

A Three-dimensional MRI Atlas of the Mouse Brain with Estimates of the Average and Variability

Although there is growing interest in finding mouse models of human disease, no technique for quickly and quantitatively determining anatomical mutants currently exists. Magnetic resonance imaging (MRI) is ideally suited to probe fine structures in mice. This technology is three-dimensional, non-destructive and rapid compared to histopathology; hence MRI scientists have been able to create detailed three-dimensional images of 60 μm resolution or better. The data is digital which lends itself to sophisticated image processing algorithms. Here we show a variational MRI atlas constructed from nine excised brains of 8 week old 129S1/SvImJ male mice. This new type of atlas is comprised of an unbiased average brain — created from alignment of the individual brains — and the mathematical descriptors of anatomical variation across the individuals. We found that the majority of internal points in the individuals never varied more than 117 μm from equivalent points in the atlas. A three-dimensional annotation of the average image was performed and used to estimate the mean and standard deviation of volumes in a variety of structures across the individual brains; these volumes never differed by more than 5%. Our results indicate that variational atlases of inbred strains represent a well-defined basis against which mutant outliers can be readily compared.

Keywords: central nervous system, image processing, magnetic resonance imaging, phenotyping

Introduction

The sequencing of the entire human and mouse genomes has allowed scientists to examine for the first time the complete set of genes that regulate the development of mammalian organisms. The relatively close evolutionary relationship between humans and mice has prompted the use of this species as a model system for investigation of mammalian organ development. To this end, both targeted and genome-wide mutagenesis projects have been undertaken at a number of major research centres throughout the world. Capitalizing upon this sizable catalog of genetic modifications will require screening mutants through functional or phenotypic assays (Nolan *et al.*, 2000). In particular, there is a widely recognized need for anatomical phenotypic screening procedures using various types of imaging technologies (Balaban and Hampshire, 2001) (see also <http://www.nbirn.net/TestBeds/Mouse/index.htm>).

Magnetic resonance imaging (MRI) has shown a substantial potential in this regard because of its ability to capture large amounts of anatomical information in a nondestructive manner. While the spatial resolution of this technique — ranging from 20 to 60 μm for fixed samples (Benveniste *et al.*, 2000; Johnson *et al.*, 2002) — is not comparable to histology (~2–5 μm), detailed

N. Kovačević¹, J.T. Henderson², E. Chan², N. Lifshitz¹, J. Bishop¹, A.C. Evans³, R.M. Henkelman^{1,4} and X.J. Chen^{1,4}

¹Mouse Imaging Centre, Hospital for Sick Children, 555 University Avenue, Toronto, Ontario, Canada M5G 1X8, ²Department of Pharmaceutical Sciences, Leslie Dan Faculty of Pharmacy, University of Toronto, Rm 315, 19 Russell Street, Toronto, Ontario, Canada M5S 2S2, ³McConnell Brain Imaging Centre, Montreal Neurological Institute, McGill University, 3801 University Street, Montreal, Quebec, Canada H3A 2B4 and ⁴Department of Medical Biophysics, University of Toronto, Toronto, Ontario, Canada

anatomical comparisons within murine systems can be made, such as in the central nervous system (CNS). In addition, this method allows the three-dimensional morphology of anatomical structures to be easily examined, thus providing significant advantages over serial histological sectioning techniques in analyzing complex structures (Dhenain *et al.*, 2001; Johnson *et al.*, 1993, 1997). Several groups have made progress in the creation of MRI atlases of individual mouse brains (e.g. <http://www.loni.ucla.edu/MAP>; Van Essen, 2002). Like conventional histological atlases (http://www.mbl.org/mbl_main/atlas.html, <http://www.hms.harvard.edu/research/brain/atlas.html>), these MRI-based atlases consist of high-resolution images of an individual specimen with anatomical labels. A crucial difference, however, is that both the data and the anatomical labels are three-dimensional and digital.

A critical step beyond obtaining the MR images is to analyze and compare the enormous amounts of data in an efficient and reliable manner. Current phenotypic assays require an average and standard deviation of the normal range of any measurement to allow identification of outliers. Likewise, we need equivalent analysis for three-dimensional images. One-dimensional metrics, such as heart rate, body weight or red blood cell count, can be averaged in a straightforward manner; however, the analogue for three-dimensional images requires more sophistication.

We propose to take the concept of an MRI atlas further by creating a *variational atlas*. The idea is to combine multiple three-dimensional images together to provide a representation of average anatomy and a range of anatomical variation within a particular population. Specifically, in this paper we register a number of genetically identical murine brains, which allows us to estimate the limits of natural variation in terms of anatomical structures, volumes, shapes and locations.

The fusing together of a set of individual images into a single image is the first component of our variational atlas, the *average image*. This average image extracts commonalities among individual brain anatomies and filters out idiosyncrasies. By manually delineating structures in the average image, we produce an annotated atlas. The second component of the variational atlas consists of a set of *deformation fields*, which is a measure of the anatomical variability across the set of individual brain images. This is the 'variational' aspect of the atlas. Here, we capture and quantify precisely the differences that are removed in the process of creating the average image.

