Send Orders for Reprints to reprints@benthamscience.net

Current Genomics, 2019, *20*, 325-331

REVIEW ARTICLE

Available Software for Meta-analyses of Genome-wide Expression Studies

Diego A. Forero[1,2,*]

[1]*PhD Program in Health Sciences, School of Medicine, Universidad Antonio Nariño, Bogotá, Colombia;* [2]*Laboratory of NeuroPsychiatric Genetics, Biomedical Sciences Research Group, School of Medicine, Universidad Antonio Nariño, Bogotá, Colombia*

ARTICLE HISTORY

Received: June 24, 2019
Revised: July 24, 2019
Accepted: August 08, 2019

DOI:
10.2174/1389202920666190822113912

Abstract: Advances in transcriptomic methods have led to a large number of published Genome-Wide Expression Studies (GWES), in humans and model organisms. For several years, GWES involved the use of microarray platforms to compare genome-expression data for two or more groups of samples of interest. Meta-analysis of GWES is a powerful approach for the identification of differentially expressed genes in biological topics or diseases of interest, combining information from multiple primary studies. In this article, the main features of available software for carrying out meta-analysis of GWES have been reviewed and seven packages from the Bioconductor platform and five packages from the CRAN platform have been described. In addition, nine previously described programs and four online programs are reviewed. Finally, advantages and disadvantages of these available programs and proposed key points for future developments have been discussed.

Keywords: Genomics, transcriptomics, bioinformatics, meta-analysis, genome-wide expression, microarray experiment.

1. INTRODUCTION

Advances in transcriptomic methods have led to a large number of published Genome-Wide Expression Studies (GWES), in humans and model organisms [1]. Broad application of international guidelines, such as the Minimum Information About a Microarray Experiment (MIAME) [2], has facilitated transparency in the report of results from GWES [3]. For several years, GWES were mainly based on the use of microarray platforms (such as the chips developed by commercial companies such as Affymetrix and Illumina, which have tens of thousands of probes targeting a large number of transcripts) [4] to compare genome-expression data for two or more groups of samples of interest [5]. In recent years, it has also involved methods built on sequencing of the transcriptome (RNA-seq), which are based on the use of next-generation sequencing platforms [6].

Repositories of GWES results, with freely available, structured and complete data, are one of the best examples of open science [7], which is beneficial for replication of initial findings and for meta-research [8, 9]. In the latest version of NCBI GEO (https://www.ncbi.nlm.nih.gov/geo), there is information for more than 3 million samples from more than 112.000 series, originated from 19.000 platforms [10, 11]. ArrayExpress is an online database (https://www.ebi.ac.uk/arrayexpress), created in 2002 and maintained by the European Bioinformatics Institute [1]. A large number of recent submissions to AE correspond to results from RNA-seq experiments and ArrayExpress has data from 72.000 experiments, for a total of 54.4 TB of available data [1]. Several available programs are useful for different steps in the bioinformatic analysis of individual GWES, such as GEOquery [12], GEO2R [11], shinyGEO [13] and Babelomics [14], among others.

Meta-analysis of GWES is a powerful approach for the identification of differentially expressed genes in biological topics or diseases of interest, combining information from multiple primary studies [3, 15-18]. Multiple bioinformatic analyses are needed to carry out a meta-analysis of GWES, such as data quality check, inclusion of data from technical replicates, annotation of probes and statistical procedures for meta-analysis [16], among others [3]. Existing statistical methods for meta-analyses of GWES have been based on combination of *p* values (such as the Stouffer's and Fisher's methods), effect sizes (such as the fixed effects and random effects models or ranks (such as the product of ranks and sum of ranks [3, 16, 19].

Several programs have been developed for carrying out meta-analysis of GWES [15, 20-23]. In addition to nominal statistical significance results, many of these programs provide the option of corrections for multiple comparisons, such as the False Discovery Rate (FDR) [24], among other bioinformatic procedures. In this article, the main features of available software for carrying out meta-analysis of GWES have been reviewed. In order to identify the available programs for meta-analysis of GWES, a search in PubMed and Google Scholar databases [25] was carried out, which was complemented with a revision of reference lists of key original and review articles [18].

*Address correspondence to this author at the Laboratory of NeuroPsychiatric Genetics, Biomedical Sciences Research Group, School of Medicine, Universidad Antonio Nariño, Bogotá, Colombia;
Tel: + 57 313 2610427; E-mail: diego.forero@uan.edu.co

1389-2029/19 $58.00+.00

2. AVAILABLE PACKAGES IN THE BIOCONDUCTOR PLATFORM

The Bioconductor platform (www.bioconductor.org) was developed as an open and collaborative resource for the development and availability of software for bioinformatics and computational biology [26]. It has been broadly used and supported by the international scientific community and the latest release contains 1649 packages. In Bioconductor, the BiocManager::install() function is useful for the installation of packages. Table **1** describes available software for meta-analyses of GWES in the Bioconductor platform, although some of those packages have not been described in articles published in indexed journals. In addition to the reference manuals, these packages have available tutorials or vignettes, which provide useful examples for the users.

