Cytokine: X 3 (2021) 100056

Contents lists available at ScienceDirect

Cytokine: X

journal homepage: www.sciencedirect.com/journal/cytokine-x

Interferon stimulated binding of ISRE is cell type specific and is predicted by homeostatic chromatin state

Sivan Leviyang

Department of Mathematics and Statistics, Georgetown University, District of Columbia 20057, USA

ARTICLE INFO

Keywords:
ISRE
ISGF3
Interferon
Homeostasis
Epigenetics
Classification

ABSTRACT

The type I interferon (IFN) signaling pathway involves binding of the transcription factor ISGF3 to IFN-stimulated response elements, ISREs. Gene expression under IFN stimulation is known to vary across cell types, but variation in ISGF3 binding to ISRE across cell types has not been characterized. We examined ISRE binding patterns under IFN stimulation across six cell types using existing ChIPseq datasets. We find that ISRE binding is largely cell specific for ISREs distal to transcription start sites (TSS) and largely conserved across cell types for ISREs proximal to TSS. We show that bound ISRE distal to TSS associate with differential expression of ISGs, although at weaker levels than bound ISRE proximal to TSS. Using existing ATACseq and ChIPseq datasets, we show that the chromatin state of ISRE at homeostasis is cell type specific and is predictive of cell specific, ISRE binding under IFN stimulation. Our results support a model in which the chromatin state of ISRE in enhancer elements is modulated in a cell type specific manner at homeostasis, leading to cell type specific differences in ISRE binding patterns under IFN stimulation.

1. Introduction

Stimulating cells with type I interferons (IFN), in particular IFNa and IFNb, activates signaling pathways that lead to the upregulation of a collection of genes known as interferon stimulated genes, ISGs. ISGs play a diverse and essential role in the innate immune response, with some acting directly as antiviral effectors while others regulate the innate and adaptive responses [1]. Cell types vary in their ISGs, presumably reflecting the need for different functional responses to infection [2–5]. For example, IFN stimulation of cardiac myocytes weakly induces some ISGs that are potently induced in cardiac fibroblasts, possibly reflecting functional restrictions on cardiac monocytes which are not replenished [5].

Despite the differences in ISGs across cell types, our understanding of the cellular factors and pathways that determine a cell type's ISGs, referred to as the cell type's IFN signature [6], is incomplete. The canonical IFN signaling pathway involves activation of STAT1 and STAT2 and the formation of the IGSF3 trimer, composed of STAT1, STAT2, and IRF9. ISGF3 binds to genomic sequence motifs called IFN-stimulated response elements, ISREs, and regulates the transcription of ISGs [7]. Over the past two decades an array of non-canonical interferon pathways and regulators have been discovered and associated with variation in IFN signatures [8,2,9–17]. However, while it is clear that IFN signaling through non-canonical pathways leads to variation in the IFN signature, the degree of variability in the IFN signature associated with canonical IFN signaling, mediated by binding of ISGF3 to ISREs, is not clear.

Previous authors have investigated ISRE binding on a genomic scale. Seminal work by Hartman et al. and Robertson et al. identified STAT1 and STAT2 binding across the genome in HeLa cells [18,19]. Testoni et al. characterized STAT2 binding at homeostasis and under IFNa stimulation over a subset of 113 ISG promoters in hepatocytes [20]. Several groups have compared the effects of ISGF3 binding and STAT2:IRF9 dimer binding to ISREs [21–24]. However, none of these works characterized ISRE binding patterns across the genome or across cell types.

An array of studies over the past decade has shown that binding of transcription factors such as ISGF3 is shaped by epigenetic factors that vary between cell type [25–30,6,31], but the extent to which this is the case for ISGF3 binding to ISRE is unclear. The importance of epigenetic state in shaping the IFN response has been demonstrated in several contexts, [32,20,33–37]. However, the association between epigenetic state and ISGF3 binding to ISRE on a genomic level has not been characterized.

In this work, we take a step upstream of the IFN signature and considered the ISRE signature. By ISRE signature, we mean the

E-mail address: Sivan.Leviyang@georgetown.edu.

https://doi.org/10.1016/j.cytox.2021.100056

Available online 17 July 2021

collection of ISREs bound by ISGF3 and its components under IFN stimulation. We consider the ISRE signature of six cell types: mouse BMDM, fibroblast, and B cells, and human HeLa, K562, and THP1 cells. We investigate the distribution of bound ISREs within and across each of the cell types. We examine the association between ISRE signatures and IFN signatures by correlating ISRE binding with differential gene expression. And, through a machine learning approach, we quantify the capacity of the homeostatic, chromatin state of an ISRE to predict ISGF3 component binding under IFN stimulation.

Our analysis provides several novel insights into the IFN response that support existing models of IFN signaling [8]. Canonically, ISG regulation has been associated with ISRE binding in promoters. Our results support a role for bound ISRE distal to transcription start sites (TSS), presumably associated with enhancers. We find that most ISGs have bound ISREs in enhancers, that variation in the ISRE signature across cell types is predominantly associated with variation in enhancers, and that bound ISREs in enhancers associate with differential expression under IFN stimulation, although more weakly than bound ISRE in promoters. We show that ISRE binding under IFN stimulation is predicted at significant accuracy by the homeostatic chromatin state at ISRE, suggesting that factors controlling the homeostatic state of a cell play a substantial role in shaping the ISRE signature and, in turn, the IFN signature.

2. Materials and methods

Accessions and specific details for all datasets, [38–42,35,43,44,3,45–47,24,33], are provided in Supplementary Tables S1-S3.

2.1. Processing ATACseq and ChIPseq Datasets

For the GEO datasets, we downloaded fastq files from the SRA dataset, used bowtie 1.2 [48] to align reads to the mouse mm10 genome and human hg38 genome, and then used MACS2, [49], to call peaks and form signals (SPMR). The number of replicates varied from 1 −4 across the datasets and when multiple replicates were available we aligned each separately using bowtie, combined the bam files, and then used MACS2 to call peaks. We used the default settings in bowtie 1.2 and filtered the output bam file for mapped and, when appropriate, paired reads, and removed duplicate reads. We used the default settings in MACS for ChIPseq data, but for ATACseq added the options `–nomodel –shift −37 –extsize 73`. We called peaks at an FDR of 0.01. We used the MACS2 `–SPMR` option to create signal output. For ENCODE datasets, we downloaded ENCODE constructed peak calls and signal files, choosing the files constructed over all replicates.

2.2. Processing Transcription Datasets

For each gene in the mm10 and hg38 genomes, we identified the transcript with highest expression level under IFN stimulation in the mouse BMDM dataset and in a human monocyte derived macrophage dataset (Andrade et al., unpublished, GSE125352). This gave us one transcript for every gene on which we performed differential expression analysis.

For microarray datasets, which were all collected using Affymetrix chips, we downloaded raw CEL files, processed the CEL files using the R oligo package [50] using the rma algorithm and quantile normalization, and then used the R limma package [51] to produce fold-change and FDR values for the differential expression of each transcript. For the RNAseq datasets, we downloaded raw fastq files, aligned them to the mm10 or hg38 transcriptome using Salmon 0.99 [52] with default setting except the addition of the flag `–validateMappings`. We normalized the Salmon, expression values using the R edgeR package's [53] calcNormFactor function with the method parameter set to TMM, processed the output using the voom function in the limma package, and used the R limma package to calculate differential expression fold-change and FDR just as for microarray data.

We defined a gene to be an ISG if the gene's expression level changed by more than 50% under IFN stimulation and if the FDR adjusted p-value for differential expression computed by the R limma package was less than 0.05 for the RNAseq datasets and 0.01 for the microarray datasets. The lower p-value cutoff for microarray datasets removed noise associated with low level transcripts, a known limitation of microarrays [54].

2.3. Classification

We built classifiers using L1 logistic regression implemented using the R glmnet package [55]. Given a collection of ISRE loci to classify, we trained the classifier using 50% of the loci. To choose the L1 penalty parameter, we applied 10-fold cross validation to the training data and using maximization of AUROC (AUC) to select the parameter. We then tested our classifier on the held out data. We used the R ROCR package [56] to compute AUC values.

We restricted classification to samples sizes sufficiently large so that an AUC greater than 0.55 was statistically significant. To determine sample sizes that were sufficiently large, we permuted the ISRE class value while keeping the homeostatic signal fixed and then considered classifications for which the permuted AUC was under 0.55 with 95% confidence.

2.4. GO Enrichment Analysis

We used the R package gprofiler2 [57] to perform GO enrichment analysis. Given a collection of genes, we called the gprofiler2 gost function, filtered for GO classes with less than 500 genes, and sorted the remaining classes in descending order of p-values.

3. Results

From the GEO and ENCODE databases [58,59], we downloaded STAT1, STAT2, and IRF9 ChIPseq datasets collected under IFN stimulation for mouse bone marrow derived macrophages (BMDM), embryonic fibroblasts (fibroblasts), and splenic B (B) cells, and human HeLa-S3 (HeLa), K562, and THP1 cells. For B and K562 cells, only STAT1 and STAT2 ChIPseq datasets were available. All datasets were collected between 1.5 −6 hours post IFN stimulation, during the time period shown to have the most ISGs upregulated [35]. Datasets did vary in the type I IFN, IFNa or IFNb, and the IFN concentration, see Methods and Supplementary Table S1 for further details.

