

Improving genomic prediction using high-dimensional secondary phenotypes: The genetic latent factor approach

This repository contains all scripts required to generate the datafiles for all replications for the hyperspectral B5IR/HEAT data as well as the random (p800) and low-rank (p800_lowrank) residual structure simulations. This repository also contains scripts to analyze these datasets using different methods and visualize the results. These scripts can be used to reproduce all intermediate results, i.e., every single accuracy for each model and replication. Finally, it contains instructions and examples that can be used to perform reproducibility spot checks that allow anyone interested to check reproducibility without having to run all analyses.

This readme contains the following sections:

- Description of top level items
- Reproducibility instructions
 - A note on reproducibility for MegaLMM
- Hyperspectral B5IR (**hyper_1415B5IR**)
 - Data generation
 - Analyses
 - Plotting
 - Single date analyses and plotting
- Hyperspectral HEAT (**hyper_1415HEAT**)
 - Data generation
 - Analyses
 - Plotting
- Simulations using random residual structure (**p800**)
 - Data generation
 - Analyses
 - Plotting
- Simulations using low-rank residual structure (**p800_lowrank**)
 - Data generation
 - Analyses
 - Plotting
- Timing (**timing**)
- List of manuscript figures
- sessionInfo() output

Note that a copy of this repository is also available at

https://github.com/KillianMelsen/glfBLUP_2025. The github repository contains all intermediate results, but excludes the datafiles. The datafiles can be generated by following the instructions in the **Reproducibility instructions** section. Note that the github repository also excludes the

intermediate results produced by MegaLMM for the single date analyses. These results can be produced using the `masterscript_hyper.R` masterscript.

Description of top level items

Folders

- `.Rproj.user` this folder contains R project files. Can be ignored.
- `genotypes` this folder contains marker data and kinship matrices for all analyses.
- `hyper_1415B5IR` this folder contains all files related to the hyperspectral B5IR data and analyses. The folder itself contains a [subfolder for the analysis scripts](#), a [subfolder for the single date analysis scripts](#), a [subfolder for scripts related to datafile generation](#), a [subfolder to store the generated datasets](#), a [subfolder that stores posterior arrays for the MegaLMM single date analyses](#), a [subfolder for temporary files that are generated by MegaLMM](#), and a [subfolder that stores the intermediate results](#).
- `hyper_1415HEAT` this folder contains all files related to the hyperspectral HEAT data and analyses. The folder itself contains a [subfolder for the analysis scripts](#), a [subfolder for scripts related to datafile generation](#), a [subfolder to store the generated datasets](#), a [subfolder for temporary files that are generated by MegaLMM](#), and a [subfolder that stores the intermediate results](#).
- `hyper_datafiles` this folder contains the raw data for the hyperspectral dataset.
- `logs` this folder is used for log files from scripts that make use of parallelization.
- `MegaLMM_rep_check_scripts` this folder contains all files related to the MegaLMM reproducibility checks. The folder itself contains a [subfolder for the hyperspectral B5IR reproducibility checks](#), another [subfolder for the hyperspectral HEAT reproducibility checks](#), a [subfolder for the reproducibility checks of the simulations with random residual structure](#), and a [subfolder for the reproducibility checks of the simulations with low-rank residual structure](#).
- `output` this folder is used for output from SLURM jobs.
- `p800` this folder contains all files related to the simulations using a random residual structure. There are subfolders for [analyses scripts](#), [data generation](#), [the generated datasets](#), [temporary MegaLMM files](#), [miscellaneous scripts \(merging intermediate results\)](#), and [the intermediate results themselves](#).
- `p800_lowrank` this folder contains all files related to the simulations using a random residual structure. There are again subfolders for [analyses scripts](#), [data generation](#), [the generated datasets](#), [temporary MegaLMM files](#), [miscellaneous scripts \(merging intermediate results\)](#), and [the intermediate results themselves](#).
- `plots` this folder contains all manuscript figures produced by the scripts.
- `rep_check_scripts` this folder contains ten annotated example scripts for reproducibility spot checks.
- `SLURM` this folder contains files for slurm jobs.
- `timing` this folder contains all files used for measuring runtimes of glfBLUP for different numbers of features. There is a single [subfolder](#) that stores the datasets of different sizes used for timing.

