



## Note

# Lateralization of gene expression in human language cortex



Guy Karlebach<sup>a</sup> and Clyde Francks<sup>a,b,\*</sup>

<sup>a</sup> Language & Genetics Department, Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands

<sup>b</sup> Donders Institute for Brain, Cognition and Behaviour, Radboud University Nijmegen, Nijmegen, The Netherlands

## ARTICLE INFO

## Article history:

Received 19 November 2014

Reviewed 12 January 2015

Revised 9 February 2015

Accepted 4 March 2015

Action editor Stefano Cappa

Published online 18 March 2015

## Keywords:

Lateralization

Gene expression

Cerebral cortex

## ABSTRACT

Lateralization is an important aspect of the functional brain architecture for language and other cognitive faculties. The molecular genetic basis of human brain lateralization is unknown, and recent studies have suggested that gene expression in the cerebral cortex is bilaterally symmetrical. Here we have re-analyzed two transcriptomic datasets derived from post mortem human cerebral cortex, with a specific focus on superior temporal and auditory language cortex in adults. We applied an empirical Bayes approach to model differential left-right expression, together with gene ontology (GO) analysis and meta-analysis. There was robust and reproducible lateralization of individual genes and GO groups that are likely to fine-tune the electrophysiological and neurotransmission properties of cortical circuits, most notably synaptic transmission, nervous system development and glutamate receptor activity. Our findings anchor the cerebral biology of language to the molecular genetic level. Future research in model systems may determine how these molecular signatures of neurophysiological lateralization effect fine-tuning of cerebral cortical function, differently in the two hemispheres.

© 2015 Elsevier Ltd. All rights reserved.

## 1. Introduction

Lateralization is an important organizing principle of the vertebrate brain, and has been elaborated in our species (Rogers, Vallortigara, & Andrew, 2012). One prominently lateralized functional network underlies the human ability to speak and understand language, and critically involves the Superior Temporal Gyrus (STG) and the auditory cortex of Heschl's Gyrus (HG), both of which show anatomical asymmetry (Agcaoglu, Miller, Mayer, Hugdahl, & Calhoun, 2014; Dikker, Silbert, Hasson, & Zevin, 2014; Lyttelton et al., 2009;

Mazoyer et al., 2014; Menenti, Gierhan, Segaert, & Hagoort, 2011; Willems, Van der Haegen, Fisher, & Francks, 2014). These regions have also shown microanatomical (Hutsler, 2003) and neurophysiological (Morillon et al., 2010) lateralization; left-sided neural oscillatory frequencies reportedly correspond to syllabic speech rhythms (Morillon et al., 2010).

We hypothesized that functional lateralization would additionally be reflected at the molecular genetic level. However, two recent studies used microarrays to obtain gene mRNA expression profiles from numerous regions of the post mortem human brain, which failed to find hemispheric differences of expression, in any cortical region (Hawrylycz et al.,

\* Corresponding author. Max Planck Institute for Psycholinguistics, Wundtlaan 1, 6525 XD Nijmegen, The Netherlands.

E-mail address: [clyde.francks@mpi.nl](mailto:clyde.francks@mpi.nl) (C. Francks).

<http://dx.doi.org/10.1016/j.cortex.2015.03.003>

0010-9452/© 2015 Elsevier Ltd. All rights reserved.

2012; Pletikos et al., 2014), using paired sample t-tests to assess left-right differences for individual genes. The dataset of Pletikos et al. (Kang et al., 2011; Pletikos et al., 2014) included 57 brains across the human lifespan, from embryo to old age, with one tissue sample excised from each brain region that was defined. For analyzing lateralized expression they grouped the brains within each set of four consecutive ages. The dataset of Hawrylycz et al. (Hawrylycz et al., 2012) comprised brains from only two adult donors, but with multiple samples excised from each brain region.

We have re-analyzed and meta-analyzed the two datasets, specifically focusing on STG and HG in all available adults aged 18–55 years. Functional lateralization of the language network is robust across this age range (Agcaoglu et al., 2014), and we excluded data from outside this range in order to reduce variance associated with neurodevelopment or advanced age. In order to further maximise sensitivity for the analysis, we then applied several approaches that have not previously been used in examining lateralization of gene expression in these datasets. Firstly, we used an empirical Bayes method which uses data from all probes on the array for an improved estimation of gene expression variances (Smyth, 2004). This method can increase power to detect differential expression when the number of samples is small. Secondly, we reasoned that subtle lateralization at the level of individual genes may translate to more robust lateralization when assessed within functional gene groups, and we therefore carried out gene set enrichment analysis (GSEA) with respect to GO sets (Hahne, 2008). Thirdly, we performed meta-analysis over the two datasets.