For each ChIPseq dataset, we called peaks using MACS2, [49], restricting our attention to peaks within 100 kb of gene transcription start sites (TSS) on the mouse and human genomes. BMDM had many more peaks than the other cell types, roughly 40,000 versus 1500 to 6000, and mouse cell types had more peaks than human cell types, even when BMDM where excluded, see Table 1 for specific values. The large number of BMDM peaks may reflect the unique role of BMDM in innate response. The difference in peak counts between mouse and human cell

Table 1
Most But Not All STAT1, STAT2, and IRF9 ChIPseq Peaks Were Near an ISRE motif. For each cell type, shown are (STAT1, STAT2, IRF9) the number of peaks called for the STAT1, STAT2 and IRF9 datasets, respectively, and (STAT1 (%), STAT2 (%), IRF9 (%)) the fraction of peaks that we associated with an ISRE.

cell type	STAT1	STAT2	IRF9	STAT1(%)	STAT2(%)	IRF9(%)
BMDM	7577	17467	14738	0.50	0.48	0.56
fibroblast	59	1582	2251	0.17	0.67	0.70
B	646	5647		0.19	0.33	
HeLa	261	794	622	0.35	0.59	0.68
K562	676	1408		0.53	0.40	
THP1	825	2888	164	0.08	0.11	0.18

types may be due to low sensitivity of ChIP antibody binding in human cells, particularly IRF9, as discussed in [24].

In order to associate ChIP peaks with ISRE, we defined an ISRE as any 10 base pair sequence that matched the canonical ISRE motif TTTCNNTTTC or differed from it by a single nucleotide (where NN allowed for any nucleotide) and then located all such sequences within 100 kb of a gene TSS. We associated an ISRE with a ChIP peak if the ISRE was within 10 base pairs of the peak summit.

Under our definition for ISREs, there are 2.5 and 3.1 million ISRE within 100 kb of TSS on the mouse and human genome, respectively, but only several thousand were associated with peaks. Conversely, a substantial percentage of peaks were associated with ISRE across all cell types, as shown in Table 1. Except for THP1 cells, greater than 30% and 56% of STAT2 and IRF9 peaks, respectively, associated with an ISRE. In contrast, STAT1 peaks associated with ISRE in a cell specific manner. Between 34% −50% of STAT1 peaks associated with ISRE for BMDM, HeLa, and K562, but less than 20% for fibroblast and B cells. Since sequences that vary by more than one nucleotide from TTTCNNTTTC may serve functionally as ISREs, a higher fraction of peaks may have associated with ISRE than our estimates.

For a given cell type, we refer to an ISRE associated with a peak as a *bound ISRE* and we distinguish bound ISRE by the combination of associated STAT1, STAT2, and IRF9 peaks: a STAT1 bound ISRE is bound only by STAT1, a STAT1.STAT2 bound ISRE is bound by STAT1 and STAT2 but not IRF9, etc. Across 3 of the 4 cell types for which we had STAT1, STAT2, and IRF9 ChIP data - namely BMDM, fibroblasts, and HeLa cells, with THP1 left out - 99% of bound ISRE were bound by the combinations, STAT2, IRF9, STAT2.IRF9, and STAT1.STAT2.IRF9. Associations involving STAT1 in the absence of STAT2 were rare. In contrast, in THP1, ISRE bound by STAT1 or STAT2 accounted for 11% and 77%, respectively, of bound ISRE, possibly reflecting sensitivity issues in ChIP antibody binding. In K562 and splenic B cells, the two cell types for which we did not have IRF9 ChIP data, 99% of bound ISRE were bound by STAT2 or STAT1.STAT2, in line with the large percentage of ISRE bound by STAT2 in BMDM, fibroblast, and HeLa cells. See Supplementary Table S4 for further details.

Overall, these associations between peaks and ISRE are consistent with existing results [23,24]: most binding of ISGF3 components was associated with ISREs, STAT1 binding to ISRE independent of STAT2 was rare, and most ISRE binding involved STAT2:IRF9 dimers or ISGF3.

3.1. ISRE Binding Patterns Within Cell Types

To characterize ISRE binding patterns within cell types, we considered the positional distribution of bound ISRE within genes. We associated each bound ISRE to the gene with the closest TSS. Since promoter and enhancer regulatory elements differ in chromatin state and functional properties [28], we split the region within 100 kb base pairs of a gene's TSS into an enhancer region that included positions 3 −100kb from the TSS and a promoter region that included positions within 500 base pairs of the TSS. We do not claim that ISRE in the enhancer and promoter regions act functionally as enhancers or promoters, respectively; instead, we use the terminology putatively and to emphasize the different characteristics of ISRE in the two regions, as we describe below. We purposely left a gap, positions 500 −3000 from the TSS, between the promoter and enhancer regions for greater separation of chromatin state and functionality.

Of genes with at least one bound ISRE, which we call *bound genes*, 82% in mouse cell types and 95% in human cell types had between 1 −6 bound ISRE. A small number of bound genes had a large number of bound ISRE, reflecting repeating, adjacent ISRE. Fig. 1 shows the number of bound ISRE in BMDM at different distances relative to the TSS. Other cell types showed a similar pattern. Bound ISRE were at their highest concentration near the TSS, in the promoter region, and fell in concentrations with distance from the TSS. However, most bound ISRE, 82% and 70% in mouse and human cell types respectively, fell in the enhancer region.

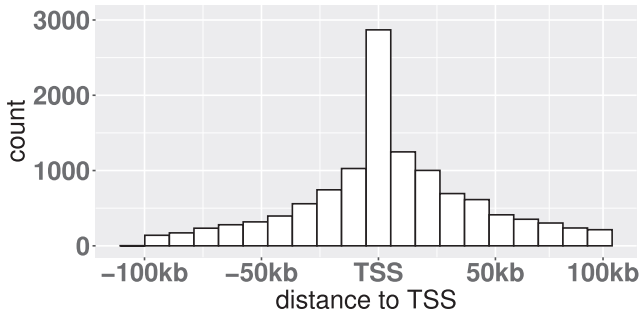

Fig. 1. Bound ISREs Were Most Dense Proximal to TSS, but Were Most Numerous Distal to TSS. Shown are the counts of bound ISRE in BMDM according to the distance of the ISRE to the gene transcription start site (TSS). Other cell types had had similar distributions with 70% −80% of bound ISRE distal to TSSs.

To further characterize the positional distribution of bound ISRE within genes, we considered three configurations of bound genes: a bound ISRE in the promoter region, in the enhancer region, and in both the promoter and enhancer region. Across all cell types, between 12 −25% of bound genes had a bound ISRE in the promoter while greater than 60% of bound genes had a bound ISRE in the enhancer region. Bound genes with bound ISRE in both the enhancer and promoter were rare, less than 5% across all cell types.

Since we may have missed a substantial number of bound ISRE due to our restrictive definition of an ISRE motif, the percentages of these three configurations may be biased. To address this potential bias, we considered two additional positional configurations: a bound ISRE just in the promoter and a bound ISRE just in the enhancer. For these two configurations, we used the absence of a STAT1, STAT2, or IRF9 peak to infer the absence of a bound ISRE (rather than the absence of a peak associated with an ISRE). Across cell types, greater than 43% of bound genes had a bound ISRE just in the enhancer, see the upper panel of Fig. 2. When we restricted genes to ISGs, bound ISREs in enhancer regions were enriched, but a substantial percentage of bound ISREs were still found in enhancer regions, see lower panel of Fig. 2.

3.2. ISRE Binding Patterns Across Cell Types

We next considered the cell specificity of ISRE binding patterns. For mouse cell types, roughly 90% of bound ISRE in fibroblasts and B cells were bound in BMDM, while about 20% of BMDM bound ISRE were bound in the other two cell types. Fibroblast and splenic B cells shared about 25% of their bound ISRE, giving an overall picture of broad ISRE binding in BMDM with fibroblasts and splenic B binding reflecting different subsets of the ISRE bound in BMDM. Splitting bound ISRE by enhancer and promoter region, we found that 20% of bound ISRE in the enhancer region were shared across all three cell types while 60% of bound ISRE in the promoter region were shared. Overall, 578 and 312 bound ISRE located in the enhancer and promoter regions, respectively, were shared across all three cell types. See Table 2 for more details.

For human cell types, there was no cell type with a broader response than the other two. Instead, HeLa and K562 shared roughly 60% of their bound ISRE while THP1 shared 20% of bound ISRE with each of the other cell types. As in mouse cell types, bound ISRE in the enhancer region were shared across all cell types at a lower percentage than bound ISRE in the promoter region, 11% and 65% respectively. Overall, 51 and 108 bound ISRE in the enhancer and promoter regions, respectively, were shared across all three cell types. The smaller number of shared bound ISRE relative to mouse likely reflects limitations in antibody sensitivity in the ChIPseq assays, as discussed above.

We performed a GO ontology enrichment analysis ([60,61]) on genes with bound ISRE that were shared across cell types, splitting the analysis

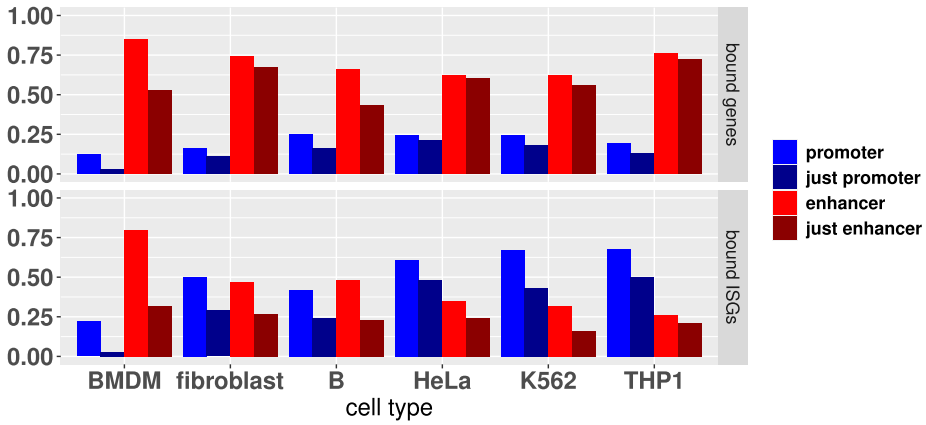

Fig. 2. In Most Genes, Bound ISRE Were Restricted to Enhancer Regions. (Upper panel) For each cell type, bars show (left to right) the fraction of bound genes with a bound ISRE in the promoter region, with bound ISRE just in the promoter region, with a bound ISRE in the enhancer region, and with a bound ISRE just in the enhancer region. (Lower panel) Same as the upper panel except that we restrict the genes to ISGs.