Materials and Methods

Specimen Preparation

Eight-week-old male inbred 129S1/SvImJ mice (Jackson Laboratory, Bar Harbor, ME) were acclimated for a period of 3 days. Animals were then anesthetized with an overdose of Avertin (2.5%) via intraperitoneal

injection. Following lack of deep tendon responses, the thoracic cavity was opened and animals perfused through the left ventricle with 10 cm³ of 0.1 M phosphate-buffered (pH 7.4) 0.9% NaCl (PBS), followed by 4% formaldehyde in PBS. Solutions were infused at room temperature (25°C). Following perfusion, the heads were removed and allowed to postfix at room temperature for an additional 60 min, at which time the cranium was opened and the brain removed in its entirety. Brains were then postfixed for an additional 24 h in 4% paraformaldehyde in PBS at room temperature.

Magnetic Resonance Imaging

A 7.0 T, 40 cm bore magnet (Magnex Scientific, Oxford, UK) connected to a Unity^{INOVA} console (Varian Instruments, Palo Alto, CA) with modified electronics for parallel imaging was used to acquire anatomical images of excised brains. Prior to imaging, the brains were removed from the fixative and placed into glass tubes filled with a proton-free susceptibility-matching fluid (Fluorinert FC-77, 3M Corp., St Paul, MN). We used two custom-built, 12 mm, non-uniform solenoid coils (Idzaiak and Haebleren, 1982) to image two brains in parallel. The parameters used in the brain scans were as follows: *T*₂-weighted, three-dimensional spin-echo sequence, with *T*_R/*T*_E = 1600/35 ms, single average, field-of-view = 12 × 12 × 24 mm and matrix size = 200 × 200 × 400 giving an image with (60 μm)³ isotropic voxels. The total imaging time was 18.5 h. The *T*_R and *T*_E settings were chosen for optimized contrast between grey matter and white matter in the mouse brain at 7 T as reported in previous studies (Guilfoyle *et al.*, 2003).

Algorithm for the Variational Atlas (Further Details in Appendix)

We used image registration to model the anatomical differences among the three-dimensional data sets. In image registration, two images are compared by performing a series of deformations in order to make one image identical to the other. The result of the process is stored in a deformation field, a vector field which records the magnitude and direction required to deform a point in the source image to come to the appropriate point in the target. In other words, the anatomical differences between the two images are encoded in the deformation field.

The algorithm takes a set of MR images acquired from a number (in this case nine) of excised brains and runs through six iterative steps. In the first step the images are registered and normalized in terms of global size, shape and MR intensity; we call these the 'globally normalized images'. At this point only extrinsic, anatomically insignificant differences are removed. The remaining steps involve detailed matching of anatomical features starting from a coarse grid and ending at the resolution of the imaging voxels.

The final output consists of an average image and the deformation fields. The deformation fields encode local positional differences between the average image and the globally normalized images. The deformation fields represent a large body of data: number of images × number of voxels × three vector components, which is ~1 GB of data. Therefore, to represent more simply and graphically the variability of the atlas, we calculate the 'mean positional distance' (MPD) image between the average and the individuals.

We ran our implementation in parallel (a single thread corresponding to each dataset) on a 192-processor Origin 3000 supercomputer (Silicon Graphics, Inc., Mountain View, CA). We used nine processors (600 MHz each) to create the average image, which took ~15 h.

Annotation

Structures that are clearly visible in the average image have been annotated by manual segmentation using the software package Display (Montreal Neurological Institute, Montreal, Canada). Every voxel of the average image has been assigned to a unique anatomical label, and the delineations were verified in three orthogonal directions. We delineated structures that were visible on MRI and nomenclature was adapted from the literature (Franklin and Paxinos, 1997). The resulting annotation is stored in the segmentation image, a three-dimensional image with the same size and coordinate system as the average image. Each voxel position of the segmentation image stores the anatomical label belonging to the corresponding voxel coordinate of the average image.

Results

The Average Image

In Figure 1, we show the average image beside an individual brain (see Supplementary Material A for a full three-dimensional comparison). Because the average image is calculated non-linearly and the noise is not well defined, we calculated the signal-to-noise ratio (SNR) to be the mean signal in a region of interest (ROI) in a homogeneous part of the brain divided by the standard deviation of the same ROI. On average, the individual brains had SNRs of 20 and the average brain of 50. We see that the contrast between grey and white matter is better than in the individual images, e.g. as seen in white matter structures like the corpus callosum, fimbria of the hippocampus and the anterior commissure. The average image on the whole exhibits greater definition than the individual images. In particular, the visibility and delineation of many anatomical structures is, on average, considerably improved (e.g. deep cerebellar structures, the globus pallidus and the lateral parabrachial nucleus). If one assumes that each of the individual specimens were identical, we would expect to see this effect as a result of averaging of repeated measurements. Thus, the improvement in the average image over that of individual specimens indicates successful alignment of all samples; clearly anatomical structures across the specimens have been 'stacked up'; otherwise averaging would have produced a blurred image (see Appendix, Fig. A1). Some fine features, however, show the reverse trend of a reduction in their level of definition within the average image. For example, white matter tracks in the striatum and small blood vessels. The extreme variability of these regions cannot be resolved by our registration with topology-preserving transformations. This qualitative comparison between the average and individual images is not a proper assessment of the accuracy

