### Original Research Article

### Characterization of Human Cortical Gene Expression in Relation to Glucose Utilization

KIRSTIN N. STERNER,<sup>1‡</sup> MICHAEL R. MCGOWEN,<sup>1</sup> HARRY T. CHUGANI,<sup>2,3</sup> ADI L. TARCA,<sup>4</sup> CHET C. SHERWOOD,<sup>5</sup> PATRICK R. HOF,<sup>6</sup> CHRISTOPHER W. KUZAWA,<sup>7</sup> AMY M. BODDY,<sup>1</sup> RYAN L. RAAUM,<sup>8</sup> AMY WECKLE<sup>1</sup> LEONARD LIPOVICH,<sup>1</sup> LAWRENCE I. GROSSMAN,<sup>1</sup> MONICA UDDIN,<sup>1</sup> MORRIS GOODMAN,<sup>1,9†</sup> AND DEREK E. WILDMAN<sup>1,10\*</sup> <sup>1</sup>Center for Molecular Medicine and Genetics, Wayne State University School of Medicine, Detroit, Michigan 48201 <sup>2</sup>Department of Pediatrics, Wayne State University School of Medicine, Detroit, Michigan 48201 <sup>3</sup>Department of Neurology, Wayne State University School of Medicine, Detroit, Michigan 48201 <sup>4</sup>Department of Computer Science, Wayne State University, Detroit, Michigan 48202 <sup>5</sup>Department of Anthropology, The George Washington University, Washington, DC 20052 <sup>6</sup>Fishberg Department of Neuroscience and Friedman Brian Institute, Mount Sinai School of Medicine, New York, New York 10029 <sup>7</sup>Department of Anthropology, Northwestern University, Evanston, Illinois 60208 <sup>8</sup>Department of Anthropology, Lehman College and The Graduate Center, City University of New York, Bronx, New York, 10468 <sup>9</sup>Department of Anatomy and Cell Biology, Wayne State University School of Medicine, Detroit, Michigan 48201 <sup>10</sup>Department of Obstetrics and Gynecology, Wayne State University School of Medicine, Detroit, Michigan 48201

**Objectives:** Human brain development follows a unique pattern characterized by a prolonged period of postnatal growth and reorganization, and a postnatal peak in glucose utilization. The molecular processes underlying these developmental changes are poorly characterized. The objectives of this study were to determine developmental trajectories of gene expression and to examine the evolutionary history of genes differentially expressed as a function of age.

Methods: We used microarrays to determine age-related patterns of mRNA expression in human cerebral cortical samples ranging from infancy to adulthood. In contrast to previous developmental gene expression studies of human neocortex that relied on postmortem tissue, we measured mRNA expression from the nondiseased margins of surgically resected tissue. We used regression models designed to identify transcripts that followed significant linear or curvilinear functions of age and used population genetics techniques to examine the evolution of these genes.

**Results:** We identified 40 transcripts with significant age-related trajectories in expression. Ten genes have documented roles in nervous system development and energy metabolism, others are novel candidates in brain development. Sixteen transcripts showed similar patterns of expression, characterized by decreasing expression during childhood. Comparative genomic analyses revealed that the regulatory regions of three genes have evidence of adaptive evolution in recent human evolution.

Conclusions: These findings provide evidence that a subset of genes expressed in the human cerebral cortex broadly mirror developmental patterns of cortical glucose consumption. Whether there is a causal relationship between gene expression and glucose utilization remains to be determined. Am. J. Hum. Biol. 25:418-430, 2013. © 2013 Wiley Periodicals, Inc.

### INTRODUCTION

Understanding the evolutionary origins of the enlarged human brain and human cognition is a topic of longstanding interest (Darwin, 1871; Holloway, 1968; Jerison, 1973). In comparison to other primates, the human brain is exceptionally large relative to body size (Gajdos et al., 2010; Jerison, 1973) and consumes more than 20% of the body's total energy (Choi et al., 2009; Holliday, 1986; Leonard et al., 2007). The intensive energetic demands of the human brain are highest during childhood (see Fig. 1; Chugani, 1998; Chugani et al., 1987), a pattern not observed in nonhuman primates (Jacobs et al., 1995). This period of enhanced energy metabolism likely reflects underlying molecular and cellular processes that require greater energy during childhood, including prolonged synapse proliferation (Glantz et al., 2007; Huttenlocher and Dabholkar, 1997) and neuronal plasticity (Galvan, 2010; Johnston, 2009). Anthropologists have long explored the evolutionary implications of this distinct pattern of brain development (Leigh, 2004), focusing on human-specific patterns of brain energy metabolism. For example, studies have explored the relationships between maternal energy availability and brain size (Martin, 1981; Ojeda et al., 2010), diet and brain metabolism (Aiello and Wheeler, 1995; Leonard and Robertson, 1992), body composition and brain metabolism (Kuzawa, 1998; Steinhoff and Vingron, 2006), and the relationship between adrenal steroids, cortical glucose utilization and brain development (Campbell, 2011).

To understand the complex developmental processes that underlie human brain development at a molecular level it is important to first identify which genes are transcriptionally active in the cerebral cortex during postnatal development and to characterize how expression of these genes changes during this dynamic period. These genes may play important roles in postnatal brain development

Received 17 September 2012; Revision Accepted 25 February 2013

DOI: 10.1002/ajhb.22394

Additional Supporting Information may be found in the online version of this article.

Contract grant sponsor: National Science Foundation; Contract grant numbers: BCS0827546 and BCS0827531. <sup>†</sup>Deceased.

<sup>&</sup>lt;sup>‡</sup>Kirstin N. Sterner is currently at Department of Anthropology, University of Oregon, Eugene, OR 97403

<sup>\*</sup>Correspondence to: Derek E. Wildman, Center for Molecular Medicine and Genetics, Wayne State University School of Medicine, Detroit, MI 48201, USA. É-mail: dwildman@wayne.edu

Published online 4 April 2013 in Wiley Online Library (wileyonlinelibrary.com).



Fig. 1. Ontogenetic changes in cerebral glucose consumption. Schematic view of developmental changes in glucose uptake (cortex) in humans and macaque monkeys. Shown are human lCMRglu trajectories in human frontal cortex from (Chugani et al., 1987) and macaque neocortex from (Jacobs et al., 1995). The y-axis shows cerebral glucose uptake as a percentage of adult levels and the x-axis shows age as a percentage of mature age.

and can help highlight biological pathways and molecular processes active in the brain during this time. In comparison to nonhuman primates, the neocortex of adult humans is characterized by more pronounced upregulation of genes involved in energy metabolism and neuronal activity (Caceres et al., 2003; Nowick et al., 2009; Oldham et al., 2006; Uddin et al., 2004). In addition, the human prefrontal cortex has an excess of changes in metabolite concentration profiles when compared with chimpanzee and rhesus monkey (Fu et al., 2011). Metabolites with human-specific concentration profiles are associated with long-term potentiation, neuroactive ligand-receptor interaction, alanine, aspartate, and glutamate metabolism and  $\beta$ -alanine metabolism (Fu et al., 2011).

Human brain tissues also show unique developmental patterns of gene expression. In the human prefrontal cortex, genes involved in energy metabolism, along with genes that code for proteins involved in protein and lipid synthesis, have a peak in expression during late adolescence (Harris et al., 2009). During adolescence, there is decreased expression of genes involved in glutamate and neuropeptide signaling and neuronal development (Harris et al., 2009). Most recently, Kang et al. (2011) reported extensive spatiotemporal heterogeneity in gene expression in a large number of regional brain samples selected from individuals spanning prenatal to adult development. There is also a postnatal delay in the expression of a subset of genes involved in neural development and neuronal activity in the human prefrontal cortex when compared with other primates, indicating that transcriptional neoteny may have played a role in human brain growth and development (Liu et al., 2012; Somel et al., 2009, 2011).

Although these studies have provided important insights into developmental changes in gene expression, all past work in this area has measured expression levels of mRNA derived from postmortem tissue. While these studies provide valuable insight, gene expression studies that utilize postmortem tissue may be more subject to biases associated with death (Gary et al., 2000) and postmortem mRNA degradation (Catts et al., 2005; Ferrer et al., 2008). In this study we examine changes in gene expression using brain tissue derived from living patients rather than tissues collected postmortem. By exploring age-related changes in mRNA expression in cerebral cortical tissue that was excised during surgery and immediately flash-frozen, we are able to evaluate and extend prior findings. We hypothesized that a subset of genes would follow age-related changes in expression and that these loci would follow trends consistent with ontogenetic changes in energy metabolism, synaptogenesis, and neuronal and synaptic plasticity observed at these ages. We examined these questions using samples from 32 individuals spanning in age from 10 months to 18 years, and revealed genes that follow expression trajectories that vary significantly with age.

### MATERIALS AND METHODS Sample information

Samples were collected from surgically resected tissue. The majority of the samples originated from surgeries performed to alleviate symptoms of epilepsy poorly controlled by medication. We specifically targeted tissue from the margin of the resection because this region showed normal electrophysiological activity in patients with epilepsy. Tissue samples were taken primarily from the temporal lobe (Supporting Information Table S1). All tissue samples were flash frozen in liquid nitrogen and stored at  $-70^{\circ}$ C at the time of surgery. This research was reviewed and approved by the Human Investigation Committee (HIC) at Wayne State University (HIC# 071608MP4X). Supporting Information Table S1 includes sample information (e.g., cortical region) and patient demographic data (e.g., age, race/ethnicity, sex).