Table 2
ISRE Binding in Enhancer Regions Was More Cell Type Specific Than ISRE Binding in Promoter Regions. Shown are the fraction of bound ISRE shared in promoter regions (top tables) and enhancer regions (bottom tables) in mouse cell types (left tables) and human cell types (right table). The diagonal of each table gives the fraction of bound ISRE in the cell type that are also bound in both of the other two cell types. The off diagonals give the fraction of bound ISRE in the row cell type that are also bound in the column cell type. (FB = fibroblast).

promoter	BMDM	FB	B	promoter	HeLa	K562	THP1
BMDM	0.41	0.5	0.6	HeLa	0.54	0.9	0.55
FB	0.97	0.79	0.8	K562	0.84	0.5	0.53
B	0.88	0.61	0.6	THP1	0.92	0.97	0.92
enhancer				enhancer			
BMDM	0.06	0.22	0.17	HeLa	0.09	0.43	0.11
FB	0.89	0.23	0.24	K562	0.46	0.1	0.13
B	0.88	0.32	0.31	THP1	0.17	0.18	0.14

into genes with shared bound ISRE in promoter and enhancer regions, respectively; see Methods for details. For promoter regions, 169 genes had bound ISRE shared across all three mouse cell types. The three GO classes with the highest p-value were *response to virus, defense response to virus, and response to interferon-beta* and 33 of the 169 genes, 20%, were in the *defense response to virus* class. In contrast, 297 genes shared bound ISRE in the enhancer region and the three GO classes with the highest p-value were *regulation of immune effector process, regulation of hemopoiesis, and regulation of chromatin binding*. 15 of the 297 genes were in the *defense response to virus* class, 5%, significantly less than what we found for promoter genes (p-value 5E-13).

GO enrichment analysis for the human cell types gave similar results. Sixty genes shared bound ISRE in the promoter region with the top GO ontology classes of *defense response to virus, response to virus, and response to type I interferon*. 27 of 60 genes, 45%, were classified in the *defense response to virus* GO class. For the enhancer region, 29 genes shared bound ISRE in the enhancer region across cell types. The top three enriched GO classes were *cellular response to type I interferon, type I interferon signaling pathway, and response to type I interferon*. 5 of the 29 genes, 17%, were in the *defense response to virus*, significantly less than for the promoter region (p-value 0.001).

Overall, these results suggest a model in which binding of ISRE in promoter regions is largely conserved across cell types and associates with effectors of viral defense, while binding of ISRE in enhancer regions controls cell-specific components of the IFN response.

3.3. ISRE Binding and Differential Expression

To investigate the correlation between bound ISRE and ISGs, we downloaded transcription datasets collected under homeostasis and under IFN simulation and performed differential expression analysis to identify the ISGs (i.e. genes differentially expressed under IFN stimulation) for each cell type, see Methods for details. BMDM had roughly 1700 ISGs, consistent with the larger number of bound ISRE and the central role of BMDM in the innate response. The other cell types had roughly 400 ISGs, except for B cells, which had roughly 700, possibly reflecting a broader interferon response in immune cells. In mouse cell types, between 40% −62% of ISGs were bound, i.e. had a bound ISRE within 100 kb of the TSS. In human cell types, the fraction was much lower, 9% −15%, see Supplementary Table S5 for further details. In THP1, the already noted sensitivity issue may explain the small percentage, and lack of sensitivity may extend to the other cell types as well. That some ISGs were unbound may reflect bound ISRE missed by our analysis, particularly in the case of the human cell types, or interferon signaling pathways independent of ISGF3 components.

To test for a significant correlation between ISRE binding and ISGs, we compared the fraction of genes that were ISGs under different positional configurations of bound ISRE. To deconvolve the contribution of bound ISRE in promoters and enhancer regions, we compared genes with no ChIP peaks to genes with a bound ISRE just in the enhancer, just in the promoter, and in the enhancer and promoter, respectively; see Fig. 3. Of genes with no peak, between 1% −3% were ISGs, depending on the cell type. For all cell types except THP1, the percentage of genes that were ISGs rose as we moved from no peaks to a bound ISRE just in the enhancer to a bound ISRE just in the promoter to a bound ISRE in the enhancer and promoter. When a bound ISRE was just in the enhancer, 2% −12% of genes were ISGs, modestly above but statistically significantly greater than the 1% −3% seen in genes with no peak. When a bound ISRE was just in the promoter, 25% −69% of genes were ISGs, significantly above the percentages for bound ISRE just in the enhancer. Relatively few genes had bound ISRE in the promoter and enhancer, limiting our statistical power. With that in mind, when a bound ISRE was in both promoter and enhancer, 33% −100% of genes were ISGs, significantly greater than the percentage for just promoters in BMDM and fibroblast, but not the other cell types, possibly due to the small samples sizes.

We next considered the association of STAT2, STAT2.IRF9, and STAT1.STAT2.IRF9 binding to ISRE with differential expression. In order to deconvolve binding of these difference peak combinations, we considered genes with a single bound ISRE, allowing us to associate a specific peak combination with differential expression. We necessarily restricted to cell types for which we had an IRF9 ChIP dataset, and we left out THP1 due to the sensitivity issues, leaving BMDM, HeLa and fibroblasts.

When the bound ISRE was in the promoter region, the percentage of

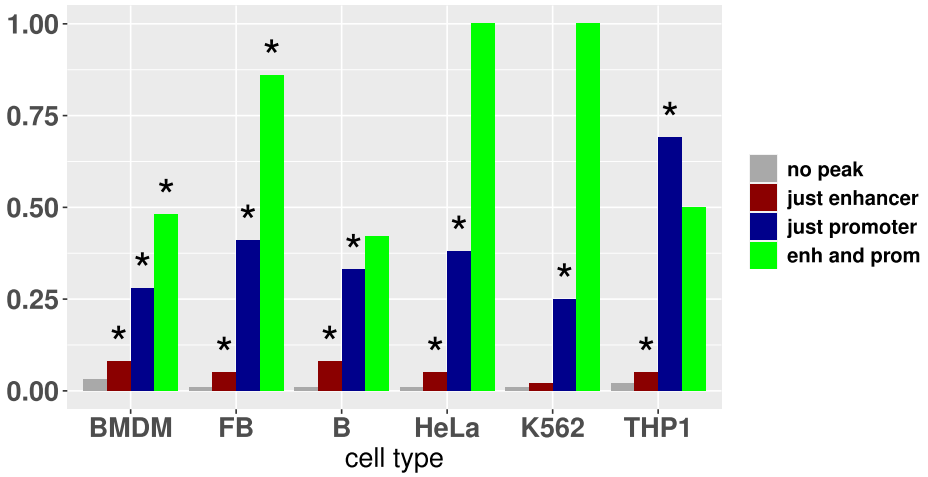

Fig. 3. Genes with Bound ISRE in Enhancer or Promoter Regions Were Significantly More Likely to be ISGs. For each cell type, bars show (left to right) the fraction of genes that are ISGs over (no peak) genes with no STAT1, STAT2, and IRF9 ChIPseq peaks, (just enhancer) genes with a bound ISRE just in the enhancer region, (just promoter), genes with a bound ISRE just in the promoter region, and (enh and prom) genes with a bound ISRE in both the enhancer and promoter regions. Asterisks identify a statistically significant increase, at greater than 95% confidence, relative to the previous (bar to the left) frequency and configuration. (FB = fibroblast).

genes that were ISGs rose sequentially as we considered STAT2, STAT2.IRF9, and STAT1.STAT2.IRF9 bound ISRE, respectively. Under STAT2 binding, 5% −12% of genes were ISGs, under STAT2:IRF9 binding, 18% −46%, and under STAT1:STAT2:IRF9 binding, 58% −60%. When the bound ISRE was in the enhancer region, percentages also rose sequentially as we considered STAT2, STAT2.IRF9, and STAT1.STAT2.IRF9 ISRE classes, respectively, but overall the percentages were lower than when the binding was in the promoter region. See Supplementary Figure 1 for further details.

Overall, our analysis shows a strong association between ISRE binding in the promoter region and differential expression under IFN stimulation. We also noted a significant, though weaker, association between ISRE binding in the enhancer regions and differential expression. However, the association between peaks and differential expression is weak or absent when we restrict to ISRE bound by STAT2, and weaker in ISRE bound by STAT2.IRF9 than ISRE bound by STAT1.STAT2.IRF9. STAT2 peaks may represents misclassifications of STAT2.IRF9 or STAT1.STAT2.IRF9 peaks. The inability of STAT2 to stably bind DNA in the absence of IRF9 supports this viewpoint [62]. In contrast, multiple reports support the existence of STAT2.IRF9 dimers that are functionally important and independent of STAT1 expression [21,23,24].

