TECHNICAL WHITE PAPER:
LASER MICRODISSECTION AND MICROARRAY PROFILING OF HUMAN PRENATAL BRAIN DEVELOPMENT

OVERVIEW

The BrainSpan 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 an ARRA-funded grant through the National Institutes of Health to a consortium consisting of the Allen Institute for Brain Science; Yale University; the University of Southern California; Massachusetts General Hospital, Harvard-MIT Health Sciences and Technology, Athinoula A. Martinos Center for Biomedical Imaging; the University of California, Los Angeles; and the University of Texas Southwestern Medical Center with strong collaborative support from the Genes, Cognition and Psychosis Program, which is part of the Intramural Research Program of NIMH, NIH. All data are publicly accessible via the Allen Brain Atlas data portal at www.brain-map.org or directly at www.brainspan.org.

A major component of BrainSpan is genome-wide transcriptional profiling aimed at the identification of transcriptional programs differentially active at different stages of brain maturation throughout development. One data modality included in this atlas is laser microdissection and subsequent microarray profiling of finely dissected tissue samples from subdivisions throughout the prenatal brain (mid-prenatal stage, 19-24 post-conception weeks (pcw)). This white paper describes the methods and processes used to generate this gene expression data. The methods for generating other data types included in BrainSpan are described in separate technical white papers accessed via the Documentation tab in the online atlas.

SAMPLE ISOLATION AND MICROARRAY DATA GENERATION

This work used post-mortem human brain specimens that were procured from the Birth Defects Research Laboratory at the University of Washington and Advanced Bioscience Resources Incorporated. All work was performed according to guidelines for the research use of human brain tissue and with approval by the Human Investigation Committees and Institutional Ethics Committees of each institute from which samples were obtained. Appropriate written informed consent was obtained and all available non-identifying information was recorded for each sample. A protocol for this project was reviewed by Western Institutional Review Board, resulting in issuance of a Determination of Exemption.

Brain tissue 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 tissue qualification and sectioning. **Table 1** lists the profiles of each donor.

Table 1. Donor profiles.

Specimen ID	Age (pcw)	Sex	Race	Average RIN	Hemisphere
H376.IV.02	21 4/7	F	AA	8.5	Left
H376.IV.03	21	F	Asian	9	Left
H376.IIIA.02	15	M	Caucasian	8.2	Left
H376.IIIB.02	16	F	Asian	8.2	Left

Abbreviations: pcw: post-conception weeks, AA: African American, RIN: RNA integrity number.

Tissue Qualification

Tissue samples were tested for RNA quality. For RNA quality assessment purposes, two to six tissue sample sections at 14-20 µm each were used for RNA extraction using the MELT Total Nucleic Acid Isolation System (Ambion, Foster City, CA). Samples were stored at -80C after lysis until processed in 96-well plates for RNA isolation. RNA was eluted with 20 µl nuclease-free water and stored at -80°C. Yields were typically between 1000-3000 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, CA) 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). In addition to RIN assessment, samples were tested for tissue quality by Nissl staining (described in detail below).

Laser microdissection and RNA isolation

An entire hemisphere (n = 2) was selected for thin sectioning and laser microdissection (LMD). For the 19-24 pcw specimens, every other tissue section was collected. Frozen tissue sections were cryosectioned at 14 µm onto polyethylene naphthalate (PEN) slides (Leica Microsystems, Inc., Bannockburn, IL) and a 1:10 Nissl series, 1:10 acetylcholinesterase series and a 1:10 GAP43 (growth associated protein 43) *in situ* hybridization series was generated for neuroanatomical reference (**Figure 1**).

