Transcriptional Profiling of STAT1 Gain-of-Function Reveals Common and Mutation-Specific Fingerprints

Simone Giovannozzi [1,2], Jonas Demeulemeester [3,4], Rik Schrijvers [1†] and Rik Gijsbers [2,5†]*

[1] Allergy and Clinical Immunology Research Group, Department of Microbiology, Immunology and Transplantation, KU Leuven, Leuven, Belgium, [2] Laboratory for Viral Vector Technology and Gene Therapy, Department of Pharmaceutical and Pharmacological Sciences, KU Leuven, Leuven, Belgium, [3] Laboratories for Computational Biology and Reproductive Genomics, Department of Human Genetics, KU Leuven, Leuven, Belgium, [4] Cancer Genomics Laboratory, The Francis Crick Institute, London, United Kingdom, [5] Leuven Viral Vector Core, KU Leuven, Leuven, Belgium

STAT1 gain-of-function (GOF) is a primary immunodeficiency typically characterized by chronic mucocutaneous candidiasis (CMC), recurrent respiratory infections, and autoimmunity. Less commonly, also immunodysregulation polyendocrinopathy enteropathy X-linked (IPEX)-like syndromes with CMC, and combined immunodeficiency without CMC have been described. Recently, our group and others have shown that different mutation-specific mechanisms underlie STAT1 GOF *in vitro*, including faster nuclear accumulation (R274W), and reduced mobility (R321, N574I) to near immobility in the nucleus (T419R) upon IFNγ stimulation. In this work, we evaluated the transcriptomic fingerprint of the aforementioned STAT1 GOF mutants (R274W, R321S, T419R, and N574I) relative to STAT1 wild-type upon IFNγ stimulation in an otherwise isogenic cell model. The majority of genes up-regulated in wild-type STAT1 cells were significantly more up-regulated in cells expressing GOF mutants, except for T419R. In addition to the common interferon regulated genes (IRG), STAT1 GOF mutants up-regulated an additional set of genes, that were in part shared with other GOF mutants or mutation-specific. Overall, R274W and R321S transcriptomes clustered with STAT1 WT, while T419R and N574I had a more distinct fingerprint. We observed reduced frequency of canonical IFNγ activation site (GAS) sequences in promoters of genes up-regulated by all the STAT1 GOF mutants, suggesting loss of DNA binding specificity for the canonical GAS consensus. Interestingly, the T419R mutation, expected to directly increase the affinity for DNA, showed the most pronounced effects on the transcriptome. T419R STAT1 dysregulated more non-IRG than the other GOF mutants and fewer GAS or degenerate GAS promotor sequences could be found in the promoter regions of these genes. In conclusion, our work confirms hyperactivation of common sets of IFNγ-induced genes in STAT1 GOF with additional dysregulation of mutation-specific genes, in line with the earlier observed mutation-specific mechanisms. Binding to more degenerate GAS sequences is proposed as a mechanism toward transcriptional dysregulation in R274W, R321S, and N574I. For T419R, an increased interaction with the DNA is suggested to result in a broader and less GAS-specific response. Our work indicates that multiple routes leading to STAT1 GOF are associated with common and private transcriptomic fingerprints, which may contribute to the phenotypic variation observed *in vivo*.

Keywords: STAT1, gain of function, mRNA sequencing, transcriptomic analysis, hypermorphic mutation, interferon gamma, gas

INTRODUCTION

Signal Transducer and Activator of Transcription 1 (STAT1) gain-of-function (GOF) is a monogenic autosomal dominant (AD) disorder first described in 2011 (1, 2). To date 105 different mutations in at least 400 patients (3) have been reported, all presenting with increased STAT1-dependent cellular responses and high levels of tyrosine phosphorylated STAT1 (pSTAT1) in different immune cells. While the common phenotypic denominator is chronic mucocutaneous candidiasis [CMC—in 98% of the patients (4)], less frequent phenotypes were also associated with STAT1 GOF, like John Cunningham (JC)-virus induced progressive multifocal leukoencephalopathy (5), immunodysregulation polyendocrinopathy enteropathy X-linked (IPEX)-like syndromes with CMC (6, 7), Orf infection (8), and combined immunodeficiency (CID) without CMC (9). Currently, treatment is symptomatic, and consists of chronic antifungal therapy for CMC and episodic antibiotics as well as immunosuppressive treatment in case of auto-immune manifestations. In addition, hematopoietic stem cell transplantation and use of jakinibs has been reported in a minority of patients with varying success (10).

STAT1 is a transcription factor of the STAT family, that plays a key role in the immune response and the Interferon (IFN) signaling pathway, modulating diverse cellular processes including proliferation, cell death and cell differentiation (11–13). In unstimulated conditions STAT1 is located in the cytoplasm as an antiparallel homodimer (14, 15). Following IFNγ stimulation, STAT1 is phosphorylated at Y701 (11) and the antiparallel homodimers change their conformation to parallel homodimers (16, 17), that in turn are imported into the nucleus to bind gamma interferon (IFNγ) activation site (GAS) sequences in promoter regions and, up- or down-regulate the transcription of interferon regulated genes (IRGs) (18–20).

Most STAT1 GOF mutations [48 out of 74 amino acid (aa) positions] are located at the antiparallel homodimer interface (**Figures 1A,B**). A small subset of mutations are located in close proximity to the parallel homodimer-DNA interface (M325, H328, Q330, S466, and T419), while 18 mutated aa were described in positions away from the antiparallel homodimer and parallel homodimer-DNA interfaces.

Recently, we and others (21–24) described that specific STAT1 GOF mutations are associated with different molecular mechanisms. We studied a subset of STAT1 GOF mutants (R274W, R321S, T419R, and N574I) that reside in different domains and at different interfaces of the STAT1 protein (24). At least three different mechanisms could be identified: (I) faster nuclear import was observed for the R274W mutation (antiparallel homodimer interface); (II) reduced nuclear mobility and delayed dephosphorylation for the R321S and the N574I mutants (distant from the homodimer or DNA interfaces); (III) slower diffusion in the nucleus, possibly due to enhanced affinity for chromatin, was observed for the T419R mutation (parallel homodimer-DNA interface).

Exploring the crystal structures of parallel and antiparallel STAT1 dimers provides hints, at least for some mutations, on the underlying molecular mechanism. For example, R274 is located at the interface of the antiparallel homodimer, in a coiled-coil domain. We hypothesized that mutating the positively charged aa R274 may destabilize the antiparallel, inactive, homodimer, resulting in a faster shift to the active parallel form (24). Alternatively, aa T419 is in close contact with the DNA backbone when STAT1 is in its active parallel conformation. Mutating threonine to a positively charged arginine (T419R GOF mutation), may increase the affinity of STAT1 for the negatively charged DNA backbone.

Interestingly, aa N574 coordinates K511 of the same monomer in both the parallel and antiparallel conformation of the homodimer via a polar contact. In addition, a recent study demonstrated that STAT1 K511 is ubiquitinated, and that K511R mutation abolished ubiquitination and resulted in enhanced transcription of Type I interferon regulated genes (one Type II interferon regulated gene, *Ifit1*, was also analyzed, but its expression was not affected) (25). Due to the close proximity and direct interaction between N574 and K511, the N574I GOF mutation might affect STAT1 ubiquitination, which in turn could result in an up-regulation of IRG (24).

While the transcriptomic fingerprints of R274Q (21) and two other GOF mutants (R274G and H629Y) were studied individually (26), a comparison between the transcriptome of different STAT1 GOF mutations and their underlying molecular mechanisms is lacking. Therefore, we employed the STAT1$^{-/-}$ U3A cell model and complemented these cells with the respective STAT1 GOF mutations using stable integration via lentiviral vector technology. This isogenic setup limited variability and to allowed direct comparison of the transcriptome following IFNγ stimulation. As controls, we included non-transduced STAT1$^{-/-}$ U3A cells (NT) and cells complemented with STAT1 WT.

METHODS

Cell Line Generation and Culture

U3A STAT1$^{-/-}$ cells (Sigma, REF 12021503-1VL) were stably complemented using lentiviral vectors carrying STAT1-GFP WT, R274W, R321S, T419R, or N574I as previously described (24). The mutants R274W, R321S, and N574I were initially chosen in the context of our research hypothesis, because they are present in different domains of the STAT1 protein and because patients bearing these mutations were in follow-up at our university hospital. The T419R mutant was additionally included as a clinical GOF mutation predicted to directly interact with the STAT1 target DNA (24). All cells were cultured at 37°C (humidified atmosphere, 5% CO_2) in DMEM (GIBCO-BRL) supplemented with 10% FBS (GIBCO, REF 10270-106), 0.01% v/v gentamicin (GIBCO-BRL). Cells were tested to be mycoplasma-free by PlasmoTest™, InvivoGen Europe.

IFNγ Stimulation and mRNA Extraction

Five replicates of each cell line were independently seeded (500.000 cells/well). After 24 h the cells were either stimulated, by substituting the medium with DMEM 10% FBS, supplemented with IFNγ 1U/μl (Roche, REF 11040596001) for 4 h, or left unstimulated. Total mRNA was isolated using Aurum™ total RNA mini kit (Biorad, REF 7326820). The 4 h time-point after

FIGURE 1 | STAT1 crystal structure and known GOF mutations (3). **(A)** Schematic representation of antiparallel STAT1 homodimer (PDB reference 1yvl). **(B)** Parallel STAT1 homodimer bound to DNA (PDB reference 1bf5). For clarity, one monomer is represented in surface mode, the other as ribbons. In red amino acids for which GOF mutations have been shown. Dark gray arrows indicate the orientation of each monomer in the respective dimer configurations.

IFNγ stimulation, was chosen because, (i) we showed that 4 h IFNγ stimulation is sufficient to upregulate ISGs (24) and (ii) to limit transcriptome effects that may result from activation of secondary pathways and later transcriptome effects.

mRNA Library Preparation and Sequencing

RNA samples were processed by the Genomics Core Leuven (Belgium). Sequence libraries were prepared with the Lexogen QuantSeq 3′ mRNA-Seq Library prep kit according to the manufacturer's protocol. Samples were indexed to allow for multiplexing. Library quality and size range was assessed using a Bioanalyzer (Agilent Technologies) with the DNA 1000 kit (Agilent Technologies, California, USA) according to the manufacturer's recommendations. Libraries were diluted to a final concentration of 2 nM and subsequently sequenced on an Illumina HiSeq4000 platform. Single-end reads of 50 bp length were produced with a minimum of 1.5 M reads per sample. Quality control of raw reads was performed with FastQC (27). FastQC: a quality control tool for high throughput sequence data. (Available online at: http://www.bioinformatics. babraham.ac.uk/projects/fastqc) Adapters were filtered with ea-utils fastq-mcf (28) (https://github.com/ExpressionAnalysis/ea-utils). Splice-aware alignment was performed with STAR v2 (29) against the human reference genome hg37 using the default parameters. Reads mapping to multiple loci in the reference genome were discarded. Resulting BAM alignment files were handled with Samtools v1.5 (30). Quantification of reads per gene was performed with HT-Seq Count v2.7.14.

