RESEARCH ARTICLE

The Frequent Network Neighborhood Mapping of the human hippocampus shows much more frequent neighbor sets in males than in females

Máté Fellner[1], Bálint Varga[1], Vince Grolmusz[1,2] *

1 PIT Bioinformatics Group, Eötvös University, Budapest, Hungary, **2** Uratim Ltd., Budapest, Hungary

* grolmusz@pitgroup.org

Abstract

In the study of the human connectome, the vertices and the edges of the network of the human brain are analyzed: the vertices of the graphs are the anatomically identified gray matter areas of the subjects; this set is exactly the same for all the subjects. The edges of the graphs correspond to the axonal fibers, connecting these areas. In the biological applications of graph theory, it happens very rarely that scientists examine numerous large graphs on the very same, labeled vertex set. Exactly this is the case in the study of the connectomes. Because of the particularity of these sets of graphs, novel, robust methods need to be developed for their analysis. Here we introduce the new method of the Frequent Network Neighborhood Mapping for the connectome, which serves as a robust identification of the neighborhoods of given vertices of special interest in the graph. We apply the novel method for mapping the neighborhoods of the human hippocampus and discover strong statistical asymmetries between the connectomes of the sexes, computed from the Human Connectome Project. We analyze 413 braingraphs, each with 463 nodes. We show that the hippocampi of men have much more significantly frequent neighbor sets than women; therefore, in a sense, the connections of the hippocampi are more regularly distributed in men and more varied in women. Our results are in contrast to the volumetric studies of the human hippocampus, where it was shown that the relative volume of the hippocampus is the same in men and women.

Introduction

While it seems to be clear for all brain scientists that the complex connection patterns of the neurons play a fundamental role in brain function [1–3], when the large-scale, macroscopic description of these connections has become available by the development of diffusion MRI techniques, it turned out that novel methods are needed to handle these large graphs [1, 2]. MRI-mapped human connectomes have only several hundred or at most one thousand vertices today [4], and, therefore, more complex, more refined graph theoretical algorithms [5–7]

study we have used exclusively the 463-node resolution graphs. S1 and S2 Tables contain the one, two three and four-element subsets of the neighbor-sets of the left- and the right hippocampi, respectively, with a minimum frequency of 0.9. The S3 Table lists those neighbor-sets of the left hippocampus, which are significantly more frequent in female connectomes; while S4 Table lists those, which are significantly more frequent in male connectomes. Similarly, S5 and S6 Tables contain the analogous data for the right hippocampus. S7 Table contains the data, analogous with Table 3 in the main text, with equal men- and women cohort sizes. The supplementary tables can be downloaded in a compressed MS Excel format from http://uratim.com/hpc/supplementary_HCP_neighbors.zip.

can be applied for their analysis than the widely followed network science approach, originally developed for tens of millions of vertices of the web graph [8].

The need for robust analytical methods

In the mathematical analysis of the large biological graphs, it is necessary to apply methods that are capable of dealing with the frequently erroneous networks, computed from high-deviation experimental data [9–11]. Our research group has developed several fault-tolerant analytical methods for biological graphs, based on a well-known robust network measure: Google's PageRank [12–15].

Today, one of the main challenges in brain science is the detailed mapping of the brain circuitry with reliable methods. While diffusion MRI yields a relatively reliable set of tools for brain imaging, the workflow of identifying the common, human brain areas across different subjects and the tractography phase of the processing have numerous difficulties [16–18]. Human braingraphs, or connectomes are very focused structures for the description of the cerebral connections: the nodes of these graphs are the anatomically identified small (1-1.5 cm^2) areas of the gray matter (called ROIs, Regions of Interests), and two nodes, corresponding to two ROIs, are connected by an edge if in the tractography phase an axonal fiber is found, connecting these two ROIs. These connectomes describe the macroscopical scale set of connections between hundreds of brain areas. Before the era of diffusion MR imaging, only very fractional knowledge was available on these connections [1]; today tens of thousands of connections can be identified and examined.

Previous work

Perhaps the most straightforward robust approach to be considered is the study of the frequently appearing cerebral connections. In work [19] we have mapped the differences in the individual variability of the connections within the lobes and some smaller brain areas.

We have constructed the Budapest Reference Connectome Server [20, 21] at the address http://pitgroup.org/connectome/, which is capable of generating consensus connectomes from the data of 477 subjects, consisting of k-frequent edges (i.e., edges that are present in at least k braingraphs), with user-selected k and other parameters.

The Budapest Reference Connectome Server is an excellent tool for generating a robust human connectome, and, additionally, its construction has led to the discovery of the phenomenon of the Consensus Connectome Dynamics (CCD) [6, 22–24], mirroring the development of the axonal connections within the human brain.

The global, or general approaches for describing the frequent (and, therefore, robust) cerebral connections are not always detailed enough for the study of specific brain regions. Additionally, the frequently appearing connections (e.g., in the Budapest Reference Connectome Server [20, 21]) may describe the frequent neighbors of the individual vertices of these graphs, but not the frequently appearing neighbor-sets of important vertices. The description of these neighbor-sets is the goal of the present work.

Our contribution: The Frequent Network Neighborhood Mapping

First, let us introduce some basic terminology. A graph $G(V, E)$ consists of a vertex-set V and an edge-set E. Edges are formed from some (un-ordered) pairs of the vertices V. Let u and v be vertices, then we say that u is a neighbor of v if the un-ordered pair $\{u, v\}$ is an edge of the graph.

In braingraphs, the elements of V correspond to anatomically identified areas of the gray matter of the brain. If we have N braingraphs, corresponding to the connectomes of N human

subjects, then in all of them, the vertex-set V is the same. However, the edge-set is typically differ from graph to graph. That is, we have N graphs on the very same vertex set V.