After drying for 30 minutes at room temperature, PEN slides were frozen at -80°C. PEN slides were lightly Nissl stained with cresyl violet to allow cytoarchitectural visualization. Slides were fixed in ice-cold 70% ethanol for 30 seconds, washed 15 seconds in nuclease-free water, stained in 0.7% cresyl violet in 0.05% NaOAc, pH 3.4 for 2 minutes, nuclease-free water for 15 seconds, followed by 15 seconds each in 50%, 75%, and 95% ethanol, followed by 20 seconds in 100% ethanol, and then a final 100% ethanol wash for 25 seconds. Slides were air-dried for 2 minutes, and dessicated in a vacuum for 1 hour, then frozen at -80°C until needed for microdissection. Laser microdissection was performed on a Leica LMD6000 (Leica Microsystems, Inc., Bannockburn, IL). Each Leica LMD6000 system included an upright research microscope fitted with a diode laser and a CCD camera to acquire live images of slides. The scope and laser were controlled via a dedicated computer running Leica LMD software (v.6.6.2.3552). Neuroanatomists drew structural boundaries on Nissl images to provide an LMD dissection guide that was also used to estimate the number and location of PEN slides required to meet minimum tissue requirements for each structure (**Figure 2**). **Table 2** provides a detailed list of the substructures collected for gene expression analysis. Structures were sampled across as many donors as possible.

RNA was isolated for each structure following the manufacturer's directions for the RNeasy Micro kit. RNA samples were eluted in 14 µl and 1 µl was run on the Agilent 2100 Bioanalyzer (Agilent Technologies, Inc., Santa Clara, CA) using the Pico assay. Due to low sample volume and incompatibility of the eluant with the Nanodrop spectrophotometer (Thermo Scientific, Wilmington, DE), samples were quantitated using the Bioanalyzer concentration output. This was done by running a 1ng/µl RNA standard on the same Pico chip and then dividing the sample concentration output by the output of the standard concentration. The average RNA Integrity Number (RIN) of all passed samples was 6.3. Samples were failed when the Bioanalyzer traces showed degraded 18S and 28S bands, with RINs typically lower than 5 failing. In most cases, 5 ng of total RNA was used as the input amount for the labeling reaction.