One of the most used packages is RankProd, which is based in the rank product method [27], being useful to integrate results from different microarray platforms [20, 28]. OrderedList measures the similarities found between gene lists and generates random scores from perturbed data to evaluate the statistical significance [21]. GeneMeta allows to use fixed-effects or random-effects models [29] for meta-analysis of GWES data. MetaArray uses two methods (based on the Markov Chain Monte Carlo techniques and the expectation-maximization algorithm) [30] for obtaining a probability of expression in a meta-analysis. Crossmeta is a package that facilitates analyses for different platforms and species, carrying out effect size and pathway meta-analyses. MetaSeq is based on the non-parametric method NOISeq [31] for carrying out meta-analysis of RNA-seq studies. Metahdep is a package that allows to carry out meta-analysis of GWES using fixed-effects or random-effects models and taking hierarchical dependence into account [22].

3. AVAILABLE PACKAGES IN THE CRAN PLATFORM

The Comprehensive R Archive Network (CRAN) (https://cran.r-project.org) was created more than 20 years ago, as a public repository of packages for the R platform contributed by the international scientific community [32]. It currently has more than 14.000 available packages and the install.packages() and library() functions are useful for installing and running packages, respectively. Table **2** describes available software for meta-analyses of GWES in the CRAN platform; all those packages have been described in articles published in indexed journals. Some of these programs have detailed tutorials available.

Table 1. Available packages in the Bioconductor platform.

Package	Article	Citations	Depends on	Rank
RankProd	Hong 2006	634	R >= 3.2.1, stats, methods, Rmpfr, gmp	200
OrderedList	Lottaz 2006	68	R>= 2.1.0, Biobase, twilight, method	332
GeneMeta	NA	NA	R>= 2.10, methods, Biobase, genefilter	601
MetaArray	NA	NA	NA	689
crossmeta	NA	NA	R>= 3.3	749
metaSeq	NA	NA	R>= 2.13.0, NOISeq, snow, Rcpp	876
metahdep	Stevens 2009	13	R>= 2.10, methods	1344

NA: Not Available.

Table 2. Available packages in the CRAN platform.

Package	Article	Citations	Tutorial	Depends on
RankAggreg	Pihur 2009	223	Yes	R≥ 2.12.0, gtools
metaMA	Marot 2009	86	Yes	R≥ 3.1.2, limma, SMVar
metaPath	Shen 2010	69	No	R≥ 3.0.0, Biobase, GSEABase, genefilter, impute
metaRNASeq	Rau 2014	34	Yes	R≥ 2.15.0
MetaIntegrator	Haynes 2017	19	Yes	R≥ 3.3, rmeta, multtest, ggplot2, parallel, Rmisc, gplots, Biobase, RMySQL, DBI, stringr, preprocessCore, GEOquery, GEOmetadb, RSQLite, data.table, ggpubr, ROCR, zoo, pracma, COCONUT, Metrics, manhattanly, snplist, DT, pheatmap, plyr, boot, dplyr, reshape2, rmarkdown, AnnotationDbi, HGNChelper, magrittr, readr

NA: Not Available.

RankAggreg allows combining gene lists from different studies and platforms, using the Genetic or the Cross-Entropy Monte Carlo algorithms [33]. metaMA is a package for meta-analysis using moderated effect sizes and *p* value combinations [34]. MetaPath facilitates the development of pathway enrichment meta-analyses, with an exploration of significance for entire pathways or for each gene [35]. MetaRNASeq is a package for meta-analysis of RNA-seq studies, using the inverse normal and Fisher combination methods for *p* value combinations [36]. MetaIntegrator allows the meta-analysis based on effect sizes and *p* values combination [37].

4. OTHER PREVIOUSLY DESCRIBED PROGRAMS

In addition to the packages available in the Bioconductor and CRAN platforms and reviewed above, there are other freely available programs for meta-analysis of GWES. Table **3** describes other freely available software for meta-analyses of GWES. Some of these programs are R packages that are not available on the CRAN or Bioconductor platforms and some other programs are not currently on the internet.

A-MADMAN is a program written in Python, available for Windows and Linux operative systems, for meta-analysis of GWES using data combination [38]. BayesPoolMicro is a program for Windows (it needs the WinBUGS software) and Linux, using a Bayesian hierarchical model for meta-analysis [39]. ICS is a program written in C++ that allows to identify the consistency of findings between GWES [40]. MAAMD runs in Windows and Mac OSX systems, allowing the preparation of data for meta-analysis of results obtained with the Affymetrix platforms [41], with help of the AltAnalyze program [42]. MAID is an R package that allows to carry out meta-analyses for two-channel microarrays, in addition to one-channel platforms [43]. metaGEM is an R package that was developed by Ramasamy *et al.* for carrying out meta-analysis of GWES data, for example, using a random-effects model [3]. MetaOmics is a pipeline with several modules for different types of computational studies, including a module with 12 methods for meta-analysis of GWES [44]. METRADISC is a program for carrying out meta-analysis of GWES, based on rank of genes between studies and using a non-parametric method (Monte Carlo permutations) [45]. MTGDR is an R package, based on the meta threshold gradient descent regularization method for meta-analysis of GWES [46].

5. AVAILABLE ONLINE PROGRAMS

There are only four available online programs for meta-analysis of GWES. NetworkAnalyst (https://www.networkanalyst.ca) is a user-friendly online platform that has the possibility of receiving as input GWE results from an important number of microarray platforms for different species [23]. In addition, expression data can be entered in the program using gene symbols as identifiers (instead of probe IDs) and it allows meta-analyses for RNA-seq studies. Users have a limit of 1000 samples for meta-analyses and NetworkAnalyst provides the tools for annotation, normalization and exploration of batch effects. It has several options for meta-analytical procedures, such as random and fixed-effects models, combining *p* values and vote counting [23]. It was previously called INMEX [47] and the two of the primary articles, [48] and [49], have 135 and 199 citations respectively. ExAtlas (https://lgsun.irp.nia.nih.gov/exatlas) is an online program for meta-analyses of GWES, including four statistical approaches: random and fixed-effects models and Fisher's method [50]. It has options for helping with extraction of data from NCBI GEO database and for carrying out other types of bioinformatic analyses, such as correlations between datasets and gene set enrichment and overlap, among others [50].