## 2. Materials and methods

### 2.1. Subjects and neuroanatomical regions

The dataset of Pletikos et al. (Pletikos et al., 2014) was previously published by Kang et al. (Kang et al., 2011) who submitted it to the Gene Expression Omnibus (GEO accession GSE25219) (Barrett et al., 2013). The dataset of Hawrylycz et al. (Hawrylycz et al., 2012) is available via the Allen Human Brain Atlas website (<http://human.brain-map.org/>). Data for healthy adults in the range 18–55 years were downloaded from these two sites. Data in this age range were available from Pletikos et al. (Pletikos et al., 2014) for 13 healthy donors aged 18, 19, 21, 22, 23, 30, 36, 37, 37, 40, 40, 42 and 55 years. The donors aged 19, 21, 30, 37 and one of those aged 40 years were female, the rest male. The dataset of Hawrylycz et al. (Hawrylycz et al., 2012) was based on two healthy males aged 24 and 39, but with multiple tissue excisions per brain region (see below).

We aimed to select data from the two source datasets which represented comparable neuroanatomical regions across the studies. Exact matching between studies was not possible because somewhat different neuroanatomical criteria and sub-divisions had been used. For the dataset of Pletikos et al., histological verification of excisions with reference to cytoarchitectonic maps and Brodmann areas was used (Kang et al., 2011). For the dataset of Hawrylycz et al., histological staining was used to confirm structure identification, but it was not reported that reference was made to

cytoarchitectonic maps. We therefore relied on the neuroanatomical regional assignments within the two datasets to identify comparable regions across datasets, based on the information given. The consistency across the datasets of lateralized gene expression that we identified (see below) supports the validity of this approach, albeit that imperfect anatomic matching across studies probably reduced this consistency. In addition, the studies described different approaches to tissue excision from within each defined cortical region. For the dataset of Pletikos et al., a small amount of white matter was included in excisions from beneath cortical layer 6, corresponding to roughly the depth of layer 6 (Kang et al., 2011). For the dataset of Hawrylycz et al., the excisions extended perpendicularly from the pial surface to the white matter, but it was not stated that white matter was deliberately included. Again, the consistency of lateralized gene expression that we found across datasets (see below) indicates that this possible difference had a limited impact on comparability across datasets.

#### 2.1.1. STG

For the dataset of Pletikos et al. (Kang et al., 2011; Pletikos et al., 2014) we downloaded microarray data (22 arrays; 13 left, 9 right) for what was defined as ‘superior temporal cortex’. The samples were described as being excised from within the posterior third of the STG, and corresponding to Brodmann Area 22 (Kang et al., 2011).

For the dataset of Hawrylycz et al. (Hawrylycz et al., 2012) we downloaded microarray data (25 arrays, 14 left, 11 right) for tissue samples that were excised from within a region defined as ‘STG, inferior bank of gyrus’ (<http://human.brain-map.org/>), avoiding the anterior end of the gyrus. We excluded data from the ‘STG, lateral bank of gyrus’, because that region was represented by only 8 samples which did not include each donor and hemisphere.

#### 2.1.2. HG

For the dataset of Pletikos et al. (Kang et al., 2011; Pletikos et al., 2014) we downloaded microarray data (23 arrays; 13 left, 10 right) for tissue samples that were excised from within what was defined as ‘primary auditory cortex (A1C)’. This was described as corresponding to Brodmann area 41 (Kang et al., 2011), and hence constitutes a close match to HG.

For the dataset of Hawrylycz et al. (Hawrylycz et al., 2012) we downloaded microarray data (11 arrays, 5 left, 6 right) for samples excised from within what was defined as ‘HG’.

### 2.2. Microarray data pre-processing

For the dataset of Pletikos et al. (Pletikos et al., 2014), expression data for Affymetrix Human Exon 1.0 ST arrays had been preprocessed with Robust Multi-array Averaging (RMA) (Hahne, 2008), and were downloaded with the corresponding metadata. In order to map probes to genes, we downloaded the Human Exon Array annotation table from the Affymetrix website ([www.affymetrix.com](http://www.affymetrix.com)), and used it for mapping Affymetrix probes to HGNC symbols, the unique symbols assigned by the HUGO Gene Nomenclature Committee.

