

# **SOLAR-Eclipse computational tools for genetic and mega-genetic analysis**



# Introduction

- Imaging genetics combines modern statistical genetics approaches with quantitative neuroimaging-derived phenotypes
- The unique nature of neuroimaging phenotypes presents certain challenge in using existing statistical genetics software

# Aims

Develop imaging genetics tools that will

- Accept neuroimaging data formats
- Implement standard genetic analyses Heritability Genetic correlation QTL GWAS
- Accept populations of random complexity
- Capable of performing Meta and Mega Genetic analysis **Methods**

Sequential Oligogenic Linkage Analysis Routines (SOLAR) is the basis for the new software – SOLAR-Eclipse (SE).

SE development in-progress New data type module for reading NIFTI and GIFTI JNI interfaces for easy pipeline integration Performance optimization for parallel processing Mega-genetic analysis capability

| 🔳 💽 heron :  |   |   |                                  |                                       | $\odot$              |
|--|---|---|----------------------------------|---------------------------------------|----------------------|
| File Edit View   | Scrollback Bookmarks Setti  | ings Help   |                                  |                                       |                      |
| änter help for help,   | exit to exit, doc to browse docu  | umentation.                                       |                                  |                                       |                      |
|  |   |   | 0XEL_31_108_104 V0XEL_31_109_103 | 8 V0XEL_31_109_104 V0XEL_31_110_103 V | 0XEL_31_111_102 V0XE |
| solar> trait VOXEL_3<br>solar> polygen                       | 1_107_105   |   |                                  |                                       |                      |
| <ul> <li>Maximize sporadic</li> </ul>                        | model   | ·····   |                                  |                                       |                      |
| *** Loglikelihoo   | d of sporadic model is -6771.1695   | 547   |                                  |                                       |                      |
|  |   |   |                                  |                                       |                      |
| <ul> <li>Maximize polygeni</li> </ul>                        | c model   | ·····   |                                  |                                       |                      |
|  | d of polygenic model is -6771.169<br>enic model is 0.0000000  | 547   |                                  |                                       |                      |
| *** Comparing po   | significance of H2r<br>lygenic and sporadic models<br>, deg = 1, p = 0.5000000  |   |                                  |                                       |                      |
|  |   |   |                                  |                                       |                      |
| •  | Summary of Results  | •••••••••••••••••••••••••••••••••••••••           |                                  |                                       |                      |
| Pedigree:<br>Phenotypes:<br>Trait:                           | safs_gwa_ped.out<br>out_0.csv<br>V0XEL_31_107_105 Individual  | ls: 814   |                                  |                                       |                      |
|  | H2r is 0.0000000 p = 0.5000   | 0000 (Not Significant)                            |                                  |                                       |                      |
| Summary resu<br>Loglikelihoo<br>Best model i<br>Final models | and models are in diractory VOXE<br>lts are in VOXEL_31_107_105/polyg<br>ds and chi's are in VOXEL_31_107_<br>s named poly and nullO (currently<br>are named poly, spor | genic.out<br>_105/polygenic.logs.out<br>/ loaded) |                                  |                                       |                      |
| Residual Kur<br>solar> 📕                                     | tosis is 0.4547, within normal ra   | Inge  |                                  |                                       |                      |
|  | 10.0.4.101 :  | peterk : R  | 🕅 10.0.4.101 ; petr              | 10.0.4.101 ;                          |                      |

Figure 1. Solar interface

• [R]-like visual interface with support for shell and TCL scripts • SE new phenotype file format was created to accommodate binary voxel-and-vertex wise data types (Figure 2) • New command lines were developed to accommodate for parallel processing of binary phenotypes using mask command

(Figure 3).

Different pedigrees are combined into a superfamily

dataset.

Heritability estimates for each population and a combined

pedigree are calculated to test significance of populationspecific difference and the combined significance of heritability or fixed effects.

2)

- solution



## Peter Kochunov, Neda Jahanshad, Charles Peterson, Thomas Nichols, Bennett Landman, Paul Thompson, David Glahn and John Blangero.

Maryland Psychiatric Research Center. University of Maryland School of Medicine Baltimore

### Progress

### Results

## • Mega-genetic analysis can be performed by combining diverse sample into a single pedigree

- Class variable is added to identify separate pedigrees Samples for which raw data is available can be combined into a single pedigree.
  - Sporadic effects are calculated individually for each sample. Inverse normal transformation is used to Z-normalize each

Solar phenotype files read binary traits (Figure

## Mask command provides for parallelization Limitations

•Memory is a limiting factor for large studies •Memory chunking is supported as ad-hoc

| ID     | Age | Sex | DTIFA: <b>NIFTI</b>      |  |  |
|--------|-----|-----|--------------------------|--|--|
| A00001 | 21  | Μ   | Gobs_fa.nii.gz <b>:1</b> |  |  |
| A00002 | 34  | F   | Gobs_fa.nii.gz <b>:3</b> |  |  |
| A00003 | 24  | Μ   | Gobs_fa.nii.gz <b>:2</b> |  |  |

Figure 2. Example for the new phenotype file CSV file format. Semicolon is used to provide identification of the file type in the header and the sequence of the volume/surface in the combined file.

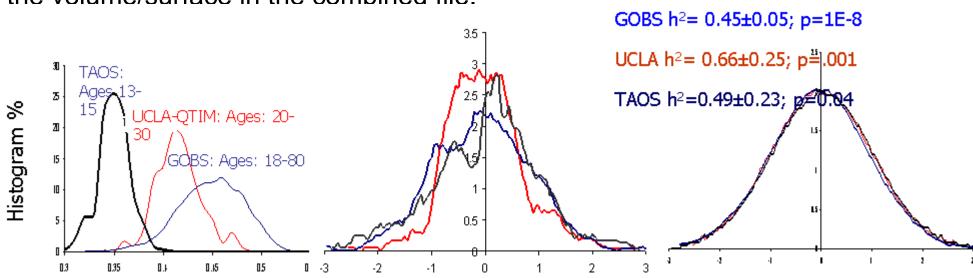


Figure 3. Mega-genetic processing Mega Analysis

h<sup>2</sup>=0.47±0.05; p=10-16

Meta Analysis SE-Weighted

h<sup>2</sup>=0.48 ±0.07; p=0.004 Meta Analysis N-Weighted h<sup>2</sup>=0.44 ±0.03; p=10<sup>-6</sup>

Figure 4. Mega- versus Metagenetic analysis

Conclusion

We are developing SOLAR-Eclipse into an imaging genetic software toolkit.

See our workshop at Imaging Genetics Conference http://www.imaginggenetics.uci.edu

Funding provided by NIH-NIBIB EB015611

# **TEXAS BIOMEDICAL RESEARCH INSTITUTE**

