Regional transcriptome analysis of AMPA and GABA<sub>A</sub> receptor subunit expression generates E/I signatures of the human brain

Kevin Shen<sup>1,4</sup>, Tommaso Zeppillo<sup>2,3,4</sup> & Agenor Limon<sup>3</sup>*

Theoretical and experimental work has demonstrated that excitatory (E) and inhibitory (I) currents within cortical circuits stabilize to a balanced state. This E/I balance, observed from single neuron to network levels, has a fundamental role in proper brain function and its impairment has been linked to numerous brain disorders. Over recent years, large amount of microarray and RNA-Sequencing datasets have been collected, however few studies have made use of these resources for exploring the balance of global gene expression levels between excitatory AMPA receptors (AMPARs) and inhibitory GABA<sub>A</sub> receptors. Here, we analyzed the relative relationships between these receptors to generate a basic transcriptional marker of E/I ratio. Using publicly available data from the Allen Brain Institute, we generated whole brain and regional signatures of AMPAR subunit gene expression in healthy human brains as well as the transcriptional E/I (tE/I) ratio. Then we refined the tE/I ratio to cell-type signatures in the mouse brain using data from the Gene Expression Omnibus. Lastly, we applied our workflow to developmental data from the Allen Brain Institute and revealed spatially and temporally controlled changes in the tE/I ratio during the embryonic and early postnatal stages that ultimately lead to the tE/I balance in adults.

Theoretical and experimental work has shown that fast excitatory (E) and inhibitory (I) currents within cortical circuits stabilize rapidly to a balanced state<sup>1-6</sup>. This E/I balance, observed from single neuron to recurrent connected network levels, produces robust neuronal learning capabilities while minimizing disturbance by output noise<sup>7</sup>. The balance's fundamental role in proper brain function has been demonstrated via pharmacological blocking of cortical inhibition which led to cortical epileptic activity and loss in stimuli features decoding<sup>8,9</sup>. Therefore, disturbances to this balance have been linked to numerous psychiatric, neurodevelopmental, and neurodegenerative disorders including schizophrenia, autism, and Alzheimer's disease<sup>10-14</sup>. At the synaptic level, the electrophysiological E/I (eE/I) balance is driven mainly by the concerted activity of excitatory α-amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid receptors (AMPARs) and inhibitory γ-aminobutyric acid type A receptors (GABA<sub>A</sub>Rs)<sup>15</sup>; however, it remains unknown to what extent the transcription levels of each receptor are also balanced across brain regions and developmental stages. This information would prove highly important and useful since direct measurement of the activity of neurotransmitter receptors at the single neuron level in healthy humans is not currently possible with available technology. Rather, differential gene expression of postmortem tissue is the most widely used approach in making inferences of neurotransmitter function in physiological and diseased states.

Our previous work demonstrated a bioinformatics approach to analyzing the organizational layout of GABA<sub>A</sub>Rs in the human brain. This approach proposed the most abundant isoforms that followed a preferential pentameric arrangement in the human brain and showed that gene co-expression patterns of GABA<sub>A</sub>R subunits are highly stereotypical within structures characterized by recurrent cytoarchitecture<sup>16</sup>. Here, we shift our focus to glutamatergic AMPARs. These receptors mediate the post-synaptic depolarization necessary for fast excitatory...
transmission and participate in the long-term potentiation and depression necessary for appropriate function of cortical circuits. Unlike the highly diverse and heterogeneous GABAARs, AMPARs are constructed as dimers of dimers from a pool of only four subunits (GluA1–GluA4) transcribed from four genes (GRIA1–GRIA4). This construction has been suggested to be highly dependent on afferent type and can change with synaptic activity. Stoichiometric alterations and single subunit defects of AMPAR tetramers have been correlated with many brain disorders. For instance, reductions in GluA2 and GluA3 subunits were found in the hippocampus of patients suffering from mild to severe forms of Alzheimer’s disease, and alterations in GluA1-4 expression level have been observed in different human brain structures of people diagnosed with schizophrenia. Therefore, understanding the organizational layout of these receptors should prove useful in determining pathological remodeling of AMPARs, and consequently alterations to the E/I balance.

Our analysis was performed at different levels using publicly available data (Fig. 1a). First, we examined major relationships between AMPAR subunit expression levels across a wide variety of structures from a microarray dataset from the Allen Brain Institute. This analysis, although low in number of subjects (n = 6), was extensive in its anatomical coverage, which ranged across 111 brain structures. To analyze inter-individual differences, we utilized RNA-sequencing (RNA-Seq) data from the Aging, Dementia, and Traumatic Brain Injury (ADTBI) study.
Results

Global and region-specific expression patterns of AMPAR subunits. We first analyzed the microarray data from the Allen Brain Atlas at the whole brain level. From this data, we selected the most representative probe for each gene (Supplementary Data 1) using exploratory factor analysis (Fig. 1a; Supplementary Fig. S1 for selection process) and divided the brain into major regions, structures, and substructures following the Allen Brain Atlas’ nomenclature (Supplementary Data 2). The initial global analysis after adjusting by age (Supplementary Data 3; Supplementary Fig. S2) revealed high expression levels (Log2 > 4) for all four subunits (Fig. 1b). GRIA2 was the most expressed AMPAR subunit in the whole brain followed by GRIA4, GRIA1 and GRIA3 (Fig. 1b). However, there was high variability across brain regions in expression levels for all four subunits, particularly for GRIA1 and GRIA4 (Supplementary Fig. S2). For example, structures such as the hippocampal formation (HiF; CA1-4 and DG) and the internal globus pallidus (GP) were outliers at the global expression level for GRIA1. A nested ANOVA analysis of gene expression, where substructures are nested within structures to account for substructure dependencies, confirmed different levels of expression for all 4 subunits across major structures and across some substructures within the same region (Supplementary Data 4). To address this regional variability, unsupervised hierarchical clustering was performed to group together structures with similar expression profiles (Fig. 1c). The strength of these clusters was examined through non-parametric bootstrap resampling to calculate both the approximately unbiased probability (AU) and bootstrap probability (BP) p values (Supplementary Fig. S3). The structures were first split between structures of telencephalic origin and those of diencephalic, mesencephalic, metencephalic, and myelencephalic origins (AU > 0.95, Fig. 1c). Within the latter group, further divisions based on structure ontogeny were observed. Regions of metencephalic origin exhibited very low levels of GRIA3 expression (Log2 < 4) and high levels of both GRIA2 and GRIA4 (Log2 > 9) while regions originating from the diencephalon, mesencephalon, and myelencephalon exhibited lower GRIA2 expression (Log2 < 9.9; Fig. 1c). There was very strong evidence (AU > 0.95) for clustering of metencephalic structures at nearly all levels. Structures from these three regions were then further split between those with GRIA1 expression levels comparable to metencephalic structures (Log2 ~ 8) and those with very low GRIA1 expression levels (Log2 < 6). However, evidence for clustering of substructures within these structures was not as robust (AU < 0.95).