Differential Expression Analysis

Differential expression analysis was performed with R-based (The R Foundation for Statistical Computing, Vienna, Austria) Bioconductor package DESeq2 (31). Reported p-values were adjusted for multiple testing with the Benjamini-Hochberg procedure, which controls false discovery rate (FDR). Genes that showed at least a 2-fold difference between the stimulated and the unstimulated condition were included in subsequent analyses. Principal component analysis (PCA) was performed using the pcaExplorer package (32).
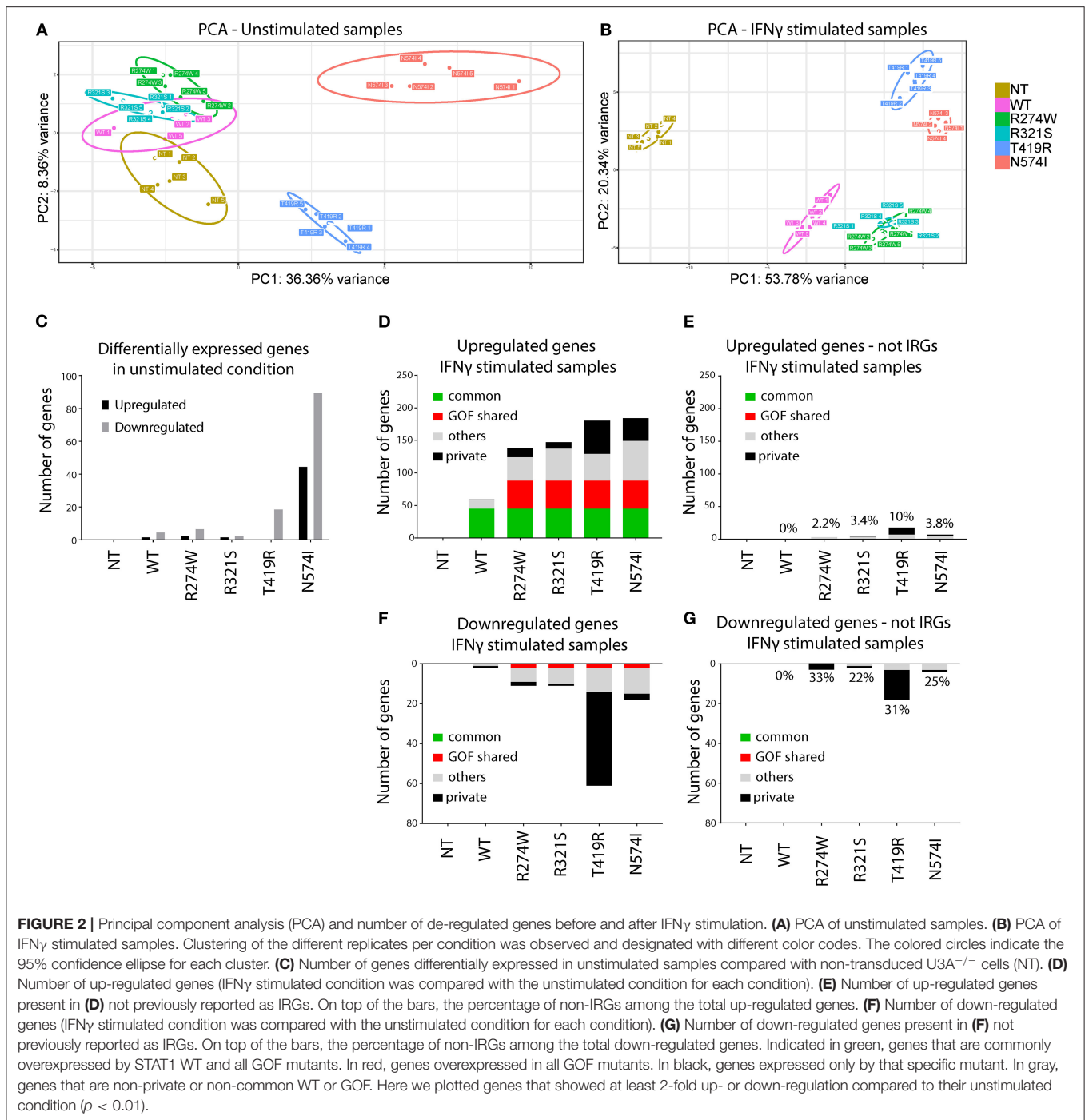
IRGs and Promoter Analysis

Interferon regulated genes (IRGs) are genes up- or down-regulated by the interferon signaling. The analysis of IRGs was performed using the Interferome database (http://www. interferome.org), a manually curated database of type I, II and III interferon-regulated genes. In our analysis we focused on type II IRGs. Promoter regions were obtained from the Eukaryotic Promoter Database (https://epd.epfl.ch/). Occurrence profile analysis and Position Weight Matrix scanning were performed on the Signal Search Analysis Server (https://ccg.epfl.ch/ssa/) (33). HOCOMOCO STAT1 consensus was used as definition of canonical GAS sequence (34). ChIP-Seq data were retrieved from the GEO database (GSE15353) (35) and correlation with promoters of genes of interest was evaluated using the ChIP-Seq online analysis tool from the Swiss Institute of bioinformatics (https://ccg.epfl.ch/chipseq/) (36).

Pathways Analysis

Pathway enrichment analysis was performed using Ingenuity Pathways Analysis (IPA, http://www.ingenuity.com/products/ ipa), using a cut-off of ±1 $\log_2$-fold change compared to the unstimulated condition and $p < 0.001$ for gene selection. Pathway activation is scored using gene expression z-scores. Briefly, a Z-score is defined as the difference between the error-weighted mean of the expression values of the genes in each pathway and the error-weighted mean of all genes in a sample after normalization. Z-scores were computed and plotted in a matrix of pathway activation scores.

Statistics

Two-way ANOVA with Dunnett's multiple comparison test was used to test the data in **Figures 2**, **3**. Friedman test with Dunn's multiple comparison was used to compare the occurrence profile of the different mutants in **Figures 7A–C**. Pearson's Chi-squared test with Yates' continuity corrections was used to compare the proportions of ChIP-seq peaks in **Figures 7D,G**. $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$, $^{****}p < 0.0001$ for all the figures.

FIGURE 2 | Principal component analysis (PCA) and number of de-regulated genes before and after IFNγ stimulation. **(A)** PCA of unstimulated samples. **(B)** PCA of IFNγ stimulated samples. Clustering of the different replicates per condition was observed and designated with different color codes. The colored circles indicate the 95% confidence ellipse for each cluster. **(C)** Number of genes differentially expressed in unstimulated samples compared with non-transduced U3A$^{-/-}$ cells (NT). **(D)** Number of up-regulated genes (IFNγ stimulated condition was compared with the unstimulated condition for each condition). **(E)** Number of up-regulated genes present in **(D)** not previously reported as IRGs. On top of the bars, the percentage of non-IRGs among the total up-regulated genes. **(F)** Number of down-regulated genes (IFNγ stimulated condition was compared with the unstimulated condition for each condition). **(G)** Number of down-regulated genes present in **(F)** not previously reported as IRGs. On top of the bars, the percentage of non-IRGs among the total down-regulated genes. Indicated in green, genes that are commonly overexpressed by STAT1 WT and all GOF mutants. In red, genes overexpressed in all GOF mutants. In black, genes expressed only by that specific mutant. In gray, genes that are non-private or non-common WT or GOF. Here we plotted genes that showed at least 2-fold up- or down-regulation compared to their unstimulated condition ($p < 0.01$).
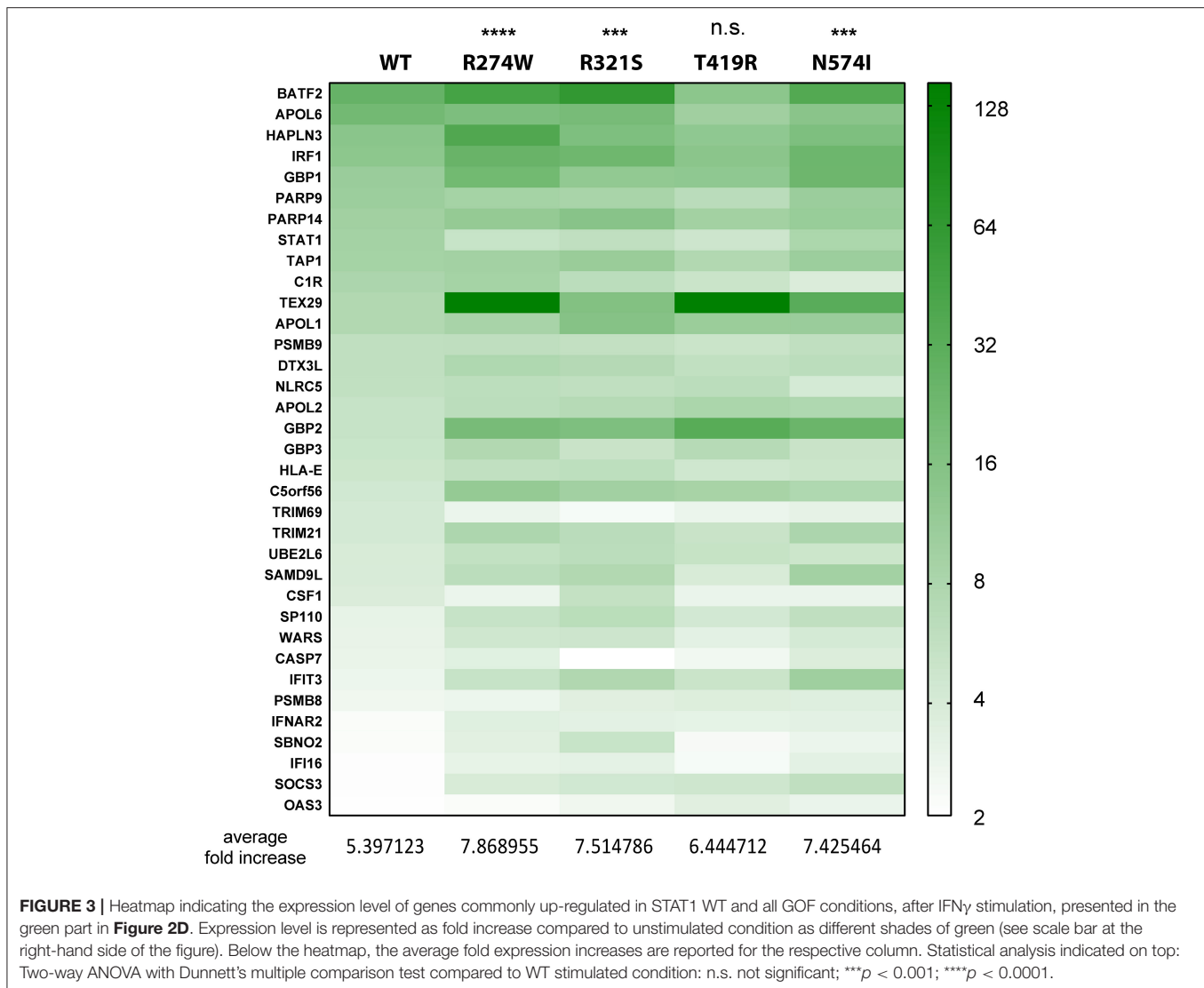
RESULTS

