1. fMRI acquisition and pre-processing

This study used a 3.0T MR scanner (Siemens Magnetom Skyra 3.0T) with a 32-channel head array coil. Subjects were scanned in a supine, head-first position. T1-weighted images were firstly obtained to exclude severe atrophy in subjects' brain. Scan parameters were as follows: MPRAGE sequence, repetition time (TR)/ echo time (TE)/inversion time (TI) = 2300/2.29/900ms, field of view (FOV) = 240 mm × 240 mm, slice thickness = 1 mm, slice gap = 0 mm, matrix = 256 × 256, number of signal averages = 1, flip angle = 8°, bandwidth = 123.36 Hz. Before resting-state fMRI scan, subjects were instructed to close their eyes and stay awake, calm breathing, keep a clear consciousness and not to engage in any specific thinking activity. fMRI data were obtained using echo planar imaging sequence with following parameters: TR / TE= 2000/30 msec, FOV = 240 mm × 240 mm, matrix = 64 × 64, slice number =33, slice thickness = 3 mm, slice gap = 1 mm, flip angle = 90°, scan duration time = 480s (240 volumes).

Data Processing and Analysis for Brain Imaging (DPABI, Version 3.0) was used for fMRI data preprocessing. Data preprocessing included the following steps [1]: (1) Data format conversion. (2) The first 10 volumes of the fMRI data were discarded due to the consideration of instability of the initial MRI signal [1]. (3) Slice timing. (4) Head motion correction by linear regression (corrected for translational and rotational misalignments). Subjects with large head motion were excluded using a threshold of maximum translation > 1.5mm, maximum head rotational motion > 1.5 degree and framewise displacement by Power et al. > 0.5mm. All the enrolled subjects passed this threshold and no data were discarded. (5) fMRI data were normalized to the Montreal Neurological Institute (MNI) space using Diffeomorphic Anatomical Registration Through Exponentiated Lie Algebra algorithm (DARTEL). During normalization, structural MRI data were co-registered to fMRI data to avoid the effect of cortical thinning or decreased brain volume on the construction of visual atlas. (6) Linear detrend of the time courses. (7) Nuisance covariates (e.g. head motion parameters, whole brain, white matter) removal. (8) Band-pass filtering (0.01-0.08 Hz). (9) Scrubbing.

2. Dice coefficient and adjusted rand index

Dice coefficient is a widely-used metric in image segmentation [1], and it is also widely used in evaluating the reproducibility of brain parcellations. It can be calculated by [1]

$$Dice = \frac{2|X \cap Y|}{|X| + |Y|} \tag{1}$$

where X and Y derives from two parcellations. A dice value of 1 represents X and Y are perfectly aligned [2].

Adjusted rand index (ARI) is also a metric to assess the reproducibility of parcellations [3]. Unlike dice coefficient, this index does not take into account clustering numbers, so it can effectively evaluate two parcellations with different cluster numbers. ARI classifies $n(n\text{-}1)/2$ pairs of voxels into one of the four sets (where n is the total number of voxels in the visual cortex), namely N_{11}, N_{00}, N_{01}, N_{10} [2,3]. For two parcellations X and Y, ARI can be calculated by [3,4]

$$ARI(U,V) = \frac{2(N_{00}N_{11} - N_{01}N_{10})}{(N_{00}+N_{01})(N_{01}+N_{11})+(N_{00}+N_{10})(N_{10}+N_{11})} \tag{2}$$

Where N_{11} represents the number of pairs from the same cluster in both X and Y, N_{00} represents the number of pairs from different clusters in X and Y, N_{01} represents the number of pairs from the same clusters in X, but different clusters in Y, and N_{10} represents the number of pairs from the same clusters in Y but different clusters in X. ARI value of 1 implies a perfect uniformity between parcellations X and Y.

3. Homogeneity index and silhouette coefficient

To measure the homogeneity index of different subregions, an average similarity between every pair of voxels within a parcel was computed which could be defined as the Pearson's correlation coefficient between the "connectivity fingerprints" of voxels [2,5,6]. For each voxel v_k, we yielded the connectivity fingerprint by correlating v_k with the rest of cortical voxels and applying Fisher's r-to-z transform to the resulting correlation coefficient [2,6].