Scripts

- `helper_functions.R` this script contains some helper functions with comments explaining the use and arguments/output of each function.
- `masterscript_hyper.R` this is a masterscript that allows anyone to conveniently reproduce all intermediate results and figures related to the hyperspectral data.
- `masterscript_MegaLMM_rep_check.R` this is a masterscript that allows anyone to conveniently produce a second, independent set of MegaLMM intermediate results. Mean accuracies can then be compared to the mean accuracies reported in the paper to confirm that the impact of Monte Carlo errors is negligible.
- `masterscript_misc.R` this is a masterscript that allows anyone to conveniently run some scripts that produce two figures related to redundancy filtering.
- `masterscript_p800.R` this is a masterscript that allows anyone to conveniently reproduce all intermediate results and figures related to the simulated data with random residual structure and $p = 800$ secondary features.
- `masterscript_p800_lowrank.R` this is a masterscript that allows anyone to conveniently reproduce all intermediate results and figures related to the simulated data with low-rank residual structure and $p = 800$ secondary features.
- `masterscript_rep_check_scripts.R` this is a masterscript that allows anyone to conveniently run all the annotated example scripts for reproducibility spot checks.
- `masterscript_timing.R` this is a masterscript that allows anyone to conveniently reproduce the timing measurements for glfBLUP.
- `plot_hyper.R` this is a script that is used to produce a figure combining intermediate results from the hyperspectral B5IR and HEAT datasets.

Other

- `.gitignore` this is a .gitignore file. Can be ignored.
- `.Rhistory` this is a .Rhistory file. Can be ignored.
- `glfBLUPold_1.3.1.tar.gz` this is an old version of the glfBLUP package required to run lsBLUP and siBLUP.
- `glfBLUP_1.0.0.tar.gz` this is the version of the glfBLUP package used in the glfBLUP analyses.
- `glfBLUP_2025.Rproj` this is the R project file. Opening this before any scripts ensures the working directory is set correctly. Otherwise, make sure that the working directory is set to the root of this folder before running any scripts.
- `README.md` this is the readme file that you are currently reading.

Reproducibility instructions

In principle, every single intermediate result present in this repository can be easily reproduced by running the seven masterscripts described above. These scripts ensure that all required datafiles are generated, training and test sets are sampled for all replications, and all analyses are run. These masterscripts also ensure that every single figure that is present in the manuscript is produced and saved in the `plots` folder. To ensure everything working correctly, set the working directory to the root of this folder, or simply open the `glfBLUP_2025.Rproj` file before opening any scripts.

Running every single script in full will take a significant amount of time, however. As a result, one may be interested in performing reproducibility spot checks, that is, confirming that the accuracies

for a specific model for specific replications of the simulations or hyperspectral data analyses are reproducible. To do this, it is **important** to make sure that all datafiles are available. All required datafiles can be generated by running the following scripts (in the presented order) if they are not yet present in the repository:

1. `hyper_1415B5IR/data_generation/hyper_preprocessing.R`
2. `hyper_1415B5IR/data_generation/generate_hyper_datasets.R`
3. `hyper_1415HEAT/data_generation/hyper_preprocessing.R`
4. `hyper_1415HEAT/data_generation/generate_hyper_datasets.R`
5. `p800/data_generation/generate_sim_p800_datasets.R`
6. `p800/data_generation/traintest_sim_p800_datasets.R`
7. `p800_lowrank/data_generation/generate_sim_p800_lowrank_datasets.R`
8. `p800_lowrank/data_generation/traintest_sim_p800_lowrank_datasets.R`

After all datafiles have been generated, the reproducibility spot checks can be performed. To do so, we have included a comment at the top of every single script in this repository. These comments contain information on what the script does, what figures it produces (if any), and instructions specific to that script that can be used to perform reproducibility spot checks. For example, the `p800/analyses/glfBLUP.R` script contains a comment with instructions to reproduce the glfBLUP CV1 and CV2 intermediate results for replication 9, a secondary feature heritability of 0.5, focal trait heritability of 0.7, and communality of 0.5. To make things easier, we have also included some annotated example scripts for reproducibility spot checks in the `rep_check_scripts` folder. For example, the `rep_check_scripts/rep_check_glfBLUP_p800.R` script is a copy of the `p800/analyses/glfBLUP.R` script mentioned before, with code commented out in such a way that the reproducibility spot check can be easily performed. All example reproducibility spot checks can be conveniently run using the `masterscript_rep_check_scripts.R` masterscript. Feel free to use the comments in all analysis scripts and the examples to perform your own spot checks.