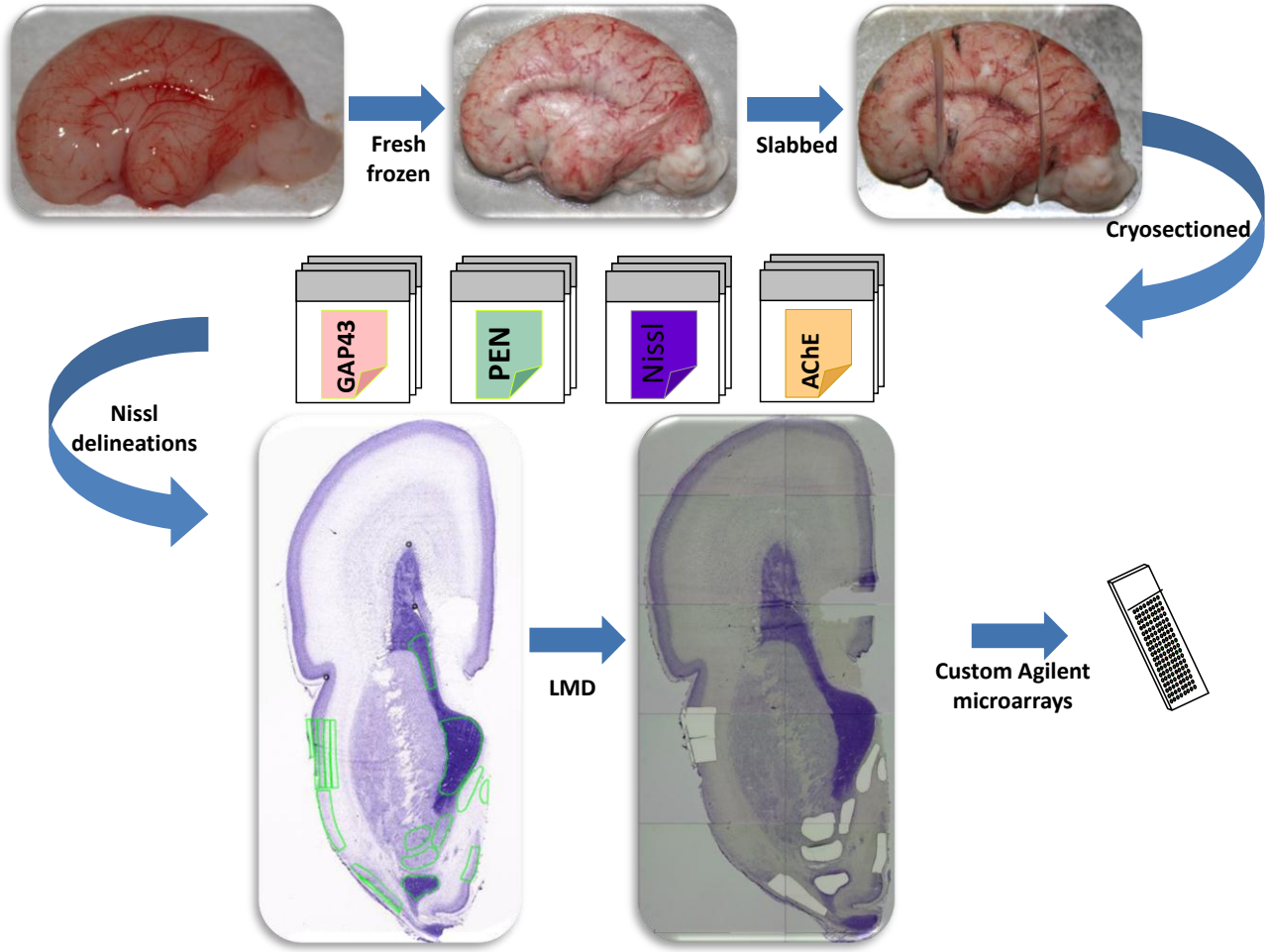

Figure 1. Overview of sample isolation process. Prenatal tissue was received, then subsequently frozen and slabbed into 1.5 to 2 cm thick coronal slabs. Each slab was cryosectioned to produce a series of slides: Nissl, PEN membrane slides (for LMD), acetylcholinesterase (AChE), and GAP43 *in situ* hybridization (GAP43). Neuroanatomical delineations were drawn on the Nissl images. These delineations were assisted by the GAP43 and acetylcholinesterase staining on adjacent tissue sections. The annotated Nissl images were used to guide the isolation of the structures on the PEN slides. RNA was isolated from the extracted tissue and then prepared for hybridization to custom Agilent microarrays.

Table 2. Substructures collected for gene expression analysis.

	Structure	Structure
Neocortex	orbital frontal	posterodorsal parietal, posteroventral parietal, dorsalmedial parietal
	posteror frontal (motor)	primary auditory
	dorsolateral prefrontal, ventrolateral prefrontal	superolateral temporal, inferolateral temporal
	frontal pole	midinferior temporal (36)
	visual (V1)	posterior parahippocampal
	midlateral extrastriate (19)	medial temporal-occipital, lateral temporal-occipital
	dorsalmedial extrastriate (V2), ventralmedial extrastriate (VP)	rostral cingulate, midcingulate, caudal cingulate
	primary somatosensory	dysgranular insular, granular insular
Transitory	lateral ganglionic eminence	caudal ganglionic eminence
	medial ganglionic eminence	rostral migratory stream
Hippocampus	stratum pyramidale, stratum radiatum, stratum oriens of rostral CA1	stratum pyramidale of caudal CA2
	stratum pyramidale of rostral CA2	stratum pyramidale of caudal CA3
	stratum pyramidale of rostral CA3	pyramidal cells of caudal CA4
	pyramidal cells of rostral CA4	granular layer, subgranular zone, polyform layer of rostral dentate gyrus
	stratum pyramidale, stratum radiatum, stratum oriens of caudal CA1	granular layer, subgranular zone, polyform layer of caudal dentate gyrus
AlloCortex	pyramidal layer of rostral subiculum	caudal part of entorhinal cortex
	pyramidal layer of caudal subiculum	rostral subdivision of area 35
	parasubiculum	caudal subdivision of area 35
	rostral presubiculum	retrosplenial cortex
	caudal presubiculum (postsubiculum)	frontal agranular insular cortex (area FI)
	rostral part of entorhinal cortex	primary olfactory (piriform)
Basal Ganglia	head of caudate, body of caudate	nucleus accumbens (shell, medial)
	claustrum, dorsal	rostral putamen
	nucleus accumbens (core)	external segment of globus pallidus
	nucleus accumbens (shell, lateral) 1	internal segment of globus pallidus
Basal Forebrain	nucleus of diagonal band	basal nucleus of Meynert
	lateral septal nucleus	bed nuclei of stria terminalis
Amygdala	rostral subdivision of medial nucleus	basomedial
	caudal subdivision of medial nucleus	anterior cortical nucleus
	central part of lateral subdivision of central nucleus	intercalated nucleus of amygdala
	lateral nucleus	ventricular zone in postamygdaloid region
	basolateral nucleus	
Thalamus	anteroventral nucleus of thalamus	pulvinar of thalamus
	ventral anterior nucleus of thalamus	lateral posterior nucleus of thalamus
	ventral lateral nucleus of thalamus	ventral posterior lateral nucleus
	lateral geniculate nuclei	ventral posterior medial nucleus
	medial geniculate nuclei	habenular nuclei
	centromedian nucleus of thalamus	zona incerta
	parafascicular nucleus of thalamus	reticular nucleus of thalamus
	lateral dorsal	subthalamus
	mediodorsal nucleus of thalamus	
Hypothalamus	dorsomedial hypothalamic nucleus	paraventricular nucleus of hypothalamus
	ventromedial hypothalamic nucleus	lateral hypothalamic area
	medial mammillary nucleus	
Midbrain	red nucleus	inferior colliculus
	substantia nigra reticulata, substantia nigra compacta	superior colliculus
	periaqueductal gray substance	oculomotor nuclear complex

