TECHNICAL WHITE PAPER:
***IN SITU* HYBRIDIZATION**

BrainSpan, an atlas of the developing human brain, is designed as a foundational resource for studying transcriptional mechanisms involved in human brain development. The resource is the outcome of three ARRA-funded grants through the National Institutes of Health to support a consortium consisting of the Allen Institute for Brain Science; Yale University (Nenad Sestan, Mark B. Gerstein); the Zilkha Neurogenetic Institute of the Keck School of Medicine of the University of Southern California (James A. Knowles, Pat Levitt); the Athinoula A. Martinos Center at Massachusetts General Hospital/Harvard Medical School and MIT HST/CSAIL (Bruce Fischl); the University of California, Los Angeles (Daniel H. Geschwind); and the University of Texas Southwestern Medical Center (Hao Huang) with strong collaborative support from the Genes, Cognition and Psychosis Program, which is part of the Intramural Research Program of NIMH, NIH (Thomas M. Hyde, Joel E. Kleinman, Daniel R. Weinberger).

One of the datasets provides gene expression data at cellular level resolution using colorimetric *in situ* hybridization (ISH) at various time points. The data are publicly accessible via a Web-based application that allows viewing of indexed image sets searchable by gene, anatomic region and tissue and subject characteristics. These data sets can be accessed through the ALLEN BRAIN ATLAS data portal or directly at www.developinghumanbrain.org.

Developmental Stage	Anatomic Regions Characterized	Gene Selection
Prenatal Development (15 - 21 pcw)	Whole brain	Canonical morphological and cell type markers based on literature and rodent ISH data sets, disease related genes associated with neocortical development. Gene selection was preferential towards data available through the Allen Developing Mouse Brain Atlas, allowing a direct phylogenetic comparison of gene expression patterns.
Postnatal Development (Birth – 19 years)	Medial prefrontal cortex Ventral striatum Amygdala Hippocampus Visual cortex	A set of functionally important and anatomically restricted cellular markers and key developmental genes designed to match the complementary NIH Blueprint Non-Human Primate (NHP) Atlas (www.blueprintnhpatlas.org), enabling a direct phylogenetic comparison of expression patterns in human and macaque, as well as mouse via the Allen Developing Mouse Brain Atlas.
Adulthood (20 – 60 years)	Medial prefrontal cortex Ventral striatum Amygdala Hippocampus Visual cortex	Genes thought to be important in human neuropsychiatric or neurologic diseases such as Alzheimer's disease, autism, depression, epilepsy, obsessive-compulsive disorder, Parkinson's disease and schizophrenia.

pcw = post-conception weeks

A list of the specific genes included in the project can be found in an accompanying document (Gene List) in the Documentation tab on the website.

POSTMORTEM HUMAN TISSUE

Subjects

Postmortem tissue samples from male and female subjects between 15 – 24 post-conception weeks were provided by the Laboratory of Developmental Biology, University of Washington and Advanced Bioscience Resources, Incorporated.

Frozen postmortem tissue samples from male and female subjects between 3 months – 53 years of age at death were provided by the brain tissue collection of the Neuropathology Section of the Clinical Disorders Branch, GCAP, IRP, National Institute of Mental Health, NIH, Bethesda, MD.

Specimens from the NIMH collection were processed and characterized as previously described (Lipska et al, 2006). Briefly, after obtaining informed consent from legal next-of-kin, demographic information and medical, substance use and psychiatric history was gathered. A postmortem clinical diagnosis was independently established by board-certified psychiatrists based on family interviews consisting of the Structured Clinical Interview for DSM-IV – clinician version (SCID-CV), the NIMH psychological autopsy interview and the Diagnostic Evaluation After Death of psychiatric, medical, police, and autopsy reports (Deep-Soboslay et al, 2005). Cases and specimens were further characterized for history of cigarette smoking, toxicology for illicit substances and prescription medications, a macro- and microscopic neuropathological examination by a board-certified neuropathologist, manner of death, verification of age at death, and postmortem interval (PMI). PMI was calculated as time elapsed, in hours, between pronounced time of death and time of tissue freezing. For cases in which time of death was not pronounced by hospital or emergency medical personnel, time of death was estimated as the midpoint between the time the individual was last seen alive and the time the decedent was discovered. Smoking history (regular use of cigarettes) was rated as "yes", “no", or “unknown” at the time of death based on medical records, family interviews, and/or nicotine levels.

