Kim *et al. BMC Genomics* (2025) 26:193
https://doi.org/10.1186/s12864-025-11356-9

BMC Genomics

SOFTWARE

Open Access

MARSweb: a fully automated web service for set-based association testing

Taegun Kim[1,2], Jaeseung Song[3] and Jong Wha J. Joo[1,2*]

Abstract

Background Despite the successes in GWAS, there is still a large gap between the known heritability and the part explained by the SNPs identified by GWAS. Set-based analysis is one of the approaches that has tried to identify associations between multiple variants in a locus a trait, leveraging allelic heterogeneity to increase power in association testing. MARS is a set-based analysis method that integrates likelihood ratio test with a recently developed fine mapping technique to accurately account for causal status of variants in a risk locus. Unfortunately, due to its complex running process, time complexity, and the requirement of high-performance computing resources, it is not widely used.

Results To address these issues, we proposed a fully automated web-based analysis service, MARSweb. By providing a web service, we minimized the effort required for initial configuration. Additionally, users can perform analyses by simply uploading their data without needing to familiarize themselves with intricate analysis procedures. Furthermore, it facilitates easier interpretation of results by integrating advanced visualization tools. We confirmed the performance of MARSweb by detecting eGenes and performing pathway analysis of the genes using a Yeast Dataset.

Conclusions MARSweb is a web-based analysis service that fully automates set-based analysis. It offers an intuitive user interface, making complex analyses more accessible while significantly reducing processing time for enhanced efficiency. MARSweb is available for use at http://cblab.dongguk.edu/MARSweb and its source code is available at https://github.com/DGU-CBLAB/MARSweb.

Keywords GWAS, Allelic heterogeneity, Set-based analysis, Likelihood ratio test, Web-based

Background

Over the past decades, Genome-wide association studies (GWAS) have discovered thousands of genetic variants associated with various traits and diseases. Unfortunately, these have turned out to explain only a small portion of phenotypic variants, and research those yet to be detected continues. Previous studies have shown that more than one causal variant may play a role in a single locus in influencing a particular disease or a trait [1–11]. This is referred to as allelic heterogeneity, and it is common for Mendelian traits [12] and known to be widespread across many other complex traits and expression quantitative trait loci [13]. For this reason, standard

*Correspondence:
Jong Wha J. Joo
jwjoo@dgu.ac.kr
[1]Division of AI Software Convergence, Dongguk University-Seoul, Seoul 04620, South Korea
[2]Department of Computer Science and Engineering, Dongguk University-Seoul, Seoul 04620, South Korea
[3]Department of Life Science, Dongguk University-Seoul, Seoul 04620, South Korea

genetic association tests that assume only one causal variant per locus may fail to identify variants with small effect sizes at loci displaying allelic heterogeneity.

While the standard association test analyzes an association between a single variant and a trait, several statistical approaches, known as set-based association tests, have been designed to analyze the association between a set of variants and a trait. Most of these set-based association test methods utilize simple statistics, such as mean or the sum of χ^2, to examine multiple variants together [14–17]. Recently, a method called Model-based Association test Reflecting causal Status (MARS) [18] has been proposed that extends a fine mapping method, CAVIAR [9], to identify associations between multiple variants in a locus and a trait. MARS incorporates the linkage disequilibrium (LD) structure into the model by utilizing a multivariate normal (MVN) distribution conditional on the causal status of the variants. It has been shown that MARS increases the power of association testing compared to other set-based association test methods by leveraging allelic heterogeneity [18].

Despite its advantages, MARS has not been widely used because it does not provide a ready-to-use program with user-friendly settings and execution processes. MARS is coded in a mixture of different programming languages and requires certain libraries to be installed for complex matrix multiplication processes. In addition, the compiling process often takes effort due to path settings and version-specific complications in which some parts of the code require higher versions of the compiler than others. Another challenge in running MARS is that, as genomics data continuously grows, high-performance computer resources are required for running complicated statistical calculations. The processing of genomic data not only costs time and memory at each step but also resources that may not be available to users.

Here, we introduce MARSweb, a fully automated web service for set-based association testing. MARSweb provides a user-friendly web service that not only allows users to run MARS without having to know all the complicated installation and execution procedures but also allows users to run MARS without having access to a high-performance computer. It reduces the running time of MARS by providing parallel processing and optimization in the I/O processes and complex statistical calculations used throughout the program, making it possible to analyze large genomics datasets in practice. In addition, it provides visual tools such as set-based eQTL map and Manhattan plot generators for visualizing the association between variant sets and a trait of interest. When applied to a yeast dataset [19], we show that MARSweb successfully discovered 2852 eGenes, of which 2210 were reported elsewhere [20]. For the newly discovered 642 eGenes, we performed a pathway analysis to find the genes are related to transcriptional regulation. Besides the web service, for those who do not want to share their data to a third party, a downloadable version of MARSweb is provided as well.

Implementation

Fully automated web service for SNP-set-based association testing

MARSweb is a web-based service that provides a user-friendly interface where a user can activate an analysis by simply clicking on the web browser. The execution process is hidden from the user, and the user does not necessarily keep track of each stage of running process. MARSweb is expected to increase the usability of the