### RNA extraction and microarray analysis

Tissue samples were homogenized in TRI Reagent (Applied Biosytems/Ambion, Austin, TX). RNA was extracted from the tissues using the MagMax-96 for Microarray kit (Applied Biosystems/Ambion) or the TRIzol protocol (Invitrogen, Carlsbad, CA) following the manufacturers' instructions. The RNA isolation was further purified with either the TURBO DNase treatment (Applied Biosytems/Ambion) or the RNeasy kit in conjunction with the RNase-Free DNase Set (Qiagen, Valencia, CA) according to the manufacturer's recommendations. Use of RNA of sufficient quality and quantity is key to successfully performing a microarray experiment. The concentration and quality of the DNA-free RNA isolations were determined using the Nanodrop 1000 (Thermo Scientific, Wilmington, DE) and the Agilent Bioanalyzer 2100 (Santa Clara, CA) respectively. As is customary, only samples with ABS 260/280 ratios (used as a measure of RNA purity) above 1.7 were selected for analysis (Ryan et al., 2004). Microarray analyses were conducted by the Applied Genomics Technology Center (AGTC; Wayne State University, Detroit, MI). AGTC followed Illumina's protocol in the TotalPrep-96 RNA Amplification Kit (Applied Biosytems/Ambion, Austin, TX) for labeling and the Whole-Genome Gene Expression with IntelliHyb Seal for the hybridization, wash and stain (Illumina, San Diego, CA), as previously described (Sterner et al., 2012).

### Data preprocessing

All samples were arrayed using the Illumina BeadChip platform (Illumina, San Diego, CA). Before data analysis, a number of preprocessing and normalization procedures were performed to control and correct for experimental

variation that may lead to false interpretations of the data (Steinhoff and Vingron, 2006). Illumina BeadStudio (V.3) was used to process the array images and obtain background corrected intensity values for all 48,803 probes in 40 arrays (corresponding to 37 unique brain samples [32 patients aged <18 years used in this study and 5 patients aged over 18 used in additional studies (Sterner et al., 2012)]) as well as the probe detection P values. A filtering step was used to retain only probes with a detection P value <0.1 in at least 16 of the total number of unique patient samples. There were 20,678 microarray probes left after this filtering step. These 20,678 probes were used in all subsequent analyses. To enable log transformation of the data, all values in the expression matrix were offset by adding a constant so that the smallest intensity value became 1.0. These expression values were then log<sub>2</sub>-transformed and quantile-normalized (Bolstad et al., 2003). Finally, the expression values for each of the two replicates for each sample were averaged.

Microarray image files collected for this research are MIAME compliant (Minimum Information about a Microarray Experiment (Brazma et al., 2001)) and are deposited in the National Center for Biotechnology Information's Gene Expression Omnibus (GEO) data repository under the series accession number GSE37721.

### Differential expression analysis

In order to focus on age-related changes that occur before adulthood, we reduced our dataset to include only samples derived from patients 18 years old or younger (n = 32). To test the association between gene expression levels and covariates (e.g., age), we fit linear models (MCCullagh and Nelder, 1989). Linear models are the preferred way to analyze microarray experiments (Kerr, 2003; Smyth, 2005). The normalized and log transformed expression levels of each probe were fit using a seconddegree polynomial function of age while adjusting for sex and developmental delay of the patients (see equation below). This model assumes a quadratic relationship between gene expression levels (on log scale) and age (Liu et al., 2005). We chose this model as a compromise between the need to fit nonlinear trends and the need to keep the number of parameters in the model low to ensure adequate power.

$$y_i = a_0 + a_1 \cdot AGE_i + a_2 \cdot AGE_i^2 + a_3 \cdot SEX_i + a_4 \cdot D_i + \epsilon_i$$

In the equation above, *i* denotes the *i*th of the 32 samples used for these analyses, AGE represents the age of the patient in years, SEX is a dichotomous variable coding for sex (0 = females, 1 = males), and *D* is a dichotomous variable coding for developmental delay (0 = no developmental delay, 1 = developmental delay). The *D* variable was used to characterize seven individuals in our dataset for whom "delay" was noted by their clinician. The exact nature of this delay (fine motor, gross motor, language, cognitive, social) was not given.

A *P* value based on an *F*-test was calculated for each probe representing the probability that both coefficients related to age  $(a_1 \text{ and } a_2)$  are null (Smyth, 2005). The nominal *P* values quantify the evidence of the association between the probe expression level and the age while adjusting for sex and developmental delay. The False Discovery Rate (Benjamini and Hochberg, 1995) was then calculated from the *P* values. Probes were sorted as a

function of the nominal P values and those with pFDRvalues <0.25 were retained as significantly related to age. The use of more relaxed thresholds (larger than 0.05) is not uncommon in the field (Dorval et al., 2012; Levula et al., 2012; Mittal et al., 2011).

The quadratic models that we used identify a range of distinct profile types depending on the sign and magnitude of the age coefficients for each probe. To help visualize the expression patterns, each probe was assigned to one of eight the profile classes based on the overall quadratic fit of their expression as a function of time (Supporting Information Table S2).

To rule out the possibility of a confounding effect between our primary variable of interest (age) and patient neuropathology, a linear model was used to test the association between the age of the patients and patient medication or neuropathology while adjusting for the sex of the patient and potential developmental delay. This was done using dummy variables coding 1 when a particular medication type or neuropathology was present in a patient and 0 when it was not present.

### Absolute expression level gene classification

In order to examine gene expression levels, we identified which probes have higher or lower absolute expression levels (across all ages) when compared with other probes on the array. Probes were classified as a function of their median expression in all available samples and sorted into one of four expression level categories corresponding to the quartiles in the distribution: expressed  $(MIN \le \bar{y} < Q_1)$ , moderate expression  $(Q_1 \le \bar{y} < Q_2)$ , high expression  $(Q_2 \le \bar{y} < Q_3)$  and very high expression  $(Q_3 \le \bar{y} < MAX)$ , where MIN = 5.99,  $Q_1 = 6.35$ ,  $Q_2 = 7.28$ ,  $Q_3 = 8.91$ , and MAX = 15.7.

#### Enrichment analyses of sets of genes

In addition to analyzing data on a gene-by-gene basis, we conducted 'self-contained' tests (Wang et al., 2010) of 24 sets of *a priori* candidate genes (Supporting Information Table S3) implicated in synapse organization and function or energy metabolism. These analyses allowed us to analyze expression patterns of groups of genes that may otherwise go undetected. We also tested for enrichment by expression level, age-association, and gene expression profile pattern. Details of these tests are described in the Supporting Information.

### Adaptive evolution of cis-regulatory regions

Given the high degree of protein-coding sequence similarity between humans and chimpanzees, King and Wilson (1975) suggested that cis-regulatory mutations might account for much of the phenotypic differences between these species. There is indeed increasing evidence of the central role that regulatory regions have played in human evolution (Torgerson et al., 2009). We investigated whether our observed developmental patterns of gene expression might be related to divergent evolution in noncoding DNA sequences between humans and our close ape relatives. We first downloaded sequences of noncoding regions of protein-coding genes identified as differentially expressed by our microarray study (n = 36). Sequences were compiled from 21 recently sequenced human genomes, including the human reference genome (hg19), 20 genomes available from the 1000 Genomes Project

(The 1000 Genomes Project Consortium, 2010), and for comparison, the *Pan troglodytes* reference genome (pan-Tro3). We substituted the *Gorilla gorilla* reference genome (gorGor3.1) for a gene when the chimpanzee sequences were missing or homology was unclear. We were unable to investigate two genes (*HYDIN* and *ZNF525*) due to ambiguity in homology with both *Pan* and *Gorilla*.

We first performed standard McDonald-Krietman (MK) tests (McDonald and Kreitman, 1991) to test for adaptive evolution on coding regions of differentially expressed genes. MK tests compare the number of synonymous sites that are polymorphic between species or fixed between species with fixed/polymorphic nonsynonymous sites. Each class of sites (synonymous vs. nonsynonymous; polymorphic vs. fixed) was statistically evaluated using a 2  $\times$ 2 contingency table and the  $\chi^2$  statistic. We then used an extension of the MK test that compares neighboring coding and noncoding sequences to examine adaptive evolution (Andolfatto, 2005) in four distinct regions that usually contain cis-regulatory elements: 5'-UTRs, 3'-UTRs, first introns, and promoter regions (defined here as sequences within 1,000 bp upstream from the transcription start site). In some cases (such as bidirectional promoters), we defined promoters as regions <1,000 bp between the transcription start site and coding sequence of a neighboring gene. Modified MK tests use fourfold degenerate sites of neighboring genes to estimate the neutral rate of evolution and compares these with the neighboring putatively nonneutral, noncoding regions. As in a standard MK test, each class of sites (neutral vs. nonneutral; polymorphic vs. divergent) is statistically evaluated using a 2 imes 2 contingency table and the  $\chi^2$  statistic. If P < 0.05, then the "nonneutral" region was inferred to have diverged from an expectation of neutrality. Additionally, this test generates the neutrality index (NI), defined as the ratio of polymorphic nonneutral sites to polymorphic neutral sites over the ratio of divergent nonneutral sites to divergent neutral sites. An NI >1 is interpreted as having more within-species variation than between-species variation and negative selection is inferred (Rand and Kann, 1996). If NI <1 then positive selection was inferred. In some situations, the neutrality index could not be defined due to the absence of neutral polymorphic sites; therefore, positive selection was inferred if nonneutral divergent sites were greater than non-neutral polymorphic sites. We performed all MK tests using the web application http://mkt.uab.es (Egea et al., 2008).

