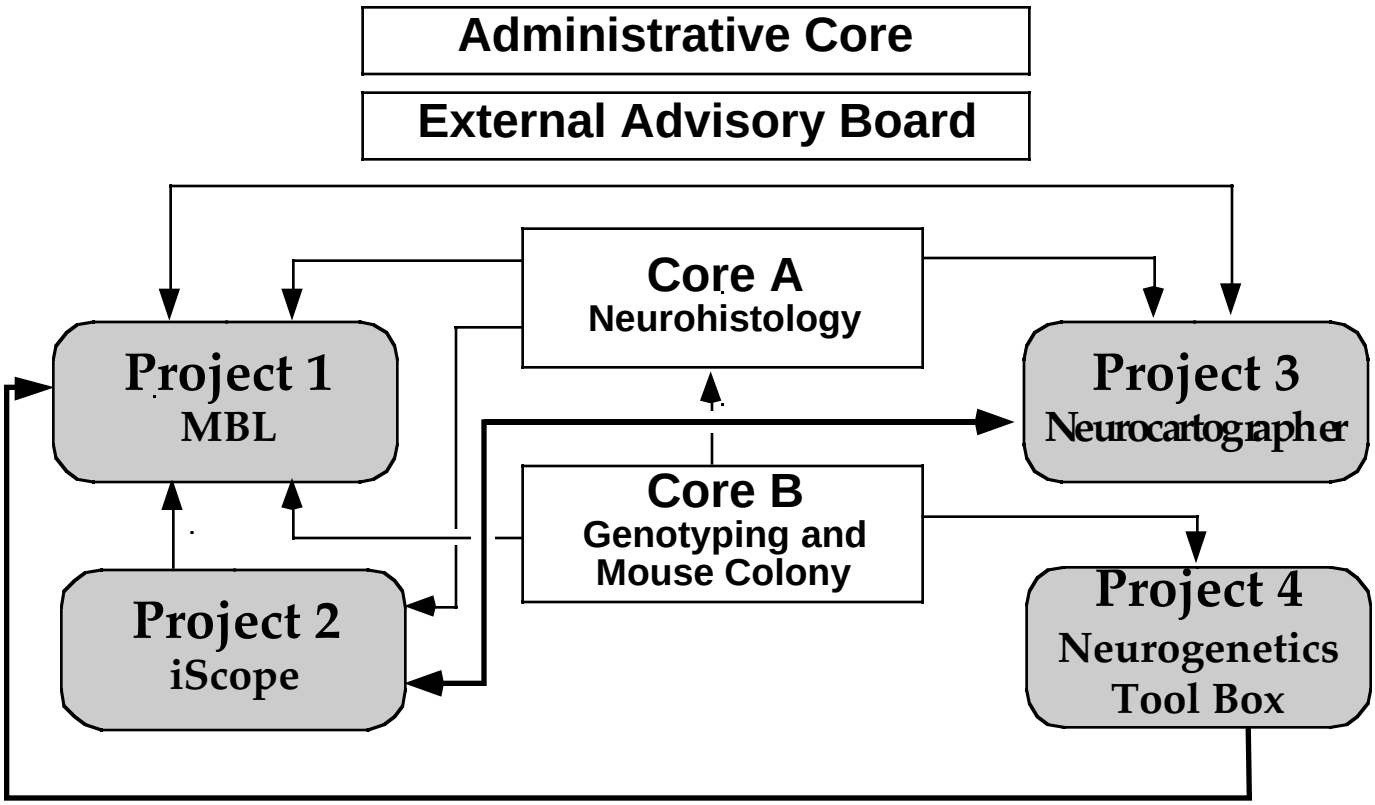
Administrative Core

Robert W. Williams, Program Director (5% effort)
University of Tennessee, Memphis, TN 38136

1. Objective

The Administrative Core will maintain all Core and Project budgets, provide renewal applications, oversee the travel budget and approve travel arrangements as they relate to the program project, and arrange and chair Executive Committee and Core Committee meetings. In addition, the Administrative Core will manage Advisory Board interactions. As chair of the Executive Committee, the Program Director will receive and distribute applications for new projects and will be responsible for new uses of the Core staff and facilities. The Director will review all requests to redistribute funds in the Program budgets and in the budgets of grant or protocol applications that include the use of Program Core units.

2. Organizational Chart



The Administrative Core will oversee all aspects of the programs and cores as detailed above. The interactions among the projects are illustrated above. An important part of the program project is the External Advisory Board. This Board will meet as a group each spring in Memphis (funds have been requested to cover travel expenses and a small honorarium). The External Advisory Board's main purpose is to guide and critique plans and our implementation. A secondary goal is to provide us with their insight into new methods and problems that the project as a whole should consider implementing. A senior member of the Human Brain Project will also be invited to participate in the annual review of the PPG.

Members of the board were selected by the PIs of the projects on the basis of their stature in major fields of bioinformatics and neuroscience. At present the board consists of

1. **Adrianne Noe,** Director of the National Museum of Health and Medicine in Washington. Dr. Noe is in charge of several major research collections, including the Yakovlev-Hakem Human Brain Collection. She is currently collaborating with Dr. Wally Welker on developing and distributing the Wisconsin Brain Collection, now known as the Brain Biodiversity Project. Dr. Noe brings a great

deal of expertise to our PPG pertaining to the organization and long-term maintenance of collections of slides and digital images.

2. **Arthur Toga**, Associate Director of the Neuroimaging Center at UCLA. Dr. Toga is one of the world's leading authorities on brain mapping and imaging. He will be advising us particularly regarding software systems and algorithms used to reconstruct and segment images as part of the NeuroCartographer project.
3. **John Belknap**, Professor of Neuroscience, Oregon Health Science Center. Dr. Belknap is one of the foremost experts on complex trait analysis as it pertains to the study CNS structure and function in mice. He is a strong statistician and behavioral neuropharmacologist who has done a great deal to advance the use of recombinant inbred strains in neuroscience. Dr. Belknap brings to this PPG a particular interest in behavioral variation among mice that will provide an excellent counterpoint and complement to our own focus on anatomical variation.
4. **Joel Richardson**, Senior Programmer, co-PI of the Mouse Genome Database, Jackson Laboratory, Bar Harbor, Maine. Joel Richardson is the senior database programmer for the Jackson Laboratory's bioinformatics program. He wrote and designed much of the Mouse Genome Database web interface. He has a strong background is large relational and object databases such as Informix and Oracle. He will provide us with criticism and advice as to how best to implement and support our informatics superstructure.
5. Unfilled. We have intentionally left one position free to be filled early in the first year. When this application is reviewed we will reassess what areas would most benefit from an external advisor: robotics, video production, web design, distributed database structure, image warehousing, cell-level image segmentation.

3. Staffing

Director (Robert W. Williams, 5% effort). The director's functions are as described above.

A half-time Administrative Assistant will be responsible for the following duties: (1) assist in scheduling meetings and progress reports, (2) help with database entry and backups, (3) help to prepare/convert documents to html for uploading onto the web, (4) maintain a listing of databases and publications, (5) order supplies, and (6) arrange for all travel specifically related to the PPG; for example Human Brain Project annual meetings, meetings with External Advisory Board members, etc.

4. Resources and Environment

The Administrative Core will be housed in the Department of Anatomy and Neurobiology at the University of Tennessee, Memphis. After renovations of the Neuroscience complex, the Administrative Office will move to the fifth floor of the Wittenborg Building (see Letter from S.T. Kitai, Chair).

As noted in the Introduction, the University of Tennessee, Memphis, has demonstrated a strong commitment to this project, demonstrated by the commitment of funds, space, and equipment.

5. Services provided

The services provided are listed in the Objectives.

Research Unit Cores

Core A (Neurohistology Core)

Glenn D. Rosen, Core Director (10% effort)
Beth Israel Deaconess Medical Center, Boston MA 02215

Core A (Neurohistology Core)

1. Objectives

The specific objective of the Neurohistology Core is to provide high-quality histology for several aspects of the research comprising the Program Project. Specifically, this work includes the basic histologic processing of brains to be part of the Mouse Brain Library (MBL, Project 1), which will be used by Project 2 (iScope) and Project 3 (Neurocartographer).

The Neurohistology Core will be responsible for ensuring proper collection, registration, and processing of all neuroanatomic materials generated by the component projects.

In consultation with other investigators of the Neuroinformatics Program Project, the Neurohistology Core will be involved in planning, application for funding, and implementing additional research utilizing neuroanatomic and pathologic analysis that can be attached to the current and future research activities of the Program Project.

2. Staffing

The Neurohistology Core will be staffed by one full-time professional and one full-time and one half-time histology technician.

A. Professional Personnel (see attached biographical sketch)

Dr. Glenn D. Rosen, Ph.D., Principal Investigator of the Neurohistology Core (10% effort), will supervise all activities of the Core.

B. Support Personnel

Stefany Palmieri, Head Technician, will histologically process all brains for the Neurohistology Core. Ms. Palmieri has been involved in processing the 600 brains currently in the MBL. She is highly skilled in all aspects of processing brains for celloidin embedding. In addition, we will hire a technician who will split his/her time between the Neurohistology Core and Project 1. This person will aid Ms. Palmieri in processing the brains and will be responsible for imaging these brains for uploading to the MBL.

3. Resources and Environment

The Neurohistology Core will be housed at the Beth Israel Deaconess Medical Center. This laboratory is located on the seventh floor of the Charles A. Dana Research Building of the Hospital.

Laboratory Facilities

The laboratory contains approximately 2100 square feet, 1500 of which are wetlab space. Half of this space will be devoted to Neurohistology Core activities. The space includes a fully equipped histology area housing a giant microtome capable of processing wholebrain human specimens, sliding microtomes, a rotary microtome, a cryostat, an oven, a refrigerator, a rocker table, a staining microscope and accessories, scales, meters, and glassware. There is a microscopy room containing a teaching comparison compound microscope, two stereomicroscopes, a compound light microscope, and equipment for macro- and microphotography. Additional facilities include cabinets for storage of serial histologic sections and a fully equipped darkroom for photographic and autoradiographic processing.

We have requested and have received commitments from Beth Israel Deaconess Medical Center for an additional 500 sq. ft. of research space allocated to us (see attached letter).

4. Administration

The technicians will report to Dr. Rosen. The staff of the Neurohistology Core will participate fully in the administrative and research meetings organized by the Administrative Core and specified elsewhere on this application. In addition, the Core Director will review materials for quality on a daily basis.

Dr. Rosen will be notified personally during a research meeting, by telephone, or by email that specimens will be shipped at a particular time and date. Dr. Rosen will alert the technical staff in the laboratory and the receiving department of the hospital. In some cases, specimens will be obtained directly by a member of the

lab at the site of origin. The specimens will be received by the head technician and each brain will be assigned a unique identifying number.

Funds for the Neurohistology Core will be managed by the Office of Research Administration of the Beth Israel Deaconess Medical Center as part of theNeuroinformatics Program Project management. Additional organizations involved in the grant administration include the Department of Neurology, Beth Israel Deaconess Medical Center.

Interaction among the Administrative Core and other Program Project cores, the different research components, and the Neurohistology Core will be facilitated by an electronic mail network as well as facsimile facilities.

5. Justification

The Neurohistology Core is necessary because a large number cases will be added to the MBL. Because tissue processing is repetitive and is not in itself hypothesis driven, it is best carried out in a core facility. A centralized Neurohistology Core will facilitate the acquisition of material for the MBL, thus ensuring that the potential of the Neuroinformatics Program Project research can be realized. The Core will process and analyze neuroanatomic materials from all component projects, obviating the need to duplicate expensive equipment and train more staff in time-consuming, difficult techniques (described below).

6. Use

The Neurohistology Core is equipped to process material of varying size, from the human brain to the brain of a mouse embryo. Personnel in this facility have carried out studies on the brains of apes, dolphins, old world monkeys, rats, and mice, and members of the staff have experience with a large number of modern neuroanatomic methods applicable to the mammalian brain. See below for details on a sampling of available techniques and protocols.

Procedures

Gross anatomic inspection of all animal tissue is always carried out, and photographs are taken when appropriate. Routine histologic procedures and imaging are the main functions of the Neurohistology Core.

A. Routine Histologic Procedures

In preparation for routine examination, animals are perfused transcardially by the Genotyping and Mouse Colony Core (Core B), then shipped to the Neurohistology Core. Once the brains arrive in the laboratory, they are given a unique identifying number and are randomly assigned to be cut in either the horizontal or coronal plane. After post-fixation for at least one week in 10% formalin, the brains are dehydrated in a series of 80%, 95%, 100% ethanol and ethanol/ether. The brains are then placed into 3% celloidin for one week followed by 12% celloidin for 2–3 days or until hard.

The celloidin block is trimmed to achieve a stable base and is notched on the left side for side orientation. The specimen is placed on a sliding microtome. The sections are cut coronally at 30 µm and are segregated and saved in 80% ethanol. Every fifth section is stained for Nissl substance. Spare sections are preserved indefinitely in 80% ethanol.

For the 3D atlas described in Project 3, we will cut the tissue at 10 µm.

Nissl Staining

Sections are washed in distilled water and placed in 1% cresyl violet acetate solution (which stains the Nissl substance) for 3–5 min. Each section is placed in distilled water for 1 minute and then differentiated and dehydrated in 70, 80, and 95% ethanol. A few drops of colophonium are added to the 95% ethanol baths. If differentiation is adequate, the sections are then cleared with terpineol and passed through xylenes. Sections are mounted with careful attention to orientation so that left and right are consistently identifiable. Coronal sections are mounted with the nick in the celloidin on the right side of the slide. Horizontal sections are mounted with the nick on the left side of the slide. The sections are then mounted in Permount. One series of every fifth section will be sent to Project 1 (MBL) and subsequently to Project 2.

Other Stains

The MBL currently is comprised solely of Nissl-stained sections. The laboratory is prepared, however, to stain adjacent sections of celloidin-embedded material using the following methods.

Myelin Staining (Loyez). Free-floating sections are washed in distilled water for 30 s and placed for 6 h in a 2% ferric ammonium sulfate solution. The sections are then washed for 30 s and incubated overnight at room temperature in a 1% hematoxylin solution in 10% ethanol/2% lithium carbonate. The next day the sections are washed twice for 30 s each and differentiated in 2% ferric ammonium sulfate until the gray matter appears. The sections are then washed in three changes of distilled water for 30 s each before differentiation in Weigert's solution (2% sodium borate/2.5% potassium ferricyanide). The sections are then washed three times in distilled water, with the second wash containing a few drops of ammonium hydroxide, before being mounted onto subbed slides, dehydrated, cleared with xylenes, and coverslipped with Permount.

Hematoxylin and Eosin. This routine neuropathological stain identifies nuclei, cytoplasm, and blood vessels. Free-floating slides stored in 80% ethanol are placed in hematoxylin for 5 min, then differentiated in 1% acid alcohol. Each step is followed by a wash in distilled water. Then the slides are placed in distilled water containing a few drops of ammonium hydroxide (bluing agent) followed by 1 min in eosin. Then the slides pass through a series of 70%, 80%, and 95% ethanol and are stored in alpha terpineol until coverslipped.

PTAH. This routine neuropathological reagent stains astrocyte gliosis. Each step in this protocol is followed by a wash in distilled water, except where noted. Free-floating slides stored in 80% ethanol are incubated in saturated mercuric chloride for 3 h and then placed in Lugol's iodine for 5 min. The sections are placed in 95% ethanol, followed by 0.25% potassium permanganate for 5 min and 5% oxalic acid for 5 min. The sections are washed 5 times in distilled water and placed in PTAH for 24 h. The sections are washed in 100% ethanol, placed in xylenes, and coverslipped.

Masson's Trichrome. This stain identifies nuclei, collagen, and blood vessels. Each step in this protocol is followed by a wash in distilled water, except where noted. Free-floating slides stored in 80% ethanol are placed in Weigert's iron hematoxylin for 30-45 s and then in Masson's fuschin OG for 5 min. Sections are then placed in 1% acetic acid for 3 min, 5% phosphotungstic acid for 5 min, 1% acetic acid for 3 min, 2% light green for 5 min, and 1% acetic acid for 1 min. There are no washes between these steps. Sections are mounted and placed in 80% ethanol, 95% ethanol, and xylenes and coverslipped.

Gallyas Protocol. Because this method yields optimal contrast for image analysis, it can be used on sections from which cell counts are made. The celloidin sections are cleared of celloidin and mounted onto glass slides. The slides are then placed into 4% formic acid for 3 h and then overnight in a solution of 10% formic acid and 9% hydrogen peroxide. The next day, the slides are washed three times for 15 min each and then placed for 15–20 min into the developer (2.5% sodium carbonate, 0.1% silver nitrate, 0.1% ammonium nitrate, 0.5% silicotungstenic acid, and 0.36% formalin). The sections are then placed in 0.5% acetic acid for 5 min, washed in distilled water for an additional 5 min, and fixed in Extaflo fixer (0.29% for 10 minutes). After another 5-min wash in distilled water, the sections are dehydrated, cleared, and coverslipped.