SMAGEXP (https://github.com/sblanck/smagexp) is an initiative [51] that incorporates the metaMA [34] and metaRNAseq [36] packages into the Galaxy online platform (https://usegalaxy.org) [52, 53]. Finally, RNA Meta Analysis (https://rnama.com) is an online program that facilitates several steps, such as preprocessing and annotation, for carrying out meta-analyses of GWES.

Table 3. Other previously described software.

Program	Article	Citations	Link
A-MADMAN	Bisognin 2009	35	compgen.bio.unipd.it/bioinfo/amadman/
BayesPoolMicro	Conlon 2006	52	www.math.umass.edu/~conlon/research/BayesPoolMicro
ICS	Rajaram 2009	5	NA
MAAMD	Gan 2014	11	www.biokepler.org/use_cases/maamd-workflow-standardize-meta-analyses-affymetrix-microarray-data
MAID	Borozan 2008	13	NA
metaGEM	Ramasamy 2008	NA	NA
MetaOmics	Ma 2018	1	https://github.com/metaOmics/metaOmics
METRADISC	Zintzaras 2008	51	NA
MTGDR	Ma 2009	49	http://www.cs.uiowa.edu/~jian/MTGDR/main.html

NA: Not Available.

6. EXAMPLES OF PUBLISHED META-ANALYSES USING AVAILABLE SOFTWARE

The use of available software, particularly those programs that have been initially well described in peer-reviewed publications, is quite helpful for the development of meta-analyses of GWES, in humans and other organisms [15, 18]. It facilitates the development of statistical analyses for genome-wide data and its replication by other researchers [15, 18]. In this section, some illustrative examples of use of available software for meta-analyses of GWES, reported in international articles, are highlighted.

Wang *et al.* carried out a meta-analysis for 7 GWES for Alzheimer´s disease and 9 GWES for Parkinson´s disease, which are available in NCBI GEO and ArrayExpress databases. The primary studies that were included used different microarray platforms and analyzed samples from several brain tissues. The authors used the RankProd package [20] for carrying the meta-analyses of GWES [54]. Jha *et al*. used 5 GWES available in the NCBI GEO database to carry out a meta-analysis for venous thrombosis, polycythemia vera and essential thrombocythemia. The authors used the metaMA package [34] to carry out the meta-analysis [55]. Piras *et al.* carried out a meta-analysis for 3 GWES that included samples from peripheral tissues from schizophrenia patients and controls; they used the GeneMeta package [56]. Forero *et al.* carried out a meta-analysis for GWES, using data from 24 previously published studies for patients with major depressive disorder and controls, with RNA extracted from different tissues and analyzed in multiple microarray platforms. They used the NetworkAnalyst program [23] for carrying out the meta-analysis of available GWES data [57]. Manchia *et al.* did a meta-analysis for 5 GWES for schizophrenia, which are available in the NCBI GEO database. One of these GWES was focused on human induced pluripotent stem cell-derived neurons and the other 4 GWES were carried out in post-mortem brain samples. They used the GeneMeta package for the meta-analysis of GWES data [58].

CONCLUSION AND FUTURE PERSPECTIVES

As discussed above, there are multiple available programs for carrying out meta-analysis of GWES. Packages that are on the Bioconductor and CRAN platforms have the advantages that they are easy to install [26] but some of them, from the perspective of the researchers in genomics, are not completely user friendly. On the other hand, available online programs have the advantage of being user friendly but have the limitations in number of samples or file sizes.

Given the size and complexity of recent GWES datasets (that in several cases have hundreds of samples), it would be important to have novel or updated programs that are both user-friendly and computationally powerful and that facilitates, in addition to analysis of microarray data, the development of meta-analysis of large RNA-seq studies [6, 59]. In this context, broad implementation of standards for reporting of RNA-seq experiments, similar to the MIAME guidelines, would be quite helpful [60].

Programs that have detailed use guidelines might have a higher possibility of being employed extensively and adequately by the international scientific community. As maintenance and update of bioinformatics software is a critical issue [61], it is important to highlight that the dependence on other programs leads to some issues when running some Bioconductor and CRAN packages. An important number of the citations of the articles describing the different CRAN and Bioconductor packages were related to methodological developments. On the other hand, several published articles describing results for meta-analysis of GWES used in-house scripts, instead of available software [62-64]. Some of these programs can also be used for meta-analysis of other types of larger -omics datasets, such as genome-wide methylation studies [65], which have a larger number of probes.

Integration of meta-analysis of GWES with experimental approaches and with other *in silico* explorations [66-70] would lead to a better and deeper understanding of multiple biological processes and of pathophysiology of diseases.

CONSENT FOR PUBLICATION

Not applicable.

FUNDING

The author is supported by research grants # 824-2017 and 2019220 from Colciencias and VCTI.

CONFLICT OF INTEREST

The author declares no conflict of interest, financial or otherwise.

ACKNOWLEDGEMENTS

Declared none.