For the dataset of Hawrylycz et al. (Hawrylycz et al., 2012), raw data from Agilent 4x44 Whole Human Genome arrays for

donors H0351.2001 and H0351.2002 were downloaded (these were the only donors with data from both hemispheres). We preprocessed the data using ‘agilp’ in Bioconductor package version 2.0, 2012-06-10 (Chain et al., 2010), with the default configuration and pipeline of that package, which constructs a local linear regression model for a baseline (mean) of all the arrays and subtracts it from each array. The HGNC gene identity mappings and metadata were obtained from the Allen Human Brain Atlas website (<http://human.brain-map.org/>): we downloaded the annotation table and used it for mapping Agilent probes to HGNC symbols.

### 2.3. Generation of lateralization statistics for individual genes

Separately in each dataset and for each cortical region (STG and HG), a linear model with two predictors, for hemisphere and donor, was fitted to the expression levels of each gene using the limma Bioconductor package (Hahne, 2008), followed by limma’s empirical Bayes smoothing (Smyth, 2004) to improve gene variance estimates. The t-statistics and two-sided p-values were obtained with respect to the hemisphere predictor. The predictor for donor, as an individual-level predictor for each brain, controlled for brain-specific technical and biological effects. Further predictors (for example for sex in the dataset of Pletikos et al.) could not be added to the model due to the relatively small number of arrays (23 in Pletikos et al.). However we investigated the possible effect of sex within the results of multidimensional scaling (MDS) analysis (see below).

### 2.4. GSEA

GO annotations were obtained using the annotation file c5.all.v4.0.symbols.gmt from the C5 collection of the Broad Institute’s Molecular Signatures Database (Subramanian et al., 2005). We used all three types of GO sets, based on: Molecular Functions, Cellular Components, and Biological Processes. GO sets that contained 10 genes or less were removed. Separately in each dataset and for each brain structure (STG and HG), and according to the approach described in this reference (Hahne, 2008), the t-statistics of the genes in each GO class were summed and divided by the square root of the number of genes in that class, and the result was compared to the standard normal distribution in order to obtain a p-value.

In order to validate the GSEA approach, we simulated microarray data with an identical sample size to the Pletikos et al. dataset (Pletikos et al., 2014), and in which any lateralization of individual genes was random. The distribution of statistics over GO sets was the same as for the real data, but without significant GO sets after false discovery rate (FDR) adjustment, which further supports the lateralization we saw in the real data as being a true feature of the data. By then grouping together differentially expressed genes with significant lateralization t-statistics in the simulated data, to construct artificial GO terms, FDR significant terms were created. Thus the approach was powered to detect lateralization at the level of GO terms, in which subtle lateralization of individual genes within those terms would combine to reveal set-level lateralization. Note that the power in the

analysis of the real data was then further increased through meta-analysis over the two datasets.

### 2.5. Meta-analysis

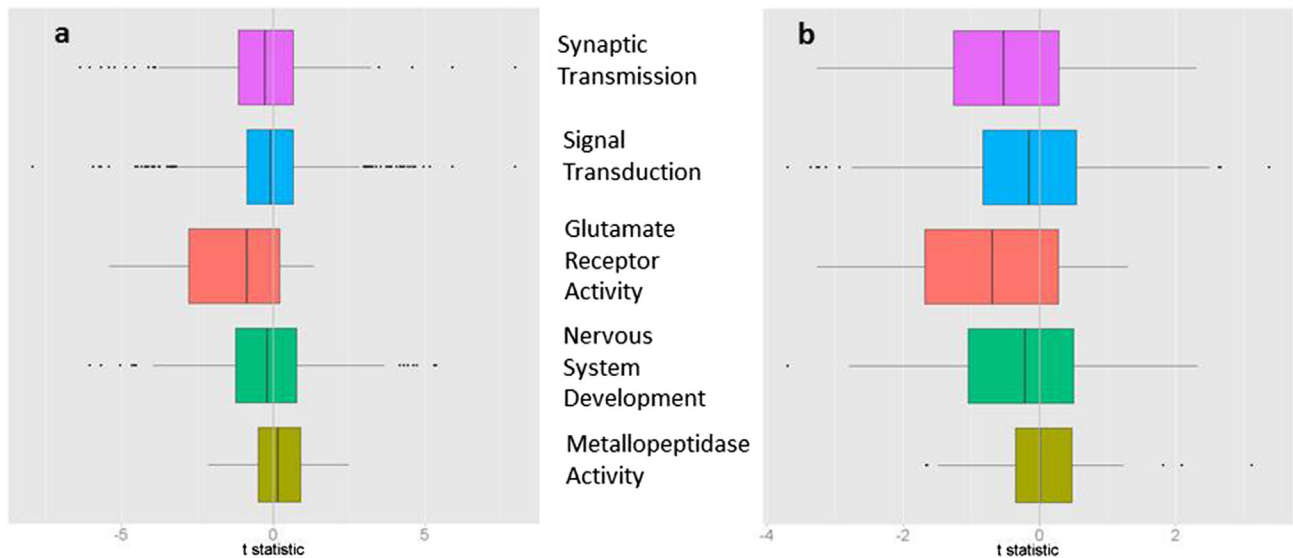
Fisher’s method (Fisher, 1932) was applied to integrate the nominal p-values obtained in the two datasets, for each individual gene, and for each GO term, separately for each region (STG or HG). The sum  $-2 \cdot (\ln(p_{\text{Pletikos}}) + \ln(p_{\text{Hawrylycz}}))$  was compared to the Chi-squared distribution with four degrees of freedom, where  $p_{\text{Pletikos}}$  and  $p_{\text{Hawrylycz}}$  were the p-values obtained within the datasets of Pletikos et al. (Pletikos et al., 2014) and Hawrylycz et al. (Hawrylycz et al., 2012), respectively. Benjamini-Hochberg (Benjamini & Hochberg, 1995) correction was then applied for obtaining the FDR separately for the two cortical regions, and separately for the individual gene analysis and the GO term analysis. Thus four separate FDRs were estimated.

