Heterozygous Cell Models of STAT1 Gain-of-Function Reveal a Broad Spectrum of Interferon-Signature Gene Transcriptional Responses

Ori Scott^{1,2,3*}, Kyle Lindsay¹, Steven Erwood^{1,4}, Chaim M. Roifman^{2,5}, Ronald D. Cohn^{1,3,4,6},
Evgueni A. Ivakine^{1,7*}

¹Genetics and Genome Biology Program, The Hospital for Sick Children Research Institute, Toronto, ON, Canada

²Division of Clinical Immunology and Allergy, Hospital for Sick Children, University of Toronto, Toronto, ON, Canada.

³Institute of Medical Science, University of Toronto, Toronto, ON, Canada.

⁴Department of Molecular Genetics, University of Toronto, Toronto, ON, Canada.

⁵Canadian Center for Primary Immunodeficiency and The Jeffrey Modell Research Laboratory for the Diagnosis of Primary Immunodeficiency, The Hospital for Sick Children and The University of Toronto, Toronto, ON, Canada

⁶Division of Clinical and Metabolic and Genetics, Hospital for Sick Children, University of Toronto, Toronto, ON, Canada.

⁷Department of Physiology, University of Toronto, Toronto, ON, Canada.

***Co-Corresponding authors:**

Dr. Evgueni A. Ivakine

zhenya.ivakine@sickkids.ca

Dr. Ori Scott

ori.scott@sickkids.ca

Abstract

Signal Transducer and Activator of Transcription 1 (STAT1) gain-of-function (GOF) is an autosomal dominant immune disorder marked by wide infectious predisposition, autoimmunity, vascular disease and malignancy. Its molecular hallmark, elevated phospho-STAT1 (pSTAT1) following interferon (IFN) stimulation, is seen consistently in all patients and may not fully account for the broad phenotypic spectrum associated with this disorder. While over 100 mutations have been implicated in STAT1 GOF, genotype-phenotype correlation remains limited, and current overexpression models may be of limited use in gene expression studies. We generated heterozygous mutants in diploid HAP1 cells using CRISPR/Cas9 base-editing, targeting the endogenous *STAT1* gene. Our models recapitulated the molecular phenotype of elevated pSTAT1, and were used to characterize the expression of five IFN-stimulated genes under a number of conditions. At baseline, transcriptional polarization was evident among mutants compared with wild type, and this was maintained following prolonged serum starvation. This suggests a possible role for unphosphorylated STAT1 in the pathogenesis of STAT1 GOF. Following stimulation with IFN α or IFN γ , differential patterns of gene expression emerged among mutants, including both gain and loss of transcriptional function. This work highlights the importance of modelling heterozygous conditions, and in particular transcription factor-related disorders, in a manner which accurately reflects patient genotype and molecular signature. Furthermore, we propose a complex and multifactorial transcriptional profile associated with various *STAT1* mutations, adding to global efforts in establishing STAT1 GOF genotype-phenotype correlation and enhancing our understanding of disease pathogenesis.

Keywords: STAT1, Interferon, Primary Immunodeficiency, Immune Dysregulation, Cell Model, Heterozygous, CRISPR/Cas9, Base-Editing

INTRODUCTION

Signal Transducer and Activator of Transcription (STAT) is a family of 7 structurally homologous transcription factors, activated downstream of various cytokine, growth factor and hormone receptors. At rest, STAT molecules are found in a latent state in the cytoplasm. After receptor ligation, canonical STAT activation follows a common sequence, starting with recruitment of tyrosine kinases from the Janus Kinase (JAK) family, which phosphorylate the cytoplasmic portion of the receptor to form a docking site for STAT. This is followed by STAT recruitment, tyrosine phosphorylation and multimerization to form active transcription factors which then migrate to the nucleus.¹⁻⁵ Within the STAT family, STAT1 is pivotal in mediating transcriptional responses to cytokines of the interferon (IFN) family, as well as interleukin-27 (IL-27). This is achieved by the formation of transcription complexes, known as interferon-stimulated gene factor 3 (ISGF3) and gamma activating factor (GAF). ISGF3 is a hetero-trimer consisting of STAT1, STAT2 and IFN-regulatory factor 9 (IRF9). It is primarily formed in the context of type I and III IFN stimulation, and binds to interferon-stimulated response element (ISRE) to regulate gene expression. In contrast, GAF is a STAT1 homo-dimer, predominantly activated in response to type II IFN and IL-27, which exerts its transcriptional activity by binding to gamma-activating sequence (GAS) within gene promoters.²⁻⁶

