Heritability Analysis of imaging traits in the HCP

The following is a step-by-step tutorial on how to perform a classic polygenic and accelerated FPHI heritability analyses in the HCP dataset using SOLAR-Eclipse Imaging Genetics software and 'solareclipser' package can be found on our github page: https://github.com/txbrain-org/solareclipser.

At the moment this procedure has only been tested on linux OS. We are working to evaluate it on MacOS. It may take some little time to have it working on Windows

Start by reviewing your data. The same subject IDs must be used in your phenotype, pedigree and genotype files to avoid confusion and error.
**Note: In the following instructions, '>' denotes a Bash prompt, while '>>' denotes the R software prompt. Commands preceded by '>' are for in the Bash terminal and should not be entered into R.

1. Installation

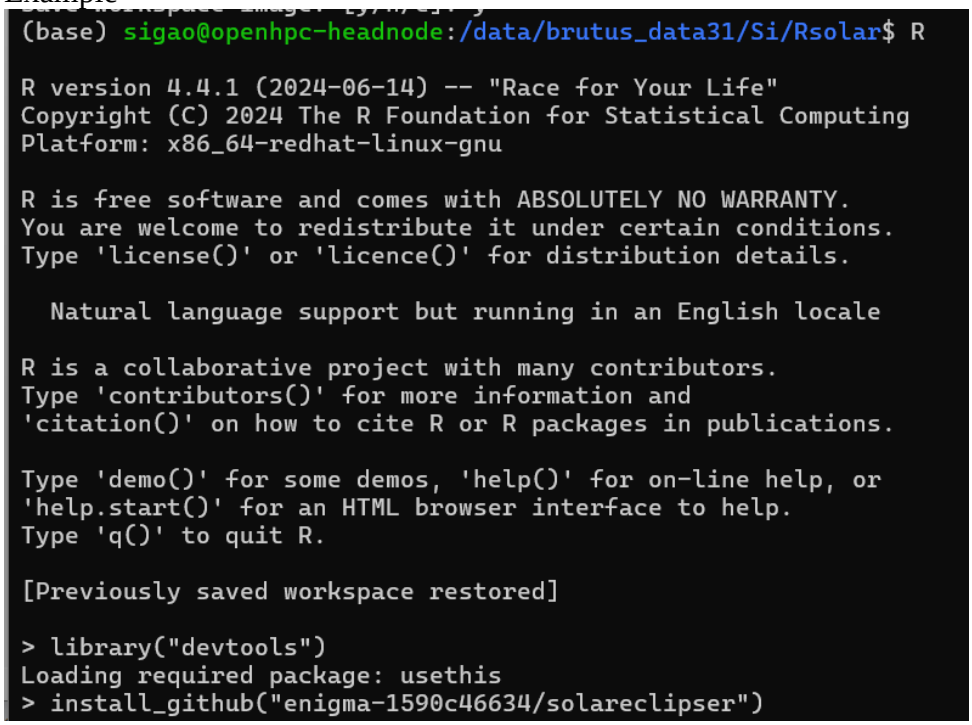
1. Download solar from https://www.nitrc.org/projects/se_linux, then install solar on your cluster and add solar directory to system path.
   > ./install_solar `pwd` `pwd`
   > export PATH=$PATH:"/data/brutus_data31/Si/Rsolar/solar/"
   - Example

```
sigao@node01:/data/brutus_data31/Si/Rsolar/solar$ ls
install_solar  README  solar.tar
sigao@node01:/data/brutus_data31/Si/Rsolar/solar$ ./install_solar `pwd` `pwd`

************************************************************
Ignore error messages (if any) from the tests below

SOLAR Eclipse version 9.0.1, last updated on January 11, 2024
Developed at Maryland Psychiatric Research Center,
University of Maryland School of Medicine, Baltimore.
Visit our documentation and tutorial website www.solar-eclipse-genetics.org
Our download page https://www.nitrc.org/projects/se_linux
Our github page https://github.com/brian09/solar-eclipse
For questions email: pkochunov@gmail.com
Enter help for help, exit to exit, doc to browse documentation.
The software development is supported by NIH grant RO1EB015611
from The National Institute for Biomedical Imaging and Bioengineering.
Enter cite to see how to cite this software.


solar.tcl is version 9.0.1
solarmain binary is version 9.0.1

Mismatched version of solar.tcl in /data/brutus_data31/Si/Rsolar/solar/lib
Ignore error messages (if any) from the tests above
************************************************************


    *** Successful Installation ***

SOLAR has been installed with the command name solar

The new documentation directory is /data/brutus_data31/Si/Rsolar/solar/doc

See README.linux if this version crashes after registering users
or fails to accept keys in parallel queing systems.  You may
need to replace solarmain with solarmain.static or solarmain.dynamic.
We can help if you email solar@txbiomedgenetics.org.

sigao@node01:/data/brutus_data31/Si/Rsolar/solar$ ls
bin  install_solar  lib  README  solar
sigao@node01:/data/brutus_data31/Si/Rsolar/solar$ export PATH=$PATH:"/data/brutus_data31/Si/Rsolar/solar/"
```

2. Install 'solareclipser' package from github, and then install the packages and its dependencies.

> R
>> install.packages("devtools", repos = "https://cran.r-project.org/")
>> library("devtools")
>> install_github("enigma-1590c46634/solareclipser")