### 2.6. MDS

For the most significant GO set in the STG meta-analysis (see below), we produced a vector of expression values per sample of the Pletikos et al. (Pletikos et al., 2014) dataset with a number of entries equal to the number of genes in the set. These vectors were then mapped into 2-dimensional space using Bioconductor’s limma (Hahne, 2008), resulting in a 2-dimensional vector for each sample. This would help visualize how the samples clustered together, in terms of their overall similarity as measured by gene expression levels within the most significant GO set. Again for the most significant set (synaptic transmission, see below), as an additional verification of lateralized expression, we applied the non-parametric Cramer test (Baringhaus & Franz, 2004) for which the null hypothesis was that samples from left and right were drawn from the same two-dimensional distribution. We also used this test to assess whether sex influenced the sample similarities of gene expression within the set ‘synaptic transmission’, within the dataset of Pletikos et al. (which contained data from both males and females).

## 3. Results

Seven individual genes showed STG lateralization with  $\text{FDR} < .01$  in meta-analysis, including Regulator of G-protein Signaling 8 (RGS8;  $\text{FDR} .00058$ ) and Glycine Receptor Subunit Alpha-2 (GLRA2;  $\text{FDR} .0015$ ) (Supplementary Table 1, Supplementary Figure 1). Striking STG lateralization was observed at the level of gene sets in meta-analysis, including two sets with corrected  $\text{FDR} < 1\text{E-}07$ ; synaptic transmission  $\text{FDR} 6.25\text{E-}09$ ; signal transduction  $\text{FDR} 3.02\text{E-}08$ , while other sets with corrected  $\text{FDR} < 1\text{E-}06$  were glutamate receptor activity, nervous system development, system development, transmission of nerve impulse, multicellular organismal development (respective FDRs  $1.33\text{E-}07$ ,  $1.33\text{E-}07$ ,  $1.33\text{E-}07$ ,  $3.97\text{E-}07$ ,  $5.39\text{E-}07$ ) (Fig. 1, Supplementary Figure 2; Supplementary Table 2). These signals were consistent across the two contributing datasets (Fig. 1, Supplementary Table 2): All of the top seven gene sets showed lateralization p



**Fig. 1** – Box plots of STG lateralization t-statistics, for genes within the four most significantly lateralized gene sets in meta-analysis, and an additional non-lateralized gene set (metallopeptidase activity) for comparison. Negative values correspond to higher expression in the left hemisphere, and vice versa. Frame a shows data from Pletikos et al. (Pletikos et al., 2014) and frame b from Hawrylycz et al. (Hawrylycz et al., 2012).

values  $< .01$  in both datasets. Other gene sets that were among the most significantly lateralized in both datasets, and significant in meta-analysis, included ‘cell surface receptor linked to signal transduction’, ‘G-protein coupled receptor protein signaling pathway’, and ‘calcium ion binding’ (Supplementary Table 2).

