



# Commentary: BRAIN NETWORKS. Correlated Gene Expression Supports Synchronous Activity in Brain Networks. *Science* 348, 1241–4

Spiro P. Pantazatos<sup>1,2\*</sup> and Xinyi Li<sup>3</sup>

<sup>1</sup> Department of Psychiatry, Columbia University, New York, NY, United States, <sup>2</sup> Molecular Imaging and Neuropathology Division, New York State Psychiatric Institute, New York, NY, United States, <sup>3</sup> Biomedical Informatics, Columbia University, New York, NY, United States

## OPEN ACCESS

### Edited by:

Ruth Luthi-Carter,  
University of Leicester,  
United Kingdom

### Reviewed by:

Detlef H. Heck,  
Department of Anatomy and  
Neurobiology, University of Tennessee  
Health Science Center, United States  
Robert W. Williams,  
University of Tennessee Health  
Science Center, United States

### \*Correspondence:

Spiro P. Pantazatos  
spiropan@gmail.com

### Specialty section:

This article was submitted to  
Neurogenomics,  
a section of the journal  
*Frontiers in Neuroscience*

**Received:** 21 June 2017

**Accepted:** 30 June 2017

**Published:** 18 July 2017

### Citation:

Pantazatos SP and Li X (2017)  
Commentary: BRAIN NETWORKS.  
Correlated Gene Expression Supports  
Synchronous Activity in Brain  
Networks. *Science* 348, 1241–4.  
*Front. Neurosci.* 11:412.  
doi: 10.3389/fnins.2017.00412

A recent report claims that functional brain networks defined with resting-state functional magnetic resonance imaging (fMRI) can be recapitulated with correlated gene expression (i.e., high within-network tissue-tissue “strength fraction,” SF) (Richiardi et al., 2015). However, the authors do not adequately control for spatial proximity. We replicated their main analysis, performed a more effective adjustment for spatial proximity, and tested whether “null networks” (i.e., clusters with center coordinates randomly placed throughout cortex) also exhibit high SF. Removing proximal tissue-tissue correlations by Euclidean distance, as opposed to removing correlations within arbitrary tissue labels as in Richiardi et al. (2015), reduces within-network SF to no greater than null. Moreover, randomly placed clusters also have significantly high SF, indicating that high within-network SF is entirely attributable to proximity and is unrelated to functional brain networks defined by resting-state fMRI. We discuss why additional validations in the original article are invalid and/or misleading and suggest future directions.

**Keywords:** functional brain networks, gene expression, allen brain atlas, resting-state fMRI, spatial proximity, confounding factors

A recent study explores relationships between gene expression and distributed spatial patterns of synchronous brain activity consistently observed in resting state (RS) fMRI (Richiardi et al., 2015) using microarray data from the Allen Brain Atlas (<http://human.brain-map.org>; Hawrylycz et al., 2012). The authors correctly state that “While functional networks are distributed spatially, meaning they cross over different tissue types, and that their sample can be spatially distant, it is important to ensure that a high strength fraction (SF) does not simply reflect the fact that tissues are the same.” They attempt to correct for spatial proximity by omitting edges between regions falling in the same “tissue class,” which are ontological labels provided by Allen Brain Atlas (Supplementary Table 4 in Richiardi et al., 2015). However, this approach inadequately controls for spatial proximity: nearby regions will fail to have their edges removed by a label boundary dividing them, while longer edges within a tissue label will be removed instead (**Figure 1A**). The issues remains even when correction uses coarser tissue classes.