Monogenic defects in the *STAT1* gene have been implicated in three distinct human disorders to date. Autosomal recessive complete loss of function (LOF) leads to severe and early-onset susceptibility to viral and Mycobacterial infections. Individuals harbouring two hypomorphic alleles display a milder form of this disease.⁷⁻⁹ A second entity, caused by heterozygous dominant negative mutations, is characterized by Mendelian Susceptibility to Mycobacterial Disease (MSMD).¹⁰⁻¹² The third disorder, STAT1 gain-of-function (GOF), was first described among a subset of individuals with chronic mucocutaneous Candidiasis and autoimmune thyroid disease, harboring heterozygous point mutations in *STAT1*.^{13,14} The molecular hallmark of the disease was defined as increased levels of

phosphorylated STAT1 (with respect to the Tyrosine 701 residue) in response to IFN stimulation.¹⁴

Since its first description in 2011, STAT1 GOF has been diagnosed in hundreds of patients, and its phenotypic spectrum expanded.^{15,16} Infectious predisposition includes fungal, bacterial, viral, opportunistic and Mycobacterial infections. Over one third of patients display autoimmune features, with hypothyroidism, type 1 diabetes, and cytopenias being common manifestations. Vascular abnormalities, notably intra-cerebral aneurysms, have been described at an increased frequency compared with the general population. Malignancies, in particular squamous cell carcinoma, are seen in up to 5% of patients.¹⁵⁻²⁰

In the decade since STAT1 GOF was first described, strides have been made in characterizing the disorder and its underlying pathophysiology. A prominent example is the impaired Th17 response observed in most patients, which has been linked to predisposition to fungal and bacterial infections.^{14,20,21} From an autoimmune standpoint, impaired type I IFN response has been proposed as a possible contributory mechanism, given the heightened IFN signature associated with other autoimmune and inflammatory conditions.²²⁻²⁴ Indeed, alterations in IFN-related gene expression have been found in some patients with STAT1 GOF and clinical features of autoimmunity.²⁵

However, many underlying disease mechanisms have remained elusive, and the genotype-phenotype correlation among patients remains poorly defined. A recent review reported that the presence of a severe complication, defined as invasive infection, cancer, symptomatic aneurysm, and in some cases severe autoimmunity, substantially worsens the prognosis and lowers survival.²⁶ Unfortunately, our ability to predict which patients would be more prone to developing such complications based on their specific mutations is greatly limited. Therefore, the need for developing tools to study the variability across *STAT1* GOF mutations is dire.

In studying the differences across *STAT1* GOF mutations, the use of cell models offers a well-controlled, accessible and non-invasive tool. Previous studies utilizing over-expression models generated in *STAT1*-null U3 fibrosarcoma cells (and more recently, HEK293 cells), have been instrumental in elucidating differences among mutations with respect to STAT1 phosphorylation kinetics, nuclear migration and accumulation.^{14,27-31} However, such models involve the expression of *STAT1* under an exogenous promoter, and do not capture the heterozygous nature of the mutation. In the context of a delicately-regulated transcription factor, over-expression models may therefore be limited in their portrayal of gene expression patterns downstream of STAT1. The current study describes our use of the Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR)/Cas9 system, and in particular CRISPR/Cas9 base-editing, to generate a series of heterozygous cell models harbouring known GOF point mutations, within the endogenous *STAT1* gene. We further use these models to show that *STAT1* GOF mutations result in different patterns of interferon-stimulated gene (ISG) expression, both at baseline and following stimulation with IFN α or IFN γ . We propose that such models may enhance our understanding of this intricate immune disorder, as well as genotype-phenotype correlation among various *STAT1* GOF mutations.