Potential Pitfalls and Problems with Celloidin Processing**.** A minor technical problem concerns tissue shrinkage during processing. For many CNS traits it is of interest to compute in vivo values. This can be difficult when the data are obtained from celloidin material where tissue shrinkage is severe. Our solution to this is to compute the degree of shrinkage for individual brains. Brain volume after processing is computed by using a uniform point counting protocol. This value is then compared to the brain weight obtained immediately after dissection. We have used this method in several cases and volumetric shrinkage is approximately 50–60%. From these volumetric shrinkage estimates we can compute approximate linear shrinkage. We have also directly estimated shrinkage by comparing MRI estimates of brain volume with those of the same brain after processing, and have obtained comparable estimates of shrinkage.

B. Immunohistochemistry

As the MBL expands, we will begin to include immunohistochemically stained tissue. A large number of immunohistochemical stains have been implemented in the Neurohistology Core laboratory. New antibodies are continually being used in response to the needs of particular research questions. For the current proposal, we plan only to stain for ChAT- and Parvalbumin-immunoreactive neurons and fibers. In the future, we may expand the MBL even further, and we therefore include protocols for some of the other candidates. Antibodies have been obtained from commercial sources and from laboratories of colleagues

For immunohistochemistry, mouse pups and adults are perfused transcardially under deep anesthesia with 0.9% saline followed by 4% paraformaldehyde. The brains are removed from their skulls and post-fixed for 24 h in 4% paraformaldehyde, then placed in a 0.1 M phosphate buffer/10% sucrose solution until the brains sink (usually 1 day). Next, the brains are placed into a 0.1M phosphate buffer/30% solution until they sink,

usually 2–3 days later. The brains are nicked on the ventral surface of either the right or left hemisphere,then cut frozen on a sliding microtome in the coronal plane at 30 µm. Consecutive sections are stored in 0.1M phosphate buffer.

Monoclonal Antibodies

Neurofilament (NF). Free-floating sections are rinsed twice in phosphate-buffered saline (PBS; pH 7.4) for 5 min each and transferred to a buffered 0.6% hydrogen peroxide solution in order to block staining of endogenous peroxidases. The sections are rinsed twice in PBS and incubated overnight at 4°C in a 1/50 dilution of mouse anti-neurofilament immunoglobulin (monoclonal antibody to the 68 kDa subunit of neurofilament from Boehringer Mannheim, Indianapolis MN). The vehicle (diluent) for all antibody incubations is 3% rabbit serum in PBS.

Sections are then placed into a solution containing the linking antibody (rabbit anti-mouse immunoglobulin, Dakopatts, Santa Barbara, CA, Z259; diluted 1/20) at room temperature for 2 h. The sections are rinsed twice with PBS and placed in a 1/250 dilution of mouse peroxidase anti-peroxidase (Dakopatts B650) at room temperature for 2 h. The tissue is rinsed twice in PBS and then twice in 50 mM Tris buffer (pH 7.6) and developed using 0.05% diaminobenzidine and 0.005% hydrogen peroxide diluted in Tris. After rinsing with Tris, sections are mounted on chrome-alum coated slides, dehydrated, counterstained with Methyl Green/Alcian Blue, and coverslipped with Permount.

The following monoclonal antibodies are suitable for use in mouse and are also available to Core A. Note again that to maintain high-throughput we initially only intend to stain for two of the following short list of cell types (ChAT and parvalbumin).

1. Radial glial fibers (Rat-401). This monoclonal antibody directed against radial glial fibers is provided by S. Hockfield, Yale University. The protocol is the same as that for NF with the exception of the dilution (1/4) of the primary antibody.
2. Vimentin. This monoclonal antibody (Boehringer Mannheim) stains radial glial fibers with less background than Rat-401. The protocol is the same as that for NF with the exception of the dilution (1/500) of the primary antibody.
3. Glutamate. This antibody (Incstar) stains glutamatergic fibers with a few cell bodies. The procedure for glutamate staining is identical to that for NF with the exception of the dilution (1/5000) of the primary antibody.
4. Parvalbumin. This antibody (ICN) stains neuronal cell bodies in the cortex (many of which are GABA positive). The staining procedure is identical to that for NF with the exception of the dilution (1/500).
5. Choline acetyltransferase (ChAT). This antibody (Chemicon) stains fibers and cells of the ascending cholinergic system of the forebrain. The procedure is the same as that for NF with the exception of the dilution (1/1000).
6. Tyrosine hydroxylase (TH)**.** Immunohistochemistry proceeds identically to parvalbumin with the exception of the primary antibody (Chemicon, 1/500).

Polyclonal antibodies

Vasoactive Intestinal Peptide (VIP). Free-floating sections are rinsed twice in phosphate-buffered saline (PBS) at pH 7.4 for 5 min each and transferred to a buffered 0.6% hydrogen peroxide solution in order to block staining of endogenous peroxidases. Prior to the first antibody incubation, sections are placed in vehicle only for 20 min at room temperature. The vehicle (diluent) for all antibody incubations is 5% goat serum in PBS. Sections are then placed into a 1/2000 solution of primary antibody (Incstar) overnight at 4°C. The next day, sections are transferred into a biotinylated goat anti-rabbit immunoglobulin solution (Vector Laboratories) diluted 1/60 for 2 h at room temperature. After two washes in PBS, the sections are placed into ABC complex (Vector Laboratories) for 2 h at room temperature. The tissue is rinsed twice in PBS and twice in 50 mM Tris buffer (pH 7.6) and developed, dehydrated, counterstained, and coverslipped as with NF (see above).

GABA. The procedure for GABA staining is identical to that for VIP with the exception of the dilution of the primary antibody (Incstar), which is 1/333.

Somatostatin. The procedure for somatostatin (Chemicon) is identical to that for VIP. The dilution is 1/100.

Glial Fibrillary Acidic Protein. The procedure for GFAP staining is identical to that for VIP with the exception of the dilution of the primary antibody (Incstar), which is 1/25.

Potential Pitfalls and Problems with Immunohistochemistry

Although antibody penetration is a potentially serious issue with immunohistochemistry, we have extensive experience with immunohistochemical procedures, and the antibodies we use have proved remarkably robust and reliable. We will pay careful attention to the quality of the tissue, and if necessary we will explore various methods to enhance antibody penetration, including the use of various detergents.

7. Financial Considerations

A. Additional Technical Support Needed

One full-time technician and one half-time technician will carry out activities related to the Neurohistology Core of the Neuroinformatics Program Project.

B. Additional Equipment Needed

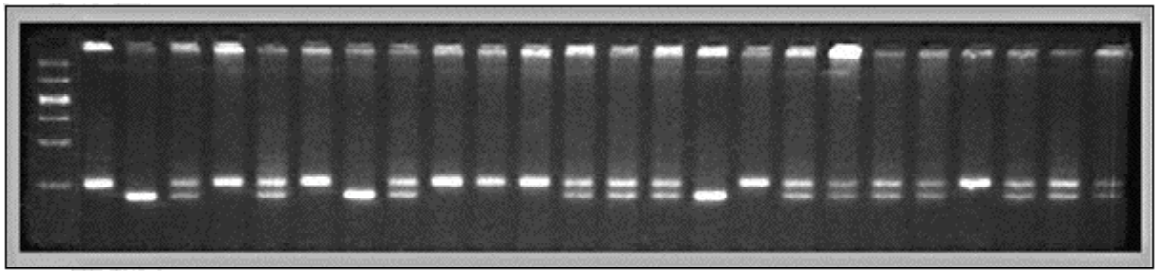
An additional microtome is needed for processing the large volume of tissue being generated by the Neurohistology Core.

C. Additional Costs Relating to Travel

Travel costs will be handled through the individual research projects.

Core B (Genotyping and Mouse Colony Core)

Robert W. Williams, Core Director
University of Tennessee, Memphis, TN 38163



Core B (Genotyping and Mouse Colony Core)

1. Objectives

The GMCC will provide essential resources that will enable the research community to greatly improve the speed and efficiency of mapping genes that modulate the size and cell composition of the mouse brain. We will generate close to 400,000 genotypes for a set of approximately 1400 mice (1200 animals between the ages of 46 and 76 days, and ~200 animals more than 2 years old) during this grant period. These genotypes will be a vital component of the Neurogenetics Tool Box. The brains of mice that have been genotyped by this core will be processed by the Neurohistology Core, and high-resolution images of all the cases will be added to the Mouse Brain Library (MBL). Genotype databases will be consolidated with the quantitative data generated as part of the MBL and the NeuroCartographer Project (Project 3). When this work is complete, the Neurogenetics Tool Box will contain a huge body of phenotypic and genotypic data for a population of over 1400 adult animals from a unique type of genetic cross. These data will make it possible to routinely map QTLs with a precision of <4 cM.

The GMCC will also be responsible for oversight of our mouse colony and for providing fixed brains to the Neurohistology Core. The maintenance of the mouse colony is not otherwise discussed in this Core description because these duties will be covered primarily by the UT Memphis Department of Comparative Medicine. Intramural funds provided by the Department of Anatomy and Neurobiology will cover minor expenses associated with perfusion and dissection of brains. We routinely process several thousand cases per year as part of other funded projects that deal with the structure of the mouse eye and retina. Additional supplies or technical assistance will not be required for the duration of this grant period.

Current status. A large number of animals have already been generated for the principal cross that will be used in this program project. More than 1100 brains have been dissected and are now being prepared for embedding in celloidin. DNAs from 1152 of these animals have been extracted and arrayed in a set of 12 96-well microtiter plates. All these cases have been processed in the same way as the current set of 600 brains in the MBL. We also have breeding colonies for numerous standard and recombinant inbred strains and will therefore be able to supply the Neurohistology Core with the material required to have at least 12 animals per strain in the MBL.

A high-resolution mapping panel. We have generated what is called a tenth-generation advanced intercross. The cross started with two common but highly divergent inbred strains of mice: C57BL/6J (B for its *b*lack coat) and DBA/2J (D for its *d*ilute brown coat). These are the same two parental strains used to produce the 36 BXD recombinant inbred strains that are already a major gene mapping resource in the MBL. These two strains differ genetically at more than 3000 genes and marker loci, greatly facilitating high-resolution QTL mapping. Of equal importance, the CNS architecture of these strains differs greatly (Williams et al. 1996, 1998, 2000); for example, weights of their brains are 420 mg and 500 mg, respectively. We have already made a high-resolution atlas of the brain of C57BL/6J (<nervenet.org/mbl/mbl.html) and expect to do the same for DBA/2J. As part of the NeuroCartographer project, we will generate consensus 3D models of the brains of these two parental strains.

The tenth-generation BD intercross we are using has many advantages over a conventional intercross or backcross. With a conventional intercross or backcross, QTLs can generally be mapped only with a precision of ±10 cM, even when 2000 or more cases are studied. To overcome this problem, Darvasi (1998) suggested the advanced intercross as a conceptually simple method to increase the precision of QTL mapping to ±2 cM.

An advanced intercross accumulates recombination events over many generations (see below), effectively stretching the genetic map. The genetic map doubles in length with each doubling of generation number. For example, a G_4 cross has twice the map length of an F_2 cross. The genetic map of G_{10} animals is five times as long as the map based on a conventional cross—7000 cM rather than 1400 cM. This means that for a given level of effort, the G_{10} panel provides five times the precision of a F_2 intercross. Thus, positional candidate gene cloning approaches become far more feasible. Dr. K. Manly has recoded the Map Manger QT program to create Map Manager QTX, which will accommodate data sets from advanced intercrosses (Manly and Olson 1999). Project 4, the Neurogenetics Tool Box, will incorporate both genotypes and quantitative neuroanatomical traits from the entire G_{10} panel.

Advantages for Neuroinformatics. Perhaps the most important advantage of the advanced intercross for neuroscientists is that large numbers of different morphometric traits will be studied in a single set of animals. This means that we will be able to explore the complex network of gene interactions that are undoubtedly important in CNS development and aging. To make this idea more concrete, consider the possibility that a QTL specifically controls variation in the proliferation of granule cells in three very different parts of the CNS—the olfactory bulb, the dentate gyrus, and the cerebellum. A single researcher might have the time and energy to study one of these populations, but not all three. The link among all three regions would remain unsuspected and undetected. But by providing a large community resource, the Genotyping and Mouse Colony Core, in concert with the Neurogenetics Tool Box, makes all information cumulative in a single experimental cross. If three investigators in succession, studying three different CNS regions, map what might be called *granule cell number modulate genes* to the same chromosomal interval, then the genetic link to a particular cell type across regions will become obvious with a look at the list of mapped quantitative trait loci. Thus, a community resource can expose genetic correlations (pleiotropy) among connected and unconnected parts of the brain.

Statistical Power. Virtually all of the CNS traits that we anticipate mapping are complex multigenic traits. The question arises as to what statistical power the G_{10} progeny set will have to detect QTLs. How small a QTL effect will this panel be able to detect? Darvasi (1998) provides an equation for the approximate number of progeny for a power of 0.5 using an advanced intercross at $75/(2d^2 + h^2)$ where d is the standardized additive allele effect size and h is the standardized dominance effect size. For a locus with a purely additive effect of 0.20 standard deviation, 934 animals are needed for a power of 0.5. Thus with our G_{10} sample we anticipate achieving a power of >0.5 for QTLs that have a standardized effect of only 0.2 SD. In other words, this is a remarkably sensitive cross for detecting and mapping QTLs.

2. Staffing and Oversight

The GMCC will be staffed by one full-time professional. Dr. Jing Gu will be responsible for the extensive genotyping. Dr. Gu currently runs the PI's genotyping lab. Dr. Gu will also be responsible for error checking and for entry of genotype data into our consolidated relational databases. She will be supported by the PPG and will be responsible for genotyping the G_{10} advanced intercross. She is highly skilled in all aspects of genotyping. The Molecular Resources Center will provide additional personnel to assist with the work.

3. Resources and Environment

The Genotyping and Mouse Colony Core will be housed at the University of Tennessee, Memphis. This Core will initially share facilities in the PI's lab (311 Wittenborg). However, late in Year 01 we will open a new UT Memphis Genotyping Center as a division of the UT Memphis Molecular Resources Center. The Genotyping Center will be directed by Dr. Robert Williams. This facility will have an initial capacity of at least 200,000 genotypes a year, of which half will be devoted to this PPG (see letter from Dr. Michael Dockter, Vice Dean for Research).

4. Administration

Dr. Jing Gu will report to Drs. Williams and Manly. Drs. Williams and Gu will meet at least once a week to review progress and will keep closely in touch with Dr. Manly by email and phone. Dr. Gu will maintain several key databases that will be used by Dr. Manly's group as part of Project 4. Dr. Manly will have direct access to the FileMaker Pro server that will host the genotype databases.

Funds for the Genotyping and Mouse Colony Core will be managed by the Department of Anatomy and Neurobiology and by the Administrative Core of the Program Project. Funds to buy equipment to support the

expanded genotyping throughput will be provided by the University of Tennessee. The limited support required to fix animals and ship brains to Dr. Rosen will be covered by intramural funds. The budget for this core does not include any equipment request. Our supplies request is limited to genotyping (~$0.20/genotype) and is relatively modest. Our cost is far lower than current commercial genotyping prices. For example, Research Genetics currently charges in excess of $4 per microsatellite genotype. We will be able to generate 400,000 genotypes for well under $80,000 in supplies.

5. Justification

Full-scale genotyping and extension of the Mouse Brain Library, as proposed in this application, will allow multiple research groups to study many traits in the same set of animals, maximizing the utility of each cross, preventing duplication of effort, and revealing the complex genetic basis of variation in different CNS compartments. This community effort also allows us to substantially increase the positional precision of QTLs that are mapped. The proposed project will greatly reduce the cost and effort of examining the basis of normal variation in CNS architecture among and within strains of mice.

Until now, it has been possible to adequately analyze only a small number of phenotypes in each animal that is genotyped. Each research group generated a unique F_2 intercross or backcross consisting of several hundred animals. Researchers observed a small number of traits, genotyped their modest sample of animals, and subsequently mapped the loci modulating the observed traits. R. Williams and colleagues have used this approach to map more than a dozen QTLs that control morphometric and quantitative variation in the architecture of the eye and brain of normal strains of mice (see *Appendix*; Williams et al. 1998; Zhou and Williams 1999; Williams 2000). Only four trait values were acquired for most mice: body weight, brain weight, eye weight, and retinal ganglion cell number. Even this was a huge undertaking for a single lab. This cottage industry approach is incredibly inefficient and yields only low-resolution estimates of the chromosomal location of QTLs and their phenotypic effects.