Let us consider an important small area, corresponding to a vertex in the connectome (called ROI, region of interest) of the brain, say the left hippocampus. It is an important question to describe those ROIs, which are directly connected to the left hippocampus since all the connections to and from the other cerebral areas go through these edges and these neighbors of the important ROI (in our case the hippocampus). It is a very interesting question whether almost all the subjects have the same neighbors of the hippocampus, or there is a considerable variability among the subjects.

If there were no any variability among the connectomes of the individual subjects, then in each connectome, the left hippocampus would be connected to the very same set of other nodes or ROIs. However, there is a considerable variability of these connections between distinct subjects [19–21]. Therefore, no such common neighbor-set exists for any vertex in the braingraphs.

Instead of the non-existent single, universal neighbor-set (with the frequency of 100%), which would have been present in the connectome of all the subjects, we can still identify at least the frequently appearing neighbor-sets of the left hippocampus (or any other given vertex of interest in the graph), with a cut-off value of, say, 80% or 90%. The formal definition is as follows:

Definition 1. *Suppose we have* N *graphs on the same vertex set V. Let $v \in V$, and let $0 \le c \le 1$. Then the vertex-set W is a c-frequent neighbor set of vertex v, if, in at least cN graphs, all the vertices of W are connected to v.*

Simple examples on Figs 1 and 2 demonstrate and clarify frequent neighbor sets.

The identification of frequent neighbor sets is a completely different task than identifying the frequently appearing edges of the connectome, which are mapped by the Budapest Reference Connectome Server at http://pitgroup.org/connectome. As an example, let us consider a vertex u, and two other vertices v and w. Suppose, that both edges $\{u, v\}$ and $\{u, w\}$ are present in 90% of all connectomes, but it may happen that the vertex-set $\{v, w\}$ appears only in the 80% of the graphs as the neighbors of vertex v: if the connectomes are indexed from 1 through 100,

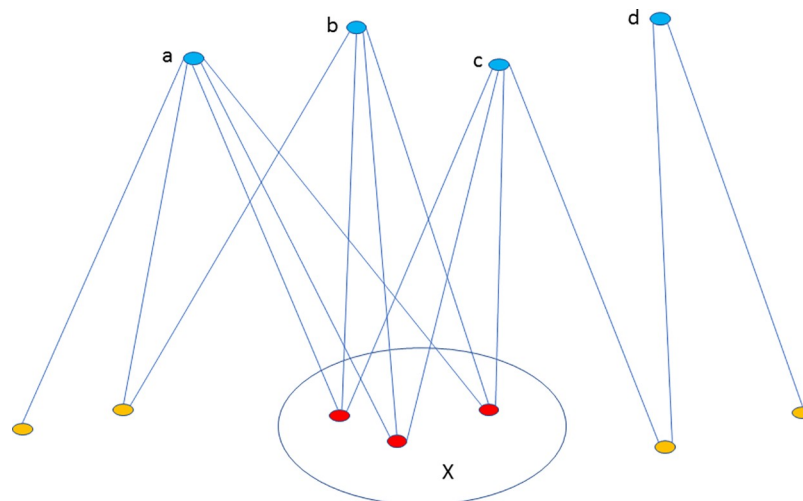**Fig 1. The visualization of the frequent neighbor sets.** Suppose we have four braingraphs from four subjects. Blue vertices a, b, c and d denote the same, interesting vertex in the braingraphs of four subjects, say the left hippocampus. Then all the red vertices of the set X are connected to a, b and c; that is, the frequency of X is 3/4. See Fig 2 for further explanation.
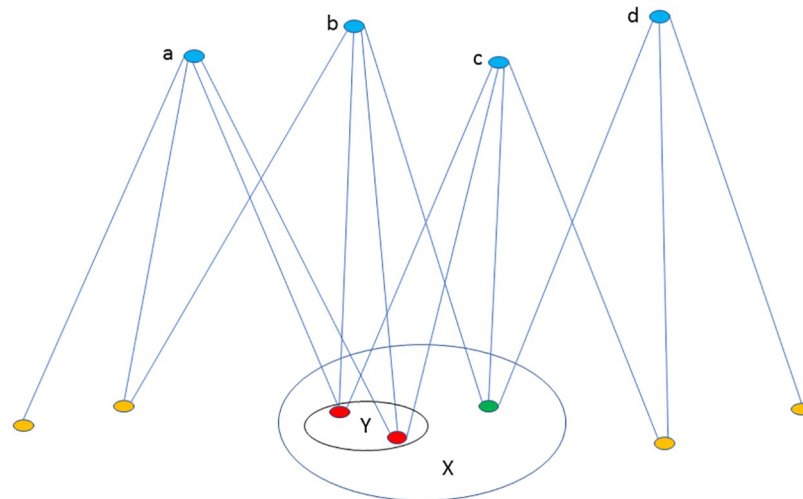
https://doi.org/10.1371/journal.pone.0227910.g001

Fig 2. This figure contains a small change, relative to Fig 1. In Fig 1, X was a frequent neighbor set, with a frequency of 3/4 = 0.75. Here, set Y with red nodes is a frequent neighbor set with a frequency of 3/4 = 0.75, but X is not, because no three-vertex subset of {a, b, c, d} are connected to all of the elements of X this time: the green vertex is connected to d instead of a. Note, that the degrees of the vertices of X are all equal to 3, both on Figs 1 and 2; therefore, the simple degree-counting does not help in the identification of frequent neighbor sets.

https://doi.org/10.1371/journal.pone.0227910.g002

it may happen that in connectomes 1 through 90 {u, v} is an edge; in connectomes 11 through 100 {u, w} is an edge; so both of them are present in 90% of the connectomes, but the set {v, w} is neighboring with u from connectomes 11 through 90, i.e., only 80 connectomes, i.e., 80%.

In the present contribution, we map the frequently appearing neighbor-sets of the left and the right hippocampi of the human connectome, and we make comparisons between the lateral and sex-differences in the frequent neighbor-sets of the left- and right hippocampi. The neighbor set sizes are bounded by 4 in our present study, since if we considered larger sets than 4, the numbers of the sets would be increased considerably. Our braingraphs have 463 vertices for all subjects.