Subjects for these studies had normal neuropathological examination results and no known history of neuropsychiatric disease, as determined by the above screening mechanism. Evidence of drug use, as determined from interviews or by toxicology tests, led to exclusion of cases. Cases in which manner of death was suicide or in which death was due to drug overdose or poisoning were also excluded.

Cerebellar pH was determined for all tissue samples in this study as described (Lipska et al, 2006) and any samples found to have pH < 6 were excluded from this study. Other tissue criteria included restriction of PMI generally to less than 36 hours, but longer PMI did not exclude samples if the pH criterion was met.

A protocol for this project was reviewed by Western Institutional Review Board, resulting in issuance of a Determination of Exemption.

Tissue Dissection and Confirmation of Anatomic Structures

Brain tissue from subjects between 15 – 24 post-conception weeks was harvested and sectioned into single hemisphere coronal slabs approximately 1.5 to 2 cm in thickness. Slabs were quick-frozen in isopentane chilled to -50°C and were stored at -80°C until the time of sectioning.

Brain tissue from postnatal and adult stages was harvested and sectioned into single hemisphere coronal slabs approximately 1 to 1.5 cm in thickness. Slabs were quick-frozen in isopentane chilled to -50°C and were stored at -80°C until the time of dissection. Specific anatomic structures to be characterized were dissected using a dental drill or dissecting blade to isolate the region or regions of interest from each coronal slab. Boundaries of each structure were defined prior to dissection in order to accommodate scannable portions of a 2” by 3” slide (tissue dimensions approximately 3.5cm x 3.5cm), and to accommodate anatomic overlap or proximity, particularly for those age groups in which genes characterized by ISH were assigned in a structure-specific manner. In general, each structure spanned multiple coronal slabs and in many instances, a starting boundary or transition point from one structure to the next occurred in the middle of a

tissue block. Therefore, a cresyl violet-based Nissl stain was utilized during cryosectioning of tissue in order to accurately determine the location of a starting or a transition point based on appearance of observable anatomic features on the Nissl-stained section. The stained tissue was also used to confirm expected cellular density and architecture for a given structure.

Anatomic regions were defined in the following manner:

- Medial prefrontal cortex (mPFC) extended from the frontal pole to the anterior aspect of the caudate nucleus. Tissue from the pole was collected once two gyri were present and connected by white matter. In more posterior aspects of the mPFC, where mPFC extends beyond the size limit of the slide, focus was placed on the medial and ventral aspects, and blocked approximately 3.5 cm lateral and superior to the ventro-medial aspect of the gyrus rectus.
- Ventral striatum extended from the anterior aspect of the caudate nucleus to the posterior aspect of the nucleus accumbens (when available).
- Amygdala extended from the anterior to the posterior portion of the amygdala. In many instances, the central nucleus was split between the amygdala block and the ventral striatum block due to its location in the tissue.
- Hippocampus extended from the posterior aspect of the amygdala to the pes hippocampi.
- Visual cortex extended from the anterior portion of the calcarine fissure and included striate cortex. Where possible, secondary visual cortex was included as well.

RNA Qualification

Prior to sectioning for ISH, tissue samples were tested for RNA quality For RNA quality assessment purposes, two tissue sample sections weighing 2 - 10 mg each were used for RNA extraction using the MELT Total Nucleic Acid Isolation System (Ambion, Foster City, CA). Samples were stored at room temperature for up to 1 week after lysis and processed in 96-well plates to enable high-throughput processing. RNA was eluted with 20 µl nuclease-free water and stored at -80°C. Yields were typically between 300-500 ng total RNA. No further purification of RNA samples was performed.