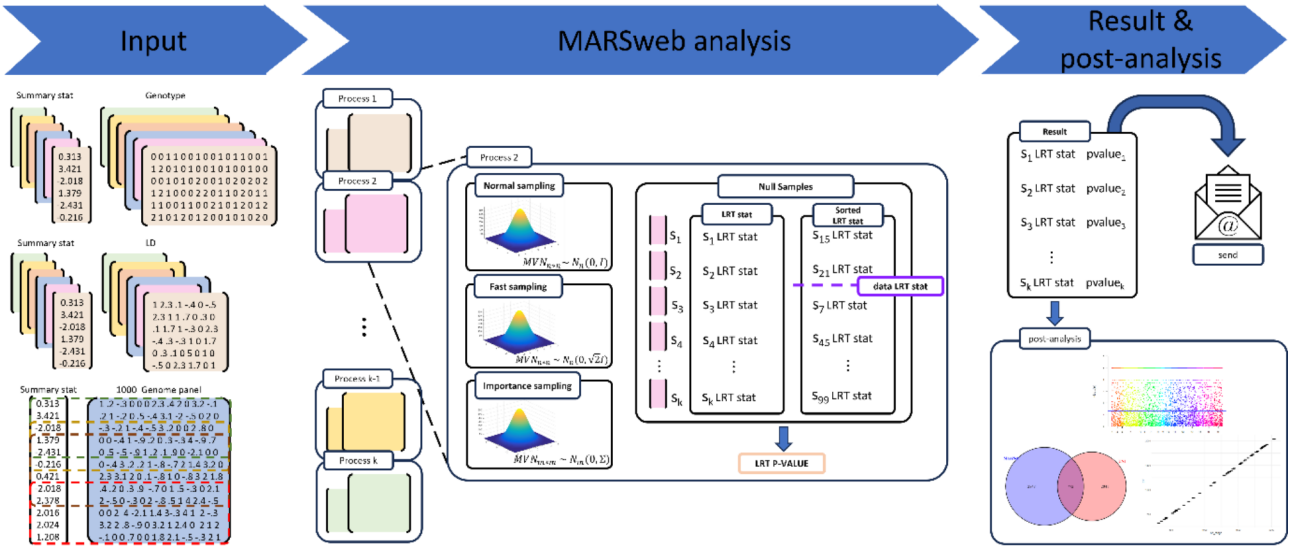

Fig. 1 Overview of MARSweb's main processes

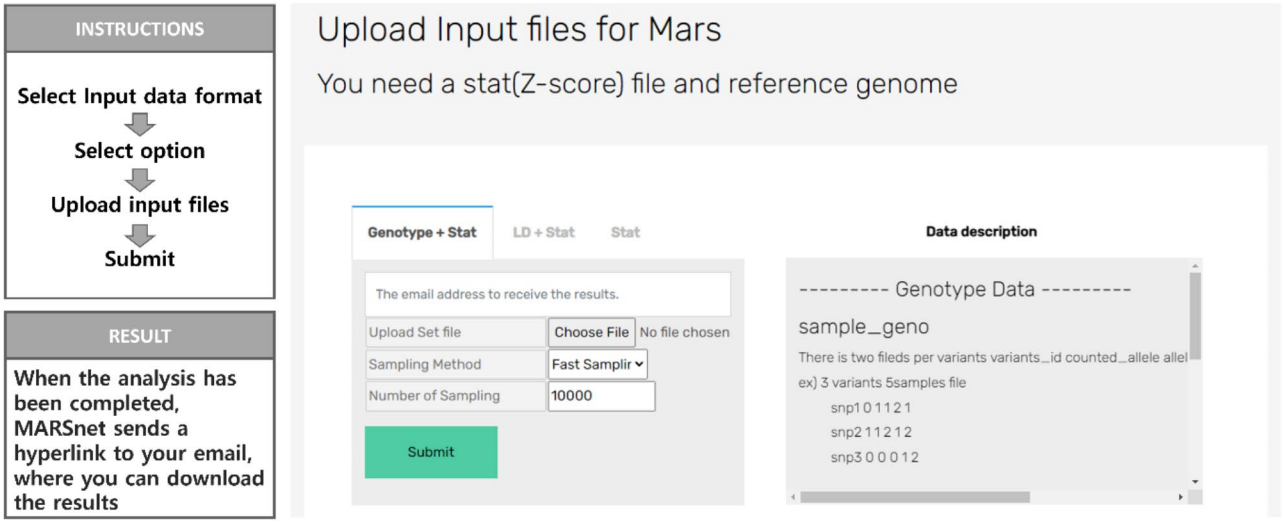

Fig. 2 MARSweb's service webpage for uploading input files

program for any user, but particularly those who lack professional experience in programming languages.

Figure 1 shows an overview of MARSweb's processes. MARSweb requires summary statistics and LD information on the genotypes. For the summary statistics, users can upload either a file with Z-scores or an association test report file from PLINK [21]. For the LD information, users can upload their own LD estimates or genotypes, from which the LD will be estimated by MARSweb automatically. In addition, for users who do not have genotype information, MARSweb provides reference LD panels not only for Human but also for other model organisms such as yeast and mouse. Genotypes from the 1000 Genome project [22], Saccharomyces Genome

