

ImmuneRegulation: A web-based tool for identifying human immune regulatory elements

Selim Kalayci^{1,2}, Myvizhi Esai Selvan^{1,2}, Irene Ramos³, Chris Cotsapas⁴, Ruth R. Montgomery⁵, Gregory Poland⁶, Bali Pulendran⁷, John Tsang^{8,9}, Robert J. Klein^{1,2} and Zeynep H. Gümüş^{1,2,*}

¹ Department of Genetics and Genomic Sciences, Icahn School of Medicine at Mount Sinai, New York, NY, 10029, USA

² Icahn Institute for Genomics and Multiscale Biology, Icahn School of Medicine at Mount Sinai, New York, NY, 10029, USA

³ Department of Microbiology and Global Health & Emerging Pathogens Institute, Icahn School of Medicine at Mount Sinai, New York, NY, 10029, USA

⁴ Department of Neurology, Yale University, New Haven, CT, 06510, USA.

⁵ Section of Rheumatology, Department of Internal Medicine, Yale School of Medicine, New Haven, CT 06520, USA

⁶ Mayo Clinic, 200 First St SW Rochester, MN 55905, USA

⁷ Emory Vaccine Center/Yerkes National Primate Research Center at Emory University, Atlanta, Georgia, United States of America.

⁸ Trans-NIH Center for Human Immunology, Autoimmunity, and Inflammation, National Institutes of Health, Bethesda, MD, 20892, USA

⁹ Systems Genomics and Bioinformatics Unit, Laboratory of Systems Biology, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, MD, 20892, USA.

* To whom correspondence should be addressed. Tel: +1-212-824-8954; Fax: +1-212-241-3310; Email: zeynep.gumus@mssm.edu

ABSTRACT

Humans can vary considerably in their healthy immune phenotypes and in their immune responses to various stimuli. We have developed an interactive web-based tool, ImmuneRegulation, to enable discovery of human regulatory elements that drive some of the phenotypic differences observed in gene expression profiles. ImmuneRegulation currently provides the largest centrally integrated resource available in the literature on transcriptome regulation in whole blood and blood cell types, including genotype data from 23,040 individuals, with associated gene expression data from 30,562 experiments, that provide genetic variant-gene expression associations on ~200 million eQTLs. In addition, it includes 14 million transcription factor (TF) binding region hits extracted from 1945 TF ChIP-seq peaks and the latest GWAS catalog of 67,230 published SNP-trait associations. Users can interactively explore ImmuneRegulation to visualize and discover associations between their gene(s) of interest and their regulators (genetic variants or transcription factors) across multiple cohorts and studies. These regulators can explain cohort or cell type dependent gene expression variations and may be critical in selecting the ideal cohorts or cell types for follow-up studies. Overall, ImmuneRegulation aims to contribute to our understanding of the effects of eQTLs and TFs on heterogeneous transcriptional responses reported in studies on the blood; in the development of molecular signatures of immune response; and facilitate their future application in patient management. ImmuneRegulation is freely available at <http://icahn.mssm.edu/immuneregulation>.

INTRODUCTION

Recent high-throughput studies are contributing to an improved understanding of immune cell function and regulation (1). Extensive datasets are now available through the Human Immunology Project Consortium (HIPC) Phase I, which include measurements of healthy and activated human immune system, coupled with detailed clinical phenotyping in well-characterized cohorts (2–5). HIPC Phase 2 is continuing to collect data. In addition, large collaborative studies, including Genotype-Tissue Expression Project (GTEx) (6), Framingham Heart Study (7) and Encyclopedia of DNA Elements (ENCODE) (8–10) are collectively generating thousands of sequencing-based, genome-wide measurements of the transcriptome, transcription regulatory regions, transcription factor binding, and others that define states of the genome in many cell types and tissues, including a significant number of immune cell types. Furthermore, GWAS catalogs (11) provide a collection of large-scale studies that yield associations between genetic variants and various traits.

The wealth of data generated by these studies can be utilized for a deeper understanding of the mechanisms by which gene expression is regulated in different immune cells and immune systems of individuals in steady state and in response to immune stimuli (e.g. pathogens; disease; vaccines). Genetic variants typically appear to localize in the regulatory regions of genes and alter gene expression levels both proximally (putatively cis-acting) and distally (putatively trans-acting) (12), while transcription factors tend to bind directly to promoter regions proximally upstream of a gene, or directly to the RNA polymerase molecule and alter gene expression (13). However, often in the discovery phase, it is difficult to assess the real biological role of these regulatory elements without extensive experimentation on a large number of humans or immune cell types. To take advantage of the already generated rich datasets in the context of metadata that accompany them, a user-friendly immune-specific platform is needed to integrate, visualize and interactively explore these datasets. Such a platform can support investigations on human immunity, infection and disease as well as in the development of predictive models and therapies.

We have built an interactive web-based visual interface, ImmuneRegulation that capitalizes on the existing knowledge on the associations between transcription regulatory elements and gene expression changes in well-defined cohorts. ImmuneRegulation enables the exploration of massive immune-system specific gene regulation datasets and provides a centralized repository that enables seamless integration of information gained from large publicly available dataset resources to help efforts in understanding the associations between the immune system and its regulation. Easy visual queries to interact with these datasets assist discovery by creating resources for hypothesis generation. Users can simultaneously query multiple resources to identify cell-type or cohort specific regulatory elements that drive the expression of genes or gene sets of interest; study potential inherited susceptibility to specific responses or disease; or design follow-up cohort response studies with improved sensitivity and accuracy of classification by excluding or including certain individuals with specific genetic variants.