### RESULTS

## Genes with expression patterns significantly associated with age

A total of 40 probes met our criteria for being considered significantly associated with age (pFDR < 0.25), indicating that these transcripts are differentially expressed during human development. Of these, 36 are annotated genes (with one gene represented by two probe sets), two are undefined transcribed loci and one is a long noncoding (lnc) RNA (Table 1; see also Supporting Information Table S2 and Dataset S1; all datasets are deposited in the Dryad repository and can be found at http://dx.doi.org/10.5061/dryad.fj8dm). None of the 40 probes found to be significantly associated with age were also significantly related to sex or maturational delay. In addition, there were no

significant associations between age and patient medication or neuropathology (Dataset S2), increasing our confidence that the genes identified are genuinely associated with age. Patient race was also included in an initial analytic model but was found not to be a significant factor.

Within the 36 identified genes, the most overrepresented biological process was nervous system development (Gene Ontology biological process GO:0007399 pFDR = 0.015; Dataset S3). Therefore, this set of loci includes more genes with known roles in nervous system development than one would expect given random chance. The genes differentially expressed by age and involved in nervous system development are myosin XVI (MYO16), neuroglycan C (CSPG5), dihydropyrimidinase-like 3 (DPYSL3), SRY (sex-determining region Y)-box 11 (SOX11), glutamate receptor ionotropic N-methyl-D-aspartate 3A (GRIN3A), vang-like 2 (VANGL2), fibroblast growth factor 11 (FGF11), and doublecortin (DCX). Expression of each of these eight genes decreased from birth to early adolescence, at which point levels were maintained (Fig. 2; average inflection  $13 \pm 1.26$ years). Eight additional genes showed this same expression pattern, identifying them as additional candidates involved in nervous system development (Supporting Information Table S2 and Fig. S1).

### Synapse- and energy metabolism-related gene enrichment

Fifteen (62.5%) of the 24 sets of *a priori* candidate genes implicated in synapse organization and function and energy metabolism (Supporting Information Table S3) had median expression levels significantly higher (pFDR < 0.05) than the remaining genes on the array (Fig. 3A and Dataset S4). For example, genes involved with *mitochondrial ATP synthesis coupled electron transport* (biological process GO:0042775) had median expression level approximately ninefold higher (pFDR  $\cong$  0) than all other genes on the array not belonging to this GO term. The only group of genes from the *a priori* sets to show significantly lower expression than other genes on the array are those involved in *neurotransmitter binding* (molecular function GO:0042165; 40% lower expression pFDR = 0.03).

Five of the 24 sets of a priori candidate genes also included genes with age-related patterns of gene expression (Fig. 3B and Dataset S4; i.e., the median P value from the age association analysis was significantly lower than the median P values of all other genes on the array). As a group, genes that function in synapse organization (GO:0050808) exhibited a gene expression profile that more often showed decreased expression with increasing age than when compared with the remainder of the genes on the array [pFDR (age) = 0.02; pFDR (profile) = 0.18; oddsratio = 2.26]. Genes involved in oxidative phosphorylation [GO:0006119; pFDR (age) = 0.02; pFDR (profile) = 0.0002; odds ratio = 3.80] and mitochondrial ATP synthesis-coupled electron transport [GO:0042775; pFDR (age) = 0.04; pFDR (profile) = 0.002, odds ratio = 4.81] more often showed increased expression with increasing age. Genes involved in energy reserve metabolic process (GO:0006112) and glycogen metabolic process (GO:0005977) more often had a "u shape" (higher in youngest and oldest individuals, but lower in between) expression pattern [pFDR (age) = 0.002; pFDR (profile) = 0.004, odds ratio = 6.64; and pFDR(age) = 0.002; pFDR (profile) = 0.01, odds ratio = 6.33, respectively].

TABLE 1. Genes	differentiall	ly expressed	by age
----------------	---------------	--------------	--------

Gene symbol	Gene name	Fold change	Expression level	P value	pFDR
S100A	S100 calcium binding protein A1	2.8011	Very high	1.2e-05	0.0724
NQO1	NAD(P)H dehydrogenase, guinone 1	4.0935	Very high	0.0002	0.1836
C13 or f16	Testis expressed 29 (TEX29)	2.8631	Moderate	3.2e-05	0.0724
CYB5D1	Cytochrome b5 domain containing 1	1.5423	High	0.0001	0.1525
DDRGK1	DDRGK domain containing 1	1.6678	Very high	0.0002	0.1836
ZNF365	Zinc finger protein 365	1 9122	Very high	0.0005	0 2469
	An unspliced putative lncRNA; EST support; shares a bidirectional promoter with the gene ARPC2	1.4509	Moderate	1.9e-05	0.0724
DCUN1D4	DCN1, defective in cullin neddylation 1, domain containing 4 (Saccharomyces cerevisiae)	1.8737	Moderate	0.0003	0.1836
STAT4	Signal transducer and activator of transcription 4	3.6449	High	1.7e-05	0.0724
TPD52L1	Tumor protein D52-like 1	3.7254	High	3.7e-05	0.0763
MED7	Mediator complex subunit 7	1.4786	High	0.0001	0.1525
TPD52L1	Tumor protein D52-like 1	3 8301	High	0.0004	0 2334
MYO16	Myosin XVI	2,9766	Very high	2.3e-05	0.0724
FARP1	FERM, RhoGEF (ARHGEF) and pleckstrin domain protein 1 (chondrocyte-derived)	1.6419	Very high	5.6e-05	0.1048
CSPG5	Chondroitin sulfate proteoglycan 5 (neuroglycan C)	2.1078	Very high	0.0001	0.1525
DPYSL3	Dihydropyrimidinase-like 3	2.0387	Very high	0.0003	0.1836
CX3CL1	Fractalkine or chemokine (C-X3-C motif) ligand 1	1.7777	Very high	0.0003	0.1939
AUTS2	Autism susceptibility candidate 2	1.6693	Very high	0.0003	0.1939
FAM57B	Family with sequence similarity 57 member B	1 7776	Very high	0.0003	0 1939
MPP3	Membrane protein, palmitoylated 3 (MAGUK p55 subfamily member 3)	1.7316	Moderate	0.0003	0.1836
SOX11	SRY (sex determining region Y)-box 11	4.6756	High	3.1e-07	0.0048
GRIN3A	Glutamate receptor, ionotropic, N-methyl-D-aspar- tate 3A	2.4637	High	2.9e-05	0.0724
VANGL2	Vang-like 2 (van gogh, Drosophila)	1.6925	High	6.5e-05	0.1120
FGF11	Fibroblast growth factor 11	2.1011	High	8.7e-05	0.1390
DCX	Doublecortin	1.9803	High	0.0002	0.1836
EMID1	EMI domain containing 1	2 4357	High	0.0003	0 1939
SH3RF3	SH3 domain containing ring finger 3	1.7010	High	0.0004	0.2349
GSK3B	Glycogen synthase kinase 3 beta	1.7666	Expressed	2.7e-05	0.0724
BCAN	Brevican	1 9912	Very high	0.0001	0 1525
ANKRD57	Ankyrin reneat domain 57	2.0884	High	4 7e-07	0.0048
TMEM108	Transmembrane protein 108	1 8238	High	0.0001	0.1525
IGSF1	Immunoglobulin superfamily, member 1	2 4555	High	0.0001	0.1525
CDI V2	Comployin 2	2.4555	High	0.0001	0.1020
TNNIS	Troponin I type 2 (cordine)	2.1411	Fyprogood	0.0004	0.2345
1101015	I 1M9 marset on X sha (UniCome He 540125)	2.0072	Expressed	0.0002	0.1000
	L1M2 repeat on A cnr (UniGene fis.340133)	1.7420	Expressed	0.0002	0.1836
DI INIA	Deviliair 4	1.4981	Expressed	0.0003	0.1836
PLIN4	Perinpin 4	1.8706	Moderate	0.0003	0.2060
HIDIN	Axonemai central pair apparatus protein	1.7998	Expressed	0.0002	0.1525
ZINF525	Linc Inger protein 525	1.5949	Expressed	0.0004	0.2334
KNH1	Ribonuclease/angiogenin inhibitor 1	1.7229	High	0.0002	0.1525

#### Validation of microarray data using qPCR

We carried out quantitative reverse transcriptase realtime PCR (qPCR) validation for six genes of interest that demonstrated distinctive expression profile classes (Fig. 4 and Supporting Information Fig. S2). We chose for validation six probes annotated to genes with expression profile classes concordant with ontogenetic changes in glucose utilization (n shape; NQ01, S100A, C13orf16 and up then flat; DDRGK1, MED7, DCUN1D4). The qPCR results for NQO1, S100A, DDRGK1, and MED7 agreed with the microarray findings in terms of the shape of the expression profile and significance assessed using the same methods and significance thresholds as used in the microarray analysis (Fig. S2). For MED7 however, we observed a poor correlation between microarray and qPCR expression levels in the subset of samples profiled using both platforms (Fig. 4). Details of the qPCR validation process are described in the Supplementary Methods.

