

easyGWAS: A cloud-based platform for comparing the results of genome-wide association studies

Dominik G. Grimm, Damian Roqueiro, Matteo Togninalli, easyGWAS Consortium and Karsten Borgwardt.

Department of Biosystems Science and Engineering, ETH Zürich, Basel, Switzerland

Introduction

Genome-wide association studies (GWAS) are a powerful tool to identify genomic regions associated to a trait in a group of individuals. To achieve this goal, GWAS exploit the abundance of high-quality genotypes and phenotypes that are continuously becoming available for a diversity of plant and animal species.

Notable examples of association studies in plants and crops include *Arabidopsis thaliana* (Atwell et al., 2010, The 1001 Genomes Consortium, 2016), rice (Zhao et al., 2011), and tomatoes (Lin et al., 2014). Examples of GWAS in animal species include fruit flies (Mackay et al., 2012), mice (Kirby et al., 2010), and humans (Scott et al., 2007; Freilinger et al., 2012).

All these independent instances of successful GWAS, many of which are conducted on related traits, provide an unprecedented wealth of information and open up new and unique possibilities for the systematic comparison of GWAS results. These GWAS results not only make it possible to detect genes with pleiotropic effects on multiple traits, but also to replicate GWAS and to increase statistical power to detect new signals through meta-analysis of several studies.

Presentation

We will present easyGWAS (Grimm et al., 2017), a cloud-based platform to perform, store, share, and compare the results of GWAS in a wide variety of both inbred and outbred plant and animal species. The three main aspects of easyGWAS that make the platform a unique online resource are: i) serving as an online repository of GWAS results, either obtained from external tools and imported into easyGWAS, or computed directly on the platform; ii) functionality to conduct GWAS with a variety of statistical methods in a standardized way to ensure a maximum degree of comparability between studies, which leads to the third key aspect iii) to compare the results of different GWAS and to conduct meta-analyses.

To conduct a genome-wide association study, meta-analysis, or comparison, easyGWAS guides the user through step-by-step procedures—or *wizards*—to select a species, phenotypes, genome build, and all appropriate algorithms.

To facilitate the interpretation of conducted analyses, easyGWAS offers on-the-fly annotations of associated hits, as well as interactive visualizations, including zoomable Manhattan plots, linkage disequilibrium plots and phenotype-phenotype correlation plots.

We will present a comprehensive online demo of easyGWAS with the following use cases: i) importing genotype and phenotype information to the platform, ii) conducting GWAS with various pre-processing steps and algorithms to detect associations, iii) visualizing and sharing of results to boost collaboration with other researchers, and iii) performing meta-analysis and comparison of results from different GWAS.

easyGWAS is available online at <https://easygwas.ethz.ch/>.

References

1001 Genomes Consortium. (2016). 1,135 genomes reveal the global pattern of polymorphism in *Arabidopsis thaliana*. *Cell*, 166(2), 481-491.

Atwell, S., et al. (2010). Genome-wide association study of 107 phenotypes in *Arabidopsis thaliana* inbred lines. *Nature* 465: 627–631.

Freilinger, T., et al. (2012). Genome-wide association analysis identifies susceptibility loci for migraine without aura. *Nat. Genet.* 44: 777–782.

Grimm, D.G., et al. (2017) easyGWAS: A Cloud-based Platform for Comparing the Results of Genome-wide Association Studies, *The Plant Cell*. tpc-00551

Kirby, A., et al. (2010). Fine Mapping in 94 Inbred Mouse Strains Using a High-Density Haplotype Resource. *Genetics* 185: 1081–1095.

Lin, T., et al. (2014). Genomic analyses provide insights into the history of tomato breeding. *Nat. Genet.* 46: 1220–1226.

Mackay, T.F., et al. (2012). The *Drosophila melanogaster* genetic reference panel. *Nature* 482: 173–178.

Scott, L.J., et al. (2007). A Genome-Wide Association Study of Type 2 Diabetes in Finns Detects Multiple Susceptibility Variants. *Science* 316: 1341–1345.

Zhao, K., et al. (2011). Genome-wide association mapping reveals a rich genetic architecture of complex traits in *Oryza sativa*. *Nat. Commun.* 2: 467.