Experimental Setup

To study the transcriptomic fingerprints of the different STAT1 GOF mutations, we complemented a STAT1$^{-/-}$ U3A cell line using lentiviral vectors to stably express STAT1-GFP WT or one of the respective STAT1 GOF mutations, R274W, R321S, T419R, or N574I, thereby generating isogenic add-back cell lines, referred to as STAT1 WT, R274W, R321S, T419R, and N574I. Non-transduced cells (hereafter termed NT) were taken along

as control. To ensure comparable STAT1 expression levels, cells were generated using the highest dilutions of a limiting dilution series of lentiviral vectors and subsequent antibiotic selection of transduced cells, ensuring a polyclonal cell population with a single LV integration. Using a stable cell line, instead of transient transfection, allowed us to eliminate variability derived from different transfection efficiencies between experiments. Detailed characterization of these cell lines, comprising quantification of the respective proteins and nuclear translocation upon IFNγ stimulation by fluorescent microscopy and increased expression

FIGURE 3 | Heatmap indicating the expression level of genes commonly up-regulated in STAT1 WT and all GOF conditions, after IFNγ stimulation, presented in the green part in **Figure 2D**. Expression level is represented as fold increase compared to unstimulated condition as different shades of green (see scale bar at the right-hand side of the figure). Below the heatmap, the average fold expression increases are reported for the respective column. Statistical analysis indicated on top: Two-way ANOVA with Dunnett's multiple comparison test compared to WT stimulated condition: n.s. not significant; ***$p < 0.001$; ****$p < 0.0001$.

of three IFNγ regulated genes (IRGs: *IRF1*, *GBP1,* and *CXCL10*) was previously reported (24).

In an effort to study the transcriptomic fingerprint of STAT1 WT, the different GOF mutations, and non-transduced U3A STAT1$^{-/-}$, each cell line was stimulated for 4 h with IFNγ or left unstimulated ($n = 5$ independent experiments per cell line). Total mRNA was extracted and sequenced on an Illumina HiSeq4000 (Methods). For each condition, we selected the genes showing at least a 2-fold difference between the unstimulated and the stimulated condition ($p < 0.01$, Benjamini-Hochberg procedure).

PCA Reveals Distinct Clustering for the Different GOF Mutants

In a first step, RNAseq data were subjected to Principal Component Analysis (PCA), either for the unstimulated conditions or after IFNγ stimulation. This analysis reduces the dimensionality of the samples and clusters mRNA sequencing results based on their similarities over the main principal components. Samples with a similar transcriptomic profile will cluster closer together. PCA evaluates the whole transcriptomic fingerprint of the different conditions to define principal components (PC1 and PC2), and plots the respective replicates of the RNAseq experiment for each specific condition accordingly. PCA demonstrated clear clustering of the different replicates per condition and allowed to identify one of the five replicates of STAT1 WT in the unstimulated condition and one of N574I in the stimulated condition as outliers (**Supplementary Figure 1**). These were removed from subsequent analyses.

PCA of the remaining replicates under non-stimulated and stimulated condition is shown in **Figures 2A,B**, respectively. In the unstimulated condition (**Figure 2A**), STAT1 WT clusters closely to R274W and R321S, and to a lesser extend to NT, differing only for PC2 (8.55% of the variance). This suggested that addition of neither STAT1 WT nor these STAT1 GOF mutations alone significantly affected the transcriptome. On the other hand, N574I and T419R formed two distinct clusters, indicating a larger

effect upon complementation with these two mutants even in the absence of IFNγ stimulation.

Upon IFNγ stimulation (**Figure 2B**), both STAT1 WT and all GOF mutants clustered away from the NT condition. Again, R274W and R321S mutant clusters overlapped and located close to the STAT1 WT cluster, while T419R and N574I formed two clusters separated from STAT1 WT and the other GOF mutants.

T419R and N574I Have Already an Impact in Unstimulated Conditions

PCA showed that, even in unstimulated conditions, STAT1 T419R and N574I addition to U3A$^{-/-}$ cells resulted in a shift relative to STAT1 WT. Therefore, we first determined the number of genes significantly ($p < 0.01$) up- or down-regulated ($\geq$2-fold) relative to STAT1 WT, in unstimulated conditions (**Figure 2C**). As expected from the PCA analysis, R274W and R321S did not show major transcriptomic differences compared to STAT1 WT, with only one gene significantly up-regulated and five down-regulated for the R274W mutant and none differentially expressed for R321S mutant. On the contrary, both T419R and N574I mutants showed more transcriptomic differences under unstimulated conditions compared to STAT1 WT, with one up- and 21 genes down-regulated (on average 2.65-fold) for T419R and 44 genes (on average 2.96-fold) up- and 90 down-regulated (on average 2.70-fold) for N574I (**Figure 2C**).

Together this underscored that the stable introduction of STAT1 WT and the respective STAT1 GOF mutants, did not affect gene expression under basal conditions, although for T419R and N574I several genes were significantly down-regulated. Interestingly, none of the genes down-regulated by N574I and 3 out of 22 genes down-regulated by T419R in unstimulated conditions were up-regulated after IFNγ stimulation.

STAT1 GOF Mutations Show a Wider Range of Up-Regulated IRGs

In a next step, we studied the impact of IFNγ stimulation on gene expression by first determining the number of up- and down-regulated genes for each condition relative to NT (**Figures 2D–G**). The NT condition did not show any significantly up- or down-regulated genes when comparing the conditions with and without IFNγ stimulation, confirming that absence of STAT1 completely abolished responsiveness to IFNγ (**Figures 2D,F**). Looking at the up-regulated genes, STAT1 WT as well as GOF showed a substantial set of up-regulated genes (**Figure 2D**): STAT1 WT showed fewer up-regulated genes ($n = 59$) than the GOF mutants ($n = 138$ for R274W, $n = 147$ for R321S, $n = 180$ for T419R, $n = 184$ for N574I). Sorting these up-regulated genes we identified 45 to be up-regulated by STAT1 WT and all the GOF mutants (referred to as commonly up-regulated; green bar, **Figure 2D**), and 43 genes exclusively up-regulated by all GOF mutants, but not by STAT1 WT (GOF shared; red bar). Of interest, only a single gene (*SMG1*) was up-regulated in STAT1 WT, but not in any of the GOF mutants, while a larger number of "private" up-regulated genes was observed for the GOF mutants ($n = 1$ for WT, $n = 14$ for R274W, $n = 10$ for R321S, $n = 51$

for T419R, $n = 35$ for N574I; private; black bar, **Figure 2D**). A smaller set of genes was up-regulated in two or more, but not all, GOF conditions ($n = 13$ for WT, $n = 36$ for R274W, $n = 49$ for R321S, $n = 41$ for T419R, $n = 61$ for N574I; mixed; gray bar, **Figure 2D**).

A small number of genes was down-regulated (**Figure 2F**) for STAT1 WT ($n = 2$, on average 2.7-fold), R274W ($n = 11$, on average 2.8-fold), R321S ($n = 11$, on average 2.6-fold) and N574I ($n = 18$, on average 2.7-fold), while T419R showed a larger number of genes down-regulated after IFNγ stimulation ($n = 61$, on average 2.6-fold).

The fact that a large set of genes was up-regulated in the GOF cells compared to STAT1 WT, led us to hypothesize that STAT1 GOF mutants may have lost specificity for IRGs. Therefore, we assessed how many dysregulated genes did not belong to the IFNγ pathway (using the Interferome.org, a manually curated database including 9,768 type II interferon genes) (**Figures 2E,G**). This analysis indicated that for the R274W and R321S mutants, only a small number of up- and down-regulated genes were non-IRGs ($n = 3$ up-regulated and $n = 3$ down-regulated for R274W, and $n = 5$ up-regulated and $n = 2$ down-regulated for R321S). For N574I, a larger set of genes up- and down-regulated were non-IRG ($n = 7$ up-regulated and $n = 4$ down-regulated). STAT1 T419R showed the largest difference by, respectively, up- and down-regulating 18 and 18 non-IRGs, respectively. The large majority of genes affected were up-regulated IRGs (**Figures 2D,E**), whereas a smaller number of genes was down-regulated (**Figures 2F,G**). Altogether, these results suggested a reduced specificity of the response to the IFNγ stimulus, resulting in more up-regulated IRGs in the GOF cells compared to WT.

Overall Expression for Commonly Up-Regulated Genes Is Higher for STAT1 GOF

As a next step, we analyzed the relative expression levels of genes up-regulated by STAT1 WT and GOF mutants. First, we performed this analysis for the common up-regulated genes (green bar, **Figure 2D**) (**Figure 3**; expression heatmap). For the 35 common up-regulated genes, we calculated the average fold increase compared to the corresponding unstimulated lines: for the WT STAT1 condition an average 5.4-fold increase was obtained, whereas R274W, R321S, and N574I mutants resulted in a significantly higher fold increase (7.9, 7.5, and 7.4 average fold increase, respectively). T419R mutant did not result in a significant difference compared to STAT1 WT, with an average 6.4-fold increase in expression compared to unstimulated cells ($p = 0.11$, 2-way ANOVA compared to WT; **Figure 3**). These results highlighted that the genes up-regulated by STAT1 WT are also up-regulated by all STAT1 GOF mutants. Still, the up-regulation of these genes was significantly stronger for all GOF (overall 1.4-fold) compared to STAT1 WT, with the exception of T419R.