By contrast, the telencephalic cluster exhibited a high degree of homogeneity, particularly within cerebral cortical regions where high levels of GRIA2 and GRIA4 (Log2 > 9.5) and lower levels of GRIA1 and GRIA3 (Log2 < 9.5) were observed (Fig. 1c). Some slight variations were introduced with the cerebral nuclei which typically exhibited lower levels of GRIA4 (Log2 < 9) and a structure-dependent increase in GRIA1 expression (Log2 > 9). However, this cluster was not as robust as that of the cortical regions with an AU ± S.E.M of only 0.942 ± 0.005 (Supplementary Fig. S3). The hippocampal formation differed greatly from all other structures with high GRIA1-3 expression (Log2 > 10) but low GRIA4 expression (Fig. 1c), and the evidence for this cluster was very strong (AU ± SE = 0.999 ± 0.001).

The hippocampal formation contained the substructures with the greatest fold enrichment of GRIA1 (CA4), GRIA2 (DG) and GRIA3 (CA1) expression (Table 1). GRIA1 was particularly enriched in the CA4 with an expression ~ 1,000% larger than the global brain average.

Proportional contributions of AMPAR subunits and inter-individual variability. To determine the contribution of all available transcripts for AMPAR subunits per brain region, we calculated the average percentage of expression for each subunit as was previously done for GABAAR subunits16; that is to calculate the sum of probe intensities across AMPAR subunits and then determine to proportional contribution of the intensity of each probe to this sum. This proportional contribution is an estimate that represents the available

| Subunit | Major region | Structure | Substructure | Expression (log2) | Fold enrichment |
|---------|--------------|-----------|--------------|------------------|-----------------|
| GRIA1   | Cerebral cortex | HiF | CA4         | 11.54            | 10.12           |
| GRIA2   | Cerebral cortex | HiF | Dentate Gyrus | 12.63            | 3.31            |
| GRIA3   | Cerebral cortex | HiF | CA1         | 10.46            | 3.75            |
| GRIA4   | Tectum         | VT | Reticular nucleus | 11.33            | 2.23            |

Table 1. Human brain substructures with greatest anatomical enrichment. The expression (log2) represents the average per brain region. The fold enrichment is the substructure’s value divided by the global average. HiF Hippocampal Formation, VT ventral thalamus.

(https://aging.brain-map.org/), which included a much higher number of individuals (n = 50), but was limited to only four regions: the hippocampus (HIP), temporal cortex (TCx), parietal cortex (PCx), and forebrain white matter (PWM). Together, these complementary analyses created a signature of AMPAR subunit expression in the healthy human brain. To determine whether transcripts for AMPAR subunits and GABA_ARs subunits were balanced, we then calculated and analyzed the ratio between AMPAR and GABA_Ar subunit expression at both the whole brain and structural levels. This generated an estimate of the most basic signature of the transcriptional E/I (E/I) ratio. Finally, to investigate the usefulness of these signatures, we applied our method to single cell data from phenotypically characterized interneurons in mice and to RNA-sequencing data from the Allen Brainspan study that ranged from early embryonic stages to adulthood in humans.
pool of AMPAR subunits mRNA in each brain region, structure, or substructure, and normalizes distinct levels of expression between different brain areas (Fig. 2). As before, the cerebral cortex was highly homogenous with a greater contribution from GRIA2. GRIA1 contributed more in the hippocampus and cerebral nuclei, and noncortical regions were more variable and dominated by GRIA4 (Fig. 2). Further analysis revealed opposing expression patterns between some subunits. For example, the hippocampus showed a gradual decrease in the proportional contribution of GRIA2 from the DG to CA4 that was paralleled by a gradual increase in the contribution of GRIA1 (Fig. 2). Similarly, GRIA4 and GRIA2 followed opposing patterns of expression in the myelencephalon and thalamus compared to the cortex (Fig. 2).

It is important to note that microarray data is highly influenced by technical factors, and intensity levels are typically compared across experimental units for the same probe, not across probes within the same experimental unit as we are doing in this study; therefore we compared the proportional contribution of GRIA subunits in the temporal cortex using microarray (n = 6 subjects with 12 temporal lobe substructures per subject) with that using RNA-Seq of the ADTBI study (n = 50 subjects), which is a better measure of mRNA expression. Proportional contributions in both datasets showed similar relationships across GRIA1, GRIA2 and GRIA3 subunits providing support for our microarray analysis of these subunits; the proportion of GRIA4 seems to be overrepresented in microarray compared to RNA-Seq (Supplementary Fig. S4), a probable consequence of choosing the probe with higher correlations across all GRIA4 probes in the whole brain, that happened to also have the highest intensity, highlighting some of the technical limitations aforementioned (Supplementary Fig. S1).

In addition to differences in level of gene expression of isolated subunits, we also explored the variability in the collective organization of AMPAR subunits between individuals. To do so, we followed the previous approach used for GABA\(_A\)Rs\(^{16}\). First, inter-individual variability, within and across substructures of the microarray study,
was quantified by calculating the Euclidian distances ($d_i$) between the expression levels of GRIA subunits in each region and correlating these results. For example, the correlation coefficient ($R$) between the superior frontal gyrus across individuals ($d_{SFG,SFG}$) was higher ($0.94 \pm 0.05$; Mean $\pm$ SD) than in the dentate gyrus ($d_{DG,DG}$ = $0.81 \pm 0.18$), indicating that the dentate gyrus is more variable across individuals than cortical regions (Supplementary Fig. S5). To extend the analysis and compare brain regions at the level of substructures the frontal operculum ($fro$), which is the frontal-most structure displayed in the Allen Atlas, was chosen to construct a reference (consensus) against which all other 110 substructures were compared ($d_{dcFro}$). Each substructure from each subject was then measured against this reference (Fig. 3a and Supplementary Fig. S6). Cerebral cortical structures were mostly homogenous with correlation coefficients close to one, indicating a high similarity to other cerebral structures and to the frontal operculum (Fig. 3a and Supplementary Data 6 for statistical analyses). Hippocampal structures, as expected, differed greatly from other cerebral regions ($r < 0.6$) and were more variable between individuals. However, the microarray study included only six subjects, so we turned to the ADTBI study ($n = 50$ subjects) to examine inter-individual variability more accurately. For this analysis, the parietal cortex was chosen as the reference structure against which the other three structures were compared. The parietal and temporal cortices were identical with virtually no global variability between all 50 subjects (Fig. 3c). Surprisingly, unlike what was previously observed for GABA$\text{\textsubscript{A}}$ subunits$^{16}$, the forebrain white matter was very similar to the cortices with low inter-individual variability (Fig. 3c). The hippocampus, in agreement with the microarray data, was very different from the other structures and highly variable between subjects (Fig. 3c). This difference arose primarily out of a higher proportional contribution of GRIA1 in the hippocampus as opposed to the other three regions.

