

MENGA 2.0: User's Notes

INTRODUCTION

MENGA (Multimodal Environment for Neuroimaging and Genomic Analysis) is a software platform that allows the investigation of the correlation patterns between neuroimaging data of any sort (both functional and structural) with mRNA gene expression profiles derived from the Allen Brain Atlas (ABA) database at high resolution. MENGA is written for Matlab 2012a version and above (The Mathworks Inc., MA, USA) and freely available. MENGA is operative system independent.

1. PREVIOUS RELEASES

If you have already installed a previous release of MENGA, do not worry: the heavy database has remained the same!

You just need to download code.zip from the website, unzip it in MENGA directory and substitute the old 'code' folder and the other main files with the new ones.

HIGHLIGHTS

- *MENGA now implements the adjusted R-squared for multivariate correlation*
- *MENGA now returns information about the directionality of the correlation*

2. GETTING STARTED

Once the program is unpackaged in its directory, start Matlab. In 'set path' add the directory MENGA_20_Export to the path and save.

Then type

```
>>MENGA
```

The menubar and the logo should appear.

3. DATA FORMATS

Region list (.mat):

The regions of interest are loaded through a .mat file, with the list of ROIs as ABA labels. Libraries are available.

Mask (.mat):

The brain mask to be applied to the image data is in a .mat file, in MNI space. Libraries are available (from FSL brain mask).

Images:

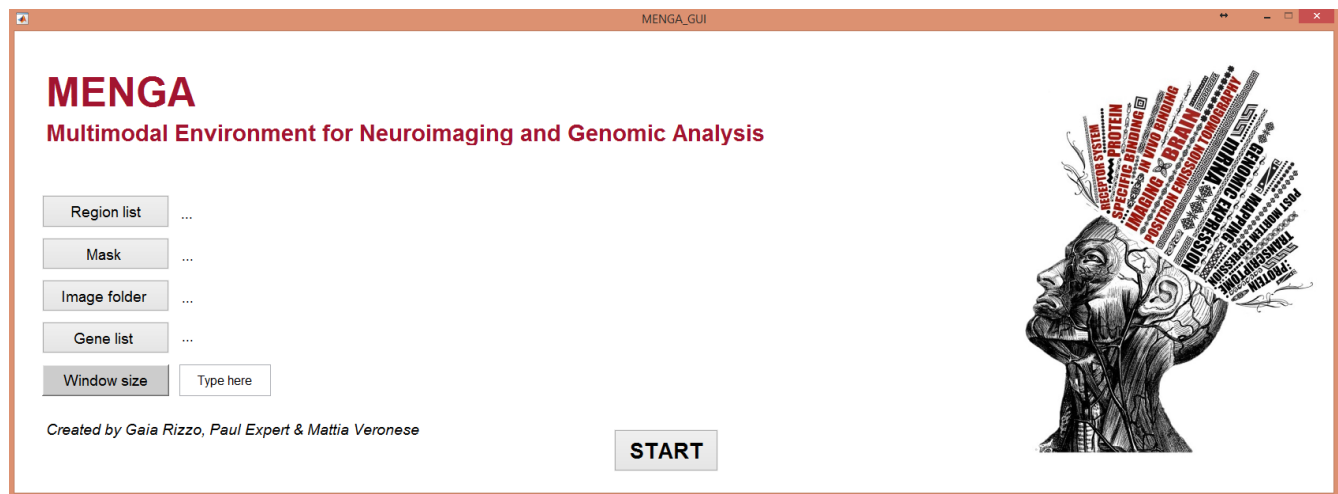
The images to be analysed must be in Analyze (.hdr/.img) or NIfTI (.nii) format. The images must be 3D, already in MNI space and oriented according to the neurological convention.

MENGA accepts images both in MNI ICBM152 coordinates (image size: 181x217x181 voxels) and FSL MNI coordinates (image size: 182x218x182) and converts them into the MNI ICBM152 convention (voxel size 1mm x 1mm x 1mm).

Gene list (.txt):

The list of genes must be a .txt file, with each gene in a new line, with no spaces. The genes are identified with their ABA abbreviation (see Allen Human Brain database or MENGA internal database). An example of gene file is available.

4. MAIN MENU & ANALYSIS



MENGA requires the user to simply fill the five fields in the GUI to run the analysis:

Region list:

It determines the level of resolution to use for the analysis. The user can choose from the already implemented solutions or extend the libraries him/herself.