3.4. *Predicting ISRE Binding from Homeostatic Chromatin State*

Given that ISRE binding patterns differed across cell types, we next asked whether a cell's homeostatic chromatin state could be used to predict binding to ISRE under IFN stimulation. Since our analysis above suggested weak or non-functional binding when ISRE were associated with STAT2 peaks alone, for both mouse and human cell types we restricted our attention to ISRE associated with a STAT2 peak and either an IRF9 or STAT1 peak in at least one cell type. Over these ISRE, we asked if a cell type's homeostatic chromatin state could be used to distinguish between bound and unbound ISRE under IFN stimulation. Importantly, all ISRE we considered were bound by multiple components of ISGF3 in some cell type, so differences in ISRE binding reflected cell specific effects. We did not include THP1 cells in this analysis, given the poor sensitivity of ChIPseq data in THP1 cells mentioned above.

To define a homeostatic state for an ISRE, we formed a 1010 base pair locus composed of the 10 base pair ISRE motif and the 500 base pairs upstream and downstream of the motif. To describe the homeostatic, chromatin state associated with the ISRE loci, we collected ATACseq and H3K4ME1, H3K4ME3, and H3K27AC ChIPseq datasets for each of our cell types. All datasets were downloaded from GEO or ENCODE and were collected at homeostasis. For each dataset, we used MACS2 to build a signal (SPMR) over each of the 1010 base pair ISRE loci and we used these signals to define a homeostatic, chromatin state for each ISRE locus.

To quantify the predictive capacity of the homeostatic state, we built binary classifiers to distinguish bound ISRE from unbound ISRE based on the chromatin state of the ISRE loci in the given cell type. To build classifiers using the homeostatic signals, we defined a set of 12 features (values) for each signal: the mean of the signal across the 1010 base pair of the ISRE locus and the fraction of the signal contained in eleven, 90 base pair windows covering the locus (the windows did not include the 10 base pairs on the far ends). We built binary classifiers to predict whether an ISRE was bound or unbound based on these features.. In all cases, we trained each classifier using half the relevant ISRE loci and used the other half to test prediction accuracy using the AUC measure. As we did above, we split our analysis based on whether the ISRE was in an enhancer or promoter region. We restricted our analysis to cell types with sufficient sample size to make AUCs greater than 0.55 statistically significant, see Methods for further details.

Fig. 4 shows the SPMR signal of the different chromatin assays averaged across ISRE loci split by cell type, promoter or enhancer region, and whether the ISRE was bound or unbound. To normalize the signals, we scaled the signals so that a background collection of loci which served as a null had a mean signal of 0. For ISRE in enhancer regions, bound ISRE showed higher level of homeostatic, accessibility and greater modification levels for both H3K4ME1 and H3K27AC than the unbound ISRE. H3K4ME3 modifications were similar for bound and unbound ISRE in BMDM, fibroblast, and K562 cell types, but were higher in B and HeLa cells. Both bound and unbound ISRE showed higher levels of accessibility and H3K4ME1 and H3K27AC modifications than the background signal. In contrast to ISRE in enhancer regions, bound ISRE in promoter regions were not consistently different in their mean chromatin signals than unbound ISRE and both bound and unbound ISRE were not above background signal levels, with the single exception of H3K27AC in B cells.

Fig. 5 gives AUC values for classifiers for bound vs unbound ISRE in enhancer regions using features from each chromatin assay separately and all chromatin assays jointly. (An AUC of 0.5 and 1.0 correspond to random assignment and perfect assignment, respectively). In line with the mean signals, prediction using H3K4ME3 was poor with AUC values of roughly 0.55 in BMDM, fibroblast, and K562 cell types, as would be expected for ISRE in enhancers, while prediction in B and HeLa cells was relatively good. Prediction using features from the other chromatin assays was more consistent. In all cell types except for fibroblasts, AUCs were roughly between 0.70 −0.80 using H3K4ME1 or H3K27AC and between 0.75 −0.90 using features across all chromatin assays.

Since a substantial fraction of ISRE were bound in multiple cell types - as shown in Table 2 - the predictive capacity of homeostatic chromatin

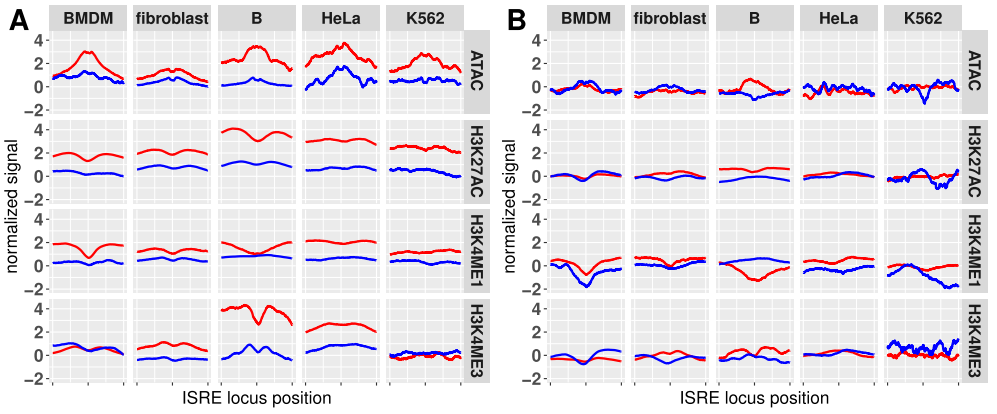

Fig. 4. In Enhancer But Not Promoter Regions, the Mean, Homeostatic Chromatin State Differed Between Bound and Unbound ISRE. Shown are the mean SPMR of bound (RED) and unbound (BLUE) ISRE in (A) enhancer and (B) the promoter regions. The signals are normalized relative to background loci. The status of an ISRE as bound and unbound was determined from assays collected under IFN stimulation while the chromatin signals were determined from assays collected at homeostasis.

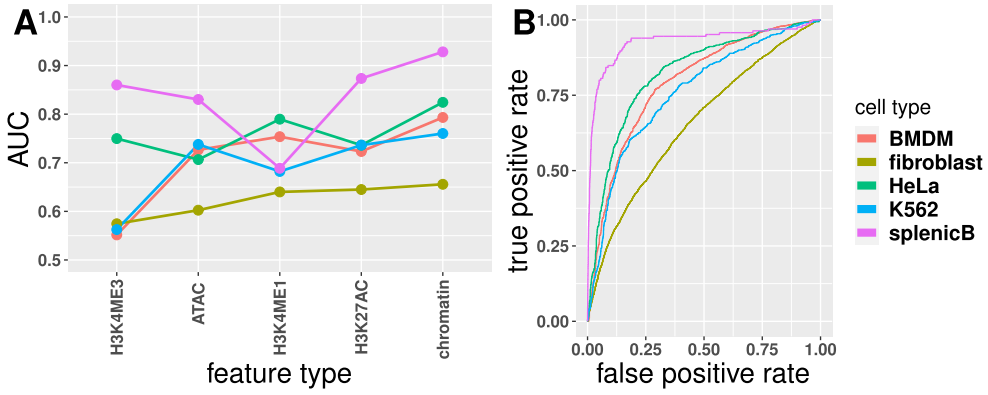

Fig. 5. Chromatin State at Homeostasis Predicts ISRE Binding Under IFN Stimulation At Significant Accuracy. (A) AUC values for classifiers using features from individual chromatin assays and using features jointly across all chromatin assays (chromatin). (B) ROC curves for the classifier using features jointly across all chromatin assays. AUC values greater than 0.55 are statistically significant.

state could derive from a shared set of ISRE rather than ISRE that are bound in only certain cell types. To address this issue, we repeated our analysis but restricting bound ISRE to ISRE that were uniquely bound in the given cell type. Across all cell types, AUC values rose or stayed roughly constant when we considered cell specific, bound ISRE, see Supplementary Figure 2 for details. Fibroblasts, in particular, had a large increase of roughly 0.10 in AUC, suggesting that fibroblast specific, bound ISRE were in activated enhancers at homeostasis, while bound ISRE shared with other cell types were in enhancers activated under non-homeostatic conditions.

As we did for ISRE in enhancer regions, we built classifiers for ISRE in promoter regions. AUCs for promoter regions were either roughly similar or lower than AUCs for enhancer regions, see Fig. 6. Using features across all chromatin assays jointly, the AUC for HeLa cells fell to 0.55, which is not significantly above random classification, but the AUC for BMDM, K562, and B cells stayed above 0.70, which is statistically significant, showing that even though mean signals were not above background in promoter regions, there were differences in the

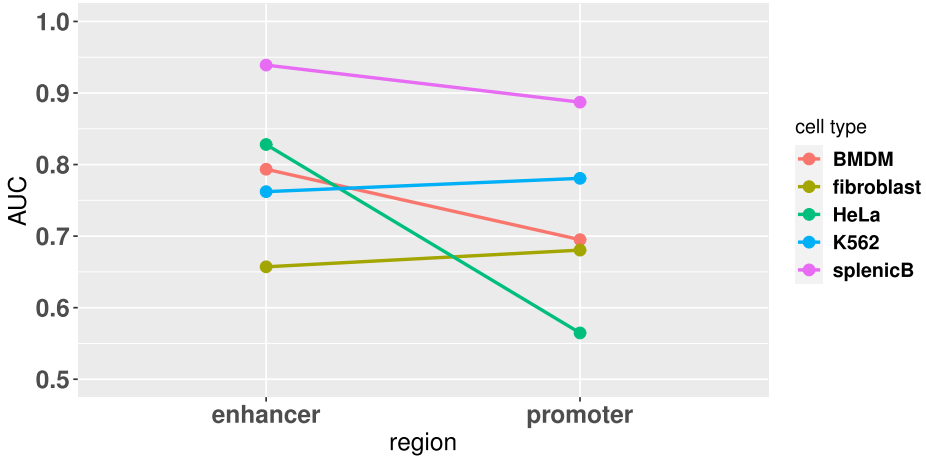

Fig. 6. Prediction Accuracy for ISRE in Promoter Regions Was Lower Than for ISRE in Enhancer Regions. AUC values for ISRE in promoter regions versus enhancer regions. All ATACseq and histone modification ChIPseq features were used for classification.

homeostatic chromatin state between bound and unbound ISRE.