**Global and region-specific patterns of transcriptional excitation–inhibition ratio markers.** To combine these results with the expression patterns in GABA$\text{\textsubscript{A}}$Rs subunits we previously uncovered, we used the microarray gene expression data as a marker of the E/I ratio by dividing the sum of probe intensities across AMPAR subunits by the sum of probe intensities across GABA$\text{\textsubscript{A}}$R subunits (Supplementary Data 7). This is the most basic and non-subunit specific estimation of the transcriptional foundations of the E/I balance based on gene expression for the principal gates of excitation and inhibition. Because age correction may change GABA$\text{\textsubscript{A}}$Rs and AMPARs differently and lead to distinct ratios in each subject, we compared the $tE/I$ ratio

![Figure 3. Inter-individual variability across brain regions. (a) Euclidian distances of AMPAR subunits per structure (111) per subject ($n = 6$; $d_i$) were correlated against a standard (frontal operculum, $fro$; $d_{i,fro}$) using microarray data corrected by age. Structures are organized in a rostro-caudal order as in Fig. 2. Correlation coefficients ($R$) closer to one indicate higher similarity in expression patterns between the substructures and the standard. Each point is a single subject containing the collective information of all four AMPAR subunits. (b) Same information as in a but grouped by structure. (c) Euclidian distances of AMPAR subunits per structure (4) and per subject ($n = 50$) were correlated against the Parietal Cortex ($PCx$) using ADTBI data corrected by age. The hippocampus ($HIP$) is drastically different from the other three regions and also highly variable between subjects. FWM, forebrain white matter, TCx, Temporal cortex. Each point is a single subject.](https://doi.org/10.1038/s41598-020-68165-1)
using gene expression levels with and without age correction. Both ratios were linearly correlated ($R^2 = 0.789$; $p < 0.0001$) and no significant difference in each substructure was observed between both ratios, although more dispersion was observed using non-corrected data ($p > 0.29$ in all cases by Welch’s testing means equal, allowing SD not equal, Fig. 4 and Supplementary Fig. S7). For simplicity, we decided to use the $tE/I$ based on non-adjusted gene expression values. Expression patterns across cerebral cortical regions were relatively homogenous with $tE/I$ ratios close to one (Fig. 4a). Similarly, cerebellar regions exhibited highly homogenous $tE/I$ ratios though these were lower than those found in the cerebral cortex (Fig. 4a). Hippocampal structures exhibited a higher $tE/I$ ratio overall compared to the rest of the brain, apart from the reticular nucleus (Supplementary Data 7). Within the hippocampus, the $tE/I$ ratio gradually decreased from DG to CA2 and then increased from CA3 to CA4 where AMPAR subunit expression was nearly double that of GABA$_A$Rs subunit expression (Fig. 4a). Cerebral nuclei and thalamic regions were more variable with more caudal regions tipping towards greater inhibition (Fig. 4a). The structure with the greatest difference was the reticular nucleus in the ventral thalamus ($tE/I_{mean} = 3.8$, Fig. 4a). These observations were further supported by the ADTBI RNA-Seq dataset. The $tE/I$ ratios using non-adjusted by age gene expression in the parietal and temporal cortices were remarkably homogenous and both were significantly different from the hippocampus, which was both higher and more spread (Fig. 4b; Supplementary Data 8 for statistical analysis). No differences in the $tE/I$ ratio in the temporal and parietal cortices, hippocampus or forebrain white matter by sex were observed ($p > 0.11$ double tailed $t$ test; Supplementary Fig. S8). Even though the $tE/I$ ratio in the forebrain white matter was more variable across individuals, the $tE/I$ ratios from these structures were remarkably constant overall, indicating that AMPARs and GABA$_A$Rs were highly correlated within brain structures, although with different slopes (Fig. 4c). Such strong correlations between AMPARs and GABA$_A$Rs suggest that transcriptional correlations start at the cellular level. Indeed, the

**Figure 4.** Structural differences in transcriptional E/I ratio in the human brain. (a) Mean ± S.E.M. of the $tE/I$ (total probe intensity of AMPAR (ΣAMPARs) subunits over total probe intensity of GABA$_A$R (ΣGABA$_A$Rs) subunits) across the whole human brain. The ratios before (red) and after age correction (blue) were calculated for each region and were not different within brain structures. Structures were ordered along the anteroposterior axis. (b) The $tE/I$ ratio calculated from the ADTBI RNA-Seq data. Each point represents a single subject. (c) Plotting total AMPAR subunit against total GABA$_A$Rs subunit expression levels from ADTBI RNA-Seq data. Strong positive linear correlations were detected within all four structures ($p < 0.001$ for all), though with different slopes. Histograms represent the data distribution for the cohort in either axis.
electrophysiological E/I (eE/I) balance has been demonstrated in excitatory pyramidal neurons as well as inhibitory interneurons\(^{31-33}\). However, analyzing the eE/I at the single cell level by single-cell RNA-Seq (sc-RNA-Seq) in humans is challenging due to the inherent low coverage of the method (approx. 10–20% of the transcriptome) and the potential mRNA degradation by postmortem interval and agonal factors\(^{34}\). Therefore, we used a high resolution sc-RNA-Seq dataset from six phenotypically identified cortical interneurons from mice\(^{35}\) to test for intra- and inter-variability of the E/I ratio at the cell-type level (Fig. 5). For this analysis, we used the dataset from Paul et al.\(^{35}\) available on the gene expression omnibus (GEO; GSE92522) which provided data from anatomically and physiologically identified cortical chandelier cells (CHCs), perisomatic fast-spiking basket cells (PVBCs), long-projecting interneurons (LPCs), Martinotti cells (MNC), interneuron-selective cells (ISCs) and CCKC basket cells (CCKCs).