A note on reproducibility for MegaLMM

Intermediate results for MegaLMM are not perfectly reproducible, even when setting a seed in R. This is due to the MCMC procedures used by MegaLMM and thus specific to the software. This is unfortunately outside of our control. We have included scripts that produce a complete, second, independent set of intermediate results in the `MegaLMM_rep_check_scripts` folder. This second set of intermediate results is stored in that folder as well. By comparing these intermediate results to the ones used to generate the figures for the manuscript, we show that the impact of Monte Carlo errors on the presented results (mean prediction accuracies in terms of correlations) is negligible. This can be confirmed by anyone by simply running the **Reproducibility checks** section of the `masterscript_MegaLMM_rep_check.R` masterscript. If you want to compare a completely new set of intermediate results, simply run the entire masterscript mentioned before.

Hyperspectral B5IR (`hyper_1415B5IR`)

Data generation

There are two scripts to generate all required datafiles. The `hyper_1415B5IR/data_generation/hyper_preprocessing.R` script takes the raw hyperspectral

B5IR data and pre-processes it. The

`hyper_1415B5IR/data_generation/generate_hyper_datasets.R` then generates the actual datafiles and samples training and test sets for the 250 replications.

Analyses

There are 13 scripts that perform the analyses on the hyperspectral B5IR data for the different models:

- `hyper_1415B5IR/analyses/glfBLUP_hyper.R` performs the CV1 and CV2 analyses for all 250 replications using glfBLUP.
- `hyper_1415B5IR/analyses/glfBLUP_hyper_CV2VEG.R` performs the CV2VEG analyses for all 250 replications using glfBLUP.
- `hyper_1415B5IR/analyses/lslBLUP_hyper.R` performs the CV1 and CV2 analyses for all 250 replications using lslBLUP.
- `hyper_1415B5IR/analyses/lslBLUP_hyper_CV2VEG.R` performs the CV2VEG analyses for all 250 replications using lslBLUP.
- `hyper_1415B5IR/analyses/MegaLMM_hyper_CV1.R` performs the CV1 analyses for all 250 replications using MegaLMM.
- `hyper_1415B5IR/analyses/MegaLMM_hyper_CV2.R` performs the CV2 analyses for all 250 replications using MegaLMM.
- `hyper_1415B5IR/analyses/MegaLMM_hyper_CV2VEG.R` performs the CV2VEG analyses for all 250 replications using MegaLMM.
- `hyper_1415B5IR/analyses/multiMLP_hyper_CV1.R` performs the CV1 analyses for all 250 replications using multiMLP.
- `hyper_1415B5IR/analyses/multiMLP_hyper_CV2.R` performs the CV2 analyses for all 250 replications using multiMLP.
- `hyper_1415B5IR/analyses/multiMLP_hyper_CV2VEG.R` performs the CV2VEG analyses for all 250 replications using multiMLP.
- `hyper_1415B5IR/analyses/siBLUP_hyper.R` performs the CV1 and CV2 analyses for all 250 replications using siBLUP.
- `hyper_1415B5IR/analyses/siBLUP_hyper_CV2VEG.R` performs the CV2VEG analyses for all 250 replications using siBLUP.
- `hyper_1415B5IR/analyses/univariate_hyper.R` performs the analyses for all 250 replications using the univariate model.

Note again that all of the above scripts contain instructions for reproducibility spot checks.

Intermediate results produced by the above scripts are saved in the `hyper_1415B5IR/results` folder.

Plotting

The hyperspectral B5IR intermediate results produced by the scripts above are used (together with the hyperspectral HEAT intermediate results) by the `plot_hyper.R` script to produce figure 3 of the main text.