ImmuneRegulation currently includes genotype data from 23,040 individuals, associated with gene expression data from 30,562 experiments, providing genetic variant- gene expression associations on ~200 million eQTLs. These enable the identification of gene(s) whose expressions are affected by human germline genetic variation in blood or blood cell types. Additionally, Transcription Factors (TFs) from publicly available ENCODE Consortium TF ChIP-seq studies in human are included. immuneRegulation currently incorporates around 14 million gene region hits extracted from 1,945 TF ChIP-seq peaks datasets. Furthermore, GWAS catalog data within immuneRegulation repository is comprised of 67,230 published SNP-trait association information.

In summary, ImmuneRegulation is a unique resource focused on regulation of gene(s) that are phenotype-dependent in expression response to immune stimuli. More specifically, it is an interactive web-based visual tool that enables the exploration of massive data on the regulation of immune-system specific genes in well-defined cohorts. Furthermore, ImmuneRegulation is currently the only visual interface for exploration of trans-eQTLs. Prior to ImmuneRegulation, these datasets were publicly available in separate resources without any exploratory features. Overall, ImmuneRegulation provides a centralized repository that enables seamless integration of information gained from HIPC with public dataset resources to help efforts in understanding the associations between the immune system and its regulation, where users can capture regulatory elements that drive some of the phenotypic differences observed in immune-related transcriptomes. ImmuneRegulation is available at <http://icahn.mssm.edu/immuneregulation>.

MATERIAL AND METHODS

ImmuneRegulation provides a unified platform for a massive collection of immune regulation data in multiple cell types and in large human cohorts. Its visually intuitive and user-friendly interface helps users query, browse and interactively explore gene regulation datasets. Users can query for regulatory elements of genes of interest in real-time, utilizing its massive dataset. We illustrate the data repository and implementation details in the following sections.

Data Collection and Processing

Publicly available eQTL datasets. We collected and curated multiple studies on blood or blood cell types from public datasets that are already available in literature. The complete list of these public datasets is provided in Table 1.

cis- and trans-eQTLs from HIPC genotype-gene expression datasets. For HIPC studies that included both genotype and gene expression data (either microarray or RNA-seq), we first identified eQTLs by utilizing the pipeline provided at <https://github.com/molgenis/systemsgenetics/wiki/eQTL-mapping-analysis-cookbook> and then added them to the ImmuneRegulation database. These studies currently include systemic profiling of influenza-vaccinated individuals at multiple timepoints (Table 1).

cis- and trans-eQTLs from HIPC RNA-seq datasets. For HIPC RNA-seq studies without accompanying genotype data, we downloaded all raw datasets from GEO (www.ncbi.nlm.nih.gov/geo)

aligned spliced reads to human genome via STAR (14), and performed SNP and indel calling by using the Genome Analysis Toolkit (15) (GATK, <https://www.broadinstitute.org/gatk/>) best practices workflow, before running the same eQTL analysis pipeline. These studies currently include datasets from Chikungunya virus infected individuals profiled at both acute and convalescent stages of the disease (Table 1).

Transcription Factors (TF) datasets. We have used all human TF datasets from the publicly available ENCODE Consortium ChIP-seq studies (16, 17). These include a total of 1,945 ChIP-seq TF-binding site datasets (in bed narrowPeak file format) from 144 different cell/tissue types and 736 different TF targets. For each dataset, we calculated gene region hits using bedtools v2.25.0 (18). Then, for each gene, we incorporated the metadata of all datasets with gene region hit(s). The specific data tracks are transferred from ENCODE portal (16) upon user request.

GWAS datasets. We have incorporated the complete NHGRI-EBI GWAS catalog (11) (www.ebi.ac.uk/gwas, downloaded in March 2018) into ImmuneRegulation data repository to be utilized in queries. Currently, this includes information about 67,230 published SNP-trait associations.

Implementation and Interface

Implementation. ImmuneRegulation utilizes multiple client-side Javascript libraries (e.g. D3.js, JQuery, etc.) to visualize and interact with large volumes of multiple data types in real time. For visualizations that are displayed in the genome browser, we incorporated igv.js (<https://igv.org/doc/doc.html>) and customized it to serve our specific needs. We also incorporate several other utility libraries (e.g. dataTables.js, underscore.js, etc.) for data manipulation and interaction. For styling of the web interface, we mainly relied on Bootstrap v3.3.7 and our own custom CSS elements. As a result, immuneRegulation runs on all modern web browsers.

Interface. We designed the immuneRegulation interface with the aim of guiding users during queries and interactions with results. Warning messages are displayed dynamically to help users construct the queries properly. Similarly, certain dynamic elements are displayed as the users interact with the page and help inform the users regarding any possible interactivity option or further information. We also incorporated multiple example queries into immuneRegulation interface for users to gain familiarity with the tool and to create customized queries of their own easily. Furthermore, multiple detailed tutorials help explain various components and functionalities within immuneRegulation. An FAQ page provides basic information pertaining to the content as well as technical aspects of the interface (e.g. missing data, load time, etc.). Users can provide feedback or ask questions regarding any technical or scientific matters directly from the immuneRegulation interface as well.

RESULTS

ImmuneRegulation centrally integrates and utilizes massive immune-specific datasets in the context of their rich metadata. In its front-end, a visually intuitive web interface enables query, browsing and interaction with large volumes of data. Users can query for regulatory elements for gene(s) of interest

in real-time. For gene(s) queried, visual, interactive summaries of regulatory elements are returned to help explore, contextualize and communicate statistical analyses and results.

Data Repository

Since expression regulation is tissue-specific, we provide an immune system focused resource to understand gene regulation in whole blood or blood cells (T-cells, macrophages and monocytes). ImmuneRegulation currently includes genotype data from 23,040 individuals, associated with gene expression data from 30,562 experiments, providing genetic variant- gene expression associations on ~200 million eQTLs. These enable the identification of gene(s) whose expressions are affected by human germline genetic variation in blood or blood cell types. Additionally, Transcription Factors (TFs) from publicly available ENCODE Consortium TF ChIP-seq studies in human are included. immuneRegulation currently incorporates around 14 million gene region hits extracted from 1,945 TF ChIP-seq peaks datasets. Furthermore, GWAS catalog data within immuneRegulation repository is comprised of 67,230 published SNP-trait association information.