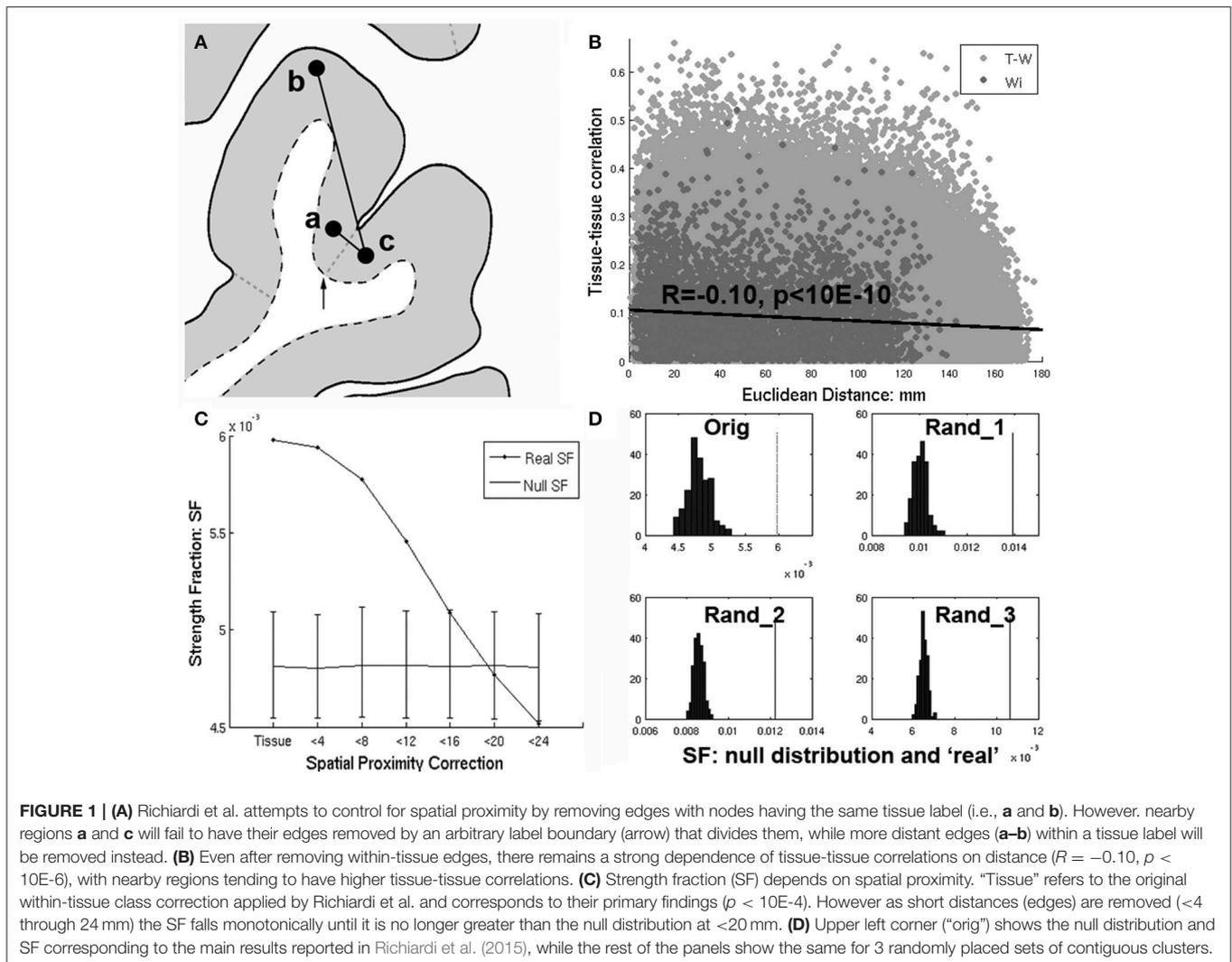
Even after removing within-tissue edges, there remains an association between tissue-tissue correlations and distance ( $R = -0.10$ ,  $p < 10E-6$ ), with nearby regions tending to have higher correlations (**Figure 1B**). Within network (Wi) edges are significantly shorter than out-of-network (T-W) edges (Wi distances vs. T-W distances 2-sample  $t$ -test:  $t_{(759,091)} = -51.1$ , Wi  $\mu = 52.9$  mm,

T-W  $\mu = 78.3$  mm). This biases the Wi SF to be greater relative to a null distribution which calculates Wi using longer connections (i.e., T-W edges which are labeled Wi as part of the shuffling).