These cells demonstrated strong linear correlations between AMPARs and GABA\(_A\)Rs with E/I ratios averaging between 0.66 to 1.8 (Fig. 5b; Table 2). Only MNCs were significantly different with a tenfold higher E/I ratio (P < 0.0001; ANOVA followed by Tukey–Kramer HSD), and CHC1 cells were significantly different from CCKC cells (P = 0.022). No other cellular types were significantly different from the others (P > 0.05).

**tE/I ratio expression patterns change during development.** To demonstrate the flexibility and capabilities of our approach, we applied our workflow to developmental data gathered by Allen Brainspan focusing on the cortex since it has more data than the other structures studied in that dataset (Supplementary Data 9). The

---

**Figure 5.** Cell-type differences in transcriptional E/I ratio in the mouse brain. (a) Transcriptional E/I ratio of phenotypically identified GABAergic cell-type\(^{35}\). Ratios were plotted on a logarithmic scale. Each point represents a single cell nucleus analyzed. (b, c) Plotting of total AMPARs expression against total GABA\(_A\)Rs subunits expression using normalized data as downloaded (b) or after Log2 transformation to normalize the distribution (c). Continuous lines show significant Pearson correlations for all cell-types (p < 0.001). Histograms represent the data distribution for the cohort in either axis.
array analysis fit the clustering scheme created by Gold et al. in the rodent model using in situ hybridization, expression patterns are highly stable well into adulthood. This shared pattern may result from a shared embryonic stable

table 2. Cell-type transcriptional E/I ratios. The mean expression represents the average per cell type. CCKC basket cells, CHC1 chandelier cells 1, CHC2 chandelier cells 2, ISC interneuron-selective cells, LPC long-projecting interneurons, MNC Martinotti cells, PVBC perisomatic fast-spiking basket cells. *Different from CCKC (p = 0.022). **Different from all other cells (p < 0.0001).

| Cell-type | Number of cells | Mean ± SD |
|-----------|-----------------|-----------|
| CCKC      | 64              | 0.66 ± 0.13 |
| CHC1      | 80              | 1.7 ± 0.38* |
| CHC2      | 52              | 1.0 ± 0.43  |
| ISC       | 63              | 0.80 ± 0.26 |
| LPC       | 136             | 1.1 ± 0.22  |
| MNC       | 62              | 10 ± 6.3**  |
| PVBC      | 127             | 0.96 ± 0.13 |

E/I ratio across cortical regions was initially spread across a large range, but all structures reached a relatively stable E/I ratio by one postnatal year (Fig. 6a). Interestingly, the E/I ratio in the cortex using the Brainspan dataset was smaller than the one in the ADTBI study, although both used RNA-Seq methods (E/I = 0.31 ± 0.0061 (mean ± S.E.M) in Brainspan vs 0.5842 ± 0.0045 in ADTBI; double tailed t test, p < 0.001). During pre-natal development, some structures, namely the dorsolateral and inferotemporal temporal cortices, auditory primary cortex, and posteroverentral parietal cortex exhibited high E/I ratios that rapidly declined approaching birth (Fig. 6a). Other structures, such as the primary somatosensory and anterior cingulate cortices rose slightly during early development and peaked at around 19 post-conception weeks (pcw) before declining towards the steady state observed in adulthood (Fig. 6a). Overall, the cortical E/I ratio stabilizes after birth. Closer inspection of the data revealed that the developmental reduction of E/I resulted from a gradual increase in GABA_A R subunit expression levels while AMPAR subunit levels remained stable (Supplementary Fig. 9). In addition to the E/I ratio, we also analyzed the patterns of expression of the sodium–potassium–chloride cotransporter (NKCC1, SLC12A2) and the neuron-specific chloride–potassium symporter 5 (KCC2, SLC12A5) which have been demonstrated to change with development and determine whether GABA_A Rs are inhibitory or excitatory. KCC2 expression increases, while NKCC1 expression decreases during development. To do so, we compared the ratio of SLC12A5 to SLC12A2 with the E/I ratio and found that during prenatal development, the higher E/I ratio corresponded to a low SLC12A5/SLC12A2 ratio while postnatal development exhibited no relationship between the two ratios (Fig. 6c). KCC2 expression is the main driver of variability in the KCC2/NKCC1 ratio. Interestingly, KCC2 gene expression was linearly correlated with the global expression of GABA_A R and AMPAR subunits.

Discussion

Our clustering analysis indicates that the co-expression patterns of AMPARs subunits is region specific according to the ontogenic origin of the brain structures. It is quite remarkable that with the expression of only 4 genes, major brain regions could be separated by unsupervised clustering analysis. This suggests that AMPARs may act as strong gene markers in synaptic components previously described by bioinformatics analysis. Our microarray analysis fit the clustering scheme created by Gold et al. in the rodent model using in situ hybridization, though with some differences. For instance, while the higher contribution of GRIA1 in the amygdala and hippocampus relative to other cerebral structures was mirrored in our analysis, it was to a much lesser extent, only 20–30% as compared to the 50% reported. Further, the cerebellum exhibited a high GRIA4 profile that is inconsistent with any of the reported clusters. On the other hand, cortical regions aligned fully with the pattern of ~ 50% GRIA2 contribution. While data from mice showed GRIA3 and GRIA4 enrichment relative to GRIA1 in the thalamus, our analysis found only GRIA4 enrichment. Both differences may be related to the high anatomical specificity of structures in hybridization studies compared to the microarray data that we used for our analysis, or it could be due to divergent characteristic between species.

One potential drawback of our microarray analysis is the low number of samples which reduces the statistical validity of our findings. However, our main aim for this microarray data was to generate a transcriptional signature of structural clusters based on the relative expressions of the different GRIA subunits. For this purpose, we utilized bootstrapping to measure the strength of our clusters. As a result, we discovered very strong evidence for upper-level divisions between cortical regions of telencephalic origin and regions of myelencephalic, metencephalic, and diencephalic origin, as well as the existence of distinct structural clusters, such as the hippocampal formation. While these results are promising, additional samples would boost the power of this study.

