Identifying Cell Type-Specific Transcription Factors by Integrating ChIP-seq and eQTL Data–Application to Monocyte Gene Regulation

Mudra Choudhury and Stephen A. Ramsey

Department of Biomedical Sciences, Oregon State University, Corvallis, OR, USA.

ABSTRACT: We describe a novel computational approach to identify transcription factors (TFs) that are candidate regulators in a human cell type of interest. Our approach involves integrating cell type-specific expression quantitative trait locus (eQTL) data and TF data from chromatin immunoprecipitation-to-tag-sequencing (ChIP-seq) experiments in cell lines. To test the method, we used eQTL data from human monocytes in order to screen for TFs. Using a list of known monocyte-regulating TFs, we tested the hypothesis that the binding sites of cell type-specific TF regulators would be concentrated in the vicinity of monocyte eQTLs. For each of 397 ChIP-seq data sets, we obtained an enrichment ratio for the number of ChIP-seq peaks that are located within monocyte eQTLs. We ranked ChIP-seq data sets according to their statistical significances for eQTL overlap, and from this ranking, we observed that monocyte-regulating TFs are more highly ranked than would be expected by chance. We identified 27 TFs that had significant monocyte enrichment scores and mapped them into a protein interaction network. Our analysis uncovered two novel candidate monocyte-regulating TFs, BCLAF1 and SIN3A. Our approach is an efficient method to identify candidate TFs that can be used for any cell/tissue type for which eQTL data are available.

KEYWORDS: eQTL, transcription factor, ChIP-seq

CITATION: Choudhury and Ramsey. Identifying Cell Type-Specific Transcription Factors by Integrating ChIP-seq and eQTL Data–Application to Monocyte Gene Regulation. *Gene Regulation and Systems Biology* 2016:10 105–110 doi: 10.4137/GRSB.S40768.

TYPE: Methodology

RECEIVED: August 22, 2016. **RESUBMITTED:** November 03, 2016. **ACCEPTED FOR PUBLICATION:** November 06, 2016.

ACADEMIC EDITOR: James Willey, Editor in Chief

PEER REVIEW: Five peer reviewers contributed to the peer review report. Reviewers' reports totaled 1139 words, excluding any confidential comments to the academic editor.

FUNDING: This work was supported by the US National Institutes of Health (award HL098807 to SAR), the Medical Research Foundation of Oregon (New Investigator Grant award to SAR), Oregon State University (Division of Health Sciences Interdisciplinary Research Grant award to SAR, University Honors College DeLoach Work Scholarship to MC, the Undergraduate Research, Innovation, Scholarship & Creativity Grant to MC, and the Honors Experiential Award to MC), and the National Science Foundation (awards 1557605-DMS and 1553728-DBI to SAR). The authors confirm that the funder had no influence over the study design, content of the article, or selection of this journal.

COMPETING INTERESTS: Authors disclose no potential conflicts of interest.

CORRESPONDENCE: stephen.ramsey@oregonstate.edu

Introduction

Mapping human cell type-specific gene regulatory networks is a significant and longstanding challenge due to the large number (>1,200) of transcription factors (TFs) and the expansiveness of noncoding regions of the genome.[1] Expression quantitative trait locus (eQTL) studies, in which expression single nucleotide polymorphisms (eSNPs) that are statistically associated with population variation in transcript abundance are systematically mapped, are an increasingly used tool for revealing functional regulatory regions in cell types of interest.[2] The genomic regions that are in linkage disequilibrium (LD) with eSNPs include cell type-specific enhancers (and other types of cis-regulatory elements) that contain TF binding sites (TFBS). However, in order to map cell type-specific regulatory regions to specific TFs, additional information is needed. One approach is to bioinformatically scan–within regulatory regions–genomic sequence for matches to known TFBS sequence patterns[3,4]; however, this approach has a significant false-positive rate[5] and TFBS sequence preferences may vary depending on the cellular context.[6] Following the recent availability of published binding site locations (based on chromatin immunoprecipitation-to-tag-sequencing [ChIP-seq]) for 133 TFs in various human cell lines through the Encyclopedia of DNA Elements (ENCODE) project,[7] we devised a novel method to directly analyze eQTL regulatory regions using ChIP-seq-based cellular TFBS measurements. Conceptually, our approach builds on previous efforts in which eQTL data were leveraged in motif-based identification of TFBS[8] and in identifying genomic correlates of regulatory SNPs.[9] We applied the method to search for TF regulators using eQTLs from two population studies of human monocytes.[10,11] Monocytes are innate immune cells that are vital in host defense, have important roles in many infectious and chronic inflammatory diseases, and are of research interest as models of cellular gene regulatory responses to microbial challenge.[12–15]

