

Supplementary Table S1: A summary of studies that systematically investigated the effects of SNPs on regulatory elements such as TFBSs. The analyses were either done by collecting experimentally supported and published data or by predicting the SNP's impact on TF binding using prediction tools.

Name	Species	DB/ tool	Website	Characteristics	Experimentally supported data or prediction
QBiC-Pred [1]	Human	Tool	http://qbic.genome.duke.edu/ (accessed on 2021-08-16)	<ul style="list-style-type: none"> TFBS prediction with regression models Prediction of changes in TF binding using ordinary least squares and evaluation of correlation between the predicted binding changes and changes in gene expression 	TFBS prediction
atSNP [2] atSNP-Search [3]	Human (atSNP: organisms from Bioconductor BSGenome package [4])	Tool, DB	http://atsnp.biostat.wisc.edu/search (accessed on 2021-08-16)	<ul style="list-style-type: none"> atSNP: R package for TF binding affinity testing for rSNPs (needs a SNP and motif set as input) atSNP Search: DB for human SNP-motif pairs and the respective significance 	TFBS prediction
INFERNO [5]	Human	Tool	http://inferno.lisanwanglab.org/index.php (accessed on 2021-08-16)	<ul style="list-style-type: none"> Inferring causal variants from GWAS results within annotated regulatory regions as enhancers including tissue context TFBS prediction with HOMER 	TFBS prediction
rSNPBASE [6], rSNPBASE 3.0 [7]	Human	DB	http://rsnp.psych.ac.cn/ (accessed on 2021-08-16) http://rsnp3.psych.ac.cn (accessed on 2021-08-16)	<ul style="list-style-type: none"> DB of rSNPs with references to regulatory elements Includes proximal and distal regulatory regions, post-transcriptional regulation, LD, and eQTL information rSNPBASE 3.0 includes regulatory element-target gene pairs for regulatory networks 	experimentally supported regulatory elements
SNP2TFBS [8]	Human	DB	https://ccg.epfl.ch/snp2tfbs/ (accessed on 2021-08-16)	<ul style="list-style-type: none"> DB of human SNPs that affect TFBSs and prediction of a consequence DB can be downloaded as text files or accessed via the website 	TFBS prediction

Ensembl Variant Effect Predictor [9]	Human, mouse, chicken, cattle, dog, goat, horse, pig, sheep,...	DB	https://www.ensembl.org/info/docs/tools/vep (accessed on 2021-08-16)	<ul style="list-style-type: none"> Collection of genomic annotation for prioritization of genomic variants in coding and non-coding regions Documented impacts of variants on, e.g., TF binding (there is almost no information available for livestock) 	TFBS prediction
HaploReg [10] HaploReg v4 [11]	Human	DB	http://pubs.broadinstitute.org/mammals/haploreg (accessed on 2021-08-16)	<ul style="list-style-type: none"> Annotation of variants based on chromatin states, conservation and regulatory motif alterations by combining variants into haplotype blocks and considering LD information 	TFBS prediction
MotifbreakR [12]	Human (organisms from Bioconductor BSGenome package [4])	Tool	No	<ul style="list-style-type: none"> R/Bioconductor package intended for human genome (adaptable to other organisms) Effects (i.e., strong, weak or neutral) on TFBSs with 3 different algorithms documented or custom SNPs, e.g., created with package funciSNP [13] 	TFBS prediction
ENlight [14]	Human	Tool	Not available anymore (access failed on 2021-08-16)	<ul style="list-style-type: none"> Annotation of GWAS variants by examining the overlays with biological annotations such as histone states, methylation patterns, TFBSs, and eQTL Returning plots for visualization 	experimentally supported regulatory elements
FunSeq2 [15]	Human	Tool	http://funseq2.gersteinlab.org/ (accessed on 2021-08-16)	<ul style="list-style-type: none"> Measurement of TF binding alteration of cancer-associated variants in non-coding regions of whole tumor genomes compared to the germline allele 	TFBS prediction
SuRFR [16]	Human	Tool	No	<ul style="list-style-type: none"> R package for the prioritization of candidate functional SNPs integrating annotations from different tools and DBs (e.g. conservation, chromatin states and TFBSs) 	experimentally supported regulatory elements
GEMINI [17]	Human	DB	Only on local machine	<ul style="list-style-type: none"> DB for coding and non-coding variant annotation by integrating chromatin information for different cell types 	experimentally supported regulatory elements