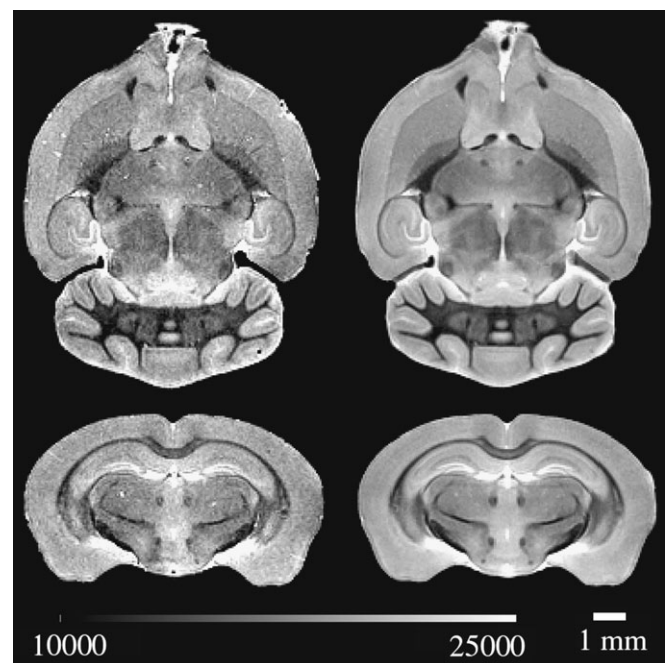


Figure 1. Images from an individual brain (left) in comparison with the average atlas (right) shown with the same contrast. A horizontal (top) and coronal (bottom) slice are shown from the full three-dimensional datasets. The individual image exhibits artefacts due to imaging (e.g. noise and residual fixative) and specimen handling.

of our registration; quantitative validation is the subject of further study.

Stereotaxic Coordinate System

Based on the average image, a three-dimensional stereotaxic coordinate system was defined (Fig. 2). The midsagittal plane that evenly divides left and right components of the brain along the anatomic midline is selected as the coordinate plane $x = 0$. Positive x -coordinates increase in the right lateral direction, while negative x -coordinates decrease in the left lateral direction. The brain was tilted so that horizontal and vertical directions are approximately the same as in the histological atlas of Franklin and Paxinos (1997). The coordinate origin was arbitrarily selected within the midsagittal plane of the cerebrum as shown in Figure 2.

Annotation

The annotated structures can be explored in several ways. For example, slices from the average image can be taken at arbitrary angles and viewed overlaid with corresponding slices of the segmentation image (Fig. 3A). In particular, this type of visualization enables slicing in coronal, sagittal or transverse directions for comparison with classical two-dimensional histological atlases. More advanced visualization in three dimensions is obtained from surface renderings of individual structures. Using interactive visualization tools (e.g. Amira, TGS, San Diego, CA) the renderings can be easily manipulated; for example, the annotations can be rotated, individually coloured and switched on or off (Fig. 3B, also see Supplementary Material B and http://www.mouseimaging.ca/var_brain_atlas.html).

The segmentation image was automatically customized for each individual brain by transforming along the appropriate deformation field. In particular, individual annotations for both raw images and globally normalized images were obtained.

Global Brain Size

In the first step of the atlas-creation algorithm, the individual brains were normalized to the average global size using the 12-parameter affine model. In order to estimate the true total brain sizes (before the normalization) we counted all brain voxels based on individual annotations of raw images. In doing so, we ensured a consistent definition across all input images of the total brain region being measured. The average \pm standard

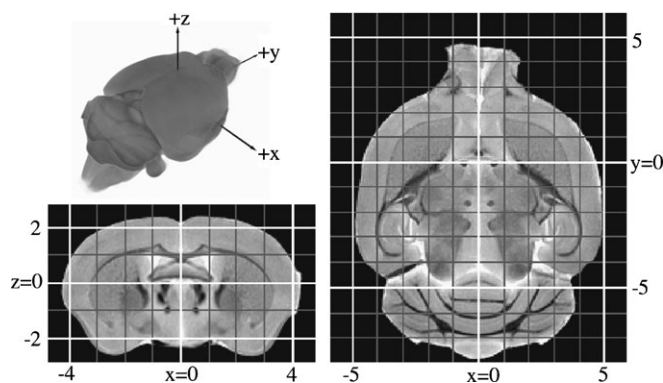


Figure 2. Stereotaxic coordinate system. Top left: three-dimensional view of the coordinate axes in relation to the average brain. Bottom left: coronal slice at $y = 0$ overlaid with the coordinate grid. Right: horizontal slice at $z = 0$ overlaid with the coordinate grid. All coordinates are given in millimetres.

deviation, minimum and maximum of the brain volume across the nine samples were 415 ± 24 , 365 and 440 mm^3 respectively.

Volumetric Measurement of Structures

The segmentation image enables volumetric measurement of the annotated structures. This is done by histogram counting, where the number of voxels with the same label determines the volume of the corresponding structure. For combined regions consisting of several structures the corresponding label counts are summed together. For example, the anterior commissure is represented as a sum of two labels, the pars anterior and the pars posterior. Table 1 shows volumes of selected structures based on the segmentation image. We estimated the variability of the calculated volumes based on the individualized annotations of the globally normalized images. Table 1 lists the standard deviation of said volumes evaluated from the nine individualized annotations.

Anatomically Significant Variability

By construction, the deformation fields capture the anatomical differences among individual brains. From the arithmetic mean of the vectors, we have the mean positional distance (MPD) image as a representation of the spatial variability across all anatomical locations. ‘Intensities’ in the MPD image represent distances in micrometres. The MPD image has the same size and coordinate system as the average image and is shown in Figure 4. Colour coding of the distances of the MPD image enables effective visualization of the variational component of the atlas.