Our results show strong differences in the neighbor-sets of the hippocampi between the sexes: we mapped the neighbor-sets, which have significant differences in frequency in men and women (we call these vertex sets "significant neighbor-sets"), and we have found that

- the number of the significant neighbor-sets of the left hippocampus is 65 times higher in males than in females;

- the number of the significant neighbor-sets of the right hippocampus is 16 times higher in males than in females;

In a sense, these results show that the neighbor sets of the hippocampus of the women are more varied between individuals, while these sets are more regular, that is, less varied in the case of men's connectomes. These results complement our general studies of the deep graph-theoretical parameters of the connectomes of men and women [5, 7, 25], where it was proven that women's connectomes are better "connected", in precisely defined mathematical and computer engineering terms.

Methods

The primary data source of the present study is Human Connectome Project's website at http://www.humanconnectome.org/documentation/S500 [26], containing the HARDI MRI

data of 413 healthy human subjects between 22 and 35 years of age. From the subjects, there were 238 women and 175 men.

We have analyzed the data of these 238 women and 175 men. Since the male and female subjects were present in different numbers in the set, and we have made comparisons between the results of the sexes below, we needed to analyze the possible effects of these differences. The details of this analysis are given in the subsection "Handling the different set sizes", below.

Construction of the graphs

The workflow for generating the connectomes has applied the CMTK toolkit [27], including the FreeSurfer tool and the MRtrix tractography processing tool [28] with randomized seeding and with the deterministic streamline method, with 1 million streamlines. For the present study we have applied the 463-vertex graph resolution.

The parcellation data (containing the ROI labels) is listed in the CMTK nypipe GitHub repository https://github.com/LTS5/cmp_nipype/blob/master/cmtklib/data/parcellation/lausanne2008/ParcellationLausanne2008.xls.

The further details of braingraph-constructions are given in [29].

The braingraphs can be downloaded from the http://braingraph.org/cms/download-pit-group-connectomes/ site. In the present study we have used the 463-node resolution graphs.

The apriori algorithm

The apriori algorithm [30] is a well-known tool in data mining for selecting the frequent item sets from a large collection of subsets of a big item set. In constructing association rules [30, 31] the selection of frequent subsets is a basic step of the rule construction. In general, an n element set has 2^n subsets, therefore for not-too-small n's it is not feasible to review *all the 2^n* subsets. However, if we want to identify only the subsets with high enough frequency (or "support", in data mining terms [32]), then we can make use of the following observation: Suppose that the set A has frequency $\alpha \geq 0$. Then all subsets of A has a frequency at least α. Therefore, first, we identify those 1-element subsets, which has a frequency at least α, it is an easy task. Then, for identifying all the 2-element subsets of frequency α we need to screen only the pairs of the frequent 1-element subsets. Next, for identifying the frequent 3-element subsets, we take all the frequent 2-element subset appears with exactly one common element, and verify if their union is frequent or not. The algorithm is continued in a similar way, and usually, it finishes quickly.

Here we have applied an adaptation of the apriori code from the website http://adataanalyst.com/machine-learning/apriori-algorithm-python-3-0/ with small modifications.

Statistical analysis

The identification of the frequent neighbor sets of the hippocampus was done by a two-step method: first we partitioned the set of the graphs into two disjoint sets, and next, the frequent neighbor sets were identified separately for each set. Only those neighbor sets were used for the study, which were frequent for both sets. Technically, the 238 female graphs were partitioned into a 130 and a 108-member set; the 175 male graphs were partitioned into two sets of size 91 and 84. The partitioning was done by the parity of the second rightmost digit of their ID (this partitioning can be considered quasi-random).

For a chosen frequent neighbor set F, we have counted its occurrences in the male dataset by $count_1(F)$ and in the female dataset by $count_2(F)$. The support was calculated as $supp_i(F) = \frac{count_i(F)}{S_i}$, where S_i, for $i = 1, 2$, is the number of male and female braingraphs,

respectively. For each set F we need to determine whether $supp_1(F)$ and $supp_2(F)$ significantly differ. For this goal we used the chi-squared test for categorical data:

	contains F	does not contain F	total
1st sample	$count_1(F)$	$S_1 - count_1(F)$	S_1
2nd sample	$count_2(F)$	$S_2 - count_2(F)$	S_2
total	$count_1(F) + count_2(F)$	$S_1 + S_2 - count_1(F) - count_2(F)$	$S_1 + S_2$

Then the test is calculated as

$$\chi^2 = \frac{(count_1(F) \cdot (S_2 - count_2(F)) - count_2(F) \cdot (S_1 - count_1(F)))^2 \cdot (S_1 + S_2)}{S_1 \cdot S_2 \cdot (count_1(F) + count_2(F)) \cdot (S_1 + S_2 - count_1(F) - count_2(F))}$$

The degree of freedom for this test is one because it is the number of samples minus one times the number of categories minus one.

Holm-Bonferroni correction [33]: After computing the p value for every frequent set, we ordered these p-values $p_1 \leq p_2 \leq p_3 \leq \ldots \leq p_m$. With a significance level $\alpha = 0.01$ let the Holm-Bonferroni value for k^{th} frequent set be $p'_k = \frac{\alpha}{m+1-k}$. Then let t be the minimal index such that $p_t > p'_t$: we have to reject the null hypotheses for i indices $i \leq t$. If $t = 1$, then we do not reject any null hypotheses thus the difference in supports is significant.

Handling the different set sizes

Our source contains the data of 413 healthy human subjects between 22 and 35 years of age. From the subjects, there were 238 women and 175 men. Since we have identified structures, which are present in 80% or 90% of the subjects of each sex, this condition is less restrictive for the men, since their cardinality is less than that of the women. In order to handle this possible problem, we have also checked our results with subsets of 175 women and the same set of 175 men. For choosing which female subjects to include, we follow a randomized choice strategy. We have taken 10 randomly chosen 175-element subsets of the 238-element female data set (i.e., we have examined 175 males and 175 females in each run), and computed the frequent neighbors for these 175-element sets, then take the average for these 10 random subsets of the female subjects.