Use of the core by individual projects. The Genotyping and Mouse Colony Core is a key adjunct of the Neurogenetics Tool Box, and all of the genotype data will be directly integrated by databases and programs described in Project 4. This core will also provide brains to the Neurohistology Core for processing and entry into the MBL. The genotypes are far more useful when combined with the extensive phenotype data that will be generated in Project 1 (images), Project 2 (cytological assays using the iScope), and Project 3 (quantitative segmentation of brains). The genotype data are a permanent tool that can be used to genetically dissect any quantitative (or qualitative) trait exposed by data sets in the MBL or exposed by the iScope or NeuroCartographer projects.

6. Procedures

Generating the AI cross. Our G_{10} advanced intercross was generated using the method outlined in Darvasi (1998) with modifications suggested by Dr. Lee Silver. We first generated (C57BL/6J x DBA/2J) F_1 mice and intercrossed them to generate a standard B6D2 F_2 generation. We also carried out the reciprocal cross to generate D2B6 F_2 animals. Approximately 100 of each of these reciprocal F_2 types were generated. Thirty mating cages were set up, and one or two F_2 individuals were selected from each litter and mated to generate the G_3 progeny. Siblings were never placed together in a mating cage; since matings were nonfilial, we use a *G* prefix rather than the *F* prefix for the third generation. The G_3 offspring from different cages were mated to generate G_4. This procedure was continued until G_{10} with a constant attempt to minimize potential fixation/inbreeding. Both sexes have been collected, and the range of ages of the current set is quite narrow (between 46 and 76 days). Two hundred animals have been set aside as part of an aging colony. They will be sacrificed at about 700 days of age. The breeding history of the advanced intercross is currently maintained in a FileMaker Pro relational database, and a large part of the database (through August 1998) is available online at <nervenet.org>. The entire database will soon be published online with full genotypes.

Scanning the genome for QTLs. To perform a genome-wide scan of the G_{10} with an average separation between markers of 20 cM, we will employ 350–400 polymorphic microsatellite markers. To choose PCR primers for whole-genome QTL scanning, we select 300–400 evenly spaced marker loci in which the product differences between parental strains are greater than 2 base pairs. Selection is not difficult because these strains have already been typed at many hundreds of microsatellite loci (Dietrich et al. 1994). The G_{10} genetic map is about 7000 cM long, so the average distance between any locus and its neighboring microsatellite loci is 18 cM. Typing 1000 G_{10} progeny with each of 350–400 markers requires

350,000–400,000 PCR reactions, which can be accomplished over 3 to 4 years (mean daily output for one technician is currently 384 reactions). We are investigating procedures that may allow us to double this output.

Genotyping the AI cross. The PCR reaction (10 μl) consists of 40 ng of genomic DNA extracted from spleen or tail using an inexpensive high-salt procedure, 1 μM of each primer, 1x reaction buffer, 200 μM dNTPs, 2.5 mM $MgCl_2$, and 0.5 U of *Taq* polymerase. Reactions are carried out in a thermocycler (M.J. Research) with a hot bonnet using a touch-down procedure in the first 5 cycles with a high annealing temperature to improve specificity. The PCR products are electrophoretically separated on either agarose (TMC MethPhor, 3.5%) or acrylamide gels and visualized with ethidium bromide under UV illumination. Agarose is preferred when alleles differ by more than 8 base pairs, acrylamide when alleles differ by 2–8 bp. We have considerable experience in the use and selection of microsatellites for typing alleles in test crosses and in running these reactions and gels efficiently. Data will be entered into a relational database (FileMaker Pro) and exported to the Neurogenetics Tool Box for statistical analysis and QTL mapping.

Data entry. All genotypes will be entered into a FileMaker Pro database. For simplicity in QTL mapping and multiple regression, markers with the *BB* homozygous genotype will be scored as +1, the *BD* heterozygote as 0, and the *DD* homozygote –1. A failed reaction will be scored as a non-numerical value (*U*). We will not score any markers with dominant molecular banding patterns. Apparent double recombinants which are defined by a single discordant typing will be retested to reduce typing errors. Animal ID number will be the "key" field in relating genotypes with MBL phenotype databases. R. Williams and K. Manly will design all genotype databases. Dr. Manly will use these databases in the Neurogenetics Tool Box project. Dr. Manly will have direct access to genotypes generated at UT Memphis using a web interface to our FileMaker Pro database (see <nervenet.org> for examples of these types of web databases).

Fixation. Many additional mice will be incorporated in the MBL collection over the period of this grant. Dr. Rosen, PI of the Neurohistology Core and the MBL, will order mice from our Administrative Core assistant, who will obtain them from the Jackson Laboratory. Mice will be anesthetized with Avertin (0.5 to 0.8 ml, i.p.) and perfused transcardially with 0.9% phosphate buffered saline followed by approximately 15 ml of 1.25% glutaraldehyde and 1.0% paraformaldehyde in 0.1 M phosphate buffer. An additional 10 to 20 ml of double-strength fixative (2.5% glutaraldehyde and 2.0% paraformaldehyde in the same buffer) is subsequently injected for 1 to 2 minutes at an increased rate and pressure. The head is removed and put in the final fixative until the brain is dissected and weighed. Brains in fixative will be shipped in groups of 50 to 100 to Dr. Rosen. We have followed this procedure without any problems for the past two years.

7. Financial Considerations

Technical support. A full-time senior technician is needed to carry out activities that are part of this Core.

Equipment. The Vice Dean of Research at UT Memphis will provide $140,000 for equipment required to perform as many as 400,000 genotypes over the duration of this grant. The PI will use these funds to assemble a new high-throughput genotyping lab. The appended letter from the Dean lists some of the key equipment that will be purchased.

Travel costs. None

8. Bibliography (Core B)

Darvasi A, Soller M (1995) Advanced intercross lines, an experimental population for fine genetic mapping. Genetics 141:1199–1207.

Darvasi A (1998) Experimental strategies for the genetic dissection of complex traits in animal models. Nat Genet 18:19–24.

Dietrich WF, Miller JC, Steen RG, et al. (1994) A genetic map of the 4006 simple sequence length polymorphisms. Nat Genet 7:220–245Doerge RW, Churchill GA (1996) Permutation tests for multiple loci affecting a quantitative character. Genetics 142:285–291.

Manly, K.F. and Olson, J.M. (1999) Overview of QTL mapping software and introduction to Map Manager QT. Mamm Gen 10: 327-334.

Roff DA (1997) Evolutionary quantitative genetics. Chapman Hall, New York.

Schmitz S., Cherny SS, Fulker DW (1998) Increase in power through multivariate analyses. Behav Genet 28:357–363.

Williams RW (1998) Neuroscience meets quantitative genetics: Using morphometric data to map genes that modulate CNS architecture. In: Morrison J, Hof P (eds) Short course in quantitative neuroanatomy. Society of Neuroscience, Washington DC, pp 66–78.

Williams RW (2000) Mapping genes that modulate mouse brain development: A quantitative genetic approach. In: Goffinet A, Rakic P (eds) Mouse Brain Development. Springer, Berlin (2000) in press.

Zhou G, Williams RW (1999) *Eye1* and *Eye2*: Gene loci that modulate eye size, lens weight, and retinal area in mouse. Invest Ophthalmol Vis Sci 40:817–825.

e. Human Subjects — none

f. Vertebrate Animals

Description of proposed use of animals

i. We will use the species *Mus domesticus* in these studies. Most animals are originally obtained from the Jackson Laboratory, Bar Harbor, Maine. Both sexes will be used. We will have breeding colonies and intercrosses of selected strains. We will have a colony of aged mice to study the genetics of CNS aging.

ii. **Justification of animal use.** Mice will be used to characterize the genetics of CNS. Over the period of this grant, we anticipate that we will use an average of approximately 1000 mice per year. Such large numbers of animals are required in order to map quantitative trait loci with precision. All of these animals will be generated in our mouse colony.

iii. **Veterinary care.** The mice will be maintained in departmental AAALAC approved facilities under the supervision of the Department of Comparative Medicine of the University of Tennessee. Only trained personnel will work with our mouse colony, and they will check room conditions, cage, and animal status daily. The Department of Comparative Medicine has new facilities in the Nash Annex. This is a state-of-the-art vivarium with extensive space for pathogen-free colonies of mice. All of our colony will be kept in this facility. The facility employs two full-time veterinarians to monitor the health and well-being of animals.

iv. **Analgesics and anesthetics.** NA

v. **Euthanasia method**. Adult mice that will be fixed by transcardial perfusion will be anesthetized deeply with Avertin. Adult mice that will be used without perfusion will be sacrificed by cervical dislocation. Neonates and fetal animals are decapitated with scissors. Some mice will be euthanized by exposure to 100% CO_2. These methods are consistent with the recommendations of the Panel on Euthanasia of the American Veterinary Medical Association.

i. Consultants — Dr. Warren Young (Scripps Institute) will serve as a consultant on Projects 1 .

h. Consortium Arrangements — see attached documents and letters of intent

REFERENCES

Airey DC, Strom RC, Williams RW (1998) Genetic architecture of normal variation in cerebellar size. Soc Neurosci Abstr 24:303.

Allen RD, Strömren Allen N, Travis JL (1981) Video-enhanced contrast, differential interference contrast microscopy. Cell Motil 1:291–302.

Amit, Y. (1997) Graphical shape templates for automatic anatomy detection with applications to MRI brain scans. *IEEE Tran Med Imaging*, 16: 28-40.

Arnold SE, Franz BR, Gur RC, Gur RE, Shapiro RM, Moberg PJ, Trojanovski JQ (1995) Smaller neuron size in schizophrenia in hippocampal subfields that mediate cortical-hippocampal interactions. Am J Psychiatry 152(5):738-48.

Atchley WR, Riska B, Kohn LAP, Plummer AA, Rutledge JJ (1984) A quantitative genetic analysis of brain and body size associations, their origin and ontogeny: data from mice. Evolution 38:1165–1179.

Bajcsy R and Kovacic S (1989) Multiresolution elastic matching. CVGIP 46: 1-21.

Bajcsy, R., Lieberson, R., and M. Reivich, (1983) A computerized system for the elastic matching of deformed radiographic images to idealized atlas images, JCAT, 7: 618-625

Baldock RA, Poole I (1993) Video camera calibration for optical densitometry. J Microsc 172:49–54.

Basten, C. J., B. S. Weir, et al. (1994). Zmap--a QTL cartographer. Proceedings of the 5th World Congress on Genetics Applied to Livestock Production: Computing Strategies and Software. C. Smith, J. S. Gavora, B. Benkelet al. Guelph, Ontario, Canada, Organizing Committee, 5th World Congress on Genetics Applied to Livestock Production.

Beach JM, Duling BR (1993) A light-emitting diode light standards for photo- and videomicroscopy. J Microsc 172:41–48.

Begleiter H, Porjesz B, Reich T, Edenberg HJ, Goate A, Blangero J, Almasy L, Foroud T, Van Eerdewegh P, Polich J, Rohrbaugh J, Kuperman S, Bauer LO, O'Connor SJ, Chorlian DB, Li TK, Conneally PM, Hesselbrock V, Rice JP, Schuckit MA, Cloninger R, Nurnberger J, Jr., Crowe R, Bloom FE (1998) Quantitative trait loci analysis of human event-related brain potentials: P3 voltage. Electroencephalogr Clin Neurophysiol 108:244-50.

Belknap JK (1992) Empirical estimates of bonferroni corrections for use in chromosome mapping studies with the bXD recombinant inbred strains. Behav Genet 22:677-684.

Belknap JK, Phillips TJ, O'Toole LA (1992) Quantitative trait loci associated with brain weight in the BXD/Ty recombinant inbred mouse strains. Brain Res Bull 29:337-44.

Belknap, J. K. (1998). Effect of within-strain sample size on QTL detection and mapping using recombinant inbred mouse strains. Behavior Genetics 28(1): 29-38.

Berry R, Walter NA, Stevens TJ, Rubano T, Wilcox AS, Hopkins JA, Sikela JM (1995) A strategy for the identification of candidate genes for alcohol- related phenotypes and other human disorders using rapid polymerase chain reaction mapping of gene-based sequence-tagged sites. Alcohol Clin Exp Res 19:821-3.

Besl, P. J. and McKay, N. D., (1992) “A method for Registration of 3D Shapes,” IEEE Trans. Pattern and Mach. Intell., PAMI, Vol. 14, No. 2, pp. 239-256

Black, K., Gado, M., Videen, T., Perlmutter, J., (1997) Baboon basal ganglia stereotaxy using internal MRI landmarks: validation and application to PET imaging. Journal of Computer Assisted Tomography, 21: 881-886

Bloom FE., Young WG, Kim YM (1990) Brain Browser. Academic Press, San Diego

Boddeke FR, van Vliet LJ, Young IT (1997) Calibration of the automated *z*-axis of a microscope using focus function. J Microsc 186:270–274.

Boddeke FR, van Vliet LJ, Young IT (1997) Calibration of the automated *z*-axis of a microscope using focus function. J Microsc 186:270–274.

Bookstein FL (1989) Principal warps: thin-plate splines and the composition of deformations. IEEE Trans. PAMI 11:567-585

Bookstein FL (1991) Morphometric tools for landmark data: geometry and biology. Cambridge (UK): CUP.

Borgefors, G., (1988) Hierarchical chamfer matching: a parametric edge matching algorithm, IEEE Trans. PAMI, 10(6): 849-865

Buck KJ, Metten P, Belknap JK, Crabbe JC (1997) Quantitative trait loci involved in genetic predisposition to acute alcohol withdrawal in mice. J Neurosci 17:3946-55.

Cabib S, Oliverio A, Ventura R, Lucchese F, Puglisi-Allegra S (1997) Brain dopamine receptor plasticity: testing a diathesis-stress hypothesis in an animal model. Psychopharmacology (Berl) 132:153-60.

Cannestra, F. A., Santori, M. E., Colin, H. J., Toga, W. A., (1997) A Three-dimensional Multimodality Brain Map of the Nemestrina Monkey. Brain research, 43: 141-148

Carbonell, E. A., T. M. Gerig, et al. (1992). “Interval mapping in the analysis of nonadditive quantitative trait loci.” Biometrics 48: 305-315.

Celio , MR, Hof, PR, Bloom, FE, Young, WG. (1998) A computerized stereotaxic atlas of the mouse brain. Soc Neurosci Abst 24:423.21

Chandrasekharaiah DS and Debnath L (1994), Continuum mechanics. Boston: Academic Press.

Chen EQ, Lam CF, Periasamy A (1991) Effect of refraction on optical microscope measurement of internal blood-vessel diameter and its correction. J Microsc 164:239–245.

Chen EQ, Lam CF, Periasamy A (1991) Effect of refraction on optical microscope measurement of internal blood-vessel diameter and its correction. J Microsc 164:239–245.

Cheverud, J. M. and E. J. Rotman (1995). “Epistasis and its contribution to genetic variance components.” Genetics 139: 1455-1461.

Christensen, G. E., R. D. Rabbitt, and M. I. Miller, (1996) Deformable templates using large transformations", IEEE Trans on Image Processing, 5: 1435-1447

Churchill, G. A. and R. W. Doerge (1994). "Empirical threshold values for quantitative trait mapping." Genetics 138: 963-971.

Cogswell CJ, Sheppard CJR (1991) Confocal differential interference contrast (DIC) microscopy: including a theoretical analysis of conventional and confocal DIC imaging. J Microsc 165:81–101.

Cogswell CJ, Sheppard CJR (1991) Confocal differential interference contrast (DIC) microscopy: including a theoretical analysis of conventional and confocal DIC imaging. J Microsc 165:81–101.

Cohen, F. S., Yang, Z., Huang, Z., and Nissanov, J., (1998) Computer Matching of Histological Rat Brain Sections," IEEE Trans. on Biomedical Engineering, Vol. 45, No. 5, pp. 642-649

Collins DL., Neelin P, Peters TM and Evans AC (1994) Automatic 3D intersubject registration of MR volumetric data in standardized talairach space. JCAT, 18(2)192-205.

Collins R (1970) Experimental modification of brain weight and behavior in mice: an enrichment study. Dev Psychobiol 3:145-55.

Collins, D. L., Evans, A. C., (1997) Animal: Validation and applications of nonlinear registration-based segmentation, Int. J. of Pattern Recogniton and AI 11(8): 1271-1294

Collins, D. L., Holmes, C. J., Peters, T. M. and Evans, A. C., (1995) Automatic 3D model-based neuroanatomical segmentation, Human Brain Mapping, Vol. 3, No. 3, pp. 190-208

Curcio, C.A., Sloan, K.R., Kalina, R.E. & Hendrickson, A.E. (1990) Human photoreceptor topography. J. Comp. Neurol. 292, 497–523.

Darvasi, A. (1998). "Experimental strategies for the genetic dissection of complex traits in animal models." Nature Genetics 18: 19-24.