Getting Started (Performing Queries)

The landing page of immuneRegulation provides the interface to construct queries as displayed in Figure 1. Users can enter gene(s) of interest by providing the associated HUGO gene symbol(s) within the text-box located at the top of the interface. (In this version, users are allowed to enter multiple genes for trans-eQTL queries; whereas only one gene input is allowed for cis-eQTL or TF queries) Then, users can select study type on the left (HIPC: cis-eQTL or trans-eQTL; or non-HIPC: cis-eQTL or trans-eQTL; or Transcription Factors) and filter relevant datasets among the collection of datasets listed on the right. Users can look through the list of datasets by referring to basic information (sample size; summary; link to publication or study) provided adjacent to each dataset. After selecting one or more datasets, user can click the Submit Query button. For each study type (cis-eQTL; trans-eQTL; Transcription Factors), users will get different visuals to display the results for their queries, which are explained below. Users can perform an updated or a completely different query by clicking the Modify Query button and making the desired changes.

Querying cis-eQTL Data

Users can enter a single gene and select one or more relevant datasets to perform a cis-eQTL query. This opens our customized Integrative Genomics Viewer (IGV) browser (19) as shown in Figure 2; cis-eQTLs for each selected dataset are displayed in separate genome browser tracks with their associated p-values. Red circles indicate that the cis-eQTL appears in multiple tracks; blue circles indicate that the cis-eQTL appears only in one track; and grey circles represent cis-eQTLs for a different gene. User can click on any specific cis-eQTL and a pop-up dialog box displays more details (position, gene, p-value, SNP ID and information link) about it. Users can zoom in and out, and also move across the genomic region to navigate other information. Additionally, two more tracks are displayed below cis-eQTL tracks for further information and reference. GWAS Catalog track displays SNPs associated with any GWAS findings in the genomic region. Clicking on any specific GWAS finding opens a pop-up dialog box displaying further details, such as disease/trait and publication link,

regarding the GWAS entry. Genes track displays gene isoforms and provides a quick view of the relative locations of other information displayed in the genome browser.

Querying *trans*-eQTL Data

Users can enter gene(s) of interest and select one or more relevant datasets to perform a *trans*-eQTL query. If multiple genes are provided as input, user is given the option, via a dynamical visual element, to apply either intersection or union operation on the gene list. In case where the user selects the intersection operation, only those SNPs that act as *trans*-eQTLs for more than one gene are displayed as a table. However, if the union operation is selected for multiple genes or if a single gene is provided as input, results are visualized similar to Figure 3. Figure 3 Panel A displays a screenshot of the resulting interactive graph of all *trans*-eQTLs that regulate the query genes, rank-ordered and visualized based on their p-values; users can also view their genomic locations in a Manhattan plot (Panel B). Users can filter *trans*-eQTL results based on genomic locations of SNPs, as well as using a subset of genes. Black circles indicate *trans*-eQTLs that have associated GWAS hits; blue circles indicate *trans*-eQTLs without any GWAS hit associations. Hovering over any specific *trans*-eQTL finding provides a quick look into its details via a tooltip text. For deeper analyses, users can click on a specific *trans*-eQTL from either graph; a text-box as in Panel C appears to the right of the graph with details about the *trans*-eQTL and links to further associated information. Clicking on the first link with the down arrow opens up a table, as in Panel D, that displays other gene targets for the specific SNP associated with the *trans*-eQTL. Similarly, clicking on the second link with the down arrow opens up a table, as in Panel E, that displays GWAS hits associated with the specific SNP.

Querying Transcription-Factor (TF) Data

Users can enter a single gene and select the ENCODE collection of datasets to query TFs. This opens a table, Figure 4 Panel A, that displays all datasets that have gene region hits for the gene queried. Each row in the table corresponds to a cell type-TF target combination and it can be filtered to look for a specific cell type or TF target using the Search box. Users can explore and select the datasets of interest as shown in Panel B. Clicking the Submit button in Panel B opens our customized IGV browser as shown in Panel C. TF ChIP-seq peaks for each selected dataset are displayed in separate genome browser tracks. GWAS Catalog and Genes tracks are also displayed to provide further information about the genomic region.

Uploading and Visualizing Private Data

immuneRegulation provides capability for users to upload datasets of their own (*cis*-eQTL and *trans*-eQTL) and integrate with the rest of the interface. The important detail here is that the user datasets are not transferred to any remote server or third-party site; the operation is handled completely within the local browser interface of the user performing the upload. Users can utilize this functionality by clicking the Upload Study button in the main query browser interface. A pop-up window opens up with descriptions of the proper format for datasets that can be uploaded, and enables the users to upload one or more datasets (*cis*-eQTL and/or *trans*-eQTL). After the user saves the changes, those datasets are added to the query browser interface similar to as shown in Figure 5. These datasets are

completely integrated within immuneRegulation interface; users can browse and query them in combination with any existing datasets, and interact with the visual results in the same way.