We hypothesized that binding sites of monocyte-regulating TFs are more enriched within monocyte eQTLs than the TFBS that are not monocyte regulators. To investigate this hypothesis, we ranked ENCODE-profiled TFBS

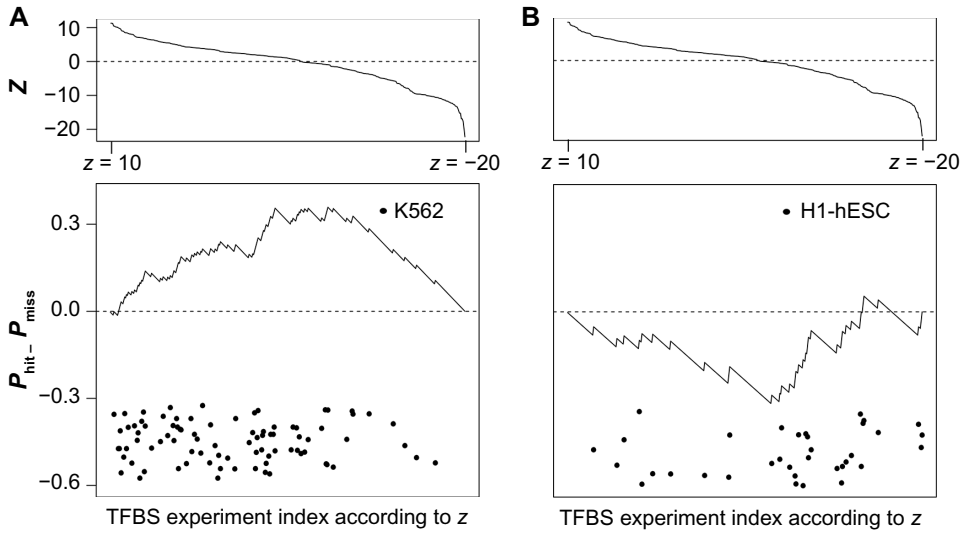

Figure 1. K562 cells are biased toward higher *z* among ChIP-seq experiments ranked by *z*-score measure of overlap with human monocyte eQTLs. (**A**) GSEA plot of K562 data sets ($E = 0.359$) with distribution of overlap *z*-scores shown above the GSEA plot. (**B**) GSEA plot of H1-hESC data sets ($E = -0.316$) with distribution of overlap *z*-scores shown above the GSEA plot (Supplementary Table 1 for a complete list of the 397 ChIP-seq experiments whose data were analyzed, including cell type and TF target for each experiment). Horizontal coordinates of the data points show rank (*z*) for ChIP-seq data sets for the indicated cell line, with lowest rank corresponding to highest *z*. Here, consistent with the definition in the original GSEA study,[17] P_{hit} is the cumulative density function (CDF) of an indicator function that is 1 for a K562 data set and 0 for a non-K562 data set, and P_{miss} is the CDF for an indicator function that is 0 for a K562 data set and 1 for a non-K562 data set.

ChIP-seq experiments based on the degrees of overlap of their genome-wide binding site profiles (in human cell lines) with monocyte eQTLs. We then tested the enrichment of myeloid cell-derived data sets in the ranked list and the enrichment of known monocyte-regulating TFs in the resulting list of TFs. We found that TFs that are known to have regulatory functions in monocytes are concentrated toward significant *P* values and are well connected to other candidate monocyte TFs in a protein interaction network. Our study uncovered two candidate monocyte-regulating TFs, BCLAF1 and SIN3A, which are previously known to have functions related to immunity.