We next considered the capacity of STAT1, STAT2, and IRF9 binding at homeostasis to predict ISGF3 component binding under IFN stimulation. In a manner analogous to our analysis of ATACseq and histone modification ChIPseq datasets above, we downloaded STAT1, STAT2, and IRF9 ChIP datasets collected at homeostasis for our cell types, used MACS2 to compute SPMR signals, and constructed classifiers based on SPMR signal features. For K562 cells, we did not have STAT1 and STAT2 ChIPseq data at homeostasis. Instead, we had K562 STAT1 and STAT2 ChIPseq datasets collected at 30 min post IFN stimulation, as opposed to the IFN stimulated datasets for K562 cells, which were collected at 6 h. See Supplementary Table S3 for dataset details.

Fig. 7 shows the mean, homeostatic STAT1, STAT2, and IRF9 SPMR signals. In contrast to results for ATACseq and histone ChIPseq signals, mean signals of bound ISRE in enhancer and promoter regions were similar. The mean signals of unbound (under IFN stimulation) ISRE were essentially zero, even for K562 cells. The bound ISRE had positive mean signals in all cell types and both promoter and enhancer regions, except for the STAT1 signal. K562 signals were particularly elevated, reflecting the 30 min of IFN stimulation, but B cells also had elevated signals, suggesting that B cells were in a relatively IFN activated state even at homeostasis.

Fig. 8 shows accuracy results for prediction of ISRE binding using the STAT1, STAT2, and IRF9 features jointly, using the chromatin features (i.e. ATACseq, H3K4ME1, H3K27AC, and H3K4ME3) jointly, as well as using both the STAT1, STAT1, IRF9 and chromatin features jointly. For ISRE in enhancer regions, prediction using STAT1, STAT2 and IRF9 features in the case of K562 cells was almost perfect, > 0.95, as might be expected given the later sampling time. Somewhat surprisingly, although in line with the mean signal profiles, B cells prediction was also nearly perfect, > 0.95. However, putting aside the special case of K562 cells, the chromatin state features were similar or better predictors of binding to ISRE under IFN stimulation than STAT1, STAT2, and IRF9 features. In contrast, for promoter regions, STAT1, STAT2, and IRF9 features were better predictors than the chromatin state.

Interestingly, combining STAT1, STAT2, and IRF9 features with chromatin state features did not substantially improve AUC values relative to using these two feature groups separately. This lack of increase in predictive capacity could reflect strong correlation between the features or non-linear effects that cannot be captured by linear predictors. We computed the correlation between our predictors using STAT1, STAT2, and IRF9 features and using chromatin state features, respectively. We found relatively low correlation levels, < 0.40. Further, we noted that many ISRE displayed high values of STAT1, STAT2, and IRF9 signals or chromatin state signals, but not both, suggesting a model in which some ISRE bound under IFN stimulation are bound by ISGF3 components at homeostasis, while other ISRE are not bound by ISGF3 components at homeostasis but instead have an activated chromatin state that leads to IFN stimulated binding.

4. Discussion

In this work, we examined ISRE binding signatures under type I IFN stimulation in six cell types. Within each cell type, bound ISREs were most dense in promoter regions, but most (70% −80%) bound ISREs were more than 3 kb away from a TSS, in what we referred to as enhancer regions. Most genes (>60%) and a substantial percentage of ISGs (>25%) with bound ISREs had bound ISREs in their enhancer region. We further found that a substantial percentage of genes and ISGs (>40% and >10% respectively) with bound ISRE had their bound ISRE restricted to enhancer regions, although this result depends on ChIPseq sensitivity. We found that bound ISRE in promoters were shared across cell types at higher levels (41% −92%) than in enhancers regions (6% −31%). Overall, these results demonstrate that a substantial portion, if not a majority, of ISGF3 binding occurs on ISRE in enhancer elements.

Presumably, ISREs function to induce or restrict expression of ISGs. In keeping with canonical models of ISG regulation, we found that 25% −69% of genes with bound ISRE in the promoter were ISGs, even in cases where bound ISRE were restricted to the promoter region. The role of enhancers in regulating ISGs is more subtle. Comparing genes with no bound ISRE within 100 kb of their TSS and genes with a bound ISRE just in the enhancer region, 1% −3% and 5% −11%, respectively, were ISGs, a relatively small effect though statistically significant. This small effect could be due to misclassification errors or a functionally weak contribution to differential expression by ISRE in enhancer regions. Alternatively, enhancer regulation may only take place in certain contexts beyond the presence of IFN stimulation, or the timing of differential expression may be delayed beyond the 2 −6 hour window post stimulation that we consider. Genes with a bound ISRE in the promoter and enhancer regions were more likely to be ISGs than genes with a bound ISRE just in the promoter region, suggesting that bound ISRE in enhancer and promoter regions work synergistically even within the first few hours of IFN stimulation. Overall, these results support a functional role for ISGF3 binding to ISRE in enhancer elements.

Using homeostatic chromatin state features, which we constructed using ATACseq and histone modification ChIPseq datasets, we built classifiers to predict whether an ISRE was bound or unbound under IFN stimulation. For ISRE in both enhancer and promoter regions, our classifiers had significant predictive capacity, with AUCs roughly in the range 0.70 −0.90. We found similar results when we restricted to ISRE that were bound only in a specific cell type. Interestingly, we found that the B cell, homeostatic state had particularly high predictive capacity, with an AUC of roughly 0.90. In fibroblasts, the homeostatic state had a higher predictive capacity for ISRE bound only in fibroblasts relative to ISRE bound in fibroblasts and other cell types. These results suggest a modulatory role for the homeostatic state in shaping the IFN response. We speculate that in B cells the homeostatic state allows for a quick response to IFN stimulation. In fibroblasts, a quick response might occur

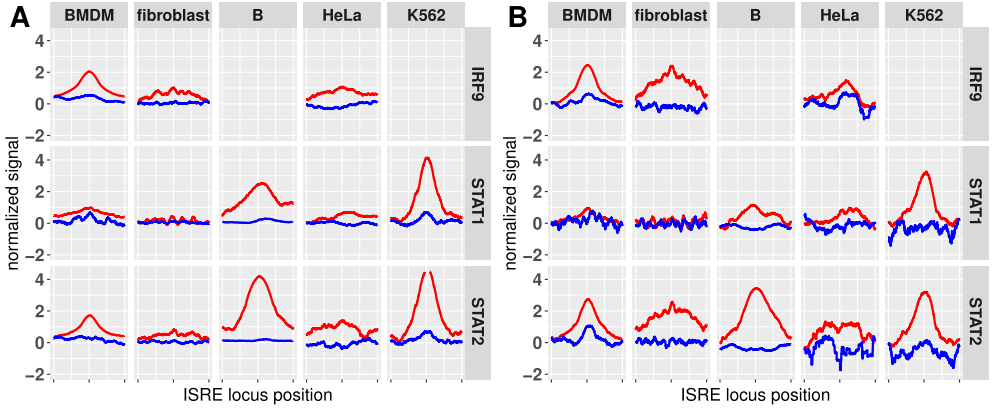

Fig. 7. In Enhancer and Promoter Regions, the Mean, Homeostatic ISGF3 Signals Differed Between Bound and Unbound ISRE. Shown are the mean SPMR signals of bound (RED) and unbound (BLUE) ISRE loci in (A) enhancer and (B) the promoter regions. An ISRE was bound or unbound based on STAT1, STAT2, and IRF9 ChIPseq datasets collected under IFN stimulation. Signals shown are from STAT1, STAT2, and IRF9 ChIPseq datasets collected under homeostasis. The exception are the K562 signals, which are from datasets collected 30 min after IFN stimulation.

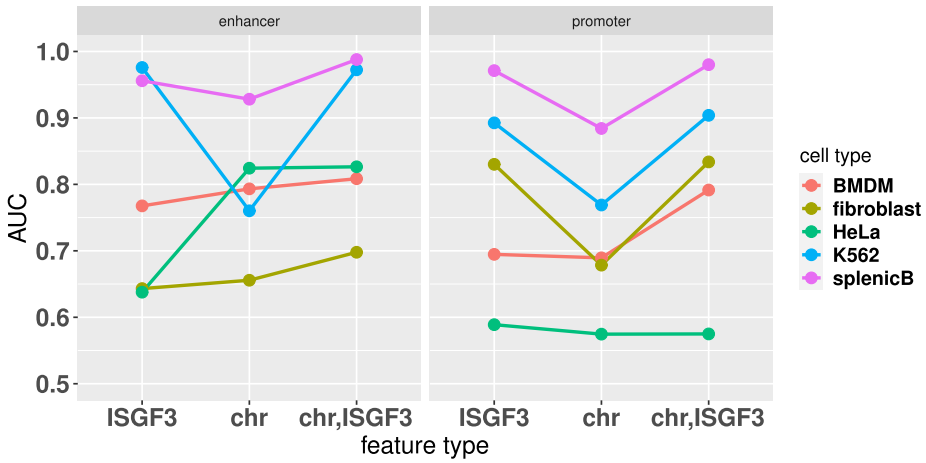

Fig. 8. For ISRE in Promoters But Not in Enhancers, STAT1, STAT2, and IRF9 ChIPseq Signals Collected at Homeostasis Were Better Predictors of IFN Stimulated Binding than Chromatin State. Shown are AUC values using the STAT1, STAT2, and IRF9 signals (ISGF3); the chromatin signals (chr); and the STAT1, STAT2, IRF9 and chromatin signals jointly (chr,ISGF3) for ISRE in enhancer and promoter regions.

in association with cell specific ISRE, but other ISRE that are bound during IFN stimulation might be associated with a slower response. Further datasets would be needed to explore these speculations.

