

MIT Open Access Articles

Parsing the Interferon Transcriptional Network and Its Disease Associations

The MIT Faculty has made this article openly available. **Please share** how this access benefits you. Your story matters.

Citation: Mostafavi, Sara et al. "Parsing the Interferon Transcriptional Network and Its Disease Associations." *Cell* 164.3 (2016): 564–578.

As Published: <http://dx.doi.org/10.1016/j.cell.2015.12.032>

Publisher: Elsevier

Persistent URL: <http://hdl.handle.net/1721.1/107116>

Version: Author's final manuscript: final author's manuscript post peer review, without publisher's formatting or copy editing

Terms of use: Creative Commons Attribution-NonCommercial-NoDerivs License





Published in final edited form as:

Cell. 2016 January 28; 164(3): 564–578. doi:10.1016/j.cell.2015.12.032.

Parsing the interferon transcriptional network and its disease associations

Sara Mostafavi^{1,2,8}, Hideyuki Yoshida^{1,8}, Devapregasan Moodley¹, Hugo LeBoité¹, Katherine Rothamel¹, Towfique Raj^{3,5}, Chun Jimmie Ye³, Nicolas Chevrier⁴, Shen-Ying Zhang⁶, Ting Feng¹, Mark Lee³, Jean-Laurent Casanova⁶, James D. Clark⁷, Martin Hegen⁷, Jean-Baptiste Telliez⁷, Nir Hacohen³, Philip L. De Jager^{3,5}, Aviv Regev³, Diane Mathis^{1,*}, Christophe Benoist^{1,*}, and the Immunological Genome Project Consortium

¹Division of Immunology, Department of Microbiology and Immunobiology, Harvard Medical School, Boston, MA

²Dept. of Statistics and Dept. Medical Genetics, University of British Columbia, Vancouver, BC

³Broad Institute of MIT and Harvard, Cambridge, MA

⁴FAS Center for Systems Biology, Harvard University, Cambridge, MA

⁵Program in Translational NeuroPsychiatric Genomics, Departments of Neurology and Psychiatry, Brigham and Women's Hospital, Boston, MA

⁶St. Giles Laboratory of Human Genetics of Infectious Diseases, The Rockefeller University, New York, NY

⁷Pfizer Immunosciences, Cambridge, MA

SUMMARY

Type-1 interferon (IFN) is a key mediator of organismal responses to pathogens, eliciting prototypical “Interferon Signature Genes” which encode antiviral and inflammatory mediators. For a global view of IFN signatures and regulatory pathways, we performed gene expression and chromatin analyses of the IFN-induced response across a range of immunocyte lineages. These distinguished ISGs by cell-type specificity, kinetics, and sensitivity to tonic IFN, and revealed underlying changes in chromatin configuration. We combined 1398 human and mouse datasets to computationally infer ISG modules and their regulators, validated by genetic analysis in both species. Some ISGs are controlled by Stat1/2 and Irf9 and the ISRE DNA motif, but others appeared dependent on non-canonical factors. This regulatory framework helped to interpret JAK1

*Address correspondence to: Diane Mathis and Christophe Benoist, Division of Immunology, Harvard Medical School, 77 Avenue Louis Pasteur, Boston, MA 02115, cbdm@hms.harvard.edu, Phone: (617) 432-7741, Fax: (617) 432-7744.

⁸Co-first author

AUTHOR CONTRIBUTIONS

HY, DM, KR, NC, SYZ, TF, NL performed experiments or contributed data; SM, HY, DM, HLB, TR, CJY, CB performed formal computational analyses; SM, HY, DM, JLC, PDJ, DM, CB wrote or edited the manuscript; JLC, JC, MH, JBT, NH, PDJ, AR, DM, CB provided supervision and secured funding.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

blockade pharmacology, different clusters being affected under tonic or IFN-stimulated conditions, and the IFN signatures previously associated with human diseases, revealing unrecognized subtleties in disease footprints, as affected by human ancestry.

INTRODUCTION

Type-1 Interferons (IFN) are a family of primordial cytokines that play a key role in responses to viral pathogens. They are a conduit between pathogen recognition by intracellular or extracellular sensors and the activation of anti-infectious mechanisms. IFN induces a set of “Interferon Signature Genes” (ISGs) that elicit a range of antiviral effectors (Schneider et al., 2014). ISGs also have pleiotropic effects on a broad array of immunological functions (Theofilopoulos et al., 2005), either directly by modulating cell-surface interaction molecules, or by perturbing signals from other cytokines or growth factors. IFN-induced responses are typically fast and have a strong positive feed-forward component, in particular the synthesis of more IFN, which ensures that a full response is activated even by low initial triggers, and spreads the alarm between neighboring cells. IFNs are thus key components of cell-intrinsic immunity, and one or more IFN types can be produced by any nucleated cells.

IFNs signal through a common heterodimeric receptor (IFNAR1/IFNAR2), or through IFNAR1 alone (de Weerd et al., 2013). The receptor subunits are associated with JAK1 and TYK2 kinases, whose activation results in phosphorylation of STAT1 and STAT2 transducers, which then associate with IRF9 to form the ISGF3 complex which activates ISG transcription (Stark and Darnell, Jr., 2012). But, besides this canonical conduit, IFN also activates several other signaling pathways: other STAT family members such as STAT3, especially in the absence of STAT1; the MAPK cascades (p38 and ERK); PI-3K and the mTOR-Akt-S6K axis (reviewed in (Uddin and Platanius, 2004; Ivashkiv and Donlin, 2014)). These additional signals directly partake in ISG induction, and likely modulate parallel cytokine or homeostatic signals. The immediate response to IFNs includes an amplification component (*Stat1*, *Irf7* and *Irf9* and some *Ifn* genes are ISGs themselves) as well as negative feedback loops: the strong negative regulators SOCS1, USP18, ISG15 are among the most strongly responsive human ISGs.

These signals elicit a canonical set of ISGs, which has been reproduced in many studies. The IFN signature includes several protein families with antiviral activity but is broader than a simple antiviral program, and the IFN signature also includes other cytokines (IL15, CSF1, TNF), chemokines, and cell-cell interaction modifiers (MHC-I, Lgals9). A common core of the IFN signature is induced in essentially all cell-types, but other ISGs are induced only in specific cell-types. It is unclear whether these cell-specific components result from differences in the balances of JAK or STAT regulators, or in the pre-existing chromatin states that enable responses at specific loci (van Boxel-Dezaire et al., 2006).

IFNs and the response they elicit are associated with a wide range of diseases. Several diseases are directly linked to an expanding group of monogenic deficiencies with enhanced or insufficient IFN activity (Crow, 2015). In addition, the interferon transcriptional signature is found elevated in blood cells in many diseases (Theofilopoulos et al., 2005; Forster,

2012), including bacterial or parasitic infections as might be expected (Berry et al., 2010) but also a range of auto-inflammatory disease like Systemic Lupus Erythematosus (SLE), Multiple Sclerosis (MS), dermatomyositis, Sjogren's disease, Scleroderma, Rheumatoid Arthritis or Sarcoidosis, and neuropsychiatric pathology. While plausible pathogenic scenarios can be proposed to account for these overexpressed IFN signatures, it is uncertain if IFN is a causative or a secondary consequence, or if a heightened IFN signature reveals over-activity of signaling pathways that are shared with the true pathogenic path.

The Immunological Genome Project (ImmGen) aims to perform a thorough dissection of gene expression and its regulation in the immune system of the mouse, including a systematic analysis of responses to cytokine triggers. Here, we have charted the transcriptional response to IFN across many hematopoietic cell-types, and performed detailed dynamic and chromatin analyses in B lymphocytes (dedicated databrowsers at www.immgen.org). We statistically integrated these data together with other large datasets that query the IFN response across genetic variation in humans and mice, and constructed a combined regulatory network for IFN. Predictions from this network, validated by genetic perturbation in both species, allowed us to parse the contributions of canonical and non-canonical signaling pathways. This novel regulatory map provides an illuminating reference to understand the association between the IFN signature and disease, as it differs across human populations, or the pharmacology of IFN-interfering drugs.

RESULTS

Cell- and time-specificity of the IFN response

We first set out to define the IFN signature across immunocyte lineages. Injection of purified IFN α *in vivo* best captured the relevant responses in unperturbed cells, and avoided unrelated responses associated with classic IFN inducers (TLR agonists) or viral infection. Type-I IFN families contain several members, whose effects largely overlap but are not completely identical (Thomas et al., 2011), but we focused this work on recombinant IFN α .

Cell-type specificity—The eleven cell-types constituting the ImmGen core lineage panel (chosen for representativity and for the possibility to purify from one mouse sufficient numbers for robust profiling) were double-sorted from C57BL/6 (B6) mice to high purity, 2 hrs after subcutaneous IFN injection. After filtering for expression and reproducibility, a set of 975 genes was found to be induced in at least one cell-type (> 2-fold induction and FDR < 0.1, induction ranging from 2 to 95-fold; heatmap of Fold Changes in Fig. 1A; Tab. S1A). A smaller set of transcripts was repressed (504 transcripts, mostly to a modest 1.5 – 3 fold; Tab. S1C). A core set of 166 ISGs was induced in all cells, which included the most robust responders (FoldChanges (FC) up to 60), highly connected in interaction databases, enriched ($p < 10^{-27}$) for antiviral effectors (*Oas*, *Ifit*, *Mx* or *Gbp* families), and included key regulators of the IFN response (e.g. *Stat1*, *Irf7/9*, *Usp18*). At the other end of the spectrum were small genesets uniquely induced in only one cell-type, mainly in granulocytes (GN) or dendritic cells (DC), but the bulk (~700) of the ISGs we uncovered were shared between several but not all cell-types (some segregated by lineage, e.g. common to all myeloid cells, or to all T lymphocytes). Cell-restricted ISGs were more modestly induced (FC 2 to 4), did not include

antiviral effectors or any specific gene family, although several of these ISGs were consistent with IFN's pro-inflammatory function (*Tnf*, *Il15*, *Csf1*, *Ccl2*, *Ccl4* in several cells, *Il23r* in GN, *Ptgs2* in DC). We calculated an "IFN score", which integrates the normalized expression of the common ISG set. Myeloid cells (GN, DC) have higher scores than lymphoid cells, at baseline and after IFN exposure (Fig. 1B).