REFERENCES

[1] Athar, A.; Füllgrabe, A.; George, N.; Iqbal, H.; Huerta, L.; Ali, A.; Snow, C.; Fonseca, N.A.; Petryszak, R.; Papatheodorou, I.; Sarkans, U.; Brazma, A. ArrayExpress update - from bulk to single-cell expression data. *Nucleic Acids Res.,* **2019**, *47*(D1), D711-D715. [http://dx.doi.org/10.1093/nar/gky964] [PMID: 30357387]

[2] Brazma, A.; Hingamp, P.; Quackenbush, J.; Sherlock, G.; Spellman, P.; Stoeckert, C.; Aach, J.; Ansorge, W.; Ball, C.A.; Causton, H.C.; Gaasterland, T.; Glenisson, P.; Holstege, F.C.; Kim, I.F.; Markowitz, V.; Matese, J.C.; Parkinson, H.; Robinson, A.; Sarkans, U.; Schulze-Kremer, S.; Stewart, J.; Taylor, R.; Vilo, J.; Vingron, M. Minimum information about a microarray experiment (MIAME)-toward standards for microarray data. *Nat. Genet.,* **2001**, *29*(4), 365-371. [http://dx.doi.org/10.1038/ng1201-365] [PMID: 11726920]

[3] Ramasamy, A.; Mondry, A.; Holmes, C.C.; Altman, D.G. Key issues in conducting a meta-analysis of gene expression microarray datasets. *PLoS Med.,* **2008**, *5*(9), e184. [http://dx.doi.org/10.1371/journal.pmed.0050184] [PMID: 18767902]

[4] Irizarry, R.A.; Warren, D.; Spencer, F.; Kim, I.F.; Biswal, S.; Frank, B.C.; Gabrielson, E.; Garcia, J.G.; Geoghegan, J.; Germino, G.; Griffin, C.; Hilmer, S.C.; Hoffman, E.; Jedlicka, A.E.; Kawasaki, E.; Martínez-Murillo, F.; Morsberger, L.; Lee, H.; Petersen, D.; Quackenbush, J.; Scott, A.; Wilson, M.; Yang, Y.; Ye, S.Q.; Yu, W. Multiple-laboratory comparison of microarray platforms. *Nat. Methods,* **2005**, *2*(5), 345-350. [http://dx.doi.org/10.1038/nmeth756] [PMID: 15846361]

[5] Allison, D.B.; Cui, X.; Page, G.P.; Sabripour, M. Microarray data analysis: From disarray to consolidation and consensus. *Nat. Rev. Genet.,* **2006**, *7*(1), 55-65. [http://dx.doi.org/10.1038/nrg1749] [PMID: 16369572]

[6] Hrdlickova, R.; Toloue, M.; Tian, B. RNA-Seq methods for transcriptome analysis. *Wiley Interdiscip. Rev. RNA,* **2017**, *8*(1), 8.

scriptome analysis. *Wiley Interdiscip. Rev. RNA,* **2017**, *8*(1), 8. [http://dx.doi.org/10.1002/wrna.1364] [PMID: 27198714]

[7] Masum, H.; Rao, A.; Good, B.M.; Todd, M.H.; Edwards, A.M.; Chan, L.; Bunin, B.A.; Su, A.I.; Thomas, Z.; Bourne, P.E. Ten simple rules for cultivating open science and collaborative R&D. *PLOS Comput. Biol.,* **2013**, *9*(9), e1003244. [http://dx.doi.org/10.1371/journal.pcbi.1003244] [PMID: 24086123]

[8] Bero, L. Meta-research matters: Meta-spin cycles, the blindness of bias, and rebuilding trust. *PLoS Biol.,* **2018**, *16*(4), e2005972. [http://dx.doi.org/10.1371/journal.pbio.2005972] [PMID: 29608562]

[9] Ioannidis, J.P.; Allison, D.B.; Ball, C.A.; Coulibaly, I.; Cui, X.; Culhane, A.C.; Falchi, M.; Furlanello, C.; Game, L.; Jurman, G.; Mangion, J.; Mehta, T.; Nitzberg, M.; Page, G.P.; Petretto, E.; van Noort, V. Repeatability of published microarray gene expression analyses. *Nat. Genet.,* **2009**, *41*(2), 149-155. [http://dx.doi.org/10.1038/ng.295] [PMID: 19174838]

[10] Barrett, T.; Wilhite, S.E.; Ledoux, P.; Evangelista, C.; Kim, I.F.; Tomashevsky, M.; Marshall, K.A.; Phillippy, K.H.; Sherman, P.M.; Holko, M.; Yefanov, A.; Lee, H.; Zhang, N.; Robertson, C.L.; Serova, N.; Davis, S.; Soboleva, A. NCBI GEO: Archive for functional genomics data sets update. *Nucleic Acids Res.,* **2013**, *41*(Database issue), D991-D995. [PMID: 23193258]

[11] Sayers, E.W.; Agarwala, R.; Bolton, E.E.; Brister, J.R.; Canese, K.; Clark, K.; Connor, R.; Fiorini, N.; Funk, K.; Hefferon, T.; Holmes, J.B.; Kim, S.; Kimchi, A.; Kitts, P.A.; Lathrop, S.; Lu, Z.; Madden, T.L.; Marchler-Bauer, A.; Phan, L.; Schneider, V.A.; Schoch, C.L.; Pruitt, K.D.; Ostell, J. Database resources of the National Center for Biotechnology Information. *Nucleic Acids Res.,* **2019**, *47*(D1), D23-D28. [http://dx.doi.org/10.1093/nar/gky1069] [PMID: 30395293]

