

ALLEN Human Brain Atlas

TECHNICAL WHITE PAPER: *IN SITU* HYBRIDIZATION IN THE ALLEN HUMAN BRAIN ATLAS

The Allen Human Brain Atlas is a publicly available online resource of gene expression information in human brain comprising multiple datasets from several projects. Colorimetric *in situ* hybridization (ISH) methods were used in several of these projects to provide gene expression data at cellular level resolution in specific brain regions. 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 datasets can be accessed through the Allen Brain Atlas data portal (www.brain-map.org) and are integrated with other Allen Brain Atlas resources.

Datasets are available from the following projects: (1) 1,000 Gene Survey in Cortex; (2) Autism Study; (3) Neurotransmitter Study; (4) Subcortex Study; and (5) Schizophrenia Study. Table 1, below, summarizes these studies.

Table 1. Summary of ISH datasets available in the Allen Human Brain Atlas.

Project	Purpose	Tissue	Gene Selection
1,000 Gene Survey in Cortex (Cortex Study)	Characterize expression of genes from multiple gene classes in two cortical regions	Visual cortex and middle temporal cortex from multiple adult control cases (n = 2 – 6 per gene)	Multiple genes (1,000) from various functional and marker gene classes, including genes related to neuropsychiatric or neurological disease and genes of interest in comparative genomics
Autism Study¹	Compare cortical microstructure of control and autism cases	Frontal, temporal, and occipital cortical regions from young postmortem control and autism cases (n=11 per condition)	Lamina- and cell type-specific molecular markers including markers for neurons, glial, and risk genes for autism (25 genes selected from an initial panel of 64 genes)

Project	Purpose	Tissue	Gene Selection
Neurotransmitter Study	Survey neurotransmitter system gene expression throughout human brain in control cases	Multiple cortical and subcortical regions, focused on: DLPFC, ventral PFC, anterior & posterior cingulate, amygdala & hippocampus, hypothalamus, thalamus, claustrum & insula, retrosplenium, ventral visual stream, and inferior frontal and parieto-temporal areas (n = 4 control cases)	Receptor, transporter, and metabolic enzyme genes for major neurotransmitter systems and the neuropeptide system (176 genes in cortex; 88 genes in subcortex)
Schizophrenia Study²	Compare controls and schizophrenics for differences in expression level or pattern in a brain region postulated to be affected in schizophrenia	Dorsolateral prefrontal cortex from control and schizophrenia cases, (n = 33 and 19, respectively)	Schizophrenia candidate genes culled from the literature, and cell-type and cortical layer markers (60 genes)
Subcortex Study	Characterize expression of neurotransmitter system genes in subcortical brain regions	Subcortical regions from front of caudate to just posterior to substantia nigra (left hemisphere); hypothalamus from right hemisphere (n = 2)	Neurotransmitter, receptor and catabolic and metabolic enzyme genes for GABA and glutamate neurotransmitter system (55) in subcortical regions; anatomic marker genes in hypothalamic tissue (10)

¹Collaboration with Dr. Eric Courchesne, University of California, San Diego.

²Collaboration with Drs. Thomas Hyde and Joel Kleinman, Section on Neuropathology, Clinical Disorders Branch, GCAP, IRP, National Institute of Mental Health, NIH.

- (1) The 1,000 Gene Survey in Cortex characterized approximately 1,000 genes in visual cortex (VC) and temporal cortex (TC) of multiple control cases with n = 2 – 6 cases characterized per gene. Genes in this dataset represented several categories of broad scientific and clinical interest: cortical cell type markers, gene families important to neural function, disease-related genes and genes important in the comparative genomics field.
- (2) The Autism Study systematically examined neocortical architecture during the early years after autism onset using *in situ* hybridization with a unique panel of lamina- and cell type- specific molecular markers to phenotype cortical microstructure. Markers for neurons, glial and risk genes for autism were tested on frontal, temporal, and occipital postmortem brain tissue from young postmortem control and autism cases. Initially, 64 candidate genes were screened to select the 25 genes chosen for analysis in 11 autistic and 11 control subjects. Staining for Nissl bodies was included for gross anatomical and cellular architectural analysis. This study was done in collaboration with Dr. Eric Courchesne, UC San Diego.
- (3) The Neurotransmitter Study surveys expression of ion channel, GPCR, transporter, and synthetic and catabolic enzyme genes for the major neurotransmitter systems, acetylcholine, dopamine, epinephrine/norepinephrine, gamma amino-butyric acid (GABA), glutamate, and serotonin, as well as a