The results are given in the S7 Table. Comparing the results with Table 1 of the main text, one may conclude that the numbers do not differ considerably. Therefore, the male-female differences we have found are due to sex, and not data-set size differences.

Discussion and results

Mapping the frequent graph-theoretical structures in the human braingraph makes possible the robust analysis of the possibly noisy data. If the frequency count is set to a high enough value, (say, the structure in question needs to be present in 80% or 90% of the connectomes of the subjects in the group analyzed), then image acquisition artifacts, data processing errors and small, random individual variabilities in the connectomes could be countermeasured.

Here we consider the hippocampus and its neighbors of the human brain. The hippocampus plays an important role in numerous brain functions, like the processing of short-time memory and turning it into long-time memory, in spatial memory and orientation [34–37]. Today, the hippocampus is perhaps the most widely studied functional and structural entity of

Table 1. The most frequent subsets of the neighbors of the left hippocampus (cut frequency value: 0.995). 413 subjects from the Human Connectome Project public release were examined. In all the subjects, the left hippocampus is connected to the following three ROIs: the Left-Putamen, Left-Thalamus-Proper and the lh.isthmuscingulate_3, and, therefore, to all 7 (= $2^3 - 1$) non-empty subsets of those (the first 7 lines of the table). The lh.superiortemporal_2 ROI is connected to the left hippocampus in all, but one subjects, so its frequency is 412/413 = 0.99758. Note that no subset, containing lh.superiortemporal_2, may have a higher frequency than the frequency of lh. superiortemporal_2 alone, i.e., 0.99758. Indeed, all the 8 (empty and non-empty) subsets of the Left-Putamen, Left-Thalamus-Proper and the lh.isthmuscingulate_3, together with the lh.superiortemporal_2 is present in 412 out of 413 subjects, i.e., with 0.99758 frequency. The ROI lh.insula_2 is connected to the left hippocampus in 411 subjects out of the 413 subjects, therefore its frequency is 411/413 = 0.99517. Similarly, all the 8 subsets of the ROIs Left-Putamen, Left-Thalamus-Proper, and the lh.isth-muscingulate_3, together with lh.insula_2 have the same 411/413 = 0.99517 frequency. Theoretically, lh.insula_2, together with the lh.superiortemporal_2 may have the same frequency as lh.insula_2 alone (if the subject, where the left hippocampus—lh.superiortemporal_2 edge is missing is one of the two subjects where the left hippocampus—lh.insula_2 edge is missing), but this is not the case: the frequency of the neighbor set {lh.superiortemporal_2,lh.insula_2} is below the cut frequency value for Table 2, i.e., 0.995.

Frequency	Subset
1	(Left-Putamen)
1	(Left-Thalamus-Proper)
1	(lh.isthmuscingulate_3)
1	(Left-Putamen)(Left-Thalamus-Proper)
1	(Left-Putamen)(lh.isthmuscingulate_3)
1	(Left-Thalamus-Proper)(lh.isthmuscingulate_3)
1	(Left-Putamen)(Left-Thalamus-Proper)(lh.isthmuscingulate_3)
.	....
0.99758	(lh.superiortemporal_2)
0.99758	(Left-Putamen)(lh.superiortemporal_2)
0.99758	(Left-Thalamus-Proper)(lh.superiortemporal_2)
0.99758	(lh.isthmuscingulate_3)(lh.superiortemporal_2)
0.99758	(Left-Putamen)(Left-Thalamus-Proper)(lh.superiortemporal_2)
0.99758	(Left-Putamen)(lh.isthmuscingulate_3)(lh.superiortemporal_2)
0.99758	(Left-Thalamus-Proper)(lh.isthmuscingulate_3)(lh.superiortemporal_2)
0.99758	(Left-Putamen)(Left-Thalamus-Proper)(lh.isthmuscingulate_3)(lh.superiortemporal_2)
....	....
0.99517	(lh.insula_2)
0.99517	(Left-Putamen)(lh.insula_2)
0.99517	(Left-Thalamus-Proper)(lh.insula_2)
0.99517	(lh.insula_2)(lh.isthmuscingulate_3)
0.99517	(Left-Putamen)(Left-Thalamus-Proper)(lh.insula_2)
0.99517	(Left-Putamen)(lh.insula_2)(lh.isthmuscingulate_3)
0.99517	(Left-Thalamus-Proper)(lh.insula_2)(lh.isthmuscingulate_3)
0.99517	(Left-Putamen)(Left-Thalamus-Proper)(lh.insula_2)(lh.isthmuscingulate_3)

https://doi.org/10.1371/journal.pone.0227910.t001

the human brain, and, consequently, the detailed study of its neighbor sets is a relevant area. Additionally, the detailed study of the cerebral circuitry is an emphasized research topic today: describing the direct neighbors of one of the most important brain areas in this sense is also a crucial question.

In our present study, we have discovered and analyzed frequent 1,2,3 and 4 element neighbor-sets of the left and the right human hippocampi.

Frequent Network Neighborhood Mapping of the hippocampus

From now on, we refer to the brain areas, or ROIs, by their resolution-250 parcellation labels, based on the Lausanne 2008 brain atlas [38] and computed by using FreeSurfer [39] and CMTK [27, 40], listed at https://github.com/LTS5/cmp_nipype/blob/master/cmtklib/data/parcellation/lausanne2008/ParcellationLausanne2008.xls. The "lh" and the "rh" prefixes abbreviate the "left-hemisphere" and "right-hemisphere" localizations.