RESULTS

Diploid HAP1 cells were chosen for heterozygous mutation modelling

For the purpose of our model generation we have used HAP1, a cell line originally derived from the KBM-7 chronic myelogenous leukemia cell line. HAP1 were previously used to study cellular responses to IFN types I, II and III, and have well-characterized transcriptional responses to IFN type I and II.³²⁻³⁴ In addition, they are readily amenable to transfection and CRISPR/Cas genome-editing.³⁵ Although HAP1 cells are originally near-haploid, like other haploid cell lines they are known to spontaneously diploidize in cell culture over time.³⁶ HAP1 cells used in this study underwent cell cycle analysis, comparing their DNA content with that of known diploid cells

[wildtype (WT) human fibroblasts]. The mean fluorescence intensity (MFI) ratios of HAP1/fibroblasts for the G₀/G₁ and G₂/M peaks were calculated to be 1.07 and 1.11, respectively, confirming our HAP1 cells to be fully diploid, and therefore suitable for heterozygous mutation modelling (**Supplementary Figure 1**).

Heterozygous *STAT1* mutants were generated using CRISPR/Cas9 base-editing

The following validated GOF transition mutations were chosen for modelling: E235G,³⁷ K278E³⁸ (both in the coiled-coil domain), P329L,^{39,40} T385M^{16,41-49} (in the DNA-binding domain), and D517G¹⁵ (in the linker domain). The dominant negative mutation Y701C,⁵⁰ affecting the JAK-phosphorylated residue Y701, was chosen for comparative modelling as well. A visual representation of modelled mutations within their respective protein domains, as well as the workflow for generating and verifying the *STAT1* mutants, is presented in **Figure 1**. For information regarding clinical manifestations described for each mutation, please refer to **Table I**.

To generate chosen mutations, single-guide RNA (sgRNA) targeting the Cas9 base-editor to the region of interest were cloned into the BPK1520_puroR plasmid. Resultant plasmids, coding for the desired sgRNA as well as a puromycin resistance cassette, were delivered by lipofection into HAP1 cells, concurrently with an additional plasmid coding the respective Cas9 base-editor (SpCas9 ABEmax, SpG CBE4max, or SpRY ABEmax). Following puromycin selection to enrich for transfected cells, base-editing efficiency was evaluated in the bulk-population by sequencing. Overall, editing efficiency ranged from 21 to 77% for the desired target nucleotide. As base editors each have a characteristic “editing window”, adjacent nucleotides to the nucleotide of interest may be prone to “bystander” editing. As an example, if two adjacent adenine residues are both within the editing window for an adenine base editor, both may be targeted and converted to guanine, albeit at different frequencies depending on their position within the editing window. In this study, bystander

mutations involving editing of adjacent bases within the editing window occurred at a frequency of 1-38%. Cells from the bulk population were single-cell sorted, and resultant single clones were screened by Sanger sequencing for presence of the desired mutation in a heterozygous state, and for absence of non-silent bystander mutations. For all mutations, the number of clones required to be screened to identify a heterozygous mutant without bystander mutations ranged from 12 to 39. For one mutation, E235G, only clones containing a second, silent mutation in an adjacent base could be identified. However, these clones recapitulated the desired amino acid change and were therefore deemed appropriate for further downstream work. Altogether, all chosen amino acid changes could be modelled using base-editing. For each mutation generated, downstream analysis was carried out on a minimum of 2 independent clones. A summary of base-editing performed in this work, including sgRNA and base-editors used, frequency of editing events in the bulk population, and number of clones required to screen to find a heterozygous mutation is provided in **Table II**.

Immunoblotting for pSTAT1 validated the molecular designation of generated *STAT1* mutants

In order to establish the validity of our newly-generated cell models, we proceeded to validate their designation as “GOF” or “LOF” based on Tyrosine-701 phosphorylation in response to IFN. To this end, cells were stimulated with IFN γ at a dose of 10ng/mL for a period of 60 minutes. pSTAT1(Y701) and total STAT1 were measured by immunoblotting at baseline in unstimulated cells, as well as following stimulation. Measurements were repeated over 5 independent experiments and densitometry analysis performed (**Figure 2; Supplementary Figure 2**). At baseline, pSTAT1 was not detectable in any of the samples. Following IFN γ stimulation, levels of pSTAT1 increased across all samples, but were significantly higher across all GOF mutants compared with WT [E235G ($p<0.05$), K278E, P329L, T385M ($p<0.001$), D517G ($p<0.01$); One-way ANOVA with Dunnett’s post-hoc test]. By the same token, pSTAT1 was lower in the Y701C LOF mutant compared with WT after stimulation ($p<0.05$). These results establish that modelled heterozygous *STAT1* mutations in

HAP1 cells lead to the same functional consequences with respect to protein phosphorylation as are seen in patients. Total STAT1 was significantly elevated only in the T385M mutant ($p < 0.05$), although a trend toward increased total STAT1, not reaching statistical significance, was seen in P329L ($p = 0.0635$) and D517G ($p = 0.0792$) as well.