limited number of representative genes from the adenosine, cannabinoid, glycine, histamine and neuropeptide systems. The dataset contains 176 genes characterized in dorsolateral prefrontal cortex, ventral prefrontal cortex, anterior & posterior cingulate, claustrum & insula, retrosplenium, ventral visual stream, inferior frontal, and parieto-temporal areas. Inferior frontal and parieto-temporal areas were included as right hemisphere homologues to Broca's and Wernicke's areas. The dataset also includes 88 genes (mostly overlapping with the 176-gene list) in amygdala, hippocampus, hypothalamus, and thalamus. A lower number of genes allowed for a higher resolution per gene (each gene repeated every 2 mm of tissue) in these regions where structural changes -- and potentially gene expression changes -- occur within a shorter anterior to posterior distance).

- (4) The Schizophrenia Study examined 60 genes in dorsolateral prefrontal cortex (DLPFC) of over 50 control and schizophrenia (SCZ) cases. This study was done in collaboration with Drs. Thomas Hyde and Joel Kleinman, Section on Neuropathology, Clinical Disorders Branch, GCAP, IRP, National Institute of Mental Health, NIH. Genes in this dataset included cell-type markers, cortical layer-specific markers and SCZ candidate genes culled from the literature. In addition to allowing gene expression comparisons between control and SCZ, the high number of individuals characterized in this dataset also provided a dataset to gauge the range of gene expression variation within each population.
- (5) The Subcortex Study characterized a set of 55 genes in subcortical regions extending from the front of the caudate through posterior substantia nigra and a second smaller set of 10 genes through the hypothalamus. The 55-gene set focused on genes encoding glutamatergic and GABAergic neurotransmitters, receptors and receptor subunits, and associated catabolic and metabolic enzymes. The 10-gene set included genes found to be useful markers for anatomic delineation, and was characterized at a relatively high sampling density in the hypothalamus. Staining for Nissl bodies, acetylcholinesterase (AChE), and cytochrome oxidase (CyOx) was included to provide anatomical context.

POSTMORTEM HUMAN TISSUE

1,000 Gene Survey in Cortex and Schizophrenia Study

Subjects

Frozen postmortem tissue samples from adult male and female control or schizophrenia subjects at least 20 years of age were provided by the brain tissue collection of the Section on Neuropathology, Clinical Disorders Branch, GCAP, IRP, National Institute of Mental Health, NIH, Bethesda, MD. Frozen postmortem tissue samples for control subjects for the 1,000 Gene Survey in Cortex were also obtained from the University of Miami Brain Endowment Bank, University of Miami Miller School of Medicine, Miami, FL.

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 were 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 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.

Control 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 from the control cohort. Cases in which manner of death was suicide or in which death was due to drug overdose or poisoning were also excluded. Positive toxicology was not an exclusion criterion for schizophrenic cases.

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. Brain tissue pH is one frequently used indicator of tissue quality and is suggested to be an objective measure of agonal state (Atz et al, 2007) which in turn has been reported to affect RNA quality (Hynd et al, 2003; Tomita et al, 2005). Although much of the research on pH and RNA quality has been focused on microarray analyses of postmortem tissue, gene expression detection using ISH in postmortem tissue is also likely subject to similar considerations. 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.

Protocols describing these projects were submitted by the Allen Institute and reviewed by Western Institutional Review Board's Regulatory Affairs Department, resulting in issuance of a Determination of Exemption.

Tissue Dissection

Brain tissue was harvested and sectioned into 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. Visual cortex was excised from frozen slabs, using the calcarine fissure as the primary landmark. Visual cortical samples contained Brodmann's areas 17 and 18, taken from the occipital pole and the slab just rostral to the occipital pole. The more rostral samples were preferentially used for clearer representation of cortical layers. Temporal cortex was also obtained from each subject, with dissection occurring at the level of the midbody of the hippocampus. The temporal cortical samples generally comprised Brodmann's areas 21, 22 and, in some cases, 20. As much as possible, visual and temporal cortices were obtained from the same hemisphere within each case; usually from the left hemisphere. DLPFC samples were dissected from the superior and middle prefrontal gyri, starting 1 cm caudal to the frontal pole extending to the rostrum of the corpus callosum, with the goal of including Brodmann's area 9 and Brodmann's area 46 in each tissue block. The majority of DLPFC samples were obtained from the right cerebral hemisphere.