In addition, we identified genes that were up-regulated in all GOF conditions following IFNγ stimulation, but not in WT (red bar, **Figure 2D**). For these genes, we analyzed the average

expression fold increase for the genes (**Figure 4**, red heatmap), which resulted in a 4.5-fold average increase for all GOF mutants (no significant difference, 2-way ANOVA). This result pointed at a subset of genes commonly up-regulated by GOF mutants, but not in WT conditions.

Pathway Analysis Indicates Distinct Fingerprints of GOF Mutants

After showing differences in the number of dysregulated genes and the magnitude of up-regulation, we investigated how these genes affect physiological pathways. The IPA software (Qiagen) was used to predict which pathways are up- or down-regulated by the different STAT1 WT or GOF mutants. In the unstimulated condition, comparing each sample to the NT sample, no significant pathway was predicted to be up- or down-regulated (result not shown). In contrast, comparing each condition before and after IFNγ stimulation, several pathways were predicted to be up- or down-regulated in both STAT1 WT and GOF mutants (**Figure 5**). The different pathways were sorted according to their activation Z-score (matrix of pathway activation scores, **Figure 5**). As expected, the highest Z-score was obtained for the interferon signaling pathway for both STAT1 WT and the different GOF mutants, with R321S and N574I showing the highest score amongst the GOF mutants. No significant Z-scores were detected when comparing IFNγ stimulated WT to IFNγ stimulated GOF samples (not shown).

The R274W mutant showed limited differences in the number of up- or down-regulated pathways when compared to STAT1 WT. On the other hand, the T419R mutant showed the highest degree of difference from STAT1 WT profile (**Figure 5**). Importantly, T419R also showed down-regulation of different pathways that remained unaffected in both STAT1 WT and in the other GOF mutants. This result mirrored the observation that, upon IFN simulation, the T419R mutant down-regulates a higher number of genes compared to STAT1 WT than the other GOF mutants (**Figures 2D,F**). Of interest, IPA for R274W, T419R, and N574I showed that STAT3 signaling was down-regulated. In line with this, the *CISH* gene, a known inhibitor of STAT3 was strongly up-regulated upon IFNγ stimulation by all GOF mutants studied (>20-fold; **Figure 6A**). Together with the known link between STAT1 GOF and STAT3 pathway down-regulation (37), this could suggest that at least part of the commonly observed phenotype for GOF patients could be due to an inhibition of STAT3 signaling. In parallel, we observed a >2-fold up-regulation of *CD274* (encoding PD-L1) (**Figure 6B**) in all GOF samples, compared to their unstimulated condition (no IFNγ). When compared to WT, all STAT1 GOF mutants showed a significant *CD274* upregulation. Up-regulation of *CD274* was described to inhibit Th17 differentiation (38) and was shown to be present in patients affected by CMC after IL27 stimulation (39). The central role of Th17 cells in the pathogenesis of CMC, suggests involvement of *CD274* up-regulation, observed in the large majority of STAT1 GOF patients.
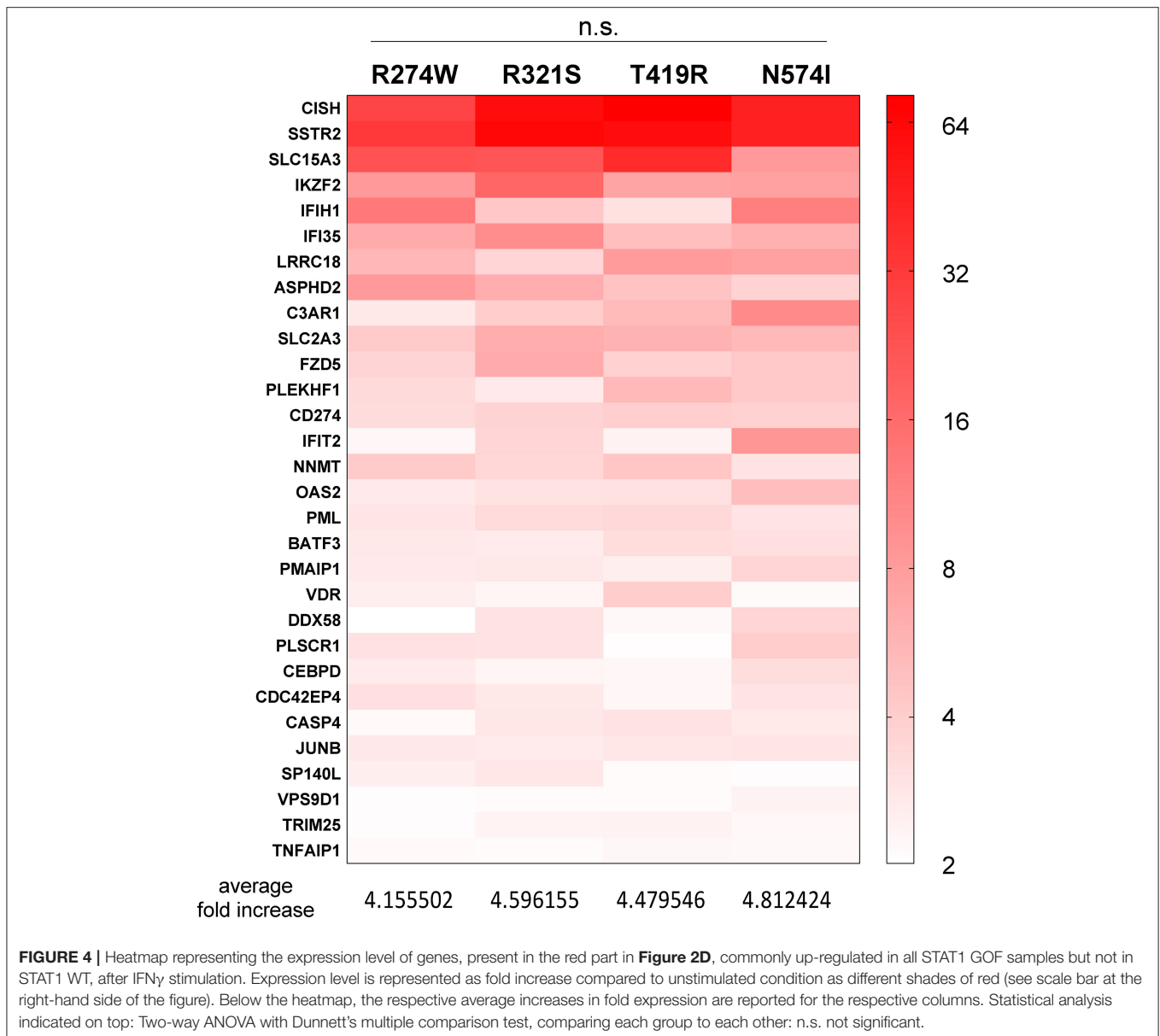
STAT1 GOF Mutants Show a Loss of Specificity for the Canonical STAT1 Binding Motifs

Together, the previous analyses indicate that GOFs hyperactivate the commonly up-regulated genes compared to STAT1 WT, with the exception of T419R (**Figure 3**). Additionally, the STAT1 GOF mutants up-regulated a set of additional genes, part of which are common for all GOF studied here (**Figure 2D**, red bar), but a substantial part is unique or shared with at least one, but not all other GOF mutants. We hypothesized that this may be explained by the binding of STAT1 GOF to more diverse sequences than the canonical STAT1 consensus motif in IRG promoters.

We therefore identified and analyzed the promoters of the genes up-regulated after IFNγ stimulation. A short promoter region [from −200 to +50 bp around the Transcription Start Site (TSS)] was retrieved for all up-regulated genes using the Eukaryotic Promoter Database (EPD) web tool (40). As a control, we used promoter sequences from 100 randomly picked genes.

We scanned the promoters of these genes using a STAT1 GAS consensus Position Weight Matrix, with different levels of stringency ($p < 0.001$, $p < 0.0001$, and $p < 0.00001$). The frequency of predicted GAS consensus sequences for each promoter region of the up-regulated genes was plotted for these respective stringencies (**Figures 7A–C**). For STAT1 WT there was a clear 20–40% peak of GAS sequence frequency between the TSS and −100bp (green line, **Figures 7A–C**). Performing the same analysis for the promoters of genes exclusively up-regulated by STAT1 GOF mutants following IFNγ stimulation, no clear peaks were observed using the highest stringency parameters ($p < 0.00001$, **Figure 7A**). In a next step, we therefore reduced the stringency of the GAS consensus definition ($p < 0.0001$ and $p < 0.001$), allowing for more degenerate GAS sequences to be included. This resulted in an increased frequency of predicted GAS in the promoter region for STAT1 WT (green line, **Figures 7B,C**), as well as the GOF mutants, albeit not to the level of STAT1 WT, whereas frequencies in the control promoter set remained low (black line, **Figure 7B**). At a confidence level of $p < 0.0001$ (**Figure 7B**), the R274W and the R321S curves were no longer significantly different from STAT1 WT (Friedman test with Dunn's multiple comparison *post-hoc* test). Lowering the stringency further to $p < 0.001$ (**Figure 7C**) also resulted in the N574I not being significantly different from STAT1 WT. In contrast, the random control set and the T419R mutant remained significantly different from STAT1 WT.

Finally, we calculated the raw number of promoters (−200, +50 bp from TSS) of up-regulated genes harboring at least one canonical GAS consensus sequence and the ones without GAS consensus (**Figures 7D,E**). Whereas, for STAT1 WT, the number of up-regulated genes with and without GAS consensus was comparable ($n = 25$ and $n = 24$, respectively), there were more GOF stimulated genes that had promoters lacking a consensus GAS sequence (**Figure 7D**). We observed a reduction of the ratio GAS present over GAS not present for all the GOF mutants, reaching statistical significance only for T419R, however (**Figure 7E**, $p = 0.025$, Pearson's Chi-squared test with Yates' continuity correction).

FIGURE 4 | Heatmap representing the expression level of genes, present in the red part in **Figure 2D**, commonly up-regulated in all STAT1 GOF samples but not in STAT1 WT, after IFNγ stimulation. Expression level is represented as fold increase compared to unstimulated condition as different shades of red (see scale bar at the right-hand side of the figure). Below the heatmap, the respective average increases in fold expression are reported for the respective columns. Statistical analysis indicated on top: Two-way ANOVA with Dunnett's multiple comparison test, comparing each group to each other: n.s. not significant.