Results and Discussion

To investigate whether anchoring the analysis of TFs using monocyte eQTL data would enable prioritizing TFBS data sets by their relevance to monocytes, we ranked the 397 ChIP-seq-derived TF or cofactor binding site (hereafter, "TFBS") location data sets based on their overlap (as measured by a *z*-score of the number of TFBS within the eQTLs; z was computed using a background model based on genome-wide-randomized eQTL placement; "Methods" section) with eQTL regions defined by LD with 80,000 monocyte eSNPs. Then, for two cell types, a cell type related to monocytes (the human myelogenous leukemia cell line K562[16]) and a negative control cell type (the human embryonic stem cell line H1-hESC), we measured the degree to which the data sets of the cell types are enriched at low ranks (corresponding to high overlap with monocyte eQTLs), using the GSEA statistical test.[17] The ChIP-seq data sets from K562 cells were biased toward low ranks (corresponding to high overlap with monocyte eQTLs), with a GSEA score $E = 0.359$ (Fig. 1A; $E \neq 0$ at $P < 0.001$); by comparison, the ChIP-seq data sets from H1-hESC cells were biased toward higher ranks (corresponding to low overlap with monocyte eQTLs), with $E = -0.316$ (Fig. 1B). From this analysis, we concluded that our method enables the prioritization of cell type-specific TFBS data sets based on eQTL data. Next, to investigate the power of our method to uncover specific TFs that regulate gene expression in a specific cell type, for each TFBS data set, we scored the significance of the number of TFBS that were within eQTLs using a one-tailed binomial test in which the background probability is based on the genome-wide-randomized background model ("Methods" section). We adjusted the 397 resulting *P* values for multiple hypothesis testing and for each of the 133 unique TF types that we analyzed (Supplementary Table 1), we selected the TFBS ChIP-seq experiment for that TF that had the largest effect size (enrichment ratio) based on the binomial test.[a] This analysis produced a list of 133 TFs that we ranked by adjusted *P* value (P_{adj}) (Supplementary Table 2). We hypothesized that TFs that regulate gene expression in monocytes would be biased toward low ranks. To test this hypothesis, we reviewed the literature in order to compile a list of 71 TFs that are known to regulate gene expression in monocytes or monocyte-derived immune cells (see Supplementary Table 3), of which 26 TFs

[a]It is acknowledged that this selection criterion will somewhat bias the P_{adj} values; however, we note that the highest P_{adj} among the top 20% of the TFs is less than 10, and thus, there is little risk of a false rejection of the null hypothesis assuming the binomial test assumptions hold.

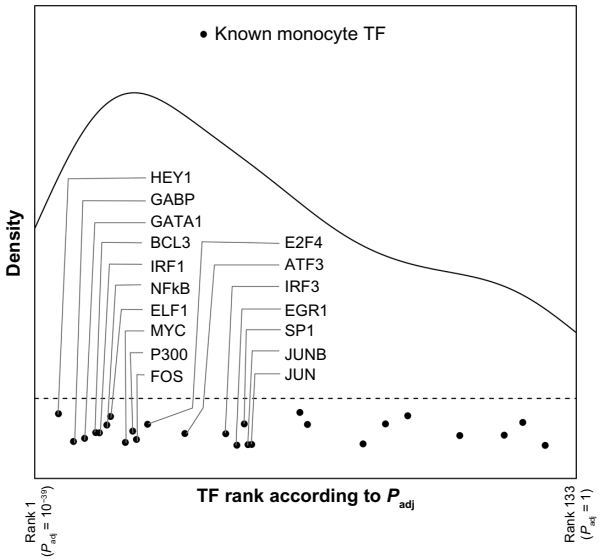

Figure 2. TFs with known functions in regulating monocyte gene expression have higher levels of TFBS overlap with monocyte eQTLs than would be expected by chance. Marks indicate the ranks of TFs with known functions in monocyte gene regulation among all TFs ranked by eQTL overlap significance (P_{adj}). Solid line indicates the kernel density estimate from the ranks of TFs that have known functions in monocyte gene regulation (density ratio between peak and lowest density value shown is ~1.57).

were among the 133 TFs profiled by ENCODE. We then tested for enrichment of the known monocyte-regulating TFs at low ranks in the list, using Gene Set Enrichment Analysis (GSEA).[17] We found that the known monocyte-regulating TFs are strongly biased toward low ranks (Fig. 2; $E = 0.329$; $P < 0.01$).