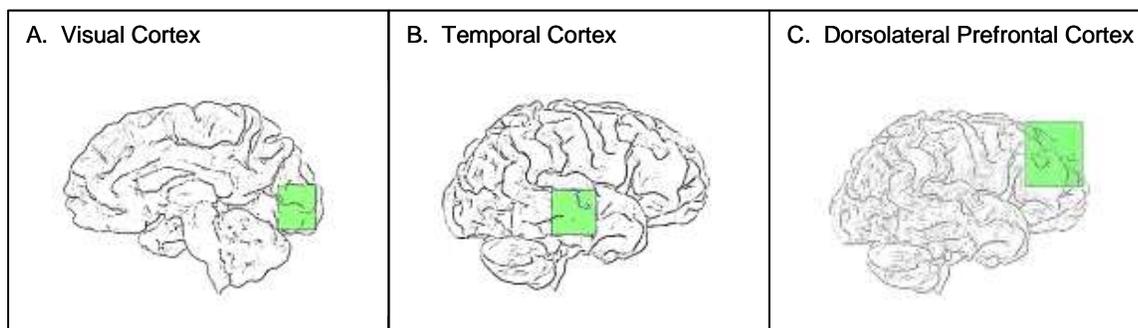


Figure 1. Schematic representations of (A) visual, (B) temporal, and (C) dorsolateral prefrontal regions excised and characterized for gene expression using colorimetric ISH. Schematic renderings are based on reference images created for the Allen Institute by Dr. Jacopo Annese, The Brain Observatory, University of California, San Diego.

Tissue Qualification

Prior to sectioning for ISH, tissue samples were tested for (a) confirmation of region of interest based on expected cytoarchitecture, (b) RNA quality, (c) absence of senile plaques and (d) absence of severe ice crystals. Tissue samples that failed these criteria were not further processed.

Two sample sections were stained for Nissl bodies to confirm region of interest based on expected cytoarchitectural features. In the DLPFC, Brodmann's areas 9 and 46 were of the greatest interest and tissue containing either one or preferably both areas was specifically chosen for further analysis by ISH.

For RNA quality assessment purposes, two tissue sample sections weighing between 2 and 5 mg and no more than 10 mg each were used for RNA extraction. RNA was extracted using the MELT Total Nucleic Acid Isolation system (Ambion, Foster City, CA), per the manufacturer's protocol. Samples were stored at room temperature for up to 1 week after lysis and processed in 96-well plates to enable high-throughput processing. The MELT extraction process captures nucleic acids using magnetic beads. RNA was eluted with 20 μ l 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 μ l nuclease-free water, standardized to a concentration of 5 ng/ μ l and 1.0 μ l run on a Pico Bioanalyzer chip. RIN values were based on the entire electrophoretic trace and range from 1 to 10, where 1 corresponds to completely degraded RNA and 10 corresponds to perfectly intact RNA (Schroeder et al, 2006).

All tissue samples were assessed for senile plaques by Thioflavin S staining and for ice crystals by hematoxylin and eosin (H&E) staining. The presence of either plaques or ubiquitously distributed large ice crystals was grounds for failing tissue for further use in ISH studies.

Autism Study

Tissue acquisition and screening

Tissue for this study was provided by the National Institute of Child Health and Human Development Brain and Tissue Bank for Developmental Disorders (Baltimore, MD) under contracts N01-HD-4-3368 and N01-HD-4-3383, the Brain and Tissue Bank for Developmental Disorders (Miami, FL), Autism Tissue Program (Princeton, NJ) and Harvard Brain Tissue Resource Center (Belmont, MA). Forty-two fresh-frozen postmortem cortical tissue blocks (1-2 cc) from the superior or middle frontal gyrus of dorsolateral prefrontal cortex (DL-PFC), the posterior superior temporal cortex (pSTC), or occipital cortex (OCC, Brodmann Area 17) were obtained from young autistic and control cases aged 2-15 years. Twenty-two blocks of DL-PFC (8 males and 3 females in each group), 5 from pSTC (2 autism males and 3 control males) and 6 (3 autism males and 3 control males) from occipital cortex of autistic and control cases passed quality control measures (mean RIN +/- SD: control 7.04 +/- 1.71, autism 7.10 +/- 1.74).