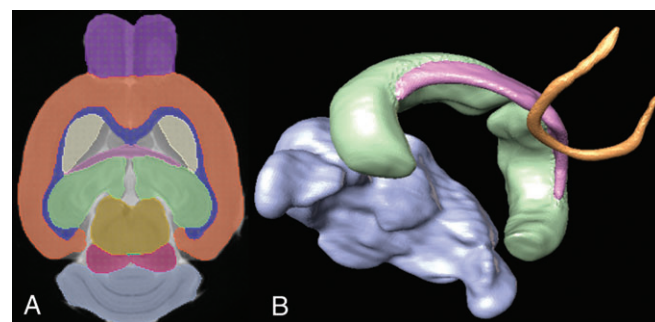


Figure 3. Three-dimensional annotation of the average image. (A) A two-dimensional view of the atlas brain overlaid with the annotations. Such views facilitate comparisons with conventional histological atlases. (B) A three-dimensional view of the anatomical structures and their spatial relationships; shown are surface renderings of cerebellum (blue), hippocampus (green), fimbria (pink) and anterior commissure (yellow).

Table 1

Volumetric measurements of selected structures

Structure	Volume $\pm \sigma$ (mm^3)
Cerebral cortex	109 ± 2
Striatum	15 ± 0.4
Hippocampus	21 ± 0.5
Brain stem	50 ± 2
Diencephalon	40 ± 0.5
Corpus callosum	13 ± 0.3
Fimbria of the hippocampus	2.2 ± 0.1
Anterior commissure	1.1 ± 0.04
Cerebellum	52 ± 1

Volumes were obtained from the three-dimensional annotation of the average image and shown in the right column. The corresponding annotations of the globally normalized individual brains were automatically obtained by transforming the atlas annotation image along the individual deformation fields. The reported standard deviations (σ) were calculated from these individualized annotations.

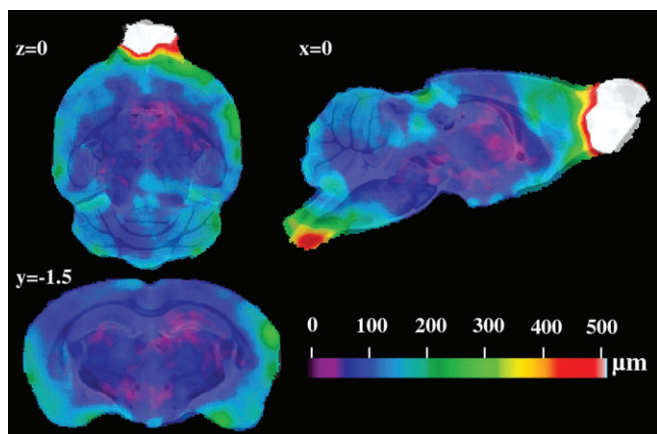


Figure 4. The mean positional difference image, MPD. Deformation magnitudes (measured in micrometres and shown in spectral colour scale) are overlaid on top of the average image (in grey colour scale). Purple and blue regions indicate low variability, while red and white indicate high variability. Shown are three orthogonal slices with stereotaxic coordinates given in millimetres.

For example, it is immediately clear that most inner structures are coloured purple and have low variability; in contrast, the olfactory bulbs and the caudal end of the brain stem are yellow, red and white, indicating high variability.

We have investigated the regions of large variability as indicated by the MPD image. Visual inspection of the individual images revealed inconsistencies that are typical of *in vitro* imaging: distortions along flexure points (e.g. midbrain and cerebellar flexure), slight differences in specimen preparation (e.g. the cutoff position at the caudal end of the brain stem), mechanical damage (e.g. one or both paraflocculi are missing from some specimens), and a small but highly variable amount of aqueous solution along the outer brain surface. In fact, we have determined that all regions with mean positional distances $> 240 \mu\text{m}$ were affected by such artefacts. The total volume of the affected regions accounts for $< 10\%$ of the brain volume.

We have calculated $117 \mu\text{m}$ as the mean spatial variability across all locations. The majority of anatomical locations (91% of the brain) have variability $< 180 \mu\text{m}$ (three voxels). In fact, the deep interior structures (comprising 53% of the brain) typically show variability $< 120 \mu\text{m}$ (two voxels). The most pronounced exception to this rule is the central part of the corpus callosum, which shows variability of $\sim 200 \mu\text{m}$. This is not surprising, given that the 129Sv inbred mouse strains are known to exhibit high callosal variability (Crawley, 2000).

Discussion

Implications for Anatomical Phenotyping

The anatomical variability within an inbred murine strain has been formalized and precisely quantified in three-dimensional space for the first time. The mathematical expression of normal variability — the deformation fields — represents a basis for detecting and quantifying subtle yet significant anatomical differences. For example, our techniques can be used to find differences between groups, such as different strains, hybrids or mutants. In this case, a variational atlas can be created for each group and comparisons between the atlases can be made. Image registration will allow precise quantification and characterization of anatomical differences.

The methodology introduced here also allows a rigorous approach to automatically detecting potential mutants. The deformation fields obtained after global normalization can be used for: (i) automated annotation and corresponding volumetric measurements and (ii) comparison with the normal average brain. A structure can be considered abnormal if found at stereotaxic locations with deformation magnitudes that are significantly larger than the corresponding MPD magnitudes. Not only will this type of analysis be useful for targeted mutagenesis, but it also has applications in screening for anatomical outliers in the context of random mutagenesis projects. In fact, coupled with the procedures we have developed for image acquisition, our image processing tools are ideally suited for high-throughput phenotypic screening. Further developments in parallel image acquisition techniques (Bock *et al.*, 2003), combined with our fully automated algorithms, will enable unsupervised, high-throughput and highly sophisticated anatomical phenotyping.