Single date analyses and plotting

There are five scripts in the `hyper_1415B5IR/analyses_single_date` folder that are used to produce several figures of the main text and supplementary material:

- `hyper_1415B5IR/analyses_single_date/glfBLUP_hyper_single_date.R` produces figure 4 of the main text.
- `hyper_1415B5IR/analyses_single_date/lSBLUP_hyper_single_date.R` produces figure S1A of the supplementary material.
- `hyper_1415B5IR/analyses_single_date/MegaLMM_hyper_single_date.R` produces some intermediate results that are used by the script described next. Note that these intermediate results are not included in the github repository due to their file size. They will have to be generated by anyone interested in reproducing the figures described below. Note that the figures will not be perfectly reproducible anyway due to the Monte Carlo errors.
- `hyper_1415B5IR/analyses_single_date/MegaLMM_hyper_single_date_plotting.R` produces the following figures of the supplementary material:
 - Figure S2B
 - Figure S2C
 - Figure S2A
 - Figure S7B
 - Figure S4
 - Figure S7C
 - Figure S5
 - Figure S6
 - Figure S7A
 - Figure S3
- `hyper_1415B5IR/analyses_single_date/siBLUP_hyper_single_date.R` produces figure S1B of the supplementary material.

Note again that the datafiles for all datasets must be present in the repository to reproduce these figures (see the Reproducibility section of this readme)!

Hyperspectral HEAT (`hyper_1415HEAT`)

Data generation

There are two scripts to generate all required datafiles. The `hyper_1415HEAT/data_generation/hyper_preprocessing.R` script takes the raw hyperspectral HEAT data and pre-processes it. The `hyper_1415HEAT/data_generation/generate_hyper_datasets.R` then generates the actual datafiles and samples training and test sets for the 250 replications.

Analyses

There are 13 scripts that perform the analyses on the hyperspectral HEAT data for the different models:

- `hyper_1415HEAT/analyses/glfBLUP_hyper.R` performs the CV1 and CV2 analyses for all 250 replications using glfBLUP.

- `hyper_1415HEAT/analyses/glfBLUP_hyper_CV2VEG.R` performs the CV2VEG analyses for all 250 replications using glfBLUP.
- `hyper_1415HEAT/analyses/lsBLUP_hyper.R` performs the CV1 and CV2 analyses for all 250 replications using lsBLUP.
- `hyper_1415HEAT/analyses/lsBLUP_hyper_CV2VEG.R` performs the CV2VEG analyses for all 250 replications using lsBLUP.
- `hyper_1415HEAT/analyses/MegaLMM_hyper_CV1.R` performs the CV1 analyses for all 250 replications using MegaLMM.
- `hyper_1415HEAT/analyses/MegaLMM_hyper_CV2.R` performs the CV2 analyses for all 250 replications using MegaLMM.
- `hyper_1415HEAT/analyses/MegaLMM_hyper_CV2VEG.R` performs the CV2VEG analyses for all 250 replications using MegaLMM.
- `hyper_1415HEAT/analyses/multiMLP_hyper_CV1.R` performs the CV1 analyses for all 250 replications using multiMLP.
- `hyper_1415HEAT/analyses/multiMLP_hyper_CV2.R` performs the CV2 analyses for all 250 replications using multiMLP.
- `hyper_1415HEAT/analyses/multiMLP_hyper_CV2VEG.R` performs the CV2VEG analyses for all 250 replications using multiMLP.
- `hyper_1415HEAT/analyses/siBLUP_hyper.R` performs the CV1 and CV2 analyses for all 250 replications using siBLUP.
- `hyper_1415HEAT/analyses/siBLUP_hyper_CV2VEG.R` performs the CV2VEG analyses for all 250 replications using siBLUP.
- `hyper_1415HEAT/analyses/univariate_hyper.R` performs the analyses for all 250 replications using the univariate model.

Plotting

The hyperspectral HEAT intermediate results produced by the scripts above are used (together with the hyperspectral B5IR intermediate results) by the `plot_hyper.R` script to produce figure 3 of the main text.

Simulations using random residual structure (p800)

Data generation

There are two scripts that generate all datafiles. The `p800/data_generation/generate_sim_p800_datasets.R` script simulates the 4500 datafiles for all combinations of secondary feature heritability (3 values), focal trait heritability (5 values), and communality (3 values). There are 100 replications for each combination ($3 \times 5 \times 3 \times 100 = 4500$). The `p800/data_generation/generate_sim_p800_datasets.R` script then samples training and test sets for all these datafiles. The produced datafiles are stored in the `p800/datasets` folder.