Our observed region-specificity in AMPAR subunit expression correlates with the differing complexity in pathways and cytoarchitecture and with embryonic origin. Indeed, previous studies have found that a selection of genes exhibit region-specific expression largely dependent on their embryonic structure of origin. For brain structures with recurrent cytoarchitecture, such as the cerebral and cerebellar cortices, AMPAR subunit expression patterns are highly stable well into adulthood. This shared pattern may result from a shared embryonic origin such as the pallial regions for cortical glutamatergic cells and the rhombic lip for the glutamatergic cells in the cerebellar cortex. In contrast, the more heterogenous patterns observed in deep cerebral nuclei, thalamic nuclei, and mesencephalic structures are congruent with more heterogenous cytoarchitecture of these regions and may underly the diverse modulatory functions of these nuclei. Recent studies have begun to demonstrate the importance of using functional rather than neuroanatomical principles for the analysis of synaptic proteins.
**Figure 6.** tE/I ratio and KCC2/NKCC1 ratio in the developing human brain. (a) tE/I ratios were calculated in different cortical regions against age (from 12 postconceptional weeks to 40 years, shown as categorical values in the plot) using data from the Allen Brainspan project (n = 40 subjects). Each point represents a single cortical region per subject. Spline smooth curves, here and in (b), show the principal trend in values for AMPA and GABA<sub>A</sub> receptors. Please see Supplementary Data 9 for detailed number of subjects per region. (b) The graph shows the mean ± S.E.M. of cortical tE/I in each subject per region. The cortical tE/I in each subject is the average of all cortical regions available for each subject. (c) Lack of correlation between the cortical tE/I ratio in each subject and the KCC2/NKCC1 ratio. KCC2 is the potassium-chloride transporter member 5 produced by the SLC12A5 gene and its expression increases during development, while NKCC1 is the Na–K–Cl cotransporter produced by SLC12A2 gene. The relationship between these transporters defines the direction of chloride permeability in GABA<sub>A</sub>Rs. (d, e) KCC2 expression is linearly correlated with the increase of GABA<sub>A</sub>Rs (d) and AMPARs (e) after birth. Points are color coded as prenatal (blue), near birth (magenta), or post-natal (orange). Linear correlations are for data grouped as pre and postnatal development.
similar analytical approaches with gene expression may help to explain the heterogenous expression of AMPA receptor subunits in the thalamic nuclei and mesencephalic structures in future studies.

Interestingly, the hippocampal formation differed greatly from the cerebral cortices, exhibited the highest expression levels of all AMPAR subunits compared to all the other brain structures, and showed the largest variability across individuals when using the correlation of Euclidian distances between subunits. The heightened expression may result from the unique neurogenesis in the hippocampus49, the rapid regulation of synaptic AMPA receptors necessary for its role in learning50-52 and the high neuronal density in this area. Previous studies have shown the necessity of GRIA1 subunit in long-term potentiation and contextual learning at the CA3-CA1 synapse52,53 and the rapid subunit specific turnover of AMPARs in response to alterations in neuronal activity54,55. Moreover, the subgranular zone of the dentate gyrus is one of the only regions in the human brain where adult neurogenesis occurs56. As a part of this process, neural stem cells enter hippocampal circuits after modulation by various environmental and cell-intrinsic factors57,58. This intense plasticity suggests that hippocampal AMPAR subunit expression varies greatly depending on the subject and environment. Our analyses of both the microarray and RNA-seq studies support this hypothesis as inter-subject variability in AMPAR subunits expression is particularly high in the hippocampal formation while AMPAR subunits expression in the cerebral cortex is highly stereotypical across individuals.

A major interest in our analysis was to determine whether the mRNA for the receptors involved in establishing the E/I ratio were balanced across brain structures in human. Because many subunits of GABA\(_{\alpha}R\)s and AMPARs follow complementary or opposing patterns of expression, we used the most conservative approach which includes all potential receptor configurations by adding transcripts for all subunits when estimating the E/I ratio. Our analysis found that the total amount of AMPARs and GABA\(_{\alpha}R\)s is highly correlated in a region-specific manner. Therefore, the E/I ratio remains highly stable across substructures within major regions such as the cerebral and cerebellar cortices but varies gradually within the hippocampus and amygdala, suggesting different balances within these regions. Our results using the ADTBI study did not find sex differences for the E/I ratio, suggesting that in healthy individuals, sex differences for particular GABA\(_{\alpha}R\)s subunits like higher cortical \(\alpha1\) in men59, may be balanced by compensatory mechanisms regulating the expression of AMPARs.

Our study unfortunately cannot distinguish between synaptic and non-synaptic receptors; therefore, our inferences are limited to estimates of total transcription levels per brain structure in the microarray and ADTBI analyses, or to the cell-type level in the single cell analysis. Nevertheless, the electrophysiological activity of synaptic and non-synaptic receptors GABA\(_{\alpha}R\)s strongly correlates within single neurons60, and AMPARs show high lateral mobility between non-synaptic and synaptic regions61 suggesting that whole transcription levels may at least partially reflect the activity-dependent need for AMPAR and GABA\(_{\alpha}R\) subunit transcription. Recent studies have also shown that electrophysiological measures of global eE/I by recording cortical synaptic AMPARs and GABAARs are highly correlated62. Future studies integrating transcriptomic with electrophysiological measures of the E/I will help to understand deviations of this balance in non-physiological states.

Our analysis also shows that the E/I ratio is established near birth. While AMPAR subunits expression is relatively stable during the embryonic stages, GABA\(_{\alpha}R\) subunit expression gradually increases until stabilizing near birth reducing the E/I ratio to a stable value. This indicates that GABA\(_{\alpha}R\)s are highly important in setting the E/I ratio. In fact, development of GABA\(_{\alpha}R\)s from immature GABA\(_{\beta2}\)- and GABA\(_{\gamma3}\)-enriched to mature GABA\(_{\alpha1}\)-enriched receptors has been observed to play a major role in synaptic plasticity in the cat visual cortex during the early post-natal phases63 and a fundamental role in the establishment of circuits and neuronal-firing properties within the mouse cortex64. On the other hand, although the amount of AMPARs remains quite stable during the development, the subunit expression and receptor assembly profile change markedly leading to a shift in receptor properties. As described by Henley and Wilkinson65, subunit adjustment is directly related to function; for instance, GRIA1 is developmentally restricted while expression of GRIA2, which produces \(\gamma\)-map.org/display/human brain/Documenta tion. Only brain regions where data for all AMPAR and GABA\(_{\alpha}R\) subunits were measured were used for the analysis (111 brain structures). For gene expression measurements
the Allen Institute used a custom design (by Beckman Coulter Genomics) Agilent 8 × 60 K array that includes the 4 × 44 K Agilent Whole Human Genome probe set supplemented with an additional 16,000 probes. The upper limit for detection was 2 and was used as part of the initial quality control as described in the white papers for microarray survey found at https://help.brain-map.org/display/humanbrain/Documentation. Normalization methods are also documented extensively in the white papers; briefly the data was normalized within each dis-

section batch through multivariate local regression fitting, each brain via quantile–quantile mapping of averages, and across all brains by aligning control samples.