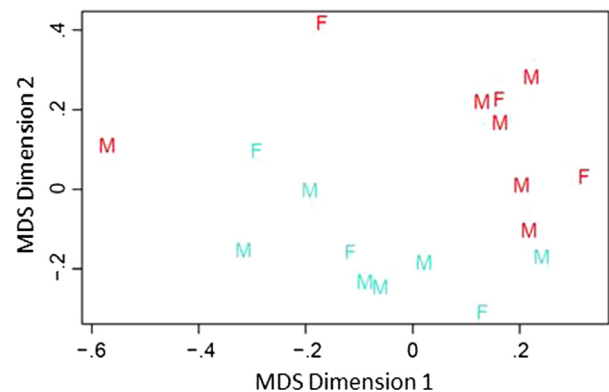
For HG, no individual genes showed significant lateralization. Gene sets showing lateralization in meta-analysis with corrected  $FDR < .01$  included again ‘G-protein coupled receptor protein signaling pathway’ ( $FDR .0023$ ), as well as various gene sets containing voltage gated channel genes: voltage gated channel activity ( $FDR .0028$ ), voltage gated cation channel activity ( $FDR .0028$ ), monovalent inorganic cation transport ( $FDR .0034$ ), voltage gated potassium channel activity ( $FDR .0046$ ) (Supplementary Table 3; Supplementary Figure 3). Gene sets coding more generally for membrane-bound proteins were also lateralized with  $FDR < .01$  (Supplementary Table 3).

Multi-dimensional scaling analysis of the Pletikos et al. STG data, for which distance between samples was defined by the similarity of genes’ expression within the set ‘synaptic transmission’ (i.e., the most significantly lateralized gene set in STG meta-analysis) showed that samples taken from the left and right cortices formed different distributions with respect to the genes in this set (Fig. 2). An additional confirmatory analysis using the Cramer test showed that the null hypothesis of left and right samples belonging to the same two-dimensional distribution was rejected ( $p = .0014$ ). We also applied the Cramer test to assess whether the sexes differed in their two-dimensional distribution, but no evidence for a sex difference was detected ( $p = .82$ , see Fig. 2). Lateralization statistics for individual genes, for the set ‘synaptic transmission’, are shown in Supplementary Table 4. Among the most strongly left-lateralized mRNAs were those coding for the dopamine receptor DRD2, the serotonin receptor subunits

HTR3B and HTR2C, and the glutamate receptor subunits GRM1, GRIK1, GRIA1 and GRIA2 (Supplementary Table 4).

#### 4. Discussion

Pletikos et al. and Hawrylycz et al., whose data we re-analyzed and meta-analyzed for the present study, concluded that gene expression in the cerebral cortex is bilaterally symmetrical and that the source hemisphere of samples made no quantitative difference that needed to be considered in data analysis. In contrast, we have detected robust and reproducible lateralization of gene expression in the STG and HG in these datasets. The contrasting results



**Fig. 2** – Multidimensional scaling (MDS) of Pletikos et al. (Pletikos et al., 2014) STG samples from the left (turquoise) and right (red) hemispheres, based on genes in the set ‘synaptic transmission’. Samples from male donors are labeled ‘M’, and those from female donors ‘F’.



have likely arisen from several aspects of our design and analysis. Firstly we specifically targeted the STG and HG as candidate regions of interest, out of all cortical and other brain regions that were sampled in these datasets. Our choice was motivated by previous literature on superior temporal lateralization in terms of function, structure, neurophysiology and microanatomy (see Introduction). There was therefore strong evidence *a priori* to expect genetic lateralization in these particular cortical regions. We then applied several approaches that were designed to maximise the power to detect subtle lateralization of gene expression, which included pooling data from all adults within the age range 18–55 years into a single analysis of the Pletikos et al. dataset, Bayesian modeling of gene expression variance, GO set analysis, and meta-analysis over the two datasets. Lateralization of gene expression was most significant for GO sets rather than at the level of individual genes, although the meta-analysis of individual genes also identified significant lateralization in the STG.

Lateralization of individual genes and GO sets was strikingly concordant across the two datasets, despite differences in the neuroanatomical regional definitions and cortical sampling procedures between the two studies. Lateralization of gene expression was therefore robust to these major sources of experimental variation. An issue with the source datasets may be imperfect anatomic matching for tissue excision between the homologous regions of the left and right hemispheres. The genetic lateralization that we detected can only be interpreted at the level of anatomical resolution defined in the source datasets. Future studies of transcriptomic lateralization may benefit from higher-resolution anatomical definitions, including defining left-right homologous excisions by cytoarchitecture as was done for the dataset of Pletikos et al., since relatively broad regions such as primary auditory cortex are known to contain various microstructurally distinct sub-regions (Fullerton & Pandya, 2007). However, attempting to define perfectly matching left-right excisions at the micro-anatomical level may prove impossible, given the well known lateralizations of function and neurophysiology for STG and HG.