In order to focus on the most probable monocyte-regulating TFs, we selected the top 20% of TFs ranked by P_{adj} ($P_{adj} \leq 5 \times 10^{-7}$). This list (Table 1) includes both 11 known monocyte-regulating TFs and 16 novel candidate monocyte-regulating TFs (NMTFs), which are highly interconnected within the protein interaction/coexpression network (Fig. 3). The network appears to be enriched for TF pairs that cooperatively regulate gene expression (e.g., ELF1–AP1,[18] IRF1–NF-κB,[19] YY1–Myc[20]).

Several of the NMTFs have functions related to immunity or interact with previously identified monocyte-regulating TFs, supporting their potential roles in monocyte gene regulation. For example, the NMTF with the smallest P_{adj}, BCLAF1, is highly expressed in various immune cell types (based on data from ImmGen and BioGPS; "Methods" section), and it plays a crucial role in signal transduction between NF-κB and C/EBPβ (both of which are known to regulate gene expression in monocytes) in multiple cell types.[22] Another significant NMTF, SIN3A, plays a role (along with the known monocyte-regulating TF, Myc) in a signaling circuit implicated in the pathogenesis of acute myeloid leukemia.[23] The NMTFs TAF1 and TAF7 may function together in monocytes, as they are coexpressed[24] and are both components of the transcription factor II D (TFIID) complex. TAF7 is highly expressed in immune cells and in monocytes in particular (ImmGen and BioGPS). The NMTF SP2 (along with other known monocyte-regulating TFs, STAT1, STAT3, IRF1, and EGR1) is among the top 10 TFs associated with response to parasitic infection.[25] The NMTF YY1 regulates the activity of the promoter of murine IRF3, which plays a key role in inflammation and immune response.[26]

THAP1, ETS1, TR4, and NRF1 have been previously identified to have specific immune regulatory functions. THAP1 is known to play a role in inducing T-cell apoptosis.[27] STAT3 is a known regulator of ETS1 in inflammation control in mouse macrophages.[28] In addition, a TR4–CD36 pathway is used by mouse miR-133a to regulate lipid accumulation in macrophages.[29] One general transcription factor, TBP whose transcript abundance is thought to be stable in human monocytes,[30] is also in the top 20%; one possible mechanism

Table 1. Top 20% of TFs or cofactors according to P_{adj} for TFBS overlap with monocyte eQTLs.

RANK	TRANSCRIPTION FACTOR	ENTREZ GENE ID	P_{adj}	RATIO
1	**HEY1**	23462	1.37×10^{-38}	1.24
2	BCLAF1	9774	4.10×10^{-35}	1.65
3	SIN3A	25942	7.17×10^{-35}	1.34
4	TAF1	6872	4.45×10^{-27}	1.37
5	**GABP/GABPA**	2551	3.19×10^{-23}	1.40
6	TBP	6908	2.10×10^{-22}	1.24
7	GTF2F1	2962	1.97×10^{-21}	1.52
8	**GATA1**	2623	2.91×10^{-18}	1.13
9	ELK4	2005	3.07×10^{-17}	1.30
10	MXI1	4601	6.11×10^{-17}	1.30
11	**BCL3**	602	3.26×10^{-16}	1.59
12	**IRF1**	3659	1.49×10^{-15}	1.18
13	**NFKB/NFKB1**	4790	8.35×10^{-14}	1.14
14	**ELF1**	1997	2.02×10^{-13}	1.14
15	SP2	6668	2.06×10^{-13}	1.33
16	NRF1	4899	8.30×10^{-13}	1.36
17	GTF2B	2959	1.04×10^{-12}	1.40
18	**MYC**	4609	2.19×10^{-12}	1.28
19	YY1	7528	1.53×10^{-11}	1.32
20	**P300/EP300**	2033	2.26×10^{-10}	1.62
21	**FOS**	2353	3.89×10^{-10}	1.38
22	THAP1	55145	1.41×10^{-09}	1.29
23	ETS1	2113	2.81×10^{-09}	1.27
24	**E2F4**	1874	3.06×10^{-09}	1.31
25	TAF7	6879	3.53×10^{-08}	1.28
26	TR4/NR2C2	7182	1.03×10^{-07}	1.58
27	E2F6	1876	4.15×10^{-07}	1.21

Note: Bolded names are TFs that have been previously identified to have regulatory functions in monocytes (Supplementary Table 3).