For analysis of variability across individuals, normalized RNA-Seq Fragments Per Kilobase Million (FPKM) data from the Aging, Dementia, and Traumatic Brain Injury study (https://aging.brain-map.org/download/index) covered four regions from 56 healthy control subjects (35 males and 21 females) between the ages of 78 and 99 years of age. Detailed documentation of the ADTBI study, including tissue collection, tissue processing, and RNA-Sequencing and quantification, can be found at: https://help.brain-map.org/display/aging/Documentation. Briefly, RNA-Sequencing was done through Illumina TruSeq (random hexamer first strand cDNA synthesis with rRNA depletion and frag-mentation) on Illumina HighSeq 2,500 using v4 chemistry, producing a minimum of 30 M 50 bp paired-end clusters per sample. FPKM gene quantifications were normalized via TBT normalization and corrected for RNA quality and batch effects.

Normalized developmental RNA-seq Reads Per Kilobase Million (RPKM) data from the BrainSpan atlas (Gencode v10 summarized to genes; https://www.brainspan.org/static/download.html) covered up to 26 regions from 42 subjects ranging from eight weeks post conception to 40 years of age. Detailed documentation of the developmental data, including Institutional Review Board approval for the tissue collection, donor and sample metadata, and transcriptome profiling methods, can be found at: https://help.brain-map.org/display/devhumanbrain/Documentation. Briefly, RNA-Sequencing was done with poly(A) selection and normalized by the addition of spike-ins.

For single cell analysis we used the processed high resolution scRNA-seq unique Transcripts Per Million (uTPM) dataset downloaded from the GEO expression omnibus accession number: GSE9252235. This dataset com-
addition of spike-ins.

prises single cell transcriptomes of anatomically and physiologically characterized GABAergic neurons. Paul et al. developed combinatorial recombine driver lines to identify and isolate 6 types of neurons. Detailed methods, including the approval for the study, can be found in the author’s manuscript35. Sequencing was performed with poly(A) linear amplification followed by Illumina TruSeq and normalized to a single cell’s total unique counts across all genes; approximately 9 k genes per cell were detected35. Our study focused on the analysis of the databases described above. All the methods to generate the original data were performed in accordance with the relevant guidelines and regulations of the respective authors’ institutions as described in the documentation accessed by the provided links.

**Analysis of AMPAR subunits.** All analysis was done as previously stated for GABA_\_Rs\^16. The normalized transformed data (Log_\_2) microarray data from the Allen Brain Atlas used 17 probes to measure the expression of 4 AMPARs genes, 4 probes for each GRIA1-3 and 5 probes for GRIA 4. To avoid redundant clustering due to collinearity between probes for the same gene, only the most representative probe for each gene as determined via principal axis Exploratory Factor Analysis with no rotation in JMP 14Pro were used. While the effects of sex and ethnicity were negligible on any of the datasets analyzed, corrections were made for age effects via linear regression in JMP 14. As before, all expression patterns were represented by the Mean (M) ± Standard Deviation (SD) unless otherwise stated. The proportional contributions of each subunit were expressed as percentages of untransformed (non-Log2) probe intensity level of each subunit to the sum of probe intensities across all subunits. Unsupervised hierarchical clustering was done using Ward’s minimum variance method in JMP 14. Bootstrap probability (BP) and approximately unbiased (AU) p-values were calculated through 10,000 boot-

straps with pvclust in RStudio using ward.D2 and Euclidean distances as the clustering method and distance measures, respectively36. Pearson product-moment correlations were used to measure both inter-structure and inter-subject variance. Euclidean distances of each structure or subject (d\_i) were correlated against a chosen consensus (d\_cons). For subjects, d\_i was calculated from the average of all subjects; for structures, the frontal operculum (fro) and parietal cortex (PCx) were used. The Euclidean distances were calculated in RStudio as follows:

\[ d_i \left( GRIAx - GRIAy \right) = \sqrt{GRIAx - GRIAy}^2 \]  

\[ d_i \left( GRIAx - GRIAy \right) = \sqrt{\sum_{i=1}^{N} (GRIAx_i - GRIAy_i)^2} \]

where d\_i is the Euclidean distance between two subunits, given by x and y, per subject or structure, represented by the integer i, and d\_cons is the consensus Euclidean distance, for all n subjects or structures.

**Analysis of the excitation–inhibition ratio.** The E/I ratio was calculated by dividing the sum of all AMPAR subunit expression by the sum of all GABA_\_Rs subunit expression. Pearson product-moment corre-

lations were then run on these ratios to measure inter-structure variance. Data from the microarray (Log2, uncorrected and corrected for age), ADTBI (normalized FPKM, uncorrected and corrected for age), single-cell (uTPM), and BrainSpan (normalized RPKM) studies were all analyzed separately. In addition to the E/I ratio, the expression of NKCC1 (SLC12A1), NKCC2, (SLC12A1) and KCC2 (SLC12A5) were also analyzed in a fashion similar to the AMPARs and GABA_\_Rs.
**Statistics and reproducibility.** All data analysis was performed, and plots constructed, in both JMP 14 and RStudio running R3.6.3 and the mosaic package for R Markdown. All R code is available upon request.

**Data availability**

All original data was downloaded from publicly available databases with links provided in the manuscript. All data after adjustments are included in supplementary data.