Dynamic analysis—Finely resolved temporal transcriptome profiling can bring valuable information concerning regulatory network structure. We thus profiled CD19⁺ B lymphocytes purified at 23 times (15mn to 15hrs) after IFN injection. Most genes were induced rapidly, peaking at ~2 hrs and declining rapidly thereafter (Fig. 1C). Despite a wide range of change (FC from 2 to 40), most ISGs reached their maximum FC in a synchronized manner (per max-normalized representation of Fig. 1D), consistent with the notion that the IFN response is rapidly curtailed by negative feedback mechanisms. This temporal profile, which was similar for common and B-cell-specific ISGs (Fig. S1D), could be fit to a simple impulse model (Chechik and Koller, 2009) with a positive induction slope (Fig. 1E). After applying weighted average smoothing to the expression profiles, we quantified several temporal metrics: time to ½ onset, magnitude (maximum FC), and induction slope (Fig. 1F, Tab. S1D). A single-parameter model that considered only the magnitude of the representative impulse function (function fit to the median expression profile) explained most (~60%) of the total variance, the remaining variance being explained by subtle differences in onset time and induction slope. Indeed, plotting the induction slope vs the max FC distinguished two response modalities, one homogenous set with a unimodal relation between slope and maxFC, and another set with a more "explosive" response (steeper slope and shorter time to ½ onset; Fig. 1F).

ISGs sensitivity to IFN was evaluated across a 10⁴ range of doses in B cells (2 hrs, *in vitro* for optimal control of exposure). There were no clear subgroups but sensitivity to IFN was notably spread, with a 50-fold difference in the "ED₅₀" for half-max response (Fig. 1G; Tab. S1B). No significant association between sensitivity and kinetic parameters was noted, although there was a trend for slower-responding genes being more sensitive.

Transcriptional consequences of IFN signaling modes

Tonic activity of the IFN signaling pathways, partially in response to IFN induced by commensal microbes (Abt et al., 2012), has a perceptible influence on immunologic activity as it enhances the sensitivity to acutely produced IFN and other cytokines, likely via sustained expression of key transducers (Gough et al., 2012). *Ifnar1*-deficient mice (Muller et al., 1994), which are completely refractory to IFN, provide a means to assess which ISGs are sensitive to tonic IFN at baseline, i.e. those with reduced expression in resting *Ifnar1*-deficient mice. Profiling B cells and MFs from these mice revealed that some ISGs were under-expressed, but not all (Fig. 2A, S2A, Tab. S1E; essentially all the transcripts affected by the *Ifnar1* deficiency were ISGs). We compared this dependence on tonic IFN signals to the amplitude of IFN reactivity (Fig. 2B). While there was a trend towards higher tonic dependence for the most responsive ISGs, it was not uniform, with a clear contingent of genes with high responsiveness to IFN but low dependence on tonic signaling (blue in Fig. 2B; these outliers were not simply those with low expression at baseline, Fig. S2B). Perhaps

counter-intuitively, transcripts least dependent on tonic IFN signals tended to have the fastest responses (Fig. 2C). Together, these data suggest that ISG expression that relies on constitutive IFN does not simply reflect the most sensitive end of the IFN signaling spectrum, but an inherently different modality with quantitatively distinct targets.

IFN binding to the IFNAR receptor activates two JAK kinases, JAK1 and TYK2, which in turn activate the STAT1/2 pathway. While JAK1 is essential, TYK2 is partially dispensable for IFN responses in mice (Vogl et al., 2010). We revisited this question by profiling cells from TYK2 deficient B10.Q mice (Shaw et al., 2003) at baseline or 2 hrs after *in vivo* IFN. Comparison of the IFN response in B10.Q and B6 mice (Fig. 2D, S2C) showed a gradation of effects: some ISGs were fully responsive in TYK2-deficient cells (along the diagonal in Fig. 2D; “TYK2-independent” hereafter), while others showed decreased responsiveness to variable degrees (captured by a TYK-2 index, Tab. S1F). We related the influence of TYK2 on ISG expression at baseline and after IFN stimulation by comparing the B10.Q/B6 FoldChange in the two conditions (Fig. 2E). Many ISGs were unaffected by the deficiency in either condition, those that were affected tended to show a greater impact at either baseline or IFN-stimulated conditions. Perhaps predictably, this distinction coincided largely with the sensitivity to tonic signaling defined above (Fig. 2F). Thus, TYK2 is required at variable levels depending on the target ISG, and this preference also varies with the context of IFN signals. Strikingly, TYK2-independent ISGs were more sensitive to low IFN doses (lower ED₅₀; Fig. 2G, left), as were the tonic-sensitive ISGs (Fig. 2G right).

We performed an “enhanced sequence motif analysis”, searching for overrepresentation of TF-binding motifs in the regions of accessible chromatin at or upstream of transcriptional start sites (TSS) defined by ATAC-seq (see below). At FDR<0.05, only ISRE and related motifs scored significantly when comparing ISG and non-IGS loci in B cells ($p=10^{-36}$). Tonic-sensitive ISGs had better scores for the ISRE motif (Fig. 2H), perhaps owing to a better representation of canonical bases at positions 5' of the core motif (Fig. 2H, top). Together, these metrics distinguish different categories of ISGs: those whose response is independent of Tyk2 are tonic-sensitive, more sensitive to low doses of IFN, possibly owing to better-fitting ISRE motifs, but have slower response kinetics.

Mechanistic aspects of ISG induction: chromatin analyses

How are activating signals evoked by IFN interpreted at the transcriptional level? IFN-induced transcription may result from *de novo* recruitment of transcription factors (TF) and RNA Polymerase-II (Pol-II) complexes to ISG promoters. In addition, transcriptional elongation may be facilitated: in several systems of rapidly inducible responses, transcripts are initiated at baseline but Pol-II is blocked by elongation inhibitory factors, and the activating signal relieves this block (Adelman and Lis, 2012).

To visualize changes in Pol-II loading, we profiled chromatin-bound Pol-II by ChIP-seq of splenic B lymphocytes over several hours after *in vivo* IFN. In general, the profiles showed a strong increase of Pol-II associated with ISG promoters and intragenic regions, waning past 8hrs (Fig. 3A, S3A, Tab. S2B). This increase in gene-associated Pol-II indicated enhanced recruitment, but inspection of more promoter regions (Fig. 2B) revealed a more variegated effect. We analyzed separately ISGs that encode regulators or effectors of the response to

IFN, as Gilchrist et al have proposed that release of paused Pol-II is more common for genes encoding regulators of inductive responses than for their targets (Gilchrist et al., 2012). While most ISGs showed a strong increase in promoter-proximal Pol-II after IFN treatment, there were a few exceptions (*Setbd2* and *Zeb2*) for which Pol-II was maximal at baseline. We calculated a “pre-loading ratio” (normalized reads over the entire gene, at baseline vs at max response). Although significantly below 1 for most ISGs, we did observe a higher value overall for regulator than for effector ISGs (Fig. 3B, Wilcoxon $p=0.001$), denoting a greater extent of pre-loaded Pol-II for regulators. Directly comparing changes in Pol-II loading at the TSS vs the gene body (Fig. 3C) showed that for fast-responding ISGs the two were largely correlated and of the same magnitude (for late-responding ISGs, preferential Pol-II loading at the TSS at 30 mn preceded that in the gene body at 1 hr). To more finely appreciate how the release of promoter-proximal paused Pol-II contributed to the IFN response, we determined the traveling ratio ((Reppas et al., 2006): the ratio of Pol-II density in the body of the gene over that near the TSS) before and after induction. Traveling ratios for ISGs matched the genome-wide distribution prior to induction, and only the fastest responders showed a slight increase in this ratio during the peak response (Fig. 3D).

As for mRNA abundance, the Pol-II loading profiles at 7 time points could be fit to an impulse function, and we used the impulse model to impute additional time-points to better compare Pol-II and mRNA time-courses. There was a good concordance between temporal Pol-II and transcript abundance profiles, the time of median $\frac{1}{2}$ max increase in Pol-II loading occurring ~ 30 mn than earlier changes in mRNA (Fig. 3E – the explosive nature of these events are illustrated in movie M1). Thus, unlike what had been proposed for fast responses in the innate immune system, the response to IFN predominantly involves Pol-II recruitment to ISG loci, likely by newly bound STAT1/2 and other activators. Faster transit from the TSS may also contribute somewhat to the response, particularly for the fastest regulatory ISGs.

IFN signaling involves the nuclear translocation of activated STAT1/2 complexes that bind DNA and activate transcription. We combined chromatin immunoprecipitation (ChIPseq) and ATACseq (Buenrostro et al., 2013) to map and quantitate Stat1/2 binding (6703 Stat2 peaks; 46% increased > 2 fold by IFN, mainly mapping to TSS and enhancers, less so to super-enhancers; Tab. S2E) and the changes in chromatin accessibility that accompany ISG activation (Fig. 4, Tab. S2C). Chromatin from splenic B lymphocytes was assessed, before or 90 mn after *in vivo* IFN treatment, the time when transcription-related changes were likely to be most prominent. The representative profiles (Fig. 4A) illustrate the strong changes in chromatin accessibility in and around ISGs, with increased accessibility at the TSS and in some other intragenic or upstream locations, while other open regions in the same loci remained unchanged. When the position of these IFN-altered regions were plotted relative to the closest ISG (Fig. S4), many fell within the -2 to -200 Kb distance generally typical of upstream ATACseq motifs, but some were found at surprisingly larger distances (> 1 Mb from the closest ISG; not shown). These induced ATAC-seq peaks were enriched in ISRE motifs ($p < 10^{-20}$) and generally corresponded to increased Stat2 binding (e.g. the *Usp18* upstream enhancer, Fig. 4A). Yet in some cases changes in Stat2 binding and chromatin accessibility were disconnected (e.g. the *Bst2* promoter, where the ATAC signal

does not follow the sharply increased Stat2 binding). Thus, the response to IFN involved rapid changes in specific locations of chromatin, although some ISGs showed no increased Stat2 binding and modest chromatin changes (e.g. *Ppa1*). In addition, there were no changes in H3-K4^{me1} methylation, which we mapped as a reference for enhancer location. For a genome-wide perspective, we computed the IFN-induced change in normalized read-count relative to baseline (ATAC-FC). Of 42,316 significant peaks, 476 regions showed IFN-increased ATACseq signals (at ATAC-FC>2 and p<0.05) and 256 decreased (Tab. S2C,D). There was a strong correlation between the increased ATAC signal around the TSS and increased mRNA levels (Fig. 4B, left), and this was also the case for some, but not all, of the intragenic or upstream ATAC-seq peaks (right). For those ISGs active in B cells (from Fig. 1A), there was also a strong correlation between increased ATAC and Stat2 binding (Fig. 4C, left); interestingly, for ISGs only active in other cell-types, some degree of Stat2 binding was noted (right panel).

