Neuroscience Letters 662 (2018) 17–21

Contents lists available at ScienceDirect

Neuroscience Letters

journal homepage: www.elsevier.com/locate/neulet

Research article

Comparative connectomics: Mapping the inter-individual variability of connections within the regions of the human brain

Csaba Kerepesi[a,c], Balázs Szalkai[a], Bálint Varga[a], Vince Grolmusz[a,b,*]

[a] PIT Bioinformatics Group, Eötvös University, H-1117 Budapest, Hungary
[b] Uratim Ltd., H-1118 Budapest, Hungary
[c] Institute for Computer Science and Control, Hungarian Academy of Sciences, Hungary

ARTICLE INFO

Keyword:
Connectome

ABSTRACT

The human braingraph, or connectome is a description of the connections of the brain: the nodes of the graph correspond to small areas of the gray matter, and two nodes are connected by an edge if a diffusion MRI-based workflow finds fibers between those brain areas. We have constructed 1015-vertex graphs from the diffusion MRI brain images of 392 human subjects and compared the individual graphs with respect to several different areas of the brain. The inter-individual variability of the graphs within different brain regions was discovered and described. We have found that the frontal and the limbic lobes are more conservative, while the edges in the temporal and occipital lobes are more diverse. Interestingly, a "hybrid" conservative and diverse distribution was found in the paracentral lobule and the fusiform gyrus. Smaller cortical areas were also evaluated: precentral gyri were found to be more conservative, and the postcentral and the superior temporal gyri to be very diverse. Similar studies concerning the human genome discovered more and less conservative sections of the DNA, opening up entirely new fields in genomics. We think that the present study is the first step in this direction in human connectomics. The clinical significance of the conservativity of a given cerebral area could be the higher sensitivity for traumas and developmental or neuro-degenerative events than the less conservative areas.

1. Introduction

Large co-operative research projects, such as the Human Connectome Project [1], produce high-quality MRI-imaging data of hundreds of healthy individuals. The comparison of the connections of the brains of the subjects is a challenging problem that may open numerous research directions. In the present work, we map the variability of the connections within different brain areas in 392 human subjects, in order to discover brain areas with higher variability in their connections or other brain regions with more conservative connections.

The braingraphs or connectomes are the well-structured discretizations of the diffusion MRI imaging data that yield new possibilities for the comparison of the connections between distinct brain areas in different subjects [2,3] or for finding common connections in distinct cerebra [4,5], forming a common, consensus human braingraph.

Here, by using the data of the Human Connectome Project [1], we describe, by their distribution functions, the inter-individual diversity of the braingraph connections in separate brain areas in 392 healthy subjects of ages between 22 and 35 years.

Since every brain is unique, the workflow that produces the braingraphs consists of several steps, including a diffeomorphism [6] of the brain atlas to the brain-image processed. After the diffeomorphism, corresponding areas of different human brains are pairwise identified through the atlas and, consequently, can be compared with one another. The braingraphs, with nodes in the corresponded brain areas, are prepared from the diffusion MRI images of the individual cerebra through a workflow detailed in the "Methods" section. Every braingraph studied contains 1015 nodes (or vertices). The vertices correspond to the subdivision of anatomical gray matter areas in cortical and subcortical regions. For the list of the regions and the number of nodes in each region, we refer to Table S1 and Figure S1 in the Supporting material.

Next, we describe the variability, or the distribution of the graph edges in each brain region, and also in each lobe. Our goal is to discover more and less conservatively "wired" areas of the brain. For example, suppose that a lobe X of the brain contains the very same connections in each of the 392 human subjects: then this lobe X is entirely conservative. If a lobe X contained graph edges that are not repeated between any two graphs, computed from the data of these 392 subjects, then this lobe is entirely diverse or non-conservative. The "diversity" or

* Corresponding author at: PIT Bioinformatics Group, Eötvös University, H-1117 Budapest, Hungary.
E-mail addresses: kerepesi@pitgroup.org (C. Kerepesi), szalkai@pitgroup.org (B. Szalkai), balorkany@pitgroup.org (B. Varga), grolmusz@pitgroup.org (V. Grolmusz).

http://dx.doi.org/10.1016/j.neulet.2017.10.003
Received 2 May 2017; Received in revised form 20 September 2017; Accepted 3 October 2017
Available online 06 October 2017
0304-3940/

"conservativeness" of lobes and ROIs are very precisely measured by cumulative distribution functions, clarified below.