SNPnexus [18]	Human	Tool	https://www.snp-nexus.org/v4/ (accessed on 2021-08-16)	<ul style="list-style-type: none"> • Functional annotation for prioritization of novel and publicly known coding and non-coding SNPs • Investigation of potential regulatory role by computing the overlaps with documented regulatory elements such as conserved TFBSs or promoter regions 	TFBS prediction
RegulomeDB [19]	Human	DB	https://regulomedb.org/ (accessed on 2021-08-16)	<ul style="list-style-type: none"> • Prioritization of functional variants from genomic sequencing or GWAS experiments • prediction of protein binding alteration within experimentally supported regulatory regions based on e.g., ChIP-seq experiments, chromatin state information, and eQTLs 	experimentally supported regulatory elements and TFBS prediction
rSNP-MAPPER [20]	Human	Tool	Not available anymore (access failed on 2021-08-16)	<ul style="list-style-type: none"> • Computation of change in TF binding predictive score for different input sequences • TFBS prediction with PWMs based on Hidden Markov Models 	TFBS prediction
is-rSNP [21]	Human	Tool	Not available anymore (access failed on 2021-08-16)	<ul style="list-style-type: none"> • Predicting SNPs affecting TFBSs by using a PWM based statistical algorithm for the calculation of significant affinity change • Input SNPs from GWAS is possible 	TFBS prediction
sTRAP [22]	Human, mouse, yeast	Tool	http://trap.molgen.mpg.de/cgi-bin/home.cgi (accessed on 2021-08-16)	<ul style="list-style-type: none"> • Comparison of two input sequences and determining the TF which is mostly affected by a SNP • TF affinity prediction with TRAP [23] 	TFBS prediction
SNPinfo [24]	Human	Tool	https://manticore.niehs.nih.gov/ (accessed on 2021-08-16)	<ul style="list-style-type: none"> • Prioritize SNPs from GWAS by considering different functional characteristics (e.g., protein structure for coding SNPs or changes in TF binding, splice sites or MicroRNA-binding sites for non-coding SNPs) gene regulation, splicing, miRNA) in order to identify disease-associated SNPs • TFBS prediction with MATCH [25] 	TFBS prediction

F-SNP [26]	Human	DB	Not available anymore (access failed on 2021-08-16)	<ul style="list-style-type: none"> Integration of different tools and databases to predict the functional effect of SNPs The biomolecular functional categories splicing, transcription, translation and post-translation are considered TFBS prediction with TFSearch [27] and Consite [28] 	TFBS prediction
SNP@Promoter [29]	Human	DB	Not available anymore (access failed on 2021-08-16)	<ul style="list-style-type: none"> Identification of SNPs within predicted TFBSs and integration of conservation and gene annotation information Computation of a conservation score for each TFBS instead of a binding score or consequence TFBS prediction with MATCH [25] 	TFBS prediction

References

- Martin, V.; Zhao, J.; Afek, A.; Mielko, Z.; Gordân, R. QBIC-Pred: quantitative predictions of transcription factor binding changes due to sequence variants. *Nucleic acids research* **2019**.
- Zuo C., Shin S., Keleş S.. atSNP: transcription factor binding affinity testing for regulatory SNP detection. *Bioinformatics* **2015**, *31*, 3353–3355.
- Shin, S.; Hudson, R.; Harrison, C.; Craven, M.; Keleş, S. atSNP Search: a web resource for statistically evaluating influence of human genetic variation on transcription factor binding. *Bioinformatics* **2018**.
- Pagès H.. BSgenome: Infrastructure for Biostrings-based genome data packages and support for efficient SNP representation. *R package* **2016**.
- Amlie-Wolf A., Tang M., Mlynarski E.E., Kuksa P.P., Valladares O., Katanic Z., Tsuang D., Brown C.D., Schellenberg G.D., Wang L.-S.. INFERNO: inferring the molecular mechanisms of noncoding genetic variants. *Nucleic acids research* **2018**, *46*, 8740–8753.
- Guo, L.; Du, Y.; Chang, S.; Zhang, K.; Wang, J. rSNPBase: a database for curated regulatory SNPs. *Nucleic acids research* **2013**, *42*, D1033–D1039.
- Guo L., Wang J.. rSNPBase 3.0: an updated database of SNP-related regulatory elements, element-gene pairs and SNP-based gene regulatory networks. *Nucleic acids research* **2017**, *46*, D1111–D1116.
- Kumar S., Ambrosini G., Bucher P. SNP2TFBS - a database of regulatory SNPs affecting predicted transcription factor binding site affinity. *Nucleic acids research* **2016**, *45*, D139–D144.
- McLaren W., Gil L., Hunt S.E., Riat H.S., Ritchie G.R.S., Thormann A., Flicek P., Cunningham F.. The ensembl variant effect predictor. *Genome biology* **2016**, *17*, 122.
- Ward, L.D.; Kellis, M. HaploReg: a resource for exploring chromatin states, conservation, and regulatory motif alterations within sets of genetically linked variants. *Nucleic acids research* **2011**, *40*, D930–D934.
- Ward, L.D.; Kellis, M. HaploReg v4: systematic mining of putative causal variants, cell types, regulators and target genes for human complex traits and disease. *Nucleic acids research* **2015**, *44*, D877–D881.
- Coetzee S.G., Coetzee G.A., Hazelett D.J.. motifbreakR: an R/Bioconductor package for predicting variant effects at transcription factor binding sites. *Bioinformatics* **2015**, *31*, 3847–3849.
- Coetzee S.G., Rhie S.K., Berman B.P., Coetzee G.A., Noushmehr H.. FunciSNP: an R/bioconductor tool integrating functional non-coding data sets with genetic association studies to identify candidate regulatory SNPs. *Nucleic acids research* **2012**, *40*, e139–e139.
- Guo, Y.; Conti, D.V.; Wang, K. Enlight: web-based integration of GWAS results with biological annotations. *Bioinformatics* **2014**, *31*, 275–276.
- Fu Y., Liu Z., Lou S., Bedford J., Mu X.J., Yip K.Y., Khurana E., Gerstein M.. FunSeq2: a framework for prioritizing noncoding regulatory variants in cancer. *Genome Biology* **2014**, *15*, 480.
- Ryan, N.M.; Morris, S.W.; Porteous, D.J.; Taylor, M.S.; Evans, K.L. SuRFing the genomics wave: an R package for prioritising SNPs by functionality. *Genome medicine* **2014**, *6*, 79.
- Paila, U.; Chapman, B.A.; Kirchner, R.; Quinlan, A.R. GEMINI: integrative exploration of genetic variation and genome