We also found that transcripts with different sensitivity to tonic signaling had distinct accessibility signals at the TSS: there was a general correlation between a gene's expression and the ATAC-seq signal, and tonic-independent ISGs followed the genome-wide trend; but tonic-sensitive ISGs showed a markedly lower TSS accessibility (Fig. 4D). In addition, tonic-sensitive loci showed higher Stat2 binding, at baseline and after IFN exposure (Fig. 4E), suggesting that these ISG loci have a distinct enhancer/promoter configuration.

Regulatory network inference

Network inference—We sought to computationally reconstruct a regulatory network of the transcriptional response to IFN by combining these mouse data with datasets from primary human CD4⁺ T cells or monocytes of the ImmVar study (Lee et al., 2014; Raj et al., 2014; Ye et al., 2014), from baseline or IFN-challenged conditions (in all, 1,398 profiles from 12 distinct data groups; Fig. 5A, Tab. S3A). Sources of perturbation in these data groups included genetic variation (between outbred humans or inbred mice), cell-type variation, and response time. We first verified the comparability of responses to IFN in mouse and human immunocytes, and estimated that 74% of the testable ubiquitous ISGs respond in both species (Fig. S5A). Using datasets that covered two species and different cell-types (Fig. 5A) increased power and ensured coverage of the most conserved elements of the IFN regulatory network.

Derivation of the IFN regulatory network consisted of three steps: 1) computing a co-expression network between a set of potential regulators and targets for each datagroup (Pearson correlation); 2) normalizing the co-expression networks (per-gene normalization); and 3) combining datagroup-specific networks using a “uniformly”-weighted network combination (Mostafavi and Morris, 2010), which was then sparsified (FDR threshold of 0.01). To enrich the search space for causal associations, only gene families of known regulatory function were considered as regulators (TFs, kinases, phosphatases). Targets were 120 ISGs common to both human and mouse, lymphoid and myeloid cells (FC>1.5 in both species and both cell-types). The inferred network consisted of 2,691 links between 92 predicted regulators and 102 ISGs (Tab.s S3B–F). To visually clarify regulator/target

relationships we applied the LAS biclustering algorithm (Shabalin et al., 2015) and identified a solution with five regulatory modules (C1 through C5).

These links applied in both human and mouse cells, with comparable scores for links computed with datasets from either species (Fig. S5B) and were not dominated by contribution from a few datasets in a drop-one analysis (Fig. S5C). Link scores were also high in both baseline and IFN-stimulated datasets (Fig. S5D), although links associated with regulatory clusters showed some quantitative bias (C3 links stronger in baseline conditions, C5 in IFN-stimulated datasets). This robust core network was later extended to include a larger set of 629 ISGs with more distributed responses (Tab. S3G), based on the strength of co-expression between these remaining ISGs and the core set of regulators.

Fig. 5B displays the scores for regulator/target pairs that form the five clusters, a striking pattern from which several observations emerged, serving both to validate the results and to open new vistas on ISG regulation. Most obviously, the main regulators thus identified were STAT1/2 and IRF9, which together form the canonical ISGF3 complex and, with IRF7, were the main predicted regulators of ISGs in clusters C3 and C4, consistent with the scores for core ISRE motif (Fig. 5C). But, in addition to these expected regulators, the inferred network predicted less anticipated links: other regulators correlated with induction of C3 ISGs included SP100, SP110, and NMI. Second, while C3 and C4 were predominantly controlled by ISGF3 and consorts, these factors were predicted to have less or no impact on other ISG clusters, in particular C1 and C5 for which different regulators were predicted (KAT2B, TBK1, LYN, ETV6; Fig. 5B). Fig 5D highlights the highly coordinated control of C3 members.

Parsing ISGs into these five clusters (using the extended gene lists, Tab. S3G) brought forth functional distinctions. C3 was composed predominantly of antiviral effectors (e.g. *OAS* and *IFIT* families, *RSDA2*; $p < 10^{-25}$), as well as the key positive and negative regulators (*USP18*, *STAT1* and *IRF9* themselves). In contrast, the other clusters contained essentially no antiviral components, but were enriched for RNA processing (C1 and C2, $p < 10^{-5}$), metabolic regulation (C4, $p < 10^{-5}$), and inflammation mediators or regulators (*IL15RA*, *CD274*, *CXCL10* in C5, $p < 10^{-9}$). C3 was strongly enriched for the ISGF3-binding ISRE motif (Fig. 5C; $p < 10^{-20}$), but ISRE was under-represented in peaks located around non-C3 loci when compared to C3 ($p < 10^{-3}$), most strikingly for C5 (Fig. 5C), concordant with its independence of STAT1/2. Another motif associated with IFN responses, the gamma-activated-sequence (GAS) motif, was only mildly enriched among these ISGs ($p < 10^{-4}$), and with no preference for any of the clusters.

Experimental Network Validation—In addition to re-discovering known regulators, we validated the IFN regulatory network using genetic perturbation. First, we used data from RNAi knockdown experiments after LPS stimulation (Chevrier et al., 2011) (a large component of LPS response is a secondary response to induced IFN). We could directly assess the accuracy of 20% of our predicted links with this dataset, observing a significant agreement between the inferred links and the RNAi results (Fig. 5E; AUROC ranging from 0.6–0.8, Fig. S5F). This applied to expected links involving STAT1/2 or IRF9, but also to more novel ones involving NMI, ETV6 or ATF3 as regulators. We also tested

transcriptional responses to IFN in *Lyn*-deficient cells (in mixed bone-marrow transfers). Although *LYN*, a predicted regulator of *C5*, has previously been associated with responses to IFN (Uddin et al., 1998), we observed only minimal consequences of its deficiency (not shown), perhaps because of compensatory activity of *FYN* or other kinases. Finally, we tested the compatibility of our predictions with differential ISG induction in fibroblasts from a *STAT1*-deficient patient, compared with four healthy controls (Chapgier et al., 2006). As illustrated in Fig. 5F, there was good concordance between our predicted *STAT1* score and the effect of the *STAT1* deficiency on individual ISGs. Interestingly, ISGs that were off-diagonal in Fig. 5F (i.e. with less actual change than predicted by the network score) were those with a high predicted *STAT2* score, suggesting that *STAT2* complements more effectively the *STAT1* deficiency for these targets. Finally, there was a strong relationship between computational network predictions regarding ISGF3 and *Stat2* binding in ChIPseq (Fig. 5G, S5G). Target genes with the 10% highest ISGF3 network scores showed strong and highly responsive *Stat2* binding (Fig. 5G, top - all with $FC > 10$; *Irf9*, the sole exception, already had strong *Stat2* at baseline); those with the lowest ISGF3 scores had much lower and predominantly non-responsive *Stat2* binding (Fig. 5G, bottom).

Implications of the IFN regulatory model for human pathology and therapeutics

Since ISG over-expression is associated with several human pathologies, and since therapeutics that modulate IFN or its signals are used in several clinical contexts (Crow et al., 2015), it was of interest to test whether the structures defined above have direct implications for human genetic variation, pathology or therapy. We collated (list and refs in Tab. S4A) genesets differentially expressed in whole blood or PBMCs studies encompassing acute or chronic microbial infection or vaccination, autoimmune diseases and major depression (MD). These genesets mainly contained ISGs induced in all cell types, consistent with their mixed-cell origin. Several interesting observations emerged from mapping these ISGs onto the IFN regulatory network (Fig. 6A). (i) The regulatory clusters were differently represented: most ISGs overexpressed in these disease-related signatures were assigned to C3 ($p < 10^{-20}$), the remaining belonging to C5 ($p < 3 \times 10^{-6}$), with an under-representation of ISGs assigned to C1, 2 or 4 ($p < 2 \times 10^{-7}$). (ii) The distribution of C3 ISGs triggered by active microbial infections (TB, Candidiasis, Yellow fever vaccine) differed from those in autoimmune diseases or MD (Figs. 6A, S6A), the latter including relatively less of the most *STAT1/2*-dependent targets at the top of the cluster. Most different were the chronic viral infections states (EBV, CVID, HIV), with even less involvement of C3 ISGs (chronic HCV appeared more like an acute infection in this regard), (iii) The ISGs over-expressed in IFN-resistant MS patients were also shifted, with an overrepresentation of the bottom of C3.