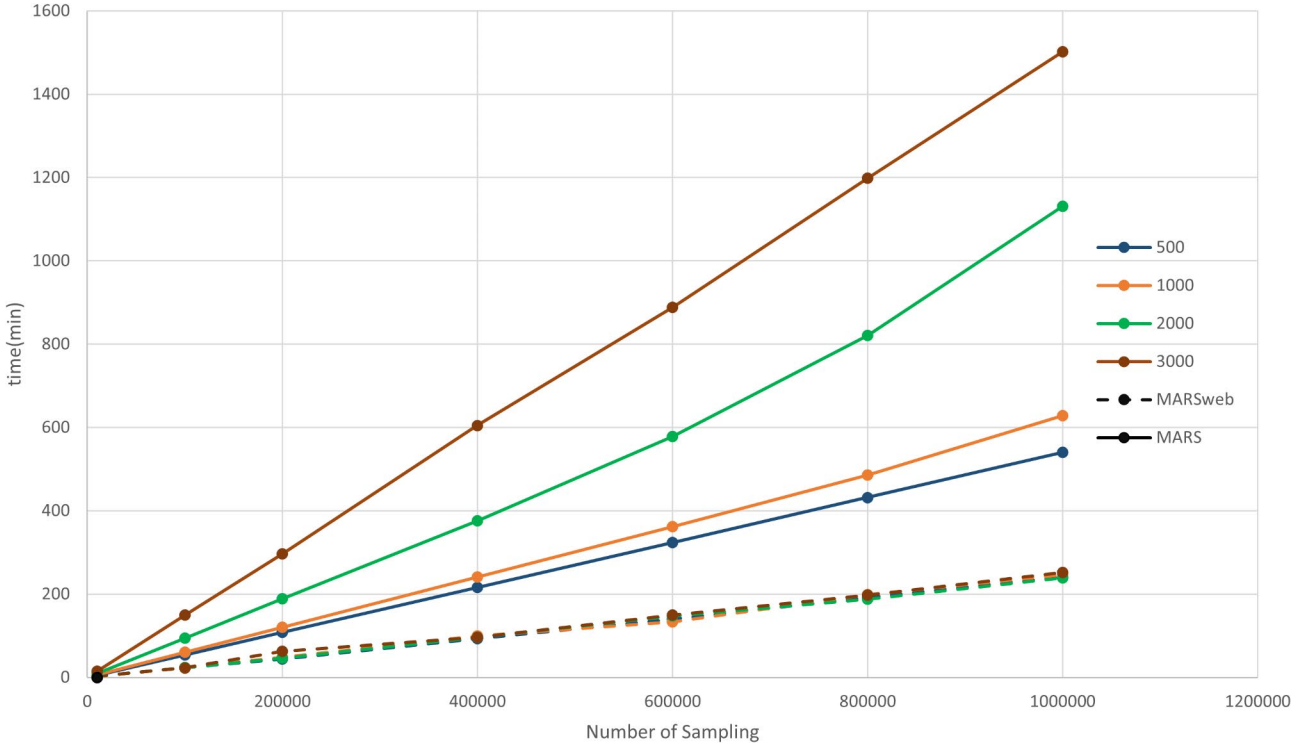

Fig. 3 Execution times for MARS (solid lines) and MARSweb (dashed lines). Yeast data was used to estimate the running time. The x-axis shows the number of samplings used for the statistical analysis and the y-axis shows the execution time. The orange, yellow, green, brown line colors indicate the number of the SNPs composing the dataset: 500, 1000, 2000, and 3000, respectively

Database [23], and Mouse Genome Project [24] were used to estimate the LD information of human, yeast, and mouse, respectively. As MARS is aimed at performing set-based analysis, set information should be provided as well. Users can either provide summary statistics and LD for each set to be analyzed or simply provide information containing chromosome and position for each set, with which MARSweb automatically extracts summary statistics and LD information for the set.

After uploading all the input files, MARSweb assigns the sets to be analyzed among the available processors. To make the run efficient, MARSweb utilizes multiple central processing units to divide tasks. To compute a test statistic, MARS performs a likelihood ratio test referred to as LRTstat. To test the statistical significance of an association using LRTstat, MARS performs a re-sampling process that samples null LRTstat statistics from an MVN distribution. There are three options for the sampling process: normal-sampling, fast-sampling, and importance-sampling. The normal-sampling option samples LRTstat from a MVN distribution with a variance-covariance matrix estimated from the summary statistics of sets, which does not require genotype information. Thus, if genotype information is not available, a user can use a reference LD panel, which is provided by MARSweb, and select the normal-sampling option. When genotype information is available, a user can choose either the fast-sampling or importance-sampling option. In most cases, the number of individuals (n) is much smaller than the number of SNPs (m). Utilizing the fact that covariance of summary statistics can be estimated from genotypes, fast-sampling use n by n identity matrix instead of m by m covariance matrix of summary statistics as variance-covariance matrix for the MVN sampling [25]. This drops the sampling time significantly. Lastly, to utilize genotypes to estimate the variance-covariance matrix in the MVN, the importance-sampling option performs the Monte Carlo simulations to reduce the number of samplings at the cost of accuracy. MARS has shown that importance-sampling well approximates the *p*-value estimated from the original sampling approach, while reducing the sampling number dramatically; from 10^8 to 10^4, in a GWAS dataset.

After the analysis has been completed, MARSweb sends a result to the email provided by the user. In addition, MARSweb provides visual tools that can be used to interpret the results or perform post-analysis.