RNA quality was determined for each specimen using high-resolution capillary electrophoresis on an Agilent Bioanalyzer 2100 (Agilent Technologies, Palo Alto, California) and Agilent's RIN software algorithm to generate RNA Integrity Numbers (RIN values). RNA was eluted in 20 µl nuclease-free water, standardized to a concentration of 5 ng/µl and 1.0 µl was assayed on a Pico Bioanalyzer chip. RIN values were based on the entire electrophoretic trace and ranged from 1 to 10, where 1 corresponds to completely degraded RNA and 10 corresponds to perfectly intact RNA (Schroeder et al, 2006).

LABORATORY PROCEDURES FOR HISTOLOGICAL STAINING AND *IN SITU* HYBRIDIZATION

Sectioning

Frozen tissue samples were sectioned in Leica CM3050 S cryostats (object temperature, -10°C; chamber temperature, -15°C) at 20 µm thickness in the coronal plane from anterior to posterior. One section was placed on each positively charged Superfrost Plus 2" x 3" microscope slide (Erie Scientific Co, Portsmouth, NH), pre-printed with a unique identifying barcode for tracking. Specimen numbers were also printed onto the slides for tracking purposes.

Following sectioning, slides designated for ISH were allowed to air dry and tissue was fixed, acetylated and dehydrated according to standard protocols as described (Lein et al, 2007). For tissue from prenatal stages, slides were baked in a 37°C oven instead of air dried prior to fixation, acetylation, and dehydration. Briefly, tissue was fixed for 20 minutes in 4% neutral buffered paraformaldehyde (PFA) and rinsed in 1x PBS, acetylated for 10 minutes in 0.1M triethanolamine with 0.25% acetic anhydride, and subsequently dehydrated using a graded series of 50%, 70%, 95% and 100% ethanol. Slides that passed section quality checks were stored at room temperature in Parafilm™-sealed slide boxes until used.

Histological Staining

Acetylcholinesterase (AChE)

A modified acetylcholinesterase protocol was used to help delineate subcortical structures at high resolution. Unlike AChE staining in fixed tissue, staining in fresh frozen tissue does not elucidate cholinergic fibers commonly seen in the cortex, but instead provides demarcation of various subcortical nuclei. AChE staining was performed using a direct coloring thiocholine method combined with a methyl green nuclear counterstain to improve tissue visibility (Karnovsky and Roots, 1964). Glass slides with fresh frozen tissue sections were removed from 4°C, allowed to equilibrate to room temperature, fixed in 10% neutral buffered formalin (NBF) and washed briefly in ultra pure water. Sections were then incubated for 30 minutes in a solution of acetylthiocholine iodide, sodium citrate, cupric sulfate, and potassium ferricyanide in a 0.1M sodium acetate buffer (pH 6.0), washed in 0.1M Tris-HCl buffer (pH 7.2), incubated with 0.5% diaminobenzidine (DAB) in 0.1M Tris-HCl with 0.03% hydrogen peroxide. Slides were incubated in 0.2% methyl green, briefly dipped in 100% EtOH, cleared with Formula 83 and coverslipped with DPX.

Nissl Staining

After brain tissue was sectioned, slides were stored at 37°C for 1 – 5 days and were removed 5 – 15 minutes prior to staining. Sections were defatted with xylene or the xylene substitute Formula 83, and hydrated through a graded series containing 100%, 95%, 70%, and 50% ethanol. After incubation in water, the sections were stained in 0.213% thionin, then differentiated and dehydrated in water and a graded series containing 50%, 70%, 95%, and 100% ethanol. Finally, the slides were incubated in xylene or Formula 83, and coverslipped with the mounting agent DPX. After drying, the slides were analyzed microscopically to ensure staining quality. Slides that passed QC were stored at room temperature in slide boxes before being cleaned in preparation for digital imaging.