We analyzed these relationships more closely for SLE, one of the diseases for which the relation between dysregulated IFN and the pathology is perhaps strongest. We retrieved available PBMC expression profiles from two recent studies that compared SLE patients and age/gender-matched controls, with cohorts of comparable sizes and age distribution, but of European and East Asian ancestry, respectively (Chiche et al., 2014; Lee et al., 2011). After modeling the effect of standard confounding factors (age, gender, and batch), we computed the association score between diagnosis of SLE and the expression levels of individual genes in the two studies (“Osaka” and “Marseille”, Fig. 6B, Tab. S4B). Most associated genes

were ISGs (enrichment $p < 10^{-20}$), mostly from C3 ($p < 10^{-6}$), only a few C5 ISGs showing a significant association. SLE-associated transcripts predominantly belonged to the tonic-sensitive category (Fig. 6C). There was good concordance between the two studies (overlap $\pi_1 = 49\%$), except that roughly half of the ISGs associated with SLE in Marseille gave little or no signal in Osaka (although we cannot formally rule out clinical differences between the cohorts, the range of disease activity was similar in the two studies). We searched the various scores and clusters described above, and the requirement for TYK2 appeared to distinguish SLE-associated gene sets: the majority of SLE-associated ISGs that scored in both studies were sensitive to TYK2 deletion in baseline conditions (consistent with their sensitivity to tonic signaling). In contrast, most of those scoring in Marseille but not Osaka showed less requirement for TYK2 for expression at baseline but more so under IFN-challenged conditions (and correspondingly belonged to the tonic-insensitive group). This dichotomy was evocative, since a *TYK2* SNP is reproducibly associated with SLE in GWAS studies of European but not Asian cohorts, even though the SNPs have a high frequency in both populations (see (Lee et al., 2012) for refs and meta-analysis). This variant did not differentially influence baseline or induced ISG levels in ImmVar donors of either ancestry (not shown). Thus, disease-associated ISGs are far from being a homogenous set that would be activated in lockstep; rather, they belong to specific regulatory clusters, with subtle variations between diseases and between human populations.

Finally, we asked if the different facets of the network would be similarly impacted by pharmacological inhibitors of JAK kinase activity (JAKi). For best relevance to the clinical setting, these experiments were performed *in vivo* and, given the indications above that targets of tonic and active IFN activation pathways are quantitatively different, we tested the effect of JAK inhibition in both contexts. The kinase inhibitor Tofacitinib (Tofa), a small molecule that preferentially blocks JAK1 activity (Clark et al., 2014), was administered in a clinically relevant regimen to mimic the human exposure, at doses that largely block phospho-Stat1 activation by IFN, alone or prior to a 2-hr parenteral IFN challenge. Under baseline conditions, a small set of transcripts was downregulated in splenic B cells after Tofa treatment, most of which were ISGs (Fig. 7A). These effects were modest (2-fold downregulation at most), also observed in other cell-types (not show), and vanished in *Ifnar1*-deficient mice (Fig. S7A). Not all ISGs were affected and fittingly since it is tonic signaling that Tofa inhibits under these conditions, these JAK1i-downregulated transcripts were those most influenced by tonic IFN (Fig. S7A). When Tofa was given in the context of an IFN challenge, a partial reduction was observed, which again did not affect all ISGs equally (captured by a drug “IFN-inhibition index” Fig. 7B). Interestingly, this inhibition was inversely related to the effect of the inhibitor on the same ISGs under tonic conditions (Fig. 7C, Tab. S5A). The distinction was even more striking when the responses were mapped onto the network structure defined above. At baseline, JAKi primarily affected tightly connected ISGs of the C3 cluster (Fig. 7D, left), but these were largely spared by the drug under IFN challenge when ISGs belonging to C5 or C2 were those most affected. Thus, inhibition of JAK1 signaling affects different targets under tonic or IFN-challenged conditions (much as the TYK2 deficiency did), and these sets correspond to distinct facets of the IFN regulatory network.

DISCUSSION

In line with ImmGen goals, we have generated a deep resource of transcriptional data surrounding the response to IFN *in vivo*, probing the regulatory network with diverse set of perturbants in a large-scale integration of datasets from mouse and human. While the results depict a response that is rather monotonic in timing, contrasting for instance with the cascading events that follow antigen-driven T cell activation, the analysis revealed a multifaceted ensemble that is far from monomorphic, involving several signaling and regulatory controls. Distinct sub-signatures are defined by the kinetic and regulatory parameters, and these facets are intertwined, such that induction kinetics, tonic effects, chromatin structure, TYK2-dependence and pharmacologic specificity are all interconnected. For instance, ISGs that are sensitive to tonic signaling at baseline are a subgroup with high sensitivity to low IFN doses but slower response kinetics, high dependence on ISRE and ISGF3, preferential association to SLE, and high sensitivity to JAKi treatment in the unchallenged state. These distinctions have direct implications for our understanding of the mechanisms of the IFN response, its association to human pathologies and for pharmacologic modulation.

Time-resolved chromatin analyses paint a clear picture of the transcriptional events that lead to fast ISG transcript accumulation. First, chromatin changes are observed at all ISGs and are largely correlated with ISG mRNA accumulation, consistent with the results of Jovanovic et al, who noted that ISGs are induced by an increased transcription rate combined with constantly fast mRNA processing and degradation rates (Jovanovic et al., 2015), with no evidence for shifts in post-transcriptional control (miRNA, translational stabilization). Second, the increase of promoter-paused Pol-II shortly preceding mRNA buildup, together with the changes in chromatin accessibility that correlate with mRNA induction, are more compatible with a model of transcriptional induction via heightened formation of initiation complexes, rather than by release of blocked Pol-II. On the other hand, facilitated elongation does take part in ISG induction, with a transiently increased traveling ratio.

Tonic sensitive genes appear to have high Stat1/2 binding at baseline and after activation, despite TSS chromatin counterintuitively less accessible than the genome-wide average. There is a relative dearth of CpG islands around the TSS of these genes (19% vs 47% genomewide, $p=2\times 10^{-5}$), in keeping with the results of Carrozzi et al (Ramirez-Carrozzi et al., 2009) who showed that many ISGs belong to a class of inducible genes that require SWI/SNF-dependent nucleosome remodeling for their activation. This may account for the slightly delayed response of these ISGs.

The detailed analysis of cell-specific responses provides a roadmap of IFN-induced responses across cell-types, beyond the core response shared by all cells (van Boxel-Dezaire et al., 2006). Indeed, the most frequent patterns are those with variegated expression, shared between some but not all immunocyte lineages, often less dramatically induced than the common antiviral ISGs, although with the same impulse response peaking at 2 hrs. For those variegated ISGs that could be considered by our network inference, we observe a preponderance of non-canonical regulators, less frequent ISRE motifs at the TSS relative to

the set of common ISGs, and correspondingly membership in clusters C1, C2, or C5 (of the 158 of these variegated ISGs, 117 were assigned to C1, C2, or C5, in the expanded cluster analysis).

These data, together with those generated in the context of the ImmVar project, allowed us to perform an integrative analysis, unusual in its scale, benefiting from the trans-species conservation of the IFN response and from the diverse sources of perturbation in the datasets: genetic variation, time course of activation, cell differentiation. There was support for the links in several datasets, and the predictions were validated with both human and mouse genetic deficiencies. However, this strategy did limit the scope to the frequently induced ISGs, and did not comprehensively extend to the most cell-specific components. Comfortingly, the co-expression-based network rediscovered the ISGF3 complex (STAT1/2, IRF9) as the dominant regulator, and the STAT1/2 scores predicted by the model concord well with the consequence of the mutation in a STAT1-deficient patient's fibroblasts. In addition, the network brings forth previously unrecognized regulators (ETV6, ATF3) that were validated by RNAi knockdown, and some (LYN, TBK1 and the NF- κ B pathway) with previously reported relevance (Uddin and Platanius, 2004; Tenover et al., 2007), but which were not followed upon as extensively as STATs in recent years. Outside of the JAK/STAT axis, IFN is documented to signal via the PI3K/mTOR, MAPK and Crk pathways, and it is plausible that the different transcriptional clusters reflect these parallel signaling routes. The results also highlight the regulatory heterogeneity of ISGs, a substantial fraction of which have little or no links to STAT1/2 and ISGF3, paralleling the fraction of the response to IFN γ that does not require STAT1 (Gil et al., 2001).

Admittedly, several known players are missing from the predicted network, e.g. known negative-feedback regulators (SOCS1 or USP18) and most strikingly the essential signal transducer, JAK1. Network inference requires some variation in the data, genetic or otherwise, to support the links, and JAK1 may not have scored because it is not responsive to activation and has limited genetic variation across healthy mice or humans (JAK1 also stands out in having no described disease association in GWAS databases). This invariance contrasts with the greater tolerance for variation in TYK2, reflected in inter-population variation and in GWAS hits.

The distinction between the regulatory structures at play during tonic signaling versus acute IFN exposure mirrors the transcript signature of pharmacologic JAK inhibition: treatment mainly affects the STAT-controlled antiviral cluster C3 at baseline. Yet, in conditions of IFN challenge that are more comparable to those encountered during acute viral infection, ISGs regulated by non-canonical factors, mostly pro-inflammatory in nature, are most affected. This difference might explain why JAK1 blockade does not result in significant exacerbation of acute infections (Cohen et al., 2014) when the antiviral program would be least affected by the drug. On the other hand, perturbation of antiviral mechanisms under tonic IFN may be linked to the herpes zoster reactivation observed in a number of patients (Clark et al., 2014), implying that tonic IFN is relevant for the control of viral latency.

Finally, the parsing of IFN targets and regulatory pathways sheds new light on the over-represented IFN signature found in a number of diseases, and provides a reference map

against which the functional and regulatory diversity of these signatures can be appreciated. With the exception of monogenic interferonopathies in which IFN is clearly the culprit (Crow, 2015), it is a matter of debate whether these signatures are causally related to the disease, or an ancillary side-effect of the true disease root or of downstream inflammation. The differences in cluster membership or ISRE scores between IFN signatures in acute/active infection, chronic infection and autoimmune conditions indicate that these signatures are not merely a generic indicator of IFN overproduction, but reflect different relationships to pathogenesis. For instance, the signature of chronic hepatitis C virus infection resembles that of recent vaccination or active TB more than that of other chronic infectious states like EBV or HIV, suggesting that chronic HCV may behave like a recurrent acute infection. The comparable signatures of autoimmune disease and major depression prompt the speculation of an unrecognized auto-inflammatory component to the latter.