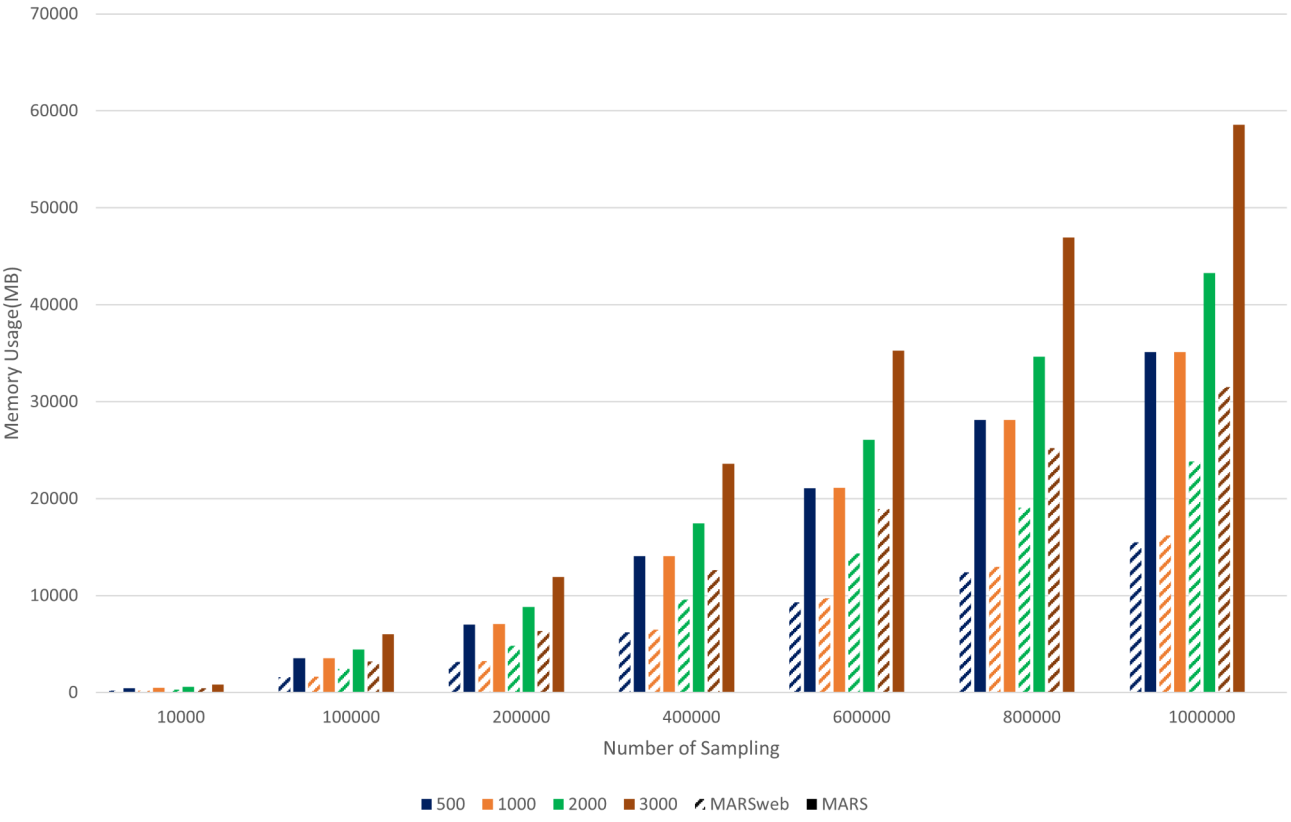

Fig. 4 Memory usage for MARS (solid bar) and MARSweb (dashed bar). Yeast data was used to estimate the memory usage. The x-axis shows the number of samples used for the statistical analysis and the y-axis shows memory usage. The orange, yellow, green, and brown colors indicate the number of the SNPs composing the dataset: 500, 1000, 2000, and 3000, respectively

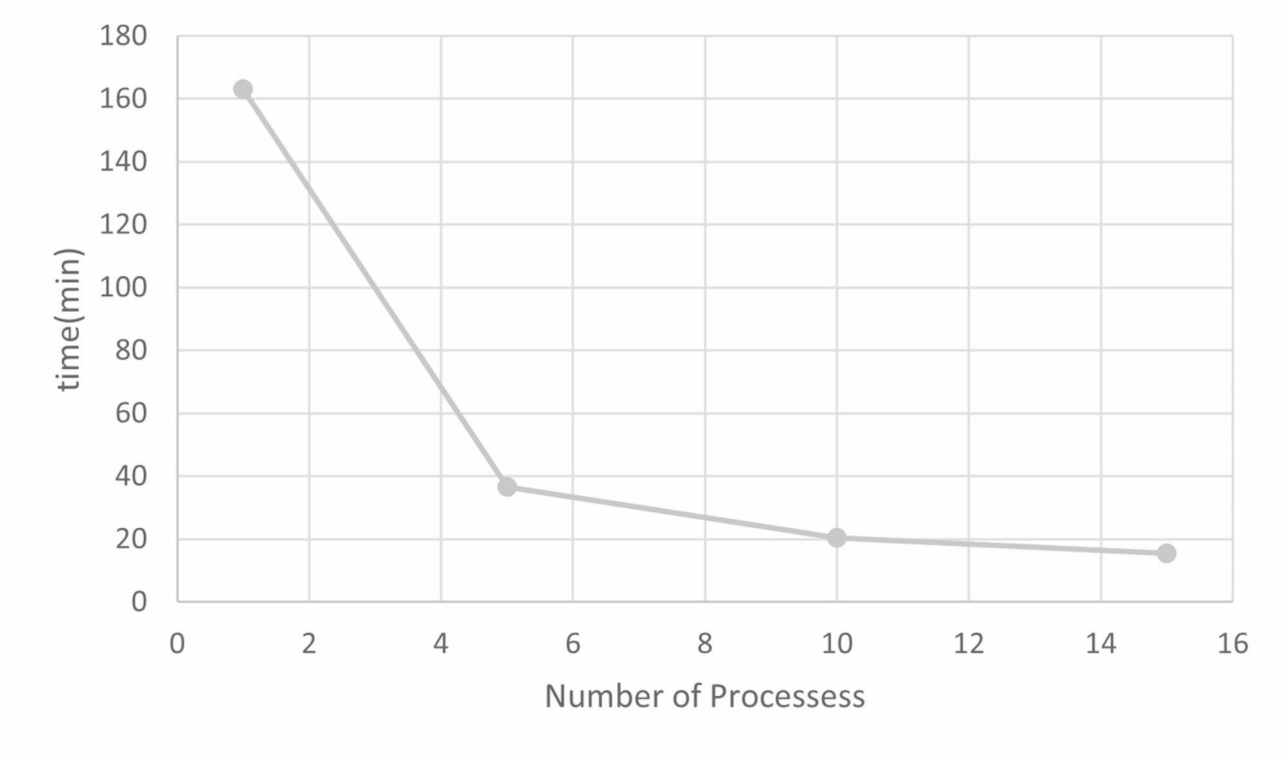

Fig. 5 Execution time for different numbers of processors used in parallel processing. 100 sets were used to estimate the execution time using 1, 5, 10, and 15 processors

MARSweb client-server web service

Apache Tomcat 9 is used as the application server for deploying the MARSweb system. HTML5 serves as the foundational framework for the web interface, with CSS and JavaScript enhancing user interaction. On the back-end, Java Servlet is employed to establish a robust client-server model.