MRI versus Histological Atlases

Despite having lower resolution than published histological atlases ($5 \mu\text{m}$ versus $60 \mu\text{m}$), our three-dimensional MRI atlas exhibits several important advantages with respect to the processing and subsequent analysis of the CNS structure. First, the MRI atlas captures a complete picture of spatial relationships within the brain. This is in contrast to histological atlases where the specimen is sliced at $\sim 150 \mu\text{m}$ intervals. In addition, histological atlases of the CNS are typically obtained from paraffin- or parlodion-embedded specimens in which the true spatial dimensions within a given section cannot be readily determined because of shrinkage during specimen dehydration and processing. Similarly, rehydration and staining following sectioning are difficult to reproduce precisely from section to section. All these factors complicate determination of true three-dimensional distances between CNS structures and make accurate three-dimensional reconstructions difficult, although there have been efforts to digitally reconstruct histological sections back into three-dimensional space (Taguchi and Chida, 2003). Overall, an MRI-based atlas is superior for measuring volumes and analysis of long, distinct morphological features such as axon tracts.

Another advantage of MRI atlases over histological atlases relates to comparisons between sample specimens with a reference atlas. For example, if a brain specimen under investigation is significantly larger or smaller than the histological image, then the stereotaxic coordinates of a two-dimensional histological atlas will be difficult to compare. Such a situation is likely even in genetically identical lines of mice, as we have shown in the Results. Matching of the slicing angle to that of a two-dimensional atlas introduces further approximations. In contrast, our variational atlas enables much more accurate comparisons in three dimensions. After performing both linear and non-linear registrations, the resulting deformation field allows the sample brain to be fully annotated and measured in terms of structural volumes and shapes. Moreover, the variational component of the atlas allows us to estimate whether or not a sample brain is within the limits of natural variation.

Finally, three-dimensional MRI datasets can serve as a 'spatial backbone' onto which higher-resolution histological sections can be overlaid. In this sense, the MR data is used as a three-dimensional organizer for histological sections. One group has

started to create a mouse brain atlas based on MRI, histology as well as other imaging techniques (MacKenzie-Graham *et al.*, 2004).

Comparison with MR Atlases Based on a Single Brain

The variational atlas overcomes many of the problems associated with atlases based on a single individual. A single MR image may have imaging artefacts, such as the magnetic field susceptibility arising from small air bubbles, which only become worse with increasing magnetic field strength (Benveniste and Blackband, 2002). The excision procedure can also cause variability among specimens. Most importantly, without a quantitative definition of 'normal anatomy' it is not possible to ensure a choice of a valid single representative; any individual animal is a potential anatomical outlier. The methodology presented here provides means for creating an atlas that is representative of many individuals, is virtually free of artefacts and has an improved signal-to-noise ratio (Fig. 1).

Comparison with Other Atlases Based on Group Average

The atlas creation algorithm uses a novel approach for extracting the average anatomical representation from a group of images. It incorporates some of the techniques for image registration and probabilistic atlases that were originally developed for the study of anatomical variations in human brains (Collins and Evans, 1997; Grenander and Miller, 1998; Woods *et al.*, 1998; Guimond *et al.*, 2000; Kochunov *et al.*, 2001; Thompson and Toga, 2002). Existing techniques for defining an average brain anatomy from a group of images can be loosely divided into two schools of thought. Methods from one group produce highly resolved, but biased average images because of dependencies on the choice of a particular subject as target for all registrations (Guimond *et al.*, 2000; Kochunov *et al.*, 2001). The methods from the other camp produce unbiased, but blurry average images because of the use of global, low-level registration models applied to genetically heterogeneous human populations (e.g. the dataset, MNI305, comprised of 305 individual human brains can be found at http://www.bic.mni.mcgill.ca/cgi/icbm_view/).

The novelty of our approach lies in using an evolving intensity average image as the source for nonlinear registrations of the individual images. In this way we avoid problems associated with other algorithms that attempt to fully localize anatomical differences of the individual images with respect to a single individual. Instead, we use a multi-resolution strategy to gradually reach an unbiased, yet highly resolved, consensus among individual images. By design, our atlas brain not only represents the average geometry, but is also located at the geometric centre of the individual brains. Consequently, the anatomical variability in three spatial dimensions is easily understood and visualized, reminiscent of one-dimensional measurements. For example, the average image and the mean positional distance image MPD can be viewed as estimates of the population mean and standard error respectively.

Conclusions

In this work, we have combined magnetic resonance imaging and image processing to measure the anatomical variability in CNS structures of an inbred mouse strain. We have measured this variability to be very low — no more than 5% volume

change in several structures — leading us to believe that we have an excellent tool to detect mutant outliers.

Supplementary Material

Supplementary material can be found at:
<http://www.cercor.oupjournals.org/>.

Notes

The authors would like to thank Lori Davidson and Jun Dazai for technical assistance and Dr John G. Sled for helpful discussion. We also gratefully acknowledge Drs David Deerfield and Art Wetzel for developing the 3-D volume browser. The research is funded by the Canada Foundation for Innovation, the Ontario Innovation Trust, the Canadian Institutes of Health Research and the Burroughs Wellcome Fund.

Address correspondence to X. Josette Chen, Mouse Imaging Centre, Hospital for Sick Children, 555 University Avenue, Toronto, Ontario, Canada M5G 1X8. Email: josette@sickkids.ca.