DISCUSSION

To comprehensively understand heterogeneity in the healthy human immune system and in response to stimuli, multiple groups are generating high-throughput measurements of gene regulation in the human immune system. These massive data create their own challenges in accurately interpreting immunophenotyping data, necessitating easy to use data mining and interpretation tools. An important goal in these efforts is phenotyping the immune states and diseases for diagnosis, prognosis and selection of therapies. ImmuneRegulation helps investigators interpret multiple data resources in their efforts to generate hypotheses on how transcriptome differences observed in different members of the same cohort can be explained by specific regulators. While such regulators (genetic variants or transcription factors) can yield insights into the immune response to stimuli, and can improve sensitivity and accuracy of classification, often in the discovery phase it is difficult to assess their real biological role without extensive experimentation. Integrating multiple datasets with existing genotype and transcriptome data on a large number of humans and immune cell types in a single, user-friendly interface supports investigations on human immunity, infection and disease. The availability of reliably curated patient cohorts and the integrative interface of ImmuneRegulation that provides exploration of massive immune regulatory datasets can help accelerate our understanding of the human immune system. We anticipate that rapid, intuitive and integrated analyses within our interface will significantly empower immune researchers in their study of the complex role genomic regulation plays in immune phenotype responses and help the translation of the findings to insights and applications in basic and clinical immunology. Specifically, some of the differences observed in the phenotypic responses may be caused by inter-individual genetic variation, and identifying the regulatory elements that lead to differences in transcriptome responses within these cohorts can yield insights into the development, remission and exacerbation of disease; they may as well improve diagnostic sensitivity and accuracy of cohort classification, and ultimately guide treatment.

AVAILABILITY

ImmuneRegulation is an open source collaborative initiative available in the GitHub repository
<https://github.com/gumuslab-mssm/immuneRegulation>

ACKNOWLEDGEMENT

We thank Ana Fernandez-Sesma, Eun-Young Kim, Andrew Kasarskis, Steven Kleinstein, Chris Cotsapas, and Steven Wolinsky for their valuable feedback in helping us to better understand and address user needs in interacting with the tool.

FUNDING

This work was supported by the National Institutes of Health [AI118610-01 to Z.H.G.]. Funding for open access charge: National Institutes of Health.

CONFLICT OF INTEREST

No conflicts.

REFERENCES

1. Furman,D. and Davis,M.M. (2015) New approaches to understanding the immune response to vaccination and infection. *Vaccine*, 10.1016/j.vaccine.2015.06.117.
<https://doi.org/10.1016/j.vaccine.2015.06.117>
<http://www.ncbi.nlm.nih.gov/pubmed/26232539>
2. Poland,G.A., Quill,H. and Togias,A. (2013) Understanding the human immune system in the 21st century: The human immunology project consortium. *Vaccine*, 10.1016/j.vaccine.2013.04.043.
<https://doi.org/10.1016/j.vaccine.2013.04.043>
<http://www.ncbi.nlm.nih.gov/pubmed/23648045>
3. Voigt,E.A., Grill,D.E., Zimmermann,M.T., Simon,W.L., Ovsyannikova,I.G., Kennedy,R.B. and Poland,G.A. (2018) Transcriptomic signatures of cellular and humoral immune responses in older adults after seasonal influenza vaccination identified by data-driven clustering. *Sci. Rep.*, 10.1038/s41598-017-17735-x.
<https://doi.org/10.1038/s41598-017-17735-x>
<http://www.ncbi.nlm.nih.gov/pubmed/29335477>
4. Cahill,M.E., Conley,S., DeWan,A.T. and Montgomery,R.R. (2018) Identification of genetic variants associated with dengue or West Nile virus disease: A systematic review and meta-analysis. *BMC Infect. Dis.*, 10.1186/s12879-018-3186-6.
<https://doi.org/10.1186/s12879-018-3186-6>
<http://www.ncbi.nlm.nih.gov/pubmed/29929468>
5. Hamlin,R.E., Rahman,A., Pak,T.R., Maringer,K., Mena,I., Bernal-Rubio,D., Potla,U., Maestre,A.M., Fredericks,A.C., Amir,E.D., et al. (2017) High-dimensional CyTOF analysis of dengue virus–infected human DCs reveals distinct viral signatures. *JCI Insight*, 10.1172/jci.insight.92424.
<https://doi.org/10.1172/jci.insight.92424>
<http://www.ncbi.nlm.nih.gov/pubmed/28679950>
6. Lonsdale,J., Thomas,J., Salvatore,M., Phillips,R., Lo,E., Shad,S., Hasz,R., Walters,G., Garcia,F., Young,N., et al. (2013) The Genotype-Tissue Expression (GTEx) project. *Nat. Genet.*, **45**, 580–585.
<https://doi.org/10.1038/ng.2653>
<http://www.ncbi.nlm.nih.gov/pubmed/23715323>
7. Framingham Heart Study (2015) Bibliography | Framingham Heart Study. *Fram. Hear. Study*.
8. Dunham,I., Kundaje,A., Aldred,S.F., Collins,P.J., Davis,C.A., Doyle,F., Epstein,C.B., Frietze,S., Harrow,J., Kaul,R., et al. (2012) An integrated encyclopedia of DNA elements in the human genome. *Nature*, 10.1038/nature11247.

<https://doi.org/10.1038/nature11247>
<http://www.ncbi.nlm.nih.gov/pubmed/22955616>

9. Myers,R.M., Stamatoyannopoulos,J., Snyder,M., Dunham,I., Hardison,R.C., Bernstein,B.E., Gingeras,T.R., Kent,W.J., Birney,E., Wold,B., *et al.* (2011) A user's guide to the Encyclopedia of DNA elements (ENCODE). *PLoS Biol.*, 10.1371/journal.pbio.1001046.

<https://doi.org/10.1371/journal.pbio.1001046>
<http://www.ncbi.nlm.nih.gov/pubmed/21526222>

10. Birney,E., Stamatoyannopoulos,J.A., Dutta,A., Guigó,R., Gingeras,T.R., Margulies,E.H., Weng,Z., Snyder,M., Dermitzakis,E.T., Thurman,R.E., *et al.* (2007) Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. *Nature*, 10.1038/nature05874.