Mapping SLE-associated ISGs against this background revealed a striking dichotomy for European vs Asian ancestry. Half of the ISGs associated with SLE in European patients showed no association whatsoever in Japanese patients. These two gene sets differed with regards to tonic IFN sensitivity and to dependence on TYK2. It may not be a coincidence that *TYK2* genetic variants are associated with SLE in European but not East Asian cohorts, when *TYK2* is associated with SLE in European but not Asian populations. A parsimonious interpretation is that two distinct ISG components are dysregulated in SLE: one shared between European and Asian patients, primarily reflecting heightened tonic IFN; another requiring TYK2 and solely over-expressed in European patients through a mechanism that remains unclear (these ISGs are equally responsive to IFN in cells from donors of Asian ancestry). Limitations to these interpretations include differences in study designs and statistical power. Regardless of the answer, however, these results suggest that the relationship between IFN and SLE is not a simple one.

In conclusion, profiling and integrating diverse genomic data types relevant to IFN transcriptional response enabled us to discover varied regulatory mechanisms for induction of ISGs. This complexity highlights, and at the same time facilitates tackling, the heterogeneity of interferon disease signatures in diverse immune diseases.

EXPERIMENTAL PROCEDURES

Data can be viewed with dedicated databrowsers on the ImmGen website (www.immgen.org) and have been deposited at NCBI Gene Expression Omnibus under GSE75306

Mice

C57BL/6J, *Tyk2* spontaneous mutant B10.Q (JAX 2024), and *Ifnar1* KO mice from the Jackson Laboratory were analyzed 2 hrs after 10,000 U IFN α sc unless otherwise specified. Tofacitinib was administered by oral gavage, 25 mg/kg twice daily.

Expression profiling

Immunocytes were double-sorted and profiled on Affymetrix ST1.0 microarrays according to ImmGen SOP (Heng et al., 2008). For the dose-response, magnetically purified splenic B cells were treated *in vitro* with graded doses of IFN α for 2hrs.

Chromatin

B splenocytes were magnetically-purified after *in vivo* IFN. ATACseq libraries were constructed exactly per (Buenrostro et al., 2013), and sequenced; for PolII and Stat1, 2 ChIPseq per (Rahl et al., 2010) and sequenced (Tab. S2A). After trimming and mapping to mm10, peaks were called by MACS (from merged datasets) and reads quantitated in corresponding regions of individual datasets.

ATACseq-enriched TF binding motif

ATAC-defined peak regions were scored for 263 PWMs (HOMER), and over-represented PWMs in peaks surrounding ISG loci identified by the hypergeometric test.

Regulatory network inference used 12 large baseline or IFN-stimulated datagroups (Tab. S3) from mouse and human. 1,152 potential regulators and 110 target ISGs in both species were defined from GeneOntology responsiveness to IFN, respectively. z-score normalized co-expression matrices (Wang et al., 2009) were defined for each datagroup. These networks were then combined with “uniform” weighting, where each edge in the final network is computed as the average of the corresponding links in all 12 networks. Final network was sparsified using a 0.01 FDR threshold (Wang et al., 2009). Interferon-related disease signatures were retrieved from relevant publications that used blood expression data (genes and refs in Tab. S4 ad Supplement)

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank study participants in the ImmVar project for their contribution, Drs S. Churchman, T. DeFranco, A.-S. Korganow, C. Lowell, D. Price, I. Sanz and ImmGen members for useful discussions, K. Hattori, A. Ortiz-Lopes, N. Asinowski, A. Rhoads, C. Araeno, M. Thorsen and K. Waraska for help with mice, cell sorting and sequencing. This work was supported by National Institutes of Health R24-AI072073 to the ImmGen consortium, RC2-GM093080 to ImmVar, and a Sponsored Research Agreement from Pfizer, Inc.

References

- Abt MC, Osborne LC, Monticelli LA, Doering TA, Alenghat T, Sonnenberg GF, Paley MA, Antenus M, Williams KL, Erikson J, et al. Commensal bacteria calibrate the activation threshold of innate antiviral immunity. *Immunity*. 2012; 37:158–170. [PubMed: 22705104]
- Adelman K, Lis JT. Promoter-proximal pausing of RNA polymerase II: emerging roles in metazoans. *Nat Rev Genet*. 2012; 13:720–731. [PubMed: 22986266]
- Amit I, Garber M, Chevrier N, Leite AP, Donner Y, Eisenhaure T, Guttman M, Grenier JK, Li W, Zuk O, et al. Unbiased reconstruction of a mammalian transcriptional network mediating pathogen responses. *Science*. 2009; 326:257–263. [PubMed: 19729616]

- Berry MP, Graham CM, McNab FW, Xu Z, Bloch SA, Oni T, Wilkinson KA, Banchereau R, Skinner J, Wilkinson RJ, et al. An interferon-inducible neutrophil-driven blood transcriptional signature in human tuberculosis. *Nature*. 2010; 466:973–977. [PubMed: 20725040]
- Buenrostro JD, Giresi PG, Zaba LC, Chang HY, Greenleaf WJ. Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nat Methods*. 2013; 10:1213–1218. [PubMed: 24097267]
- Chapgier A, Wynn RF, Jouanguy E, Filipe-Santos O, Zhang S, Feinberg J, Hawkins K, Casanova JL, Arkwright PD. Human complete Stat-1 deficiency is associated with defective type I and II IFN responses in vitro but immunity to some low virulence viruses in vivo. *J Immunol*. 2006; 176:5078–5083. [PubMed: 16585605]
- Chechik G, Koller D. Timing of gene expression responses to environmental changes. *J Comput Biol*. 2009; 16:279–290. [PubMed: 19193146]
- Chevrier N, Mertins P, Artyomov MN, Shalek AK, Iannacone M, Ciaccio MF, Gat-Viks I, Tonti E, DeGrace MM, Clauser KR, et al. Systematic discovery of TLR signaling components delineates viral-sensing circuits. *Cell*. 2011; 147:853–867. [PubMed: 22078882]
- Chiche L, Jourde-Chiche N, Whalen E, Presnell S, Gersuk V, Dang K, Anguiano E, Quinn C, Burtey S, Berland Y, et al. Modular transcriptional repertoire analyses of adults with systemic lupus erythematosus reveal distinct type I and type II interferon signatures. *Arthritis Rheumatol*. 2014; 66:1583–1595. [PubMed: 24644022]
- Clark JD, Flanagan ME, Telliez JB. Discovery and development of Janus kinase (JAK) inhibitors for inflammatory diseases. *J Med Chem*. 2014; 57:5023–5038. [PubMed: 24417533]
- Cohen S, Radominski SC, Gomez-Reino JJ, Wang L, Krishnaswami S, Wood SP, Soma K, Nduaka CI, Kwok K, Valdez H, et al. Analysis of infections and all-cause mortality in phase II, phase III, and long-term extension studies of tofacitinib in patients with rheumatoid arthritis. *Arthritis Rheumatol*. 2014; 66:2924–2937. [PubMed: 25047021]
- Crow MK, Olfieriev M, Kirou KA. Targeting of type I interferon in systemic autoimmune diseases. *Transl Res*. 2015; 165:296–305. [PubMed: 25468480]
- Crow YJ. Type I interferonopathies: mendelian type I interferon up-regulation. *Curr Opin Immunol*. 2015; 32:7–12. [PubMed: 25463593]
- de Weerd NA, Vivian JP, Nguyen TK, Mangan NE, Gould JA, Braniff SJ, Zaker-Tabrizi L, Fung KY, Forster SC, Beddoe T, et al. Structural basis of a unique interferon-beta signaling axis mediated via the receptor IFNAR1. *Nat Immunol*. 2013; 14:901–907. [PubMed: 23872679]
- Forster S. Interferon signatures in immune disorders and disease. *Immunol Cell Biol*. 2012; 90:520–527. [PubMed: 22491066]
- Gil MP, Bohn E, O'Guin AK, Ramana CV, Levine B, Stark GR, Virgin HW, Schreiber RD. Biologic consequences of Stat1-independent IFN signaling. *Proc Natl Acad Sci U S A*. 2001; 98:6680–6685. [PubMed: 11390995]
- Gilchrist DA, Fromm G, dos SG, Pham LN, McDaniel IE, Burkholder A, Fargo DC, Adelman K. Regulating the regulators: the pervasive effects of Pol II pausing on stimulus-responsive gene networks. *Genes Dev*. 2012; 26:933–944. [PubMed: 22549956]
- Gough DJ, Messina NL, Clarke CJ, Johnstone RW, Levy DE. Constitutive type I interferon modulates homeostatic balance through tonic signaling. *Immunity*. 2012; 36:166–174. [PubMed: 22365663]
- Heng TS, Painter MW. Immunological Genome Project Consortium. The Immunological Genome Project: networks of gene expression in immune cells. *Nat Immunol*. 2008; 9:1091–1094. [PubMed: 18800157]
- Ivashkiv LB, Donlin LT. Regulation of type I interferon responses. *Nat Rev Immunol*. 2014; 14:36–49. [PubMed: 24362405]
- Jovanovic M, Rooney MS, Mertins P, Przybylski D, Chevrier N, Satija R, Rodriguez EH, Fields AP, Schwartz S, Raychowdhury R, et al. Immunogenetics. Dynamic profiling of the protein life cycle in response to pathogens. *Science*. 2015; 347:1259038. [PubMed: 25745177]
- Lee HM, Sugino H, Aoki C, Nishimoto N. Underexpression of mitochondrial-DNA encoded ATP synthesis-related genes and DNA repair genes in systemic lupus erythematosus. *Arthritis Res Ther*. 2011; 13:R63. [PubMed: 21496236]