Table 2. The summary of the results for sex differences. The first column list the minimum support, or, in other words, the frequency cut-off values: there are two values: 0.8 or 0.9, i.e., 80% 90%. The second column denotes the righ-, left- or both hippocampi; the abbreviation HPC stands for the word "hippocampus". In the third column the sex is given; the next four columns contain the number size 1, 2, 3 and 4 frequent neighbor-sets of the hippocampus considered. The next column gives the number of the neighbor-sets, which have significantly different frequencies (p = 0.001) in male and female connectomes. The last, ninth column gives the number of neighbor-sets, which are significantly more frequent in male or in female connectomes: the sum of the two numbers in the ninth column is equal to the number in the eighth column. For example, in the first row, we can see that in males, the left hippocampus has 45 frequent 1-element neighbor sets; 844 frequent 2-element neighbor sets, 9102 3-element neighbor sets and 65150 frequent 4-element neighbor sets, where the frequency cut-off is 0.8. Moreover, one can see that there are 15732 sets, differing significantly in frequency in males and in females; and the last column says that from these 15732 sets, 15497 are present in the braingraph of males and only 235 in the braingraphs of females.

Support	Node	sex	Size 1	Size 2	Size 3	Size 4	# significant differences	Sign. diff. for whom
0.8	HPC left	male	45	844	9102	65150	15732	15497
0.8	HPC left	female	38	679	7128	48717		235
0.8	HPC right	male	50	1023	11989	91498	1762	1659
0.8	HPC right	female	50	997	10725	73424		103
0.9	HPC	male	64	1781	28823	309659	19828	17688
0.9	HPC	female	62	1406	17976	150167		2140
0.8	HPC left	all	43	759	7863	54861		
0.8	HPC right	all	52	1059	11963	85624		
0.9	HPC	all	68	1728	24741	233548		

We have mapped separately the direct neighbor sets of the left- and the right hippocampi.

We need to recall an important property of the frequencies of the neighbor sets of a given vertex, in our case the left- or the right hippocampus. We say that a set U is a subset of set V, if every element of U is, at the same time, the element of V. V is called the superset of U. We denote this relation as follows: $U \subseteq V$ or, equivalently, $V \supseteq U$.

Let us consider the left hippocampus. For every neighbor-set U of the left hippocampus, we assign a frequency value as follows: we count the graphs of the subjects, where every element of set U is connected to the left hippocampus, and divide this number by the number of all the graphs. Let us denote this frequency value by $\phi(U)$. Clearly, if V is a superset of U, then the frequency of V cannot be larger than the frequency of U: $\phi(U) \geq \phi(V)$.

This observation concerning the frequencies, of course, holds for the neighbor-sets of any vertex of our graphs.

Table 2 lists the neighbor-sets of the left hippocampus with a minimum frequency of 0.995. Three ROIs (the Left-Putamen, Left-Thalamus-Proper and the lh.isthmuscingulate_3) are connected to the left hippocampus in all braingraphs, while several other ROI (such as the lh. superiortemporal_2 and the lh.insula_2), together with some subsets of the first three ROIs form the remaining neighbor sets with a minimum frequency of 0.995.

The S1 and S2 Tables contain the one, two three and four-element subsets of the neighbor-sets of the left- and the right hippocampi, respectively, with a minimum frequency of 0.9.

Sex differences

In what follows we compare the neighbor sets of the left- and right hippocampi in braingraphs, computed from the male and female subjects of the dataset of the Human Connectome Project [26]. We identify those neighbor sets of the hippocampus that are significantly more frequent in male- and in female connectomes. We will see that male connectomes have much more frequent hippocampal neighbor sets than female connectomes.

By our knowledge, this is the first observed significant sex difference in the connections of the hippocampus in the literature.

Anatomical sex differences in the volume of the hippocampus were studied in [41]: it was found that males have larger absolute hippocampus volumes, but the relative hippocampus volume, compared to both the total brain volume or the intracranial volume, are the same in the two sexes. Here we demonstrate statistically significant differences in the number of the frequent neighbor-sets of the hippocampus in males and females.

Here we show that there are 65 times more frequent neighbor sets of size at most 4 of the left hippocampus in males than in females; and there are 16 times more frequent neighbor sets of size at most 4 of the right hippocampus in males than in females. These results show that the variability of the neighbor sets of the hippocampus is smaller in the case of males than in females: in males there are much more frequent neighbor sets than in females. In other words, the variability of the neighbor sets of the hippocampus in the connectomes of women is greater than in the case of men.

Table 1 summarizes the results for the sex differences in the frequent neighbor sets of the hippocampus. In size 1, 2, 3 and 4 neighbor-sets, males have more frequent sets than females. The only exception is the 1-element frequent neighbors of the right hippocampus, where both males and females have 50 frequent singleton sets. Since all the elements of the frequent 2,3 and 4 element subsets need to be present also as a frequent 1-element set, it is surprising that there is such a big difference in the 4 element frequent neighbors of the right hippocampus in males (91498 sets) and females (73424 sets).

The following observation is much more surprising: There is a definitive, but not too large difference between the numbers of the frequent size-1,2,3 and 4 neighbor sets of the left and the right hippocampi between the sexes. If we consider, however, the number of neighbor sets with frequencies *statistically significantly differing* (with p = 0.001) between the two sexes, we have got that 15732 sets differ, and from these, 15497 is significantly more frequent in male-, and 235 is significantly more frequent in female connectomes, in the case of the neighbors of the left hippocampus. Neighbor sets of the right hippocampus have 1762 significant differences, from which 1659 is significantly more frequent in males, and 103 in female connectomes.

If we take the union of the neighbor sets of the left- and right hippocampi, then the number of the neighbor-sets with significant differences is 19828, from which 17688 are more frequent in male connectomes, and 2140 in female connectomes (p = 0.001).

Table 3 lists 10 neighbor-sets of the left hippocampus with the most significantly different frequencies in the sexes, where the higher frequency appears at the male (5 sets) and also at the female subjects (5 sets). The S3 Table lists those neighbor-sets of the left hippocampus, which are significantly more frequent in female connectomes; while S4 Table lists those, which are significantly more frequent in male connectomes.