[12] Davis, S.; Meltzer, P.S. GEO query: A bridge between the Gene Expression Omnibus (GEO) and Bio conductor. *Bioinformatics,* **2007**, *23*(14), 1846-1847. [http://dx.doi.org/10.1093/bioinformatics/btm254] [PMID: 17496320]

[13] Dumas, J.; Gargano, M.A.; Dancik, G.M. Shiny GEO: A web-based application for analyzing gene expression omnibus datasets. *Bioinformatics,* **2016**, *32*(23), 3679-3681. [http://dx.doi.org/10.1093/bioinformatics/btw519] [PMID: 27503226]

[14] Alonso, R.; Salavert, F.; Garcia-Garcia, F.; Carbonell-Caballero, J.; Bleda, M.; Garcia-Alonso, L.; Sanchis-Juan, A.; Perez-Gil, D.; Marin-Garcia, P.; Sanchez, R.; Cubuk, C.; Hidalgo, M.R.; Amadoz, A.; Hernansaiz-Ballesteros, R.D.; Alemán, A.; Tarraga, J.; Montaner, D.; Medina, I.; Dopazo, J. Babelomics 5.0: Functional interpretation for new generations of genomic data. *Nucleic Acids Res.,* **2015**, *43*(W1), W117-21. [http://dx.doi.org/10.1093/nar/gkv384] [PMID: 25897133]

[15] Walsh, C.J.; Hu, P.; Batt, J.; Santos, C.C. Microarray meta-analysis and cross-platform normalization: Integrative genomics for robust biomarker discovery. *Microarrays (Basel),* **2015**, *4*(3), 389-406. [http://dx.doi.org/10.3390/microarrays4030389] [PMID: 27600230]

[16] Kontou, P.I.; Pavlopoulou, A.; Bagos, P.G. Methods of analysis and meta-analysis for identifying differentially expressed genes. *Methods Mol. Biol.,* **2018**, *1793*, 183-210. [http://dx.doi.org/10.1007/978-1-4939-7868-7_12] [PMID: 29876898]

[17] Waldron, L.; Riester, M. Meta-analysis in gene expression studies. *Methods Mol. Biol.,* **2016**, *1418*, 161-176. [http://dx.doi.org/10.1007/978-1-4939-3578-9_8] [PMID: 27008014]

[18] Forero, D.A.; Lopez-Leon, S.; González-Giraldo, Y.; Bagos, P.G. Ten simple rules for carrying out and writing meta-analyses. *PLOS Comput. Biol.,* **2019**, *15*(5), e1006922. [http://dx.doi.org/10.1371/journal.pcbi.1006922] [PMID: 31095553]

[19] Chang, L.C.; Lin, H.M.; Sibille, E.; Tseng, G.C. Meta-analysis methods for combining multiple expression profiles: Comparisons, statistical characterization and an application guideline. *BMC Bioinformatics,* **2013**, *14*, 368. [http://dx.doi.org/10.1186/1471-2105-14-368] [PMID: 24359104]

[20] Hong, F.; Breitling, R.; McEntee, C.W.; Wittner, B.S.; Nemhauser, J.L.; Chory, J. Rank Prod: A bio conductor package for detecting differentially expressed genes in meta-analysis. *Bioinformatics,* **2006**, *22*(22), 2825-2827. [http://dx.doi.org/10.1093/bioinformatics/btl476] [PMID: 16982708]

[21] Lottaz, C.; Yang, X.; Scheid, S.; Spang, R. Ordered list: A bio conductor package for detecting similarity in ordered gene lists. *Bioinformatics,* **2006**, *22*(18), 2315-2316. [http://dx.doi.org/10.1093/bioinformatics/btl385] [PMID: 16844712]

[22] Stevens, J.R.; Nicholas, G. metahdep: Meta-analysis of hierarchically dependent gene expression studies. *Bioinformatics,* **2009**, *25*(19), 2619-2620. [http://dx.doi.org/10.1093/bioinformatics/btp468] [PMID: 19648140]

[23] Zhou, G.; Soufan, O.; Ewald, J.; Hancock, R.E.W.; Basu, N.; Xia, J. Network analyst 3.0: A visual analytics platform for comprehensive gene expression profiling and meta-analysis. *Nucleic Acids Res.,* **2019**, *47*(W1), W234-W241. [http://dx.doi.org/10.1093/nar/gkz240] [PMID: 30931480]

[24] Tusher, V.G.; Tibshirani, R.; Chu, G. Significance analysis of microarrays applied to the ionizing radiation response. *Proc. Natl. Acad. Sci. USA,* **2001**, *98*(9), 5116-5121. [http://dx.doi.org/10.1073/pnas.091062498] [PMID: 11309499]

[25] Falagas, M.E.; Pitsouni, E.I.; Malietzis, G.A.; Pappas, G. Comparison of PubMed, scopus, web of science and google scholar: Strengths and weaknesses. *FASEB J.,* **2008**, *22*(2), 338-342. [http://dx.doi.org/10.1096/fj.07-9492LSF] [PMID: 17884971]

[26] Gentleman, R.C.; Carey, V.J.; Bates, D.M.; Bolstad, B.; Dettling, M.; Dudoit, S.; Ellis, B.; Gautier, L.; Ge, Y.; Gentry, J.; Hornik, K.; Hothorn, T.; Huber, W.; Iacus, S.; Irizarry, R.; Leisch, F.; Li, C.; Maechler, M.; Rossini, A.J.; Sawitzki, G.; Smith, C.; Smyth, G.; Tierney, L.; Yang, J.Y.; Zhang, J. Bioconductor: Open software development for computational biology and bioinformatics. *Genome Biol.,* **2004**, *5*(10), R80. [http://dx.doi.org/10.1186/gb-2004-5-10-r80] [PMID: 15461798]