### Adaptive evolution of cis-regulatory regions

We found evidence suggestive of adaptive evolution in the human lineage in *cis*-regulatory regions of three of the

differentially expressed genes (Supporting Information Table S4). Two of these genes (BCAN and GRIN3A) showed evidence of positive selection in their promoter regions (*BCAN*,  $\chi^2 = 4.739$ , df = 1, *P* = 0.029; *GRIN3A*,  $\chi^2 = 7.866$ , df = 1, *P* = 0.005). *GRIN3A* showed evidence of positive selection in the 5' and 3' UTR regions (5'UTR,  $\chi^2 = 7.510$ , df = 1, P = 0.006; 3'UTR,  $\chi^2 = 6.639$ , df = 1, P = 0.009), and BCAN showed evidence of positive selection in its 3'UTR and first intron (3'UTR,  $\chi^2 = 6.572$ , df = 1, *P* = 0.010; intron,  $\chi^2 = 6.898$ , df = 1, P = 0.008). BCAN and GRIN3A show human-specific expression changes in the prefrontal cortex when compared with chimpanzee and rhesus macaque, which is consistent with our findings (Liu et al., 2012). Human-specific expression changes may also be explained by *trans*-factors (e.g., transcription factors or microRNAs) (Somel et al., 2011). The promoter region of ANKRD57 showed statistically significant evidence of negative selection ( $\chi^2 = 6.239$ , df = 1, P = 0.012), as evidenced by a high neutrality index (NI=15.3). This implies evolutionary pressure for conservation at ANKRD57 between humans and chimpanzees. In addition, one gene, STAT4, shows evidence of positive selection ( $\chi^2 = 5.176$ , df = 1, P = 0.022) in its coding sequence in the human lineage.



Fig. 2. Expression plots for differentially expressed genes annotated to nervous system development. The  $\log_2$  transformed and normalized mRNA expression levels as a function of age for genes annotated to the Gene Ontology (GO) term *nervous system development*. The black curve represents a quadratic fit of the expression levels on age. The *y*-axis shows log2 normalized gene expression signal intensities and the x-axis shows age in years. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

# Intersection of differentially expressed genes with knock out mouse data

Sixteen of our 36 genes have associated phenotype data derived from knock out or knock down experiments in mice (Austin et al., 2004). Disruption of 11 of these 16 genes resulted in abnormal phenotypes assigned to the following broad categories: nervous system (MP:0003631) or behavior/neurological (MP:0005386) [Mammalian Genome Informatics (Blake et al., 2009)], suggesting they are important for nervous system development and function in mice (Dataset S5). These genes included: BCAN (Brakebusch et al., 2002), CPLX3 (Reim et al., 2009; Xue et al., 2008), CSPG5 (Juttner et al., 2005), CX3CL1 (Soriano et al., 2002), DCX (Kappeler et al., 2006; Nosten-Bertrand et al., 2008; Tuy et al., 2008), GRIN3A (Brody et al., 2005; Das et al., 1998; Larsen et al., 2011; Nakanishi et al., 2009; Roberts et al., 2009),  $GSK3\beta$  (Kapfhamer et al., 2010; Kim et al., 2006), HYDIN (Davy and Robinson, 2003), *SOX11* (Lin et al., 2011; Potzner et al., 2010; Thein et al., 2010), *S100A* (Ackermann et al., 2006) and *VANGL2* (Kibar et al., 2001; Torban et al., 2007). Future work will be necessary to understand whether the function of these genes in the human brain is comparable to their function in the mouse brain.

### DISCUSSION

### Gene expression in the context of energy metabolism

In light of the size and energetic costs of the human brain, energy metabolism genes might be important in human neocortical development. Indeed, previous studies have shown that genes coding for proteins involved in biological processes that supply the brain with energy display evidence of adaptive evolution in primates (Adkins and Honeycutt, 1994; Adkins et al., 1996; Andrews et al., 1998; Doan et al., 2004; Grossman et al., 2001, 2004; K.N. STERNER ET AL.



### A Significant difference in expression level.

### **B** Significant change in expression by age.

GOTerm	pFDR (Age)	Profile	pFDR (Profile)	Odds Ratio (Profile)
Synapse Organization	0.02	down then flat	0.18	2.26
Mitochondrial ATP Synthesis Coupled Electron Transport	0.04	linear up	0.002	4.81
Glycogen Metabolic Process	0.002	u shape	0.01	6.33

Fig. 3. Synapse- and energy metabolism-related gene enrichment. GO terms associated with synapse and energy metabolism related functions and modified gene lists derived from the Genes to Cognition Database http://www.genes2cognition.org/cgi-bin/GeneListView. Further information about GO Terms and Gene Lists can be found in Supporting Information Table S3. (a) As a group, these sets of genes have significantly (pFDR < 0.05) higher median expression level than all other probes called present on the array. \*Mitochondrial ATP synthesis coupled electron transport (GO:00042775) has been abbreviated as "electron transport." (b) These sets of genes show enrichment in genes more likely to be associated with age (pFDR < 0.05). Oxidative phosphorylation (GO:0006119) and energy reserve metabolic process (GO:0006112) are not shown because they are the parent terms of mitochondrial ATP synthesis coupled electron transport (GO:00042775) and glycogen metabolic process (GO:0005977), respectively. The gene expression profile classes most enriched in these gene lists are given with pFDR values and odds ratios.

Schmidt et al., 2002; Uddin et al., 2008a,b). Additionally, more pronounced upregulation of genes involved in energy metabolism and neuronal activity has been found in human neocortical tissue compared with nonhuman primates (Caceres et al., 2003; Fu et al., 2011; Nowick et al., 2009; Oldham et al., 2006; Uddin et al., 2004). We hypothesized that expression of genes involved in the process of energy metabolism would increase or decrease with developmental changes in energy consumption in the brain. Surprisingly, many genes that encode proteins known to play a role in energy metabolism (e.g., components of NADH dehydrogenase) did not show age-related patterns of gene expression, but instead had higher expression levels across all samples compared with the expression levels of the remaining genes on the microarray.

Of the 36 genes found to be differentially expressed by age, only two are known to encode proteins with functions related to energy metabolism: NQO1 and  $GSK3\beta$ . Glycogen synthase kinease- $3\beta$  ( $GSK3\beta$ ), a serine-threonine kinase involved in glucose homeostasis (Buller et al., 2008; Hoshi et al., 1996; Plyte et al., 1992) and numerous

other functions, showed overall low expression that decreased during childhood. NAD(P)H dehydrogenase [quinone] 1 is an enzyme that functions in antioxidant defense by limiting production of reactive oxygen species (Ross et al., 2000) and has *electron carrier activity* (molecular function GO:0009055). *NQO1* expression more closely mirrored known patterns of glucose consumption (Chugani et al., 1987) and had overall expression levels in the upper quartile. Disruption of *NQO1* in mice led to decreased circulating glucose levels (Gaikwad et al., 2001).

When we analyzed energy metabolism genes as a group (designated by such gene ontology terms as oxidative phosphorylation and mitochondrial ATP synthesiscoupled electron transport) we found this group to be enriched with genes that showed increased expression with age (Fig. 3B). This finding suggests that some of these transcripts may also be developmentally regulated but failed to meet the strict criteria of our differential expression analyses described above. These genes as a whole showed a trend of increased expression until adulthood, rather than a peak in expression earlier in development, as is the case with cortical glucose uptake. The

### AGE-RELATED GENE EXPRESSION IN CEREBRAL CORTEX



**Microarray Expression** 

Fig. 4. Test of the correlation between microarray and qPCR data. y-axis shows—DCt values ( $C_t$  reference –  $C_t$  target), while the x-axis shows the log2 normalized microarray expression.

difference between the expression profile we predicted and the profile we observed may be attributed in part to (1) the fact that glucose is not just an energy substrate, but during growth can also be diverted into biochemical pathways that promote nucleotide and lipid biosyntheses (McKenna et al., 2006); (2) the possibility that glucose metabolism efficiency increases with age; (3) the observation that the majority of glucose is used to maintain resting membrane potentials, and with the increase in synaptic pruning during adolescence energy requirements diminish; and/or (4) the fact that mitochondria use ATP for axonal transport, which should use more energy in the adult than the neonate, and for Ca buffering, whose needs should increase with the number of active synapses.

### Developmentally regulated genes

Although aerobic energy metabolism-related genes were not overrepresented in our dataset of differentially expressed genes, we did identify a subset of the genome that is developmentally regulated in the brain during approximately the first two decades of life and therefore may be critical for postnatal brain development. In this section we discuss some of the major patterns observed in our dataset and highlight potential candidate genes for future studies.

Genes involved in nervous system development, general developmental processes, and biological regulation were overrepresented in our dataset (Dataset S3). The majority of our differentially expressed genes (n = 28) showed

greater than median expression. In addition, of the eight profile classifications we used (Supporting Information Table S2), more transcripts fit the "down then flat" profile than the other seven profiles. Genes annotated to nervous system development in particular, all showed this pattern of expression (i.e., highest at birth with decreasing expression until adolescence). This finding is consistent with previous studies of postmortem human cortex tissue that have found that genes associated with axon guidance, long-term synaptic depression, regulation of transcription, and cell division show a continually declining pattern of expression over the course of postnatal development (Somel et al., 2009). Many of these genes have well described roles in the nervous system and our findings are consistent with studies that highlight their importance during development (e.g., GRIN3A encodes a subunit of a glutamate receptor which plays a critical role in synaptic plasticity (Traynelis et al., 2010)). In addition, some of these genes have been associated with disorders of the nervous system. For example, reduced expression of DPYSL3, neuroglycan C (encoded by CSPG5), and DCX early in development has been shown to correlate with Down syndrome (Weitzdoerfer et al., 2001), fetal growth retardation in rodents (Saito et al., 2009), and abnormal neural growth and differentiation (Qin et al., 2000a,b), respectively. Our findings suggest that these genes also alter their expression levels as a normal part of postnatal life.