When a more direct correction for distance is applied (removing proximal edges), within-network SF is no longer greater than null. **Figure 1C** shows dependence of SF on proximity. “Tissue” refers to the within-tissue class correction applied by Richiardi et al. and demonstrates their primary findings ( $p < 10E-4$ ). However, as short-range edges are removed (4–24 mm), SF falls monotonically until it is no greater than null at  $<20$  mm. In addition, applying linear regression to adjust for distance (French and Pavlidis, 2011) results in a large *negative* SF (SF =  $-0.61$ ,  $p = 1$ , data not shown, see Section *Supplementary Discussion* below for pitfalls of this approach and omitting negative correlations prior to SF calculation). Thus, the claim of the original article: “Given that we used only cortical samples, that we removed edges linking tissues of the same class, and that functional networks are spatially distributed, this finding cannot emerge from spatial proximity or gross tissue similarity” is false.

Moreover, the null distribution derived in Richiardi et al. is flawed because the permutation strategy assumes all regions are independent and equally exchangeable, which is not true given the spatial autocorrelation and distance bias.

Although, not reported in the original article, the authors claim that SF remains significant after a linear regression-based distance correction is applied and only positive connections are included (personal communication). However, there are two problems with this: (1) The assumption that tissue-tissue correlation strength varies linearly with distance is too strong. A plot of the tissue-tissue correlations vs. distance shows that the best-fit curve is steep for short edges and less steep at around 20 mm: after adjusting for the best-fit line there will still be a distance bias. Model-based correction will not be as optimal as simply removing proximal connections. (2) Applying a cutoff of zero for connections contributing to the SF is not well justified (this applies to the main analyses as well). What, biologically, distinguishes a correlation of 0.1 vs.  $-0.1$  other than i.e., noise in the expression vector? Furthermore, after regression, about half of the connections (that were included in the original



main analyses) will be negative due to mean centering and omitted in the new analysis, making the cutoff of zero even more arbitrary.

Short (<16 mm) connections account entirely for the significant SF reported in Richiardi et al. (2015) (**Figure 1C**). Given that the main claim of Richiardi et al. is that correlated gene expression relates specifically to RS functional networks, a crucial question is “is high local SF specific to the RS networks”? If so, then the SF of distributed clusters with centers randomly placed throughout cortex (with size and total number of Wi nodes similar to RS networks) should be non-significant. However, for 1,000 randomly selected cluster sets,  $p$ -values were *all 0* ( $<0.001$ ). Panel D (“Orig,” upper left corner) shows the null distribution and real SF corresponding to the main results reported in the original article, while the rest of the panels show the same for three randomly selected sets of clusters. Thus the significant SF reported in the original article is entirely attributable to spatial proximity *and* is unrelated to RS fMRI networks. Note that SF of Wi RS networks cannot be compared to SF of randomly selected clusters since SF is a function of total number of Wi connections.

Matlab code replicating the primary results presented in Richiardi et al. (2015) and results presented here are available at [https://github.com/spiropan/ABA\\_functional\\_networks](https://github.com/spiropan/ABA_functional_networks). See Section *Supplementary Discussion* below for why the additional validation analyses in Richiardi et al. (2015) (Figures 2, 3) are invalid and/or misleading, and the relationship of their results with differentially stable genes identified in Hawrylycz et al. (2015).

## SUPPLEMENTARY DISCUSSION

An alternative/complementary approach to distance correction is to regress out the effects of distance on tissue-tissue correlations prior to computing SF. As mentioned in the main text, linear regression-based distance correction leads to a large negative and non-significant SF. However, SF remains significant after this correction is applied and only positive connections are included when calculating SF. However, there are two problems with and explanations for this: (1) The assumption that tissue-tissue correlation strength varies linearly with distance is too strong. A plot of the tissue-tissue correlations vs. distance shows that the best-fit curve is steep for short edges and less steep at around 20 mm: after adjusting for the best-fit line there will still be a distance bias. No matter what model is adjusted for, the correction will not be as optimal as simply removing proximal connections. (2) Applying an arbitrary cutoff of zero for connections contributing to the SF is not well justified (this applies to the main analyses as well). What, biologically, distinguishes a correlation of 0.01 vs.  $-0.01$  other than i.e., noise in the expression vector that nudges the correlation across zero? Furthermore, after regression, about half of the connections (that were included in the original main analyses) will be negative due to mean centering and omitted

in the new analysis, making the cutoff of zero even more arbitrary.