- Lee MN, Ye C, Villani AC, Raj T, Li W, Eisenhaure TM, Imboywa SH, Chipendo PI, Ran FA, Slowikowski K, et al. Common genetic variants modulate pathogen-sensing responses in human dendritic cells. *Science*. 2014; 343:1246980. [PubMed: 24604203]
- Lee YH, Choi SJ, Ji JD, Song GG. Associations between PXX and TYK2 polymorphisms and systemic lupus erythematosus: a meta-analysis. *Inflamm Res*. 2012; 61:949–954. [PubMed: 22592861]
- Mostafavi S, Morris Q. Fast integration of heterogeneous data sources for predicting gene function with limited annotation. *Bioinformatics*. 2010; 26:1759–1765. [PubMed: 20507895]
- Muller U, Steinhoff U, Reis LF, Hemmi S, Pavlovic J, Zinkernagel RM, Aguet M. Functional role of type I and type II interferons in antiviral defense. *Science*. 1994; 264:1918–1921. [PubMed: 8009221]
- Rahl PB, Lin CY, Seila AC, Flynn RA, McCuine S, Burge CB, Sharp PA, Young RA. c-Myc regulates transcriptional pause release. *Cell*. 2010; 141:432–445. [PubMed: 20434984]
- Raj T, Rothamel K, Mostafavi S, Ye C, Lee MN, Replogle JM, Feng T, Lee M, Asinovski N, Frohlich I, et al. Polarization of the effects of autoimmune and neurodegenerative risk alleles in leukocytes. *Science*. 2014; 344:519–523. [PubMed: 24786080]
- Ramirez-Carrozzi VR, Braas D, Bhatt DM, Cheng CS, Hong C, Doty KR, Black JC, Hoffmann A, Carey M, Smale ST. A unifying model for the selective regulation of inducible transcription by CpG islands and nucleosome remodeling. *Cell*. 2009; 138:114–128. [PubMed: 19596239]
- Reppas NB, Wade JT, Church GM, Struhl K. The transition between transcriptional initiation and elongation in *E. coli* is highly variable and often rate limiting. *Mol Cell*. 2006; 24:747–757. [PubMed: 17157257]
- Schneider WM, Chevillotte MD, Rice CM. Interferon-stimulated genes: a complex web of host defenses. *Annu Rev Immunol*. 2014; 32:513–545. [PubMed: 24555472]
- Shabalin AA, Weigman VJ, Perou CM, Nobel AB. Finding large average submatrices in high dimensional data. *The Annals of Applied Statistics*. 2015; 3:985–1012.
- Shaw MH, Boyartchuk V, Wong S, Karaghiosoff M, Ragimbeau J, Pellegrini S, Muller M, Dietrich WF, Yap GS. A natural mutation in the Tyk2 pseudokinase domain underlies altered susceptibility of B10. Q/J mice to infection and autoimmunity. *Proc Natl Acad Sci U S A*. 2003; 100:11594–11599. [PubMed: 14500783]
- Stark GR, Darnell JE Jr. The JAK-STAT pathway at twenty. *Immunity*. 2012; 36:503–514. [PubMed: 22520844]
- Tenoever BR, Ng SL, Chua MA, McWhirter SM, Garcia-Sastre A, Maniatis T. Multiple functions of the IKK-related kinase IKKepsilon in interferon-mediated antiviral immunity. *Science*. 2007; 315:1274–1278. [PubMed: 17332413]
- Theofilopoulos AN, Baccala R, Beutler B, Kono DH. Type I interferons (alpha/beta) in immunity and autoimmunity. *Annu Rev Immunol*. 2005; 23:307–336. [PubMed: 15771573]
- Thomas C, Moraga I, Levin D, Krutzik PO, Podoplelova Y, Trejo A, Lee C, Yarden G, Vleck SE, Glenn JS, et al. Structural linkage between ligand discrimination and receptor activation by type I interferons. *Cell*. 2011; 146:621–632. [PubMed: 21854986]
- Uddin S, Grumbach IM, Yi T, Colamonici OR, Plataniias LC. Interferon alpha activates the tyrosine kinase Lyn in haemopoietic cells. *Br J Haematol*. 1998; 101:446–449. [PubMed: 9633884]
- Uddin S, Plataniias LC. Mechanisms of type-I interferon signal transduction. *J Biochem Mol Biol*. 2004; 37:635–641. [PubMed: 15607020]
- van Boxel-Dezaire AH, Rani MR, Stark GR. Complex modulation of cell typespecific signaling in response to type I interferons. *Immunity*. 2006; 25:361–372. [PubMed: 16979568]
- Vogl C, Flatt T, Fuhrmann B, Hofmann E, Wallner B, Stiefvater R, Kovarik P, Strobl B, Muller M. Transcriptome analysis reveals a major impact of JAK protein tyrosine kinase 2 (Tyk2) on the expression of interferon-responsive and metabolic genes. *BMC Genomics*. 2010; 11:199. [PubMed: 20338026]
- Wang K, Narayanan M, Zhong H, Tompa M, Schadt EE, Zhu J. Meta-analysis of inter-species liver co-expression networks elucidates traits associated with common human diseases. *PLoS Comput Biol*. 2009; 5:e1000616. [PubMed: 20019805]

Ye CJ, Feng T, Kwon HK, Raj T, Wilson MT, Asinovski N, McCabe C, Lee MH, Frohlich I, Paik HI, et al. Intersection of population variation and autoimmunity genetics in human T cell activation. *Science*. 2014; 345:1254665. [PubMed: 25214635]

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

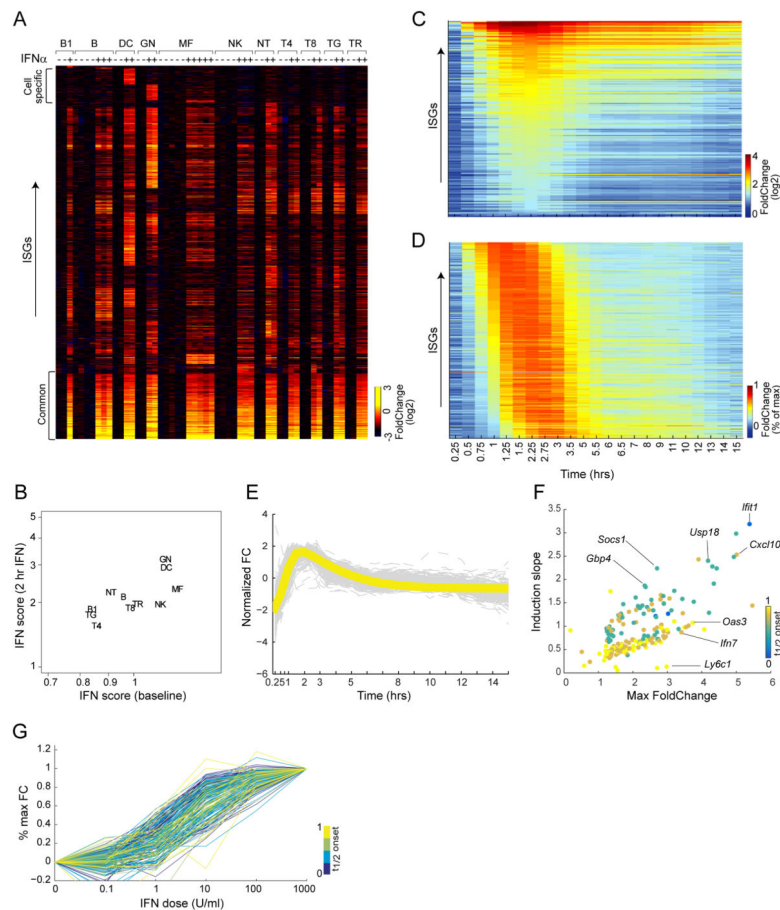


Fig. 1. Breadth and kinetics of the transcriptional response to IFN

Mice were challenged *in vivo* with IFN α , and cells sorted for gene expression profiling. **(A)** Expression changes for 975 ISGs that are significantly induced in at least one of 11 cell-types (ubiquitously-induced core set at bottom). **(B)** Mean IFN signature score for each cell-types before or after IFN challenge. **(C)** Temporal foldchange (vs untreated) for 182 ISGs induced in B cells *in vivo*. **(D)** As C, scaled to maximum induction for each gene. **(E)** Same data, yellow line is the impulse function fit to average standardized temporal profile (each ISG's standardized z-score in grey). **(F)** Maximum FC vs Slope of Onset, each point is a gene, color-coded according to its $t_{1/2\text{onset}}$. **(G)**. Dose-response in splenic B cells IFN-stimulated *in vitro*, as a proportion of the maximal response. See also Fig S1 and Tab. S1.