Gene expression profiles were obtained from the structures listed above. For cortical regions, an attempt was made to capture 9 layers. The 9 layers from superficial to deep are 1) subpial granular zone, 2) marginal zone, 3) cortical plate outer, 4) cortical plate inner, 5) subplate zone, 6) intermediate zone, 7) subventricular zone outer, 8) subventricular zone inner, and 9) ventricular zone. Other structures not listed in this table were also captured.

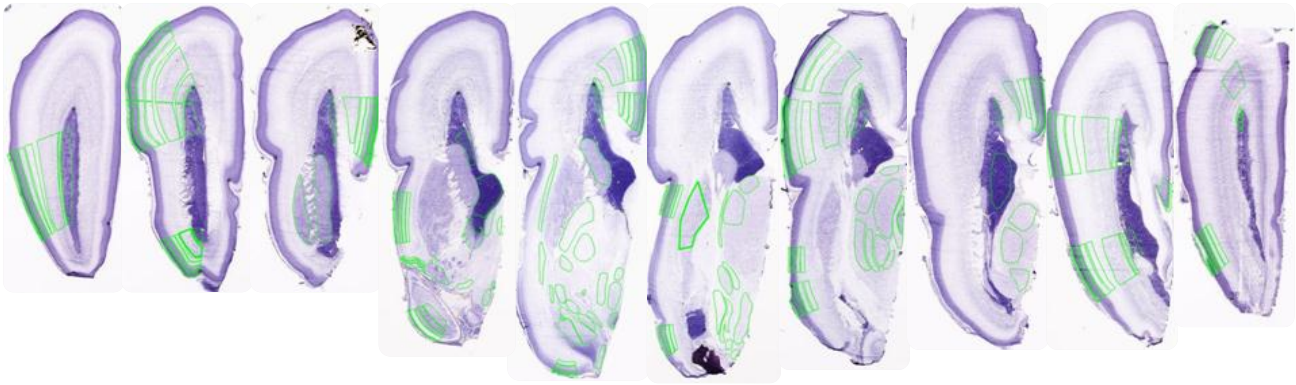

Figure 2. Panel of neuroanatomical delineations spanning tissue sections throughout a prenatal specimen. Coronal Nissl sections with regions destined for laser microdissection are delineated in green. Tissue within the green boundary is isolated on adjacent PEN membrane slides.

Nissl

For the Nissl neuroanatomical reference slides, slides were stored at 37°C for 1-5 days 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 with 0.213% thionin, then differentiated and dehydrated in water and a graded series containing 50%, 70%, 95% and 100% ethanol. Finally, slides were incubated in xylene or Formula 83, and coverslipped with the mounting agent DPX. After drying, slides were cleaned prior to digital imaging.

Acetylcholinesterase

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% ethanol, cleared with Formula 83 and coverslipped with DPX.