Together, these results suggested that the promoters of STAT1 GOF up-regulated genes presented fewer canonical GAS motifs than WT STAT1 up-regulated genes, leading us to hypothesize that the STAT1 GOF mutants have reduced specificity for the canonical GAS consensus and activated a wider array of genes. In addition, in promoters of T419R up-regulated genes, GAS sequence frequencies remained very low, even when lowering the stringency of GAS consensus definition to $p < 0.001$ (**Figure 7C**, orange line). Interestingly, the T419R mutant was the only one not significantly different from the control group at both $p < 0.001$ and $p < 0.0001$ stringency levels and harbored fewer canonical GAS motifs in the promoters of its up-regulated genes.

STAT1 ChIP-Seq Peaks Suggests Altered GAS Consensus Recognition

The previous analysis showed that the STAT1 GOF mutants recognized the canonical GAS sequences (the genes up-regulated by STAT1 WT were also up-regulated by STAT1 GOF) as well as more degenerated GAS sequences. Interestingly, Robertson et al. (35) performed ChIP-Seq to identify mammalian DNA target sequences bound by endogenous STAT1 in unstimulated and IFNγ-stimulated HeLa-S3 cells (30 min). This dataset compiled endogenous STAT1 binding sites and contains >18.000 peaks. We used these ChIP-Seq data as a proxy for STAT1 GAS sequences, and determined the frequency of ChIP-Seq reads for all promoter sequences of the genes up-regulated in STAT1 WT
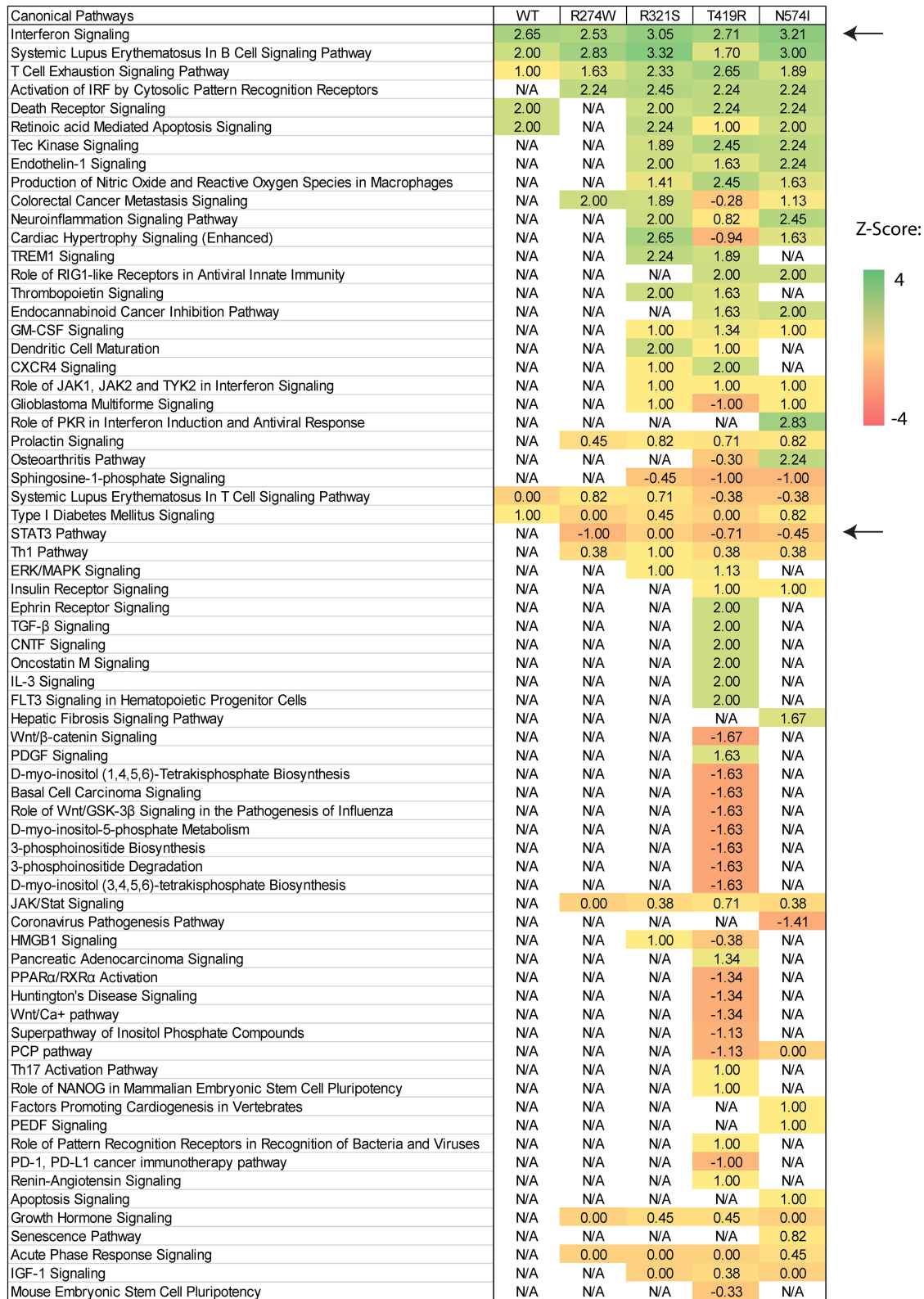
Canonical Pathways	WT	R274W	R321S	T419R	N574I
Interferon Signaling	2.65	2.53	3.05	2.71	3.21
Systemic Lupus Erythematosus In B Cell Signaling Pathway	2.00	2.83	3.32	1.70	3.00
T Cell Exhaustion Signaling Pathway	1.00	1.63	2.33	2.65	1.89
Activation of IRF by Cytosolic Pattern Recognition Receptors	N/A	2.24	2.45	2.24	2.24
Death Receptor Signaling	2.00	N/A	2.00	2.24	2.24
Retinoic acid Mediated Apoptosis Signaling	2.00	N/A	2.24	1.00	2.00
Tec Kinase Signaling	N/A	N/A	1.89	2.45	2.24
Endothelin-1 Signaling	N/A	N/A	2.00	1.63	2.24
Production of Nitric Oxide and Reactive Oxygen Species in Macrophages	N/A	N/A	1.41	2.45	1.63
Colorectal Cancer Metastasis Signaling	N/A	2.00	1.89	-0.28	1.13
Neuroinflammation Signaling Pathway	N/A	N/A	2.00	0.82	2.45
Cardiac Hypertrophy Signaling (Enhanced)	N/A	N/A	2.65	-0.94	1.63
TREM1 Signaling	N/A	N/A	2.24	1.89	N/A
Role of RIG1-like Receptors in Antiviral Innate Immunity	N/A	N/A	N/A	2.00	2.00
Thrombopoietin Signaling	N/A	N/A	2.00	1.63	N/A
Endocannabinoid Cancer Inhibition Pathway	N/A	N/A	N/A	1.63	2.00
GM-CSF Signaling	N/A	N/A	1.00	1.34	1.00
Dendritic Cell Maturation	N/A	N/A	2.00	1.00	N/A
CXCR4 Signaling	N/A	N/A	1.00	2.00	N/A
Role of JAK1, JAK2 and TYK2 in Interferon Signaling	N/A	N/A	1.00	1.00	1.00
Glioblastoma Multiforme Signaling	N/A	N/A	1.00	-1.00	1.00
Role of PKR in Interferon Induction and Antiviral Response	N/A	N/A	N/A	N/A	2.83
Prolactin Signaling	N/A	0.45	0.82	0.71	0.82
Osteoarthritis Pathway	N/A	N/A	N/A	-0.30	2.24
Sphingosine-1-phosphate Signaling	N/A	N/A	-0.45	-1.00	-1.00
Systemic Lupus Erythematosus In T Cell Signaling Pathway	0.00	0.82	0.71	-0.38	-0.38
Type I Diabetes Mellitus Signaling	1.00	0.00	0.45	0.00	0.82
STAT3 Pathway	N/A	-1.00	0.00	-0.71	-0.45
Th1 Pathway	N/A	0.38	1.00	0.38	0.38
ERK/MAPK Signaling	N/A	N/A	1.00	1.13	N/A
Insulin Receptor Signaling	N/A	N/A	N/A	1.00	1.00
Ephrin Receptor Signaling	N/A	N/A	N/A	2.00	N/A
TGF-β Signaling	N/A	N/A	N/A	2.00	N/A
CNTF Signaling	N/A	N/A	N/A	2.00	N/A
Oncostatin M Signaling	N/A	N/A	N/A	2.00	N/A
IL-3 Signaling	N/A	N/A	N/A	2.00	N/A
FLT3 Signaling in Hematopoietic Progenitor Cells	N/A	N/A	N/A	2.00	N/A
Hepatic Fibrosis Signaling Pathway	N/A	N/A	N/A	N/A	1.67
Wnt/β-catenin Signaling	N/A	N/A	N/A	-1.67	N/A
PDGF Signaling	N/A	N/A	N/A	1.63	N/A
D-myo-inositol (1,4,5,6)-Tetrakisphosphate Biosynthesis	N/A	N/A	N/A	-1.63	N/A
Basal Cell Carcinoma Signaling	N/A	N/A	N/A	-1.63	N/A
Role of Wnt/GSK-3β Signaling in the Pathogenesis of Influenza	N/A	N/A	N/A	-1.63	N/A
D-myo-inositol-5-phosphate Metabolism	N/A	N/A	N/A	-1.63	N/A
3-phosphoinositide Biosynthesis	N/A	N/A	N/A	-1.63	N/A
3-phosphoinositide Degradation	N/A	N/A	N/A	-1.63	N/A
D-myo-inositol (3,4,5,6)-tetrakisphosphate Biosynthesis	N/A	N/A	N/A	-1.63	N/A
JAK/Stat Signaling	N/A	0.00	0.38	0.71	0.38
Coronavirus Pathogenesis Pathway	N/A	N/A	N/A	N/A	-1.41
HMGB1 Signaling	N/A	N/A	1.00	-0.38	N/A
Pancreatic Adenocarcinoma Signaling	N/A	N/A	N/A	1.34	N/A
PPARα/RXRα Activation	N/A	N/A	N/A	-1.34	N/A
Huntington's Disease Signaling	N/A	N/A	N/A	-1.34	N/A
Wnt/Ca+ pathway	N/A	N/A	N/A	-1.34	N/A
Superpathway of Inositol Phosphate Compounds	N/A	N/A	N/A	-1.13	N/A
PCP pathway	N/A	N/A	N/A	-1.13	0.00
Th17 Activation Pathway	N/A	N/A	N/A	1.00	N/A
Role of NANOG in Mammalian Embryonic Stem Cell Pluripotency	N/A	N/A	N/A	1.00	N/A
Factors Promoting Cardiogenesis in Vertebrates	N/A	N/A	N/A	N/A	1.00
PEDF Signaling	N/A	N/A	N/A	N/A	1.00
Role of Pattern Recognition Receptors in Recognition of Bacteria and Viruses	N/A	N/A	N/A	1.00	N/A
PD-1, PD-L1 cancer immunotherapy pathway	N/A	N/A	N/A	-1.00	N/A
Renin-Angiotensin Signaling	N/A	N/A	N/A	1.00	N/A
Apoptosis Signaling	N/A	N/A	N/A	N/A	1.00
Growth Hormone Signaling	N/A	0.00	0.45	0.45	0.00
Senescence Pathway	N/A	N/A	N/A	N/A	0.82
Acute Phase Response Signaling	N/A	0.00	0.00	0.00	0.45
IGF-1 Signaling	N/A	N/A	0.00	0.38	0.00
Mouse Embryonic Stem Cell Pluripotency	N/A	N/A	N/A	-0.33	N/A