Of the remaining 28 differentially expressed genes, another eight (FARP1, CX3CL1, AUTS2, FAM57B, MPP3,

*EMID1*, SH3RF3, and GSK3 $\beta$ ) showed an expression pattern similar to that seen for the genes specifically annotated to *nervous system development*, suggesting they may be equally important for the developing brain and are candidates for further study.  $GSK3\beta$ , as discussed above, is involved in glucose homeostasis and disruption of  $GSK\beta3$ in mice resulted in abnormal behavior/neurological phenotypes. FARP1, fractalkine (CX3CL1), and MPP3 are all annotated to biological regulation and therefore may regulate biological process, quality or function during postnatal brain development. Autism susceptibility candidate 2 (AUTS2) has been associated with autism (Sultana et al., 2002) and mental retardation (Kalscheuer et al., 2007). Although the precise function of AUTS2 is currently unknown, recent evidence suggests that in mice it may be important for neuronal development (Bedogni et al., 2010). Our study suggests that this gene is developmentally regulated in humans. Less is known about the function of FAM57B, EMID1, and SH3RF3 in the brain. In addition, there is much less information about the functions of the majority of the genes with expression trajectories that fit the remaining 7 profile classifications (Supporting Information Table S2). Although some of these genes have been described in other tissues (e.g., troponin I type 3 in the heart, TNNI3), their role in the brain is unknown. Our study suggests these genes are differentially expressed in the human brain during childhood and adolescence and may make interesting candidates for further study.

## Gene expression in the context of human DNA sequence evolution

There is increasing evidence that *cis*-regulatory changes can have large effects on gene expression patterns, and thus morphological and behavioral differences, between closely-related species (Wray, 2007). One key difference between humans and nonhuman primates is that humans have a longer period of postnatal development (Leigh, 2001; Leigh and Park, 1998) during which there is extensive brain growth and reorganization and elevated levels of glucose consumption. We found evidence of genes that likely play important roles in brain development during this time period. Although we do not currently have directly comparable expression data for nonhuman primate species, we were able test the regulatory regions of our candidate genes for evidence of adaptive evolution in humans. We found evidence for adaptive evolution in the cis-regulatory regions of three differentially expressed genes identified in this study. Two of the genes (BCAN and GRIN3A) showed evidence of positive selection in the human lineage since the last common ancestor shared with the chimpanzee. BCAN is a gene implicated in longterm potentiation and stabilization of synapses during postnatal brain development (Dityatev and Schachner, 2003). GRIN3A is part of the NMDA receptor and directly involved in synaptic transmission and plasticity (Travnelis et al., 2010). Adaptive evolution or divergence of the coding region of GRIN3A between humans and chimpanzees has been implicated in other studies (Goto et al., 2009; Toll-Riera et al., 2011). Significant human divergence in the *cis*-regulatory regions of these genes suggests that age-related expression patterns may have also diverged greatly between humans and chimpanzees. In support of our findings, BCAN and GRIN3A show humanspecific expression changes in the prefrontal cortex when

compared with nonhuman primates (Liu et al., 2012). More experimental work is needed to determine if the *cis*regulatory changes we identify here drive changes observed in the expression of these genes. In contrast, *ANKRD57* shows evidence of strong purifying selection, and may show evidence of a retained expression pattern in this gene. One gene, *STAT4*, shows evidence of positive selection in the coding sequence. *STAT4* is a transcription factor involved in immune function and cell proliferation; its role in the brain development is unclear although it shows human-specific expression changes in the prefrontal cortex as well (Liu et al., 2012).

### Potential limitations

Our study includes four potential limitations. First, our study is more targeted and conservative than previous studies, which may result in fewer detected differentially expressed transcripts. Second, our study design is crosssectional, not longitudinal. The ideal study would sample human individuals multiple times during development, but current technology does not enable this approach given the invasive nature of tissue collection and other ethical considerations. Third, our samples were opportunistically collected and are derived from patients with known neuropathologies (e.g., epilepsy); and therefore, some gene expression levels may be representative of neuropathology. Fourth, gene expression patterns may not mirror patterns of protein expression. The following section discusses these potential limitations in the context of our findings.

Our study is more conservative than previously published studies. Recently, a developmental human transcriptome dataset was published that included over 1,500 microarray experiments sampling individuals from prenatal to adult life from eleven neocortical regions (Kang et al., 2011). The study demonstrated that nearly 24% of well-annotated genes are differentially expressed within a particular brain region in at least 1 of 15 age class bins. This result suggests many more genes are differentially expressed than we report here. One reason for this discrepancy is that our analytical approach is quite different from that taken by Kang et al. (2011). We calculated quadratic regressions, which require differential expression values at many developmental time points in order for significance to be achieved. In that sense, our study is much more conservative. However, our goal was to find genes that showed age-related patterns of expression difference that parallel previously published glucose utilization curves (Chugani et al., 1987), and underlying processes of synaptogenesis (Glantz et al., 2007; Huttenlocher and Dabholkar, 1997), whereas Kang et al. (2011) focused on characterizing gene expression patterns across brain regions and developmental stages. As such, the two studies are not directly comparable. Despite these differences some results from the two studies are similar. For example, we corroborate the finding that doublecortin (DCX)shows higher expression in children than in adults.

In addition, the stringency of our multiple comparisons criteria made our regression analyses conservative; therefore, many transcripts that follow developmental trends but are not strong correlates of age did not pass the significance threshold in our analyses. Although the stringency of our methods gives us confidence that we have identified the genes with the strongest developmental trajectories, it is likely that additional gene transcripts are involved in cortical development.

### The present study design is cross-sectional, which requires that we infer developmental changes from samples obtained from different individuals varying in age. Although this could lead us to conflate individual differences with age-related changes, the statistical methods that we used only identified relatively consistent trends across the full set of samples and included rigorous adjustment for multiple hypothesis testing, thus reducing the likelihood of individual data points having large effects on the findings.

In order to avoid potential biases associated with utilizing tissue samples derived postmortem (see above), we utilized tissue derived mainly from patients undergoing surgery to treat epilepsy. As part of these procedures, a small amount of healthy tissue is often removed to ensure full excision of the lesion. We used this healthy (i.e., electrophysiologically inactive) tissue as in previous studies that demonstrate gene expression differences in nonepileptic (electrophysiologically inactive) vs. epileptic (electrophysiologically active) tissue (Rakhade et al., 2005). Furthermore, brain glucose metabolism in epileptic infants, children, and adults was found to follow a similar developmental curve to patients without epilepsy (Chugani et al., 1987) despite partial epilepsy and patient treatment with anti-epileptic medication (Bentourkia et al., 1998). Unfortunately we did not have patient information about duration of disease, although all surgical candidates shared the characteristic of having epilepsy that was unresponsive to standard medications at the time the tissue was collected. Although inferring normal developmental patterns of gene expression from surgically resected tissue derived from patients with neurological pathologies (e.g., epilepsy) is not ideal, use of the healthy surrounding tissue provides invaluable information about differential gene expression in human neocortical tissue during ontogeny. We have found some correspondence between postmortem and resected gene expression patterns (this study and Sterner et al., 2012). In addition, we note that some of the genes [e.g., GRIN3A (Choi et al., 2009; Harris et al., 2009; Henson et al., 2008) and BCAN (Gary et al., 2000)] included in our study have also been shown to be developmentally regulated in other studies that measure mRNA levels in tissues collected postmortem from individuals without epilepsy.

Pubertal stage may serve as a better indicator of age for some individuals in our study. It would also be interesting to explore gene expression changes associated with the onset of puberty (Ojeda et al., 2010) and examine the relationship between hormone production and the developmental changes we describe here (Campbell, 2011). Unfortunately we were unable to consider pubertal stage for the individuals included in this study; however, only a few individuals included in our analyses were taken from individuals at or around the time of puberty onset.

In this study we determined age-related patters of mRNA expression in human cerebral cortical samples. Studies suggest that mRNA levels may only partly correspond to protein concentrations (Plotkin, 2010). It will be interesting to determine if genes involved in energy metabolism are developmentally regulated at the protein level and if there is correspondence between mRNA and protein levels in our brain samples.

### CONCLUSIONS AND FUTURE PROSPECTS

In sum, our investigation of patterns of gene transcription in human cortical samples representing individuals from 10 months to 18 years of age identified 40 transcripts (38 unique genes) that followed distinct developmental changes in expression that are similar to glucose utilization curves (Chugani et al., 1987) and processes of synapse formation (Glantz et al., 2007; Huttenlocher and Dabholkar, 1997). Among genes following significant developmental/ontogenetic trajectories, those involved in nervous system development were significantly overrepresented. We found comparatively little evidence for developmental trajectories in genes annotated to function in synapse or energy metabolism related processes. Instead, many of these genes were found to have high expression levels at all ages when compared with expression levels of the remaining genes on the microarray. Analyses of cis-regulatory sequence changes provides evidence of adaptive evolution in three of the 36 age-specific differentially expressed genes identified in this study. Two of these genes (BCAN and GRIN3A) showed evidence of positive selection in the human lineage since the last common ancestor shared with the chimpanzee, while ANKRD57 showed evidence of strong purifying selection in this same lineage. Closer examination of how the 40 transcripts described here, in particular those with evidence suggesting positive selection in humans (e.g., BCAN and GRIN3A), are regulated in nonhuman primates and other large brained mammals will help determine more precisely whether the developmental patterns of gene expression we observed are evolutionarily derived in humans or conserved among other primates.