Appendix

Image Registration Methods

For affine registrations we used the methodology of Woods *et al.* (1998a,b) and the corresponding software package AIR5.2.2 developed at the University of California, Los Angeles. We employed a 12-parameter affine transformation model with Levenberg-Marquardt minimization of the ratio image uniformity cost function. Complete details of AIR methodology and validation have been given previously (Woods *et al.*, 1998a,b).

For nonlinear registrations we used the multi-resolution, multi-scale ANIMAL methodology developed at the Montreal Neurological Institute (MNI) (Collins and Evans, 1997). Briefly, this methodology treats spatial transformations as deformation vector fields encoded as grid transforms. In the standard registration framework, where the source image deforms to align with the target image, grid transforms store the displacement vectors on regular three-dimensional grids of variable resolutions. In a low-resolution grid transform the displacement vectors are defined sparsely (e.g. for every 10th voxel coordinate); in this case the displacements of the remaining source voxels are obtained through interpolation. By contrast, a grid at the highest resolution records a unique displacement vector for every source voxel. The registration algorithm is designed to minimize the value of an objective function, which is constructed as a weighted sum of the similarity measure (a local correlation statistic) and a cost function, which constrains change in magnitude of the local deformation vectors. Continuity of deformation fields is achieved through moving window averaging. The multi-scale aspect of the methodology is implemented through the use of variable scale feature extraction (blurred image intensity and image gradient magnitude); the spatial scale of the extracted features is determined by the size of the convolving three-dimensional isotropic Gaussian kernel (full-width-half-maximum, FWHM = $2.35 \times$ standard deviation). Complete details and validation of this technique have been given previously (Collins and Evans, 1997).

Schedule of the Variational Atlas Algorithm

The schedule is graphically shown in Figure A1.

Step 0 (Affine Average)

In the initial step, the brains were first spatially normalized to share the same location, orientation and global size, and then intensity normalized to have equivalent brightness.

We arbitrarily selected one image, say I_n . The remaining images, I_1, \dots, I_{n-1} were registered to I_n using a global affine transformation model. The unbiased common affine space was defined by matrix averaging and reconciliation (Woods *et al.*, 1998a,b).

We next performed intensity normalization of the images J_1, \dots, J_n . We first corrected for image intensity non-uniformity, an MR artefact associated with radiofrequency field inhomogeneities, with an automated algorithm (Sled *et al.*, 1998). We then used another automated algorithm (Kovacevic *et al.*, 2002) to estimate mean grey matter intensities in J_1, \dots, J_n . The average value of the mean grey matter

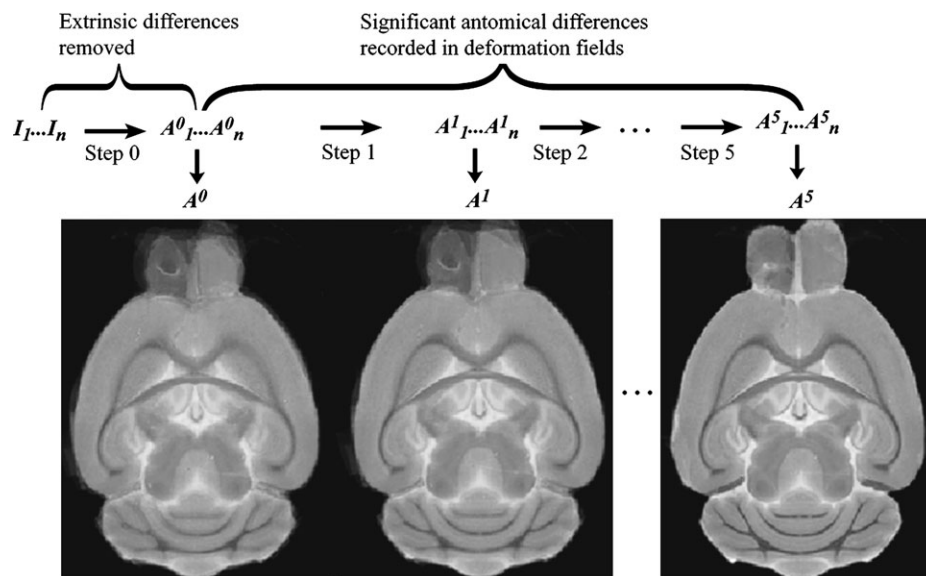


Figure A1. Atlas creation algorithm. Step 0 normalizes individual images in terms of the global brain size and gross shape. At the end of each step, an updated average image is produced. The evolution of the average image is marked by an increase in sharpness of CNS substructures.

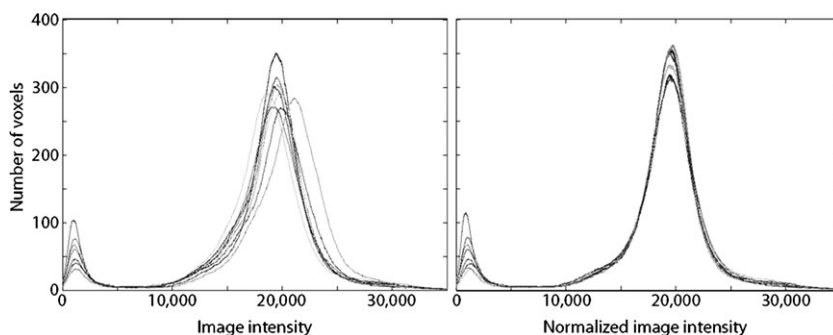


Figure A2. Intensity normalization. Left: histograms of non-normalized images J_1, \dots, J_n . Right: histograms of unhomogeneity corrected and intensity normalized images; the small peak corresponding to the intensity level ~ 1500 represents the mean background intensity, as calculated from the small amount of background voxels included in the measurements; the large peak corresponding to the intensity level $\sim 20\,000$ represents the mean grey matter intensity.