It is likely that the optimization approach in Richiardi et al. used to derive the 136 consensus genes (i.e., multiplying each gene’s expression by 10 and recalculating the strength fraction) identified genes with both high local spatial autocorrelation and variability across the cortex. This is consistent with the observation that >75% of these 136 consensus genes are in the top 10% of genes found to have consistently high region-to-region variability (so called differentially stable, DS genes) across the cortex identified in Hawrylycz et al. (2015). Furthermore, GO functions related to potassium channels (featured prominently in Richiardi et al. Supplemental Table S3) were most over-represented among high-DS genes ( $P < 1.70 \times 10^{-12}$ ) in Hawrylycz et al. (2015). Given that genes high in DS (i.e., consistent region-to-region variability), *irrespective of belonging to resting state functional networks*, are more likely to be involved in brain functioning (Hawrylycz et al., 2015), this could account for the enrichment ( $p = 0.006$ ) of SNPs associated with functional network SF observed in the IMAGEN portion of the Richiardi et al. analyses.

Figure 2 in Richiardi et al. is misleading, and does not constitute evidence for “definite differences in functional connectivity strength mostly within the functional networks themselves.” Given that the authors used a *post-hoc*, biased approach to generate the loosely thresholded functional connectivity difference matrices and maps, it is unclear whether comparable results could be generated when applying their scoring procedure to 136 genes randomly selected from the background set or from the top 10% of genes showing variability across the cortex (i.e., cortical DS genes reported in Hawrylycz et al. (2015)). Finally, the results from mouse tractography data ( $p = 0.011$  Mantel correlation, Figure 3 in Richiardi et al.) does not make any adjustment for spatial proximity, and is likely also confounded by spatial proximity.

The Richiardi et al. study is an important step toward identifying genes whose spatial pattern of cortical expression relate to distributed functional networks consistently observed in resting state fMRI. However, we are not quite there yet. Further work will be required to adequately control for the confounding effects of spatial proximity. While here distances were computed in 3D MNI space, computing distances in flattened cortical surface (2D) space would make distance measurements more accurate. While distance correction using 3D Euclidean distance is suboptimal compared to 2D Euclidean, it is optimal compared to using region labels. 2D Euclidean distance would be more accurate than 3D distance, and while the slope and shape of the curve in **Figure 1C** might change, SF would still fall monotonically as short-range edges are removed and be no greater than the null distribution at around 24–32 mm.

Future studies to relate gene expression with resting-state functional networks will require valid and more appropriate null distributions, and could benefit from “non-parametric” approaches to correct for distance (i.e., calculating outcome

measures such as within-network SF across distance bins to directly visualize distance effects etc.).

## AUTHOR CONTRIBUTIONS

Conceived of the project (SP), analyzed data in Matlab (SP), analyzed data in R (XL) and wrote the manuscript (SP, XL).

## REFERENCES

- French, L., and Pavlidis, P. (2011). Relationships between gene expression and brain wiring in the adult rodent brain. *PLoS Comput. Biol.* 7:e1001049. doi: 10.1371/journal.pcbi.1001049
- Hawrylycz, M., Miller, J. A., Menon, V., Feng, D., Dolbeare, T., Guillozet-Bongaarts, A. L., et al. (2015). Canonical genetic signatures of the adult human brain. *Nat. Neurosci.* 18, 1832–1844. doi: 10.1038/nn.4171
- 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, 391–399. doi: 10.1038/nature11405
- Richiardi, J., Altmann, A., Milazzo, A. C., Chang, C., Chakravarty, M. M., Banaschewski, T., et al. (2015). BRAIN NETWORKS. Correlated gene

## ACKNOWLEDGMENTS

This work was supported in part by a Janssen Translational Neuroscience Postdoctoral Research Fellowship and a K01MH108721 (SP). We thank Paul Pavlidis and Leon French for helpful comments and suggestions.

expression supports synchronous activity in brain networks. *Science* 348, 1241–1244. doi: 10.1126/science.1255905

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2017 Pantazatos and Li. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.