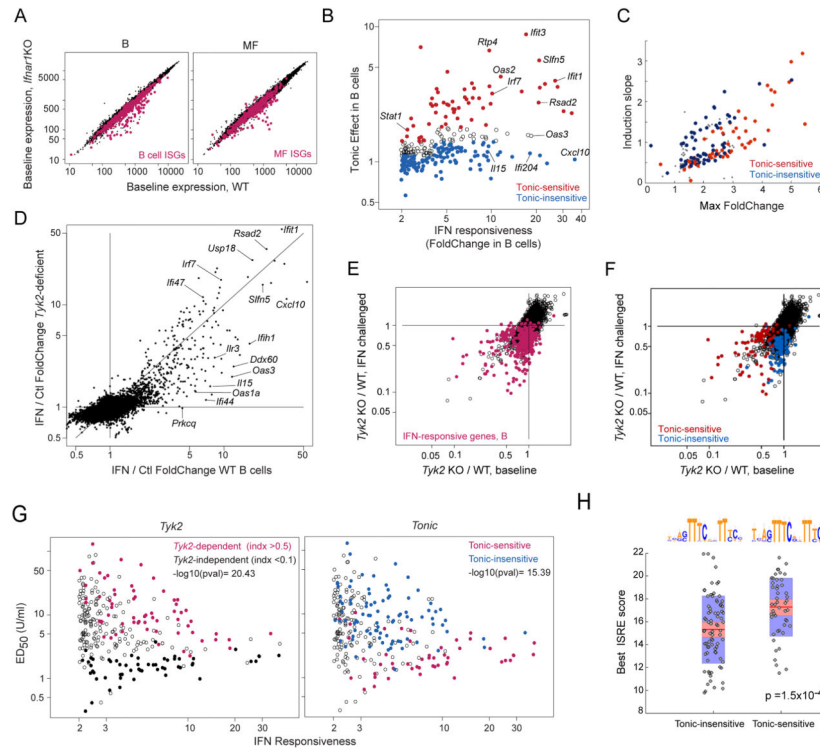


Fig. 2. Signal requirements for tonic and induced ISG expression in B cells

(A). Comparison of B cell and macrophage profiles of WT or *Ifnar1*-deficient mice. Red: common ISGs per Fig. 1. (B) ISG response intensity (change 2 hrs after IFN) and dependence on tonic IFN signaling at baseline (Foldchange between untreated B cells, WT and *Ifnar1*-KO). (C) Positioning of “Tonic-hi” and “Tonic-lo” ISGs (per B) relative to response kinetics of Fig. 1F. (D) Foldchange/Foldchange plot comparing ISG induction in B cells from WT or *Tyk2*-deficient mice. (E) Expression ratios (*Tyk2*-deficient / WT) in resting or in IFN B cells. B cell ISGs in red. (F) As E, highlighting tonic-sensitive and tonic-insensitive (per B). (G) ISG response intensity vs ED₅₀ (Fig. 1F), highlighting TYK2-dependency (left) or sensitivity to tonic signaling (right). (H) Max ISRE motif scores (JASPAR) in TSS and 5' upstream regions of ISGs distinguished by sensitivity to tonic signaling; consensus sequence for each set, ignoring outliers, shown at top (Wilcoxon rank sum test p.vals). See Fig S2 and Tab. S1E.

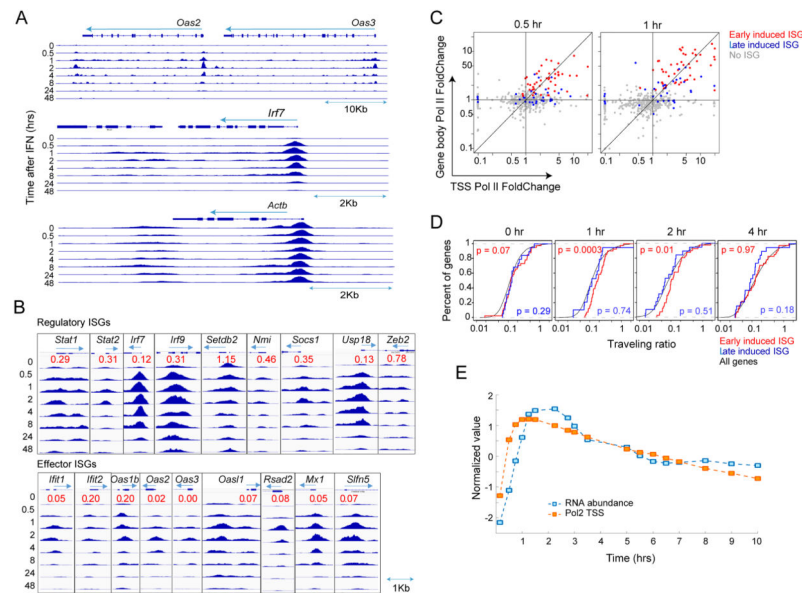


Fig. 3. Recruitment of RNA Pol-II on ISGs during the IFN response in B cells
 ChIPseq of bound Pol-II in B cells from IFN α treated mice (0–48hrs). **(A)** Pol-II accumulation at ISG and control loci (same scale). **(B)** Gallery of normalized Pol-II on TSS of ISGs with primarily regulatory or effector functions. “Pre-loading” ratio (red) for each gene is the ratio of Pol-II accumulation (integrated read counts through entire gene) at baseline vs maximum after IFN α . **(C)** Relation between fold increase in Pol-II (ratio over baseline levels) at TSS and within gene body at two times. Early- ($t_{1/2\text{onset}} < 45\text{min}$) and late- ($45\text{min} < t_{1/2\text{onset}} < 2\text{hr}$) induced ISGs are shown, $r=0.32$, $p=5 \times 10^{-3}$ (0.5hr); $r=0.44$, $p=7 \times 10^{-5}$ (1hr) **(D)** Traveling ratio (intragenic/TSS-associated Pol-II) cumulative density at different times; Mann-Whitney U test p.vals for early or late-induced ISGs vs genome-wide **(E)** Computational fit to impulse function of ISG RNA abundance and Pol-II at TSS. See Fig S3.

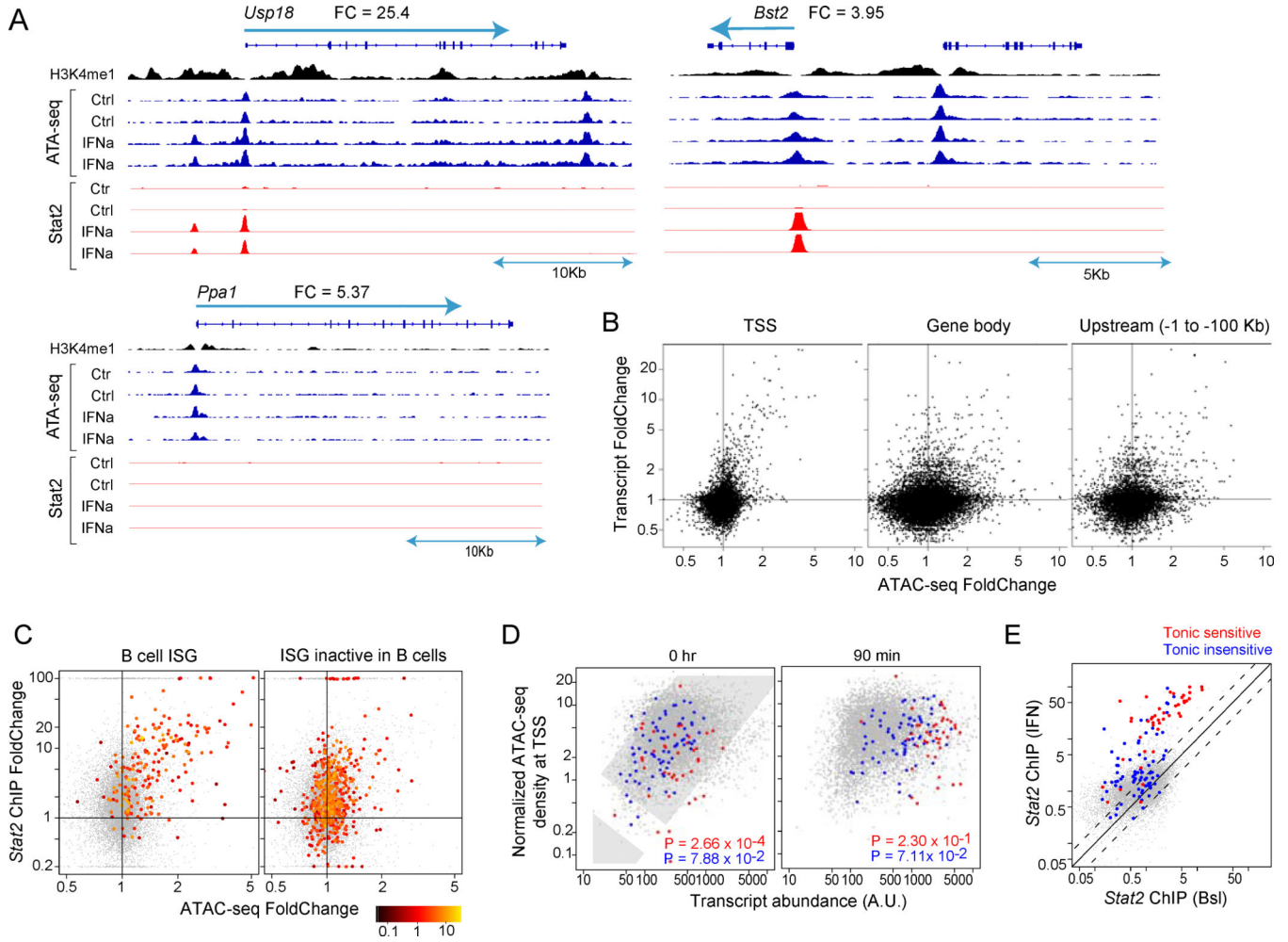


Fig. 4. Chromatin modifications during the IFN response in B cells
 ATACseq of accessible chromatin regions in B cells (baseline or 90 mn) after IFN α *in vivo*. **(A)** ATAC-seq and Stat2 ChIPseq read tracks (fixed scale) with H3K4me1 ChIPseq. **(B)** Change in ATACseq peaks vs changes in transcript level for peaks located in different gene regions. $r=0.39$, $p=2\times 10^{-16}$ (TSS), $r=0.17$, $p=10^{-3}$ (Gene body); $r=0.20$, $p=9\times 10^{-6}$ (Upstream) **(C)**. Relation between IFN-induced changes in ATAC peak intensity and Stat2 binding for ISGs responsive in B cells (left) or only active in non-B cells (right), color-coded according to its ATAC-seq signal (IFN). $r=0.51$, $p=10^{-14}$ (B cell ISG); $r=0.21$, $p=4\times 10^{-6}$ (ISG inactive in B cells) **(D)** Relation between TSS ATAC signal and expression (tonic-sensitive, red, and tonic-insensitive transcripts, blue) (per Fig. 2B). $r=0.25$, $p=2\times 10^{-16}$ (0hr) and $r=0.23$, $p=2\times 10^{-16}$ (90min) **(E)** Stat2 ChIPseq signal intensity, according to sensitivity to tonic IFN. See Fig S4 and Tab. S2C.