Importantly, our analysis did not depend on precise left-right matching within any given subject. Rather, the lateralization that we detected was an average property of all left excisions versus all right excisions, while we also added a predictor variable for 'subject' to improve the accuracy of the model as regards possible subject-specific biological and technical variation. Our findings therefore depend on appropriate anatomic matching of left-right homologous regions over all excisions, on average, and not at the individual subject level. Although detailed information on the precise location of excisions was not available for both of the source studies, both had performed analysis of gene expression lateralization with their data, which indicates that the authors regarded the left and right excisions as being sufficiently comparable to support this approach. The concordance of lateralization over both studies, for specific gene sets, indicates that our findings are relatively robust to the precise locations of excisions.

That lateralization of gene expression was not detected by Pletikos et al. and Hawlyrycz et al., whereas they did

detect substantial differences between cortical regions ipsilaterally, indicates the relatively subtle nature of lateralized gene expression in the cerebral cortex. This is also consistent with data from human embryonic and fetal material, where little evidence for lateralization has been detected in four post mortem studies, with the possible exception of a small number of genes in the first trimester of gestation (Johnson et al., 2009; Lambert et al., 2011; Pletikos et al., 2014; Sun, Collura, Ruvolo, & Walsh, 2006). However, those embryonic data may also benefit from the kinds of analyses that we have applied here, and the results could then be compared to the adult genetic lateralization that we have identified.

It is striking that many of the lateralized gene sets that we detected, including those that were most strongly lateralized, have clear neuronal functions likely to affect signaling, learning and information-processing properties of circuitry differently on the two sides. Lateralized gene sets included those involved in synaptic transmission (STG), glutamate receptor activity (STG), voltage gated cation channels (HG), and the G-protein coupled receptor protein signaling pathway (STG and HG). These findings occurred despite our having tested all GO sets without reference to their neuronal relevance, while performing correction for multiple testing (~1,000 sets, after excluding sets with fewer than ten member genes). Our findings of gene mRNA lateralization may complement studies of receptor mapping within the cerebral cortex, at the protein level (Zilles, Bacha-Trams, Palomero-Gallagher, Amunts, & Friederici, 2015). Interestingly, lateralization within gene sets that are defined for their developmental roles was also detected, including 'nervous system development' (STG). This indicates that transcriptional factors and other developmentally important proteins continue to have important roles in maintaining lateralized function in the adult brain.

Given the diversity of gene sets that showed evidence for lateralization, it is not obvious how to interpret our findings in terms of specific implications for lateralized neurophysiology, or regarding the differences of lateralization found between STG and HG. These topics must be subjects for future research. Our data indicate a robust, lateralized fine-tuning of gene expression at the genomic-level, which may have wide-ranging but unpredictable implications for diverse neurophysiological mechanisms and properties. It is likely that the combinatorial effects of small quantitative differences over many genes determine neurophysiological outcomes that underlie lateralized function in the superior temporal and auditory cortex. The future modeling of these effects in cell and animal models may require the simultaneous manipulation of many genes, which will probably depend on improvements in technology for carrying out these studies. Lateralization of gene expression will also need to be studied at the level of individual neuronal and glial subtypes, and with respect to both grey and white matter. Although we understand that the focus of both source studies was to collect cortical tissue, at least one of the studies also included a small proportion of white matter in excisions (see section 2.1 above). However, there is no reason to suspect that either study systematically included more white matter in excisions from one hemisphere than the other, which could have affected the

transcriptional lateralization that we detected. Again, because our findings stem from averages over all samples, they should be relatively robust to random sample heterogeneities of composition.

A common limitation of post mortem studies is small sample sizes, and this applies also to the source datasets that we have analyzed here. We took several steps to mitigate this, including the pooled analysis of all available adults aged 18–55 years in the dataset of Pletikos et al., the use of Bayesian smoothing, and meta-analysis over two datasets. Reassuringly, our QQ plots (Supplementary Figures 2 and 3) indicated no general deviation of  $p$  values from the null expectations throughout the bulk of the distribution of results, with only a deviation at the tails, as would be expected if a minority of gene sets are significantly lateralized. The lateralized gene sets include both relatively small sets (e.g., glutamate receptor activity, 20 genes) and relatively large sets (e.g., signal transduction, 1572 genes).