Z-Score:

4

-4

FIGURE 5 | Canonical pathway prediction via IPA on differentially regulated genes for each condition after IFNγ stimulation. Pathway activation is scored using gene expression Z-scores and plotted in an activation Z-score matrix. Z-score represents a measure of how much the specific pathway is predicted to be up- or
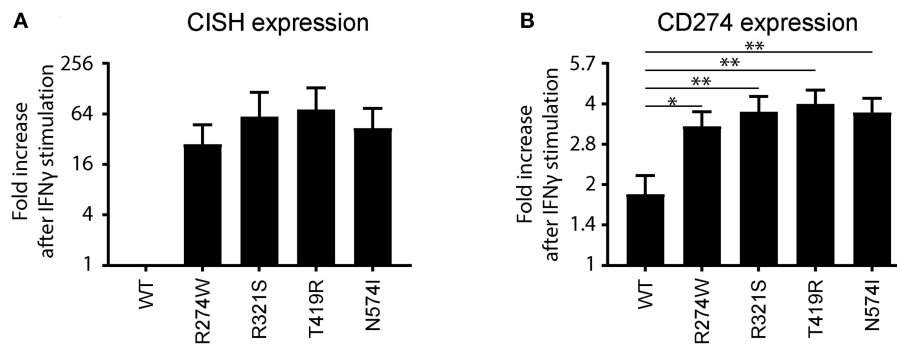
(Continued)

FIGURE 6 | (A) *CISH* and (B) *CD274* (PD-L1) mRNA expression fold increase upon IFNγ stimulation relative to the non-stimulated condition for STAT1 WT or the respective GOF mutants represented as means ± standard errors. The fold increase for CISH in the IFNγ stimulated WT sample was found not to be significantly different from the non-stimulated STAT1 WT, and was therefore set to 1. Statistical analysis in (B) compares fold increase after IFNγ stimulation for each GOF sample to WT: one-way ANOVA with Dunnett's multiple comparison test compared to WT stimulated condition: *$p < 0.05$; **$p < 0.01$.

and GOF cell lines (**Figure 7F**). For each STAT1 cell line, we determined which of the respective promoters harbored a ChIP-Seq peak and plotted the ratio between the number of genes harboring at least one ChIP-Seq peak in their promoter, over the number of genes not harboring any ChIP-Seq peak. For the GOF mutants, we analyzed only the genes not already up-regulated by STAT1 WT, to highlight the GOF-specific impact on GAS recognition. All GOF mutants showed a higher proportion of promoters lacking a ChIP-Seq peak, compared to STAT1 WT, suggesting that the promoter of these genes under standard conditions does not get frequented by STAT1 upon IFNγ stimulation. As these genes are nonetheless up-regulated upon IFNγ stimulation, STAT1 GOF mutants likely recognize more divergent, non-canonical GAS sequences. For STAT1 R274W and R321S the differences with STAT1 WT were less pronounced, suggesting that these two mutants may be less promiscuous.

DISCUSSION

To date, 105 different mutations in STAT1 have been associated with a GOF and more are identified each year. Most STAT1 GOF patients are affected by CMC, but a wider range of phenotypes has been associated with this disease. Symptomatic treatment has been inconsistent, with varying response to jakinibs (22, 41–43). It remains unknown whether this clinical heterogeneity can, at least in part, be explained by differences in STAT1 GOF molecular mechanisms or whether this is mainly driven by the genetic background of the patient.

Recently, we demonstrated that distinct dynamics can underlie STAT1 GOF for a subset of STAT1 GOF mutations that spread over the protein and associate with diverse GOF phenotypes in patients (STAT1 GOF R274W, R321S, T419R, and N574I), including faster nuclear accumulation and reduced

mobility in the nucleus following IFNγ stimulation (24). These results were in line with other reports (21–23).
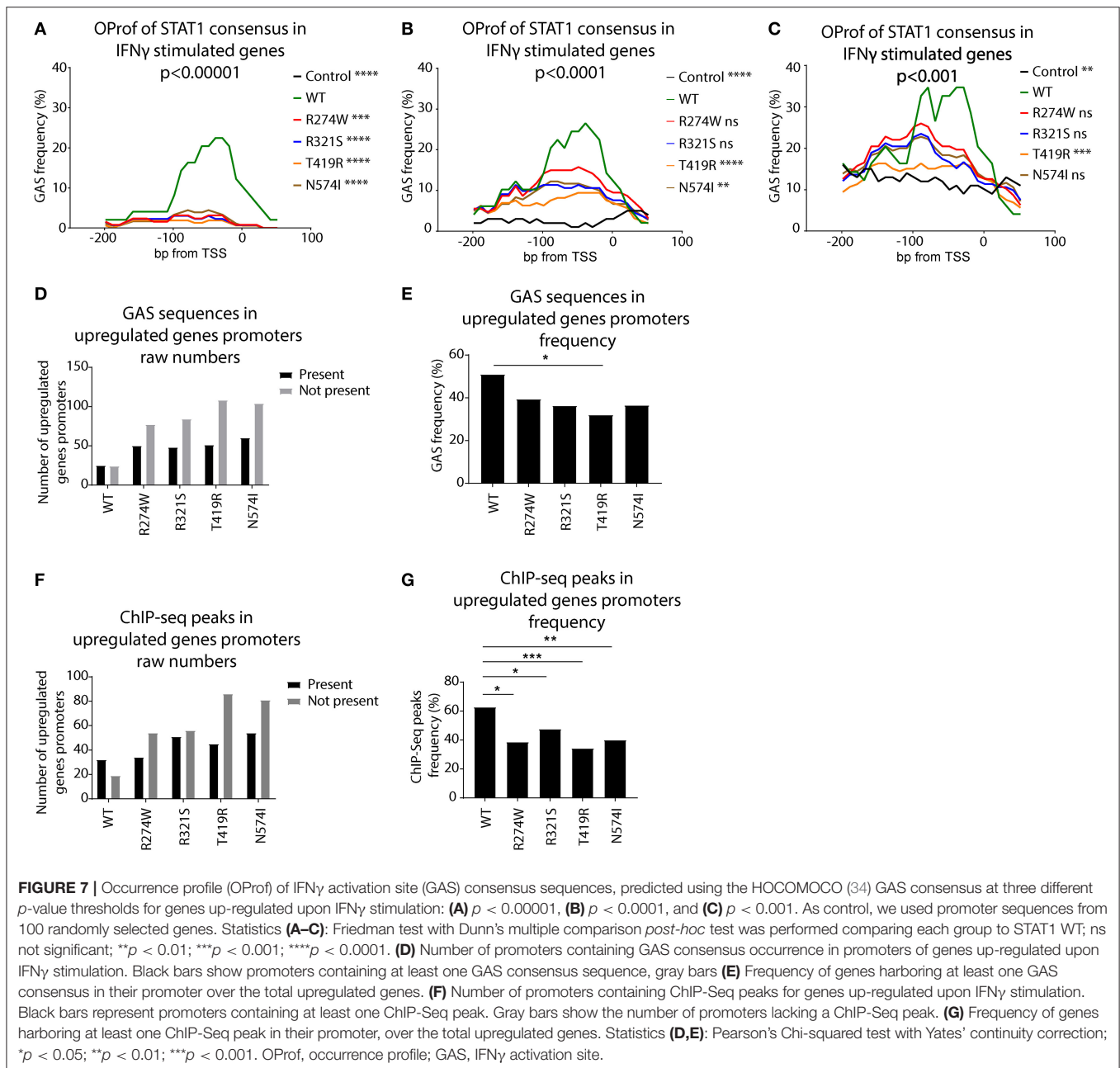
In the present study, we hypothesized that the different molecular mechanisms identified for various GOF mutants may also affect the transcriptomic fingerprint of each mutant, in an effort to shed light on a potential link between molecular mechanism and varying phenotypes.

We performed mRNA sequencing in an isogenic U3A $STAT1^{-/-}$ cell model complemented with STAT1 wild type or STAT1 GOF R274W, R321S, T419R, and N574I and compared their transcriptomes, with and without IFNγ stimulation, to identify dysregulated pathways extending beyond the group of immediate IFNγ-response genes.

Our study highlights that STAT1 GOF mutants (with the exception of T419R) up-regulated a subset of genes upon IFNγ that is common to STAT1 WT. In addition, another subset of genes up-regulated (such as *CISH, CD274*) was shared by all the GOF mutants, possibly associated with the common GOF phenotype. Finally, each mutant also presented a private subset of up- and down-regulated genes, possibly associated with mutation specific differences. Our analysis indicated that loss of specificity for the canonical GAS consensus sequences might drive this broader activation pattern. The most outspoken effect was observed for the T419R mutant, which up- and down-regulated multiple non-IRGs. We hypothesize that this is a result of the altered contact with genomic DNA.

The demonstration of distinct transcriptomic fingerprints for specific STAT1 GOF mutants upon IFNγ stimulation, may be explained by differences in the molecular mode of action of the respective GOF mutations, described previously by our group and others (21–24).