Analyses

There are 15 scripts that perform the analyses on the simulated datasets with random residual structure:

- `p800/analyses/benchmark.R` performs the CV1 and CV2 benchmark analyses for all 45 combinations and 100 replications.
- `p800/analyses/CV1_multiMLP_part1.R` performs the CV1 multiMLP analyses for all 45 combinations and the first 10 replications.
- `p800/analyses/CV1_multiMLP_part2.R` performs the CV1 multiMLP analyses for all 45 combinations and the second 10 replications (11 - 20).
- `p800/analyses/CV2_multiMLP_part1.R` performs the CV2 multiMLP analyses for all 45 combinations and the first 10 replications.
- `p800/analyses/CV2_multiMLP_part2.R` performs the CV2 multiMLP analyses for all 45 combinations and the second 10 replications (11 - 20).
- `p800/analyses/glfBLUP.R` performs the CV1 and CV2 glfBLUP analyses for all 45 combinations and 100 replications.
- `p800/analyses/lisBLUP.R` performs the CV1 and CV2 lisBLUP analyses for all 45 combinations and 100 replications.
- `p800/analyses/MegaLMM_part1.R` performs the CV1 and CV2 MegaLMM analyses for all 45 combinations and the first 4 replications.
- `p800/analyses/MegaLMM_part2.R` performs the CV1 and CV2 MegaLMM analyses for all 45 combinations and the second 4 replications (5 - 8).
- `p800/analyses/MegaLMM_part3.R` performs the CV1 and CV2 MegaLMM analyses for all 45 combinations and the third 4 replications (9 - 12).
- `p800/analyses/MegaLMM_part4.R` performs the CV1 and CV2 MegaLMM analyses for all 45 combinations and the fourth 4 replications (13 - 16).
- `p800/analyses/MegaLMM_part5.R` performs the CV1 and CV2 MegaLMM analyses for all 45 combinations and the fifth 4 replications (17 - 20).
- `p800/analyses/siBLUP_part1.R` performs the CV1 and CV2 siBLUP analyses for all 45 combinations and the first 50 replications.
- `p800/analyses/siBLUP_part2.R` performs the CV1 and CV2 siBLUP analyses for all 45 combinations and the second 50 replications (51 - 100).
- `p800/analyses/univariate.R` performs the univariate analyses for all 45 combinations and 100 replications.

Note that the multiMLP, MegaLMM, and siBLUP analyses were split in several parts to make runtimes of individual scripts a little bit more manageable. After running the above scripts, results for these three models can be merged with `p800/misc/CV1_multiMLP_merge.R`, `p800/misc/CV2_multiMLP_merge.R`, `p800/misc/MegaLMM_merge.R`, and `p800/misc/siBLUP_merge.R`.

All intermediate results produced by the above scripts are stored in the `p800/results` folder.

Plotting

The `p800/plot_p800_results.R` script produces figures 2 and 5 of the main text.

Simulations using low-rank residual structure (`p800_lowrank`)

Data generation

There are two scripts that generate all datafiles. The

`p800_lowrank/data_generation/generate_sim_p800_lowrank_datasets.R` script simulates the 4500 datafiles for all combinations of secondary feature heritability (3 values), focal trait heritability (5 values), and communality (3 values). There are 100 replications for each combination ($3 \times 5 \times 3 \times 100 = 4500$). The

`p800_lowrank/data_generation/generate_sim_p800_lowrank_datasets.R` script then samples training and test sets for all these datafiles. The produced datafiles are stored in the `p800_lowrank/datasets` folder.