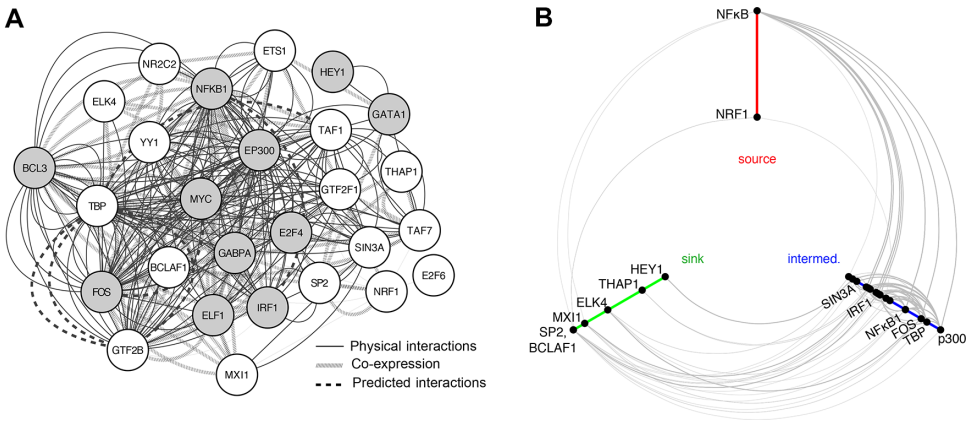

Figure 3. Interaction network of TFs and cofactors identified by overlap analysis with human monocyte eQTLs. (**A**) Graphical representation of the network. Each node represents a TF (Table 1); shading denotes a known role for the TF in monocyte gene regulation (Supplementary Table 3). Edges denote interactions or coexpression ("Methods" section) as indicated in the legend. A dashed edge means that the two TFs are predicted (by the GeneMANIA database) to physically interact ("predicted interactions"). (**B**) Hive plot[21] representation of the network. TFs are arranged into "source" (red), "intermediate" (blue), and "sink" (green) axes based on the graph topology. Edge thickness represents the interaction type (thick arc = physical interaction; medium = predicted interaction; thin = coexpression). For example, in the hive plot, NF-κB is identified as an upstream regulator in the monocyte gene regulatory network.

could be TFIID recruitment to TATA-less promoters by cell type-specific regulators.[31]

Discussion and Conclusions

We have demonstrated an efficient method to identify candidate TFs that regulate gene expression in a specific cell type of interest, by leveraging publicly available human genome-wide location data sets and eQTL data sets. A key advantage of this approach is that it can leverage the extensive array of publicly accessible ChIP-seq data from the ENCODE project (1,485 TF ChIP-seq data sets as of the current date). We note that one limitation of the approach is that human ChIP-seq data sets for many TFs are at present primarily available from cancerous cell lines,[7] whose gene regulatory networks may be altered compared to noncancerous cells of the same lineage type. Nevertheless, the strong result for enrichment of known monocyte TF regulators in the list of TFs ranked using our method suggests that this method is useful to screen for TF regulators for noncancerous cell types. Undoubtedly, confirmatory evidence regarding a specific function for a candidate TF in a given cell type of interest, (such as gene expression as we used in this work), would have an important role in avoiding false-positive identification of a TF due to a cancer cell line-specific artifact.