We used homeostatic STAT1, STAT2, and IRF9 ChIPseq datasets to explore the capacity of ISGF3 component binding at homeostasis to predict binding under IFN stimulation. For ISRE in enhancer regions, chromatin state at homeostasis was a similar or a better predictor than ISGF3 component binding, while in promoter regions ISGF3 component binding was a much better predictor than chromatin state. Interestingly, in enhancers, prediction based on chromatin state and homeostatic ISGF3 binding were not strongly correlated, suggesting a model in which some ISRE are at least weakly bound by ISGF3 at homeostasis, while others have an active chromatin state but require IFN stimulation to allow for binding.

The main limitation of this work is the dependence on six STAT1, STAT2, and IRF9 ChIPseq datasets, one for each cell type. We used the STAT1, STAT2, and IRF9 ChIPseq datasets collected under IFN stimulation to determine the bound ISRE, which in turn affected every other analysis we did. We cannot rule out that some of our results reflect the bias of a particular ChIPseq dataset. Further, more datasets would allow us to better investigate cell type specific differences. For example, prediction for B cells was particularly accurate, but whether this was associated with the specific dataset or reflective of homeostatic regulation in B cells is unclear. The relative scarcity of STAT1, STAT2, and IRF9 ChIPseq datasets, for example in the GEO and ENCODE databases, has been previously noted [63] and limited our analysis.

Overall, we have shown that ISRE signatures both overlap and vary between cell types, reflecting cell type specificity in enhancer regions and a conserved antiviral response in promoter regions. We have shown that ISRE signatures correlate with ISG signatures, with bound ISREs in promoters strongly correlated with differential expression of ISGs and with bound ISREs in enhancers also correlating with differential expression, although more weakly. Finally, we have shown that chromatin state at homeostasis is predictive of ISRE binding under IFN stimulation. Together, these results support a significant role for enhancer elements in the IFN response and suggest that cell types modulate ISGF3 binding to ISRE through homeostatic pathways that shape chromatin state.

Funding

This research received no external funding.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

We are deeply indebted to the many groups that collected the datasets used in this study. In particular, this work would not have been possible without the BMDM, fibroblast, and THP1 data collected by Platanitis et al. in [24] and the HeLa data collected by Au-Yeung and Horvath in [33], and we thank T. Decker and E. Platanitis for answering our questions relating to their datasets. This work benefited from data assembled by the ImmGen consortium and the ENCODE consortium.

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at https://doi.org/10.1016/j.cytox.2021.100056.

References

[1] W.M. Schneider, M.D. Chevillotte, C.M. Rice, Interferon-stimulated genes: a complex web of host defenses., Annual review of immunology 32 (2014) 513–45. arXiv:NIHMS150003, doi:10.1146/annurev-immunol-032713-120231. http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4313732&tool=pmcentrez&rendertype=abstract.

[2] A.H.H. van Boxel-Dezaire, J.A. Zula, Y. Xu, R.M. Ransohoff, J.W. Jacobberger, G. R. Stark, Major Differences in the Responses of Primary Human Leukocyte Subsets to IFN-β, J. Immunol. 185 (10) (2010) 5888–5899, https://doi.org/10.4049/jimmunol.0902314.

[3] S.W. Cho, J. Xu, R. Sun, M.R. Mumbach, A.C. Carter, Y.G. Chen, K.E. Yost, J. Kim, J. He, S.A. Nevins, S.F. Chin, C. Caldas, S.J. Liu, M.A. Horlbeck, D.A. Lim, J. S. Weissman, C. Curtis, H.Y. Chang, Promoter of lncRNA Gene PVT1 Is a Tumor-Suppressor DNA Boundary Element, Cell 173 (6) (2018) 1398–1412.e22, https://doi.org/10.1016/j.cell.2018.03.068. URL https://doi.org/10.1016/j.cell.2018.03.068.

[4] X. Wu, V.L. Dao Thi, Y. Huang, E. Billerbeck, D. Saha, H.H. Hoffmann, Y. Wang, L. A. Silva, S. Sarbanes, T. Sun, L. Andrus, Y. Yu, C. Quirk, M. Li, M.R. MacDonald, W. M. Schneider, X. An, B.R. Rosenberg, C.M. Rice, Intrinsic Immunity Shapes Viral Resistance of Stem Cells, Cell 172 (3) (2018) 423–438.e25. doi:10.1016/j.cell.2017.11.018. doi: 10.1016/j.cell.2017.11.018.

[5] J. Zurney, K.E. Howard, B. Sherry, Basal expression levels of IFNAR and Jak-STAT components are determinants of cell-type-specific differences in cardiac antiviral responses., Journal of virology 81 (24) (2007) 13668–80. doi:10.1128/JVI.01172-07. http://www.ncbi.nlm.nih.gov/pubmed/17942530%0Ahttp://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC2168836.

[6] F.J. Barrat, M.K. Crow, L.B. Ivashkiv, Interferon target-gene expression and epigenomic signatures in health and disease, Nat. Immunol. 20 (12) (2019)