Analyses

There are 15 scripts that perform the analyses on the simulated datasets with random residual structure:

- `p800_lowrank/analyses/benchmark.R` performs the CV1 and CV2 benchmark analyses for all 45 combinations and 100 replications.
- `p800_lowrank/analyses/CV1_multiMLP_part1.R` performs the CV1 multiMLP analyses for all 45 combinations and the first 10 replications.
- `p800_lowrank/analyses/CV1_multiMLP_part2.R` performs the CV1 multiMLP analyses for all 45 combinations and the second 10 replications (11 - 20).
- `p800_lowrank/analyses/CV2_multiMLP_part1.R` performs the CV2 multiMLP analyses for all 45 combinations and the first 10 replications.
- `p800_lowrank/analyses/CV2_multiMLP_part2.R` performs the CV2 multiMLP analyses for all 45 combinations and the second 10 replications (11 - 20).
- `p800_lowrank/analyses/glfBLUP.R` performs the CV1 and CV2 glfBLUP analyses for all 45 combinations and 100 replications.
- `p800_lowrank/analyses/lisBLUP.R` performs the CV1 and CV2 lisBLUP analyses for all 45 combinations and 100 replications.
- `p800_lowrank/analyses/MegaLMM_part1.R` performs the CV1 and CV2 MegaLMM analyses for all 45 combinations and the first 4 replications.
- `p800_lowrank/analyses/MegaLMM_part2.R` performs the CV1 and CV2 MegaLMM analyses for all 45 combinations and the second 4 replications (5 - 8).
- `p800_lowrank/analyses/MegaLMM_part3.R` performs the CV1 and CV2 MegaLMM analyses for all 45 combinations and the third 4 replications (9 - 12).
- `p800_lowrank/analyses/MegaLMM_part4.R` performs the CV1 and CV2 MegaLMM analyses for all 45 combinations and the fourth 4 replications (13 - 16).
- `p800_lowrank/analyses/MegaLMM_part5.R` performs the CV1 and CV2 MegaLMM analyses for all 45 combinations and the fifth 4 replications (17 - 20).
- `p800_lowrank/analyses/siBLUP_part1.R` performs the CV1 and CV2 siBLUP analyses for all 45 combinations and the first 50 replications.
- `p800_lowrank/analyses/siBLUP_part2.R` performs the CV1 and CV2 siBLUP analyses for all 45 combinations and the second 50 replications (51 - 100).
- `p800_lowrank/analyses/univariate.R` performs the univariate analyses for all 45 combinations and 100 replications.

Note that the multiMLP, MegaLMM, and siBLUP analyses were split in several parts to make runtimes of individual scripts a little bit more manageable. After running the above scripts, results

for these three models can be merged with `p800_lowrank/misc/CV1_multiMLP_merge.R`, `p800_lowrank/misc/CV2_multiMLP_merge.R`, `p800_lowrank/misc/MegaLMM_merge.R`, and `p800_lowrank/misc/siBLUP_merge.R`.

All intermediate results produced by the above scripts are stored in the `p800_lowrank/results` folder.

Plotting

The `p800_lowrank/plot_p800_lowrank_results.R` script produces figure 6 of the main text.

Timing (**timing**)

There are four scripts that together produce the glfBLUP timing figure (S8 of the supplementary material):

1. The `timing/generate_timing_data.R` script generates simulated datasets with random residual structures and different numbers of secondary features.
2. The `timing/traintest_timing_data.R` script samples training and test sets for the generated datasets.
3. The `timing/timing.R` performs the actual timing of the different components of the glfBLUP method and stores the results in the `timing` folder.
4. The `timing/plotting.R` then uses those results to produce figure S8 of the supplementary material.

List of manuscript figures

- Figure 2 of the main text is produced by `p800/plot_p800_results.R`.
- Figure 3 of the main text is produced by `plot_hyper.R`.
- Figure 4 of the main text is produced by `hyper_1415B5IR/analyses_single_date/glfBLUP_hyper_single_date.R`.
- Figure 5 of the main text is produced by `p800/plot_p800_results.R`.
- Figure 6 of the main text is produced by `p800_lowrank/plot_p800_lowrank_results.R`.
- Figure S1A of the supplementary material is produced by `hyper_1415B5IR/analyses_single_date/siBLUP_hyper_single_date.R`.
- Figure S1B of the supplementary material is produced by `hyper_1415B5IR/analyses_single_date/siBLUP_hyper_single_date.R`.
- Figure S2A of the supplementary material is produced by `hyper_1415B5IR/analyses_single_date/MegaLMM_hyper_single_date_plotting.R`.
- Figure S2B of the supplementary material is produced by `hyper_1415B5IR/analyses_single_date/MegaLMM_hyper_single_date_plotting.R`.
- Figure S2C of the supplementary material is produced by `hyper_1415B5IR/analyses_single_date/MegaLMM_hyper_single_date_plotting.R`.
- Figure S3 of the supplementary material is produced by `hyper_1415B5IR/analyses_single_date/MegaLMM_hyper_single_date_plotting.R`.