Gene expression studies demonstrated baseline polarization among *STAT1* mutants

To evaluate the transcriptional impact of the various *STAT1* mutations, we used quantitative real-time PCR (qRT-PCR) to assess differences in interferon-stimulated gene (ISG) expression under a number of conditions, including baseline, serum starvation, and stimulation with IFN type I and II. A visual summary of STAT1 signaling associated with the above conditions is presented in **Figure 3**. A set of five ISG was chosen consisting of *GBP1*, *IFIT2*, *IRF1*, *APOL6* and *OAS1*. These genes were selected as they are known to increase in human cells at least 2-fold following either IFN α or IFN γ stimulation.⁵¹ Moreover, previous studies specifically done in HAP1 cells showed these genes to increase at least 2-fold following stimulation with either type I or II IFN stimulation.³⁴

At baseline, significant differences in gene expression among WT and some of the mutants were already noted, involving a mixed pattern of both increased and decreased expression (**Figure 4a, Table III**). The most prominent mutants showing elevated expression (2 genes each) were E235G and P329L; E235G showed increased expression of *GBP1* ($p < 0.01$) and *APOL6* ($p < 0.001$), while P329L demonstrated increased *APOL6* ($p < 0.001$) and *OAS1* expression ($p < 0.01$). Other changes noted included reduced *OAS1* expression in Y701C ($p < 0.05$), and elevated *IFIT2* expression in D517G ($p < 0.01$). No baseline differences were noted between T385M and WT, or between K278E and WT. In order to ensure that the observed transcriptional differences were inherent to the mutants, rather than a result of external cell-culture cytokine/growth factor stimuli, gene expression was measured following 24 hours of serum starvation (**Figure 4b, Table III**). In the context of serum

starvation, the mutants E235G and P329L maintained a profile of elevated gene expression. E235G demonstrated increased expression of *GBP1* ($p<0.0001$) and *APOL6* ($p<0.01$) compared to WT, while P329L showed enhanced expression of *APOL6* ($p<0.0001$), *OAS1* ($p<0.001$) and in addition, *IRF1* ($p<0.0001$). Mutants which were previously no different than WT (K278E and T385M) with respect to all genes measured, remained so under serum starvation. In the context of these results, and given that serum starvation in and of itself may impact STAT1 activation and gene transcription in an IFN-independent manner,⁵² we elected to proceed with IFN stimulation experiments under normal cell culture conditions, as described by others.^{14,16,25,27,28,30,31}

STAT1 mutants displayed a differential response to IFN α stimulation involving both loss and gain of transcriptional function

After establishing baseline expression levels, transcriptional responses (fold increase in expression after stimulation) were measured following a 6-hour stimulation with IFN α (10ng/mL) (**Figure 5a; Table III**). Of all GOF mutants, only T385M showed an elevated fold change (FC) across all genes measured compared to WT [*GBP1* ($p<0.0001$), *IFIT2* ($p<0.01$), *IRF1* ($p<0.0001$), *APOL6* ($p<0.01$), *OAS1* ($p<0.05$)]. E235G, which had an elevated baseline expression of *GBP1* and *APOL6*, showed an elevated FC in the expression of *IFIT2* ($p<0.0001$) and *IRF1* ($p<0.0001$), with no difference in FC with respect to other genes. In contrast, P329L, which previously showed increased baseline expression of *APOL6* and *OAS1*, showed a reduced FC in the expression of *GBP1* ($p<0.05$), *APOL6* ($p<0.05$) and *OAS1* ($p<0.01$) compared with WT. Decreased FC compared to WT was also seen in K278E [*IRF1* ($p<0.05$), *APOL6* ($p<0.05$), *OAS1* ($p<0.0001$)] and D517G [*APOL6* ($p<0.05$), *OAS1* ($p<0.001$)]. The LOF mutant Y701C was marked by reduced FC across all but 1 gene compared to WT [*GBP1* ($p<0.01$), *IFIT2* ($p<0.001$), *IRF1* ($p<0.05$), *APOL6* ($p<0.01$)].