From our analysis using monocyte eQTL data, we conclude that TFs with known functions in regulating gene expression in monocytes tend to have binding sites located in the vicinity of monocyte eQTLs. Of the NMTFs that were identified in our analysis, the two most significant TFs were BCLAF1 and SIN3A. As revealed by protein network reconstruction, the top 20% of TFs identified by our approach are highly interconnected. Further targeted experiments would be useful to elucidate their specific roles in monocyte gene regulation. The approach is generally applicable to other cell types or tissues in that it could be used to identify candidate TFs for any cell type or tissue type for which eQTLs have been mapped from human population studies.

Methods

Software: We used the R statistical computing software (v3.2.2) and Bioconductor (v3.2) packages on OS X 10.11.4.

Databases. We obtained 397 TFBS location data files (derived from ChIP-seq analysis of 133 TFs in 68 human cell lines) from the public repository of the ENCODE project[7] via the University of California Santa Cruz Genome Browser portal. Additionally, we obtained data files containing eSNPs and associated eQTL *P* values from online supplementary data for two large-cohort monocyte eQTL studies involving 1,773 unrelated individuals.[10,11] We carried out all genome analyses using coordinates from the GRCh37/hg19 human reference genome assembly. We identified mouse orthologs expressed in monocytes and other immune cell types of the top TFs using the Immunological Genome (ImmGen) project database.[32] We used the gene annotation portal, BioGPS,[33] to obtain human microarray-based gene expression measurements for the top-ranked TFs.

eQTL intervals. We mapped genomic intervals containing SNPs that are in LD with each eSNP using the SNP Annotation and Proxy (SNAP) tool[34] using haplotype data from the 1,000 Genomes Project,[35] including only SNPs with $R^2 \geq 1.0$ and distance ≤10 kbp.

TFBS/eQTL overlap count significance testing. We created a set of allowed regions for random eQTL LD block placement (for use as our background model for TFBS/eQTL overlap counts) by combining peak regions from all 397 TF ChIP-seq data sets with the monocyte eQTLs and merging together any regions whose edges were within 1.9 kbp of one

another (selected so that the resulting set of allowed regions covered ~25% of the genome). As a background model for the counts of TFBS/eQTL overlaps, we placed all eQTL LD blocks into random nonoverlapping locations within the set of allowed regions. We carried out 1,000 iterations of the LD block randomization. For each TF ChIP-seq data set, the count of peak center coordinates that overlapped with an eQTL LD block was computed, for both the human eQTL LD blocks and (as a background model) 1,000 iterations of the randomly placed LD blocks. For each ChIP-seq data set, we computed z-scores for the overlap counts using the mean and standard deviation of the background model-based counts (where $z >> 0$ indicates that the count of regions of overlap between the eQTLs and the ChIP-seq peaks is higher than would be expected based on the overlap count distribution of the null model where the eQTL LD blocks are randomly placed). For scoring the enrichment of cell types among ChIP-seq data sets ranked by z (in descending order) as well as scoring the enrichment of monocyte-regulating TFs among TFs ranked by adjusted P value (P_{adj}) (in ascending order), we used the gene set enrichment analysis (GSEA) statistical test[17] with $P = 0$ (unweighted) to obtain an enrichment score (E score). For testing cell types among ChIP-seq data sets, a higher E score means that the ChIP-seq experiments corresponding to the cell type or TF have lower ranks for their z-scores (ie, higher z-scores) than would be expected by chance. For testing monocyte-regulating TFs among all TFs ranked by P_{adj}, a higher E score means that the monocyte-regulating TFs have lower P_{adj} value ranks than would be expected by chance. To test for $E \neq 0$ significance, we used a permutation-based approach[36] with at least 1,000 iterations. For filtering TFs, we computed the average (over the 1,000 iterations) fraction of ChIP-seq peaks whose centers overlap one of the randomly placed LD blocks; the probability was used in a one-tailed binomial test. We adjusted the binomial P values for multiple hypothesis testing using the Benjamini–Hochberg false discovery rate method.[37]

TF network construction and visualization. We used GeneMANIA[38] (4/16/2016 release) to construct the interaction network for the top 27 TFs (by P value). We used Cytoscape[39] v3.4.0 for network layout and visualization. We used the HiveR R package[40] for generating the hive plot.