- annotations. *PLoS computational biology* **2013**, *9*, e1003153.
18. Dayem Ullah, A.Z.; Lemoine, N.R.; Chelala, C. SNPnexus: a web server for functional annotation of novel and publicly known genetic variants (2012 update). *Nucleic acids research* **2012**, *40*, W65–W70.
 19. Boyle, A.P.; Hong, E.L.; Hariharan, M.; Cheng, Y.; Schaub, M.A.; Kasowski, M.; Karczewski, K.J.; Park, J.; Hitz, B.C.; Weng, S.; others. Annotation of functional variation in personal genomes using RegulomeDB. *Genome research* **2012**, *22*, 1790–1797.
 20. Riva, A. Large-scale computational identification of regulatory SNPs with rSNP-MAPPER. *BMC genomics*. BioMed Central, 2012, Vol. 13, p. S7.
 21. Macintyre, G.; Bailey, J.; Haviv, I.; Kowalczyk, A. is-rSNP: a novel technique for in silico regulatory SNP detection. *Bioinformatics* **2010**, *26*, i524–i530.
 22. Manke, T.; Heinig, M.; Vingron, M. Quantifying the effect of sequence variation on regulatory interactions. *Human mutation* **2010**, *31*, 477–483.
 23. Thomas-Chollier, M.; Hufton, A.; Heinig, M.; O'keeffe, S.; El Masri, N.; Roeder, H.G.; Manke, T.; Vingron, M. Transcription factor binding predictions using TRAP for the analysis of ChIP-seq data and regulatory SNPs. *Nature protocols* **2011**, *6*, 1860.
 24. Xu Z., Taylor J.A.. SNPinfo: integrating GWAS and candidate gene information into functional SNP selection for genetic association studies. *Nucleic Acids Research* **2009**, *37*, W600–W605.
 25. Kel A.E.*, Gößling E., Cheremushkin E., Kel-Margoulis O.V., Wingender E.. MATCH: a tool for searching transcription factor binding sites in DNA sequences. *Nucleic Acids Research* **2003**, *31* (13), 3576–3579.
 26. Lee P.H., Shatkay H.. F-SNP: computationally predicted functional SNPs for disease association studies. *Nucleic acids research* **2007**, *36*, D820–D824.
 27. Akiyama, Y. TFSEARCH: searching transcription factor binding sites. <http://diyhlpl.us/bryan/irc/protocol-online/protocol-cache/TFSEARCH.html> **2006**.
 28. Sandelin, A.; Wasserman, W.W.; Lenhard, B. ConSite: web-based prediction of regulatory elements using cross-species comparison. *Nucleic acids research* **2004**, *32*, W249–W252.
 29. Kim, B.C.; Kim, W.Y.; Park, D.; Chung, W.H.; Shin, K.s.; Bhak, J. SNP@ Promoter: a database of human SNPs (single nucleotide polymorphisms) within the putative promoter regions. *BMC bioinformatics*. BioMed Central, 2008, Vol. 9, p. S2.