Transcriptional response of *STAT1* mutants to IFN γ stimulation differed from IFN α responses

We sought to determine whether transcriptional responses of *STAT1* mutants to IFN α could accurately predict their responses to stimulation with IFN γ (**Figure 5b; Table III**). Following a 6-hour stimulation with IFN γ (10ng/mL), 3 mutants showed vastly different transcriptional responses compared with those seen following IFN α stimulation. T385M, which previously showed a transcriptional GOF with respect to all genes following IFN γ stimulation, now showed a reduced FC of *APOL6* ($p < 0.0001$), with no other differences compared with WT. E235G previously demonstrated elevated FC in expression of *IFIT2* and *IRF1* in response to IFN α , whereas no differences from WT were seen in these genes with IFN γ stimulation. In contrast, FC of *GBP1* ($p < 0.05$) and *APOL6* ($p < 0.0001$) were now decreased in E235G, and that of *OAS1* increased ($p < 0.001$) compared to WT. One more mutant showing substantial differences in responses to IFN γ and IFN α was K278E; while IFN α stimulation resulted in reduced FC of *GBP1*, *APOL6* and *OAS1* compared with WT, IFN γ stimulation caused an increased FC in *IFIT2* ($p < 0.01$), with no significant differences with respect to other genes. Mutants showing more similar trends in response to both IFN γ and IFN α included P329L, D517G, and the LOF mutant Y701C. P329L again showed reduced FC of *GBP1* ($p < 0.05$) and *APOL6* ($p < 0.0001$), though FC for *OAS1* was no different than WT. D517G demonstrated reduced FC in *APOL6* ($p < 0.0001$), but no difference in FC of *OAS1*. Y701C showed consistent responses to IFN γ and IFN α , with reduced FC seen again for *GBP1* ($p < 0.001$), *IFIT2* ($p < 0.001$), *IRF1* ($p < 0.001$) and *APOL6* ($p < 0.0001$) but no difference in FC of *OAS1*.

DISCUSSION

The current study demonstrated for the first time the implementation of CRISPR/Cas9 base-editing in creating heterozygous cells models of STAT1 GOF and LOF. Previously, most studies of STAT1 GOF were performed in patient samples, or in overexpression models. Work done in patient-derived

samples, be they primary or immortalized cells, has provided a wealth of information regarding pathway alterations associated with STAT1 GOF. However, such samples are a limited resource necessitating access to patients, their obtaining can be invasive, and no perfectly isogenic control is available for comparison. Moreover, as sample collection is typically done after patients have already become symptomatic, it is challenging to exclude variability relating to factors such as concurrent systemic inflammation, infection or immunosuppressive/modulatory treatments. In regards to overexpression models, the majority of studies have employed the *STAT1*-null U3 fibrosarcoma cells (and more recently, HEK293 cells on a WT background).^{14,27-31} These models have been instrumental in studying STAT1 phosphorylation and de-phosphorylation kinetics, as well as migration of STAT1 between the cytoplasm and the nucleus. However, such models are characterized by expression of *STAT1* under an exogenous promoter, and an inaccurate gene dosage. These factors considerably limit the application of overexpression models to the study of precise gene expression and signaling pathway alterations. This limitation is particularly substantial in the case of STAT1, a transcription factor which is under delicate transcriptional control, impacts the expression of other transcription factors, and in itself regulates its own expression.

The current approach of mutant generation via base-editing offers an opportunity to model *STAT1* mutations in a heterozygous manner, and under control of the endogenous gene promoter, resulting in highly relevant cell models for dissecting the molecular pathogenesis of the disease from a transcriptional standpoint. Base-editing is efficient, quick, and enables modelling of a rapidly-expanding repertoire of point mutations.⁵³⁻⁵⁵ As with most CRISPR/Cas9-based applications, the use of base editing may be limited by the need for a protospacer adjacent motif (PAM) in close proximity to the area of interest. However, with the advent of newly engineered base-editors with extended sequence recognition (such as SpG and SpRY editors used in this work), a wider array of PAMs may now be used in targeting sites for base-editing.⁵⁶ Furthermore, while base-editing was previously

limited in its ability to create transversion mutations, recent works have expanded the arsenal of base-
editors, now allowing the generation of certain transversions in addition to transitions.⁵⁷