Supplementary information. The software source code for this method is available under a free-software, open-source (Apache 2.0) license at github.com/ramseylab/eqtlchiptest.

Acknowledgments

MC thanks Alvin Yu for technical advice. The authors thank Dr Thomas Sharpton for helpful suggestions.

Author Contributions

SAR and MC conceived and designed the experiments. MC and SAR analyzed the data. MC and SAR wrote the manuscript. All authors reviewed and approved the final manuscript.

Supplementary Material

Supplementary Table 1.
Supplementary Table 2.
Supplementary Table 3.

REFERENCES

1. Vaquerizas Juan M, Kummerfeld Sarah K, Teichmann Sarah A, Luscombe Nicholas M. A census of human transcription factors: function, expression and evolution. *Nat Rev Genet*. 2009;10:252–63.
2. GTEx Consortium. Human genomics. The Genotype-Tissue Expression (GTEx) pilot analysis: multi-tissue gene regulation in humans. *Science*. 2015;348:648–60.
3. Levo M, Segal E. In pursuit of design principles of regulatory sequences. *Nat Rev Genet*. 2014;15:453–68.
4. Ramsey S. An empirical prior improves accuracy for Bayesian estimation of transcription factor binding site frequencies within gene promoters. *Bioinform Biol Insights*. 2016;9:59–69.
5. Bulyk Martha L. Computational prediction of transcription-factor binding site locations. *Genome Biol*. 2003;5:201.
6. Sharmin M, Bravo HC, Hannenhalli S. Heterogeneity of transcription factor binding specificity models within and across cell lines. *Genome Res*. 2016;26:1110–23.
7. Gerstein MB, Kundaje A, Hariharan M, et al. Architecture of the human regulatory network derived from ENCODE data. *Nature*. 2012;489:91–100.
8. von Rohr P, Friberg MT, Kadarmideen HN. Prediction of transcription factor binding sites using genetical genomics methods. *J Bioinform Comput Biol*. 2007;5:773–93.
9. Wang D, Rendon A, Wernisch L. Transcription factor and chromatin features predict genes associated with eQTLs. *Nucleic Acids Res*. 2013;41:1450–63.
10. Zeller T, Wild P, Szymczak S, et al. Genetics and beyond–the transcriptome of human monocytes and disease susceptibility. *PLoS One*. 2009;5:e10693.
11. Fairfax BP, Makino S, Radhakrishnan J, et al. Genetics of gene expression in primary immune cells identifies cell type-specific master regulators and roles of HLA alleles. *Nat Genet*. 2012;44:502–10.
12. Gilchrist M, Thorsson V, Li B, et al. Systems biology approaches identify ATF3 as a negative regulator of Toll-like receptor 4. *Nature*. 2006;441:173–8.
13. Ramsey Stephen A, Klemm Sandy L, Zak Daniel E, et al. Uncovering a macrophage transcriptional program by integrating evidence from motif scanning and expression dynamics. *PLoS Comput Biol*. 2008;4:e1000021.
14. Medzhitov R, Horng T. Transcriptional control of the inflammatory response. *Nat Rev Immunol*. 2009;9:692–703.
15. Ramsey SA, Knijnenburg TA, Kennedy KA, et al. Genome-wide histone acetylation data improve prediction of mammalian transcription factor binding sites. *Bioinformatics*. 2010;26:2071–5.
16. Lozzio BB, Lozzio CB, Bamberger EG, Feliu AS. A multipotential leukemia cell line (K-562) of human origin. *Proc Soc Exp Biol Med*. 1981;166:546–50.
17. Subramanian A, Tamayo P, Mootha VK, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Nat Acad Sci U S A*. 2005;102:15545–50.
18. Nishiyama C, Yokota T, Okumura K, Ra C. The transcription factors Elf-1 and GATA-1 bind to cell-specific enhancer elements of human high-affinity IgE receptor alpha-chain gene. *J Immunol*. 1999;163:623–30.
19. Iwanaszko M, Kimmel M. NF-κB and IRF pathways: cross-regulation on target genes promoter level. *BMC Genomics*. 2015;16:307.
20. Wang J, Zhuang J, Iyer S, et al. Sequence features and chromatin structure around the genomic regions bound by 119 human transcription factors. *Genome Res*. 2012;22:1798–812.
21. Krzywinski M, Birol I, Jones SJ, Marra MA. Hive plots–rational approach to visualizing networks. *Brief Bioinform*. 2012;13:627–44.
22. Shao A-W, Sun H, Geng Y, et al. Bclaf1 is an important NF-κB signaling transducer and C/EBPβ regulator in DNA damage-induced senescence. *Cell Death Differ*. 2016;23:865–75.
23. Jiang X, Hu C, Arnovitz S, et al. miR-22 has a potent anti-tumour role with therapeutic potential in acute myeloid leukaemia. *Nat Commun*. 2016;7:11452.
24. Devaiah BN, Lu H, Gegonne A, et al. Novel functions for TAF7, a regulator of TAF1-independent transcription. *J Biol Chem*. 2010;285:38772–80.
25. He J-J, Ma J, Elsheikha HM, Song H-Q, Zhou D-H, Zhu X-Q. Proteomic profiling of mouse liver following acute *Toxoplasma gondii* infection. *PLoS One*. 2016;11:e0152022.
26. Xu H-G, Liu L, Gao S, Jin R, Ren W, Zhou G-P. Cloning and characterizing of the murine IRF-3 gene promoter region. *Immunol Res*. 2016;64:969–77.
27. Lu C, Li J-Y, Ge Z, Zhang L, Zhou G-P. Par-4/THAP1 complex and Notch3 competitively regulated pre-mRNA splicing of CCAR1 and affected inversely the survival of T-cell acute lymphoblastic leukemia cells. *Oncogene*. 2013;32:5602–13.