Darvasi, A. and M. Soller (1995). "Advanced intercross lines, an experimental population for fine genetic mapping." Genetics 141: 1199-1207.

Darvasi, A. and M. Soller (1997). "A simple method to calculate resolving power and confidence interval of QTL map location." Behavior Genetics 27(2): 125-132.

Davidson D, Bard J, Brune R, Burger A, Dubreuil C, Hill W, Kaufman M, Quinn J, Stark M, Baldock R. (1997) The mouse atlas and graphical gene-expression database. Seminars in Cell and Developmental Biology 8:509-17.

Davies, R. B. (1977). "Hypothesis testing when a nuisance parameter is present only under the alternative." Biometrika 64: 247-254.

de Boor C (1978) A practical guide to splines, New York: Springer-Verlag.

Doerge, R. W. and G. A. Churchill (1996). "Permutation tests for multiple loci affecting a quantitative character." Genetics 142: 285-294.

Duchon J (1977) Splines minimizing rotation-invariant norms in Sobolev spaces. In: Constructive theory of functions of several variables, (Schempp W and Zeller K, eds), pp 85-100. Berlin: Springer-Verlag.

Eilbert, J., Gallistel, C., McEachron, D., (1990) The variation in user drawn outlines on digital images: effects on quantitative autoradiography. Computerized Medical Imaging and Graphics, 14: 331-339

Ellisman MH (1995) Design of an intermediate high voltage EM for 3D studies of biological material and its integration with a system for remote access. Proc 53rd Ann Meet MSA. Jones & Begell, New York, pp. 66–67

Entwistle A (1998) A comparison between the use of a high-resolution CCD camera and 35 mm film for obtaining coloured micrographs. J Microsc 192:81–89.

Entwistle A (1998) A comparison between the use of a high-resolution CCD camera and 35 mm film for obtaining coloured micrographs. J Microsc 192:81–89.

Erwin VG, Markel PD, Johnson TE, Gehle VM, Jones BC (1997a) Common quantitative trait loci for alcohol-related behaviors and central nervous system neurotensin measures: hypnotic and hypothermic effects. J Pharmacol Exp Ther 280:911-8.

Erwin VG, Radcliffe RA, Gehle VM, Jones BC (1997b) Common quantitative trait loci for alcohol-related behaviors and central nervous system neurotensin measures: locomotor activation. J Pharmacol Exp Ther 280:919-26.

Evans AC, Collins DL, Neelin P, MacDonald D, Kamber M, Marrett TS. (1994) Three-dimensional correlative imaging: Application in human brain mapping. In: Functional Neuroimaging Technical Foundations (Thatcher RW, Hallet M, Zeffirot T, John ER, Huerta M, eds) Academic Press, San Diego,; 145-62.

Fan GY, Mercurio PJ, Young SJ, Ellisman MH (1993) Telemicroscopy. Ultramicrosc 52:499–503.

Farkas DL, Baxter G, DeBiasio RL, Gough A, Nederlof MA, Pane D, Pane J, Patek DR, Ryan KW, Taylor DL (1993) Mulitmodal light microscopy and the dynamics of molecules, cells, and tissues. Annu Rev Phsyiol 55:785–817.

Fireston L, Cook K, Culp K, Talsania N, Preston K (1991) Comparison of autofocus methods for automated microscopy. Cytometry 12:195–206.

Fox, P.T. S. Mikiten, G. Davis and J. Lancaster, (1994) BrainMap: A database of Human Functional Brain Mapping in "Functional Neuroimaging" (R.W. Thatcher, M. Hallett, T. Zeffiro, E. Roy John and M. Huerta, eds.). Academic Press, San Diego. Pp. 95-105

Frankel WN (1995) Taking stock of complex trait genetics in mice. Trends Gen 11:471–477.

Frankel,WN, NJ Schork (1997). Who's afraid of epistasis? Nature Genetics 14: 317-373.

Friston, K. J., J. Ashbuner, C. D. Frith, J.-B. Poline, J. D. Heather, and R. S. J. Frackowiak, (1995) Spatial registration and normalization of images", Human Brain Mapping 2: 165-189

Fuller JL (1979) Fuller BWS lines: history and results. In: Development and evolution of brain size (Hahn ME, Jensen C, Dudek BC, ed), pp 187–204. New York: Academic.

Gabrani, M and Tretiak OJ (1999) Surface-based matching using elastic transformations. Pattern Recognition 32:87-97.
Galaburda AM, Aboitiz F, Rosen GD, Sherman GF (1986) Histological asymmetry in the primary visual cortex of the rat: Implications for mechanisms of cerebral asymmetry. Cortex 22:151-160.
Galaburda AM, Corsiglia J, Rosen GD, Sherman GF (1987) Planum temporale asymmetry: Reappraisal since Geschwind and Levitsky. Neuropsychologia 25:853-868.
Gallaher EJ, Jones GE, Belknap JK, Crabbe JC (1996) Identification of genetic markers for initial sensitivity and rapid tolerance to ethanol-induced ataxia using quantitative trait locus analysis in BXD recombinant inbred mice. J Pharmacol Exp Ther 277:604-12.
Gehle VM, Erwin VG (1998) Common quantitative trait loci for alcohol-related behaviors and CNS neurotensin measures: voluntary ethanol consumption. Alcohol Clin Exp Res 22:401-8.
Geschwind N, Levitsky W (1968) Human brain: Left-right asymmetries in temporal speech region. Science 161:186-187.
Ghosh, P., O'Dell, M., Narasimhan, P., Fraser, S., Jacobs, R., (1994) Mouse lemur microscopic MRI brain atlas. Neuroimage, 1: 345-9
Gilissen E, Williams RW (1997) Genetic dissection and QTL analysis of forebrain, hindbrain, olfactory bulb, and cerebellum. Soc Neurosci Abstr 23:864.
Gilissen E, Zilles K (1996) The calcarine sulcus as an estimate of the total volume of the human striate cortex: A morphometric study of reliability and intersubject variability. J Brain Res 37: 57–66.
Gill K, Liu Y, Deitrich RA (1996) Voluntary alcohol consumption in BXD recombinant inbred mice: relationship to alcohol metabolism. Alcohol Clin Exp Res 20:185-90.
Glasbey CA, Martin NH (1995) Multimodal microscopy by digital image processing. J Microsc 181:225–237.
Glaser EM, Wilson PD (1998) The coefficient of error of optical fractionator population size estimates: a computer simulation comparing three estimators. J Microsc 192:163–171.
Grisel JE, Belknap JK, O'Toole LA, Helms ML, Wenger CD, Crabbe JC (1997) Quantitative trait loci affecting methamphetamine responses in BXD recombinant inbred mouse strains. J Neurosci 17:745-54.
Gundersen HJG (1980) Stereology--or how figures for spatial shape and content are obtained by observation of structures in sections. Microsc Acta 83:409-26.
Gur RE, Mozley PD, Resnick SM, Mozley LH, Shtasel DL, Gallacher F, Arnold SE, Karp JS, Alavi A, Reivich M, Gur RC (1995) Resting cerebral glucose metabolism in first-episode and previously treated patients with schizophrenia relates to clinical features. Arch Gen psychiatry 52(80:657-67.
Gustafson, C (1995) BRAIN v 1.4 User Manual. Computer Vision Center for Vertebrate Brain Mapping. Technical Report.
Gustafson, C (1997) BRAIN v 2.0 User Manual. Computer Vision Center for Vertebrate Brain Mapping. Technical Report.
Gustafson, C (1999) BRAIN v 3.0 User Manual. Computer Vision Center for Vertebrate Brain Mapping. Technical Report.
Hahn ME, Haber SB (1978) A diallel analysis of brain and body weight in male inbred laboratory mice (Mus musculus). Behav Genet 8:251-60.
Haldane, J. B. S. and C. H. Waddington (1931). “Inbreeding and Linkage.” Genetics 16: 357-374.
Haley, C. S. and S. A. Knott (1992). “A simple regression method for mapping quantitative trait loci in line crosses using flanking markers.” Heredity 69: 315-324.
Hamilton D (1999) Net speed. Mac Addict 37: 26–30.
Harris RM (1985) Light microscopic depth measurement of thick sections. J Neurosci Meth 14:97–100.
Henderson ND (1979) Dominance for large brains in laboratory mice. Behav Genet 9:45-9.
Hibbard LS, McCasland JS, Brunstrom JE, Pearlman AL (1996) Automated recognition and mapping of immunolabelled neurones. J. Microsc 18:241–256.
Hibbard, L. S., and Hawkins, R. A., (1988) Objective Image Alignment for Three-Dimensional Reconstruction of Digital Autoradiogram, Journal of Neuroscience Methods 26, pp. 55- 74
Hibbard, L., Hawkins, R., (1987) Three-dimensional representation and analysis of brain energy matabolism. Science, 236
Horton JC, Hocking DR (1996) Intrinsic variability of ocular dominance column periodicity in normal macaque monkeys. J Neurosci 16:7228–7339.
Huang, Z. and Cohen, F. S., (1996) Affine-Invariant Moments and B-Splines for Object Recognition from Image Curves, IEEE Trans. on Image Processing, Vol. 5, No. 10, pp. 1473-1480
Ibrahim, W. S. and Cohen, F. S., (1998) Registering Coronal Histological 2-D Sections of a Rat Brain with Coronal Sections of a 3D Brain Atlas Using Geometric Curve Invariants, IEEE Trans. on Medical Imaging, Vol. 17, No. 6, pp. 957-966
Inoué S (1986) Video microscopy. Plenum Press, New York.
Jansen, R. C. (1993). “Interval mapping of multiple quantitative trait loci.” Genetics 135: 205-211.
Jeffs B, Clark JS, Anderson NH, Gratton J, Brosnan MJ, Gauguier D, Reid JL, Macrae IM, Dominiczak AF (1997) Sensitivity to cerebral ischaemic insult in a rat model of stroke is determined by a single genetic locus. Nat Genet 16:364-7.
Jiang, C. and Z.-B. Zeng (1995). “Multiple trait analysis of genetic mapping for quantitative trait loci.” Genetics 140: 111-1127.
Jiang, C. and Z.-B. Zeng (1997). “Mapping quantitative trait loci with dominant and missing markers in various crosses from two inbred lines.” Genetica 101: 47-58.
Jiang, H, R. A. Robb, and K. S. Holton, (1992) A new approach to 3-D registration of multimodality medical images by surface matching, Proc. SPIE 1808: 196-213

Katz HB, Davies CA (1983) The separate and combined effects of early undernutrition and environmental complexity at different ages on cerebral measures in rats. Dev Psychobiol 16:47-58.
Kearsey MJ, Pooni HS (1996) The genetical analysis of quantitative traits. London ; New York: Chapman & Hall.
Kim, B., Boes, J., Frey, K., Meyer, C., (1997) Mutual information for automated unwarping of rat brain autoradiographs. Neuroimage, 5: 31-40
Knapp, S. J., W. C. Bridges, et al. (1990). "Mapping quantitative trait loci using molecular marker linkage maps." Theoretical and Applied Genetics 79: 583-592.
Koslow, SH, Huerta, MF. 1997. Neuroinformatics: An Overview of the Human Brain Project. New York: Lawrence Erlbaum Associates, Publishers
Kozinska, D., J. Nissanov, O. Tretiak and C. Ozturk, (1996) Surface guided reconstruction of 3D volume from histological material. Soc. Neurosci. Abs. 22: 493.2
Kozinska, D., Tretiak, O. J., Nissanov, J. and Ozturk, C., (1997) "Multidimensional Alignment Using the Euclidean Distance Transform," Computer Vision, Graphics and Image Processing, Vol. 59, No. 6, pp. 373-387
Kruglyak, L. and E. S. Lander (1995). "A nonparametric approach for mappng quantitative trait loci." Genetics 139: 1421-1428.
Lai C, Lyman RF, Long AD, Langley CH, Mackay TF (1994) Naturally occurring variation in bristle number and DNA polymorphisms at the scabrous locus of Drosophila melanogaster. Science 266:1697-702.
Lande R (1979) Quantitative genetic analysis of multivariate evolution, applied to brain:body size allometry. Evolution 33:234–251.
Lander ES, Schork NJ (1994) Genetic dissection of complex traits. Science 265:2037-48.
Lander, E. and L. Kruglyak (1995). "Genetic dissection of complex traits: guidelines for interpreting and reporting linkage results." Nature Genetics 11: 241-247.
Lander, E. S. and D. Botstein (1989). "Mapping Mendelian factors underlying quantitative traits using RFLP linkage maps." Genetics 121: 185-199.
Levy, A., Schaewe, T., Miller, M., Smith, K., Hammoud, A., Henderson, J. et al., (1997) An Internet-connected, patient-specific, deformable brain atlas integrated into a surgical navigation system. Journal of Digital Imaging, 10: 231-7
Lincoln, S., M. Daly, et al. (1992). Mapping genes controlling quantitative traits with MAPMAKER/QTL.
Lipp HP, Schwegler H, Crusio WE, Wolfer DP, Leisinger-Trigona MC, Heimrich B, Driscoll P (1989) Using genetically-defined rodent strains for the identification of hippocampal traits relevant for two-way avoidance behavior: a non-invasive approach. Experientia 45:845–859.
Lu L, Airey, DA, Gilissen E, Zhou G, Williams RW (1999) Int Mamm Genome Soc 13: in press.Allen RD, Strömren Allen N, Travis JL (1981) Video-enhanced contrast, differential interference contrast microscopy. Cell Motil 1:291–302.
Lumsdon T, Dickson M, Gladden MH (1995) A system for aligning video-recorded visual data andhigh-quality multichannel recordings of simultaneously occurring signals. J Microsc 178:182–184.
Lynch M, Walsh B (1998) Genetics and analysis of quantitative traits. Sunderland, Ma.: Sinauer.
Mackay TF (1996) The nature of quantitative genetic variation revisited: lessons from Drosophila bristles. Bioessays 18:113-21.
Maes, F., Collignon, A., Vandermeulen, D., Marchal, G., and Suetens, P., (1997) Multimodality image registration by maximzation of mutual information, IEEE Trans. Med. Im. 16(2): 187-198
Mangin, B., B. Goffinet, et al. (1994). "Constructing confidence intervals for QTL location." Genetics 138(4): 1301-1308.
Manly, K. F. (1993). "A Macintosh program for storage and analysis of experimental genetic mapping data." Mammalian Genome 4: 303-313.
Manly, K. F. and J. M. Olson (1999). "Overview of QTL mapping software and introduction to Map Manager QT." Mammalian Genome 10: 327-334.
Manly, K. F. and R. W. Elliott (1991). "RI Manager, a microcomputer program for analysis of data from recombinant inbred strains." Mammalian Genome 1: 123-126.
Martinez, O. and R. N. Curnow (1992). "Estimating the locations and the sizes of the effects of quantitative trait loci using flanking markers." Theoretical and Applied Genetics 85: 480-488.
Martinez, O. and R. N. Curnow (1994). "Missing markers when estimating quantitative trait loci using regression mapping." Heredity 73: 198-206.
McEachron, D.L., Tretiak, O.J., and Feingold, E. 1987. Autoradiography and video densitometry: Image processing with DUMAS. *Functional Photography*: 30-44.
Meinguet J (1979) An intrinsic approach to multivariate spline interpolation at arbitrary points. In: NATO Advanced Study Institute on Polynomial and Spline Approximation , (Sahney BN, ed), pp. 163-190. Dordrecht: D. Reidel Publ. Co..
Moreno-Gonzalez, J. (1992). "Genetic models to estimate additive and non-additive effects of marker-associated QTL using multiple regression techniques." Theoretical and Applied Genetics 85: 435-444.
Morrison J, Hof P (1998) Introduction to the short course in quantitative neuroanatomy: a picture is worth a thousand words, but a number is worth a thousand pictures. Soc Neurosci, Washington, p vii–viii.
Nissanov J, Bertrand L (1998a) .2D Digital Atlas of the Adult Rat Brain. Computer Vision Center for Vertebrate Brain Mapping, a NIH Biomedical Technology Resource, Drexel University.
Nissanov J, Bertrand L (1998b) 3D Anatomy of the Rat Brain: A 42 _m Isotropic Atlas, V 1. Computer Vision Center for Vertebrate Brain Mapping, a NIH Biomedical Technology Resource, Drexel University.
Nissanov J, McEachron DL (1991) Advances in image processing for autoradiography. J. Chem. Neuroanat. 4: 329-42.
Nissanov, J. 1998. Brain 3.0: Exercise Manual. Imaging & Computer Vision Center #03-31-1998

Nissanov, J., Bertrand, L., Gustafson, C., Lo, J., and O. Tretiak. Submitted. A 42 _m isotropic rat brain atlas reconstructed from 2D histological sections. J. Neurosci Methods.

Oldenbourg R, Terada H, Tiberio R, Inoué S (1993) Image sharpness and contrast transfer in coherent confocal microscopy. J Microsc 172:31–39.