Table 4 lists 10 neighbor-sets of the right hippocampus with the most significantly different frequencies in the sexes, where the higher frequency appears at the male (5 sets) and also at the female subjects (5 sets). The S5 Table lists those neighbor-sets of the right hippocampus, which are significantly more frequent in female connectomes; while S6 Table lists those, which are significantly more frequent in male connectomes. The significance threshold is p = 0.01 (corrected by the Holm-Bonferroni method).

As we can see in Table 1, there is not a large difference between the frequent 1-element neighbors of the hippocampus in men and women. However, there is a very significant difference in neighbor-set frequencies of higher cardinality, as it is described in the last two columns of Table 1. One possible reason for this could be that the neighbor sets of the hippocampus in general and the left hippocampus, in particular, have less variability in the case of men than in the case of women: men have the more regularly appearing neighbor-sets, while women have more varied neighbor-sets.

Table 3. Several neighbor-sets of the left hippocampus with the most significant differences in the frequencies between the sexes. The first five rows list five subsets, which are more frequent in the braingraphs of men than of women. It is interesting that the lh.fusiform_4 ROI is present in the five sets. The next five are the most significant sets with frequencies higher in females than in males. The lh.bankssts_2 and the Brain-Stem ROIs are present in all five sets. The S3 Table lists those neighbor-sets of the left hippocampus, which are significantly more frequent in female connectomes; while S4 Table lists those, which are significantly more frequent in male connectomes, $p \leq 0.01$. Similarly, S5 and S6 Tables contain the analogous data for the right hippocampus.

frequency	frequency	name
male	**female**	
0.823	0.556	(lh.fusiform_4)(lh.precuneus_10)(lh.precuneus_11)(lh.superiortemporal_6)
0.823	0.561	(lh.fusiform_4)(lh.isthmuscingulate_2)(lh.precuneus_6)(lh.superiortemporal_6)
0.823	0.561	(lh.fusiform_4)(lh.lingual_6)(lh.precuneus_10)(lh.superiortemporal_6)
0.823	0.561	(lh.fusiform_4)(lh.lingual_8)(lh.precuneus_10)(lh.superiortemporal_6)
0.823	0.561	(lh.fusiform_4)(lh.isthmuscingulate_2)(lh.precuneus_10)(lh.superiortemporal_6)
0.880	0.967	(Brain-Stem)(lh.bankssts_2)
0.880	0.967	(Brain-Stem)(Left-Thalamus-Proper)(lh.bankssts_2)(lh.isthmuscingulate_3)
0.880	0.967	(Brain-Stem)(Left-Putamen)(lh.bankssts_2)(lh.isthmuscingulate_3)
0.880	0.967	(Brain-Stem)(Left-Thalamus-Proper)(lh.bankssts_2)
0.880	0.967	(Brain-Stem)(lh.bankssts_2)(lh.isthmuscingulate_3)

https://doi.org/10.1371/journal.pone.0227910.t003

Table 4. Several neighbor-sets of the right hippocampus with the most significant differences in the frequencies between the sexes. The first five rows list five subsets, which are more frequent in the braingraphs of men than of women. It is interesting that the rh.precuneus_4 ROI is present in all five sets. The next five sets are the most significant with frequencies higher in females than in males. The rh.fusiform_8 and the rh.middletemporal_9 ROIs are present in all five sets. The S5 Table lists those neighbor-sets of the right hippocampus, which are significantly more frequent in female connectomes; while S6 Table lists those, which are significantly more frequent in male connectomes, $p \leq 0.01$.

frequency	frequency	name
male	**female**	
0.840	0.682	(rh.entorhinal_1)(rh.parahippocampal_2)(rh.precuneus_4)(rh.transversetemporal_1)
0.909	0.774	(rh.entorhinal_1)(rh.parahippocampal_2)(rh.precuneus_2)(rh.precuneus_4)
0.914	0.782	(Brain-Stem)(rh.entorhinal_1)(rh.parahippocampal_2)(rh.precuneus_4)
0.914	0.787	(Brain-Stem)(rh.entorhinal_1)(rh.precuneus_2)(rh.precuneus_4)
0.914	0.787	(rh.bankssts_2)(rh.entorhinal_1)(rh.parahippocampal_2)(rh.precuneus_4)
0.736	0.862	(rh.fusiform_8)(rh.inferiortemporal_1)(rh.middletemporal_9)
0.769	0.887	(rh.fusiform_8)(rh.inferiorparietal_4)(rh.lingual_7)(rh.middletemporal_9)
0.769	0.887	(rh.fusiform_8)(rh.middletemporal_9)(rh.supramarginal_9)
0.791	0.900	(rh.fusiform_8)(rh.lingual_7)(rh.middletemporal_9)
0.791	0.900	(Right-Amygdala)(rh.fusiform_8)(rh.middletemporal_9)

https://doi.org/10.1371/journal.pone.0227910.t004

Conclusions

First in the literature, we have mapped the frequent neighbor sets of the human hippocampus, by applying the Frequent Network Neighborhood Mapping method. We have identified the frequent neighbor sets of the human hippocampus, and we have also compared the data of healthy young men and women in respect to the neighborhood of the hippocampus. We have found that men have much more significantly more frequent neighbor sets of the hippocampus than women. We have repeated the computations for equal numbers of men and women (c.f. S7 Table), and have found almost the same results. Our results are in contrast with the generally much better connection properties of the braingraphs of women than of men, as reported in [5, 7, 25]. Our results also need to be compared to the volumetric studies [41], where it was shown that hippocampal volumes, relative to the intracranial volume and also to

the total brain volume are the same in the two sexes. Therefore, we have shown that the neighbors of the human hippocampus significantly differ in men and women, while there are no relative volumetric differences in the hippocampus.

Supporting information

S1 Table.
(XLSX)

S2 Table.
(XLSX)

S3 Table.
(XLSX)

S4 Table.
(XLSX)

S5 Table.
(XLSX)

S6 Table.
(XLSX)

S7 Table.
(XLSX)

Author Contributions

Conceptualization: Vince Grolmusz.