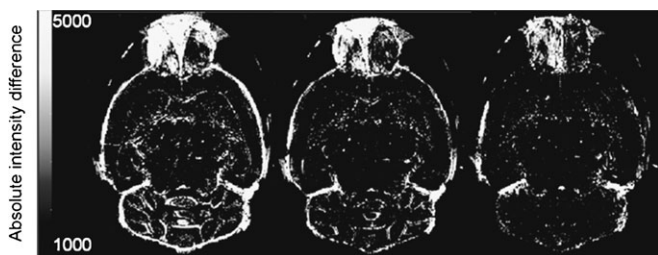


Figure A3. The absolute difference in image intensities between the average image and one of its components (Steps 0, 2 and 5). Left: the difference between A_k^0 and A^0 ; the outlines of anatomical features (e.g. corpus callosum and arbor vita) indicate local misalignments between the two images. Middle: the difference between A_k^2 and A^2 ; the outlines are still present, though less pronounced. Right: the difference between A_k^5 and A^5 ; no outlines remain within the brain interior, as a result of a group-wide consensus being reached. The bright spots along the outer surface and in olfactory bulbs are due to artefacts (see Results/Anatomically significant variability).

intensity across all images was calculated next. The images were normalized to the average grey matter peak using linear intensity rescaling. Figure A2 illustrates the effect of intensity normalization on image histograms.

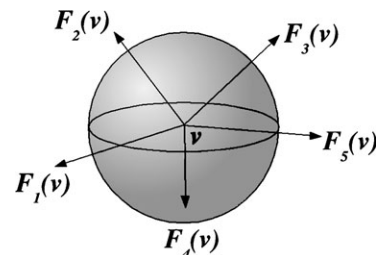


Figure A4. The centroid from the individual images. Every anatomical location v in the average image is at the centroid of the corresponding locations in individual images.

Images at this point were spatially and intensity normalized, inhomogeneity corrected versions of the initial raw images and are denoted A_1^0, \dots, A_n^0 . The initial average image A^0 is created as a voxel-by-voxel intensity average of A_1^0, \dots, A_n^0 . Intensity normalization is an important step; without it, voxel-by-voxel intensity averaging would be biased towards brighter images. This is undesirable because brightness of MR images is more likely to depend on variations in the MR experiment than on real biological differences.

Table A1

Schedule of nonlinear registrations in steps 1-5

	Grid resolution (μm)	Gaussian FWHM (μm)/feature
Step 1.1	720	240/intensity blur
Step 1.2	600	240/gradient magnitude
Step 2.1	480	180/intensity blur
Step 2.2	360	180/gradient magnitude
Step 3.1	240	180/intensity blur
Step 3.2	180	180/gradient magnitude
Step 4.1	180	120/intensity blur
Step 4.2	120	120/gradient magnitude
Step 5.1	120	120/intensity blur
Step 5.2	60	120/gradient magnitude

The middle column represents grid resolutions in terms of the nodal distance. The right column lists feature extracted images used in each step; the feature scales are measured by the size of the convolving Gaussian kernels.

Step 1 (First Generation Nonlinear Average)

The purpose of this step is to account for large-scale nonlinear anatomical differences. We started with a two-step nonlinear registration of the average A^0 to all individual images A_1^0, \dots, A_n^0 (steps 1.2 and 1.3 in Table A1). Using the corresponding inverse transforms, the individual images were resampled to produce images A_1^1, \dots, A_n^1 . This new set of individual images was then intensity averaged to produce a new average image A^1 .

Steps 2-5 (Second-Fifth Generation of Nonlinear Average)

The remaining steps were similar to step 1: the most recent average image was nonlinearly registered to its constituents and an updated set of individual images was produced. In order to avoid repeated resamplings that tend to accumulate interpolation error, we instead resampled the initial images using the concatenated transforms. In all resampling steps we used a windowed sinc interpolation with window size of 11 voxels in all three directions.

At the end of the last step, the concatenated individual deformation fields $\mathbf{g}_1, \dots, \mathbf{g}_n$ were defined so that $\mathbf{g}_k(A_k^0) = A_k^5$. More precisely, $\mathbf{g}_k = \mathbf{g}_k^5 \cdots \mathbf{g}_k^1$, where \mathbf{g}_k^i denotes the transform that was obtained in step i such that $\mathbf{g}_k^i(A_k^{i-1}) = A_k^i$. The last average image A^5 is an intensity average of A_1^5, \dots, A_n^5 . The registration throughout the nonlinear steps is scheduled so that the full resolution is reached at the end of step 5. Details of the registration parameters used in all non-linear steps (1-5) are given in Table A1. Note how each non-linear step contains two substeps to allow gradual progression in spatial matching. By the end of step 5, the full spatial normalization among individual images is achieved (Fig. A3).