The molecular hallmark of STAT1 GOF has been designated as elevated pSTAT1 (Tyr701) following type I or II IFN stimulation.¹⁴ However, the uniformity of this finding across all patients is perplexing in the context of high clinical variability. This suggests that elevated pSTAT1 does not fully account for STAT1 GOF disease pathogenesis. The notion that pSTAT1 may be a secondary feature of STAT1 GOF, has received support in recent years. In this regard, some studies in patient samples found STAT1 itself to be elevated, suggesting that total STAT1, rather than pSTAT1, is the primary disease driver of STAT1 GOF.^{58,59} In our current study, pSTAT1 was elevated in all GOF mutants following stimulation (as described in patients). In addition, increased total STAT1 was clearly seen in one GOF mutant, with a trend toward increased total STAT1 in two others. It is possible that elevated STAT1 levels would develop across all mutants over time following repeated stimuli.

Our analysis of gene expression at baseline and following serum starvation further supports the notion of total STAT1, rather than pSTAT1, as driving the transcriptional abnormalities seen in STAT1 GOF. Our cell models showed baseline polarization in terms of ISG expression among certain mutants, particularly E235G and P329L, suggesting that transcriptional homeostasis for some *STAT1* GOF mutants is different than that of WT. These results are recapitulated under conditions of serum starvation, suggesting that this baseline polarization may occur in a manner which is independent of external cytokine stimuli, and possibly independent (or only partially dependent) of pSTAT1. While cytokine-dependent activation of pSTAT1 has been regarded as the canonical pathway of STAT1 signaling, a well-established transcriptional role exists for unphosphorylated STAT1 (U-STAT1).⁶⁰⁻⁶³ U-STAT1 mediates the constitutive baseline activation of many ISG, in a

295 manner which may be both cytokine-dependent and independent. It does so by acting both as a
 296 homodimer, and in complex with other transcription factors such as U-STAT2, and IRF9.^{7,60-64} It is
 297 therefore possible that mutated, unphosphorylated STAT1 molecules may result in differential
 298 transcriptional activity at baseline, causing an abnormal pattern of ISG expression even in “naïve”
 299 cells prior to cytokine stimulation. Taken together, our current findings support the notion that
 300 STAT1 GOF pathogenesis may not be fully attributed to canonical, cytokine-related pSTAT1
 301 activation. Further work would be required to understand the mechanisms leading to baseline gene
 302 expression polarization among STAT1 GOF mutants, and what, if any, is the role of U-STAT1.
 303
 304 Treatment of STAT1 GOF mutants in our study with either IFN γ or IFN α , has shown differential
 305 stimulation responses with a mixed pattern of increased, decreased, or similar fold change in ISG
 306 expression compared to WT. Differences were noted among WT and mutants, between IFN γ or IFN α
 307 stimulation, and also within the same genotype and stimulation group across different genes. A case
 308 in point would be T385M, which showed no baseline differences in ISG expression compared to WT,
 309 a gain of transcriptional function with respect to all genes following IFN α stimulation, and no
 310 difference (and even reduced fold change for one gene) after IFN γ treatment. P329L, which was
 311 marked by increased baseline ISG expression, demonstrated reduced responsiveness to both IFN γ
 312 and IFN α stimuli. Comparatively, E235G, which was also characterized by enhanced baseline ISG
 313 expression, showed reduced transcriptional responses to IFN γ but increased fold change following
 314 stimulation with IFN α . The notion of differential transcriptional response to stimuli in STAT1 GOF
 315 is supported by previous evidence from *in-vitro* work done in patient cells. Kobbe et al showed that
 316 in T-cells from patients with the F172L mutation, fold change in *GBP1* expression was elevated
 317 compared to WT following IFN α , but not IFN γ stimulation, while *MIG1* fold change was elevated
 318 after treatment with either IFN α or IFN γ , but not combined treatment with both.⁶⁵ Meesilpavikkai et

al demonstrated that in T-cells harbouring the V653I mutation, *CXCL10* and *CD274* fold change was increased compared to WT after stimulation with IL-27, but not with IFN γ .⁶⁶ Similar findings were reported by groups studying patient PBMC assessing the expression of various other genes.^{67,68}

The intricate patterns of gene expression seen in STAT1 GOF mutants may result in part from altered STAT1-DNA binding status. For instance, the 329 residue lies one amino acid away from the site of direct DNA interaction. It is therefore conceivable that P329L could result in more constitutive baseline DNA binding and transcriptional activation, but with less potential for further transcriptional enhancement following stimulation. However, another mutation affecting the DNA binding domain, T385M, showed a vastly different gene expression pattern compared with P329L, both in terms of baseline expression and in regards to IFN reactivity, suggesting that additional factors may come into play. One such factor may relate to differential activation of STAT1-dependent and independent transcription factor complexes among the various mutants, leading to differential ISG expression both at baseline and following stimulation.