In situ hybridization

For the GAP43 (growth associated protein 43) *in situ* hybridization slides, details of the *in situ* hybridization process are available in the *In Situ* Hybridization White Paper under the Documentation tab in the online atlas. In brief, a colorimetric, digoxigenin-based method for labeling target mRNA was used to detect GAP43 gene expression on human prenatal tissue sections.

mRNA Profiling

Samples passing RNA quality control (QC) were amplified and profiled. Briefly, 5 ng of total RNA was the input for amplification using a modified version of Ambion's MessageAmp aRNA kit. Following the first round amplification, 400 ng was input into a second round amplification, using a modified version of Ambion's MessageAmp aaRNA Kit. 5 µg of second round product was then used as input into a Cy3-coupling reaction. Typical yields ranged from 2,000 ng to 5,000 ng of Cy3-coupled product, with Cy3 specific activity ranging between 8 and 14. Product size is typically 3,000 bases as determined by Bioanalyzer Nanochips. 600ng of Cy3-coupled product was then fragmented and hybridized to the Agilent 8X60K array according to manufacturer's protocols.

The cRNA goes through a quality control step in which the Bioanalyzer trace of the Cy3-coupled product was analyzed before fragmentation and each amplified product was scored for Cy3 incorporation, yield, and cRNA size spread. Samples were failed when determined to be outliers in any of the above metrics. Failed samples were typically re-amplified using another RNA aliquot. Each RNA sample was spiked with a specific combination of 2 ERCC transcripts that correspond to 2 of the ERCC probes on the array. In this way, the samples were barcoded molecularly and ERCC probe signal was used to ensure sample integrity during array processing.

Hybridization was to an Agilent 8x60K array, custom-designed by Beckman Coulter Genomics in conjunction with the Allen Institute that was specifically designed for the Allen Human Brain Atlas project (http://human.brain-map.org/). The array design included the existing 4x44K Agilent Whole Human Genome probe set supplemented with an additional 16,000 probes. At least two different probes were available for 93% of genes with Entrez Gene IDs (21,245 genes). Probes were located on different exons as much as possible when multiple probes were available for a gene. Other probes on the microarray were for transcripts with UCSC IDs (1,852 transcripts) and Agilent IDs (1,268 transcripts). A small set of probes mapped to contigs (253) and not to any well-defined transcript. An additional set of probes were included to overlap with the 1,000- and 60-gene sets that were characterized by ISH for the 1,000 Gene Survey in Cortex and the Subcortex Study, respectively, both of which are integrated into the Allen Human Brain Atlas. Finally, 40 probes (5 replicates per probe) specific to the ERCC transcripts were included for the sample tracking molecular barcodes. Details about Affymetrix terminology and definitions for exon probe set annotations and transcript cluster groupings can be found in Affymetrix technical white papers.

For each specimen, two batches of RNA samples (ranging from approximately 125 samples to 196 samples) were processed for hybridization. Each batch of RNA also included 5% technical replicates. Labeling and scanning were completed following the manufacturer's recommendations and profiles were normalized using COMBAT within each batch. Sample amplification, labeling, and microarray processing were performed by Covance in Seattle, WA.

Data Processing

Upon receipt, all microarray data was subjected to a standard set of QC steps that included recording and assessment of: (1) the 99% non-control probe signal, (2) visual inspection of array thumbnails, (3) %CV of non-control probes, and (4) outlier detection. All metrics used in (1, 2, 3) are available in Agilent's Feature Extraction (FE) output. Outliers were identified by examining the hierarchical clustering results to see whether samples from the related structures were grouped together. Also calculated was Inter-Array Connectivity (IAC), which was used to numerically measure the dis-similarity among the arrays, which was used to identify outliers.

ERCC transcripts were used to detect possible errors in sample handling related to mixing samples in wells or inadvertently loading an incorrect sample onto a particular array. Microarray data for ERCC specific probes were assessed using a heat map to track placement of samples. The quality control process included confirming that only the expected spiked transcripts showed signal above background. Samples were failed if they did not contain exclusively the ERCC transcripts spiked into the RNA well. **Figure 3** below shows a heat map in which an example cross-contamination event was detected.