Received: 11 February 2020; Accepted: 8 June 2020
Published online: 09 July 2020

**References**

1. Van Vreeswijk, C. & Sompolinsky, H. Chaos in neuronal networks with balanced excitatory and inhibitory activity. *Science* **274**, 1724–1726 (1996).
2. Van Vreeswijk, C. & Sompolinsky, H. Chaotic balanced state in a model of cortical circuits. *Neural Comput.* **10**, 1321–1371 (1998).
3. Froemke, R. C., Merzenich, M. M. & Schreiner, C. E. A synaptic memory trace for cortical receptive field plasticity. *Nature* **450**, 425–429 (2007).
4. Dorrn, A. L., Yuan, K., Barker, A. J., Schreiner, C. E. & Froemke, R. C. Developmental sensory experience balances cortical excitation and inhibition. *Nature* **465**, 932–936 (2010).
5. Sun, Y. J. et al. Fine-tuning of pre-balanced excitation and inhibition during auditory cortical development. *Nature* **465**, 927–931 (2010).
6. Li, Y. T., Ma, W. P., Pan, C. J., Zhang, L. I. & Tso, H. W. Broadening of cortical inhibition mediates developmental sharpening of orientation selectivity. *J. Neurosci.* **32**, 3981–3991 (2012).
7. Rubin, R., Abbott, L. F. & Sompolinsky, H. Balanced excitation and inhibition are required for high-capacity, noise-robust neuronal selectivity. *Proc. Natl. Acad. Sci. USA* **114**, E9366–E9375 (2017).
8. Sillito, A. M. The contribution of inhibitory mechanisms to the receptive field properties of neurons in the striate cortex of the cat. *J. Physiol.* **250**, 305–329 (1975).
9. Dichter, M. A. & Ayala, G. F. Cellular mechanisms of epilepsy: a status report. *Science* **237**, 157–164 (1987).
10. Gao, R. & Penzes, P. Common mechanisms of excitatory and inhibitory imbalance in schizophrenia and autism spectrum disorders. *Curr. Mol. Med.* **15**, 146–167 (2015).
11. Nelson, S. B. & Valakh, V. Excitatory/inhibitory balance and circuit homeostasis in autism spectrum disorders. *Sci. Rep.* **8**, 684–698 (2018).
12. Foss-Feig, J. H. et al. Searching for cross-diagnostic convergence: neural mechanisms governing excitation and inhibition balance in schizophrenia and autism spectrum disorders. *Biol. Psychiatry* **81**, 848–861 (2017).
13. Busche, M. A. & Konnerth, A. Impairments of neural circuit function in Alzheimer’s disease. *Philos. Trans. R. Soc. B* **371**, 20150429 (2016).
14. Ren, S. Q. et al. Amyloid β causes excitation/inhibition imbalance through dopamine receptor 1-dependent disruption of fast-spiking GABAAergic input in anterior cingulate cortex. *Sci. Rep.* **8**, 1–10 (2018).
15. Tätti, R., Haley, M. S., Swanson, O. K., Tsila, T. & Maffei, A. Neurophysiology and regulation of the balance between excitation and inhibition in neocortical circuits. *Biol. Psychiatry* **81**, 821–831 (2017).
16. Sequeira, A., Shen, K., Gottlieb, A. & Limon, A. Human brain transcriptome analysis finds region- and subject-specific expression signatures of GABAAR subunits. *Commun. Biol.* **2**, 1–14 (2019).
17. Malenka, R. C. & Bear, M. F. LTP and LTD: an embarrassment of riches. *Neuron* **44**, 5–21 (2004).
18. Olsen, R. W. & Sieghart, W. GABAA receptors: subtypes provide diversity of function and pharmacology. *Neuropharmacology* **56**, 141–148 (2009).
19. Tichelaar, W. A., Waffling, M., Keinänen, K., Stark, H. & Madden, D. R. The three-dimensional structure of an ionotropic glutamate receptor reveals a dimer-of-dimers assembly. *J. Mol. Biol.* **405**, 1–18 (2000).
20. Zhao, Y., Chen, S., Swensen, A. C., Qian, W.-J. & Gouaux, E. Architecture and subunit arrangement of native AMPA receptors elucidated by cryo-EM. *Science* **360**, 355–362 (2019).
21. Hollman, M. & Heinemann, S. Cloned glutamate receptors. *Annu. Rev. Neurosci.* **17**, 31–108 (1994).
22. Töth, K. & McBain, C. J. Afferent-specific innervation of two distinct AMPA receptor subtypes on single hippocampal interneurons. *Nat. Neurosci.* **1**, 572–578 (1998).
23. Rubio, M. E. & Wenth Holden, R. J. Glutamate receptors are selectively targeted to postsynaptic sites in neurons. *Neuron* **18**, 939–950 (1997).
24. Liu, S. Q. J. & Cull-Candy, S. G. Synaptic activity at calcium- permeable AMPA receptors induces a switch in receptor subtype. *Nature* **405**, 454–458 (2000).
25. Liu, S. J. & Cull-Candy, S. G. Activity-dependent change in AMPA receptor properties in cerebellar stellate cells. *J. Neurosci.* **22**, 3881–3889 (2002).
26. Salpietro, V. et al. AMPA receptor GluA2 subunit defects are a cause of neurodevelopmental disorders. *Nature Communications* **10**, 1–16 (2019).
27. Thorns, V., Mallory, M., Hansen, L. & Masliah, E. Alterations in glutamate receptor 2/3 subunits and amyloid precursor protein expression during the course of Alzheimer’s disease and Lewy body variant. *Acta Neuropathol.* **94**, 339–548 (1997).
28. Rubio, M. D., Drummond, J. B. & Meador-Woodruff, J. H. Glutamate receptor abnormalities in schizophrenia: implications for innovative treatments. *Biomol. Ther.* **25**, 1–18 (2012).
29. Eastwood, S. L. et al. Decreased expression of mRNAs encoding non-NMDA glutamate receptors GluRI and GluRII in medial temporal lobe neurons in schizophrenia. *Mol. Brain Res.* **29**, 211–223 (1995).
30. Hawrylycz, M. J. et al. An anatomically comprehensive atlas of the adult human brain transcriptome. *Nature* **489**, 391–399 (2012).
31. Xue, M., Atallah, B. V. & Scanziani, M. Equalizing excitation–inhibition ratios across visual cortical neurons. *Nature* **511**, 596–600 (2014).
32. He, H., Shen, W., Zheng, L., Guo, X. & Cline, H. T. Excitatory synaptic dysfunction cell-autonomously decreases inhibitory inputs and disrupts structural and functional plasticity. *Nat. Commun.* **9**, 1–14 (2018).
33. Campanac, E. et al. Enhanced intrinsic excitability in basket cells maintains excitatory–inhibitory balance in hippocampal circuits. *Neuron* **77**, 712–722 (2013).
34. Crow, M. & Gillis, J. Co-expression in single-cell analysis: saving grace or original sin?. *Trends Genet.* **34**, 833–831 (2018).