The workflow involves users selecting input data and running options through the front-end. Subsequently, this information is transmitted to the server, which conducts the analysis, including data pre-processing. Following the completion of the analysis, the server sends a link to the user's provided email address, allowing them to download the results.

Figure 2 shows the webpage for running MARSweb, where users can upload input files, choose options for their analysis, and get description for each stage.

The web service and a step-by-step manual are provided on http://cblab.dongguk.edu/MARSweb.

Program optimization

MARS is composed of multiple running stages, for which different programming languages have been used. Communication between different languages produces I/O burdens for each stage, resulting in high memory usage and time delay. MARSweb unifies the programming language to C++, which is fast and efficient in many aspects and reduces the unnecessary I/O burdens. Furthermore, MARSweb utilized Eigen library to optimize the matrix calculations to reduce the running time.

Parallel processing for high-throughput analysis

The set-based association test is mostly used for high-throughput data such as GWAS or eQTL, which requires hundreds to tens of thousands of sets to be analyzed. At most of the running stages, each set runs independently, thus sets could be assigned to different processors. MARSweb utilizes high performance multiple core servers for the parallel processing to increase the efficiency of high-throughput analyses. When a user uploads an input dataset, MARSweb utilizes the available number of cpu cores to divide the tasks among multiple processes, the results of which are merged and provided to the user.

Post analysis tools

MARSweb provides post-analysis tools that visualize results from a set-based analysis. One generates set-based eQTL maps that show the association between sets of SNPs and gene expressions in a map, allowing that a users can systematically view a set-based eQTL analysis result in a sight. In the map, a user can compare results of the set-based association test to those of the standard eQTL analysis, referred to as the univariate test, which is

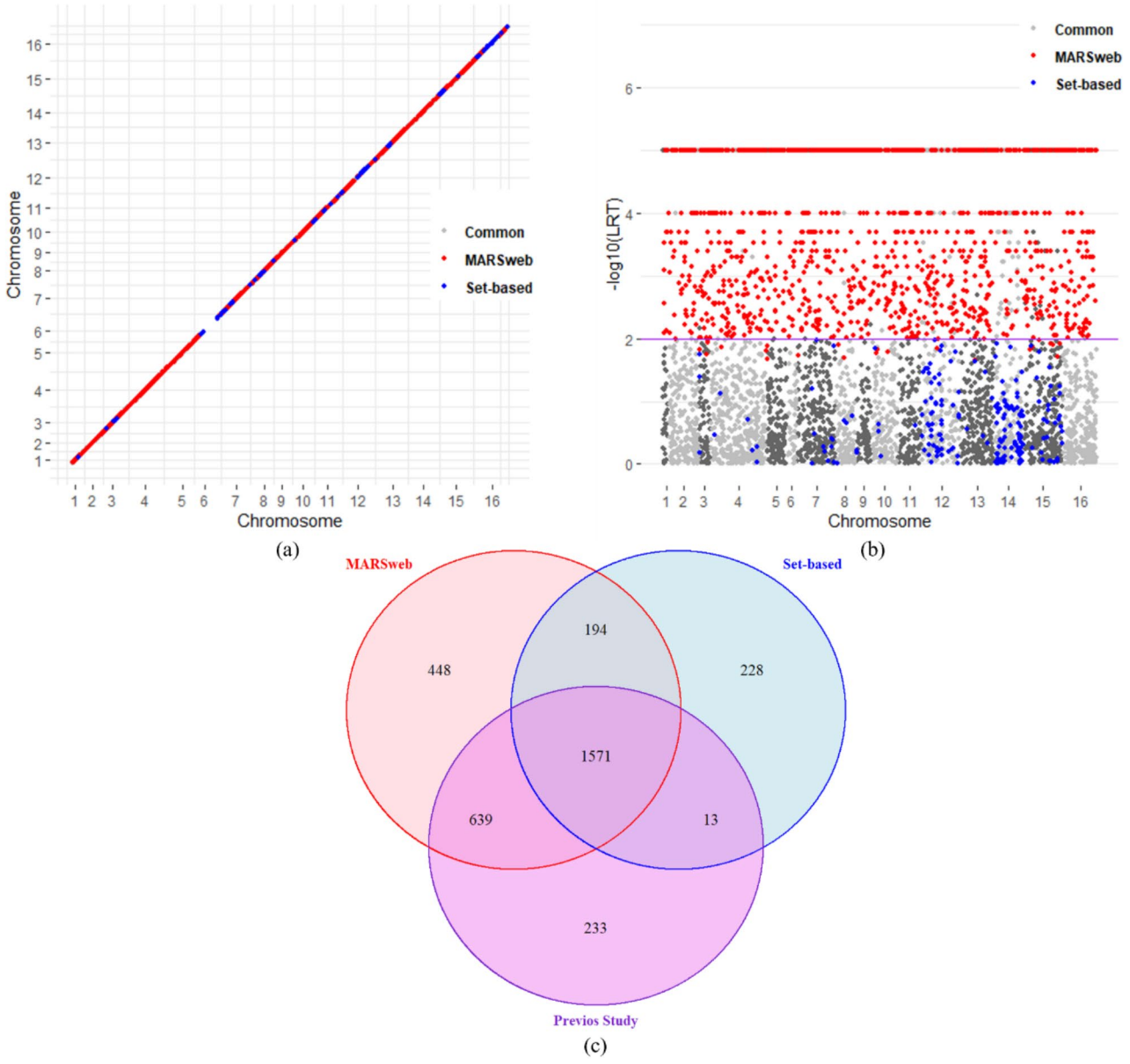

Fig. 6 eGenes identification in a yeast dataset using MARSweb. (**a**) A set-based eQTL map. The x-axis shows the locations of SNPs and the y-axis shows the location of genes. (**b**) A set-based manhattan plot. The x-axis shows the locations of genes and the y-axis shows the −log10 converted *p*-values. Red dots indicate genes identified by MARSweb but not identified by traditional methods in the previous study, blue dots show the results only from the set-based test, and gray dots show the results from both univariate SNP and set-based tests. (**c**) A Venn diagram, comparing eGenes identified by MARSweb (red), the ones identified by the univariate test (blue), and the ones reported in the previous study (purple) [11]

based on an association between a single SNP and a trait [18].