- Example



2. Perform polygenic in R

1. Load library
   >> library(solareclipser)
2. Specify output directory
   >> sc <- SolarCommand$new(save_output_dir = "/data/brutus_data31/Si/Rsolar/")
3. Read pedigree file
   >> sc$load(obj = "pedigree", fpath = "HCP_imputed_filtered_ped.csv", cond = "-t 0")
4. Read phenotype file
   >> sc$load(obj = "phenotypes", fpath = "HCP_WM_ave_norm.csv")
5. Select a trait from the phenotype file
   >> sc$trait("CC")$polygenic()
6. Run polygenic
   >> sc$run()
   - Example

```
> library(solareclipser)
> sc <- SolarCommand$new(save_output_dir = "/data/brutus_data31/Si/Rsolar/")
> sc$load(obj = "pedigree", fpath = "HCP_imputed_filtered_ped.csv", cond = "-t 0")
> sc$load(obj = "phenotypes", fpath = "HCP_WM_ave_norm.csv")
> sc$trait("CC")$polygenic()
> sc$run()
------------------------------------------------------------

proc file7bfc46212070 {} {
  load pedigree HCP_imputed_filtered_ped.csv -t 0
  load phenotypes HCP_WM_ave_norm.csv
  trait CC
  polygenic
}

------------------------------------------------------------
Unloading current pedigree data ...
Loading Empirical Pedigree...
There are 2284 people.  If not correct, be sure ID's are unique.
Determining families (use -1 option to skip this)
Warning if number of families is greater than one it may
cause an issue when loading phi2 matrix.  Use -1 option to
avoid any potential issues.
*********************************************************************
*  Maximize sporadic model                                          *
*********************************************************************

    *** Loglikelihood of sporadic model is -487.520945

*********************************************************************
*  Maximize polygenic model                                         *
*********************************************************************

    *** Loglikelihood of polygenic model is -332.229219
    *** H2r in polygenic model is 0.9344430

    *** Determining significance of H2r
    *** Comparing polygenic and sporadic models
    *** chi = 310.5835, deg = 1, p = 8.148806e-70

*****************************************************************************
*                          Summary of Results                               *
*****************************************************************************

        Pedigree:    HCP_imputed_filtered_ped.csv empirical
        Phenotypes:  HCP_WM_ave_norm.csv
        Trait:       CC                    Individuals:  999

                         H2r is 0.9344430  p = 8.148806e-70  (Significant)
               H2r Std. Error:  0.0104121

        Warning.  Unexpectedly high heritabilities might result from
        numerical problems, especially if mztwins are present.


        Output files and models are in directory CC/
        Summary results are in CC/polygenic.out
        Loglikelihoods and chi's are in CC/polygenic.logs.out
        Best model is named poly and null0 (currently loaded)
        Final models are named poly, spor

        Residual Kurtosis is -0.0900, within normal range
>
```

3. Perform Fast and Powerful Heritability Inference (FPHI) in R.

1. Load library
   >> library(solareclipser)
2. Specify output directory
   >> sc <- SolarCommand$new(save_output_dir = "/data/brutus_data31/Si/Rsolar/")
3. Read pedigree file. No need to load pedigree again if pedigree has been loaded
   >> sc$load(obj = "pedigree", fpath = "HCP_imputed_filtered_ped.csv", cond = "-t 0")
4. Read phenotype file
   >> sc$load(obj = "phenotypes", fpath = "HCP_WM_ave_norm.csv")
5. Select a trait from the phenotype file
   >> sc$trait("CC")
6. Create eigenvalue and eigenvector files for the phenotypes
   >> sc$create_evd_data(output_fbasename = "CC_evd")
7. Create fphi with evd data
   >> sc$fphi(evd_data = "CC_evd")
8. Run FPHI
   >> sc$run()
   - Example

```
------------------------------------------------------------
> library(solareclipser)
> sc <- SolarCommand$new(save_output_dir = "/data/brutus_data31/Si/Rsolar/")
> sc$load(obj = "phenotypes", fpath = "HCP_WM_ave_norm.csv")
> sc$trait("CC")
> sc$create_evd_data(output_fbasename = "CC_evd")
> sc$fphi(evd_data = "CC_evd")
> sc$run()
------------------------------------------------------------

proc file7bfc764c31a7 {} {
  load phenotypes HCP_WM_ave_norm.csv
  trait CC
  create_evd_data --o CC_evd
  fphi --evd_data CC_evd
}

------------------------------------------------------------
********************************************************************************
*        Fast Permutation Heritability Inference (FPHI) Summary of Results     *
********************************************************************************
  Pedigree:    HCP_imputed_filtered_ped.csv
  Phenotypes:  HCP_WM_ave_norm.csv
  Trait:   CC H2r = 0.934443010898 SE = 0.0104086242211  Individuals:  999
  polygenic loglik:   -332.22921885  sporadic loglik:   -487.520945344  p = 8.14880208133e-70 (Significant)
Fully Converged Parameters

      Name               Value           Std Error


 Parameter           Fit Value      Standard Error
      mean    0.00404827444118      0.118458228889
        e2     0.0655569891024     0.0104086242211
       h2r      0.934443010898     0.0104086242211
        sd       1.04560390379     0.0260963338527

>
```