### ACKNOWLEDGMENTS

The authors thank four anonymous reviewers for their insightful comments and helpful suggestions. The authors also thank William Leonard (Northwestern University) and Otto Muzik (Wayne State University) for insightful discussion and Mei-Li Lee (Wayne State University), Henriette Mair-Meijers, and Lucie Grégoire (Wayne State University) for providing research support.

### LITERATURE CITED

- Ackermann GE, Marenholz I, Wolfer DP, Chan WY, Schafer B, Erne P, Heizmann CW. 2006. S100A1-deficient male mice exhibit increased exploratory activity and reduced anxiety-related responses. Biochim Biophys Acta 1763:1307–1319.
- Adkins RM, Honeycutt RL. 1994. Evolution of the primate cytochrome c oxidase subunit II gene. J Mol Evol 38:215-231.
- Adkins RM, Honeycutt RL, Disotell TR. 1996. Evolution of eutherian cytochrome c oxidase subunit II: Heterogeneous rates of protein evolution and altered interaction with cytochrome c. Mol Biol Evol 13:1393–1404.
- Aiello LC, Wheeler P. 1995. The expensive-tissue hypothesis: the brain and the digestive system in human and primate evolution. Curr Anthropol 36:199.
- Andolfatto P. 2005. Adaptive evolution of non-coding DNA in Drosophila. Nature 437:1149–1152.
- Andrews TD, Jermiin LS, Easteal S. 1998. Accelerated evolution of cytochrome b in simian primates: adaptive evolution in concert with other mitochondrial proteins? J Mol Evol 47:249–257.
- Austin CP, Battey JF, Bradley A, Bucan M, Capecchi M, Collins FS, Dove WF, Duyk G, Dymecki S, Eppig JT, Grieder FB, Heintz N, Hicks G, Insel TR, Joyner A, Koller BH, Lloyd KC, Magnuson T, Moore MW, Nagy A, Pollock JD, Roses AD, Sands AT, Seed B, Skarnes WC, Snoddy J, Soriano P, Stewart DJ, Stewart F, Stillman B, Varmus H, Varticovski L, Verma IM, Vogt TF, von Melchner H, Witkowski J, Woychik RP, Wurst W, Yancopoulos GD, Young SG, Zambrowicz B. 2004. The knockout mouse project. Nat Genet 36:921–924.

- Bedogni F, Hodge RD, Nelson BR, Frederick EA, Shiba N, Daza RA, Hevner RF. 2010. Autism susceptibility candidate 2 (Auts2) encodes a nuclear protein expressed in developing brain regions implicated in autism neuropathology. Gene Expr Patterns 10:9–15.
- Benjamini Y, Hochberg Y. 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J Royal Stat Soc B 57:289-300.
- Bentourkia M, Michel C, Ferriere G, Bol A, Coppens A, Sibomana M, Bausart R, Labar D, De Volder AG. 1998. Evolution of brain glucose metabolism with age in epileptic infants, children and adolescents. Brain Dev 20:524–529.
- Blake JA, Bult CJ, Eppig JT, Kadin JA, Richardson JE. 2009. The mouse genome database genotypes: phenotypes. Nucleic Acids Res 37(Database issue):D712–D719.
- Bolstad BM, Irizarry RA, Astrand M, Speed TP. 2003. A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. Bioinformatics 19:185–193.
- Brakebusch C, Seidenbecher CI, Asztely F, Rauch U, Matthies H, Meyer H, Krug M, Bockers TM, Zhou X, Kreutz MR, Montag D, Gundelfinger ED, Fassler R. 2002. Brevican-deficient mice display impaired hippocampal CA1 long-term potentiation but show no obvious deficits in learning and memory. Mol Cell Biol 22:7417–7427.
- Brazma A, Hingamp P, Quackenbush J, Sherlock G, Spellman P, Stoeckert C, Aach J, Ansorge W, Ball CA, Causton HC, Gaasterland T, Glenisson P, Holstege FC, Kim IF, Markowitz V, Matese JC, Parkinson H, Robinson A, Sarkans U, Schulze-Kremer S, Stewart J, Taylor R, Vilo J, Vingron M. 2001. Minimum information about a microarray experiment (MIAME)-toward standards for microarray data. Nat Genet 2:365–371.
- Brody SA, Nakanishi N, Tu S, Lipton SA, Geyer MA. 2005. A developmental influence of the N-methyl-d-aspartate receptor NR3A subunit on prepulse inhibition of startle. Biol Psychiatry 57:1147–1152.
- Buller CL, Loberg RD, Fan MH, Zhu Q, Park JL, Vesely E, Inoki K, Guan KL, Brosius FC III. 2008. A GSK-3/TSC2/mTOR pathway regulates glucose uptake and GLUT1 glucose transporter expression. Am J Physiol Cell Physiol 295:C836-C843.
- Caceres M, Lachuer J, Zapala MA, Redmond JC, Kudo L, Geschwind DH, Lockhart DJ, Preuss TM, Barlow C. 2003. Elevated gene expression levels distinguish human from non-human primate brains. Proc Natl Acad Sci USA 100:13030–13035.
- Campbell B. 2011. Adrenarche in comparative perspective. Am J Hum Biol 23:44–52.
- Catts VS, Catts SV, Fernandez HR, Taylor JM, Coulson EJ, Lutze-Mann LH. 2005. A microarray study of post-mortem mRNA degradation in mouse brain tissue. Brain Res Mol Brain Res 138:164–177.
- Choi KH, Zepp ME, Higgs BW, Weickert CS, Webster MJ. 2009. Expression profiles of schizophrenia susceptibility genes during human prefrontal cortical development. J Psychiatry Neurosci 34:450–458.
- Chugani HT. 1998. A critical period of brain development: studies of cerebral glucose utilization with PET. Prev Med 27:184–188.
- Chugani HT, Phelps ME, Mazziotta JC. 1987. Positron emission tomography study of human brain functional development. Ann Neurol 22: 487–497.
- Darwin C. 1871. The descent of man and selection in relation to sex. London: John Murray.
- Das S, Sasaki YF, Rothe T, Premkumar LS, Takasu M, Crandall JE, Dikkes P, Conner DA, Rayudu PV, Cheung W, Chen HS, Lipton SA, Nakanishi N. 1998. Increased NMDA current and spine density in mice lacking the NMDA recentor subunit NR3A. Nature 393:377-381
- lacking the NMDA receptor subunit NR3A. Nature 393:377–381. Davy BE, Robinson ML. 2003. Congenital hydrocephalus in hy3 mice is caused by a frameshift mutation in Hydin, a large novel gene. Hum Mol Genet 12:1163–1170.
- Dityatev A, Schachner M. 2003. Extracellular matrix molecules and synaptic plasticity. Nat Rev Neurosci 4:456–468.
- Doan JW, Schmidt TR, Wildman DE, Uddin M, Goldberg A, Huttemann M, Goodman M, Weiss ML, Grossman LI. 2004. Coadaptive evolution in cytochrome c oxidase: 9 of 13 subunits show accelerated rates of nonsynonymous substitution in anthropoid primates. Mol Phylogenet Evol 33:944-950.
- Dorval V, Smith PY, Delay C, Calvo E, Planel E, Zommer N, Buee L, Hebert SS. 2012. Gene network and pathway analysis of mice with conditional ablation of Dicer in post-mitotic neurons. PLoS One 7:e44060.
- Egea R, Casillas S, Barbadilla A. 2008. Standard and generalized McDonald-Kreitman test: a website to detect selection by comparing different classes of DNA sites. Nucleic Acids Res 36:W157–W162.
- Ferrer I, Martinez A, Boluda S, Parchi P, Barrachina M. 2008. Brain banks: benefits, limitations and cautions concerning the use of post-mortem brain tissue for molecular studies. Cell Tissue Bank 9:181–194.
- Fu X, Giavalisco P, Liu X, Catchpole G, Fu N, Ning ZB, Guo S, Yan Z, Somel M, Paabo S, Zeng R, Willmitzer L, Khaitovich P. 2011. Rapid metabolic evolution in human prefrontal cortex. Proc Natl Acad Sci USA 108:6181–6186.