Neurotransmitter Study and Subcortex Study

Tissue Acquisition and Screening

Tissue criteria for the Neurotransmitter Study and Subcortex Study were similar to those applied for the Whole Brain Microarray Survey for the Allen Human Brain Atlas. Postmortem brains from males and females between 18 and 68 years of age, with no known neuropsychiatric or neuropathological history ('control' cases) were eligible for inclusion in the study. Consent from next-of-kin was obtained in all cases. A screening process was employed to validate, as much as possible, a 'control' diagnosis, as well as to qualify the tissue for cytoarchitectural integrity and acceptable RNA quality. In addition to evidence of known history of neuropathological or neuropsychiatric disease, cases with long-term illnesses or events that resulted in hypoxic conditions lasting more than an hour were also excluded. As a safety precaution, a serology screen was performed to exclude cases with Hepatitis B, Hepatitis C, or HIV-1/HIV-2 (ViroMed, Minnetonka, MN). Blood samples from each case were submitted for toxicology screening (NMS Labs, Willow Grove, PA) to determine presence of medications that might indicate neuropathology or neuropsychiatric disorder, including substance abuse or addiction. Gross neuropathology reports were provided by radiologists based on MR data. Microneuropathology assessment was performed by pathology consultants who viewed selected tissue sections stained with hematoxylin & eosin (H&E), thionin-based Nissl stain, or Thioflavin S or silver to determine: (1) presence of local ischemic events; (2) evidence of plaques or tangles; or (3) other

cytoarchitectural abnormalities. All screening and validation data were reviewed by a Case Review Committee (CRC) consisting of Allen Institute staff and external advisors with varying expertise in imaging, genetics, and neuroanatomy.

Tissue Dissection and Freezing

Initial collection, dissection and freezing of whole brain were performed at the NICHD Brain and Tissue Bank at the University of Maryland, Baltimore, or at the University of California Irvine Psychiatry Brain Donor Program and Functional Genomics Laboratory, Department of Psychiatry and Human Behavior, resulting in frozen coronal slabs specifically cut to anatomic landmarks to preserve structural integrity of regions of interest as much as possible. Cerebellum and brainstem were removed, with the brainstem cut at the junction of the pons and medulla, leaving the pons attached to the cerebrum. A small sample of the frontal pole was taken for pH measurement. An alginate (Cavex BV, Holland) mold was made by pouring freshly mixed liquid alginate (250 g alginate powder in 2L ultra-pure water) around a plastic brain model placed ventral side down, taking care to leave the dorsal portion of the model uncovered. Once the alginate hardened, the fresh cerebral tissue was placed in the space created by the model, and cooled to 4°C. For the Subcortex Study, the cerebrum was dissected into full coronal slabs approximately 4 cm in thickness, using the mammillary bodies as the initial landmark, resulting in two slabs containing the subcortical regions of interest. One of these slabs extended from the front of the caudate to the back of the mammillary bodies while the second slab extended from back of mammillary bodies to just posterior to the SN. The two remaining slabs were a slab containing frontal pole and a slab containing occipital pole. For the Neurotransmitter Study, the cerebrum was hemisected and the right hemisphere cut into coronal slabs, using the mammillary bodies and the genu and splenium of the corpus callosum as primary anatomic landmarks. Additional coronal cuts were made at 1 – 1.5 cm intervals from the frontal and occipital poles. Slabs were immediately frozen in a bath of dry ice and isopentane and stored at -80°C until shipping via overnight courier.

Upon receipt of tissue, slabs were photographed and inspected, then placed in -80°C until future use. Photographic images taken at the time of slabbing and tissue receipt were reviewed to assess gross brain morphology and anatomic landmarks. RNA quality assessment was performed on RNA extracted from tissue using TRI Reagent and magnetic bead extraction, followed by RIN assessment on a Bioanalyzer. The Case Review Committee reviewed all tissue screening data as it became available and approved or rejected a case for further processing.

For the Subcortex Study, slabs from approved cases containing subcortical regions were further divided at the midline and dissected into a 4 cm x 4 cm block extending from the caudate ventrally to mid-amygdala and from midline lateral to insula. For the Neurotransmitter Study, slabs were divided into approximately 4.2 cm x 3.7 cm blocks, each blocked to capture at least one primary structure of interest.

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 or 25 µm (Schizophrenia Study) thickness in the coronal plane from anterior to posterior. One section was placed on each positively charged Superfrost Plus™ 1" x 3" or 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. Samples too large to fit on slides were trimmed prior to sectioning.