<https://doi.org/10.1038/nature05874>
<http://www.ncbi.nlm.nih.gov/pubmed/17571346>

11. Welter,D., MacArthur,J., Morales,J., Burdett,T., Hall,P., Junkins,H., Klemm,A., Fliceck,P., Manolio,T., Hindorff,L., *et al.* (2014) The NHGRI GWAS Catalog, a curated resource of SNP-trait associations. *Nucleic Acids Res.*, **42**.

<https://doi.org/10.1093/nar/gkt1229>
<http://www.ncbi.nlm.nih.gov/pubmed/24316577>

12. Gilad,Y., Rifkin,S.A. and Pritchard,J.K. (2008) Revealing the architecture of gene regulation: the promise of eQTL studies. *Trends Genet.*, 10.1016/j.tig.2008.06.001.

<https://doi.org/10.1016/j.tig.2008.06.001>
<http://www.ncbi.nlm.nih.gov/pubmed/18597885>

13. Kadonaga,J.T. (2004) Regulation of RNA Polymerase II Transcription by Sequence-Specific DNA Binding Factors. *Cell*, 10.1016/S0092-8674(03)01078-X.

[https://doi.org/10.1016/S0092-8674\(03\)01078-X](https://doi.org/10.1016/S0092-8674(03)01078-X)
<http://www.ncbi.nlm.nih.gov/pubmed/14744435>

14. Dobin,A., Davis,C.A., Schlesinger,F., Drenkow,J., Zaleski,C., Jha,S., Batut,P., Chaisson,M. and Gingeras,T.R. (2013) STAR: Ultrafast universal RNA-seq aligner. *Bioinformatics*, 10.1093/bioinformatics/bts635.

<https://doi.org/10.1093/bioinformatics/bts635>
<http://www.ncbi.nlm.nih.gov/pubmed/23104886>

15. McKenna,A., Hanna,M., Banks,E., Sivachenko,A., Cibulskis,K., Kernytsky,A., Garimella,K., Altshuler,D., Gabriel,S., Daly,M., *et al.* (2010) The genome analysis toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.*, **20**, 1297–1303.

<https://doi.org/10.1101/gr.107524.110>
<http://www.ncbi.nlm.nih.gov/pubmed/20644199>

16. Sloan,C.A., Chan,E.T., Davidson,J.M., Malladi,V.S., Strattan,J.S., Hitz,B.C., Gabdank,I., Narayanan,A.K., Ho,M., Lee,B.T., *et al.* (2016) ENCODE data at the ENCODE portal. *Nucleic Acids Res.*, **44**, D726–D732.

<https://doi.org/10.1093/nar/gkv1160>
<http://www.ncbi.nlm.nih.gov/pubmed/26527727>

17. Landt,S.G., Marinov,G.K., Kundaje,A., Kheradpour,P., Pauli,F., Batzoglou,S., Bernstein,B.E., Bickel,P., Brown,J.B., Cayting,P., *et al.* (2012) ChIP-seq guidelines and practices of the ENCODE and modENCODE

consortia. *Genome Res.*, **22**, 1813–1831.
<https://doi.org/10.1101/gr.136184.111>
<http://www.ncbi.nlm.nih.gov/pubmed/22955991>

18. Quinlan,A.R. and Hall,I.M. (2010) BEDTools: A flexible suite of utilities for comparing genomic features. *Bioinformatics*, 10.1093/bioinformatics/btq033.

<https://doi.org/10.1093/bioinformatics/btq033>
<http://www.ncbi.nlm.nih.gov/pubmed/20110278>

19. Robinson,J.T., Thorvaldsdóttir,H., Winckler,W., Guttman,M., Lander,E.S., Getz,G. and Mesirov,J.P. (2011) Integrative genomics viewer. *Nat. Biotechnol.*, **29**, 24–26.

<https://doi.org/10.1038/nbt.1754>
<http://www.ncbi.nlm.nih.gov/pubmed/21221095>

TABLE AND FIGURES LEGENDS

Table 1. List of eQTL datasets provided in immuneRegulation.

Figure 1. Main visual interface for performing queries.

Figure 2. cis-eQTL results for two different datasets displayed within the customized IGV browser. GWAS Catalog and Genes tracks are also displayed by default.

Figure 3. All trans-eQTL results from query **A** p-value ordered; **B** Manhattan plot; For a specific trans-eQTL **C** details table; **D** eQTL target genes table; **E** GWAS results table.

Figure 4. Transcription Factor gene hits results **A** table sorted by cell type; **B** table sorted by target, and visualization list is populated; **C** studies added to the list visualized within the customized IGV browser.

Figure 5. Query interface displaying user uploaded datasets.

Table 1.