- Gaikwad A, Long DJ, Stringer JL, Jaiswal AK. 2001. In vivo role of NAD(P)H:quinone oxidoreductase 1 (NQO1) in the regulation of intracellular redox state and accumulation of abdominal adipose tissue. J Biol Chem 276:22559–22564.
- Gajdos ZK, Henderson KD, Hirschhorn JN, Palmert MR. 2010. Genetic determinants of pubertal timing in the general population. Mol Cell Endocrinol 324:21-29.
- Galvan A. 2010. Neural plasticity of development and learning. Hum Brain Mapp 31:879-890.
- Gary SC, Zerillo CA, Chiang VL, Gaw JU, Gray G, Hockfield S. 2000. cDNA cloning, chromosomal localization, and expression analysis of human BEHAB/brevican, a brain specific proteoglycan regulated during cortical development and in glioma. Gene 256:139–147.
- Glantz LA, Gilmore JH, Hamer RM, Lieberman JA, Jarskog LF. 2007. Synaptophysin and postsynaptic density protein 95 in the human prefrontal cortex from mid-gestation into early adulthood. Neuroscience 149:582-591.
- Goto H, Watanabe K, Araragi N, Kageyama R, Tanaka K, Kuroki Y, Toyoda A, Hattori M, Sakaki Y, Fujiyama A, Fukumaki Y, Shibata H. 2009. The identification and functional implications of human-specific "fixed" amino acid substitutions in the glutamate receptor family. BMC Evol Biol 9:224.
- Grossman LI, Schmidt TR, Wildman DE, Goodman M. 2001. Molecular evolution of aerobic energy metabolism in primates. Mol Phylogenet Evol 18:26-36.
- Grossman LI, Wildman DE, Schmidt TR, Goodman M. 2004. Accelerated evolution of the electron transport chain in anthropoid primates. Trends Genet 20:578–585.
- Harris LW, Lockstone HE, Khaitovich P, Weickert CS, Webster MJ, Bahn S. 2009. Gene expression in the prefrontal cortex during adolescence: implications for the onset of schizophrenia. BMC Med Genomics 2:28.
- Henson MA, Roberts AC, Salimi K, Vadlamudi S, Hamer RM, Gilmore JH, Jarskog LF, Philpot BD. 2008. Developmental regulation of the NMDA receptor subunits, NR3A and NR1, in human prefrontal cortex. Cereb Cortex 18:2560–2573.
- Holliday M. 1986. Body composition and energy needs during growth. In: Falkner F, Tanner J, editors. Human Growth: A Comprehensive Treatise. Vol. 2, 2nd Ed. New York: Plenum Press. pp. 101–115.
- Holloway RL Jr. 1968. The evolution of the primate brain: some aspects of quantitative relations. Brain Res 7:121–172.
- Hoshi M, Takashima A, Noguchi K, Murayama M, Sato M, Kondo S, Saitoh Y, Ishiguro K, Hoshino T, Imahori K. 1996. Regulation of mitochondrial pyruvate dehydrogenase activity by tau protein kinase I/glycogen synthase kinase 3beta in brain. Proc Natl Acad Sci USA 93:2719–2723.
- Huttenlocher PR, Dabholkar AS. 1997. Regional differences in synaptogenesis in human cerebral cortex. J Comp Neurol 387:167–178.
- Jacobs B, Chugani HT, Allada V, Chen S, Phelps ME, Pollack DB, Raleigh MJ. 1995. Developmental changes in brain metabolism in sedated rhesus macaques and vervet monkeys revealed by positron emission tomography. Cereb Cortex 5:222–233.
- Jerison H. 1973. Evolution of the brain and intelligence. Academic Press, New York. p 482.
- Johnston MV. 2009. Plasticity in the developing brain: implications for rehabilitation. Dev Disabil Res Rev 15:94–101.
- Juttner R, More MI, Das D, Babich A, Meier J, Henning M, Erdmann B, Mu Ller EC, Otto A, Grantyn R, Rathjen FG. 2005. Impaired synapse function during postnatal development in the absence of CALEB, an EGF-like protein processed by neuronal activity. Neuron 46:233–245.
- Kalscheuer VM, FitzPatrick D, Tommerup N, Bugge M, Niebuhr E, Neumann LM, Tzschach A, Shoichet SA, Menzel C, Erdogan F, Arkesteijn G, Ropers HH, Ullmann R. 2007. Mutations in autism susceptibility candidate 2 (AUTS2) in patients with mental retardation. Hum Genet 121:501–509.
- Kang HJ, Kawasawa YI, Cheng F, Zhu Y, Xu X, Li M, Sousa AM, Pletikos M, Meyer KA, Sedmak G, Guennel T, Shin Y, Johnson MB, Krsnik Z, Mayer S, Fertuzinhos S, Umlauf S, Lisgo SN, Vortmeyer A, Weinberger DR, Mane S, Hyde TM, Huttner A, Reimers M, Kleinman JE, Sestan N. 2011. Spatio-temporal transcriptome of the human brain. Nature 478:483–489.
- Kapfhamer D, Berger KH, Hopf FW, Seif T, Kharazia V, Bonci A, Heberlein U. 2010. Protein Phosphatase 2a and glycogen synthase kinase 3 signaling modulate prepulse inhibition of the acoustic startle response by altering cortical M-Type potassium channel activity. J Neurosci 30:8830–8840.
- Kappeler C, Saillour Y, Baudoin JP, Tuy FP, Alvarez C, Houbron C, Gaspar P, Hamard G, Chelly J, Metin C, Francis F. 2006. Branching and nucleokinesis defects in migrating interneurons derived from doublecortin knockout mice. Hum Mol Genet 15:1387–1400.
- Kerr MK. 2003. Linear models for microarray data analysis: hidden similarities and differences. J Comput Biol 10:891–901.

- Kibar Z, Vogan KJ, Groulx N, Justice MJ, Underhill DA, Gros P. 2001. Ltap, a mammalian homolog of Drosophila Strabismus/Van Gogh, is altered in the mouse neural tube mutant Loop-tail. Nat Genet 28: 251–255.
- Kim WY, Zhou FQ, Zhou J, Yokota Y, Wang YM, Yoshimura T, Kaibuchi K, Woodgett JR, Anton ES, Snider WD. 2006. Essential roles for GSK-3s and GSK-3-primed substrates in neurotrophin-induced and hippocampal axon growth. Neuron 52:981–996.
- King MC, Wilson AC. 1975. Evolution at 2 levels in humans and chimpanzees. Science 188:107–116.
- Kuzawa CW. 1998. Adipose tissue in human infancy and childhood: an evolutionary perspective. Am J Phys Anthropol Suppl 27:177–209.
- Larsen RS, Corlew RJ, Henson MA, Roberts AC, Mishina M, Watanabe M, Lipton SA, Nakanishi N, Perez-Otano I, Weinberg RJ, Philpot BD. 2011. NR3A-containing NMDARs promote neurotransmitter release and spike timing-dependent plasticity. Nat Neurosci 14:338–344.
- Leigh SR. 2001. Evolution of human growth. Evol Anthropol 10:223–236. Leigh SR. 2004. Brain growth, life history, and cognition in primate and
- human evolution. Am J Primatol 62:139–164. Leigh SR, Park PB. 1998. Evolution of human growth prolongation. Am J
- Phys Anthropol 107:331–350. Leonard WR, Robertson ML. 1992. Nutritional requirements and human
- evolution: a bioenergetics model. Am J Hum Biol 4:179–195. Leonard WR. Snodgrass JJ. Robertson ML, 2007. Effects of brain evolution
- on human nutrition and metabolism. Annu Rev Nutr 27:311–327.
- Levula M, Oksala N, Airla N, Zeitlin R, Salenius JP, Jarvinen O, Venermo M, Partio T, Saarinen J, Somppi T, Suominen V, Virkkunen J, Hautalahti J, Laaksonen R, Kahonen M, Mennander A, Kytomaki L, Soini JT, Parkkinen J, Pelto-Huikko M, Lehtimaki T. 2012. Genes involved in systemic and arterial bed dependent atherosclerosis--Tampere Vascular study. PLoS One 7:e33787.
- Lin L, Lee VM, Wang Y, Lin JS, Sock E, Wegner M, Lei L. 2011. Sox11 regulates survival and axonal growth of embryonic sensory neurons. Dev Dyn 240:52–64.
- Liu H, Tarima S, Borders AS, Getchell TV, Getchell ML, Stromberg AJ. 2005. Quadratic regression analysis for gene discovery and pattern recognition for non-cyclic short time-course microarray experiments. BMC Bioinformatics 6:106.
- Liu X, Somel M, Tang L, Yan Z, Jiang X, Guo S, Yuan Y, He L, Oleksiak A, Zhang Y, Li N, Hu Y, Chen W, Qiu Z, Paabo S, Khaitovich P. 2012. Extension of cortical synaptic development distinguishes humans from chimpanzees and macaques. Genome Res 22:611–622.
- Martin RD. 1981. Relative brain size and basal metabolic rate in terrestrial vertebrates. Nature 293:57–60.
- MCCullagh P, Nelder J. 1989. Generalized linear models. Boca Raton: Chapmanand Hall/CRC.
- McDonald JH, Kreitman M. 1991. Adaptive protein evolution at the Adh locus in Drosophila. Nature 351:652–654.
- McKenna MC, Gruetter R, Sonnewald U, Waagepetersen HS, Schousboe A. 2006. Energy metabolism of the brain. In: Siegel GJ, Albers RW, Brady ST, Price DL, editors. Basic neurochemistry: molecular, cellular and medical aspects, 7th ed. San Diego, CA: Elsevier Academic Press. p 531–557.
- Mittal P, Romero R, Tarca AL, Draghici S, Nhan-Chang C-L, Chaiworapongsa T, Hotra J, Gomez R, Kusanovic JP, Lee D-C, Kim CJ, Hassan SS. 2011. A molecular signature of an arrest of descent in human parturition. Am J Obstet Gynecol 204:177.e115–177.e133.
- Nakanishi N, Tu S, Shin Y, Cui J, Kurokawa T, Zhang D, Chen HS, Tong G, Lipton SA. 2009. Neuroprotection by the NR3A subunit of the NMDA receptor. J Neurosci 29:5260–5265.
- Nosten-Bertrand M, Kappeler C, Dinocourt C, Denis C, Germain J, Phan Dinh Tuy F, Verstraeten S, Alvarez C, Metin C, Chelly J, Giros B, Miles R, Depaulis A, Francis F. 2008. Epilepsy in Dcx knockout mice associated with discrete lamination defects and enhanced excitability in the hippocampus. PLoS One 3:e2473.
- Nowick K, Gernat T, Almaas E, Stubbs L. 2009. Differences in human and chimpanzee gene expression patterns define an evolving network of transcription factors in brain. Proc Natl Acad Sci USA 106: 22358–22363.
- Ojeda SR, Dubay C, Lomniczi A, Kaidar G, Matagne V, Sandau US, Dissen GA. 2010. Gene networks and the neuroendocrine regulation of puberty. Mol Cell Endocrinol 324:3–11.
- Oldham MC, Horvath S, Geschwind DH. 2006. Conservation and evolution of gene coexpression networks in human and chimpanzee brains. Proc Natl Acad Sci USA 103:17973–17978.
- Plotkin JB. 2010. Transcriptional regulation is only half the story. Mol Syst Biol 6:406.
- Plyte SE, Hughes K, Nikolakaki E, Pulverer BJ, Woodgett JR. 1992. Glycogen synthase kinase-3: functions in oncogenesis and development. Biochim Biophys Acta 1114:147–162.