Data curation: Bálint Varga, Vince Grolmusz.

Formal analysis: Máté Fellner, Bálint Varga, Vince Grolmusz.

Funding acquisition: Vince Grolmusz.

Investigation: Vince Grolmusz.

Methodology: Máté Fellner, Bálint Varga, Vince Grolmusz.

Project administration: Vince Grolmusz.

Resources: Vince Grolmusz.

Software: Máté Fellner, Bálint Varga, Vince Grolmusz.

Supervision: Vince Grolmusz.

Validation: Vince Grolmusz.

Visualization: Bálint Varga.

Writing – original draft: Vince Grolmusz.

Writing – review & editing: Vince Grolmusz.

References

1. Seung HS. Reading the book of memory: sparse sampling versus dense mapping of connectomes. Neuron. 2009 Apr; 62(1):17–29. Available from: http://dx.doi.org/10.1016/j.neuron.2009.03.020. PMID: 19376064

2. Sporns O, Tononi G, Kötter R. The human connectome: A structural description of the human brain. PLoS Computational Biology. 2005 Sep; 1(4):e42. Available from: http://dx.doi.org/10.1371/journal.pcbi.0010042. PMID: 16201007

3. Lichtman JW, Livet J, Sanes JR. A technicolour approach to the connectome. Nature Review Neuroscience. 2008 Jun; 9(6):417–422. Available from: http://dx.doi.org/10.1038/nrn2391.

4. Hagmann P, Grant PE, Fair DA. MR connectomics: a conceptual framework for studying the developing brain. Front Syst Neurosci. 2012; 6:43. Available from: http://dx.doi.org/10.3389/fnsys.2012.00043. PMID: 22707934

5. Szalkai B, Varga B, Grolmusz V. Graph Theoretical Analysis Reveals: Women's Brains Are Better Connected than Men's. PLoS One. 2015; 10(7):e0130045. Available from: http://dx.doi.org/10.1371/journal.pone.0130045. PMID: 26132764

6. Kerepesi C, Szalkai B, Varga B, Grolmusz V. How to Direct the Edges of the Connectomes: Dynamics of the Consensus Connectomes and the Development of the Connections in the Human Brain. PLOS One. 2016 June; 11(6):e0158680. Available from: http://dx.doi.org/10.1371/journal.pone.0158680. PMID: 27362431

7. Szalkai B, Varga B, Grolmusz V. Brain Size Bias-Compensated Graph-Theoretical Parameters are Also Better in Women's Connectomes. Brain Imaging and Behavior. 2018; 12(3):663–673. Available from: http://dx.doi.org/10.1007/s11682-017-9720-0. PMID: 28447246

8. Faloutsos M, Faloutsos P, Faloutsos C. On power-law relationships of the internet topology. In: ACM SIGCOMM Computer Communication Review. vol. 29. ACM; 1999. p. 251–262.

9. Goll J, Uetz P. The elusive yeast interactome. Genome Biology. 2006; 7(223):1–6.

10. Gavin AC, Aloy P, Grandi P, Krause R, Boesche M, Marzioch M, et al. Proteome survey reveals modularity of the yeast cell machinery. Nature. 2006 Mar; 440(7084):631–636. Available from: http://dx.doi.org/10.1038/nature04532. PMID: 16429126

11. Krogan NJ, Cagney G, Yu H, Zhong G, Guo X, Ignatchenko A, et al. Global landscape of protein complexes in the yeast Saccharomyces cerevisiae. Nature. 2006 Mar; 440(7084):637–643. Available from: http://dx.doi.org/10.1038/nature04670. PMID: 16554755

12. Grolmusz VI. Identifying diabetes-related important protein targets with few interacting partners with the PageRank algorithm. Royal Society Open Science. 2015; 2(4):140252. https://doi.org/10.1098/rsos.140252 PMID: 26064627

13. Grolmusz V. A note on the pagerank of undirected graphs. Information Processing Letters. 2015; 115(6):633–634. https://doi.org/10.1016/j.ipl.2015.02.015

14. Bánky D, Iván G, Grolmusz V. Equal opportunity for low-degree network nodes: a PageRank-based method for protein target identification in metabolic graphs. PLoS One. 2013; 8(1):e54204. https://doi.org/10.1371/journal.pone.0054204 PMID: 23382878

15. Iván G, Grolmusz V. When the Web meets the cell: using personalized PageRank for analyzing protein interaction networks. Bioinformatics. 2011; 27(3):405–407. https://doi.org/10.1093/bioinformatics/btq680 PMID: 21149343

16. Zhan L, Zhou J, Wang Y, Jin Y, Jahanshad N, Prasad G, et al. Comparison of nine tractography algorithms for detecting abnormal structural brain networks in Alzheimer's disease. Frontiers in Aging Neuroscience. 2015; 7:48. https://doi.org/10.3389/fnagi.2015.00048 PMID: 25926791

17. Jbabdi S, Johansen-Berg H. Tractography: where do we go from here? Brain Connectivity. 2011; 1(3):169–183. Available from: http://dx.doi.org/10.1089/brain.2011.0033. PMID: 22433046

18. Bastiani M, Shah NJ, Goebel R, Roebroeck A. Human cortical connectome reconstruction from diffusion weighted MRI: the effect of tractography algorithm. Neuroimage. 2012 Sep; 62(3):1732–1749. Available from: http://dx.doi.org/10.1016/j.neuroimage.2012.06.002. PMID: 22699045

19. Kerepesi C, Szalkai B, Varga B, Grolmusz V. Comparative Connectomics: Mapping the Inter-Individual Variability of Connections within the Regions of the Human Brain. Neuroscience Letters. 2018; 662(1):17–21. https://doi.org/10.1016/j.neulet.2017.10.003 PMID: 28988973

20. Szalkai B, Kerepesi C, Varga B, Grolmusz V. The Budapest Reference Connectome Server v2. 0. Neuroscience Letters. 2015; 595:60–62. https://doi.org/10.1016/j.neulet.2015.03.071 PMID: 25862487