The development of axonal connections through axon growth between different neurons is a stochastic procedure with numerous interacting factors, commonly described in the field called "axonal guidance" [7–11]. The axons in the connections can retract and re-connect elsewhere, depending on the activity of the target neuron [10], and the neurons, forming the axonal connections can entirely be removed by synaptic pruning, mostly between birth and adolescence [12–14].

Therefore, it is an interesting question to describe the brain areas where the stochastic process of axonal growth leads to higher and lower diversity, i.e., where the individual variability of the graph edges in the braingraph is higher and lower. In this exploratory study, we cannot aim to explain the possible causes and implications of the higher and lower diversity in cerebral connections, we intend to describe the human brain regions with more and less variable connections. In this study, we have chosen the distribution function (defined in details later) for the description of the diversity. It is clear for us that this choice is not the only possible way for the characterization of the diversity of the cerebral connections.

Fig. 1A and B contains a simplified example on three small graphs (1, 2, 3) each with only two regions (A & B). The example clarifies the method, the way the results are presented through a distribution function, and the diagrams describing these functions. In order to better follow our results, it is suggested consulting Fig. 1A and B.

For any fixed brain area, and for any $x : 0 \leq x \leq 1$, let $F(x)$ denote the fraction of the edges (i.e., the number of the edges in question, divided by the number of all edges in the fixed area) in the fixed area (i.e., with both vertices in the fixed area) that are present in at most the fraction x of all braingraphs, (for a more exact definition of $F(x)$ we refer to the "Methods" section). We note that $F(x)$ is a cumulative distribution function [15] of a random variable described in the "Methods" section and in Fig. 1A and B.

2. Results and discussion

Table 1 and Fig. 2 summarize the edge diversity results for the 392 graphs for the lobes of the brain, described by the distribution functions $F(x)$. The last column of Table 1 contains the data for the whole brain with 1015 nodes and 70,652 edges. The sum of the edges of the lobes in Table 1 is 30,326: these edges have both endpoints in the same lobe. More than forty thousand edges are present and accounted for only in the last column because these edges connect nodes from different lobes. Therefore, the values in the last column cannot be derived from the other columns, since that column contains the contribution of edges that do not contribute to any other columns.

We want to find out which brain areas are more conservative and which are more diverse than the others. For clarifying the terms