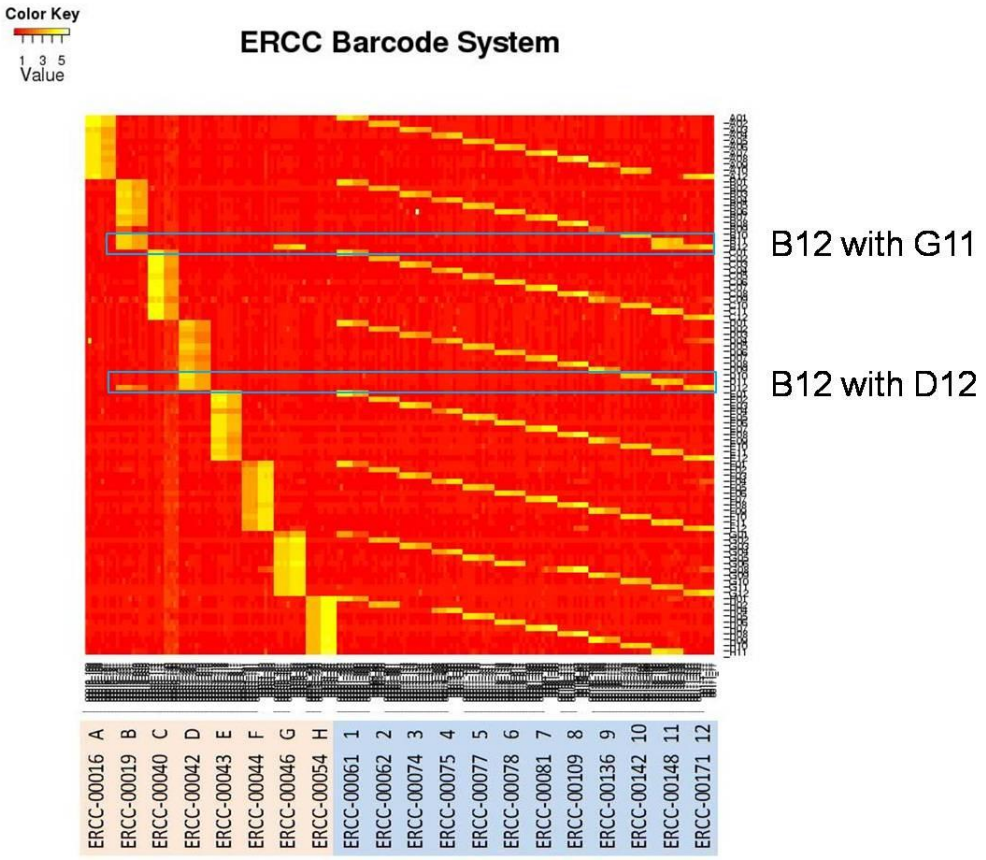

Figure 3. ERCC molecular barcode system for sample tracking throughout the hybridization process. The 20 ERCC transcripts are listed along the horizontal axis. Well positions on a 96-well plate are listed on the right of the heat map. Highly expressed probes in each sample identify the sample's position in the 96-well plate specified by row (A-H) and column (1-12). Sample swaps or cross-contamination events are detected by the presence of signal in an unexpected column or row. This heatmap shows an example cross-contamination event that would be detected by the ERCC transcripts.

Gene expression data for samples passing quality control were normalized in three steps. First, "within-batch" normalization was performed on samples of each batch using a 75% centering algorithm. Expression distributions of all samples in a single batch were normalized to have the same 75th percentile expression values. Second, across all batches of a single brain, "cross-batch" normalization was done to reduce the systematic batch bias by applying the ComBat (http://statistics.byu.edu/johnson/ComBat/) method. After these two steps, microarray data was normalized brain by brain. Raw expression values, normalized expression values, and present/absent calls were all uploaded into the database for further data analyses and data visualization.

REFERENCES

Johnson WE, Li C, Rabinovic A (2007) Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostatistics* 8(1):118-127 (PMID:16632515).

Karnovsky MJ, Roots L (1964) A "Direct Coloring" Thiocholine Method for Cholinesterases. *J Histochem Cytochem* 12:219-221.

Schroeder A, Mueller O, Stocker S, Salowsky R, Leiber M, Gassman M, Lightfoot S, Menzel W, Granzow M, Ragg T (2006) The RIN: an RNA integrity number for assigning integrity values to RNA measurements. *BMC Mol Biol* 7, 3.