1574–1583, https://doi.org/10.1038/s41590-019-0466-2. URL https://doi.org/10.1038/s41590-019-0466-2.
[7] J.E. Darnell, I.M. Kerr, G.R. Stark, Jak-STAT pathways and transcriptional activation in response to IFNs and other extracellular signaling proteins, Science 264 (5164) (1994) 1415–1421, https://doi.org/10.1126/science.8197455.
[8] L.B. Ivashkiv, L.T. Donlin, Regulation of type i interferon responses, Nature Reviews Immunology 14 (1) (2014) 36–49. arXiv:NIHMS150003, doi:10.1038/nri3581. doi: 10.1038/nri3581.
[9] H.H. Ho, L.B. Ivashkiv, Role of STAT3 in type I interferon responses: Negative regulation of STAT1-dependent inflammatory gene activation, J. Biol. Chem. 281 (20) (2006) 14111–14118, https://doi.org/10.1074/jbc.M511797200.
[10] Y. Tanabe, T. Nishibori, L. Su, R.M. Arduini, D.P. Baker, M. David, Cutting Edge: Role of STAT1, STAT3, and STAT5 in IFN-$\alpha\beta$Responses in T Lymphocytes, J. Immunol. 174 (2) (2005) 609–613, https://doi.org/10.4049/jimmunol.174.2.609.
[11] W.-B. Wang, D.E. Levy, C.-K. Lee, STAT3 Negatively Regulates Type I IFN-Mediated Antiviral Response, J. Immunol. 187 (5) (2011) 2578–2585, https://doi.org/10.4049/jimmunol.1004128.
[12] M.P. Gil, M.J. Ploquin, W.T. Watford, S.H. Lee, K. Kim, X. Wang, Y. Kanno, J. J. O'Shea, C.A. Biron, Regulating type 1 IFN effects in CD8 T cells during viral infections: Changing STAT4 and STAT1 expression for function, Blood 120 (18) (2012) 3718–3728, https://doi.org/10.1182/blood-2012-05-428672.
[13] L. Icardi, S. Lievens, R. Mori, J. Piessevaux, L. De Cauwer, K. De Bosscher, J. Tavernier, Opposed regulation of type I IFN-induced STAT3 and ISGF3 transcriptional activities by histone deacetylases (HDACS) 1 and 2, FASEB J. 26 (1) (2012) 240–249, https://doi.org/10.1096/fj.11-191122.
[14] A. Yarilina, K.H. Park-Min, T. Antoniv, X. Hu, L.B. Ivashkiv, TNF activates an IRF1-dependent autocrine loop leading to sustained expression of chemokines and STAT1-dependent type I interferon-response genes, Nat. Immunol. 9 (4) (2008) 378–387, https://doi.org/10.1038/ni1576.
[15] S.L. Ng, B.A. Friedman, S. Schmid, J. Gertz, R.M. Myers, B.R. TenOever, T. Maniatis, IκB kinase ϵ(IKKϵ) regulates the balance between type I and type II interferon responses, Proceedings of the National Academy of Sciences of the United States of America 108 (52) (2011) 21170–21175. doi:10.1073/pnas.1119137109.
[16] S. Wienerroither, P. Shukla, M. Farlik, A. Majoros, B. Stych, C. Vogl, H.J. Cheon, G. R. Stark, B. Strobl, M. Müller, T. Decker, Cooperative Transcriptional Activation of Antimicrobial Genes by STAT and NF-κB Pathways by Concerted Recruitment of the Mediator Complex, Cell Reports 12 (2) (2015) 300–312, https://doi.org/10.1016/j.celrep.2015.06.021.
[17] L.B. Ivashkiv, IFNγ: signalling, epigenetics and roles in immunity, metabolism, disease and cancer immunotherapy, Nat. Rev. Immunol. 18 (9) (2018) 545–558, https://doi.org/10.1038/s41577-018-0029-z. URL https://doi.org/10.1038/s41577-018-0029-z.
[18] S.E. Hartman, P. Bertone, A.K. Nath, T.E. Royce, M. Gerstein, S. Weissman, M. Snyder, Global changes in STAT target selection and transcription regulation upon interferon treatments, Genes and Development 19 (24) (2005) 2953–2968, https://doi.org/10.1101/gad.1371305.
[19] G. Robertson, M. Hirst, M. Bainbridge, M. Bilenky, Y. Zhao, T. Zeng, G. Euskirchen, B. Bernier, R. Varhol, A. Delaney, N. Thiessen, O.L. Griffith, A. He, M. Marra, M. Snyder, S. Jones, Genome-wide profiles of STAT1 DNA association using chromatin immunoprecipitation and massively parallel sequencing, Nat. Methods 4 (8) (2007) 651–657, https://doi.org/10.1038/nmeth1068.
[20] B. Testoni, C. Völlenkle, F. Guerrieri, S. Gerbal-Chaloin, G. Blandino, M. Levrero, Chromatin dynamics of gene activation and repression in response to interferon α(IFNα) reveal new roles for phosphorylated and unphosphorylated forms of the transcription factor STAT2, J. Biol. Chem. 286 (23) (2011) 20217–20227, https://doi.org/10.1074/jbc.M111.231068.
[21] K. Fink, N. Grandvaux, STAT2 and IRF9 beyond ISGF3, Jak-Stat 2 (4) (2013) e27521, https://doi.org/10.4161/jkst.27521.
[22] K. Blaszczyk, A. Olejnik, H. Nowicka, L. Ozgyin, Y.L. Chen, S. Chmielewski, K. Kostyrko, J. Wesoly, B.L. Balint, C.K. Lee, H.A. Bluyssen, STAT2/IRF9 directs a prolonged ISGF3-like transcriptional response and antiviral activity in the absence of STAT1, Biochem. J. 466 (3) (2015) 511–524, https://doi.org/10.1042/BJ20140644.
[23] K. Blaszczyk, H. Nowicka, K. Kostyrko, A. Antonczyk, J. Wesoly, H.A. Bluyssen, The unique role of STAT2 in constitutive and IFN-induced transcription and antiviral responses, Cytokine and Growth Factor Reviews 29 (2016) 71–81, https://doi.org/10.1016/j.cytogfr.2016.02.010. URL https://doi.org/10.1016/j.cytogfr.2016.02.010.
[24] E. Platanitis, D. Demiroz, A. Schneller, K. Fischer, C. Capelle, M. Hartl, T. Gossenreiter, M. Müller, M. Novatchkova, T. Decker, A molecular switch from STAT2-IRF9 to ISGF3 underlies interferon-induced gene transcription, Nature Communications 10 (1) (2019) 1–17, https://doi.org/10.1038/s41467-019-10970-y. URL https://doi.org/10.1038/s41467-019-10970-y.
[25] N.D. Heintzman, G.C. Hon, R.D. Hawkins, P. Kheradpour, A. Stark, L.F. Harp, Z. Ye, L.K. Lee, R.K. Stuart, C.W. Ching, K.A. Ching, J.E. Antosiewicz-Bourget, H. Liu, X. Zhang, R.D. Green, V.V. Lobanenkov, R. Stewart, J.A. Thomson, G.E. Crawford, M. Kellis, B. Ren, Histone modifications at human enhancers reflect global cell-type-specific gene expression, Nature 459 (7243) (2009) 108–112, https://doi.org/10.1038/nature07829.
[26] L. Song, Z. Zhang, L.L. Grasfeder, A.P. Boyle, P.G. Giresi, B.K. Lee, N.C. Sheffield, S. Gräf, M. Huss, D. Keefe, Z. Liu, D. London, R.M. McDaniell, Y. Shibata, K. A. Showers, J.M. Simon, T. Vales, T. Wang, D. Winter, Z. Zhang, N.D. Clarke, E. Birney, V.R. Iyer, G.E. Crawford, J.D. Lieb, T.S. Furey, Open chromatin defined by DNaseI and FAIRE identifies regulatory elements that shape cell-type identity, Genome Res. 21 (10) (2011) 1757–1767, https://doi.org/10.1101/gr.121541.111.
[27] M. Kellis, B. Wold, M.P. Snyder, B.E. Bernstein, A. Kundaje, G.K. Marinov, L. D. Ward, E. Birney, G.E. Crawford, J. Dekker, I. Dunham, L.L. Elnitski, P. J. Farnham, E.A. Feingold, M. Gerstein, M.C. Giddings, D.M. Gilbert, T.R. Gingeras, E.D. Green, R. Guigo, T. Hubbard, J. Kent, J.D. Lieb, R.M. Myers, M.J. Pazin, B. Ren, J.A. Stamatoyannopoulos, Z. Weng, K.P. White, R.C. Hardison, Defining functional DNA elements in the human genome, Proceedings of the National Academy of Sciences of the United States of America 111 (17) (2014) 6131–6138, https://doi.org/10.1073/pnas.1318948111.
[28] I. Dunham, A. Kundaje, S.F. Aldred, P.J. Collins, E. Al., An integrated encyclopedia of DNA elements in the human genome, Nature 489 (7414) (2012) 57–74. doi: 10.1038/nature11247.
[29] Roadmap Epigenomics Consortium, A. Kundaje, W. Meuleman, J. Ernst, E. Al., Integrative analysis of 111 reference human epigenomes, Nature 518 (7539) (2015) 317–329. doi:10.1038/nature14248.
[30] G. Natoli, S. Ghisletti, I. Barozzi, The genomic landscapes of inflammation, Genes and Development 25 (2) (2011) 101–106, https://doi.org/10.1101/gad.2018811.
[31] S. Sun, L.B. Barreiro, The epigenetically-encoded memory of the innate immune system, Curr. Opin. Immunol. 65 (2020) 7–13, https://doi.org/10.1016/j.coi.2020.02.002. URL https://doi.org/10.1016/j.coi.2020.02.002.
[32] T.C. Fang, U. Schaefer, I. Mecklenbrauker, A. Stienen, S. Dewell, M.S. Chen, I. Rioja, V. Parravicini, R.K. Prinjha, R. Chandwani, M.R. MacDonald, K. Lee, C. M. Rice, A. Tarakhovsky, Histone H3 lysine 9 di-methylation as an epigenetic signature of the interferon response, J. Exp. Med. 209 (4) (2012) 661–669, https://doi.org/10.1084/jem.20112343.
[33] N. Au-Yeung, C.M. Horvath, Histone H2A.Z Suppression of Interferon-Stimulated Transcription and Antiviral Immunity Is Modulated by GCN5 and BRD2, iScience 6 (2018) 68–82. doi:10.1016/j.isci.2018.07.013. doi: 10.1016/j.isci.2018.07.013.
[34] S. Kadota, K. Nagata, Silencing of IFN-stimulated gene transcription is regulated by histone H1 and its chaperone TAF-I, Nucleic Acids Res. 42 (12) (2014) 7642–7653, https://doi.org/10.1093/nar/gku485.
[35] S. Mostafavi, H. Yoshida, D. Moodley, H. Leboité, K. Rothamel, T. Raj, C.J. Ye, N. Chevrier, S.Y. Zhang, T. Feng, M. Lee, J.L. Casanova, J.D. Clark, M. Hegen, J. B. Telliez, N. Hacohen, P.L. De Jager, A. Regev, D. Mathis, C. Benoist, Parsing the Interferon Transcriptional Network and Its Disease Associations, Cell 164 (3) (2016) 564–578, https://doi.org/10.1016/j.cell.2015.12.032.
[36] S.H. Park, K. Kang, E. Giannopoulou, Y. Qiao, K. Kang, G. Kim, K.H. Park-Min, L. B. Ivashkiv, Type i interferons and the cytokine TNF cooperatively reprogram the macrophage epigenome to promote inflammatory activation, Nat. Immunol. 18 (10) (2017) 1104–1116, https://doi.org/10.1038/ni.3818.
[37] R. Kamada, W. Yang, Y. Zhang, M.C. Patel, Y. Yang, R. Ouda, A. Dey, Y. Wakabayashi, K. Sakaguchi, T. Fujita, T. Tamura, J. Zhu, K. Ozato, Interferon stimulation creates chromatin marks and establishes transcriptional memory, Proceedings of the National Academy of Sciences of the United States of America 115 (39) (2018) E9162–E9171. doi:10.1073/pnas.1720930115.
[38] Z. Czimmerer, B. Daniel, A. Horvath, D. Rückerl, G. Nagy, M. Kiss, M. Peloquin, M. M. Budai, I. Cuaranta-Monroy, Z. Simandi, L. Steiner, B. Nagy, S. Poliska, C. Banko, Z. Bacso, I.G. Schulman, S. Sauer, J.F. Deleuze, J.E. Allen, S. Benko, L. Nagy, The Transcription Factor STAT6 Mediates Direct Repression of Inflammatory Enhancers and Limits Activation of Alternatively Polarized Macrophages, Immunity 48 (1) (2018), https://doi.org/10.1016/j.immuni.2017.12.010, 75–90.e6.
[39] R. Ostuni, V. Piccolo, I. Barozzi, S. Polletti, A. Termanini, S. Bonifacio, A. Curina, E. Prosperini, S. Ghisletti, G. Natoli, Latent enhancers activated by stimulation in differentiated cells, Cell 152 (1–2) (2013) 157–171, https://doi.org/10.1016/j.cell.2012.12.018. URL https://doi.org/10.1016/j.cell.2012.12.018.
[40] R. Fan, A. Toubal, S. Goñi, K. Drareni, Z. Huang, F. Alzaid, R. Ballaire, P. Ancel, N. Liang, A. Damdimopoulos, I. Hainault, A. Soprani, J. Aron-Wisnewsky, F. Foufelle, T. Lawrence, J.F. Gautier, N. Venteclef, E. Treuter, Loss of the co-repressor GPS2 sensitizes macrophage activation upon metabolic stress induced by obesity and type 2 diabetes, Nat. Med. 22 (7) (2016) 780–791, https://doi.org/10.1038/nm.4114.
[41] V.M. Link, S.H. Duttke, H.B. Chun, I.R. Holtman, E. Westin, M.A. Hoeksema, Y. Abe, D. Skola, C.E. Romanoski, J. Tao, G.J. Fonseca, T.D. Troutman, N.J. Spann, T. Strid, M. Sakai, M. Yu, R. Hu, R. Fang, D. Metzler, B. Ren, C.K. Glass, Analysis of Genetically Diverse Macrophages Reveals Local and Domain-wide Mechanisms that Control Transcription Factor Binding and Function, Cell 173 (7) (2018), https://doi.org/10.1016/j.cell.2018.04.018, 1796–1809.e17, URL https://doi.org/10.1016/j.cell.2018.04.018.
[42] J. Brumbaugh, B. Di Stefano, X. Wang, M. Borkent, E. Forouzmand, K.J. Clowers, F. Ji, B.A. Schwarz, M. Kalocsay, S.J. Elledge, Y. Chen, R.I. Sadreyev, S.P. Gygi, G. Hu, Y. Shi, K. Hochedlinger, Nudt21 Controls Cell Fate by Connecting Alternative Polyadenylation to Chromatin Signaling, Cell 172 (1–2) (2018), https://doi.org/10.1016/j.cell.2017.11.023, 106–120.e21, URL https://doi.org/10.1016/j.cell.2017.11.023.
[43] J.A. Daniel, M.A. Santos, Z. Wang, C. Zang, K.R. Schwab, M. Jankovic, D. Filsuf, H. T. Chen, A. Gazumyan, A. Yamane, Y.W. Cho, H.W. Sun, K. Ge, W. Peng, M. C. Nussenzweig, R. Casellas, G.R. Dressler, K. Zhao, A. Nussenzweig, PTIP promotes chromatin changes critical for immunoglobulin class switch recombination, Science 329 (5994) (2010) 917–923, https://doi.org/10.1126/science.1187942.
[44] A. Sabò, T.R. Kress, M. Pelizzola, S. De Pretis, M.M. Gorski, A. Tesi, M.J. Morelli, P. Bora, M. Doni, A. Verrecchia, C. Tonelli, G. Fagà, V. Bianchi, A. Ronchi, D. Low, H. Müller, E. Guccione, S. Campaner, B. Amati, Selective transcriptional regulation by Myc in cellular growth control and lymphomagenesis, Nature 511 (7510) (2014) 488–492, https://doi.org/10.1038/nature13537.