21. Szalkai B, Kerepesi C, Varga B, Grolmusz V. Parameterizable Consensus Connectomes from the Human Connectome Project: The Budapest Reference Connectome Server v3.0. Cognitive Neurodynamics. 2017 feb; 11(1):113–116. https://doi.org/10.1007/s11571-016-9407-z PMID: 28174617

22. Kerepesi C, Varga B, Szalkai B, Grolmusz V. The Dorsal Striatum and the Dynamics of the Consensus Connectomes in the Frontal Lobe of the Human Brain. Neuroscience Letters. 2018 March; 673:51–55. https://doi.org/10.1016/j.neulet.2018.02.052 PMID: 29496609

23. Szalkai B, Varga B, Grolmusz V. The Robustness and the Doubly-Preferential Attachment Simulation of the Consensus Connectome Dynamics of the Human Brain. Scientific Reports. 2017; 7(16118). https://doi.org/10.1038/s41598-017-16326-0 PMID: 29170405

24. Szalkai B, Kerepesi C, Varga B, Grolmusz V. High-Resolution Directed Human Connectomes and the Consensus Connectome Dynamics. PLoS ONE. 2019 Sep; 14(4). Available from: https://doi.org/10.1371/journal.pone.0215473. PMID: 30990832

25. Szalkai B, Varga B, Grolmusz V. The Graph of Our Mind. arXiv preprint arXiv:160300904. 2016.

26. McNab JA, Edlow BL, Witzel T, Huang SY, Bhat H, Heberlein K, et al. The Human Connectome Project and beyond: initial applications of 300 mT/m gradients. Neuroimage. 2013 Oct; 80:234–245. Available from: http://dx.doi.org/10.1016/j.neuroimage.2013.05.074. PMID: 23711537

27. Daducci A, Gerhard S, Griffa A, Lemkaddem A, Cammoun L, Gigandet X, et al. The connectome mapper: an open-source processing pipeline to map connectomes with MRI. PLoS One. 2012; 7(12): e48121. Available from: http://dx.doi.org/10.1371/journal.pone.0048121. PMID: 23272041

28. Tournier J, Calamante F, Connelly A, et al. MRtrix: diffusion tractography in crossing fiber regions. International Journal of Imaging Systems and Technology. 2012; 22(1):53–66. https://doi.org/10.1002/ima.22005

29. Kerepesi C, Szalkai B, Varga B, Grolmusz V. The braingraph. org Database of High Resolution Structural Connectomes and the Brain Graph Tools. Cognitive Neurodynamics. 2017; 11(5):483–486. https://doi.org/10.1007/s11571-017-9445-1 PMID: 29067135

30. Agrawal R, Srikant R. Fast algorithms for mining association rules in large databases. In: Bocca JB, Jarke M, Zaniolo C, editors. Proc. of the 20th International Conference on Very Large Data Bases (VLDB'94),. vol. 1215. Kaufmann Publishers Inc.,; 1994. p. 487–499.

31. Agrawal R, Imielinski T, Swami AN. Mining Association Rules between Sets of Items in Large Databases. In: Buneman P, Jajodia S, editors. Proceedings of the 1993 ACM SIGMOD International Conference on Management of Data, Washington, D.C., May 26-28, 1993. ACM Press; 1993. p. 207–216.

32. Han J, Kamber M. Data Mining: Concepts and Techniques. Morgan Kaufmann Publishers; 2000.

33. Holm S. A simple sequentially rejective multiple test procedure. Scandinavian Journal of Statistics. 1979; 6(2):65–70. Available from: https://www.jstor.org/stable/4615733.

34. Santin LJ, Rubio S, Begega A, Miranda R, Arias JL. Spatial learning and the hippocampus. Revista de Neurologia. 2000; 31:455–462. PMID: 11027098

35. Voineskos AN, Winterburn JL, Felsky D, Pipitone J, Rajji TK, Mulsant BH, et al. Hippocampal (subfield) volume and shape in relation to cognitive performance across the adult lifespan. Human Brain Mapping. 2015 Aug; 36:3020–3037. https://doi.org/10.1002/hbm.22825 PMID: 25959503

36. Nees F, Pohlack ST. Functional MRIstudies of the hippocampus. Frontiers of Neurology and Neuroscience. 2014; 34:85–94. https://doi.org/10.1159/000356427 PMID: 24777133

37. Iglesias JE, Augustinack JC, Nguyen K, Player CM, Player A, Wright M, et al. A computational atlas of the hippocampal formation using ex vivo, ultra-high resolution MRI: Application to adaptive segmentation of in vivo MRI. Neuroimage. 2015 Jul; 115:117–137. https://doi.org/10.1016/j.neuroimage.2015.04.042 PMID: 25936807

38. Hagmann P, Cammoun L, Gigandet X, Meuli R, Honey CJ, Wedeen VJ, et al. Mapping the structural core of human cerebral cortex. PLoS Biol. 2008 Jul; 6(7):e159. Available from: http://dx.doi.org/10.1371/journal.pbio.0060159. PMID: 18597554

39. Fischl B. FreeSurfer. Neuroimage. 2012; 62(2):774–781. https://doi.org/10.1016/j.neuroimage.2012.01.021 PMID: 22248573

40. Gerhard S, Daducci A, Lemkaddem A, Meuli R, Thiran JP, Hagmann P. The connectome viewer toolkit: an open source framework to manage, analyze, and visualize connectomes. Frontiers in Neuroinformatics. 2011; 5(3):1–15. Available from: http://dx.doi.org/10.3389/fninf.2011.00003.

41. Tan A, Ma W, Vira A, Marwha D, Eliot L. The human hippocampus is not sexually-dimorphic: Meta-analysis of structural MRI volumes. NeuroImage. 2016 Jan; 124:350–366. https://doi.org/10.1016/j.neuroimage.2015.08.050 PMID: 26334947