28. Zhang H, Hu H, Greeley N, et al. STAT3 restrains RANK- and TLR4-mediated signalling by suppressing expression of the E2 ubiquitin-conjugating enzyme Ubc13. *Nat Commun*. 2014;5:5798.
29. Peng X-P, Huang L, Liu Z-H. miRNA-133a attenuates lipid accumulation via TR4-CD36 pathway in macrophages. *Biochimie*. 2016;127:79–85.
30. Piehler Armin P, Grimholt Runa M, Ovstebø R, Berg Jens P. Gene expression results in lipopolysaccharide-stimulated monocytes depend significantly on the choice of reference genes. *BMC Immunol*. 2010;11:21.
31. Valledor AF, Borràs FE, Cullell-Young M, Celada A. Transcription factors that regulate monocyte/macrophage differentiation. *J Leukoc Biol*. 1998;63:405–17.
32. Heng Tracy SP, Painter Michio W; Immunological Genome Project Consortium. The Immunological Genome Project: networks of gene expression in immune cells. *Nat Immunol*. 2008;9:1091–4.
33. Wu C, Orozco C, Boyer J, et al. BioGPS: an extensible and customizable portal for querying and organizing gene annotation resources. *Genome Biol*. 2009;10:R130.
34. Johnson AD, Handsaker RE, Pulit SL, Nizzari MM, O'Donnell CJ, Bakker PIW. SNAP: a web-based tool for identification and annotation of proxy SNPs using HapMap. *Bioinformatics*. 2008;24:2938–9.
35. 1000 Genomes Project Consortium, Auton A, Brooks LD, et al. A global reference for human genetic variation. *Nature*. 2015;526:68–74.
36. Yu Alvin Z, Ramsey Stephen A. A computational systems biology approach for identifying candidate drugs for repositioning for cardiovascular disease. *Interdiscip Comput Life Sci*. 2016:1–6. doi:10.1007/s12539-016-0194-3.
37. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc B*. 1995;57:289–300.
38. Mostafavi S, Ray D, Warde-Farley D, Grouios C, Morris Q. GeneMANIA: a real-time multiple association network integration algorithm for predicting gene function. *Genome Biol*. 2008;9:S4.
39. Shannon P, Markiel A, Ozier O, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res*. 2003;13:2498–504.
40. Hanson Bryan A. *HiveR: 2D and 3D Hive Plots for R* 2016. R Package Version 0.2.55. Available at: cran.r-project.org. Published 2016.