Another tool provided by MARSweb draws set-based Manhattan plots that show the association between sets of SNPs and a trait in a dot plot. With these plots, given a threshold, a user can identify significant associations with a quick look. As in the set-based eQTL map, a user can compare results of the set-based association test to those of the univariate association test.

MARSweb also provides Venn diagrams that compare the number of associations identified by the set-based test and the univariate test. A user can download the list of sets identified by either or both tests.

Downloadable docker image

Genetic data contains sensitive health-related information, and thus, sharing it represents a privacy risk for individuals or data subjects. Some data are protected by the law of genetic privacy [26], and it cannot be legally

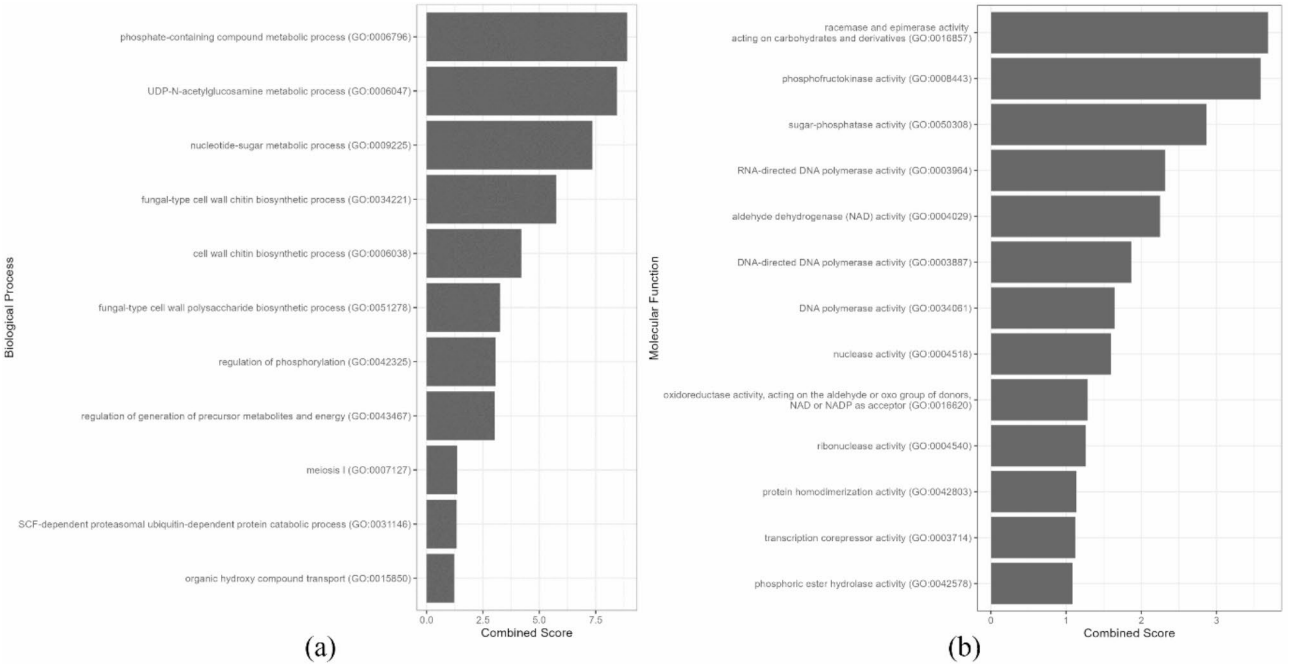

Fig. 7 Functional annotation results for the eGenes newly identified by MARSweb. The bar plots depict the significantly enriched pathways for (**a**) GO-Biological Process and (**b**) GO-Molecular Function. The lengths of the bars present the combined score of each GO-term

shared with others. For those who do not want to provide their data to a third-party, a downloadable version of MARSweb is available as a web software docker image file along with the source code. Users can easily download and install the docker image to run the MARSweb service on their local computer, cloud server, etc. A user does not need to consider any preprocessing steps, such as setting up the running environment, installing necessary packages, preparing running scripts, etc. For those not familiar with using a docker image, the necessary information and step-by-step commands are provided in a detailed user manual. Furthermore, MARSweb provides a source code for the analysis tool, thus a user can make any necessary changes for their analysis if required. The docker image, source code, and its manual are provided on https://github.com/DGU-CBLAB/MARSweb.

Results and discussion

MARSweb improves runtime performance of MARS

To evaluate the performance of MARSweb, we estimated the execution time for different SNP sample sizes—500, 1000, 2000, and 3000 SNPs, with sampling sizes ranging from 10^5 to 10^6 chosen from yeast data [19]. The biggest bottleneck for MARS with respect to the execution time is the MVN distribution re-sampling process. As the size of the analyzed dataset increases, the computation time increases by up the cube of the increase in the size of the dataset. Figure 3 compares execution time of MARS and MARSweb for the datasets with increasing numbers of SNP sampling. For both MARS and MARSweb, the fast-sampling option was used, and no parallel processing was applied for the analysis. The figure shows that MARSweb displays a dramatically higher execution time compared to MARS. As the number of SNPs increased, the discrepancy in execution time between MARS and MARSweb increases. Considering that the most used number of samplings sized used in GWAS was 10^8, which took over a day for MARS to complete, this represents a huge advantage for MARSweb. Notably, increasing the resampling size the performance of MARSweb, while for MARS, the set size is critical for the analysis in respect to the running cost.