Colorimetric *In Situ* Hybridization

A colorimetric, digoxigenin-based method for labeling target mRNA was used to detect gene expression on human tissue sections.

Probe Design and Synthesis

For labeling target mRNA in tissue sections using ISH, digoxigenin-labeled riboprobes were designed and synthesized according to specific criteria. Briefly, using sequences obtained from RefSeq and a semi-automated process based on Primer3 software (Rozen and Skaletsky, 2000), probes were designed to be between 400-1000 bases in length (optimally > 600 bases) and to contain no more than 200 bp with > 90% homology to non-target transcripts. In addition, to allow comparability of mouse and human gene expression datasets, human probes were designed to have > 50% overlap with the existing Allen Mouse Brain Atlas probe when the mouse and human genes were orthologous. Riboprobes were synthesized using standard *in vitro* transcription (IVT) reactions based on PCR templates prepared from human cDNA clones (NIH Mammalian Gene Collection, Open Biosystems, Huntsville, AL) or pooled cDNA synthesized from human brain total RNA. cDNA was prepared from human brain RNA from prefrontal, temporal, parietal, occipital, and frontal cortical areas as well as medulla and cerebellum (Ambion, Austin, TX) using Superscript III RTS First-Strand cDNA Synthesis Kit (Invitrogen, Carlsbad, CA), then pooled in equal amounts to provide templates for PCR.

PCR primers were obtained from Integrated DNA Technologies (Coralville, IA) at a final concentration of 10 µM, and designed with GC content between 42% – 62% and an optimal size of 22nt with lower and upper limits of 18 and 26 nt, respectively. For cDNA clones, the clone sequence was compared with RefSeq sequences, and consensus sequences with > 98% homology across 80% of the total length were used to develop probes. When a clone was used as a template, a single PCR was used requiring only a forward and reverse primer with an additional SP6 RNA polymerase binding sequence (GCGATTTAGGTGACACTATAG). When using brain cDNA as a template, probes were generated against sequences within a region 3000 bp from the 3' end using 3 primers: forward, reverse, and a nested reverse primer containing the SP6 RNA polymerase binding sequence. cDNA primers underwent a BLAST analysis to verify amplification of only target sequence. All cDNA reactions were run on the Bioanalyzer for quality control and all PCR products generated from cDNA template were sequenced from both ends using MegaBACE and ABI3700 capillary instruments at the University of Washington High Throughput Genomics Unit.

Standard conditions for PCR and IVT reactions were as described by Lein et al (2007). IVT reactions were diluted to working stocks of 30ng/µl with THE (0.1mM Sodium Citrate pH 6.4, Ambion). Aliquots were stored in one- or two-use volumes to minimize freeze/thaw cycles. IVT dilutions were stored at -80°C. For hybridization, the probe was diluted 1:100 (to 300ng/ml) into *in situ* hybridization buffer (Ambion) in 96-well ISH Probe Plates. Each well provides probe for one ISH slide. Probe plates were stored at -20°C until used in an ISH run.

All PCR and IVT products were run on the Bioanalyzer for size and morphology quality control. Specifically, PCR products that were not of the correct size (+/- 100bp) or that showed multiple products were not used to generate riboprobes. IVT products that were shorter than their predicted size were not used. It is common to see IVT products that run slightly larger than their predicted molecular weight, or as multiple peaks, due to secondary structure of the RNA. IVT products with multiple bands were not used for ISH unless the additional bands were determined to result from secondary structure.