List of eQTL studies included in immuneRegulation

Public Studies

Study	Cell type	# of individuals	Ethnicity	Cis or Trans	Metastudy	Link
Framingham Consortium	Whole Blood	5,257	European	Cis	No	https://www.framinghamheartstudy.org/
GTEx Consortium	Whole Blood	338	Mostly European	Cis; Trans	Yes	https://gtexportal.org/home/
Westra	Peripheral Blood	5,311	European	Cis; Trans	No	http://genenetwork.nl/bloodeqtlbrowser/
Jager	Monocytes	162	African American	Cis; Trans	No	http://science.sciencemag.org/content/344/6183/519
Jager	Monocytes	155	East Asian	Cis	No	http://science.sciencemag.org/content/344/6183/520
Jager	Monocytes	377	European	Cis	No	http://science.sciencemag.org/content/344/6183/521
Jager	Monocytes	696	Mixed	Cis	No	http://science.sciencemag.org/content/344/6183/522
Fairfax	Monocytes	288	European	Cis; Trans	No	http://www.nature.com/ng/journal/v44/n5/full/ng.2205.html
Fairfax	Monocytes	432	European	Cis; Trans	No	http://science.sciencemag.org/content/343/6175/1246949
Jager	CD4+ T-cells	162	African American	Cis	No	http://science.sciencemag.org/content/344/6183/519
Jager	CD4+ T-cells	155	East Asian	Cis	No	http://science.sciencemag.org/content/344/6183/519
Jager	CD4+ T-cells	377	European	Cis	No	http://science.sciencemag.org/content/344/6183/519
Jager	CD4+ T-cells	696	Mixed	Cis; Trans	Yes	http://science.sciencemag.org/content/344/6183/519
Fairfax	B-cells	288	European	Cis; Trans	No	http://www.nature.com/ng/journal/v44/n5/full/ng.2205.html
Naranbhai	CD16+ Neutrophils	101	European	Cis; Trans	No	https://www.nature.com/articles/ncomms8545
Fairfax	Monocytes: LPS exposure, 2hr	432	European	Cis; Trans	No	http://science.sciencemag.org/content/343/6175/1246949
Fairfax	Monocytes: LPS exposure, 24hr	433	European	Cis; Trans	No	http://science.sciencemag.org/content/343/6175/1246949
Fairfax	Monocytes: INF γ exposure	434	European	Cis; Trans	No	http://science.sciencemag.org/content/343/6175/1246949
Lude	Peripheral Blood	2,116	European	Cis	No	https://www.nature.com/articles/ng.3737
Heijmans - meQTL	Whole Blood	3,841	European	Cis; Trans	No	https://www.nature.com/articles/ng.3721
Heijmans - eQTM	Whole Blood	2,101	European	Cis	No	https://www.nature.com/articles/ng.3721
Pierce	Peripheral Blood	1,799	East Asian	Cis; Trans	No	http://journals.plos.org/plosgenetics/article?id=10.1371/journal.pgen.1004818
Yamada	Peripheral Blood	298	East Asian	Cis; Trans	No	http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0100924
Dermitzakis	lymphoblastoid cell lines	373	European	Cis	No	https://www.nature.com/articles/nature12531
Dermitzakis	lymphoblastoid cell lines	89	African	Cis	No	https://www.nature.com/articles/nature12531

HIPC Studies

Study	Cell type	# of individuals	Ethnicity	Cis or Trans	Metastudy	Study Type	Timepoints
HIPC - Emory	Peripheral Blood	55	Mixed	Cis; Trans	No	Influenza vaccination	(Day): 0, 1, 3, 7, 14
HIPC - Mayo	Peripheral Blood	120	Mixed	Cis; Trans	No	Influenza vaccination	(Day): 0, 28
HIPC - Yale	Peripheral Blood	69	Mixed	Cis; Trans	No	Influenza vaccination	(Day): -0, 2, 7, 28
HIPC - NIH	Peripheral Blood	57	Mixed	Cis; Trans	No	Influenza vaccination	(Day): -7, 0, 1, 7, 70
HIPC - DHIPC	Peripheral Blood	43	Hispanic or Latino	Cis; Trans	No	Chikungunya infection	Acute stage (2-4 days), Convalescent stage (14-17 days)

Enter Gene(s):

Enter HUGO Gene symbols

Study List:

0 studies selected

Study Type: Not Selected

↓ HIPC Studies 22

↳ cis-eQTL 11

↳ trans-eQTL 11

↓ Non-HIPC Studies 37

↳ cis-eQTL 23

↳ trans-eQTL 14

↓ Transcription Factors 1945

Non-HIPC Studies

Whole Blood

- Whole Blood - European American (Framingham Heart Study, NHLBI) 5257 samples  
- Whole Blood - Mixed ethnicity - GTEx Consortium (Science, 2015) 338 samples  
- Whole Blood - Meta-analysis - European (Nature Genetics, 2013) 5311 samples  
- Whole Blood - European (Nature Genetics, 2017) 2166 samples  
- Whole Blood - East Asian (PLoS Genetics, 2014) 1799 samples  
- Whole Blood - East Asian (PLoS One, 2014) 298 samples  

Monocytes

Have private studies:

Upload Study

Submit Query

IGV

chr9:4,938,257-5,010,725

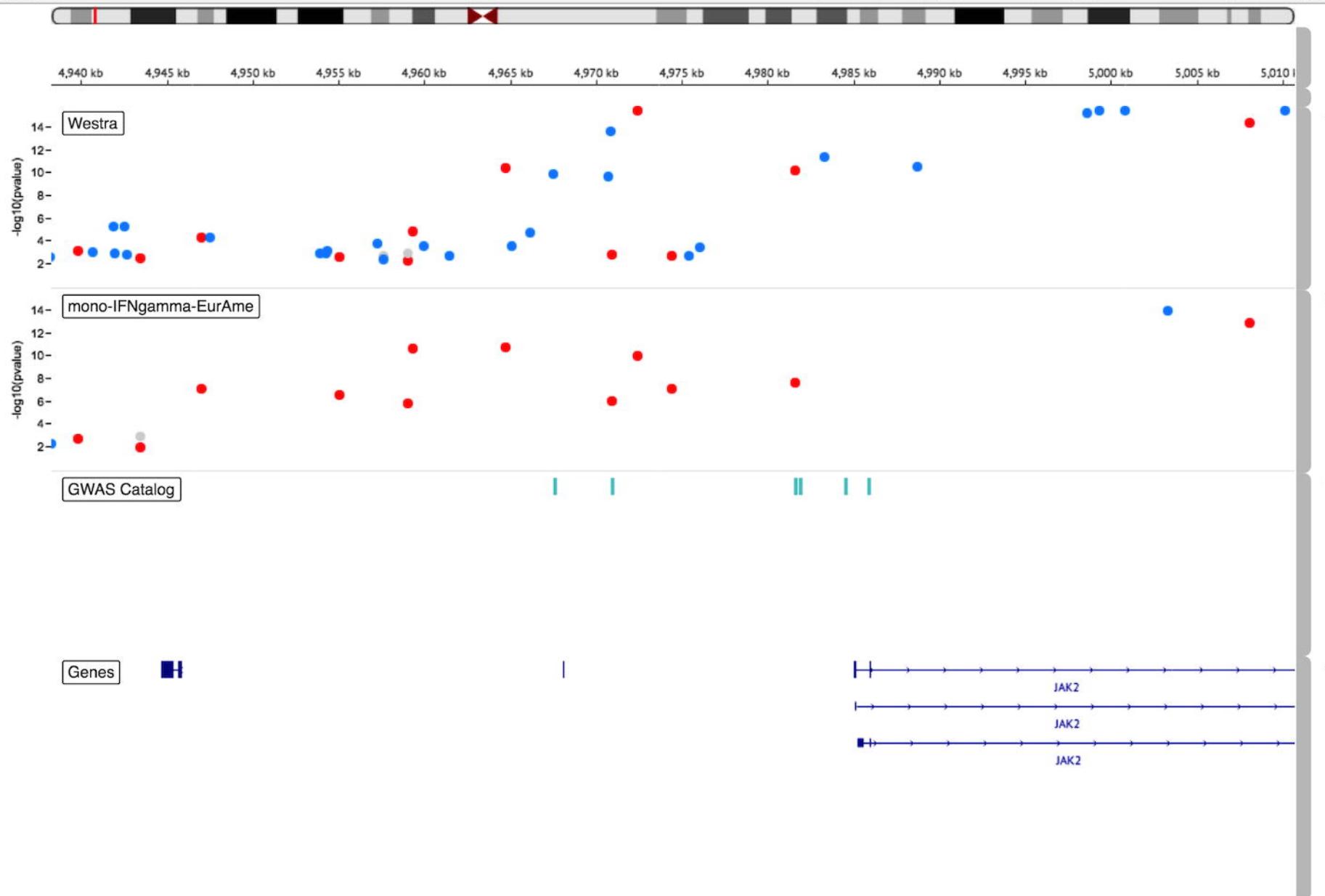


72 kb

Hide Labels

Center Line

Show Cursor Guide



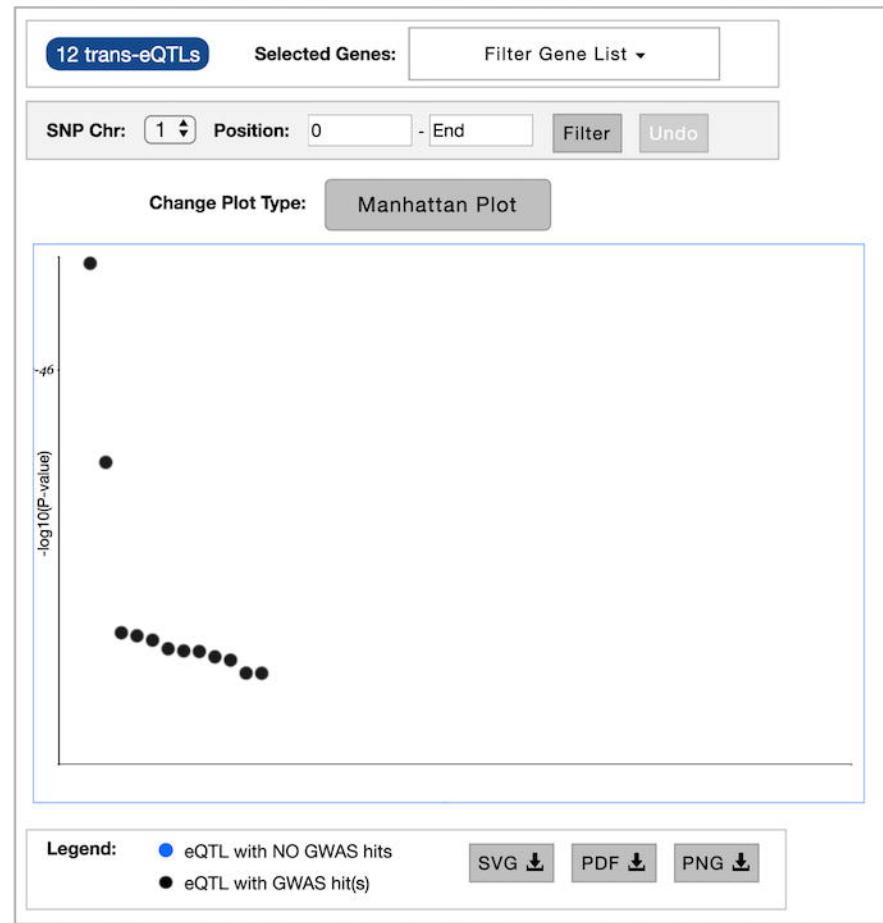
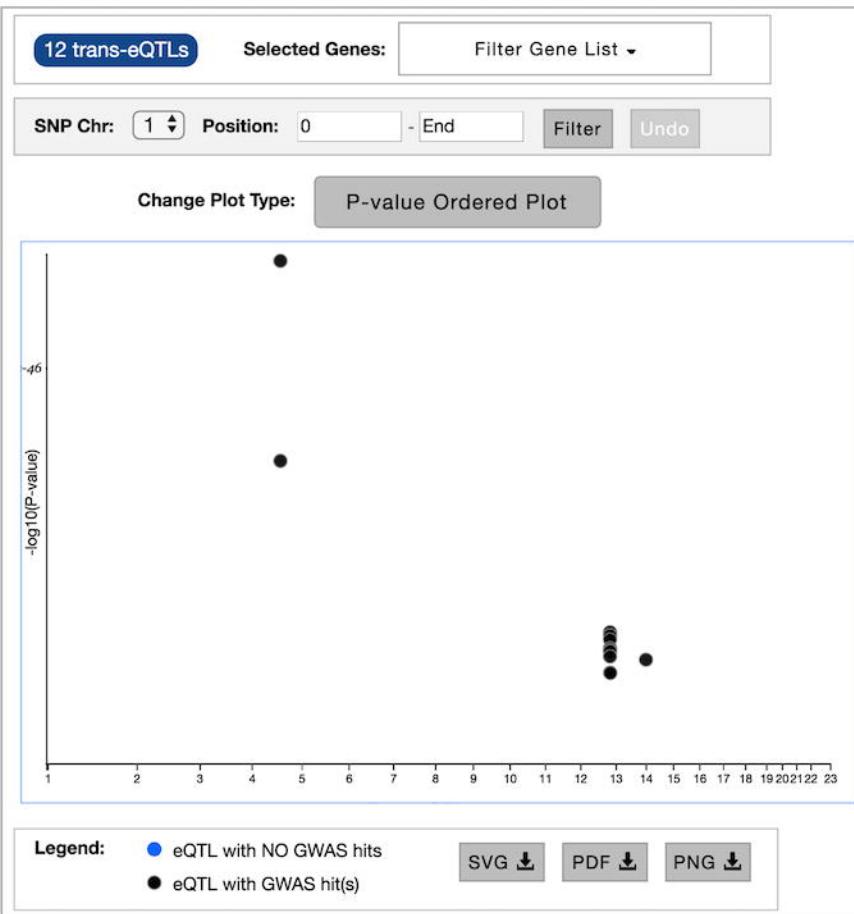
Legend:

- cis-eqtl for the queried gene common in multiple tracks
- cis-eqtl for the queried gene in a single track
- cis-eqtl for a different gene in the genomic region

SVG

PDF

PNG

A**B****C**

Trans-eQTL Details:

P-value: 2.2565e-16
 SNP-Pos: 4:103797685
 Gene: LTA Ensembl
 Gene-Pos: 6:31650043
 rsid: rs228614 dbSNP
 Other Gene Targets for: 'rs228614'
 GWAS Results (2) for: 'rs228614'

D

Show 10 entries Search:

Table: eQTL Target Genes

Gene	P-Value	Chr	Center-Pos
COL9A2	2.6773e-10	1	40538988
LTA	2.2565e-16	6	31650043
OPTN	6.0855e-10	10	13220141

Showing 1 to 3 of 3 entries Previous 1 Next

E

Show 10 entries Search:

Table: GWAS Results

Disease/Phenotype	P-Value	Odds-ratio/beta	Pubmed ID
Multiple sclerosis	1E-7	1.09	21833088
Primary sclerosing cholangitis	1E-7	1.13	27992413

Showing 1 to 2 of 2 entries Previous 1 Next

Target Gene: IL17A

23 total hits

8 cell types

19 targets

A

Show 10 entries

Search:

Table: Transcription Factor ChIP-Seq Studies

Select	Cell Type	Transcription Factor	Study Details
Select	A549	MAFK	View: ENCODE
Select	A549	NFE2L2	View: ENCODE
Select	GM12874	CTCF	View: ENCODE
Select	GM12878	SPI1	View: ENCODE
Select	GM12878	MEF2B	View: ENCODE
Select	GM12878	ATF2	View: ENCODE
Select	GM12878	RUNX3	View: ENCODE
Select	GM12878	NFIC	View: ENCODE
Select	GM12878	RAD51	View: ENCODE
Select	GM12878	EP300	View: ENCODE

Showing 1 to 10 of 23 entries

Previous [1](#) [2](#) [3](#) Next

[Submit](#)

Target Gene: IL17A

23 total hits:

8 cell types

19 targets

Show 10 entries

Search:

Table: Transcription Factor ChIP-Seq Studies

Select	Cell Type	Transcription Factor	Study Details
Select	GM12878	MEF2A	View: ENCODE
Select	GM12878	MEF2B	View: ENCODE
Select	K562	MITF	View: ENCODE
Select	GM23338	NANOG	View: ENCODE
Deselect	A549	NFE2L2	View: ENCODE
Deselect	HepG2	NFE2L2	View: ENCODE
Deselect	IMR-90	NFE2L2	View: ENCODE
Select	GM12878	NFIC	View: ENCODE
Select	GM12878	RAD51	View: ENCODE
Select	GM12878	RUNX3	View: ENCODE

Showing 11 to 20 of 23 entries

Previous [1](#) [2](#) [3](#) Next

[Submit \(3\)](#)

C

IGV

chr6:52,050,185-52,056,436



6,252 bp

[Hide Labels](#)

[Center Line](#)

[Show Cursor Guide](#)



52,050,500 52,051,000 52,051,500 52,052,000 52,052,500 52,053,000 52,053,500 52,054,000 52,054,500 52,055,000 52,055,500 52,056,000 52,056,436

Cell: A549
Target: NFE2L2

Cell: IMR-90
Target: NFE2L2

Cell: HepG2
Target: NFE2L2

GWAS Catalog

Genes

IL17A

Enter Gene(s):

Enter HUGO Gene symbols

Study List:

↓ User Studies

2

↳ cis-eQTL

1

↳ trans-eQTL

1

↓ HIPC Studies

22

↳ cis-eQTL

11

↳ trans-eQTL

11

↓ Non-HIPC Studies

37

0 studies selected

Study Type: Not Selected

User Studies

- myCisData.txt - (Cis)
- myTransData.txt - (Trans)

Non-HIPC Studies

Whole Blood

- Whole Blood - European American (Framingham Heart Study, NHLBI) 5257 samples  
- Whole Blood - Mixed ethnicity - GTEx Consortium (Science, 2015) 338 samples  
- Whole Blood - Meta-analysis - European (Nature Genetics, 2013) 5311 samples  

Have private studies: [Upload Study](#)

[Submit Query](#)