Ooijen, J. W. V. (1992). "Accuracy of mapping quantitative trait loci in autogamous species." Theoretical and Applied Genetics 84: 803-811.

Ozturk, C., Bertrand, L and J. Nissanov. Submitted. Interbrain variability in the rat: implications for automatic parcellation. Neuroimage

Pakkenberg B, Gundersen HJG (1997) Neocortical neuron number in humans: effect of sex and age. J Comp Neurol 384:312–320.

Paxinos, G., Watson, C., Pennisi, M., Topple, A., (1985) Bregma, lambda and the interaural midpoint in stereotaxic surgery with rats of different sex, strain and weight. J Neurosci Methods 13(2): 139-143

Pechura C, Martin, JB. 1991. Mapping the Brain and its Functions: Integrating Enabling Technologies into Neuroscience Research. Washington D.C.:National Academy Press

Pelizzari, C.A., G.T.Y. Chen, D. R. Spelbring, R. R. Weichselbaum, C. T. Chen, (1989) "Accurate three-dimensional registration of CT, PET, and/or MR images of the brain", JCAT, 13(1): 20-26

Pluta M (1989) Advanced light microscopy, Vol 2. Elsevier, Amsterdam.

Preza C, Snyder DL, Conchello JA (1996) Imaging modesl for three-dimensional transmitted-light DIC microscopy. Proc SPIE Symp in Three-Dimensional Microscopy: Image Acquisition and Processing IV 2984:220–231.

Rasband W (1995) NIH Image 1.59 Manual. National Institutes of Health, Bethesda MD

Rebai, A., B. Goffinet, et al. (1995). "Comparing power of different methods for QTL detection." Biometrics 51(1): 87-99.

Reed, CL, Nissanov, J., Bertrand, L and EJ Gallaher. 1999. Computerized 3-D Mouse Brain Atlas. Soc. Neurosci. Abst. 25:823.1

Rice DS, Williams RW, Goldowitz D (1995) Genetic control of retinal projections in inbred strains of albino mice. J Comp Neurol 354:459-69.

Roderick TH, Wimer RE, Wimer CC (1976) Genetic manipulation of neuroanatomical traits. In: Knowing, thinking, and believing (Petrinovich L, McGaugh L, ed), pp New York: Plenum.

Roderick TH, Wimer RE, Wimer CC, Schwartzkroin PA (1973) Genetic and phenotypic variation in weight of brain and spinal cord between inbred strains of mice. Brain Res 64:345-53.

Rosen GD, Harry JD (1990) Brain volume estimation from serial section measurements: A comparison of methodologies. J Neurosci Meth 35:115–124.

Rosen GD, Sherman GF, Galaburda AM (1989a) Interhemispheric connections differ between symmetrical and asymmetrical brain regions. Neuroscience 33:525–533.

Rosen GD, Sherman GF, Galaburda AM (1991) Ontogenesis of neocortical asymmetry: A [^{3}H]thymidine study. Neuroscience 41:779–790.

Rosen GD, Sherman GF, Galaburda AM (1993) Neuronal subtypes and anatomic asymmetry: Changes in neuronal number and cell-packing density. Neuroscience 56:833–839.

Rosen GD, Sherman GF, Mehler C, Emsbo K, Galaburda AM (1989b) The effect of developmental neuropathology on neocortical asymmetry in New Zealand Black mice. Int J Neurosci 45:247–254.

Roysam B, Ancin H, Bhattacharjya AK, Chisti MA, Seegal R, Turner JN (1994) Algorithms for automated characterization of cell populations in thick specimens from 3-D confocal fluorescence microscopy data. J Microsc 173:115–126

Rubattu S, Volpe M, Kreutz R, Ganten U, Ganten D, Lindpaintner K (1996) Chromosomal mapping of quantitative trait loci contributing to stroke in a rat model of complex human disease [see comments]. Nat Genet 13:429-34.

Santos A, Ortiz de Solórzano C, Vaquero JJ, Peña JM, Malpica N, Del Pozo F (1997) Evaluation of autofocus functions in cytogenetic analysis. J Microsc 188:264–272.

Schormann T, Jovin TM (1991) Contrast enhancement and depth perception in three-dimensional representations of differential interference contrast and confocal scanning laser microscope images. J Microsc 166:155–168.

Schormann, T., Zilles, K., (1998) Three-dimensional linear and nonlinear transformations: an integration of light microscopical and MRI data. Hum Brain Mapp 6(5-6): 339-347

Seyfried TN, Daniel WL (1977) Inheritance of brain weight in two strains of mice. J Hered 68:337–338.

Slotnick, B M., Brown, BL, (1980) Variability in the stereotaxic position of cerebral points in the albino rat. Brain research, 5: 135-139

St-Jean, P., Sadikot, A., Collins, L., Clonda, D., Kasrai, R., Evans, A. et al., (1998) Automated atlas integration and interactive three-dimensional visualization tools for planning and guidance in functional neurosurgery. IEEE Transaction on Medical Imaging, 17: 672-80

Strom RC, Williams RW (1997) Mapping genes that control variation in brain weight using F2 intercross progeny. Soc Neurosci Abst 23:864.

Talairach J, Tournoux P. Co-Planar Stereotaxic Atlas of the Human Brain. 3-Dimensional Proportional System: An Approach to Cerebral Imaging. Georg. Theime Verlag, New York, 1988

Tanksley SD (1993) Mapping polygenes. Annu Rev Genet 27:205-33.

Taylor BA (1978) Recombinant inbred strains. Use in gene mapping. In: Origins of inbred mice (Morse H, ed), pp 423–438. New York: Academic.

Taylor BA, Wnek C, Kotlus BS, Roemer N, MacTaggart T, Phillips SJ (1999) Genotyping new BXD recombinant inbred mouse strains and comparison of BXD and consensus maps. Mamm Genome 10:335-48.

Thévenaz P and Unser M (1995) Efficient geometric transformations and 3-D image registration, (invited paper) Proc.IEEE Int. Conference on Acoustics, Speech, and Signal Processing, Detroit, MI, USA, May 8-12, vol. V, pp. 2919-2922

Thévenaz, P. and M. Unser, (1996) A pyramid approach to sub-pixel image fusion based on mutual information, Proc. IEEE Int. Conf. on Image Processing, Lausanne, Switzerland, September 16-19, vol. I, pp. 265-268

Thompson, P. M., MacDonald, D., Mega, M. S., Holmes, C. J., Evans, A. C., Toga A. W., (1997) Detection and mapping of abnormal brain structure with a probabilistic atlas of cortical surfaces. J. Comput. Assist. Tomogr. 21(4): 567-581

Thompson, P., Toga, A., (1997) Detection, visualization and animation of abnormal anatomic structure with a deformable probabilistic brain atlas based on random vector field transformations. Medical Image Analysis, 1: 271-94

Toga AW, Samaie M, Payne BA (1989) Digital rat brain: a computerized atlas. Brain Res Bull 22:323-33

Toga AW, Ambach, K. L., and Schluender, S., (1994) High-resolution anatomy from in situ human brain. Neuroimage 1(4): 334-344

Toga AW, Banerjee, P. K., (1993) Registration revisited. Journal of Neuoscience Methods, 48: 1-13

Toga AW, Santori, E. M., Hazani, R., Ambach, K., (1995) A 3D digital map of rat brain. Brain Res. Bull. 38(1): 77-85

Toga, AW, Thompson, P (1999) An introduction to Brain Warping. In: Brain Warping (Toga, AW, ed.) pp 1-26. San Diego: Academic Press.

Tolliver BK, Belknap JK, Woods WE, Carney JM (1994) Genetic analysis of sensitization and tolerance to cocaine. J Pharmacol Exp Ther 270:1230-8.

Visscher, P. M., R. Thompson, et al. (1996). "Confidence intervals in QTL mapping by bootstrapping." Genetics 143(2): 1013-1020.

Voelkl E, Allard LF, Bruley J, Williams DB (1997) Undergraduate TEM instruction by telepresence microscopy over the Internet. J Microsc 187:138–142.

Wahlsten D (1983) Maternal effects on mouse brain weight. Dev Brain Res 9:215-221.

Wells, W. W., Viola, P., Atsumi, H., Nakajima, S. and Kikinis, R., (1996) Multi-modal volume registration by maximization of mutual information, Medical Image Analysis, 1(1): 35-51

West MJ, Gundersen HJG (1990) Unbiased stereological estimation of number of neurons in the human hippocampus. J Comp Neurol 296:1–22.

Whittaker, J. C., R. Thompson, et al. (1996). "On the mapping of QTL by regression of phenotype on marker-type." Heredity 77: 23-32.

Williams RW, D Goldowitz,, Strom R (1997) Brain weight in relation to body weight, age, and sex: A multiple regression analysis. Soc Neurosci Abstr 864.

Williams RW, Rakic P (1988) Three-dimensional counting: An accurate and direct method to estimate numbers of cells in sectioned material. J Comp Neurol 278:344-352.

Williams RW, Strom RC, Goldowitz D (1998) Natural variation in neuron number in mice is linked to a major quantitative trait locus on Chr 11. J Neurosci 18:138-46.

Williams RW, Strom RC, Goldowitz DG (1996a) Mapping quantitative trait loci that control normal variation in brain weight in the mouse. Soc Neurosci Abstr 22:518.

Williams RW, Strom RC, Rice DS, Goldowitz D (1996b) Genetic and environmental control of variation in retinal ganglion cell number in mice. J Neurosci 16:7193-7205.

Willis B, Roysam B, Turner JN, Holmes TJ (1993) Iterative, constrained 3-D image reconstruction of transmitted light bright-field micrographs based on maximum-likelihood reconstruction. J Microsc 169:347–361.

Wimer CC, Wimer RE (1989) On the sources of strain and sex differences in granule cell number in the dentate area of house mice. Brain Res Dev Brain Res 48:167-76.

Woods, R. P., Grafton, S. T., Holmes, C. J., Cherry, S. R., Mazziotta, J. C., (1998a) Automated image registration: I. General methods and intrasubject, intramodality validation. J. Comput. Assist. Tomogr. Jan-Feb; 22: 139-52

Woods, R. P., J. C. Mazziotta, and S. R. Cherry, (1993) "MRI-PET registration with automated algorithm", JCAT 17: 536-546

Wu SH, Barba J, Gil J (1996) A focusing algorithm for high magnification. J Microsc 184:133–142.

Young D, Glasbey CA, Gray AJ, Martin NJ (1998) Toward automatic cell identification in DIC microscopy. J Microsc 192:186–193.

Zaluzec NJ (1995) Tele-presence microscopy: an interactive multi-user environment for collaboratrive researhc using high speed networks and the Internet. Proc 53rd Ann Meet MSA, Jones & Begell, New York, pp. 14–15.

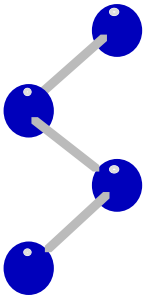
Zeng, Z.-B. (1993). "Theoretical basis for separation of multiple linked gene effects in mapping quantitative trait loci." Proc Nat Acad Sci USA 90: 10972-10976.

Zeng, Z.-B. (1994). "Precision mapping of quantitative trait loci." Genetics 136: 1457-1468.

Zhou G, Williams RW (1999) Mouse models for the analysis of myopia: an analysis of variation in eye size of adult mice. Optom Vis Sci 76:408-18.

Zilles K (1985) The Cortex of the Rat: A Stereotaxic Atlas. Berlin: Springer-Verlag.

Zilles K, Zilles B, Schleicher A (1980) A quantitative approach to cytoarchitectonics. VI. The areal pattern of the cortex of the albino rat. Anat Embryol 159:335–360.



THE SCRIPPS RESEARCH INSTITUTE

Warren G. Young Ph.D.
Director, NP Computing
10550 North Torrey Pines Road
Mail Code: SBR1
La Jolla, CA 92037

Nov. 18, 1999

Robert W. Williams
Center for Neuroscience
Department of Anatomy and Neurobiology
University of Tennessee
855 Monroe Avenue
Memphis TN 38163

Dear Rob,

I am writing to indicate my willingness and enthusiasm to serve as a consultant on your program project grant titled “Informatics Center for Mouse Neurogenetics” should it be funded. As developer of NeuroZoom, a microscope software application for mapping and quantifying neuroanatomic structures, I would try to apply the features of that application to the datasets that you are developing. I am also actively working on an information database system that covers the entire spectrum of scientific data, from the raw, newly acquired data from lab instruments, to the final published paper. Applying your Mouse Brain Library to these kinds of data would be very useful to the scientific community at large.

I look forward to a strong collaborative effort with you and your team.

Sincerely,

Warren G. Young

Warren G. Young, Ph.D.
Dir, NP Computing
Department of Neuropharmacology
The Scripps Research Institute

Application Update:

To the reviewer:

1. In the last four months we have made progress that we think could be helpful in assessing this application and prospects for attaining our goals. In this application we have proposed as part of the Genotyping and Mouse Colony Core to generate ~400,000 genotypes, an imposing number even for a well-funded PPG core. This December Dr. Williams, Manly, and Rosen applied to the Mammalian Genotyping Service (an NHLBI-funded facility run by Dr. James Weber) to genotype a group of our advanced intercross mice. They have accepted our application and have agreed to carry out 180,000 genotypes in the next year on our behalf (500 mice at 350 markers). This still leaves us with a very significant amount of genotyping locally (including our aging population of mice in year 03 and 04), but we are now virtually guaranteed to have a large number of genotypes with which to populate our databases even before the grant period would begin. Thus, we expect that in year 01 we will at minimum have genotypes and some phenotypes (e.g., body and brain weight) for no less than 500 case. Dr. M. Dockter, Dean for Research at UT, has purchased a 96-channel ABI 3700 DNA sequencer that we will be using for genotyping. The 3700 will be placed in Genotyping Core and will allow us to significantly improve our output.

2. In the last three months a group of us (Dr. Manly, PI of Project 4, Dr. David Threadgill from Vanderbilt University, and myself) have developed an improved method to map QTLs. The method is statistically much more powerful than other QTL mapping methods and Monte Carlo analysis has shown that QTLs with effects as small as 0.2 SD should be consistently detected. The method is referred to as RIX mapping (recombinant inbred intercross mapping). This method has very high positional precision compared to an F2 intercross and equivalent to that of our G10 cross. RIX mapping is based on the simple procedure of intercrossing RI strains in all possible combinations. For example, the set of 35 BXD stains can be crossed to generate 595 RIX lines. The method can also be generalized across RI sets; for example BXD strains can be crossed to AXB strains. In this way as many as 5000 lines segregating for alleles inherited from C57BL/6J can be generated using commercially available stocks maintained by the Jackson Laboratory.

   How does this development affect our grant application? First, as part of the Mouse Genotyping and Colony Core, we now propose to generate at least 50 RIX lines per

year, and to add these lines to the MBL collection. This addition does not affect our budget request. Funds for colony support will come in part from the Mouse Colony Core and in part from institutional funds for mouse genomics provide to our co-investigator Dr. Dan Goldowitz. Dr. Goldowitz and I are now in the process of generating the complete set of 76 RIX strains from the CXB RI set. The second consequence is that we will now need to carry out additional genotyping of four commonly used RI sets (BXD, AXB, CXB, and BXH). This genotyping is already in progress in my lab and will be completed in year 01 or early in year 02 of this grant. Our goal is to genotype all of the major RI sets with a common set of approximately 500 microsatellite markers. This will make it relatively straightforward to pool data across RI and RIX sets.

3. Recent successes in using the MBL for neurogenetic investigation. Dr. Rosen has used the MBL to study the striatum of several strains of mice and of an F2 intercross. All of the material that he has studies (over 60 cases) is now a part of the MBL and is accessible to anyone via the net.

4. Dr. Nissanov and colleagues have produced a much improved high resolution 3D atlas of the mouse brain. Their new atlas has an isotropic resolution of 15 μm. They are now busy segmenting and labeling nuclei. The new techniques that they have adapted to align sections has resulted in a representation in which it is difficult to determine the original plane of section.

List of Appendix Material

Printouts from the Mouse Brain Library

Printouts from the Internet Microscopy Project

Printouts from Drexel

Printouts from Map Manager QTb28

Manly KF, Cudmore Jr RH, Kohler G (1997) User's Manual for Map Manager Classic and Map Manager QT.

Manly KF, Olson JM (1999) Overview of QTL mapping software and introduction to Map Manager QT. Mamm Genome 10:327–334.

Williams RW (2000) Mapping genes the modulate mouse brain development: A quantitative genetic approach. In Goffinet A, Rakic P, eds. Mouse Brain Development. Springer-Verlag Berlin, pp. 21–49.

Williams RW (1998) Neuroscience meets quantitative genetics: Using morphometric data to map genes that modulate CNS architecture. Short Course Syllabus: Quantitative Neuroanatomy, Morrison J, Hof P, eds. pp. 66–78. Soc Neurosci, Washington DC.