35. Paul, A. et al. Transcriptional architecture of synaptic communication delineates GABAergic neuron identity. *Cell* **171**, 522–539 (2017).
36. Ben-Ari, Y., Gaiarsa, J.-L., Tsyag, R. & Khazipov, R. GABA: a pioneer transmitter that excites immature neurons and generates primitive oscillations. *Physiol. Rev.* **87**, 1215–1284 (2007).
37. Kang, H. J. et al. Spatio-temporal transcriptome of the human brain. *Nature* **478**, 483–489 (2011).
38. Kiroch, L. & Chechik, G. On expression patterns and developmental origin of human brain regions. *PLoS Comput. Biol.* **12**, e1005064 (2016).
39. Colantuoni, C. *et al.* Temporal dynamics and genetic control of transcription in the human prefrontal cortex. *Nature* **478**, 519–523 (2011).
40. Pletikos, M. *et al.* Temporal specification and bilaterality of human neocortical topographic gene expression. *Neuron* **81**, 321–332 (2014).
41. Gold, S. J., Ambros-Ingersen, J., Horowitz, J. R., Lynch, G. & Gall, C. M. Stoichiometries of AMPA receptor subunit mRNAs in rat brain fall into discrete categories. *J. Comp. Neurol.* **385**, 491–502 (1997).
42. Hodge, R. D. *et al.* Conserved cell types with divergent features in human versus mouse cortex. *Nature* **573**, 61–68 (2019).
43. Zapala, M. A. *et al.* Adult mouse brain gene expression patterns bear an embryologic imprint. *PNAS* **102**, 10337–10362 (2005).
44. Nieuwenhuys, R., Voogd, J. & van Huijzen, C. *The Human Central Nervous System: A Synopsis and Atlas* (Steinkopff, Heidelberg, 2008).
45. Montiel, J. F. & Aboitz, F. Pallial patterning and the origin of the isocortex. *Front. Neurosci.* **9**, 377 (2015).
46. Leto, K. *et al.* Consensus paper: cerebellar development. *Cerebellum* **15**, 789–828 (2016).
47. Sherman, S. M. Thalamus plays a central role in ongoing cortical functioning. *Nat. Neurosci.* **19**, 533–541 (2016).
48. Roy, M. *et al.* Proteomic analysis of postsynaptic proteins in regions of the human neocortex. *Nat. Neurosci.* **21**, 130–138 (2018).
49. Lee, E. & Hygon, S. Adult hippocampal neurogenesis and related neurotrophic factors. *BMB Rep.* **42**, 239–244 (2009).
50. Malenka, R. C. Synaptic plasticity and AMPA receptor trafficking. *Ann. N. Y. Acad. Sci.* **1003**, 1–11 (2003).
51. Lee, H.-K. & Kirkwood, A. AMPA receptor regulation during synaptic plasticity in hippocampus and neocortex. *Semin. Cell Dev. Biol.* **22**, 514–520 (2011).
52. Mitsushima, D., Ishihara, K., Sano, A., Kessels, H. W. & Takahashi, T. Contextual learning requires synaptic AMPA receptor delivery in the hippocampus. *PNAS* **108**, 12503–12508 (2011).
53. Terashima, A., Suh, Y. H. & Isaac, J. T. R. The AMPA receptor subunit GluA1 is required for CA1 hippocampal long-term potentiation but is not essential for synaptic transmission. *Neurochem. Res.* **44**, 549–561 (2019).
54. Hou, Q., Gilbert, J. & Man, H.-Y. Homeostatic regulation of AMPA receptor trafficking and degradation by light-controlled single synaptic activation. *Neuron* **72**, 806–818 (2011).
55. De Paola, V., Arber, S. & Caroni, P. AMPA receptors regulate dynamic equilibrium of presynaptic terminals in mature hippocampal networks. *Nat. Neurosci.* **6**, 491–500 (2003).
56. Ming, G. & Song, H. Adult neurogenesis in the mammalian brain: significant answers and significant questions. *Neuron* **70**, 687–702 (2011).
57. Kempermann, G., Song, H. & Gage, F. H. Neurogenesis in the adult hippocampus. *Cold Spring Harb. Perspect. Biol.* **7**, a018812 (2015).
58. Toda, T., Pary Jak, S. L., Linker, S. B. & Gage, F. H. The role of adult hippocampal neurogenesis in brain health and disease. *Mol. Psychiatry* **24**, 67–87 (2019).
59. Pandya, M. *et al.* Sex- and age-related changes in GABA signaling components in the human cortex. *Biol. Sex Diff.* **10**, 5 (2019).
60. Glykys, J. & Mody, I. The main source of ambient GABA responsible for tonic inhibition in the mouse hippocampus. *J. Physiol.* **582**, 1163–1178 (2007).
61. Chater, T. E. & Goda, Y. The role of AMPA receptors in postsynaptic mechanisms of synaptic plasticity. *Front. Cell. Neurosci.* **8**, 401 (2014).
62. Scaduto, P., Sequeira, A., Vawter, M. P., Bunney, W. & Limon, A. Preservation of global synaptic excitatory to inhibitory ratio during long postmortem intervals. *Sci. Rep.* **10**, 8626 (2020).
63. Chen, L., Yang, C. & Mower, G. D. Developmental changes in the expression of GABAA receptor subunits (a1, a2, a3) in the cat visual cortex and the effects of dark rearing. *Mol. Brain Res.* **88**, 135–143 (2001).
64. Fagioliini, M. *et al.* Specific GABAA circuits for visual cortical plasticity. *Science* **303**, 1681–1683 (2004).
65. Henley, J. M. & Wilkinson, K. A. Synaptic AMPA receptor composition in development, plasticity and disease. *Nat. Rev. Neurosci.* **17**, 337–350 (2016).
66. Suzuki, R. & Shimodaira, H. Pclust: an R package for assessing the uncertainty in hierarchical clustering. *Bioinformatics* **22**, 1540–1542 (2006).

**Acknowledgements**
This work was supported by NIH Research Grants R21MH113177 and R21AG053740 (AL) and the Borsa di Mobilita extra Erasmus fellowship (TZ). We thank Adolfo Sequeira and Marquis P. Vawter for fruitful discussions, and the reviewers for their comments that helped to improve this work.

**Author contributions**
A.L. conceived and designed the study. K.S. performed the Euclidian distances analyses. K.S., T.Z. and A.L. performed bioinformatic, statistical analysis of data and preparation of figures. All authors contributed to data interpretation, co-authored and have approved the final manuscript.

**Competing interests**
The authors declare no competing interests.

**Additional information**
**Supplementary information** is available for this paper at [https://doi.org/10.1038/s41598-020-68165-1](https://doi.org/10.1038/s41598-020-68165-1).

**Correspondence** and requests for materials should be addressed to A.L.

**Publisher’s note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.