- Potzner MR, Tsarovina K, Binder E, Penzo-Mendez A, Lefebvre V, Rohrer H, Wegner M, Sock E. 2010. Sequential requirement of Sox4 and Sox11 during development of the sympathetic nervous system. Development 137:775–784.
- Qin J, Mizuguchi M, Itoh M, Takashima S. 2000a. Immunohistochemical expression of doublecortin in the human cerebrum: comparison of normal development and neuronal migration disorders. Brain Res 863: 225–232.
- Qin J, Mizuguchi M, Itoh M, Takashima S. 2000b. A novel migrationrelated gene product, doublecortin, in neuronal migration disorder of fetuses and infants with Zellweger syndrome. Acta Neuropathol 100:168-173.
- Rakhade SN, Yao B, Ahmed S, Asano E, Beaumont TL, Shah AK, Draghici S, Krauss R, Chugani HT, Sood S, Loeb JA. 2005. A common pattern of persistent gene activation in human neocortical epileptic foci. Ann Neurol 58:736–747.
- Rand DM, Kann LM. 1996. Excess amino acid polymorphism in mitochondrial DNA: contrasts among genes from Drosophila, mice, and humans. Mol Biol Evol 13:735–748.
- Reim K, Regus-Leidig H, Ammermuller J, El-Kordi A, Radyushkin K, Ehrenreich H, Brandstatter JH, Brose N. 2009. Aberrant function and structure of retinal ribbon synapses in the absence of complexin 3 and complexin 4. J Cell Sci 122:1352–1361.
- Roberts AC, Diez-Garcia J, Rodriguiz RM, Lopez IP, Lujan R, Martinez-Turrillas R, Pico E, Henson MA, Bernardo DR, Jarrett TM, Clendeninn DJ, Lopez-Mascaraque L, Feng G, Lo DC, Wesseling JF, Wetsel WC, Philpot BD, Perez-Otano I. 2009. Downregulation of NR3A-containing NMDARs is required for synapse maturation and memory consolidation. Neuron 63:342–356.
- Ross D, Kepa JK, Winski SL, Beall HD, Anwar A, Siegel D. 2000. NAD(P)H:quinone oxidoreductase 1 (NQO1): chemoprotection, bioactivation, gene regulation and genetic polymorphisms. Chem Biol Interact 129:77–97.
- Ryan MM, Huffaker SJ, Webster MJ, Wayland M, Freeman T, Bahn S. 2004. Application and optimization of microarray technologies for human postmortem brain studies. Biol Psychiatry 55:329–336.
- Saito A, Matsui F, Hayashi K, Watanabe K, Ichinohashi Y, Sato Y, Hayakawa M, Kojima S, Oohira A. 2009. Behavioral abnormalities of fetal growth retardation model rats with reduced amounts of brain proteoglycans. Exp Neurol 219:81–92.
- Schmidt TR, Goodman M, Grossman LI. 2002. Amino acid replacement is rapid in primates for the mature polypeptides of COX subunits, but not for their targeting presequences. Gene 286:13–19.
- Smyth G. 2005. Limma: linear models for microarray data. In: Gentleman R, Carey V, Dudoit S, Irizarry RA, Huber W, editors. Bioinformatics and computational biology solutions using R and bioconductor. New York: Springer. p 397–420.
- Somel M, Franz H, Yan Z, Lorenc A, Guo S, Giger T, Kelso J, Nickel B, Dannemann M, Bahn S, Webster MJ, Weickert CS, Lachmann M, Paabo S, Khaitovich P. 2009. Transcriptional neoteny in the human brain. Proc Natl Acad Sci USA 106:5743–5748.
- Somel M, Liu X, Tang L, Yan Z, Hu H, Guo S, Jiang X, Zhang X, Xu G, Xie GLi N, Hu Y, Chen W, Paabo S, Khaitovich P. 2011. MicroRNA-driven developmental remodeling in the brain distinguishes humans from other primates. PLoS Biol 9:e1001214.
- Soriano SG, Amaravadi LS, Wang YF, Zhou H, Yu GX, Tonra JR, Fairchild-Huntress V, Fang Q, Dunmore JH, Huszar D, Pan Y. 2002. Mice deficient in fractalkine are less susceptible to cerebral ischemia-reperfusion injury. J Neuroimmunol 125:59–65.
- Steinhoff C, Vingron M. 2006. Normalization and quantification of differential expression in gene expression microarrays. Brief Bioinform 7:166–177.
- Sterner KN, Weckle A, Chugani HT, Tarca AL, Sherwood CC, Hof PR, Kuzawa CW, Boddy AM, Abbas A, Raaum RL, Gregoire L, Lipovich L, Grossman LI, Uddin M, Goodman M, Wildman DE. 2012. Dynamic gene expression in the human cerebral cortex distinguishes children from adults. PLoS One 7:e37714.
- Sultana R, Yu CE, Yu J, Munson J, Chen D, Hua W, Estes A, Cortes F, de la Barra F, Yu D, Haider ST, Trask BJ, Green ED, Raskind WH, Disteche CM, Wijsman E, Dawson G, Storm DR, Schellenberg GD, Villacres EC. 2002. Identification of a novel gene on chromosome 7q11.2 interrupted by a translocation breakpoint in a pair of autistic twins. Genomics 80:129–134.
- Thein DC, Thalhammer JM, Hartwig AC, Crenshaw EB III, Lefebvre V, Wegner M, Sock E. 2010. The closely related transcription factors Sox4 and Sox11 function as survival factors during spinal cord development. J Neurochem 115:131–141.
- Toll-Riera M, Laurie S, Alba MM. 2011. Lineage-specific variation in intensity of intensity of natural selection in mammals. Mol Biol Evol 28:383– 398.

- Torban E, Wang HJ, Patenaude AM, Riccomagno M, Daniels E, Epstein D, Gros P. 2007. Tissue, cellular and sub-cellular localization of the Vangl2 protein during embryonic development: effect of the Lp mutation. Gene Expr Patterns 7:346–354.
- Torgerson DG, Boyko AR, Hernandez RD, Indap A, Hu X, White TJ, Sninsky JJ, Cargill M, Adams MD, Bustamante CD, Clark AG. 2009. Evolutionary processes acting on candidate cis-regulatory regions in humans inferred from patterns of polymorphism and divergence. PLoS Genet 5:e1000592.
- Traynelis SF, Wolmuth LP, McBain CJ, Menniti FS, Vance KM, Ogden KK, Hansen KB, Yuan H, Myers SJ, Dingledine R. 2010. Glutamate receptor ion channels: structure, regulation, and function. Pharmacol Rev 62:405–496.
- Tuy FP, Saillour Y, Kappeler C, Chelly J, Francis F. 2008. Alternative transcripts of Dclk1 and Dclk2 and their expression in doublecortin knockout mice. Dev Neurosci 30:171–186.
- Uddin M, Goodman M, Erez O, Romero R, Liu G, Islam M, Opazo JC, Sherwood CC, Grossman LI, Wildman DE. 2008a. Distinct genomic signatures of adaptation in pre- and postnatal environments during human evolution. Proc Natl Acad Sci USA 105:3215-3220.

- Uddin M, Opazo JC, Wildman DE, Sherwood CC, Hof PR, Goodman M, Grossman LI. 2008b. Molecular evolution of the cytochrome c oxidase subunit 5A gene in primates. BMC Evol Biol 8:8.
- Uddin M, Wildman DE, Liu G, Xu W, Johnson RM, Hof PR, Kapatos G, Grossman LI, Goodman M. 2004. Sister grouping of chimpanzees and humans as revealed by genome-wide phylogenetic analysis of brain gene expression profiles. Proc Natl Acad Sci USA 101:2957–2962.
- Wang K, Li M, Hakonarson H. 2010. Analysing biological pathways in genome-wide association studies. Nat Rev Genet 11:843–854.
- Weitzdoerfer R, Fountoulakis M, Lubec G. 2001. Aberrant expression of dihydropyrimidinase related proteins-2,-3 and [minus]4 in fetal Down syndrome brain. J Neural Transm Suppl 61:95–107.
- Wray GA. 2007. The evolutionary significance of cis-regulatory mutations. Nat Rev Genet 8:206–216.
- Xue M, Stradomska A, Chen H, Brose N, Zhang W, Rosenmund C, Reim K. 2008. Complexins facilitate neurotransmitter release at excitatory and inhibitory synapses in mammalian central nervous system. Proc Natl Acad Sci USA 105:7875–7880.
- 1000 Genomes Project Consortium. 2010. A map of human genome variation from population-scale sequencing. Nature 467: 1061-1073.