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). 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 use.

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.

Cytochrome Oxidase (CyTox)

The method for CyTox staining was adapted from Keirnan, 2008. Slides were removed from the freezer and incubated in chilled acetone and chilled 0.5% glutaraldehyde for 5 min in each solution, washed 3x in 10% sucrose in 0.1M NaPO₄ at 42°C, incubated in Tris-CoCl₂ for 10 min and returned to 0.1M NaPO₄ buffer. Slides were then incubated with equine Cytochrome C in DAB solution for 90 min at 37°C, fixed in 10% NBF, and dehydrated in a graded series of 50%, 70%, 95% and 100% ethanol. Slides were incubated in Formula 83 and coverslipped with the mounting agent DPX.

Hematoxylin and Eosin (H&E)

To assess ice crystal presence and severity, a representative section from each tissue sample was subjected to a regressive H&E stain (Culling et al, 1985). After sectioning, tissue was fixed with neutral buffered formaldehyde or fixed, acetylated and dehydrated as described above. Slides were processed either manually or with autostainers depending on the required throughput. Sections were stained first with commercially prepared Harris hematoxylin, differentiated in 1% HCl in 70% ethanol, blued with 1% lithium carbonate and stained in 1% eosin Y in 1% aqueous calcium chloride. Sections were then dehydrated in a graded series of 70%, 95% and 100% ethanol, cleared in xylene and coverslipped with either DPX or CureMount mounting media.

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.

Thioflavin S

Thioflavin S staining protocols were based on previously described methods (Sun et al, 2002). Slides were fixed, acetylated and dehydrated and were stained for 7 minutes with a mixture of Thioflavin S (Sigma; 0.125 mg/mL) and Sytox Orange (Invitrogen; 0.1 µl/mL) in 50% EtOH. Slides were rinsed in fresh 50% ethanol and water, coverslipped using Hydromatrix and viewed using a fluorescence microscope to visualize stained plaques (Thioflavin S) and cell nuclei and neurites (Sytox Orange).

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 μ M, and designed with GC content between 42% – 62 % and an optimal size of 22nt with lower and upper limits of 18nt and 26nt, 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 30 ng/ μ l 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.

High-Throughput In Situ Hybridization

ISH processes were performed on robotic platforms that allowed processing of approximately 100 2" x 3" or 200 1" x 2" slides at a time. 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 to accommodate 2" x 3" slides 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 (TSA™, Perkin Elmer, Waltham, MA) 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. For the Neurotransmitter Study, the protocol was further modified to increase sensitivity by using TSA™ Plus Biotin (Perkin Elmer) in place of TSA and changing the anti-DIG-POD antibody concentration to 0.10 U/ml.

To reduce background signal, an acid alcohol wash step was performed 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 to ensure that all slides were rinsed in clean solution.

Each ISH run included several controls. For all studies, a *Drd1a* positive control in mouse tissue was used to provide verification of a successful ISH run, and a negative control (no probe) in mouse tissue was included as an indication of background for each ISH run. Positive controls in human tissue were also included, with slight variations depending on the study (see Table 2).

Table 2. Summary of human gene probes used as run controls for each ISH study.

Study	ISH Controls
1,000 Gene Survey in Cortex	GAP 43, CTNND2 PCP4 in visual cortex CART in temporal cortex
Autism Study	GAP43, CTNND2
Neurotransmitter Study	GAP43, CALB1
Schizophrenia Study	GAP43, CTNND2
Subcortex Study	PCP4, CTNND2

Human GAP43 was included in almost all studies to gauge signal from a probe with typically high expression. CTNND2 and CALB1 were used to gauge signal of a probe with typically moderate expression. In the 1,000 Gene Survey, human PVALB was originally included as a control gene but was discontinued because it was less informative as a control gene than initially anticipated. PCP4 and CART were included as regional indicators for visual cortex (VC) and temporal cortex (TC), respectively. The PCP4 expression pattern delineates the boundary of Brodmann's area 17 and Brodmann's area 18. CART was selected for detection of Brodmann's area 21 boundaries in TC based on prior literature (Hurd and Fagergren, 2000) but validation of this gene as a marker for this region was difficult due to by the subtle cytoarchitectural boundaries between Brodmann's areas 21, 22 and 20. No ISH-based regional indicator was used for DLPFC.