MENGA already includes several default lists of regions, derived from the structure (169 ROIs) and coarse level (26 ROIs) defined in the ABA. We also included in MENGA a simplified version of the coarse level, with 15 ROIs. A summary of ABA labels is reported in the Excel file included in the manual zip package.

If no region list is selected, MENGA uses the simplified coarse list as default.

Mask:

The user has the choice to select an image mask, i.e. a binary 3D image to limit the analysis in the areas of interest removing the background. The voxels outside the mask are set to "not-a-number" (NaN) and then discarded from further analysis. Mask definition is user-dependent, and some libraries are available.

If no mask is selected, MENGA uses the FSL brain mask (limited to the left hemisphere) as default.

Image folder:

The user needs to select the folder containing the images to be analysed. All the images contained in the folder will be analysed separately. In the same folder, a new folder will be created for each image (named as the image itself) containing the MENGA results. See Output section.

If no image folder is selected, MENGA exits.

Gene list:

The user loads the gene list. There is no limit on the number of genes that can be included in the analysis, however the user needs to consider the computational load required by an analysis including hundreds of genes.

If no gene list is selected, MENGA exits.

Window size:

This integer represents the size (in mm) of the window used for the image import. This should reflect the effective spatial resolution of the image being analysed.

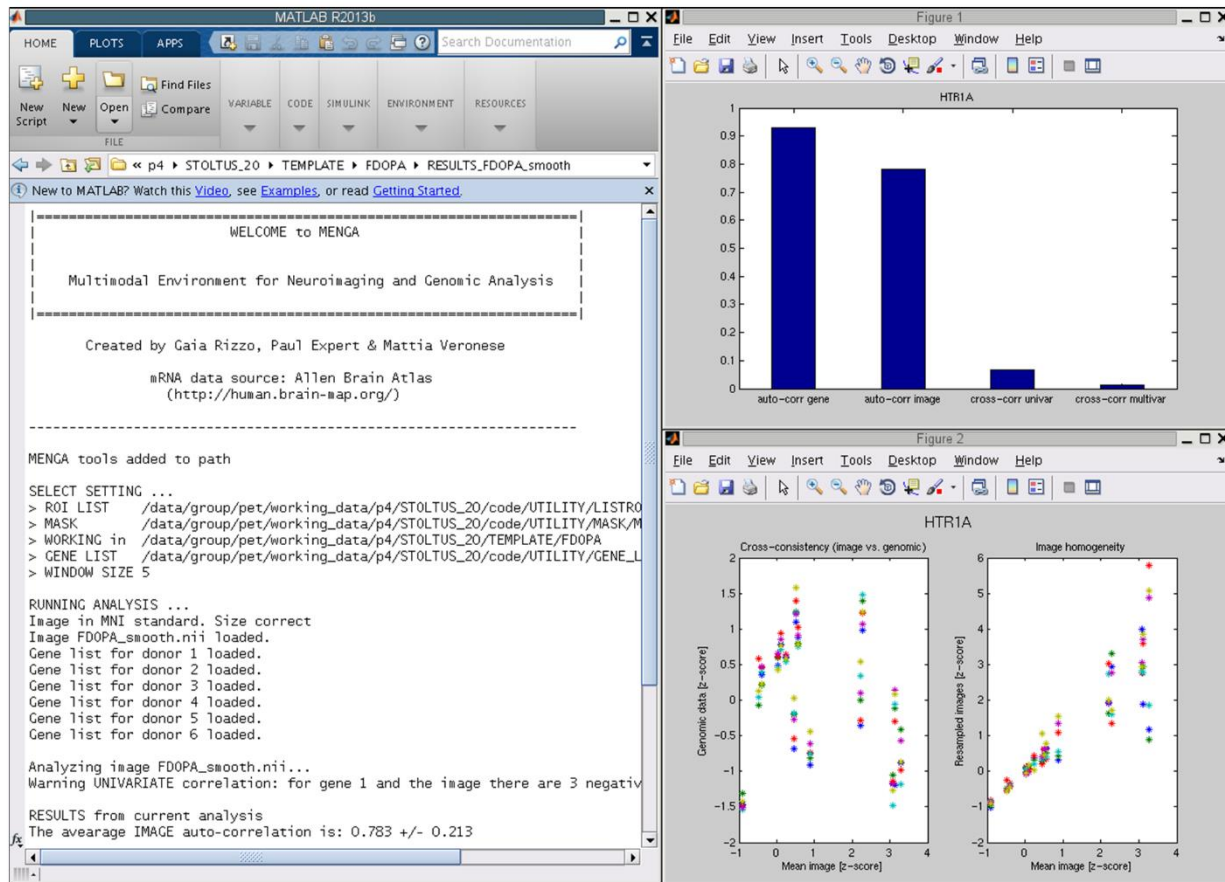
If no value is inserted, MENGA uses 5 as default.

