

The type of heritability you're calculating in the provided R script is **narrow-sense heritability** (h^2). Narrow-sense heritability estimates the proportion of phenotypic variance explained by the additive genetic variance.

Key points confirming this:

1. Phenotypic variance breakdown:

- The script decomposes the phenotypic variance into additive genetic variance (`varU`) and residual/environmental variance (`varE`).
- Heritability is calculated as $h^2 = \frac{\sigma^2_G}{\sigma^2_G + \sigma^2_E}$.

2. Bayesian Linear Regression (BGLR):

- The model used is "`BayesA`", a Bayesian method that estimates genetic variance components and fits marker effects.
- The `BGLR` package is commonly used for genomic prediction and heritability estimation from genomic data.

Issues and Recommendations:

1. Variance Component File (`varUFilename`):

- The script references "`ETA_1_ScaleBayesA.dat`" for `varU`, which may not correspond directly to the additive genetic variance (`varB`).
- Consider verifying the contents of "`ETA_1_ScaleBayesA.dat`" and if it represents the variance of marker effects.
- The commented-out line (`varUFilename = paste0(PREFIX, "ETA_1_varB.dat")`) seems more appropriate for additive genetic variance.

2. Scaling and Standardization:

The genotype matrix `geno` is scaled using `scale(geno)/sqrt(ncol(geno))`. This standardization is typical but ensure it aligns with your experimental design.

3. Iterations and Burn-in:

`nIter=5000` and `burnIn=1000` are standard but could be adjusted depending on data size and convergence diagnostics.

Conclusion:

The procedure calculates narrow-sense heritability but double-check the variance component file (`varUFilename`) for accuracy.

But, for step : "1. Variance Component File (varUFilename):" I don't have "ETA_1_varB.dat" values



If "ETA_1_varB.dat" is unavailable, but "ETA_1_ScaleBayesA.dat" is being used instead, you might be estimating a transformed or scaled version of the genetic variance specific to the "BayesA" model rather than the direct additive genetic variance.

Clarification of "ETA_1_ScaleBayesA.dat" in BayesA:

- "ETA_1_ScaleBayesA.dat" stores the posterior scale parameter for the variance of marker effects under the BayesA model.
- BayesA assumes a heavy-tailed prior on marker effects, which can inflate the variance. Therefore, this file may not directly represent the additive genetic variance (`varU`).

How to Properly Estimate Additive Genetic Variance in BayesA:

1. If "ETA_1_varB.dat" is missing, but you have "ETA_1_ScaleBayesA.dat", the latter provides a scaled variance parameter.
2. The variance of marker effects in BayesA can be recovered from the scale parameter, but this requires using the degrees of freedom and scale parameter together.
3. If you can't extract `varB`, an alternative is using "BayesC" or "RKHS" models where variance components are more directly available.

Suggested Fix:

- If "ETA_1_varB.dat" is unavailable, you could still use "ETA_1_ScaleBayesA.dat" cautiously but note it may overestimate variance.
- Consider verifying the output files generated by the BGLR model run (`fm`) for any direct variance components.

Would you like help modifying the script to ensure a more direct narrow-sense heritability calculation?

Then give a better approach to calculate the heretability



To calculate **narrow-sense heritability** more accurately using the **BGLR** package in R, a more standard approach involves estimating variance components directly. Here's a corrected version using the **"RKHS"** (Reproducing Kernel Hilbert Space) model, which estimates genetic variance more directly:

R

```
calculateHeritabilityTrait <- function(traitName, geno, pheno, outputDir) {
  message(">>> Calculating Heritability for ", traitName)

  # Prepare data
  X = scale(geno) / sqrt(ncol(geno)) # Standardize genotypes
  Y = pheno[, traitName]             # Extract phenotype vector

  # Set model and output prefix
  MODEL = "RKHS"
  PREFIX = paste0(outputDir, "/tmp/", MODEL, "-", traitName, "-")

  # Create kernel matrix for genetic relationship
  K = tcrossprod(X) / ncol(X) # Centered GRM (Genomic Relationship
Matrix)

  # Fit the model using BGLR
  eta = list(list(K=K, model="RKHS"))
  fm = BGLR(y=Y, ETA=eta, nIter=5000, burnIn=1000, verbose=FALSE,
saveAt=PREFIX)

  # Extract variance components directly from the model object
  varU = fm$ETA[[1]]$varU # Genetic variance
  varE = fm$varE          # Residual variance

  # Calculate narrow-sense heritability
  h2 = varU / (varU + varE)

  # Return heritability estimate
  return(h2)
}
```

Key Changes and Improvements:

1. Kernel-Based Approach (RKHS):

- The "RKHS" model fits a genomic relationship matrix (GRM) instead of per-marker variances.
- This model directly estimates additive genetic variance (`varU`).

2. Variance Extraction:

- The variance components are directly retrieved from the fitted model object (`fm$ETA[[1]]$varU` and `fm$varE`).
- No need to read files externally.

3. Genomic Relationship Matrix (GRM):

The matrix `K` is constructed as `tcrossprod(X) / ncol(X)` for a standardized GRM.

✓ Why It's Better:

- Avoids file dependency issues.
- Direct extraction of variance components.
- More robust and commonly used model (RKHS).

Let me know if you need further adjustments for other types of heritability!

ChatGPT can make mistakes. Check important info.