ISH processes were performed on robotic platforms that allowed processing of approximately 100 2" x 3" or 200 1" x 3" slides. Detailed descriptions of the high-throughput platform, protocols, and reagent preparation are available elsewhere for 1" x 3" slides (Lein et al, 2007). Minor protocol modifications were made to accommodate the larger slides and were primarily with respect to delivery of reagent volumes. In general, slides containing tissue sections were placed in flow-through chambers on temperature-controlled racks on computer-controlled Tecan Genesis liquid handling platforms for addition of solutions. The first steps blocked endogenous peroxidase activity and permeabilized the tissue, followed by subsequent hybridization of digoxigenin-labeled probes to target mRNA. After a series of washes to eliminate excess probe, the remaining bound probe was subjected to a series of enzymatic reaction steps to detect and amplify digoxigenin signal. First, a horseradish peroxidase (HRP)-conjugated anti-digoxigenin antibody was added, followed by biotin-coupled tyramide that is converted by HRP to an intermediate that binds to cell-associated proteins at or near the HRP-linked probe. Neutravidin conjugated with alkaline phosphatase (AP) was then bound to biotin and BCIP/NBT was added. A blue/purple particulate precipitate forms as a result of the enzymatic cleavage of BCIP by AP and subsequent indole reaction with NBT. Finally, the colorimetric reaction was stopped by washing with EDTA and fixed with 4% PFA. This entire process occurred over the course of approximately 23.5 hours on the Tecan automated platform. A minor modification of the ISH protocol was made for the prenatal tissue: a decrease in proteinase K concentration to 0.0135 U/ml.

To reduce background signal, an acid alcohol wash step was performed on a Leica Autostainer after completion of the hybridization process. Slides were rinsed 4 times (1 minute each) in acid alcohol (70%, adjusted to pH = 2.1 with 12N HCl) and rinsed 4 times in milliQ water (1 minute each). Acid alcohol and water solutions were refreshed every fourth rack of slides to ensure that all slides were rinsed in clean solution.

Some ISH slides (all prenatal ISH slides and ~25% of the adult ISH slides) were stained with 1N feulgen-HP yellow counterstain (Anatech Ltd.) on a Leica Autostainer to improve image acquisition. After the completion of the ISH protocol (fixation with 4% PFA), slides were rinsed in milliQ water, then rinsed twice (1 minute each) in acid alcohol (70%, adjusted to pH = 2.1 with 12N HCl), rinsed once in DI water (~2 minutes), rinsed three times in 1N HCl (~6 minutes each), rinsed twice in DI water (2-3 minutes each), rinsed three times in HP yellow (~6 minutes each), rinsed twice in DI water (2-3 minutes each), rinsed twice in acid alcohol (1 minute each), and finally rinsed in milliQ water.

Each postnatal and adult ISH run contained several controls: GAP43 and Drd1a were used as positive controls on human and mouse tissue, respectively, to provide verification of a successful ISH run. In addition, GAP43 on prenatal tissue was tested on two slides as an indicator of fetal ISH quality and internal consistency within the ISH run. The postnatal and adult human tissue for the positive control was consistently from the same subject each time. In addition, a moderately expressed gene, CTNND2, was tested for each individual tissue block, as an indicator of tissue block quality. Extremely high or extremely low CTNND2 expression levels beyond what would normally be expected for human variation in expression could be in indicator of compromised tissue or RNA quality that was otherwise missed. A negative control (no probe) slide was also included as an indication of background for each ISH run.

IMAGE ACQUISITION AND DATA PROCESSING

Digital imaging of ISH and stained slides was performed using ScanScope or ScanScopeXT scanners (Aperio Technologies, Inc; Vista, CA). The line scan camera continually adjusts for focus based on a variable number of focus points and is an advantage for large human tissue sections that tend to have more variation in height. The ScanScope scanner uses a 20x objective downsampled in software to minimize data volume with a final image resolution of approximately 1 µm/pixel.

Once images were acquired, the Informatics Data Pipeline (IDP) managed image preprocessing, image QC, ISH expression detection and measurement, Nissl processing, annotation QC and public display of information via the Web application. The IDP has been described in detail previously (Dang et al., 2007), and was modified slightly for processing human images.

IDP Processing Modules

Four modules constituted the processing pipeline for human cortex images: (1) image preprocessing, (2) ISH expression detection, (3) Nissl processing and (4) AChE processing. In image preprocessing, scanned ISH, Nissl and AChE images were converted and processed to provide more consistent white background intensities and orientation across samples. Aperio ScanScope images were first converted to JPEG 2000 format, then orientation adjusted and white balanced.