To our knowledge, this is the first study comparing the effect of different STAT1 GOFs on gene expression in an isogenic cell

FIGURE 7 | Occurrence profile (OProf) of IFNγ activation site (GAS) consensus sequences, predicted using the HOCOMOCO (34) GAS consensus at three different p-value thresholds for genes up-regulated upon IFNγ stimulation: **(A)** $p < 0.00001$, **(B)** $p < 0.0001$, and **(C)** $p < 0.001$. As control, we used promoter sequences from 100 randomly selected genes. Statistics **(A–C)**: Friedman test with Dunn's multiple comparison post-hoc test was performed comparing each group to STAT1 WT; ns not significant; **$p < 0.01$; ***$p < 0.001$; ****$p < 0.0001$. **(D)** Number of promoters containing GAS consensus occurrence in promoters of genes up-regulated upon IFNγ stimulation. Black bars show promoters containing at least one GAS consensus sequence, gray bars **(E)** Frequency of genes harboring at least one GAS consensus in their promoter over the total upregulated genes. **(F)** Number of promoters containing ChIP-Seq peaks for genes up-regulated upon IFNγ stimulation. Black bars represent promoters containing at least one ChIP-Seq peak. Gray bars show the number of promoters lacking a ChIP-Seq peak. **(G)** Frequency of genes harboring at least one ChIP-Seq peak in their promoter, over the total upregulated genes. Statistics **(D,E)**: Pearson's Chi-squared test with Yates' continuity correction; *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$. OProf, occurrence profile; GAS, IFNγ activation site.

model. In 2017, Fujiki et al. (21) performed mRNA sequencing of STAT1 R274Q mutant in U3C cells. Similar to our results, the authors observed an overlap of genes modulated by both STAT1 WT and the R274Q mutant. In contrast to our results, they did not observe a significant difference in expression magnitude between STAT1 WT and R274Q. This discrepancy may be attributed to the different conditions used: 24 h of IFNγ stimulation using 2 biological replicates compared to 4 h of IFNγ stimulation using 5 biological repeats in our study. We specifically chose the 4 h time-point in order to monitor mRNA status early after IFNγ addition: we wanted to study the direct effect of the STAT1 GOF mutations on gene transcription, and aimed to limit the activation of secondary genes/pathways.

Additional time points would be valuable to study. Still, in line with our results, they showed a similar TSS-proximal ChIP-Seq peak distribution at IFNγ modulated genes for STAT1 WT and R274Q mutant. In 2018, Ovadia et al. (26) described a patient with a H629Y STAT1 GOF mutation, located in the STAT1 SH2 domain. RNA-seq analysis on U3A cells transiently transfected with STAT1 WT or H629Y, and R274G STAT1 GOF mutants and revealed a minority of common GOF IFNγ-induced genes and a larger amount of mutation-specific up-regulated genes. In line with this, Meesilpavikkai et al. (39) described a V653I STAT1 GOF mutation, also residing in the SH2 domain resulting in a common STAT1 GOF phenotype (CMC and autoimmunity), together with atypical diverse infections and impaired cytokine

regulation. This might allude to a combination of dysregulation of a common trunk of GOF genes together with mutation-specific dysregulations.

Our study has several limitations. First, we leveraged a U3A cancer cell line, depleted for endogenous STAT1 and complemented with STAT1 WT or GOF mutants using lentiviral vector expression. This resulted in cell lines homozygous for the GOF mutations, in contrast to their heterozygous context in patients. Second, the use of a cancer cell line could affect the transcriptomic fingerprint and limit extrapolation toward patient-derived cells and to the patients' phenotype. Indeed, this study does not explain the phenotypic variation observed between STAT1 GOF patients, and additional experiments in immunologically more relevant cells are required to verify this in more detail. However, complementing a $STAT1^{-/-}$ cell line with the respective GOF mutations, allowed us to evaluate the mutations in an isogenic setting and to address additional confounders, with the only difference between each of our groups being a single point mutation in STAT1, unaffected by donor-to-donor variability. Moreover, the limited number of patients that can be sampled per STAT1 mutation (105 different mutants for over 400 patients described), currently complicates genotype-phenotype association studies, especially for the more divergent clinical presentations and even more rare associated features (such as vascular abnormalities). We believe that binning these STAT1 GOF mutations into mechanistically homogenous groups might help organize the observed phenotypic heterogeneity. In our work, each GOF mutant up- and/or down-regulated a subset of private genes, possibly associated with mutation-specific phenotypes, although pathway analysis did not show a clear correlation. Third, the STAT1 ChIP-Seq analysis derived from HeLa-S3 cells, was used to evaluate data obtained in U3A cells, similarly to Fujiki et al. (21), possibly affecting the correlation. Nevertheless, the IFNγ-stimulated STAT1 WT condition demonstrated a high proportion of ChIP-Seq peaks in stimulated genes, indicating an important degree of concordance between these models. We analyzed small TSS-proximal regions of the up- or down-regulated genes (−200 to +50 bp), to focus on the regulatory sequences strictly correlated with these genes, although regulatory regions located further upstream have been implicated in regulation by STAT1 (44, 45). Finally, we studied only the effect of IFNγ, and therefore STAT1 homodimerization, on the transcriptome. STAT1 is known to also form heterodimers with STAT2 upon IFNα stimulation, and STAT1/STAT2 heterodimers may also be involved in the STAT1 GOF pathogenesis. Likewise, STAT1/STAT3 heterodimers can be formed upon IL6 or IL27 stimulation (46). Neither of these heterodimers were addressed in this study. The effect of the different STAT1 GOF mutants on the transcriptome upon type I IFN, or IL6/IL27 stimulation would therefore certainly be worth analyzing.

In conclusion, our work further expands the knowledge on STAT1 and STAT1 GOF mutations. Our data are in line with previous findings of different molecular routes toward a STAT1 GOF, but also provide evidence for common and GOF mutation-specific impacts on the transcriptome, which in part

may explain STAT1 GOF phenotypic presentation, but this requires additional experiments in physiologically more relevant cells. In-depth analysis of gene expression patterns associated with STAT1 GOF mutations enhances our understanding of the complexity of STAT1 GOF pathophysiology and genotype-phenotype correlations. In addition, the identification of specific STAT1 GOF up-regulated genes, can improve our prognostic capacity to identify and possibly stratify patients affected by this disorder, and may ultimately open up new avenues for disease interrogation and targeting.

Our work indicates how some STAT1 GOF mutants have lost specificity for canonical GAS sequences. All GOF mutants up-regulated genes for which a higher proportion of promoters lacked a ChIP-Seq peak compared to STAT1 WT, suggesting that the promoter of these genes is normally not bound by STAT1 upon IFNγ stimulation. The fact that their gene expression is up-regulated in STAT1 GOF underscored that STAT1 GOF recognized non-canonical GAS sequences as well. For the mutant affecting a DNA-interacting aa, T419R, even promoters lacking GAS sequences showed IFNγ-mediated stimulation. Most importantly, this work revealed how single point mutations, even when resulting in a similar common phenotype, may affect the transcriptome differently, helping to explain the pathogenesis of the wide range of STAT1 GOF phenotypes.

DATA AVAILABILITY STATEMENT

The datasets presented in the study were submitted to GEO database. This data can be found here: https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE166392.

AUTHOR'S NOTE

While the first STAT1 gain-of-function (GOF) mutations were described in 2011, a mechanistic correlation between the patients' mutations and their phenotype is still missing. The major cellular hallmarks of this disorder is increased phosphorylation of STAT1 protein and increased expression of interferon regulated genes. In this study, we analyzed different STAT1 GOF in an isogenic U3A-cell model to correlate the transcriptome with the underlying molecular mechanism, previously characterized by our group. We show how STAT1 GOF hyper-activate the physiological STAT1-regulated genes and on top dysregulate an additional set of GOF shared and mutation-specific genes upon interferon gamma (IFNγ) stimulus, including CISH and CD274, two genes involved in the inhibition of the STAT3 pathway and Th-17 differentiation. Following analysis of GAS sequence occurrence in the respective promoter regions, we observed different degree of loss of specificity for the canonical STAT1 consensus GAS sequence by STAT1 GOF mutants, depending on the specific GOF mutation. Overall, this work contributes to define STAT1 GOF mutations as a heterogeneous group of mutations, sharing common features and, at the same time, different underlying molecular mechanisms and degrees of transcriptional dysregulation.

AUTHOR CONTRIBUTIONS

FUNDING

ACKNOWLEDGMENTS

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fimmu.2021.632997/full#supplementary-material

REFERENCES

1. Liu L, Okada S, Kong XF, Kreins AY, Cypowyj S, Abhyankar A, et al. Gain-of-function human STAT1 mutations impair IL-17 immunity and underlie chronic mucocutaneous candidiasis. *J Exp Med.* (2011) 208:1635–48. doi: 10.1084/jem.20110958

2. van de Veerdonk FL, Plantinga TS, Hoischen A, Smeekens SP, Joosten LAB, Gilissen C, et al. STAT1 mutations in autosomal dominant chronic mucocutaneous candidiasis. *N Engl J Med.* (2011) 365:54–61. doi: 10.1056/NEJMoa1100102

3. Okada S, Asano T, Moriya K, Boisson-Dupuis S, Kobayashi M, Casanova JL, et al. Human STAT1 gain-of-function heterozygous mutations: chronic mucocutaneous candidiasis and type I interferonopathy. *J Clin Immunol.* (2020) 40:1065–81. doi: 10.1007/s10875-020-00847-x

4. Toubiana J, Okada S, Hiller J, Oleastro M, Lagos Gomez M, Aldave Becerra JC, et al. Heterozygous STAT1 gain-of-function mutations underlie an unexpectedly broad clinical phenotype. *Blood.* (2016) 127:3154–64. doi: 10.1182/blood-2015-11-679902

5. Zerbe CS, Marciano BE, Katial RK, Santos CB, Adamo N, Hsu AP, et al. Progressive multifocal leukoencephalopathy in primary immune deficiencies: STAT1 gain of function and review of the literature. *Clin Infect Dis.* (2016) 62:986–94. doi: 10.1093/cid/civ1220

6. Uzel G, Sampaio EP, Lawrence MG, Hsu AP, Hackett M, Dorsey MJ, et al. Dominant gain-of-function STAT1 mutations in FOXP3WT IPEX-like Syndrome. *J Allergy Clin Immunol.* (2013) 131:1611–23.e3. doi: 10.1016/j.jaci.2012.11.054

7. Maeshima K, Ishii K, Shibata H. An adult fatal case with a STAT1 gain-of-function mutation associated with multiple autoimmune diseases. *J Rheumatol.* (2019) 46:325–7. doi: 10.3899/jrheum.180210

8. Kilic SS, Puel A, Casanova J-L. Orf infection in a patient with Stat1 gain-of-function. *J Clin Immunol.* (2015) 35:80–3. doi: 10.1007/s10875-014-0111-7