IMAGE ACQUISITION

Two image acquisition platforms were used for scanning and digitizing human cortex data. Initially, images were generated using a previously described (Slaughterbeck et al, 2007) high-throughput automated microscopy platform developed at the Allen Institute for Brain Science. Each stand-alone image capture system (ICS) unit consisted of a Leica DM6000B automated brightfield microscope, Leica DC500 camera, Ludl BioPrecision stage with automated slide loader, Microscan CCD-3 barcode imager, and an HP wx6200 dual-processor workstation. Custom ScopeController software combined with ImagePro Plus commercial software handled all system integration, configuration and image acquisition processes. Following transfer from 25-slide cassette holders to a microscope stage, each slide was barcode scanned for data tracking

purposes, and a color, white balance and background correction check was performed. For image acquisition, each slide was initially pre-scanned with a 1.25x objective and tissue boundaries defined. The slide was subsequently scanned using a 10x objective for higher resolution. Five autofocus points were required to maintain focus across the human tissue section. Four of these points were located at the top, bottom, left and right edges of the tissue border determined by a tissue masking algorithm. The fifth point was located at the mathematical center of the four edge points. Image resolution of this system is 1.05 $\mu\text{m}/\text{pixel}$.

Most of the image acquisition for these projects was completed 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 tended to have more variation in height. The ScanScope scanner uses a 20x objective downsampled in software to minimize data volume acquired for this project. The downsampling provides similar image resolution (1.00 $\mu\text{m}/\text{pixel}$) to the ICS scanning systems.

DATA PROCESSING

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), with slight modifications for processing human tissue images.

IDP Cluster Computation Requirements

One major challenge for the processing pipeline was the large image size that resulted from a human tissue section. To support the processing of human cortex tissue images generated by the ICS or by the ScanScope, the informatics processing platform was migrated to a 64-bit Linux platform, including cluster hardware, system software and IDP applications. Cluster blades were configured to operate in 64-bit mode with at least 8GB of main memory each to provide dedicated blades with 8 – 14GB of working memory to execute the processing modules.

IDP Processing Modules

Three modules constituted the processing pipeline for human cortex images: image preprocessing, ISH expression detection and Nissl processing. In image preprocessing, scanned ISH and Nissl images were converted and processed to provide more consistent white background intensities and orientation across samples. The preprocessing steps differed slightly for images acquired using the ICS vs. the Aperio system. Specifically, ICS images were first stitched from tiles in Tiff format, white balanced and finally converted to JPEG 2000 file format. Aperio ScanScope images were first converted to JPEG 2000 format, then orientation adjusted and white balanced. In either case, the final products were images in JPEG 2000 compressed format for further pipeline processing and analysis.

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 expressers (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 (Lein et al., 2007; 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 and CyTox stained subcortical tissue sections were preprocessed for white balance for online presentation without additional analysis.

QUALITY CONTROL AND PUBLIC DISPLAY

After image acquisition, and during the course of data processing, two additional quality control steps occurred. Once image preprocessing was complete, image quality control ensured focus criteria were met 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 proceeded through the IDP for ISH expression detection.

A second phase of quality control comprised verification of anatomic region and verification that gene expression data were not obscured or adversely affected by technical or tissue artifacts. Images that met anatomic and expression detection criteria were passed for public release. For each set of gene images available in the online viewer, the nearest set of Nissl-stained sections (and other histological data, when available) can be accessed and viewed.

REFERENCES

- Atz M, et al (2007) Methodological considerations for gene expression profiling of human brain. *J Neurosci Methods* 163:295-309.
- Culling CFA, Allison RT, Barr WT (1985) Haemotoxylin and its counterstains. In: *Cellular Pathology Technique, 4th Edition*, pp 160-161. Butterworths, London.
- 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, Kleinman JE (2005) Reliability of psychiatric diagnosis in postmortem research. *Biol Psychiatry* 57:96-101.
- Hurd YL, Fagergren P (2000) Human cocaine- and amphetamine-regulated transcript (CART) mRNA is highly expressed in limbic- and sensory-related brain regions. *J Comp Neurol* 425:583-98.
- Hynd MR, Lewohl JM, Scott HL, Dodd PR (2003) Biochemical and molecular studies using human autopsy brain tissue. *J Neurochem*. 85:543-562.
- 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).

Slaughterbeck CR, et al (2007) High-throughput automated microscopy platform for the Allen Brain Atlas. *J Assoc Lab Automation* 12:377-383.

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 microarray analyses of postmortem human brain. *Biol Psychiatry* 55:346-52.