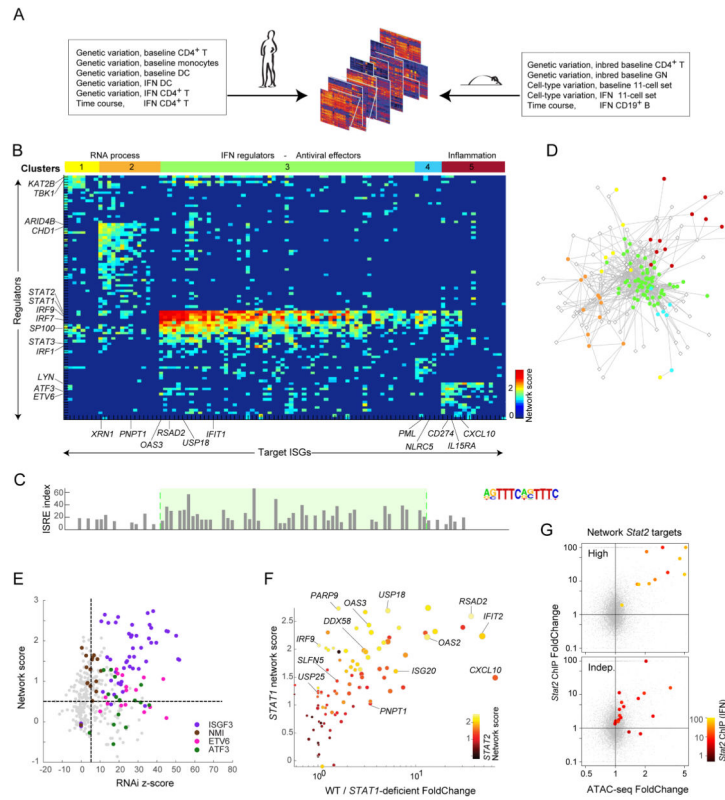


Fig. 5. Trans-species regulatory network inference

(A) Datasets used for network inference. (B) Heatmap of inferred regulatory scores (sparsified at FDR<0.01) connecting regulators in rows (kinases, phosphatases, TFs) and targets (columns), color-coded to strength of predicted association. Regulators and targets biclustered into 5 clusters, color-coded at top. (C) FIMO scores in TSS of target ISGs for ISRE. (D) Network-based representation of regulator-target links, color-coded by cluster membership (per B); only showing links passing Bonferroni $p < 10^{-6}$. (E) Each testable regulator-target link shown as a dot, placed by its z-score from regulator RNAi knockdown ((Amit et al., 2009)) and network score. Predicted links for ISGF3, NMI, ETV6, and ATF3 in color. (F) Correspondence between predicted STAT1 regulator score for each ISG and expression Foldchange in fibroblasts from a STAT1-deficient patient. Predicted STAT2 regulatory strength for these ISGs color-coded. Dot size shows magnitude of response in wild-type condition. See Fig S5 and Tab.s S3B-G. (G) Relation between IFN-induced changes in ATAC peak intensity and Stat2 binding for ISGs with top or bottom 10% Stat1/2 network scores (color-coded for Stat2 signal intensity).

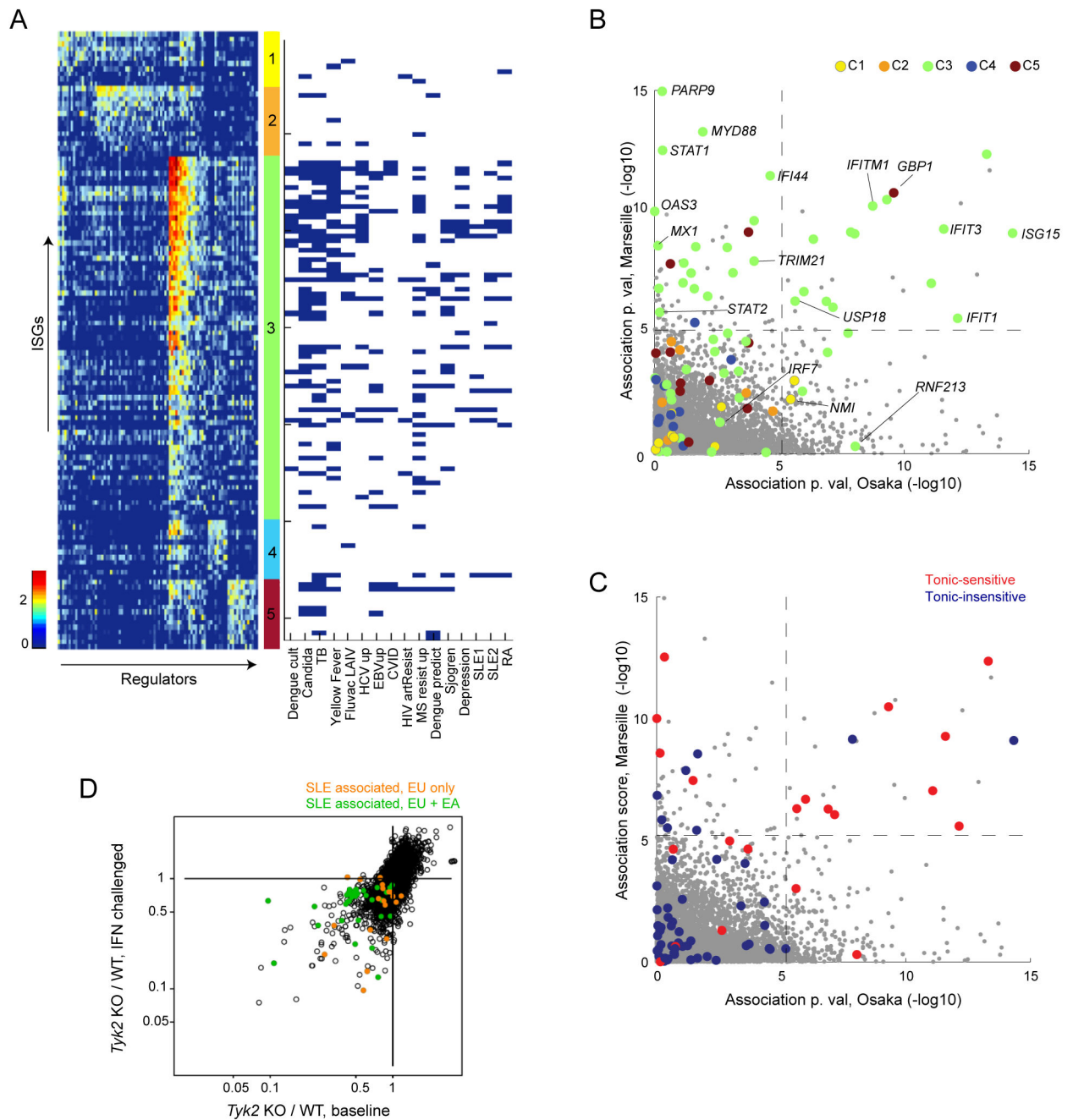


Fig. 6. Projection of the IFN response network onto human disease signatures

(A) Predicted IFN regulatory network, per Fig. 5 (left), arranged into 5 biclusters (middle) and their presence in IFN signatures reported in 16 human diseases/conditions (in columns, right). (B) Association between expression and SLE diagnosis in case/control Europeans or East Asian cohorts. Cluster membership color-coded per 5B. (C) As B, color-coded for tonic IFN sensitivity per Fig. 2B. (D). Tyk2 dependence for ISG expression at baseline or after IFN, color-coded for association of expression with SLE in both (green) or only European (orange) SLE patients. See Fig S6 and Tab.s S4A and S4B.

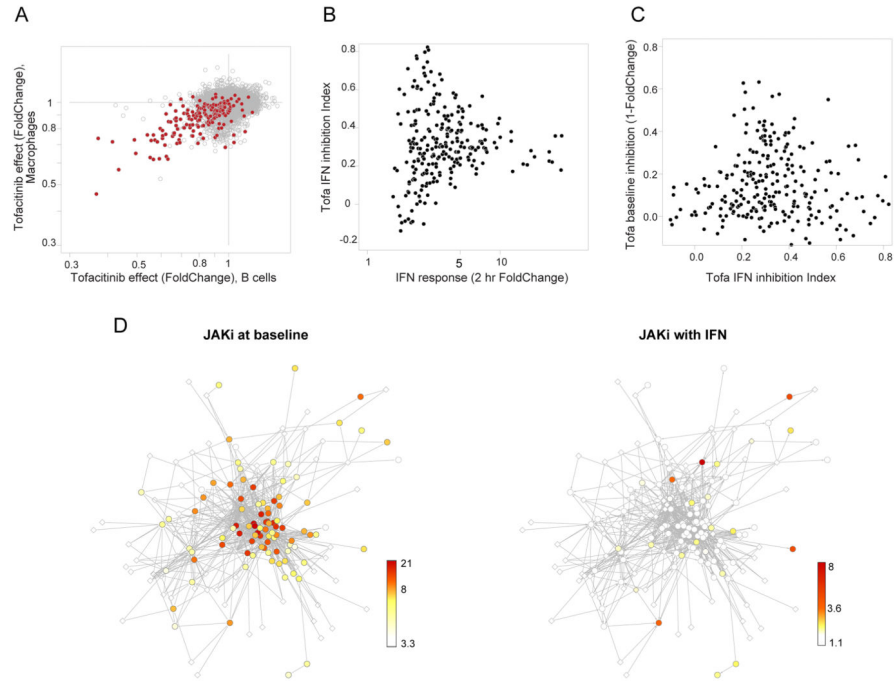


Fig.7. Pharmacological inhibition of different facets of the ISG network
 Mice gavaged with JAKi inhibitor Tofacitinib, and immunocytes profiled directly or 2 hrs after IFN α treatment. **(A)** Inhibition of baseline gene expression in macrophages or B cells. Common ISGs in red. **(B)** Profiling of JAK inhibition with 2hr IFN challenge, relating intensity of IFN response in untreated mice and JAK1-inhibition index (0 = no inhibition, 1 = full inhibition of response). **(C)** Negative relation between inhibition of acute IFN response by JAK1i (inhibition index from B) and inhibition at baseline (1-foldchange of panel A). **(D)** Effect of JAK1 inhibition (color-coded by pvals for observed foldchanges) in two conditions plotted on framework of regulatory network of Fig. 5. See Fig. S7 and Tab. S5A.