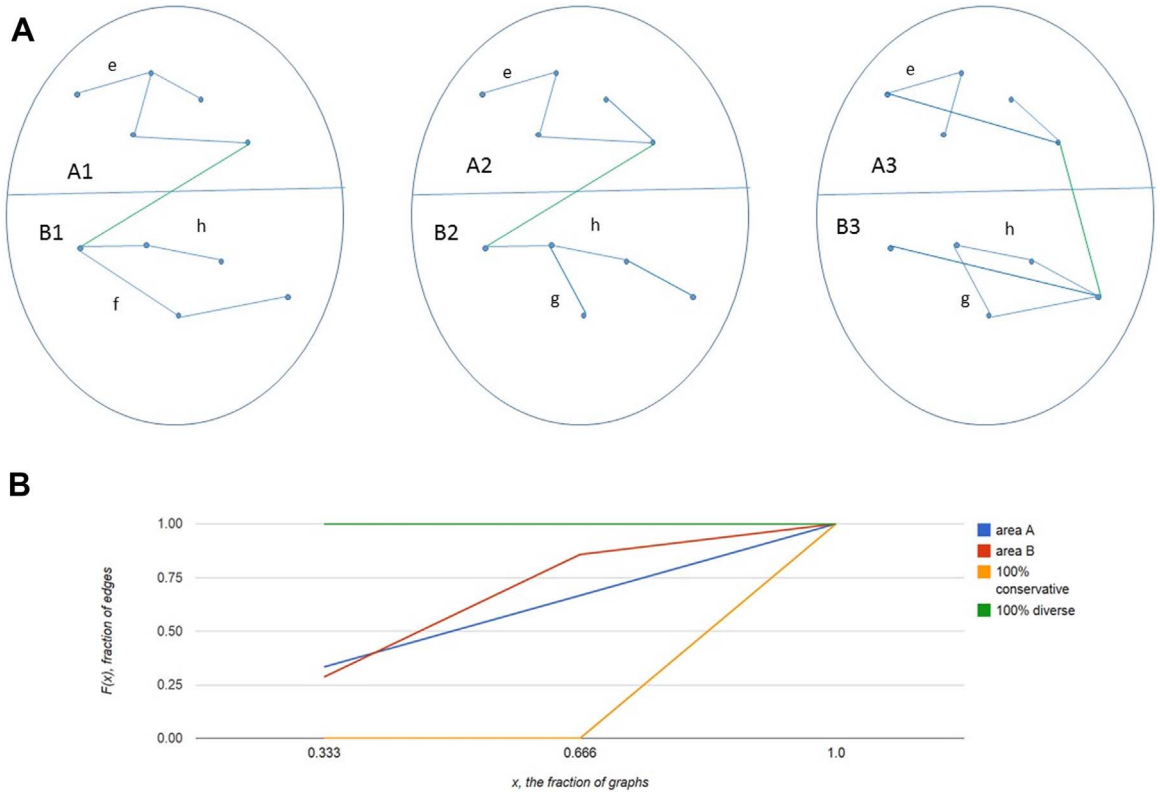

Fig. 1. (A) A simple example of computing the edge distribution between brain areas. In the example, there are three "braingraphs", each with two areas: *A* and *B*. We intend to count the edges that are present in all three graphs, only in two graphs and only in a single graph, respectively (between the same nodes, but in different graphs). For example, the copies of edge *e* are present in all three *A* areas, copies of edge *h* in all three *B* areas, copies of edge *g* in two *B* areas and edge *f* is present only in *B*1. The edges crossing the boundary of *A* and *B* (colored green) are ignored when counting the edge distribution within the areas *A* and *B*. In area *A*, two edges are present once, two edges twice and also two edges (including edge *e*) exactly three times. In area *B*, two edges (including *f*) are present once, four edges (including *g*) twice and one edge – *h* – three times. (B) The $F(x)$ distribution functions for areas *A* and *B* in A. On axis *x*, the fractions of the graphs are given, 1/3 correspond to one graph, 2/3 for two and 1.0 for all three graphs. $F(x)$ is defined as the fraction of the edges in the fixed area that are present in at most the fraction *x* of all braingraphs. Data points corresponding to area *A* are on the same blue line (1/3, 2/3, 1) and those, corresponding to area *B* are on the broken, red line (2/7, 6/7, 1). We remark that if all three graphs are the same, then the data points are (0,0,1) (the extremely conservative case, orange line). Similarly, if no two graphs have the same edges, the data points are (1,1,1) (that is the extremely diverse case, green line). This type of diagram is used for the presentation of the results of the distribution of the edges in separate areas of the brain: The faster the line reaches the top $F(x) = 1$ value, the more diverse is the edge set in the corresponding brain area. We also note that in the diagram the lines connect the data points corresponding to the discrete values on axis *x*, and *do not* describe the step-function $F(x)$ *between* the data points: we have chosen this visualization method because of its clarity even if a higher number of areas are shown (cf. Figs. 2 and 3 with numerous crossing lines). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)