[45] D.H. Phanstiel, K. Van Bortle, D. Spacek, G.T. Hess, M.S. Shamim, I. Machol, M. I. Love, E.L. Aiden, M.C. Bassik, M.P. Snyder, Static and Dynamic DNA Loops form AP-1-Bound Activation Hubs during Macrophage Development, Mol. Cell 67 (6) (2017) 1037–1048.e6, https://doi.org/10.1016/j.molcel.2017.08.006. URL https://doi.org/10.1016/j.molcel.2017.08.006.

[46] N. Mohaghegh, D. Bray, J. Keenan, A. Penvose, K.K. Andrilenas, V. Ramlall, T. Siggers, NextPBM: A platform to study cell-specific transcription factor binding and cooperativity, Nucleic Acids Res. 47 (6) (2019), https://doi.org/10.1093/nar/gkz020.

[47] V. Nurminen, A. Neme, S. Seuter, C. Carlberg, The impact of the vitamin D-modulated epigenome on VDR target gene regulation, Biochimica et Biophysica Acta - Gene Regulatory Mechanisms 1861 (8) (2018) 697–705, https://doi.org/10.1016/j.bbagrm.2018.05.006. URL https://doi.org/10.1016/j.bbagrm.2018.05.006.

[48] B. Langmead, C. Trapnell, M. Pop, S.L. Salzberg, Ultrafast and memory-efficient alignment of short DNA sequences to the human genome, Genome Biol. 10 (3) (2009), https://doi.org/10.1186/gb-2009-10-3-r25.

[49] Y. Zhang, T. Liu, C.A. Meyer, J. Eeckhoute, D.S. Johnson, B.E. Bernstein, C. Nussbaum, R.M. Myers, M. Brown, W. Li, X.S. Shirley, Model-based analysis of ChIP-Seq (MACS), Genome Biol. 9 (9) (2008), https://doi.org/10.1186/gb-2008-9-9-r137.

[50] B.S. Carvalho, R.A. Irizarry, A framework for oligonucleotide microarray preprocessing, Bioinformatics 26 (19) (2010) 2363–2367, https://doi.org/10.1093/bioinformatics/btq431.

[51] M.E. Ritchie, B. Phipson, D. Wu, Y. Hu, C.W. Law, W. Shi, G.K. Smyth, limma powers differential expression analyses for RNA-sequencing and microarray studies, Nucleic acids research 43 (7) (2015) e47, https://doi.org/10.1093/nar/gkv007, http://www.ncbi.nlm.nih.gov/pubmed/25605792%0Ahttp://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC4402510.

[52] R. Patro, G. Duggal, M.I. Love, R.A. Irizarry, C. Kingsford, Salmon provides fast and bias-aware quantification of transcript expression, Nat. Methods 14 (4) (2017) 417–419, https://doi.org/10.1038/nmeth.4197.

[53] M.D. Robinson, D.J. McCarthy, G.K. Smyth, edgeR: A Bioconductor package for differential expression analysis of digital gene expression data, Bioinformatics 26 (1) (2009) 139–140, https://doi.org/10.1093/bioinformatics/btp616.

[54] C. Wang, B. Gong, P.R. Bushel, J. Thierry-Mieg, D. Thierry-Mieg, J. Xu, H. Fang, H. Hong, J. Shen, Z. Su, J. Meehan, X. Li, L. Yang, H. Li, P.P. Łabaj, D.P. Kreil, D. Megherbi, S. Gaj, F. Caiment, J. Van Delft, J. Kleinjans, A. Scherer, V. Devanarayan, J. Wang, Y. Yang, H.R. Qian, L.J. Lancashire, M. Bessarabova, Y. Nikolsky, C. Furlanello, M. Chierici, D. Albanese, G. Jurman, S. Riccadonna, M. Filosi, R. Visintainer, K.K. Zhang, J. Li, J.H. Hsieh, D.L. Svoboda, J.C. Fuscoe, Y. Deng, L. Shi, R.S. Paules, S.S. Auerbach, W. Tong, The concordance between RNA-seq and microarray data depends on chemical treatment and transcript abundance, Nat. Biotechnol. 32 (9) (2014) 926–932, https://doi.org/10.1038/nbt.3001.

[55] J. Friedman, T. Hastie, R. Tibshirani, Regularization Paths for Generalized Linear Models via Coordinate Descent, Journal of Statistical Software 33 (1) (2010) 1–22. arXiv:0908.3817, doi:10.1016/j.expneurol.2008.01.011.

[56] T. Sing, O. Sander, N. Beerenwinkel, T. Lengauer, ROCR: Visualizing classifier performance in R, Bioinformatics 21 (20) (2005) 3940–3941, https://doi.org/10.1093/bioinformatics/bti623.

[57] U. Raudvere, L. Kolberg, I. Kuzmin, T. Arak, P. Adler, H. Peterson, J. Vilo, G: Profiler: A web server for functional enrichment analysis and conversions of gene lists (2019 update), Nucleic Acids Res. 47 (W1) (2019) W191–W198, https://doi.org/10.1093/nar/gkz369.

[58] T. Barrett, S.E. Wilhite, P. Ledoux, C. Evangelista, I.F. Kim, M. Tomashevsky, K. A. Marshall, K.H. Phillippy, P.M. Sherman, M. Holko, A. Yefanov, H. Lee, N. Zhang, C.L. Robertson, N. Serova, S. Davis, A. Soboleva, NCBI GEO: Archive for functional genomics data sets - Update, Nucleic Acids Res. 41 (D1) (2013) 991–995, https://doi.org/10.1093/nar/gks1193.

[59] C.A. Davis, B.C. Hitz, C.A. Sloan, E.T. Chan, J.M. Davidson, I. Gabdank, J.A. Hilton, K. Jain, U.K. Baymuradov, A.K. Narayanan, K.C. Onate, K. Graham, S.R. Miyasato, T.R. Dreszer, J.S. Strattan, O. Jolanki, F.Y. Tanaka, J.M. Cherry, The Encyclopedia of DNA elements (ENCODE): Data portal update, Nucleic Acids Res. 46 (D1) (2018) D794–D801, https://doi.org/10.1093/nar/gkx1081.

[60] T.G.O. Consortium, The Gene Ontology Resource: 20 years and still GOing strong, Nucleic Acids Res. 47 (D1) (2019) D330–D338, https://doi.org/10.1093/nar/gky1055.

[61] H. Mi, X. Huang, A. Muruganujan, H. Tang, C. Mills, D. Kang, P.D. Thomas, PANTHER version 11: Expanded annotation data from Gene Ontology and Reactome pathways, and data analysis tool enhancements, Nucleic Acids Res. 45 (D1) (2017) D183–D189, https://doi.org/10.1093/nar/gkw1138.

[62] H.A. Bluyssen, D.E. Levy, Stat2 is a transcriptional activator that requires sequence-specific contacts provided by Stat1 and p48 for stable interaction with DNA, J. Biol. Chem. 272 (7) (1997) 4600–4605, https://doi.org/10.1074/jbc.272.7.4600.

[63] N. Au-Yeung, R. Mandhana, C.M. Horvath, Transcriptional regulation by STAT1 and STAT2 in the interferon JAK-STAT pathway, Jak-Stat 2 (3) (2013) e23931, https://doi.org/10.4161/jkst.23931.