Finally, we applied geometric centring to produce a 'minimal deformation target', as introduced by Kochunov *et al.* (2001). Briefly, we calculate the average deformation field \mathbf{h} by averaging grid transforms $\mathbf{g}_1^{-1}, \dots, \mathbf{g}_n^{-1}$ for each voxel in the

A^5 -space

The concatenated deformation fields $\mathbf{F}_k = \mathbf{g}_k \mathbf{h}$, were then used to produce $A_k = \mathbf{F}_k(A_k^0)$, $k = 1, \dots, n$. The intensity average of A_k s was defined as the final atlas average image A . By construction, A is geometrically centred with respect to the individual images A_1^0, \dots, A_n^0 . For every anatomical location v in A , its n deformation vectors point towards homologous individual locations $\mathbf{F}_1(v), \dots, \mathbf{F}_n(v)$ in A_1^0, \dots, A_n^0 in such a way that v is at the centroid of the individual locations (Fig. A4). As discussed in the main text, the arithmetic mean of the vector magnitudes $|\mathbf{F}_1(v)|, \dots, |\mathbf{F}_n(v)|$ is the mean positional distance, $\text{mpd}(v)$ (MPD in the main text). The sphere centred at v , with radius equal to $\text{mpd}(v)$ represents the spatial variability of the location v .

References

Balaban RS, Hampshire VA (2001) Challenges in small animal non-invasive imaging. *ILAR J* 42:248-62.

- Benveniste H, Blackband S (2002) MR microscopy and high resolution small animal MRI: applications in neuroscience research. *Prog Neurobiol* 67:393-420.
- Benveniste H, Kim K, Zhang L, Johnson GA (2000) Magnetic resonance microscopy of the C57BL mouse brain. *Neuroimage* 11: 601-611.
- Bock NA, Konyer NB, Henkelman RM (2003) Multiple-mouse MRI. *Magn Res Med* 49:158-167.
- Collins DL, Evans AC (1997) ANIMAL: validation and applications of nonlinear registration based segmentation. *Int J Pattern Recogn* 11:1271-1294.
- Crawley JN (2000) What's wrong with my mouse? Behavioral phenotyping of transgenic and knockout mice, pp. 9-29. New York: Wiley-Liss.
- Dhenain M, Ruffins SW, Jacobs RE (2001) Three-dimensional digital mouseatlas using high-resolution MRI. *Dev Biol* 232:458-470.
- Franklin K, Paxinos G (1997) The mouse brain in stereotaxic coordinates. San Diego: Academic Press.
- Grenander U, Miller MI (1998) Computational anatomy: an emerging discipline. *Q Appl Math* 56:617-694.
- Guilfoyle DN, Dyakin VV, O'Shea J, Pell GS, Helpert JA (2003) Quantitative measurements of proton spin-lattice (T1) and spin-spin (T2) relaxation times in the mouse brain at 7.0 T. *Magn Reson Med* 49:576-580.
- Guimond A, Meunier J, Thirion JP (2000) Average brain models: a convergence study. *Comput Vis Image Understand* 77:192-210.
- Idziak S, Haerberlen U (1982) Design and construction of high homogeneity RF coil for solid-state multiple-pulse NMR. *J Magn Reson* 50:281-288.
- Johnson GA, Benveniste H, Black RD, Hedlund LW, Maronpot RR, Smith BR (1993) Histology by magnetic resonance microscopy. *Magn Reson Q* 9:1-30.
- Johnson GA, Benveniste H, Engelhardt RT, Qiu H, Hedlund LW (1997) Magnetic resonance microscopy in basic studies of brain structure and function. *Ann N Y Acad Sci* 820:139-147.
- Johnson GA, Cofer GP, Fubara B, Gewalt SL, Hedlund LW, Maronpot RR (2002) Magnetic resonance histology for morphologic phenotyping. *J Magn Reson Imaging* 16:423-429.
- Kochunov P, Lancaster JL, Thompson P, Woods R, Mazziotta J, Hardies J, Fox P (2001) Regional spatial normalization: towards an optimal target. *J Comput Assist Tomogr* 25:805-816.
- Kovacevic N, Lobaugh NJ, Bronskill MJ, Levine B, Feinstein A, Black SE (2002) A robust method for extraction and automatic segmentation of brain images. *Neuroimage* 17:1087-1100.
- MacKenzie-Graham A, Jones ES, Shattuck DW, Dinov ID, Bota M, Toga AW (2004) A multimodal, multidimensional atlas of the C57BL/6j mouse brain. *J Anat* 204:93-102.
- Nolan PM, Peters J, Strivens M, Rogers D, Hagan J, Spurr N, Gray IC, Vizor L, Brooker D, Whitehill E, *et al.* (2000) A systematic genome-wide, phenotype-driven mutagenesis programme for gene function studies in the mouse. *Nat Genet* 25:440-443.
- Sled JG, Zijdenbos AP, Evans AC (1998) A nonparametric method for automatic correction of intensity nonuniformity in MRI data. *IEEE Trans Med Imag* 17:87-97.
- Taguchi M, Chida K (2003) Computer reconstruction of the three-dimensional structure of mouse cerebral ventricles. *Brain Res Brain Res Protoc* 12:10-15.
- Thompson PM, Toga AW (2002) A framework for computational anatomy. *Comput Vis Sci* 5:13-34.
- Van Essen, DC. (2002) Surface-based atlases of cerebellar cortex in the human, macaque, and mouse. *Ann NY Acad Sci* 978:468-479.
- Woods RP, Grafton ST, Collins JH, Cherry SR, Mazziotta JC (1998a) Automated image registration: I. General methods and intra-subject, intramodality validation. *J Comput Assist Tomogr* 22: 139-152.
- Woods RP, Grafton ST, Watson JD, Sicotte NL, Mazziotta JC (1998b) Automated image registration: II. Intersubject validation of linear and nonlinear models. *J Comput Assist Tomogr* 22:153-165.