- Figure S4 of the supplementary material is produced by `hyper_1415B5IR/analyses_single_date/MegaLMM_hyper_single_date_plotting.R`.
- Figure S5 of the supplementary material is produced by `hyper_1415B5IR/analyses_single_date/MegaLMM_hyper_single_date_plotting.R`.
- Figure S6 of the supplementary material is produced by `hyper_1415B5IR/analyses_single_date/MegaLMM_hyper_single_date_plotting.R`.
- Figure S7 of the supplementary material is produced by `hyper_1415B5IR/analyses_single_date/MegaLMM_hyper_single_date_plotting.R`.
- Figure S8 of the supplementary material is produced by `timing/plotting.R`.
- Figure S9 of the supplementary material is produced by `hyper_1415B5IR/plot_RF.R`.
- Figure S10 of the supplementary material is produced by `hyper_1415HEAT/plot_RF.R`.

sessionInfo() output

```

R version 4.4.1 (2024-06-14 ucrt)
Platform: x86_64-w64-mingw32/x64
Running under: Windows 11 x64 (build 26100)

Matrix products: default

locale:
[1] LC_COLLATE=English_Netherlands.utf8  LC_CTYPE=English_Netherlands.utf8
LC_MONETARY=English_Netherlands.utf8 LC_NUMERIC=C
[5] LC_TIME=English_Netherlands.utf8

time zone: Europe/Amsterdam
tzcode source: internal

attached base packages:
[1] grid      parallel  stats      graphics  grDevices  utils      datasets
methods    base

other attached packages:
[1] ggforce_0.4.2      statgenGWAS_1.0.9  statgenHTP_1.0.6.1
LMMsolver_1.0.8     lubridate_1.9.3    forcats_1.0.0      stringr_1.5.1
[8] dplyr_1.1.4        purrr_1.0.2        readr_2.1.5        tidyr_1.3.1
tibble_3.2.1        tidyverse_2.0.0    ggplot2_3.5.1
[15] rrBLUP_4.6.3       tensorflow_2.16.0  keras_2.15.0       MegaLMM_0.9.5
MCMCglmm_2.36       ape_5.8-1          coda_0.19-4.1
[22] Matrix_1.7-0       glfBLUP_1.0.0      gfBLUPold_1.3.1
doParallel_1.0.17   iterators_1.0.14   foreach_1.5.2      tictoc_1.2.1
[29] rlist_0.4.6.2

loaded via a namespace (and not attached):
[1] dotCall64_1.1-1    spam_2.10-0        gtable_0.3.5        tensorA_0.36.2.1
lattice_0.22-6      tzdb_0.4.0          vctrs_0.6.5         tools_4.4.1
[9] tfruns_1.5.3        generics_0.1.3      fansi_1.0.6          SpATS_1.0-18
pkgconfig_2.0.3     data.table_1.16.0  lifecycle_1.0.4     cubature_2.1.1
[17] farver_2.1.2        compiler_4.4.1      munsell_0.5.1        codetools_0.2-20
pillar_1.9.0        nloptr_2.1.1        whisker_0.4.1        MASS_7.3-61
[25] boot_1.3-31         abind_1.4-8         nlme_3.1-166         tidyselect_1.2.1
digest_0.6.37       stringi_1.8.4       reshape2_1.4.4       splines_4.4.1
[33] polyclip_1.10-7     rprojroot_2.0.4     here_1.0.1           colorspace_2.1-1
cli_3.6.3           magrittr_2.0.3      base64enc_0.1-3      utf8_1.2.4
[41] corpcor_1.6.10      withr_3.0.1         scales_1.3.0         rappdirs_0.3.3
timechange_0.3.0    lme4_1.1-35.5       reticulate_1.39.0    hms_1.1.3
[49] png_0.1-8           rlang_1.1.4         Rcpp_1.0.13          zeallot_0.1.0
glue_1.7.0          tweenr_2.0.3         rstudioapi_0.16.0    minqa_1.2.8
[57] jsonlite_1.8.9      R6_2.5.1            plyr_1.8.9

```