[27] Breitling, R.; Armengaud, P.; Amtmann, A.; Herzyk, P. Rank products: A simple, yet powerful, new method to detect differentially regulated genes in replicated microarray experiments. *FEBS Lett.,* **2004**, *573*(1-3), 83-92. [http://dx.doi.org/10.1016/j.febslet.2004.07.055] [PMID: 15327980]

[28] Del Carratore, F.; Jankevics, A.; Eisinga, R.; Heskes, T.; Hong, F.; Breitling, R. Rank Prod 2.0: A refactored bioconductor package for detecting differentially expressed features in molecular profiling datasets. *Bioinformatics,* **2017**, *33*(17), 2774-2775. [http://dx.doi.org/10.1093/bioinformatics/btx292] [PMID: 28481966]

[29] Choi, J.K.; Yu, U.; Kim, S.; Yoo, O.J. Combining multiple microarray studies and modeling inter study variation. *Bioinformatics,* **2003**, *19*(Suppl.1), i84-i90. [http://dx.doi.org/10.1093/bioinformatics/btg1010] [PMID: 12855442]

[30] Choi, H.; Shen, R.; Chinnaiyan, A.M.; Ghosh, D. A latent variable approach for meta-analysis of gene expression data from multiple microarray experiments. *BMC Bioinformatics,* **2007**, *8*, 364. [http://dx.doi.org/10.1186/1471-2105-8-364] [PMID: 17900369]

[31] Tarazona, S.; García-Alcalde, F.; Dopazo, J.; Ferrer, A.; Conesa, A. Differential expression in RNA-seq: A matter of depth. *Genome Res.,* **2011**, *21*(12), 2213-2223. [http://dx.doi.org/10.1101/gr.124321.111] [PMID: 21903743]

[32] Polanin, J.R.; Hennessy, E.A.; Tanner-Smith, E.E. A review of meta-analysis packages in R. *J. Educ. Behav. Stat.,* **2017**, *42*, 206-242. [http://dx.doi.org/10.3102/1076998616674315]

[33] Pihur, V.; Datta, S.; Datta, S. Rank Aggreg, an R package for weighted rank aggregation. *BMC Bioinformatics,* **2009**, *10*, 62. [http://dx.doi.org/10.1186/1471-2105-10-62] [PMID: 19228411]

[34] Marot, G.; Foulley, J.L.; Mayer, C.D.; Jaffrézic, F. Moderated effect size and p-value combinations for microarray meta-analyses. *Bioinformatics,* **2009**, *25*(20), 2692-2699. [http://dx.doi.org/10.1093/bioinformatics/btp444] [PMID: 19628502]

[35] Shen, K.; Tseng, G.C. Meta-analysis for pathway enrichment analysis when combining multiple genomic studies. *Bioinformatics,* **2010**, *26*(10), 1316-1323.

[http://dx.doi.org/10.1093/bioinformatics/btq148] [PMID: 20410053]

[36] Rau, A.; Marot, G.; Jaffrézic, F. Differential meta-analysis of RNA-seq data from multiple studies. *BMC Bioinformatics,* **2014**, *15*, 91. [http://dx.doi.org/10.1186/1471-2105-15-91] [PMID: 24678608]

[37] Haynes, W.A.; Vallania, F.; Liu, C.; Bongen, E.; Tomczak, A.; Andres-Terrè, M.; Lofgren, S.; Tam, A.; Deisseroth, C.A.; Li, M.D.; Sweeney, T.E.; Khatri, P. Empowering multi-cohort gene expression analysis to increase reproducibility. *Pac. Symp. Biocomput.,* **2017**, *22*, 144-153. [http://dx.doi.org/10.1142/9789813207813_0015] [PMID: 27896970]

[38] Bisognin, A.; Coppe, A.; Ferrari, F.; Risso, D.; Romualdi, C.; Bicciato, S.; Bortoluzzi, S. A-MADMAN: Annotation-based microarray data meta-analysis tool. *BMC Bioinformatics,* **2009**, *10*, 201. [http://dx.doi.org/10.1186/1471-2105-10-201] [PMID: 19563634]

[39] Conlon, E.M.; Song, J.J.; Liu, J.S. Bayesian models for pooling microarray studies with multiple sources of replications. *BMC Bioinformatics,* **2006**, *7*, 247. [http://dx.doi.org/10.1186/1471-2105-7-247] [PMID: 16677390]

[40] Rajaram, S. A novel meta-analysis method exploiting consistency of high-throughput experiments. *Bioinformatics,* **2009**, *25*(5), 636-642. [http://dx.doi.org/10.1093/bioinformatics/btp007] [PMID: 19176547]

[41] Gan, Z.; Wang, J.; Salomonis, N.; Stowe, J.C.; Haddad, G.G.; McCulloch, A.D.; Altintas, I.; Zambon, A.C. MAAMD: A work flow to standardize meta-analyses and comparison of affy-metrix microarray data. *BMC Bioinformatics,* **2014**, *15*, 69. [http://dx.doi.org/10.1186/1471-2105-15-69] [PMID: 24621103]