An important question arising from our study is whether genes showing lateralized expression in STG and HG also contain polymorphisms that are associated with quantitative variation in structural and functional measures of lateralization within these regions. This question may be approached using gene-set enrichment analysis and polygenic risk scores, in brain imaging genetics datasets produced for genome-wide association scanning (Cai et al., 2014; Guadalupe et al., 2015; Hibar et al., 2015; Thompson et al., 2014). Genetic association studies of individual candidate genes, that were not so far motivated by data on lateralized gene expression, have indicated that measures of language-related functional lateralization may associate with common genetic variants (Ocklenburg et al., 2013; Pinel et al., 2012).

In summary, we have detected transcriptional differences between the cerebral hemispheres in the superior temporal cortex and auditory cortex of human adults, particularly for genes that are crucially involved in neuronal electrophysiology and neurotransmission, and which were reproducible across two datasets. These findings anchor the cerebral biology of language to the molecular level. Future research in model systems may determine how these molecular signatures of neurophysiological lateralization effect fine-tuning of cerebral cortical function, differently in the two hemispheres. Polymorphisms within these gene sets may also combine to influence language-related cognitive variability and disorders.

## Acknowledgments

G.K. and C.F. are funded by the Max Planck Society and a grant to C.F. from the Netherlands Organization for Scientific Research (NWO, Open Programme grant 824.14.005).

## Supplementary data

Supplementary data related to this article can be found at <http://dx.doi.org/10.1016/j.cortex.2015.03.003>.

## REFERENCES

- Agcaoglu, O., Miller, R., Mayer, A. R., Hugdahl, K., & Calhoun, V. D. (2014). Lateralization of resting state networks and relationship to age and gender. *NeuroImage*.
- Allen Human Brain Atlas. <http://human.brain-map.org/>.
- Baringhaus, L., & Franz, C. (2004). On a new multivariate two-sample test. *Journal of Multivariate Analysis*, 88(1), 190–206.
- Barrett, T., Wilhite, S. E., Ledoux, P., Evangelista, C., Kim, I. F., Tomashevsky, M., et al. (2013). NCBI GEO: archive for functional genomics data sets—update. *Nucleic Acids Research*, 41(Database issue), D991–D995.
- Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate – a practical and powerful approach to multiple testing. *Journal of Royal Statistical Society Series B Methodology*, 57(1), 289–300.
- Cai, D. C., Fonteijn, H., Guadalupe, T., Zwiers, M., Wittfeld, K., Teumer, A., et al. (2014). A genome-wide search for quantitative trait loci affecting the cortical surface area and thickness of Heschl's gyrus. *Genes, Brain, and Behavior*, 13(7), 675–685.
- Chain, B., Bowen, H., Hammond, J., Posch, W., Rasaiyaah, J., Tsang, J., et al. (2010). Error, reproducibility and sensitivity: a pipeline for data processing of Agilent oligonucleotide expression arrays. *BMC Bioinformatics*, 11, 344.
- Dikker, S., Silbert, L. J., Hasson, U., & Zevin, J. D. (2014). On the same wavelength: predictable language enhances speaker-listener brain-to-brain synchrony in posterior superior temporal gyrus. *The Journal of Neuroscience*, 34(18), 6267–6272.
- Fisher, R. A. (1932). *Statistical methods for research workers* (4th ed.). Edinburgh etc: Oliver and Boyd.
- Fullerton, B. C., & Pandya, D. N. (2007). Architectonic analysis of the auditory-related areas of the superior temporal region in human brain. *The Journal of Comparative Neurology*, 504(5), 470–498.
- Guadalupe, T., Zwiers, M. P., Wittfeld, K., Teumer, A., Vasquez, A. A., Hoogman, M., et al. (2015). Asymmetry within and around the human planum temporale is sexually dimorphic and influenced by genes involved in steroid hormone receptor activity. *Cortex*, 62(0), 41–55.
- Hahne, F. (2008). *Bioconductor case studies*. New York, NY: Springer.
- Hawrylycz, M. J., Lein, E. S., Guillozet-Bongaarts, A. L., Shen, E. H., Ng, L., Miller, J. A., et al. (2012). An anatomically comprehensive atlas of the adult human brain transcriptome. *Nature*, 489(7416), 391–399.
- Hibar, D. P., Stein, J. L., Renteria, M. E., Arias-Vasquez, A., Desrivieres, S., Jahanshad, N., et al. (2015). Common genetic variants influence human subcortical brain structures. *Nature*.
- Hutsler, J. J. (2003). The specialized structure of human language cortex: pyramidal cell size asymmetries within auditory and language-associated regions of the temporal lobes. *Brain and Language*, 86(2), 226–242.
- Johnson, M. B., Kawasawa, Y. I., Mason, C. E., Krsnik, Z., Coppola, G., Bogdanovic, D., et al. (2009). Functional and evolutionary insights into human brain development through global transcriptome analysis. *Neuron*, 62(4), 494–509.
- Kang, H. J., Kawasawa, Y. I., Cheng, F., Zhu, Y., Xu, X., Li, M., et al. (2011). Spatio-temporal transcriptome of the human brain. *Nature*, 478(7370), 483–489.
- Lambert, N., Lambot, M. A., Bilheu, A., Albert, V., Englert, Y., Libert, F., et al. (2011). Genes expressed in specific areas of the human fetal cerebral cortex display distinct patterns of evolution. *PloS One*, 6(3), e17753.
- Lytelton, O. C., Karama, S., Ad-Dab'bagh, Y., Zatorre, R. J., Carbonell, F., Worsley, K., et al. (2009). Positional and surface area asymmetry of the human cerebral cortex. *NeuroImage*, 46(4), 895–903.
- Mazoyer, B., Zago, L., Jobard, G., Crivello, F., Joliot, M., Percey, G., et al. (2014). Gaussian mixture modeling of hemispheric