Another widely used technique to assess the homogeneity of clustering algorithms and quantify parcellation reliability is silhouette coefficient (SC) [7]. SC could assess the uniformity of parcels as well as the degree of separation between them. It can be computed by [2,7]

$$SC = \frac{b_i - a_i}{max(a_i, b_i)} \tag{3}$$

where a_i and b_i could be calculated severally by [1,6]

$$a_i = \frac{1}{n_k - 1} \sum_{j \in U_K, i \neq j} d(v_i, v_j) \tag{4}$$

$$b_i = \frac{1}{M} \sum_{j \in \aleph(u_k)} d(v_i, v_j) \tag{5}$$

For a given parcellation U{(U_1, U_2, ..., U_k }, a_i and b_i correspond to within-parcel and inter-parcel dissimilarity with respect to voxel $v_i \in U_k$, respectively; n_k denotes the voxel number of U_k; $\aleph(U_k)$ denotes the adjacent parcel set of U_k, in which the number of voxels is M; $d(v_i, v_j)$ denotes the Pearson distance between v_i and v_j, which could be computed by 1-r, and r is Pearson correlation [2]. The value range of SC is [-1,1], and a SC value of 1 implies a high clustering quality. The calculation of SC will return a $n \times 1$ matrix, where n is the number of parcellation. We averaged this matrix to get a global score for each parcellation.

4. Calculation of network properties

In a binary network, the degree of a node is the number of edges that are connected to the node, which reflects the nodal information communication ability in the functional network [8]. The averaged nodal degree is calculated as [8]

$$k = \frac{1}{N} \sum_i k_i \tag{6}$$

Where N is the number of nodes (number of parcels), k_i is the degree of the ith node (parcel).

Nodal efficiency for a node represents the ability of information transfer for that given node in the network [8]. It can be calculated by formula (7) [8]

$$E_{nodal\_i} = \frac{1}{N-1}\sum_{j,j\neq i}\frac{1}{d_{ij}} \tag{7}$$

In Formula (7), N denotes the number of nodes in the network, in this study, the number of clusters in the visual cortex, d_{ij} denotes the distance between the ith node and the jth node.

In graph theory, a small-world network is a type of network with many short links connecting neighboring nodes, and few distant links creating shortcuts across the network [4,8]. As a result, a small-world network is more efficient in information transfer than regular networks, and has greater local connectivity than random networks [8]. For a given network, the small-world index σ can be calculated as [4,8]:

$$\sigma = \frac{c}{l} \tag{8}$$

Where c is the clustering coefficient of the network, and l is the shortest path length of the network. And a small-world index above 1 indicates a small-world topology of the given network [4].

In graph theory, the rich chub organization characterizes an architecture with densely connected nodes forming a hub in the network [8]. In contrast, the links of non-hub nodes are fewer than hub nodes. For a given network, rich club index $\emptyset$ can be calculated by the following formula [8]:

$$\emptyset(k) = \frac{2E_k}{N_k(N_k-1)} \tag{9}$$

Where k is a degree threshold, N_k is the number of nodes with nodal degree larger than k, and E_k is the number of edges among those nodes.

Reference

[1] Wang Y, Lu WZ, Yan TQ, et al. Functional MRI reveals effects of high intraocular pressure on central nervous system in high-tension glaucoma patients. Acta Ophthalmol. 2019;97:E341-E348.

[2] Arslan S, Ktena SI, Makropoulos A, et al. Human brain mapping: A systematic comparison of

parcellation methods for the human cerebral cortex. NeuroImage. 2018;170:5-30.

[2] Hubert L, Arabie P. Comparing partitions. J Classif. 1985;2:193-218.

[3] Arslan S, Ktena S I, Makropoulos A, et al. Human brain mapping: A systematic comparison of parcellation methods for the human cerebral cortex. Neuroimage. 2018;170:5-30.

[4] Craddock RC, James GA, Holtzheimer PE, Hu XP, Mayberg HS. A whole brain fMRI atlas generated via spatially constrained spectral clustering. Hum Brain Mapp. 2012;33:1914-1928

[5] Power JD, Cohen AL, Nelson SM, et al. Functional Network Organization of the Human Brain. Neuron. 2011;72:665-678.

[6] Rousseeuw PJ. Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. J Comput Appl Math. 1987;20:53-65.

[7] Jinhui W, Xindi W, Mingrui X, et al. GRETNA: a graph theoretical network analysis toolbox for imaging connectomics. Front Hum Neurosci. 2015;9:458.

Supplementary Table S1. Demographic and clinical information of POAG patients and HCs.