[42] Emig, D.; Salomonis, N.; Baumbach, J.; Lengauer, T.; Conklin, B.R.; Albrecht, M. AltAnalyze and DomainGraph: Analyzing and visualizing exon expression data. *Nucleic Acids Res.,* **2010**, *38*(Web Server issue), W755-W762. [http://dx.doi.org/10.1093/nar/gkq405] [PMID: 20513647]

[43] Borozan, I.; Chen, L.; Paeper, B.; Heathcote, J.E.; Edwards, A.M.; Katze, M.; Zhang, Z.; McGilvray, I.D. MAID: An effect size based model for microarray data integration across laboratories and platforms. *BMC Bioinformatics,* **2008**, *9*, 305. [http://dx.doi.org/10.1186/1471-2105-9-305] [PMID: 18616827]

[44] Ma, T.; Huo, Z.; Kuo, A.; Zhu, L.; Fang, Z.; Zeng, X.; Lin, C.W.; Liu, S.; Wang, L.; Liu, P.; Rahman, T.; Chang, L.C.; Kim, S.; Li, J.; Park, Y.; Song, C.; Oesterreich, S.; Sibille, E.; Tseng, G.C. MetaOmics: Analysis pipeline and browser-based software suite for transcriptomic meta-analysis. *Bioinformatics,* **2018**, *35*(9):1597-1599. [PMID: 30304367]

[45] Zintzaras, E.; Ioannidis, J.P. Meta-analysis for ranked discovery datasets: Theoretical framework and empirical demonstration for microarrays. *Comput. Biol. Chem.,* **2008**, *32*(1), 38-46. [http://dx.doi.org/10.1016/j.compbiolchem.2007.09.003] [PMID: 17988949]

[46] Ma, S.; Huang, J. Regularized gene selection in cancer microarray meta-analysis. *BMC Bioinformatics,* **2009**, *10*, 1. [http://dx.doi.org/10.1186/1471-2105-10-1] [PMID: 19118496]

[47] Xia, J.; Fjell, C.D.; Mayer, M.L.; Pena, O.M.; Wishart, D.S.; Hancock, R.E. INMEX--a web-based tool for integrative meta-analysis of expression data. *Nucleic Acids Res.,* **2013**, *41*(Web Server issue), W63-W70. [http://dx.doi.org/10.1093/nar/gkt338] [PMID: 23766290]

[48] Xia, J.; Benner, M.J.; Hancock, R.E. NetworkAnalyst--integrative approaches for protein-protein interaction network analysis and visual exploration. *Nucleic Acids Res.,* **2014**, *42*, W167-W174. [http://dx.doi.org/10.1093/nar/gku443]

[49] Xia, J.; Gill, E.E.; Hancock, R.E. NetworkAnalyst for statistical, visual and network-based meta-analysis of gene expression data. *Nat. Protoc.,* **2015**, *10*(6), 823-844. [http://dx.doi.org/10.1038/nprot.2015.052] [PMID: 25950236]

[50] Sharov, A.A.; Schlessinger, D.; Ko, M.S. ExAtlas: An interactive online tool for meta-analysis of gene expression data. *J. Bioinform. Comput. Biol.,* **2015**, *13*(6), 1550019. [http://dx.doi.org/10.1142/S0219720015500195] [PMID: 26223199]

[51] Blanck, S.; Marot, G. SMAGEXP: A galaxy tool suite for transcriptomics data meta-analysis. *Gigascience,* **2019**, *8*(2), 8. [http://dx.doi.org/10.1093/gigascience/giy167] [PMID: 30698691]

[52] Goecks, J.; Nekrutenko, A.; Taylor, J.; Galaxy, T. Galaxy: A comprehensive approach for supporting accessible, reproducible, and transparent computational research in the life sciences. *Genome Biol.,* **2010**, *11*(8), R86. [http://dx.doi.org/10.1186/gb-2010-11-8-r86] [PMID: 20738864]

[53] Hillman-Jackson, J.; Clements, D.; Blankenberg, D.; Taylor, J.; Nekrutenko, A.; Galaxy, T. Using galaxy to perform large-scale interactive data analyses. In: *Curr. Protoc. Bioinformatics*, **2012**, Unit10.5.

[54] Wang, Q.; Li, W.X.; Dai, S.X.; Guo, Y.C.; Han, F.F.; Zheng, J.J.; Li, G.H.; Huang, J.F. Meta-analysis of parkinson's disease and alzheimer's disease revealed commonly impaired pathways and dysregulation of NRF2-dependent genes. *J. Alzheimer's Dis.,* **2017**, *56*(4), 1525-1539. [http://dx.doi.org/10.3233/JAD-161032] [PMID: 28222515]

[55] Jha, P.K.; Vijay, A.; Sahu, A.; Ashraf, M.Z. Comprehensive Gene expression meta-analysis and integrated bioinformatic approaches reveal shared signatures between thrombosis and myeloproliferative disorders. *Sci. Rep.,* **2016**, *6*, 37099. [http://dx.doi.org/10.1038/srep37099] [PMID: 27892526]

[56] Piras, I.S.; Manchia, M.; Huentelman, M.J.; Pinna, F.; Zai, C.C.; Kennedy, J.L.; Carpiniello, B. Peripheral biomarkers in schizophrenia: A meta-analysis of microarray gene expression datasets. *Int. J. Neuropsychopharmacol.,* **2019**, *22*(3), 186-193. [http://dx.doi.org/10.1093/ijnp/pyy103] [PMID: 30576541]

[57] Forero, D.A.; Guio-Vega, G.P.; González-Giraldo, Y. A comprehensive regional analysis of genome-wide expression profiles for major depressive disorder. *J. Affect. Disord.,* **2017**, *218*, 86-92. [http://dx.doi.org/10.1016/j.jad.2017.04.061] [PMID: 28460316]