- lateralization for language in a large sample of healthy individuals balanced for handedness. *PLoS One*, 9(6), e101165.
- Menenti, L., Gierhan, S. M., Segaert, K., & Hagoort, P. (2011). Shared language: overlap and segregation of the neuronal infrastructure for speaking and listening revealed by functional MRI. *Psychological Science*, 22(9), 1173–1182.
- Morillon, B., Lehongre, K., Frackowiak, R. S., Ducorps, A., Kleinschmidt, A., Poeppel, D., et al. (2010). Neurophysiological origin of human brain asymmetry for speech and language. *Proceedings of the National Academy of Sciences USA*, 107(43), 18688–18693.
- Ocklenburg, S., Arning, L., Gerding, W. M., Epplen, J. T., Gunturkun, O., & Beste, C. (2013). FOXP2 variation modulates functional hemispheric asymmetries for speech perception. *Brain and Language*, 126(3), 279–284.
- Pinel, P., Fauchereau, F., Moreno, A., Barbot, A., Lathrop, M., Zelenika, D., et al. (2012). Genetic variants of FOXP2 and KIAA0319/TTRAP/THEM2 locus are associated with altered brain activation in distinct language-related regions. *The Journal of Neuroscience*, 32(3), 817–825.
- Pletikos, M., Sousa, A. M., Sedmak, G., Meyer, K. A., Zhu, Y., Cheng, F., et al. (2014). Temporal specification and bilaterality of human neocortical topographic gene expression. *Neuron*, 81(2), 321–332.
- Rogers, L. J., Vallortigara, G., & Andrew, R. J. (2012). *Divided brains: The biology and behaviour of brain asymmetries*. Cambridge, England; New York: Cambridge University Press.
- Smyth, G. K. (2004). Linear models and empirical Bayes methods for assessing differential expression in microarray experiments. *Statistical Applications in Genetics and Molecular Biology*, 3. Article3.
- Subramanian, A., Tamayo, P., Mootha, V. K., Mukherjee, S., Ebert, B. L., Gillette, M. A., et al. (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proceedings of the National Academy of Sciences USA*, 102(43), 15545–15550.
- Sun, T., Collura, R. V., Ruvolo, M., & Walsh, C. A. (2006). Genomic and evolutionary analyses of asymmetrically expressed genes in human fetal left and right cerebral cortex. *Cerebral Cortex*, 16(Suppl. 1), i18–25.
- Thompson, P. M., Stein, J. L., Medland, S. E., Hibar, D. P., Vasquez, A. A., Renteria, M. E., et al. (2014). The ENIGMA Consortium: large-scale collaborative analyses of neuroimaging and genetic data. *Brain Imaging and Behavior*, 8(2), 153–182.
- Willems, R. M., Van der Haegen, L., Fisher, S. E., & Francks, C. (2014). On the other hand: including left-handers in cognitive neuroscience and neurogenetics. *Nature Reviews. Neuroscience*, 15(3), 193–201. [www.affymetrix.com](http://www.affymetrix.com). Affymetrix Web Site.
- Zilles, K., Bacha-Trams, M., Palomero-Gallagher, N., Amunts, K., & Friederici, A. D. (2015). Common molecular basis of the sentence comprehension network revealed by neurotransmitter receptor fingerprints. *Cortex*, 63, 79–89.