For ISH expression detection, adaptive image processing techniques were applied to 10x full-resolution ISH images to detect and quantify gene expression. As an example of the scale of the engineering challenge, the number of detected expressors (cells) can reach the level of nearly 2 million in a typical 3 GB cortex image. With large tissue sizes, image quality with regard to uniformity was compromised and robust detection became much more dependent on stain intensity, contrast and focus. An algorithm based on techniques used for the Allen Mouse Brain Atlas (1; Supplemental Methods 2) was significantly redesigned to accommodate the long addressing mode and full image resolution needs that resulted from large image sizes. The resulting module produced a mask of detected expressor objects and a set of numerical values describing the statistical attributes of gene expression. The mask image with measured intensity of expression was then pseudo-color coded and converted to AFF file format for Web display.

Nissl stained sections were processed to determine tissue area size and estimate grey matter area as a reference for computation of the relative expressing density indicator available in the Web application when viewing images. Computed expression measurements were normalized by estimated tissue area from the closest Nissl section to provide a relative indication of expression within the whole tissue or gray matter area. AchE stained subcortical tissue sections were preprocessed for white balance for online presentation without additional analysis.

QUALITY CONTROL AND WEB-BASED DATA PRESENTATION

During the course of data processing additional quality control steps were implemented to ensure high quality data images. Once image preprocessing was complete, image quality control ensured the images were in good focus and provided an initial indication of the presence of signal. If focus criteria were not met, the images were failed and the appropriate slides were rescanned. If focus criteria were met, the images were passed and processed through the IDP as described above.

Raw ISH data, quantified heat mask representations, and accompanying histological data generated in this project are publicly accessible via the Allen Brain Atlas data portal at www.brain-map.org or directly at www.developinghumanbrain.org.

REFERENCES

Dang CN et al (2007) The Allen Brain Atlas: Delivering neuroscience to the Web on a genome wide scale. In *Data Integration in the Life Sciences.* pp17-26. Springer-Berlin.

Deep-Soboslay A, Akil M, Martin CE, Bigelow LB, Herman MM, Hyde TM, KleinmanJE (2005) Reliability of psychiatric diagnosis in postmortem research. *Biol Psychiatry* 57:96-101.

Karnovsky MJ, Roots L (1964) A “direct coloring” thiocholine method for cholinesterases. *J Histochem Cytochem* 12:219-221.

Kiernan, JA (2008) *Histological and Histochemical Methods: Theory and Practice, 4th Ed.* Bloxham, UK: Scion Publishing (pp. 415- 427).

Lein ES et al (2007) Genome-wide atlas of gene expression in the adult mouse brain. *Nature* 445:168-176.

Lipska BK, Deep-Soboslay A, Weickert CS, Hyde TM, Martin CE, Herman MM, Kleinman JE. (2006) Critical factors in gene expression in postmortem human brain: Focus on studies in schizophrenia. *Biol Psychiatry* 60:650-658.

Rozen S, Skaletsky, HJ (2000) Primer3 on the WWW for general users and for biologist programmers. In *Bioinformatics Methods and Protocols: Methods in Molecular Biology*. pp365-386. Humana Press.

Schroeder A, Mueller O, Stocker S, Salowsky R, Leiber M, et al (2006) The RIN: an RNA integrity number for assigning integrity values to RNA measurements. *BMC Mol Biol* 7, 3 (2006).

Sun A, Nguyen XV, Bing G (2002) Comparative analysis of an improved thioflavin-s stain, Gallyas silver stain, and immunohistochemistry for neurofibrillary tangle demonstration on the same sections. *J Histochem Cytochem* 50:463-472.

Tomita H, Vawter MP, Walsh DM, Evans SJ, Choudary PV, et al (2004) Effect of agonal and postmortem factors on gene expression profile: quality control in microarrayanalyses of postmortem human brain. *Biol Psychiatry* 55:346-52.