9. Hartono SP, Vargas-Hernández A, Ponsford MJ, Chinn IK, Jolles S, Wilson K, et al. Novel STAT1 gain-of-function mutation presenting as combined immunodeficiency. *J Clin Immunol.* (2018) 38:753–6. doi: 10.1007/s10875-018-0554-3

10. Forbes LR, Vogel TP, Cooper MA, Castro-Wagner J, Schussler E, Weinacht KG, et al. Jakinibs for the treatment of immune dysregulation in patients with gain-of-function signal transducer and activator of transcription 1 (STAT1) or STAT3 mutations. *J Allergy Clin Immunol.* (2018) 142:1665–9. doi: 10.1016/j.jaci.2018.07.020

11. Au-Yeung N, Mandhana R, Horvath CM. Transcriptional regulation by STAT1 and STAT2 in the interferon JAK-STAT pathway. *JAKSTAT.* (2013) 2:e23931. doi: 10.4161/jkst.23931

12. Darnell JE, Kerr IM, Stark GR. Jak-STAT pathways and transcriptional activation in response to IFNs and other extracellular signaling proteins. *Science.* (1994) 264:1415–21. doi: 10.1126/science.8197455

13. O'Shea JJ, Holland SM, Staudt LM. JAKs and STATs in immunity, immunodeficiency, and cancer. *N Engl J Med.* (2013) 368:161–70. doi: 10.1056/NEJMra1202117

14. Zhong M, Henriksen MA, Takeuchi K, Schaefer O, Liu B, Hoeve J ten, et al. Implications of an antiparallel dimeric structure of nonphosphorylated STAT1 for the activation–inactivation cycle. *Proc Natl Acad Sci USA.* (2005) 102:3966–71. doi: 10.1073/pnas.0501063102

15. Zheng J, Veerdonk FL van de, Crossland KL, Smeekens SP, Chan CM, Shehri TA, et al. Gain-of-function STAT1 mutations impair STAT3 activity in patients with chronic mucocutaneous candidiasis (CMC). *Eur J Immunol.* (2015) 45:2834–46. doi: 10.1002/eji.201445344

16. Wenta N, Strauss H, Meyer S, Vinkemeier U. Tyrosine phosphorylation regulates the partitioning of STAT1 between different dimer conformations. *Proc Natl Acad Sci USA.* (2008) 105:9238–43. doi: 10.1073/pnas.0802130105

17. Mertens C, Zhong M, Krishnaraj R, Zou W, Chen X, Darnell JE. Dephosphorylation of phosphotyrosine on STAT1 dimers requires extensive spatial reorientation of the monomers facilitated by the N-terminal domain. *Genes Dev.* (2006) 20:3372–81. doi: 10.1101/gad.1485406

18. Katsoulidis E, Li Y, Mears H, Platanias LC. The p38 mitogen-activated protein kinase pathway in interferon signal transduction. *J Interferon Cytokine Res.* (2005) 25:749–56. doi: 10.1089/jir.2005.25.749

19. Kim HS, Lee M-S. STAT1 as a key modulator of cell death. *Cell Signal.* (2007) 19:454–65. doi: 10.1016/j.cellsig.2006.09.003

20. Varinou L, Ramsauer K, Karaghiosoff M, Kolbe T, Pfeffer K, Müller M, et al. Phosphorylation of the Stat1 transactivation domain is required for full-fledged IFN-γ-dependent innate immunity. *Immunity.* (2003) 19:793–802. doi: 10.1016/S1074-7613(03)00322-4

21. Fujiki R, Hijikata A, Shirai T, Okada S, Kobayashi M, Ohara O. Molecular mechanism and structural basis of gain-of-function of STAT1 caused by pathogenic R274Q mutation. *J Biol Chem.* (2017) 292:6240–54. doi: 10.1074/jbc.M116.753848

22. Zimmerman O, Rösler B, Zerbe CS, Rosen LB, Hsu AP, Uzel G, et al. Risks of ruxolitinib in STAT1 gain-of-function-associated severe fungal disease. *Open Forum Infect Dis.* (2017) 4:ofx202. doi: 10.1093/ofid/ofx202

23. Petersen J, Staab J, Bader O, Buhl T, Ivetic A, Meyer T. Identification of a distinct subset of disease-associated gain-of-function missense mutations in the STAT1 coiled-coil domain as system mutants. *Mol Immunol.* (2019) 114:30–40. doi: 10.1016/j.molimm.2019.07.008

24. Giovannozzi S, Lemmens V, Hendrix J, Gijsbers R, Schrijvers R. Live cell imaging demonstrates multiple routes toward a STAT1 gain-of-function phenotype. *Front Immunol.* (2020) 11:1114. doi: 10.3389/fimmu.2020.01114

25. Zuo Y, Feng Q, Jin L, Huang F, Miao Y, Liu J, et al. Regulation of the linear ubiquitination of STAT1 controls antiviral interferon signaling. *Nat Commun.* (2020) 11:1146. doi: 10.1038/s41467-020-14948-z

26. Ovadia A, Sharfe N, Hawkins C, Laughlin S, Roifman CM. Two different STAT1 gain-of-function mutations lead to diverse IFN-γ-mediated gene expression. *NPJ Genomic Medicine.* (2018) 3:1–5. doi: 10.1038/s41525-018-0063-6

27. Andrews S. *FastQC: A Quality Control Tool for High Throughput Sequence Data.* (2010). Available online at: http://www.bioinformatics.babraham.ac.uk/projects/fastqc/ (accessed February 01, 2021).

28. Aronesty E. *ea-utils: Command-Line Tools for Processing Biological Sequencing Data.* (2011). Available online at: https://github.com/ExpressionAnalysis/ea-utils (accessed February 01, 2021).

29. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics.* (2013) 29:15–21. doi: 10.1093/bioinformatics/bts635

30. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The sequence alignment/map format and SAMtools. *Bioinformatics.* (2009) 25:2078–9. doi: 10.1093/bioinformatics/btp352

31. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* (2014) 15:550. doi: 10.1186/s13059-014-0550-8

32. Marini F, Binder H. pcaExplorer: an R/Bioconductor package for interacting with RNA-seq principal components. *BMC Bioinformatics.* (2019) 20:331. doi: 10.1186/s12859-019-2879-1

33. Ambrosini G, Praz V, Jagannathan V, Bucher P. Signal search analysis server. *Nucleic Acids Res.* (2003) 31:3618–20. doi: 10.1093/nar/gkg611

34. Kulakovskiy IV, Vorontsov IE, Yevshin IS, Sharipov RN, Fedorova AD, Rumynskiy EI, et al. HOCOMOCO: towards a complete collection of transcription factor binding models for human and mouse via large-scale ChIP-Seq analysis. *Nucleic Acids Res.* (2018) 46:D252–9. doi: 10.1093/nar/gkx1106

35. Robertson G, Hirst M, Bainbridge M, Bilenky M, Zhao Y, Zeng T, et al. Genome-wide profiles of STAT1 DNA association using chromatin immunoprecipitation and massively parallel sequencing. *Nat Methods.* (2007) 4:651–7. doi: 10.1038/nmeth1068

36. Ambrosini G, Dreos R, Kumar S, Bucher P. The ChIP-Seq tools and web server: a resource for analyzing ChIP-seq and other types of genomic data. *BMC Genom.* (2016) 17:938. doi: 10.1186/s12864-016-3288-8

37. van Zelm MC, Bosco JJ, Aui PM, De Jong S, Hore-Lacy F, O'Hehir RE, et al. Impaired STAT3-dependent upregulation of IL2Rα in B cells of a patient with a STAT1 gain-of-function mutation. *Front Immunol.* (2019) 10:768. doi: 10.3389/fimmu.2019.00768

38. Romberg N, Morbach H, Lawrence MG, Kim S, Kang I, Holland SM, et al. Gain-of-function STAT1 mutations are associated with PD-L1 overexpression and a defect in B-cell survival. *J Allergy Clin Immunol.* (2013) 131:1691–3. doi: 10.1016/j.jaci.2013.01.004

39. Meesilpavikkai K, Dik WA, Schrijver B, Nagtzaam NMA, van Rijswijk A, Driessen GJ, et al. A novel heterozygous mutation in the STAT1 SH2 domain causes chronic mucocutaneous candidiasis, atypically diverse infections, autoimmunity, and impaired cytokine regulation. *Front Immunol.* (2017) 8:274. doi: 10.3389/fimmu.2017.00274

40. Dreos R, Ambrosini G, Périer RC, Bucher P. The eukaryotic promoter database: expansion of EPDnew and new promoter analysis tools. *Nucleic Acids Res.* (2015) 43:D92–6. doi: 10.1093/nar/gku1111

41. Bloomfield M, Kanderová V, Paračková Z, Vrabcová P, Svaton M, Fronková E, et al. Utility of ruxolitinib in a child with chronic mucocutaneous candidiasis caused by a novel STAT1 gain-of-function mutation. *J Clin Immunol.* (2018) 38:589–601. doi: 10.1007/s10875-018-0519-6

42. Moriya K, Suzuki T, Uchida N, Nakano T, Katayama S, Irie M, et al. Ruxolitinib treatment of a patient with steroid-dependent severe autoimmunity due to STAT1 gain-of-function mutation. *Int J Hematol.* (2020) 112:258–62. doi: 10.1007/s12185-020-02860-7

43. Mössner R, Diering N, Bader O, Forkel S, Overbeck T, Gross U, et al. Ruxolitinib induces interleukin 17 and ameliorates chronic mucocutaneous candidiasis caused by STAT1 gain-of-function mutation. *Clin Infect Dis.* (2016) 62:951–3. doi: 10.1093/cid/ciw020

44. Abou El Hassan M, Huang K, Eswara MBK, Xu Z, Yu T, Aubry A, et al. Properties of STAT1 and IRF1 enhancers and the influence of SNPs. *BMC Mol Biol.* (2017) 18:6. doi: 10.1186/s12867-017-0084-1

45. Yuasa K, Hijikata T. Distal regulatory element of the STAT1 gene potentially mediates positive feedback control of STAT1 expression. *Genes Cells.* (2016) 21:25–40. doi: 10.1111/gtc.12316

46. Hirahara K, Onodera A, Villarino AV, Bonelli M, Sciumè G, Laurence A, et al. Asymmetric action of STAT transcription factors drives transcriptional outputs and cytokine specificity. *Immunity.* (2015) 42:877–89. doi: 10.1016/j.immuni.2015.04.014