Williams RW, Strom RC, Goldowitz D (1998) Natural variation in neuron number in mice is linked to a major quantitative trait locus on Chr 11. J Neurosci 18:138–146.

Rosen GD, Harry JD (1990) Brain volume estimation from serial section measurements: a comparison of methodologies. J Neurosci Meth 35:115–124.

Rosen GD, Sherman GF, Emsbo K, Mehler C, Galaburda AM (1990) The midsagittal area of the corpus callosum and total neocortical volume differ in three inbred strains of mice. Exp Neurol 107:271–276.

Rosen GD, Sherman GF, Galaburda AM (1993) Neuronal subtypes and anatomic asymmetry: changes in neuronal number and cell-packing density. Neuroscience 56:833–839.

Nissanov paper 1

MICHAEL D. HIRSCH, PH.D.
(301) 443-1815
MHIRSCH@NIH.GOV

SUMMARY STATEMENT
(Privileged Communication)

Application Number: 1 P20 MH62009-01
DUAL PROGRAM CLASS CODE: RR/JDP DUAL: DAAGHDLMTWAAHLDE
ZMH1 BRB-P (1)
Review Group: NIMH SPECIAL EMPHASIS PANEL

Meeting Dates: IRG: FEB/MARCH 2000 COUNCIL: MAY 2000 NI
Requested Start Date: 07/01/2000
PAR99-138

WILLIAMS, ROBERT W, PHD
UNIV OF TENNESSEE, MEMPHIS
DEPT OF ANATOMY & NEUROBIO
855 MONROE AVENUE
MEMPHIS, TN 38163

Project Title: INFORMATICS CENTER FOR MOUSE NEUROGENETICS

IRG Action: Priority Score: 138
Human Subjects: 10-NO HUMAN SUBJECTS INVOLVED
Animal Subjects: 30-ANMLS INV.-VERIFIED, NO IRG CONCERNS OR COMMENT

GENDER, MINORITY, CHILDREN, & CLINICAL TRIAL CODES NOT ASSIGNED

PROJECT YEAR	DIRECT COSTS REQUESTED	DIRECT COSTS RECOMMENDED	ESTIMATED TOTAL COST
01	1,134,107	1,134,107	1,244,877
02	1,094,723	1,094,723	1,201,646
03	1,102,538	1,102,538	1,210,225
04	1,039,670	1,039,670	1,141,216
05	1,067,723	1,067,723	1,172,009
TOTAL	5,438,761	5,438,761	5,969,973

NOTE TO APPLICANT FOLLOWS SUMMARY STATEMENT

RESUME AND SUMMARY OF DISCUSSION: This is a new application for a P20 exploratory center being submitted by Dr. Robert Williams from the University of Tennessee. There are four research projects and a neurohistology / genotyping core in this proposal. The overall goal is to develop informatics tools for mouse neurogenetics focusing on the structural variation in the mouse CNS through the mapping of quantitative trait loci (QTLs).

PROJECT 1 "Mouse Brain Library" (PI: Dr. Glenn Rosen, Harvard University). The MBL consists of a large-volume library of brain sections suitable for morphometric investigation. This is a strong application where the breadth of its collection is the key factor in its utility. The attention to detail in this application is impressive. Aside from minor weaknesses such as the need for plans for publicizing the database, no serious concerns were identified in this exciting application. (PS = 170)

PROJECT 2 "Internet Microscope Systems" (PI: Dr. Robert Williams, University of Tennessee). The iScope captures and displays detailed image stacks over the web through the use of robotic slide handlers allowing stereological studies of the brain scans in the MBL. This project represents a forward reaching approach to dealing with the problem of how to maximize use of valuable brain collections.

Both hardware and software components of the iScope are based on solid foundations. Aside from some suggestions on how to realize the full capability of the iScope system, the Panel was highly enthusiastic about its promise. (PS = 154)

PROJECT 3 "NeuroCartographer" (PI: Dr. Jonathan Nissanov, Drexel University). This project will develop a suite of software tools and 3D models of neuroanatomical structures allowing users to reconstruct and digitally dissect material in the MBL. The research team has compiled an impressive record of accomplishment with respect to the automation of the multiple laborious and complex operations that are involved in the construction of digital atlases and in the use of the atlases for assemblage of VOIs (volumes of interest) for QTL mapping. The amount of work proposed is enormous especially given the inherent difficulty in achieving accurate spatial transformations of 2D data and converting 2D into a 3D format. However, the research team appears to be well-positioned to make headway toward the stated goals. (PS = 154)

PROJECT 4 "Neurogenetics Tool Box" (PI: Dr. Kenneth Manly, Roswell Park Cancer Institute). The NTB is composed of a set of gene mapping programs and aims to develop a state of the art information and reporting database for QTL analysis for the most common panel of recombinant inbred mouse strains and various intercrosses. This will provide for a very user-friendly utility for investigators of brain and behavioral QTLs. The choice of mapping algorithms is solid. Moreover, the free and open access to both programs and code is a significant strength. Only minor weaknesses were identified with the NTB. (PS = 165)

CORE "A. Neurohistology" (PI: Dr. Glenn Rosen) "B. Genotyping" (PI: Dr. Robert Williams). Core A, Neurohistology Core is adequately staffed with personnel and facilities for execution of needed project functions. The histology Core is fully justified on grounds that it will avoid duplication of facilities in other project components. Core B, Genotyping and Mouse Colony Core is also adequately equipped for fixing mouse brains for histology and for extraction of genetic information to be entered into the Neurogenetic database. As with Core A, Core B is an integral and essential component of the program project. (PS = 157)

Overall, this is an outstanding proposal that is highly significant to the goals and objectives of the Human Brain Project. The four research projects are solid proposals that can stand on their own yet demonstrate strong interdependence arguing for the utility of a P20 exploratory center mechanism. The investigators are excellent researchers with demonstrated records of productivity. The PI, Dr. Robert Williams, is very qualified to lead the Center. There is evidence of good institutional support. The proposed research plans are well thought out and demonstrate convincing evidence of feasibility. No major concerns were identified in this application. Only relatively minor weaknesses were raised, and even then, those issues centered on suggestions for exploiting the full potential and promise of this Center. It was the Committee's opinion that the strengths of the application far outweigh its weaknesses. Consequently, the Panel was very enthusiastic about this highly meritorious application.

OVERALL DESCRIPTION (Applicant's Abstract): The purpose of this

Neuroinformatics project is to develop and exploit a suite of image databases, motorized Internet microscopes, and software to study the genetic basis of structural variation of the mouse CNS. Resources are open to the research community through an integrated web interface at <nervenet.org>. The focus of this project is to provide a collaborative research environment for mapping quantitative trait loci (QTLs). These genes are responsible for the extraordinary variation in CNS structure among mice and humans. QTL analysis is a burgeoning field that tackles complex biological traits modulated by many genes. We will develop four significant new resources and technologies: (1) The Mouse Brain Library (MBL) consists of a huge, well-organized library of brain sections suitable for morphometric investigation. Thousands of images can be rapidly searched, sorted, and downloaded at a resolution of 5 microns per pixel using an intuitive and powerful web interface. (2) The Internet Microscope System (iScope) captures and displays extremely detailed movies -z-axis image stacks-suitable for sophisticated stereological study of all brains in the MBL. The iScope includes robotic slide handlers controlled over the web 24 hours a day, 7 days a week (3) The NeuroCartographer Project will develop a suite of software tools and 3D models of hundreds of neuroanatomical structures that will enable researchers to reconstruct and digitally dissect; material in the MBL. (4) The Neurogenetics Tool Box (NTB) comprises a set of gene mapping programs that will enable neuroscientists to rapidly identify and evaluate QTLs responsible for the astonishing variation in CNS architecture. The NTB will include genotypes from an unusually large advanced intercross designed to map loci with sufficient precision to enable a candidate gene approach to cloning QTLs. Achieving the aims of these four projects will catalyze a new era in the structural analysis of the adult mammalian nervous system and will lead to a large number of novel lines of research on the development, aging, and pathology of the human brain.

PROJECT 1

Mouse Brain Library

PI: Dr. Glenn Rosen

CRITIQUE 1:

The PI proposes to take an extant and impressive database and extend it by adding on additional animals stained for Nissl substance, sections stained with or for other substances, and a greater range of images. Of these the first is easiest to justify, as a greater subject number for any RI strain and a greater range of animals, whether intercrosses or standard inbreds, would increase the probability that a quantitative variation is picked up and documented quantitatively. The decision to focus on RIs and standard inbreds is proper and elegant, as these are likely to produce both the most easily established and most generally useful bodies of data.

The second aim; production of data from localizing something other than basophilic material, gets to the heart of the one potential weakness of this project: what is it that is being quantified? As set up currently and as apparently planned for the immediate future the best quantitative data appears

to deal with the size of large neuronal ensembles rather than with estimates of population size or density. That is a conclusion by this reviewer rather than a fact stated by the PI, however, and it derives from the design of the project and of data published previously. Specifically, the proposal to capture images of each section at 4.5 micrometer per pixel at what would be a single focal plane precludes use of the best appropriate stereological method, the optical dissector, in this material. The related proposal to capture images of a few dozen regions at 1 micrometer per pixel could conceivably be adapted for the dissector approach, although that does not appear explicitly in the project outline. Finally, the quantitative approach taken in the submitted publication of immunostained material (Rosen et al., 1993) is neither stereological nor design-based and data generated by such an approach would be of limited use to other investigators. One concludes then that the data are to be collected so that the relative size and volume of large regions (e.g. cortical areas or subcortical nuclei) can be quantified. These are by no means trivial things to determine and, it is worth pointing out, that macro-variations identified in MBL could be further probed in the iScope to generate hard estimates of neuronal populations.

Choices of what to examine other than Nissl substance boil down, then, to those things that permit an investigator to draw more accurate boundaries, such as those between cortical areas or thalamic nuclei. Some of the choices listed in presentation of the Histology Core are more than reasonable for this purpose (e.g. myelin stains and parvalbumin immunoreactivity) whereas others (choline acetyltransferase) appear to be somewhat idiosyncratic. Therefore, it is worth reinvestigating the issue of what is most useful to the field in general and to employ those methods.

The proposal to use immunostained and histologically stained material raises other issues of providing sufficient subject numbers and of matching material prepared by frozen sectioning with that prepared by celloidin embedding. The PI has done an excellent job of adapting methods for sectioning brains to maximize output for Nissl stains. However, for histochemically and immunocytochemically prepared material the investigators are back to cutting and processing brains one at a time. Moreover processing for these methods produces very different patterns of tissue shrinkage than celloidin embedding. Whereas the former is not tractable, but accents the need to choose stains wisely, the latter may be as counterstains such as the Giemsa stain are appropriate for immunostained material. There seems little reason to get particularly concerned about any of these issues, however, since the PI of this project and the PI of the proposal are experienced in the application of chemical methods to study the CNS.

Dr. Rosen is an established investigator who employs principally anatomical methods to study animal models of human neurological diseases. The PI has developed each of the skills necessary to accomplish the outlined experiments. The facilities at Beth Israel are appropriate for the work that is proposed.

CRITIQUE 2:

The Mouse Brain Library is a serious undertaking to make an Internet resource of mouse neurogenetics available to the widest range of researchers in the most user-friendly manner. Judging on both past accomplishments and plans for the future, it clearly succeeds in this.

Strengths: The breadth of the collection is a key factor in its utility. The PI tends to focus on the utility of the material for the conducting of prospective genetic studies. An equally valuable feature of the library, however, is the occasional user who would want to compare an altered neurological structure in a transgenic or knockout mouse with a well accepted control animal. This user will appreciate that no matter how many strains the mutants were crossed to, the background brain structures will likely be available in the collection.

The attention to the informatics dimension of the project is excellent. There is a significant amount of attention to detail that is apparent here and the result is a powerful and useful tool for morphometric studies.

The structure of the web access is intuitive and well designed.

The plans to implement searches based on region of interest and to enhance the resolution at these locations will be sure to increase use of the site.

Weaknesses: Despite these important strengths and overall strong points of the application, there are still some areas that might benefit from some attention.

The utility of the database will be directly proportional to its degree of usage, and the function of proportionality is probably exponential. No effort is spelled out to generate publicity for the site and to encourage its use by geneticists, neurobiologists and molecular biologists. If this proposal is implemented as is, the only way for a naïve user to find it is by accident or word of mouth. The latter can be an effective promotional tool, but some explicit attempt to "market" the site would have been useful. In the same vein, some technophobic users might appreciate a "help line". The goal is to increase the use of the resource thus expanding its power.

A related problem is the lack of background material provided to the reviewers on the specifics of the approach. It is difficult to find a compelling rationale for the use of QTL analysis by researchers other than specialized quantitative geneticists. The literature to date is not vast and thus, in addition to summarizing the fruits of past work, it would have been a good idea to outline how much power a non-expert would have at their disposal. This reviewer could not even find a detailed description of the 'advanced intercross'. The jest of the approach is clear, but not the details. These will be crucial to a first time user (and non-geneticist).

The choice of immunocytochemistry stains is capricious to a general user and probably not worth pursuing. A random neurobiologist might come up with a list of two or three other "favorite" stains. The focus on the cholinergic system is never justified and this aspect should probably be dropped.

The availability of color at the site is of questionable value given the way the resource is likely to be used. The requisite bandwidth and storage facility could well be put to other uses.

RECOMMENDATION: Priority Score = 170

PROJECT 2

Internet Microscope Systems

PI: Dr. Robert Williams

CRITIQUE 1:

The iScope project is an innovative marriage of robotic hardware systems with software programs that will enable investigators at large to utilize the extensive and expanding mouse tissue library created by Project l. This project embodies many of the goals of the Neuroinformatics initiative as iScope capitalizes on ground breaking technology and takes it a step farther to acquire, store, and disseminate a large amount of neuroanatomic data. Moreover, the development of a virtual library of histological sections for quantitative analysis via the Web will establish a prototype that may be adapted for widespread application to the study of brain morphometry and neuropathologic correlates of human disease.

Significance: The iScope represents a forward reaching solution to the dilemma of how to maximize use of valuable brain collections. As histologic sections generated by the Mouse Brain Library (MBL) in Project l become available for researchers world-wide via an automated online microscope, the easy accessibility of tissue for quantitative study will have a profound impact in bolstering the field of mouse brain morphometry and in particular the search for quantitative trait loci (QTLs) associated with morphologic phenotypes. Correlational analyses of strain-related brain morphology will proliferate exponentially as numerous investigators, each pursuing their own line of study on the same databank of histologic sections, are able to compare results with other iScope users. Moreover, the availability of this resource undoubtedly will result in a reduction in overall research costs, as other laboratories will not need to generate inbred mouse crosses or perform histological processing, and will curtail the number of animals sacrificed for research via elimination of redundancy in competing laboratories around the world. Implementation of the iScope is also likely to drive the fields of robotic microscopy and computer aided image capture to improve in resolution and speed.

Approach: Both hardware and software components of the iScope are based on a solid foundation of methods that have already been applied to previous research problems. The robotic microscope is patterned after the slide-handling system developed by Dr. Nissanov at Drexel University and has been duplicated with modifications by Dr. Williams at the University of Tennessee. The programs for stereologic analysis of cell populations in the image stacks will be based on the Counting Box program developed at Yale, and streaming video technology will utilize Apple's QuickTime4. Thus, the iScope appears to be entirely feasible although all aspects of this methodology will be pushed to new limits by the enormity of the task at hand.

Overall, Dr. Williams has done a very thorough job of considering potential

problems and pitfalls associated with microscopy over the Internet. The live image stacks or archived files of these stacks which will cover approximately 200 sample sites per brain will satisfy most users' needs. However, to make the iScope truly the equivalent of having slides on the scope in front of the user, the investigator should be able to select a region of interest and select a custom sampling paradigm. This ability requires two modifications. The first is that image stacks can be generated at any point on the sections; the second that the software package allows the user to specify the spacing of the counting boxes over the region of interest. Customization of the sampling grid for stereological analysis will ensure that the program has universal utility and maximal efficiency in estimation of different cell populations, e.g., neurons versus glia in the same nucleus, or different size nuclei, e.g., the striatum versus the locus coeruleus. The applicants should be encouraged to address this technical limitation in order to realize the full potential of the iScope system.

One potential limitation of the iScope is that there will be a significant backlog in usage due largely to the labor intensive nature of stereological research even with the availability of five scopes. To address this concern, the applicants should estimate the average time that a single investigator would require online to analyze a hypothetical sample, e.g., 25 image stacks (which is a very modest number for stereological research) in each of 8 animals in 4 different mouse strains, as well as how many hours per week an individual researcher can expect to have access to the iScope. Given these two parameters, will use of the iScope be a practical alternative to conventional in laboratory investigation? Another related problem is that the utility of the iScope for stereological analysis of regions of interest is largely dependent on the generation of maps for different strains in the Project 3. At the end of the 1st year, only 8 atlases will be available, i.e. 2 strains will be mapped with two different stains in males and females.