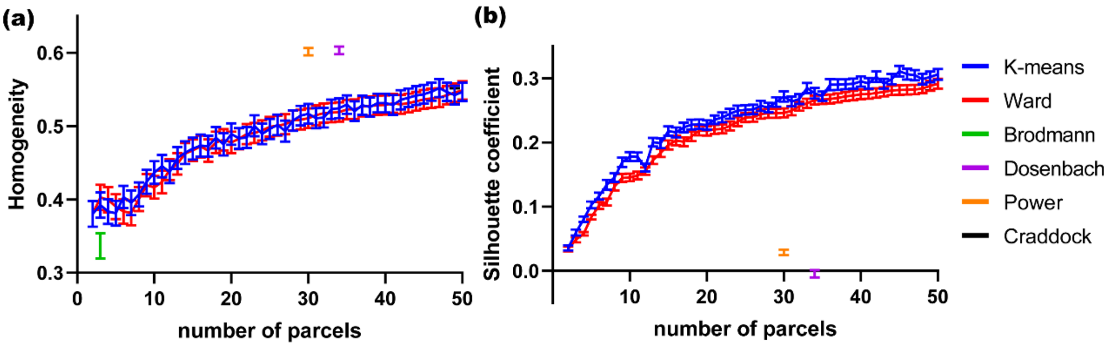
	POAG	HC
Age (years)	49.22 ± 15.26	49.90 ± 5.62
Sex (M / F)	15 / 21	7 / 13
IOP_L (mmHg)	35.42 ± 16.16	17.06 ± 3.63
IOP_R (mmHg)	33.94 ± 16.04	15.42 ± 3.68
Mean IOP (mmHg)	36.42± 12.15	16.42± 1.79
RNFL_L (μm)	74.77 ± 33.06	NA
RNFL_R (μm)	86.25 ± 22.85	NA
Disease duration (day)	136.39 ± 160.38	NA

POAG: primary open angle glaucoma; HC: healthy control; IOP: intraocular pressure; RNFL: retinal nerve fiber layer. NA represents not applicable. M and F represent male and female. L and R represent left and right.

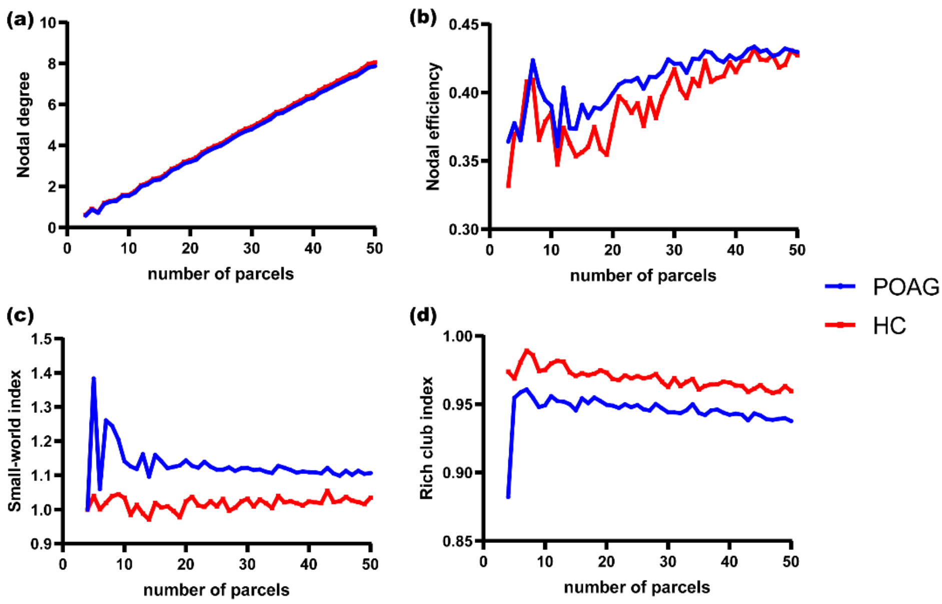
Supplementary Table S2. Two clustering approaches.

Name	Resolution	Description	Application in this study
K-means	Varying	K-means clustering algorithm first assumes X is a data set including N elements, $X=\{x_1,x_2,...,x_n\}$, and after that, this algorithm needs to find a partition $P=\{C_1,C_2,...,C_k\}$ which could minimize the function $$f(P_k) = \sum_{i=1}^{k} \sum_{xi\in ci} d(x_i, m_i)$$ Where $\sum_{xi\in ci} x_i$ denotes the clustering center of ith cluster, i=1,...,k; n_i denotes the number of data items of C_i; $m_i=1/n_i$; $d(x_i, m_i)$ denotes the distance from x_i to m_i, such as Euclidean distance.	K-means clustering was applied to the group-averaged connectivity matrix with spatial voxel coordinates to improve spatial contiguity of the parcellations.
Ward	Varying	The Ward clustering combines the two most similar data points among all the data points by calculating the similarity between the two kinds of data points, and iterates this process repeatedly. To put it simply, the hierarchical clustering is to determine the similarity	Ward clustering was applied to the group-averaged connectivity matrix with spatial voxel coordinates to improve spatial contiguity of the

between each category of data points and all data points by calculating the distance between them. The smaller the distance, the higher the similarity. The two nearest data points or categories are combined to generate a cluster tree.	parcellations.



Supplementary Figure S1. Group-level homogeneity results in HCs using parcellations generated from POAG group. (a) Homogeneity index. (b) SC.



Supplementary Figure S2. Network properties between POAG patients and HCs using the data-driven atlas (K-means algorithm) generated from 36 POAG patients. (a) Nodal degree. (b) Nodal efficiency. (c)

Small-world index. (d) Rich club index.