Now just START the analysis.

5. OUTPUT

After the analysis a report with all the correlation statistics is displayed in the Matlab workspace and saved as text file (.txt) in the result folder (left panel in the figure below). The file includes also the setting information used for the analysis and a copy of the raw genomic data (in z-score).

Additionally, MENGA generates two figures (right panel in the figure below), each one saved as both .jpg and .fig files), summarizing genomic and image correlation statistics and the corresponding scatter plots.



A Matlab file for expert users is also stored in the result directory recording all of the internal results.

These include:

- *gen2don*: a structure with the genomic values in the ROIs for each donor (as mean, standard deviation, number of samples, etc.). Different genes are reported in different rows.
- *img2don*: a structure with the re-sampled image values in the ROIs for each donor (as mean, standard deviation, number of samples, etc.). Different images are reported in different rows. There are 6 fields because each **single** image is resampled in the genomic space of each donor (in order to have an exact 1:1 mRNA-image match).
- *genelabel*: list of genes
- *listroi*: list of ABA labels of ROIs
- *namelistroi*: list of ROI names
- *mask*: mask used
- *windowSize*: size of the window for image import
- *tabauto**: matrix with the genomic and image univariate autocorrelation
- *statauto**: matrix with the summary statistics of the genomic and image univariate autocorrelation (as mean, standard deviation, coefficient of variation, minimum and maximum)
- *tabcorr*: matrix with the genomic/image univariate crosscorrelation
- *statcorr*: matrix with the summary statistics of the genomic/image univariate crosscorrelation (as mean, standard deviation, coefficient of variation, minimum and maximum)
- *tabcorrMulti*: matrix with the value of multivariate crosscorrelation (one value for each gene)
- *pvalueMulti*: value of chance likelihood (one value for each gene)
- *direc*: matrix with information about the directionality of the univariate correlation (1 when positive correlation, -1 where negative)
- *direcMulti*: matrix with information about the directionality of the multivariate correlation (1 when positive correlation, -1 where negative)

If you need assistance just contact us!

gaia.rizzo@dei.unipd.it
mattia.veronese@kcl.ac.uk
paul.expert@kcl.ac.uk

6. LICENSE, DISCLAIMER AND LIMITATION OF LIABILITY

MENGA Software Suite, Release 2.0 © 2016, (the "Software").

The Software remains the property of the University of Padua and the King's College ("the Universities") and of its owners. The Software is distributed "AS IS" under this Licence solely for non-commercial use in the hope that it will be useful, but in order that the Universities protect its assets for the benefit of its educational and research purposes, the Universities make it clear that no condition is made or to be implied, nor is any warranty given or to be implied, as to the accuracy of the Software, or that it will be suitable for any particular purpose or for use under any specific conditions.

Furthermore, the Universities disclaim all responsibility for the use which is made of the Software. They further disclaim any liability for the outcomes arising from using the Software.

The Licensee agrees to indemnify the Universities and hold the Universities harmless from and against any and all claims, damages and liabilities asserted by third parties (including claims for negligence) which arise directly or indirectly from the use of the Software or the sale of any products based on the Software.

No part of the Software may be reproduced, modified, transmitted or transferred in any form or by any means, electronic or mechanical, without the express permission of the Universities. The permission of the Universities is not required if the said reproduction, modification, transmission or transference is done without financial return, the conditions of this Licence are imposed upon the receiver of the product, and all original and amended source code is included in any transmitted product.

You may be held legally responsible for any copyright infringement that is caused or encouraged by your failure to abide by these terms and conditions.

You are not permitted under this Licence to use this Software commercially. Use for which any financial return is received shall be defined as commercial use, and includes (1) integration of all or part of the source code or the Software into a product for sale or license by or on behalf of Licensee to third parties or (2) use of the Software or any derivative of it for research with the final aim of developing software products for sale or license to a third party or (3) use of the Software or any derivative of it for research with the final aim of developing non-software products for sale or license to a third party, or (4) use of the Software to provide any service to an external organization for which payment is received.

If you are interested in using the Software commercially, please contact the primary co-ordinators Gaia Rizzo, Mattia Veronese or Paul Expert.