Innovation: The establishment of a virtual stereology laboratory is in itself an innovative concept and a timely one as the steadily declining cost of digital storage space has made the digitization of images by this automated system cost efficient. This project is truly the next logical step in application of modern computer technology to neuroanatomical study. In addition, the sheer magnitude of the volume of slides that will be handled and the number of through-focus series stacks that will be generated will demand novel solutions to the practical problems of how to store and access slides rapidly and how to provide the complex image data to clients worldwide using different platforms and different bandwidth connections.

Investigator: Dr. Williams is highly qualified to carry out this project. The PI is an accomplished neuroanatomist and stereologist who has also worked extensively in developing electronic circuitry and computer software for neuromorphometry. Furthermore, Dr. Williams has assembled an impressive team of researchers for Project 2, each with the expertise in a different aspect of the project. For example, Dr. Nissanov was the original developer of the automated slide-handling system that will be used for the iScope. Dr. Goldowitz has been instrumental in development of the high resolution through-put methods to analyze mouse CNS, and Dr. Park has extensive expertise in software programming especially as related to acquiring and processing image stacks.

Environment: The physical space set aside for the iScope laboratory at the University of Tennessee, Memphis, appears to be adequate, and institutional support for the project is both enthusiastic and financially supportive.

CRITIQUE 2:

Introduction: The investigators propose to continue to develop resources and techniques for high-magnification quantitative analysis of the Mouse Brain Library collection (MBL). There are four aims. Aim 1 is to develop four new video microscopes over a five-year period. Two of these will be optimized for uninterrupted and automatic acquisition of through-focus image stacks (optical sections). The video clips of these stacks will become part of the MBL collection (aim 4). The other two instruments will be inverted microscopes equipped with robotic slide handlers and will be used to allow outside investigators to roam through slide collections and, in conjunction with data in the MBL collection database and analytical tools, perform comparisons - referencing to the data in the archive and employing a variety of software tools for quantitative analyses. Aim 2 is to develop the robotic slide handler system to be installed in the microscopes commissioned conjunction with aim 1. Aim 3 is to develop an automated system to acquire large archives of through-focus images stacks. Aim 4 is to implement standard video streaming packages as part of this system.

Significance: A general rationale for use of slide collection capable robotic microscope systems is the very large amount of 3 dimensional data represented within a single microscope slide. There are similar efforts to provide access to slide collections in the area of digital clinical pathology and embryology, both in the US and abroad. Commercial systems for telepathology and telemicroscopy are becoming available (both Nikon and Leica market robotic microscopes). There are many technical challenges that need to be addressed in order to make the proposed capability useful. Many of these challenges are considered by the applicants. However, it is not clear that the work proposed will develop this capability as a robust and easily distributed system before other concurrent technologies make the approach proposed much less useful. During the next few years there will be significant advances in network capacity, software infrastructure for managing federated data and scaleable distributed-shared storage technologies. Developments in these areas may make direct access to real slide collections via robotic telemicroscopes somewhat less attractive.

The linkage between key questions in neuroscience and the technology being developed is reasonable. While there are strengths in the technology development competent - there are also some aspects of the work that are clearly not aligned with the leading wave of the technologies involved.

Approach: iScope and slide handling: The approach is mainly to use non-infinity corrected light microscope systems and bright field Nomarski optics to record optical section images with robotic microscopes. As noted above, two of these will be equipped with custom "jukebox" devices to change slides. The investigators will also have computer-controlled stages (x,y,z control) and focus tools to help produce image stacks etc. The robotic handling system is intended to deliver users to an area in a real slide that is within 200um of an area selected from the MBL database. The applicants target an improvement to

better than 50um by the end of the project. While the project is interesting and potentially useful the investigators are coupling the engineering to rather outdated optical systems and microscope imaging technologies. Microscopy methods proposed will likely be supplanted by more recently developed light optical systems employing laser-based specimen exploration methods such now commonly used in confocal and pulsed laser muliphoton based imaging systems.

Web archive of image stacks and software for stack analysis and Video Streaming: An image stack acquisition system will be developed which will acquire and automatically archive these stacks. The system will allow users to batch process image stacks (through focus series) and coordinate the registration and verification of the data in the database with the simultaneous remote computer control of the microscope. The system will also allow the user to "order" data from a piece of the collection, both the slide and repository of already digitized stacks. Lastly the system will provide stack analysis applications to users as Java applets and compiled C programs. These will provide for simultaneous viewing and analysis of multiple stacks. To support this and the direct use of the microscope, the investigators basically propose to keep their software development pace in sync with the rapid pace of development of video streaming technologies. These last two aims, batch processing and video streaming, are more adaptations of technologies more fully developed elsewhere.

Investigator: Dr. Williams is an Associate Professor of Anatomy and Neurobiology and is well trained in Psychobiology, Physiology and Developmental Neuroscience. The PI is an accomplished investigator and is a frequent contributor to the literature in applied mouse molecular genetics and the associated neuroanatomy. From the proposal it is clear that Dr. Williams has considerable talent in the area of technology development, as is evident from that which has been a component of his own research activity.

Environment: The Neurosciences Institute and Anatomy and Neurobiology Department at UT Memphis is a first rate environment. The space and resources committed are limited but sufficient.

RECOMMENDATION: Priority Score = 154

PROJECT 3

NeuroCartographer

PI: Dr. Jonathan Nissanov

CRITIQUE 1:

Significance: This project affords potential solutions to several very important problems in Neuroinformatics and Neuroscience. Facilities developed by this project will be made available to the scientific community for visualizing and quantifying a wide array of gross morphologic and cellular

features of mouse brains. In addition, the proposed efforts should lead to the development of useful and extensible solutions to neuroinformatics problems of aligning, matching, segmenting and quantifying 2D and 3D digital images of brain tissue. If the approach is successful, users will be empowered to find and quantify homologous brain segments in a variety of genetically distinct brains and to relate variation across brains with corresponding genetic variation of the mice. Large scale use of the facilities to be developed for QTL mapping of a wide array of brain structural characteristics has potential to greatly enhance our understanding of the genetic origins and the structural and functional relevance of many aspects of brain morphology. The tools provided will also be applicable, potentially, to many non-genetic analyses, such as for example, the assessment of neural plasticity via analyses of early stimulation, enrichment and learning effects on cellular and morphometric brain parameters.

Approach: Past accomplishments of the PI and colleagues and the planned work on Neurocartographer are impressive. The applicants have made excellent progress on construction of 2D and 3D atlases of the rat brain, which involved challenges similar to those posed by the plan of work. A description of one of their atlases has been submitted for publication. Thus the proposed work in many ways extends work already completed, for which this team has developed several useful tools and approaches. The description of the proposed work is exemplary in its expression of awareness of the possible pitfalls and alternative strategies. The plan to segment and align such a large number of brains is extremely ambitious.

Although the proposal is enormously detailed and thorough, it could have been improved by expansion of certain topics. Thus, it is difficult to get a sense of the correspondence between the precision of morphometric quantification that will be afforded by the tools to be developed and the requirements of quantification that are demanded by QTL analysis. How does this correspondence vary with respect to the size of the brain compartment being examined and what is the size threshold at which the tools will become ineffective? It would have been useful had the applicants provided a description of a realistic sample problem, wherein reasonable functionality assumptions concerning as yet undeveloped tools were applied to QTL mapping of a particular morphometric trait, and expected results were estimated. An additional and possibly troublesome issue has to do with the impact of the 50%-60% shrinkage that routinely occurs as a result of celluloidin embedding of brain material. Little is said as to the impact of this shrinkage on the accuracy of VOI quantification.

Innovation: The research team for Project 3 has imported or developed tools that have the potential to greatly facilitate the proposed development of large numbers of brain atlases and tools for searching the MBL database for varied instances of particular neuroanatomical compartments to be analyzed. Examples include the program AIR, to be used for automatic section-to-section alignment in the atlas construction phase of the research. The research team is adept in developing creative approaches to problems that arise, such as the proposed supplementation of the AIR program with the use of fiducial points and correlation peaks to drive nonlinear transformation in the section-to-section alignment process during atlas construction. Other examples include the program Brain, that will be used to delineate the atlases in three dimensions, the 3DBA

algorithm for cross-atlas alignment, and the supplementation of this process by spline transformation and fiducial alignment as developed by Drs. Gabani and Tretiak. There are many additional examples, indicating that the research team is at the cutting edge of technological development with respect to these issues. Thus, although the proposed task is huge, the experience, knowledge and creativity of the research team and their large arsenal of extant tools are likely to yield important progress toward implementing Neurocartographer.

Investigator: After completing an NIH postdoctoral fellowship in Neurobiology at UC San Diego, Dr. Jonathan Nissanov has held several positions at Drexel University, Philadelphia. The PI was the Associate Director of the Computer Vision Center for Vertebrate Brain Mapping, an NIH Biomedical Technology Resource at Drexel University and he currently holds the position of Research Professor in the School of Biomedical Engineering. Dr. Nissanov has supervised atlas construction and software development for release to the research community, as well as a large number of collaborative projects. Dr. Jonathan Nissanov has conducted extensive research on image analysis and segmentation tools that are directly relevant to the aims of this project. The PI and his colleagues at Drexel have designed and implemented a robotic slide handling system, that will make it possible to automate many aspects of the MBL and the Internet Microscopy project. Dr. Nissanov will work closely with Dr. Tretiak, who is an expert in computer vision and the alignment and segmentation of brain material.

Environment: The work of the investigators will be centered in the Drexel University Computer Vision Center for Vertebrate Brain Mapping. This laboratory affords an extensive collection of state-of-the-art computer facilities, software and histological facilities expressly assembled to enable construction and manipulation of 2D and 3D brain image sets.

CRITIQUE 2:

Significance: If successful, this project promises to build a set of tools that will allow the access and quantification of a very large and important data resource by the community. The data itself is very valuable, but these tools will enhance the usefulness of the information and provide a framework to generate new knowledge by comparisons across data sets.

Approach: Aim 1: Construction of mouse 3D atlases and navigation software. The project builds on the previous successful efforts of the investigators. The construction of the atlases from the parental strains of mice is a critical aspect of the entire program project application. The PI has laid out a clear set of objectives for creating the atlases and incorporated mechanisms to circumvent possible pitfalls. A strength of the approach is that the new software application for construction and navigation of the atlases will be developed based on an existing prototype MacOStat. The extensions to the prototype to build Neuroterrain will endow it with important new functionality, by improving the ability to navigate through the atlases and to compare equivalent sections from different brains. The applicants' description of the software design is very brief and limited to invoking JAVA as the software environment. There is little discussion of possible pitfalls of using JAVA as a 3D visualization environment (i.e. it may be slow on some machines and slower over the net) or whether other existing platforms, NeuroZoom or ImageJ might be

more useful alternatives for the community. Also, there is no mention or critique of the efforts of Dr. Larry Swanson and his colleagues in the development of NeuARt, a rat brain atlas with manually segmented brain regions that stores connectivity and labeling data.

Aim 2: Affine mapping into standard coordinates: A strength of this approach is the use of a well accepted method for affine registration, AIR, developed by other investigators in the Human Brain Project. The PI is well acquainted with most of the relevant literature and methodology for both linear and non linear transformation algorithms. The PI acknowledges the possible limitations of this approach and proposes adequate solutions, albeit extremely time consuming ones.

Aim 3: Segmentation of the MBL and extraction of quantitative traits: This is by far the most difficult and problematic portion of the proposal. The goal of segmenting all of the brains into many different anatomical regions is important and the usefulness of the data depends on the accuracy of the segmentation. The PI has written a careful analysis of possible solutions to this difficult problem, however since it will take 3 years to develop and test the algorithm, this aspect of the project is somewhat of a risk. The plan for testing the accuracy of the methods is well thought out and insures a level of confidence in the data. The section of the proposal describing quantitative analysis routines is quite brief and somewhat speculative. This is surprising since the delivery of "numbers" to the neurogenetics tool box was described as a primary goal of the project. There is a concern that users of the quantitative tools may not understand the mathematical basis of the morphological comparisons. The usefulness of the quantitative measures depends not only on the validity of the mathematical approach but also how well the users understand what is being measured. Changes in volume are an obvious measure, however a more thorough description of the types of comparisons and statistical measures to be implemented would strengthen the proposal. It appears that these issues are out of the scope of the current proposal. The plan to allow users of the database to define their own regions or volumes of interest is somewhat problematic. How will the applicants ensure the quality of the segmentation performed by the user? What standards will be put in place for user based data entry? Making the data model plastic is an ambitious goal which could degrade the quality of the information without some oversight.

Innovation: Most of the approaches are adaptations of existing methods, or build on previous work of the investigators. This work follows on the efforts of the BrainMap application and is a logical and useful extension of those ideas.

Investigator: The PI is well qualified to do this work. Dr. Nissanov is the associate director of the Computer Vision Center for Vertebrate Brain Mapping an NIH funded Biotechnology Resource. Dr. Nissanov has extensive experience in the design and implementation of software for the neuroscience community and an established record for documentation and support of the software. Dr. Nissanov has previous experience with creating 2D atlases from brain sections and has published several papers in this area.

Environment: The environment for this work is excellent. The facilities are well equipped to handle the volume and the accuracy required to generate the data. All members of the group have expertise in the required research areas.

Computational facilities are very good.

Summary: This proposal is an ambitious application to develop visualization tools for segmentation and quantification of anatomical data from mouse brain atlases. The major strengths of the application are the detailed experimental plan for construction of the atlases, the registration process and segmentation of neuroanatomical features from sections. This project will allow users of the Mouse Brain Library access to a huge data set aligned to a common coordinate system that can be queried and the results quantified using a set of measurement tools. Users will have web access to this data resource and the analysis tools. This component of the overall project is very well integrated with the other efforts. Weaknesses of the work are the inherent difficulty in achieving accurate spatial transformations of 2D data and converting 2D into a 3D format. The value of the atlas and the analysis tools depend critically on the accuracy of the alignment and segmentation procedures. The description of the analysis tools for quantifying morphological attributes was very brief and somewhat vague. The PI has a great deal of experience in the experimental techniques, a firm grasp of the algorithms for segmentation and transformation and experience in producing software applications for release in the public domain.

RECOMMENDATION: Priority Score = 154

PROJECT 4

Neurogenetics Tool Box

PI: Dr. Kenneth Manly

CRITIQUE 1:

The goal of this project is to provide a set of algorithms for the analysis of quantitative trait loci in any of the MBL collection of mouse RI lines or G10 intercrosses. Described as a 'toolbox' the project will develop four components over the proposed five years of the project.

First, data that is submitted by an investigator will be subjected to biostatistical evaluation. One goal of this component is to provide the user with information on the suitability of the data for the analysis packages offered. In addition, the data will be analyzed for key descriptive statistics, correlation matrices between submitted traits and others in the data base (or submitted), estimates of heritability and of the number of genes likely to be involved in controlling the trait.

Second, and most complex, is a gene- and QTL-mapping module. This will initially be based on existing algorithms (Map Manager QT and QTX) and refined later. All coding will be done in C++ for computation and Python for web interaction. The goal is to make accessible an on-line tool for investigators to map either single loci or quantitative trait loci with user friendly input.

The third module will allow clients to permanently archive their results to enrich the database and increase its power. The fourth tool will be an extensive and detailed user manual.

This component of the program has significant strengths. The choice of mapping algorithms is solid. The PI is a very experienced investigator.

The fact that the program will come 'pre-filled' with previous phenotypic data is another strength. This will allow correlations to be uncovered that were not in the original experimental plan of the client. The free and open access to both the programs and the code is a significant strength as well. Also, it is clear that the project leader had put considerable thought into documentation and on-line help.

There are underlying weaknesses, however. These do little to dampen overall enthusiasm for the project but should be attended to as implementation moves forward.

Despite attention and significant effort, the project would benefit from a specific program to promote its use. Among geneticists there needs to be a plan to promote the value of morphological analysis. Among anatomists it would benefit to have a plan to promote genetics.

The program is well designed to take a novice through the steps of getting a QTL plot, but there is little description of how the program would aid the user after that. In the final analysis, the novice will have to publish the results and it would be advantageous to have a plan to assist the user in the analysis of his/her data and its interpretation.

Specifics are missing in the way in which submitted data is treated. Many users will want to keep their hard earned measurements private, but the community benefits from the results being as public as possible. There are many plausible solutions to this dilemma, but none are spelled out here.

It would bolster the applicants' case significantly if a 'preliminary data' section were included that illustrated a typical run from data collection to analysis. Too much is assumed about the user's familiarity with the specific protocols of the system.