In addition to the execution time, we have compared memory usage of MARS and MARSweb for the datasets with increasing numbers of SNP sampling (Fig. 4). The figure demonstrates that MARSweb is significantly more memory efficient.

In addition, MARSweb is expected to increase in performance utilizing parallel processing. To measure the effectiveness of parallel processing, we have increased the number of cores and estimated the running time. Figure 5 shows the execution time using different numbers of processors up to analyze 100 number of sets using MARSweb. The execution time drops as the number of processors increases.

The experiments were conducted on a server with 267 GB of RAM and a dual-socket Intel® Xeon® E5-2630 v4 CPU running at 2.20 GHz under a CentOS Linux 7 environment. For evaluating the running time and memory usage in Figs. 3 and 4 and a single core was used and for

evaluating the effectiveness of parallel processing (Fig. 5), depending on the number of processes, the required number of 20 cores was allocated.

Yeast dataset analysis using MARSweb

We performed an eQTL set-based analysis utilizing a yeast dataset containing 1012 meiotic segregants with 5720 genes and 42,052 SNPs [19]. The R package Matrix eQTL [27] was used for mapping SNP and gene locations and to estimate test statistics. Based on Saccharomyces cerevisiae reference genome SGD R64-1-1 [28], SNPs within ± 10Kb of a transcription starting site (TSS) were used to define each set, resulting in 5661 gene sets. Applying MARSweb, we identified 2852 eGenes, including 1087 not captured by the conventional set-based method. Among them, 642 are newly identified eGenes not reported in previous studies [20] (Fig. 6). A significance threshold level of $p < 0.01$ (FDR adjust p-value) was used for the analysis. Figure 6 shows the set-based eQTL map, SNP-based Manhattan plot, and a Venn diagram drawn by MARSweb.

To validate the credibility and biological meanings of the eGenes that were only detected using MARSweb, we conducted a functional annotation analysis for the 642 genes. There were 11 and 13 significantly enriched (combined score > 1) pathways for "Biological Processes" (BP) and "Molecular Function" (MF), respectively. The majority of the pathways identified in GO-BP were related to metabolic or synthetic processes, which implies that MARSweb can explain genetically regulated gene expression for fundamental biological mechanisms (Fig. 7A). Interestingly, these genes are known to have enzymatic activities involved in transcriptional regulation, supporting that MARSweb can successfully identify novel SNP-gene associations (Fig. 7B).

Conclusions

In this paper, we introduced a fully automated set-based analysis tool, referred to as MARSweb, providing a user-friendly web-based service. This hides all the complicated installation and running procedures from the user. After uploading the input data and choosing execution options, users do not even have to keep their computers running while performing the analysis and will be notified about the results via e-mail. The web service provides some visualization tools that can help the post-analysis interpretation and presentation of the results as well. In addition, we provide a downloadable docker image, which could be useful for those who do not want to expose their data to a third party.

MARSweb upgrades the previous MARS software in many aspects by optimizing the code and execution processes, including parallel processing. Utilizing various simulated data, we have compared the performance of MARSweb and MARS and show that MARSweb has improved greatly in respect to the running time, which, considering the sizes of modern genomic data available, is a major advantage. Utilizing yeast data, we discovered 642 new eGenes, not identified using a traditional analysis in previous research, and confirmed their validity through functional annotation analysis. This analysis shows that MARSweb can successfully identify the unknown SNP-gene associations in practice.

Considering the fact that the majority of MARS users are expected to be scientists with Biology or Statistics background, we believe the tool could be used by everyone, including those unfamiliar with computer programming. The MARSweb can be obtained from https://github.com/DGU-CBLAB/MARSweb, where installation instructions are provided, and the MARSweb service is available at http://cblab.dongguk.edu/MARSweb.

Abbreviations

GWAS	Genome-wide association study
SNP	Single-nucleotide polymorphisms
MARS	Model-based association test reflecting causal status
LD	Linkage disequilibrium
MVN	Multivariate normal
TSS	Transcription starting site
FDR	False discovery rate
eQTL	Expression quantitative trait loci
BP	Biological processes
MF	Molecular function

Acknowledgements

Not applicable.

Author contributions

TK and JWJJ designed this study. TK developed the MARSweb and performed experi-ments to evaluate it. JS performed the yeast dataset analysis. TK, JS, and JWJJ wrote the paper. All authors have read and approved the final manuscript.

Funding

This research was supported by the MSIT(Ministry of Science and ICT), Korea, under the ITRC(Information Technology Research Center) support progr am(IITP-2025-2020-0-01789), and the Artificial Intelligence Convergence Innovation Human Resources Development(IITP-2025-RS-2023-00254592) supervised by the IITP(Institute for Information & Communications Technology Planning & Evaluation).

Data availability

We utilized the processed dataset provided at https://figshare.com/s/83bddc1ddf3f97108ad4 for our experiments with yeast data.