[58] Manchia, M.; Piras, I.S.; Huentelman, M.J.; Pinna, F.; Zai, C.C.; Kennedy, J.L.; Carpiniello, B. Pattern of gene expression in different stages of schizophrenia: Down-regulation of NPTX2 gene revealed by a meta-analysis of microarray datasets. *Eur. Neuropsychopharmacol.,* **2017**, *27*(10), 1054-1063. [http://dx.doi.org/10.1016/j.euroneuro.2017.07.002] [PMID: 28732597]

[59] Conesa, A.; Madrigal, P.; Tarazona, S.; Gomez-Cabrero, D.; Cervera, A.; McPherson, A.; Szcześniak, M.W.; Gaffney, D.J.; Elo, L.L.; Zhang, X.; Mortazavi, A. A survey of best practices for RNA-seq data analysis. *Genome Biol.,* **2016**, *17*, 13. [http://dx.doi.org/10.1186/s13059-016-0881-8] [PMID: 26813401]

[60] Brazma, A. Minimum Information About a Microarray Experiment (MIAME)--successes, failures, challenges. *Scientific World J.,* **2009**, *9*, 420-423. [http://dx.doi.org/10.1100/tsw.2009.57] [PMID: 19484163]

[61] Taschuk, M.; Wilson, G. Ten simple rules for making research software more robust. *PLOS Comput. Biol.,* **2017**, *13*(4), e1005412. [http://dx.doi.org/10.1371/journal.pcbi.1005412] [PMID: 28407023]

[62] Wang, N.; Zhang, Y.; Xu, L.; Jin, S. Relationship between alzheimer's disease and the immune system: A meta-analysis of differentially expressed genes. *Front. Neurosci.,* **2019**, *12*, 1026. [http://dx.doi.org/10.3389/fnins.2018.01026] [PMID: 30705616]

[63] Naz, S.; Khan, R.A.; Giddaluru, J.; Battu, S.; Vishwakarma, S.K.; Subahan, M.; Satti, V.; Khan, N.; Khan, A.A. Transcriptome meta-analysis identifies immune signature comprising of RNA binding proteins in ulcerative colitis patients. *Cell. Immunol.,* **2018**, *334*, 42-48. [http://dx.doi.org/10.1016/j.cellimm.2018.09.003] [PMID: 30327138]

[64] Li, M.D.; Burns, T.C.; Morgan, A.A.; Khatri, P. Integrated multi-cohort transcriptional meta-analysis of neurodegenerative diseases. *Acta Neuropathol. Commun.,* **2014**, *2*, 93. [http://dx.doi.org/10.1186/s40478-014-0093-y] [PMID: 25187168]

[65] Ratanatharathorn, A.; Boks, M.P.; Maihofer, A.X.; Aiello, A.E.; Amstadter, A.B.; Ashley-Koch, A.E.; Baker, D.G.; Beckham, J.C.; Bromet, E.; Dennis, M.; Garrett, M.E.; Geuze, E.; Guffanti, G.; Hauser, M.A.; Kilaru, V.; Kimbrel, N.A.; Koenen, K.C.; Kuan, P.F.; Logue, M.W.; Luft, B.J.; Miller, M.W.; Mitchell, C.; Nugent, N.R.; Ressler, K.J.; Rutten, B.P.F.; Stein, M.B.; Vermetten, E.; Vinkers, C.H.; Youssef, N.A.; Uddin, M.; Nievergelt, C.M.; Smith, A.K.; Nievergelt, C.M.; Smith, A.K. Epigenome-wide association of PTSD from heterogeneous cohorts with a common multi-site analysis pipeline. *Am. J. Med. Genet. B. Neuropsychiatr. Genet.,* **2017**, *174*(6), 619-630.

[http://dx.doi.org/10.1002/ajmg.b.32568] [PMID: 28691784]

[66] Ramanan, V.K.; Shen, L.; Moore, J.H.; Saykin, A.J. Pathway analysis of genomic data: Concepts, methods, and prospects for future development. *Trends Genet.,* **2012**, *28*(7), 323-332. [http://dx.doi.org/10.1016/j.tig.2012.03.004] [PMID: 22480918]

[67] Tranchevent, L.C.; Capdevila, F.B.; Nitsch, D.; De Moor, B.; De Causmaecker, P.; Moreau, Y. A guide to web tools to prioritize candidate genes. *Brief. Bioinform.,* **2011**, *12*(1), 22-32. [http://dx.doi.org/10.1093/bib/bbq007] [PMID: 21278374]

[68] Evangelou, E.; Ioannidis, J.P. Meta-analysis methods for genome-wide association studies and beyond. *Nat. Rev. Genet.,* **2013**, *14*(6), 379-389. [http://dx.doi.org/10.1038/nrg3472] [PMID: 23657481]

[69] Huang, W.; Sherman, B.T.; Lempicki, R.A. Bioinformatics enrichment tools: Paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res.,* **2009**, *37*(1), 1-13. [http://dx.doi.org/10.1093/nar/gkn923] [PMID: 19033363]

[70] Guio-Vega, G.P.; Forero, D.A. Functional genomics of candidate genes derived from genome-wide association studies for five common neurological diseases. *Int. J. Neurosci.,* **2017**, *127*(2), 118-123. [http://dx.doi.org/10.3109/00207454.2016.1149172] [PMID: 26829381]