A specific example of this and an issue in its own right is some set of comments about the swiftness or sluggishness of the web-based format. Would the tool box be just as useful with submitted information (the throughput will not likely be too high to enable this).

CRITIQUE 2:

Introduction: The Neurogenetics Toolbox, a genetic database of marker genotypes and map distances for crosses and strains in the MBL, fulfills an important Phase II aim of the Neuroinformatic directive in allowing community access to a large, valuable database. However, this project goes well beyond other Web-based databases in providing services that normally would require the establishment of collaborative arrangements, as for example users may submit

potential quantitative phenotypes for statistical evaluation and quantitative trait loci (QTL) mapping. In addition, phenotypes generated by Project 3 will automatically be available for analysis to supplement the phenotypes generated by outside investigators. Furthermore, the rapid advancement of QTL mapping will most certainly lead to insight into the genetic basis of brain diseases involving quantitative alterations such as schizophrenia, depression, and dementia.

Significance: This is a model prototype of Web-based sharing of scientific expertise as it will allow the scientist with limited material resources and genetic background to use the genetic database as a research tool much as the personal computer is used by scientists with little in depth knowledge of programming. Thus, the Neurogenetic Toolbox will allow for complete access to mouse phenotypes and genotypes.

Approach: This ambitious plan to make powerful statistical and gene mapping utilities available via the Web is well conceived and carefully planned. The algorithms employed in this project are direct derivatives of Map Manager software which was developed by Dr. Manly, and therefore are readily adaptable to the functions of the present proposal. Furthermore, both Drs. Manly and Williams have experience in software development and therefore have planned for extensive testing, feedback, and modification in implementation of the Neurogenetics Toolbox.

Innovation: The Neurogenetics Toolbox is distinguished from other databases by virtue of its links with morphometric data generated via Projects 1 and 3. Rather than just accessing legacy data, investigators will be able to conduct original research projects via the Net and therefore the Neurogenetics Toolbox will quickly become an invaluable resource for the neuroscience community. The addition of an advanced hybrid cross, G10, will allow for a fine-grained mapping of QTL enabling the researcher to distinguish between several closely linked genes that may influence quantitative traits.

Investigator:: Dr. Manly possesses a high level of expertise in computer programming related to gene mapping applications. He has already written a series of programs in the Map Manager series that have been widely adopted for use by the genetics community. The Map Manager's user manual is easy to read and follow even for the novice. Dr. Williams is an established mouse geneticist. Dr. Williams also has experience in computer programming and is able to translate his knowledge into practical applications that benefit the research community.

Environment: The facilities at the University of Tennessee, Memphis, are entirely adequate for state of the art genotyping. The PI has the support of both the Director of the Neuroscience Institute and the Associate Dean of Research. There is also ample computer space and network support.

CRITIQUE 3:

Significance: QTL analysis of brain and behavioral phenotypes has increased over the past decade. There are a number of methods to calculate measures of association, p-values and LOD scores, some more useful than others. The PI proposes to develop an analytical center, available to QTL researchers, that is

easy to use and provides state or the art analysis. This will be a great help for many researchers, especially those who have expertise in brain and behavioral phenotypes but less in quantitative genetics.

Approach: The proposal is well conceived and is based on conventional procedures for QTL analysis. It is flexible to address different genetic models. Not only will clients be able to conduct QTL analysis, but genetic correlational analysis among phenotypes will also be possible. Private data sets will be treated confidentially. For many researchers, this service will be very useful. This component fits well with the other components and could also stand-alone.

Innovation: The concept of a central clearinghouse for genotypic and phenotypic analysis is novel and complemented by on-line tutorials, access to extant data and a menu of analytical techniques. There are a number of features available for the analysis of complex traits, including a means to analyze epistasis. This is a very important feature.

Investigator: The PI is known as one of the early researchers in application of QTL analysis to the mouse. Dr. John Belknap, one of the leading experts worldwide in QTL analysis, will ably assist him.

Environment: Much of the service proposed by the PI will take place on the web and as such, computing equipment becomes a major concern. Computing facilities at UT Memphis and at Roswell Park seem to be adequate and the investigators have addressed data and file transfer protocols. What is not known is the kind of traffic to expect and whether the systems will be able to keep up with demand.

Strengths:

This project holds promise to make QTL and genetic correlational analysis accessible and available to a wide contingent of researchers.

Researchers will have a large database of genotypic and phenotypic data available to help guide their own investigations.

A number of analytical techniques will be available and the PI has indicated an effort to provide instruction in basic and comparative analyses.

Concerns:

There are two concerns. First, although the PI states that he and Dr. Williams will maintain the database, there is no indication of how this will be done, means for testing data accuracy etc. Both are active research scientists. Do the applicants really mean that they will do this large task, or are there plans to have other individuals help?

The other concern is minor. Have the investigators thought of how to inform new investigators about some of the real pitfalls as well as the strengths of QTL analysis?

RECOMMENDATION: Priority Score = 165

CORE

A. Neurohistology. B. Genotyping

PIs: Dr. Glenn Rosen, Dr. Robert Williams

DESCRIPTION (adapted from the applicant's introduction):

Core A provides histology for the MBL, i.e., collection, registration and processing of all neuroanatomic materials generated by the component projects. The facility is housed at the Beth Israel Deaconess Medical Center of Harvard Medical School. The personnel will be one professional (Dr. Rosen), one full-time and one half-time histological technician. The full-time technician is currently active having already processed 600 brains for the MBL . Dr. Rosen will oversee quality control of the procedures.

Core B will provide personnel and facilities for the extraction of genetic data of a large collection of mouse strains. The data will supply the Neurogenetics Tool Box (Project 4) component of the program project. This Core will also oversee the project mouse colony and provide fixed brains to the Neurohistology Core for further processing.

CRITIQUE 1:

Significance: The Core A facility is needed to carry out processing of the large number of brains that will be added to the MBL. The facility avoids duplication of facilities and personnel for tissue preparation efforts that would otherwise be required for other component projects.

Approach: The procedures of celluloidin embedding of tissue, sectioning, staining and immunohistochemistry appear to be state-of-the-art. A variety of stains and antibody reactions can be employed, although Nissl staining and AChT/Parvoalbumin immunoreactivity are the only assays planned a-priori for all tissue. Key personnel of Core A have accumulated ample experience and demonstrable competence with respect to the proposed procedures.

Investigator: Dr. Glenn Rosen received his PhD in Biobehavioral Sciences and Developmental Psychobiology at the University of Connecticut in 1982, and is currently an associate Professor in Neurology at Beth Israel Deaconess Medical Center and Harvard Medical School. Dr. Rosen received the National Dyslexia research Foundation Distinguished Young Investigator Award, and has many publications in neuroscience journals of high-quality concerning functional neuroanatomy and neurochemistry.

Environment: The histology section of Beth Israel Deaconess Medical Center has served well as the venue for the initial processing of brains for this project.

CRITIQUE 2:

Core A will provide histological service to analyze the genetics of morphology, neurochemistry (by immunohistochemistry, etc) and cytology in service to the overall Program. This is the major phenotyping service.

What is called for and what is to be provided, according to the proposal is a standard, up to date histological and imaging service. The PI is well qualified and technical staff appear to be competent to carry out the technical aspects of this work.

Coordination between the laboratory in Memphis and the histology lab in Boston seems to be appropriate. What is not presented in detail is the expected throughput and estimated times for processing.

CRITIQUE 3:

Significance: The extracted genetic information that Core B will supply to the Neurogenetic Toolbox Component of the project in combination with morphometric data obtained using tools afforded by the Neurocartographer component will be used by researchers to map the genes that determine and modulate various structural features of the brain. A substantial amount of this work has already been completed. In addition, Core B will extract and fix mouse brains for processing by Core A (Neurohistology). The use of a tenth-generation intercross by Core B scientists yields a large number of different morphometric traits that can be studied by neuroscientists and it increases the precision of QTL mapping, relative to results that could be obtained with conventional intercross or backcross procedures. Also, access to many genotypes and phenotypes simultaneously will enhance the ability of single investigators to uncover pleiotropy (genetically based morphological correlations in various parts of the brain). These features can substantially enhance the efficiency and statistical power of gene mapping. Full scale genotyping and extension of the MBL as proposed will allow multiple research groups to study many traits in the same set of animals, maximizing the utility of each cross, preventing duplication of effort and revealing the genetic basis of CNS variation. Committed institutional matching financial support and planned new building facilities are noteworthy.

Approach: The tenth generation advanced intercross carried out by Core B researchers will yield genetic differences at more than 3000 genes and marker loci. The typing of 1000 tenth generation progeny with each of 350-400 markers will result in 350-400K PCR reactions each of which will be entered into a database and exported to the Neurogenetics Tool Box. Thus large array of genetic information will be made available in support of QTL mapping of many brain morphological characteristics.

Innovation: The application of the tenth generation advanced intercross is an innovative aspect of Core B.

Investigator: Dr. Williams, PI of the program project and director of Core B received his PhD in Physiology and did postdoctoral work in Developmental Neuroscience at Yale University and is currently a professor in Anatomy and Neurobiology, Center for Neuroscience, at the University of Tennessee. Dr. Williams has published abundantly in sources of the highest quality concerning

genetical modulation and QTL mapping of parameters of the mammalian visual system.

Environment: The work of Core B will be carried out initially in Dr. Williams' laboratory at the University of Tennessee, Memphis. It is not clear whether sufficient facilities and space will be available given other demands on this resource. A new UT Memphis Genotyping Center is scheduled for opening late in 2001 and will be directed by Dr. Williams. This facility will have a capacity of 2000 genotypes per year of which half will be devoted to this program project.

CRITIQUE 4:

Core B is a critical core component for genotyping and production of advanced intercross BXD mice.

The advanced intercross method is recognized as an effective way to pinpoint more precisely chromosomal areas containing QTL or candidate genes that affect quantitative phenotypes. The PI maintains a colony of these and information about these animals is already available on the Internet.

The PI's laboratory is capable of a reasonable throughput of genotyping and plans are underway to try to double the production. The methods are state of the art and should serve the needs of the entire project.

One minor concern is that most of this project is centered on identifying polymorphisms in the genome relevant to neurobehavioral characteristics. Are there any plans by the group to expand their efforts to more functional genomics like gene expression?

The University of Tennessee, Memphis will provide $140K for equipment to support the genotyping requirements for the full program period.

RECOMMENDATION: Priority Score = 157

ANIMAL SUBJECTS: No comments or concerns.

BUDGET: The budget seems appropriate.

NOTICE: The NIH has modified its policy regarding the receipt of amended applications. Detailed information can be found by accessing the following URL address: http://grants.nih.gov/grants/policy/amendedapps.htm

NIH announced implementation of Modular Research Grants in the December 18, 1998 issue of the NIH Guide to Grants and Contracts. The main feature of this concept is that grant applications (R01, R03, R21, R15) will request direct costs in $25,000 modules, without budget detail for individual categories.

Further information can be obtained from the Modular Grants Web site at

http://grants.nih.gov/grants/funding/modular/modular.htm

MEETING ROSTER
NATIONAL INSTITUTE OF MENTAL HEALTH
ZMH1 BRB-P (01)
March 22 - 24, 2000

CHAIRPERSON
BONNIE L. WEBBER, PHD
PROFESSOR OF INTELLIGENT SYSTEMS
DIVISION OF INFORMATICS
UNIVERSITY OF EDINBURGH
EDINBURGH EHI I-HN
UNITED KINGDOM

MARK H. ELLISMAN, PHD
PROFESSOR AND DIRECTOR
CENTER FOR RESEARCH ON
BIOLOGICAL STRUCTURE
UNIVERSITY OF CALIFORNIA, SAN DIEGO
LA JOLLA, CA 92093-0608

THOMAS D ALBRIGHT, PHD
PROFESSOR
VISION CENTER LABORATORY
THE SALK INSTITUTE
LA JOLLA, CA 92037-1099

MIKE FOLK, PHD
COMPUTER SCIENTIST
NATIONAL CENTER FOR
SUPERCOMPUTING APPLICATION
UNIVERSITY OF ILLINOIS AT URBANA
CHAMPAIGN, IL 61820

HELENE BENVENISTE, PHD
ASSOCIATE PROFESSOR
DEPARTMENT OF ANESTHESIOLOGY
DUKE UNIVERSITY MEDICAL CENTER
DURHAM, NC 27710

MICHAEL GABRIEL, PHD
PROFESSOR
DEPARTMENT OF PSYCHOLOGY
BECKMAN INSTITUTE
UNIVERSITY OF ILLINOIS
URBANA, IL 61801

JAMES F. BRINKLEY, PHD
ASSOCIATE PROFESSOR
DEPARTMENT OF BIOLOGICAL STRUCTURE
SCHOOL OF MEDICINE
UNIVERSITY OF WASHINGTON
SEATTLE, WA 98195

KARL F HERRUP, PHD
PROFESSOR
ALZHEIMERS RESEARCH LABORATORY
CASE WESTERN RESERVE UNIVERSITY
CLEVELAND, OH 44106

JOHN H. BYRNE, PHD
PROFESSOR AND CHAIR
DEPT OF NEUROBIOLOGY AND ANATOMY
UNIV. OF TEXAS MEDICAL SCHOOL
HOUSTON, TX 77030

EDWARD H HERSKOVITS, PHD
ASSISTANT PROFESSOR
DEPARTMENT OF RADIOLOGY
JOHNS HOPKINS SCHOOL OF MEDICINE
BALTIMORE, MD 21287

LAWRENCE HUNTER, PHD
COMPUTER SCIENTIST
NATIONAL CANCER INSTITUTE
NATIONAL INSTITUTES OF HEALTH
BETHESDA, MD 20892

NICHOLAS T. LANGE, PHD
ASSOCIATE PROFESSOR
DEPARTMENT OF PSYCHIATRY
MCLEAN HOSPITAL
HARVARD MEDICAL SCHOOL
BELMONT, MA 02178

STEWART H. HENDRY, PHD
ASSOCIATE PROFESSOR
DEPARTMENT OF NEUROSCIENCE
JOHN HOPKINS UNIVERSITY
BALTIMORE, MD 21218

RICHARD M LEAHY, PHD
PROFESSOR
SIGNAL AND IMAGE PROCESSING INSTITUTE
UNIVERSITY OF SOUTHERN CALIFORNIA
LOS ANGELES, CA 90089

GWEN A. JACOBS, PHD
ASSOCIATE PROFESSOR AND CO-DIRECTOR
DEPARTMENT OF LETTERS & SCIENCE
CENTER FOR COMPUTATIONAL BIOLOGY
MONTANA STATE UNIVERSITY
BOZEMAN, MT 59717

BERTHA K MADRAS, PHD
PROFESSOR
DEPARTMENT OF PSYCHIATRY
NEW ENGLAND REGIONAL PRIMATE RES CTR
HARVARD MEDICAL SCHOOL
SOUTHBOROUGH, MA 01772-9102

TERRY L. JERNIGAN, PHD
PROFESSOR
BRAIN IMAGE ANAYSIS LABORATORY
UNIVERSITY OF CALIFORNIA, SAN DIEGO
LA JOLLA, CA 92093-0949

ALLEN D MALONY, PHD
ASSOCIATE PROFESSOR
DEPARTMENT OF COMPUTER AND INFORMATION
UNIVERSITY OF OREGON
EUGENE, OR 97403

STEVEN KORNGUTH, PHD
PROFESSOR OF NEUROLOGY
INSTITUTE ADVANCED TECHNOLOGY
UNIVERSITY OF TEXAS
AUSTIN, TX 78759-5329

BENJAMIN REESE, PHD
PROFESSOR
NEUROSCIENCE RESEARCH INSTITUTE
UNIVERSITY OF CALIFORNIA
SANTA BARBARA, CA 93106-5060

LYNN D SELEMON, PHD
RESEARCH SCIENTIST
DEPARTMENT OF NEUROBIOLOGY
YALE UNIVERSITY
NEW HAVEN, CT 06510

KATHLEEN A. WELSH-BOHMER, PHD
ASSOCIATE PROFESSOR
BRYAN ALZHEIMER'S CENTER
DUKE UNIVERSITY MEDICAL CENTER
DURHAM, NC 27705

FREDERIC L WIGHTMAN, PHD
PROFESSOR AND DIRECTOR
DEPARTMENT OF PSYCHOLOGY
UNIVERSITY OF WISCONSIN
MADISON, WI 53705

SCIENTIFIC REVIEW ADMINISTRATOR
HOUMAM H ARAJ, PHD
SCIENTIFIC REVIEW ADMINISTRATOR
DIVISION OF EXTRAMURAL ACTIVITIES
NATIONAL INSTITUTE OF MENTAL HEALTH
NATIONAL INSTITUTES OF HEALTH
BETHESDA, MD 20892-9608

Consultants are required to absent themselves from the room during the review of any application if their presence would constitute or appear to constitute a conflict of interest.