Availability and requirements

Project name: MARSweb.
Project home page: http://cblab.dongguk.edu/MARSweb.
Operating system: CentOS7.
Programming Language: Java, C++.
Other requirements: Apache Tomcat 9.
License:
Any restrictions to use by non-academics: permission by the author.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Received: 3 October 2024 / Accepted: 11 February 2025
Published online: 24 February 2025

References

1. Broekema RV, Bakker OB, Jonkers IH. A practical view of fine-mapping and gene prioritization in the post-genome-wide association era. Open Biol. 2020;10(1):190221.
2. Haiman CA, Le Marchand L, Yamamato J, Stram DO, Sheng X, Kolonel LN, et al. A common genetic risk factor for colorectal and prostate cancer. Nat Genet. 2007;39(8):954–6.
3. Lango Allen H, Estrada K, Lettre G, Berndt SI, Weedon MN, Rivadeneira F, et al. Hundreds of variants clustered in genomic loci and biological pathways affect human height. Nature. 2010;467(7317):832–8.
4. Galarneau G, Palmer CD, Sankaran VG, Orkin SH, Hirschhorn JN, Lettre G. Fine-mapping at three loci known to affect fetal hemoglobin levels explains additional genetic variation. Nat Genet. 2010;42(12):1049–51.
5. Stahl EA, Raychaudhuri S, Remmers EF, Xie G, Eyre S, Thomson BP, et al. Genome-wide association study meta-analysis identifies seven new rheumatoid arthritis risk loci. Nat Genet. 2010;42(6):508–14.
6. Trynka G, Hunt KA, Bockett NA, Romanos J, Mistry V, Szperl A, et al. Dense genotyping identifies and localizes multiple common and rare variant association signals in celiac disease. Nat Genet. 2011;43(12):1193–201.
7. Yang J, Ferreira T, Morris AP, Medland SE, Genetic Investigation of ATC, Replication DIG et al. Conditional and joint multiple-SNP analysis of GWAS summary statistics identifies additional variants influencing complex traits. Nat Genet. 2012;44(4):369– 75, S1-3.
8. Flister MJ, Tsaih SW, O'Meara CC, Endres B, Hoffman MJ, Geurts AM, et al. Identifying multiple causative genes at a single GWAS locus. Genome Res. 2013;23(12):1996–2002.
9. Hormozdiari F, Kostem E, Kang EY, Pasaniuc B, Eskin E. Identifying causal variants at loci with multiple signals of association. Genetics. 2014;198(2):497–508.
10. Kichaev G, Yang WY, Lindstrom S, Hormozdiari F, Eskin E, Price AL, et al. Integrating functional data to prioritize causal variants in statistical fine-mapping studies. PLoS Genet. 2014;10(10):e1004722.
11. Hormozdiari F, Kichaev G, Yang WY, Pasaniuc B, Eskin E. Identification of causal genes for complex traits. Bioinformatics. 2015;31(12):i206–13.
12. Hormozdiari F, Zhu A, Kichaev G, Ju CJ, Segre AV, Joo JWJ, et al. Widespread allelic heterogeneity in Complex traits. Am J Hum Genet. 2017;100(5):789–802.
13. Jansen R, Hottenga JJ, Nivard MG, Abdellaoui A, Laport B, de Geus EJ, et al. Conditional eQTL analysis reveals allelic heterogeneity of gene expression. Hum Mol Genet. 2017;26(8):1444–51.
14. Ionita-Laza I, Lee S, Makarov V, Buxbaum JD, Lin X. Sequence kernel association tests for the combined effect of rare and common variants. Am J Hum Genet. 2013;92(6):841–53.
15. Bakshi A, Zhu Z, Vinkhuyzen AA, Hill WD, McRae AF, Visscher PM, et al. Fast set-based association analysis using summary data from GWAS identifies novel gene loci for human complex traits. Sci Rep. 2016;6:32894.
16. Liu JZ, McRae AF, Nyholt DR, Medland SE, Wray NR, Brown KM, et al. A versatile gene-based test for genome-wide association studies. Am J Hum Genet. 2010;87(1):139–45.
17. Li MX, Kwan JS, Sham PC. HYST: a hybrid set-based test for genome-wide association studies, with application to protein-protein interaction-based association analysis. Am J Hum Genet. 2012;91(3):478–88.
18. Hormozdiari F, Jung J, Eskin E, JW JJ. MARS: leveraging allelic heterogeneity to increase power of association testing. Genome Biol. 2021;22(1):128.
19. Bloom JS, Ehrenreich IM, Loo WT, Lite TL, Kruglyak L. Finding the sources of missing heritability in a yeast cross. Nature. 2013;494(7436):234–7.
20. Albert FW, Bloom JS, Siegel J, Day L, Kruglyak L. Genetics of trans-regulatory variation in gene expression. Elife. 2018;7.
21. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet. 2007;81(3):559–75.
22. Genomes Project C, Auton A, Brooks LD, Durbin RM, Garrison EP, Kang HM, Korbel JO, Marchini JL, McCarthy S, McVean GA, et al. A global reference for human genetic variation. Nature. 2015;526(7571):68–74.
23. Cherry JM, Adler C, Ball C, Chervitz SA, Dwight SS, Hester ET, Jia Y, Juvik G, Roe T, Schroeder M, et al. SGD: Saccharomyces Genome Database. Nucleic Acids Res. 1998;26(1):73–9.
24. Lilue J, Doran AG, Fiddes IT, Abrudan M, Armstrong J, Bennett R, Chow W, Collins J, Collins S, Czechanski A, et al. Sixteen diverse laboratory mouse reference genomes define strain-specific haplotypes and novel functional loci. Nat Genet. 2018;50(11):1574–83.
25. Joo JW, Hormozdiari F, Han B, Eskin E. Multiple testing correction in linear mixed models. Genome Biol. 2016;17:62.
26. Clayton EW, Evans BJ, Hazel JW, Rothstein MA. The law of genetic privacy: applications, implications, and limitations. J Law Biosci. 2019;6(1):1–36.
27. Shabalin AA. Matrix eQTL: ultra fast eQTL analysis via large matrix operations. Bioinformatics. 2012;28(10):1353–8.
28. Cherry JM, Adler C, Ball C, Chervitz SA, Dwight SS, Hester ET, et al. SGD: Saccharomyces Genome Database. Nucleic Acids Res. 1998;26(1):73–9.

Publisher's note