One important finding of our study relates to the designation of mutations as “GOF” or “LOF”. All in all, each STAT1 GOF mutant in our study showed evidence for transcriptional GOF with respect to at least 1 of the 5 genes measured compared to WT, either by means of increased baseline expression, or increased fold change following stimulation. However, some mutants, such as K278E and D517G showed more evidence for transcriptional LOF rather than GOF. The LOF mutant, Y701C, predictably showed reduced transcriptional responses to both stimuli across 4 of the 5 genes measured, with the exception of *OAS1* which is known to be co-regulated by non-STAT1 dependent transcriptional complexes.⁶ Our findings in STAT1 GOF mutants suggest that the molecular designation of GOF (as it relates to tyrosine phosphorylation), may not always indicate heightened gene expression. The notion of diminished ISG transcriptional responses in some mutants is

supported by recent studies reporting transcriptional LOF in STAT1 GOF. Ovadia et al reported that compared with WT, the H629Y mutation showed either diminished or unchanged fold increase in gene expression in response to IFN γ .²⁹ More recently, work done in a mouse model of the R274Q mutation demonstrated a reduction in the expression of the ISG *Cxcl10* and *Irf1* following viral infection *in-vivo*.⁶⁹

The current study is constrained by a few limitations. These include a relatively small number of genes tested, and the measurement of fold change following single stimulation of naïve cells. Future work will involve larger-scale gene studies, including pathways which extend beyond the immediate group of IFN-response genes. In regards to stimulation of naïve cells, previous work showed that STAT1 GOF cells had an impaired transcriptional response not only upon initial stimulation, but also to re-stimulation.⁴⁵ It would therefore be important in the future to study how the transcriptional responses change over time and following repeated or different stimuli. Such work may further help understand the evolution of transcriptional responses as they occur *in-vivo*. Finally, further work may be merited with regard to mechanisms resulting in differential gene expression across mutants and stimuli, and in particular elucidating the possible contribution of total or U-STAT1 to the abnormal gene expression patterns in STAT1 GOF.

In conclusion, we present a series of heterozygous *STAT1* cell models generated using CRISPR/Cas9 base-editing, showing the utility of this technique in modelling heterozygous immune-mediated disease. Our cell models demonstrate intricate patterns of ISG expression, involving transcriptional abnormalities at baseline, following serum starvation, and after stimulation with type I or II IFN. Taken together, our findings are in line with a growing body of literature suggestive of complex and multi-factorial transcriptional responses in STAT1 GOF, which cannot be simply and generally classified as either gain or loss of function. Moreover, our findings may indicate an important role for

total and U-STAT1 in disease pathogenesis, in addition to the role of elevated pSTAT1. Continued investigation of gene expression patterns associated with *STAT1* mutations, may enhance our understanding of both disease pathophysiology and genotype-phenotype correlation in STAT1 GOF. This, in turn, has the potential to improve our prognostic capacity of patients affected by this disorder, and may ultimately open up new avenues for disease interrogation and targeting.

MATERIALS AND METHODS

Mutation selection

Previously published mutations were selected for modelling according to the following criteria: (1) transition mutations (A•G or C•T) for generation by CRISPR/Cas9 base-editing; (2) patients carrying the mutations met clinical criteria for STAT1 GOF diagnosis; and (3) previously published *in-vitro* analysis confirmed the presence of elevated Tyr701 phosphorylated-STAT1 (pSTAT1) following IFN stimulation. An additional heterozygous loss of function transition mutation, Y701C, was chosen for modelling as well.

Cell culture

HAP1 cells (kind gift of Dr. Aleixo Muise, Toronto) were cultured in IMDM medium (Wisent Bioproducts 319-105-CL) supplemented with 10% heat-inactivated FBS (Wisent Bioproducts 080-150) and 1% penicillin-streptomycin (Wisent Bioproducts 450-201-EL). For serum-starvation experiments, HAP1 cells were placed in IMDM containing 0.25% FBS and 1% penicillin-streptomycin. Fibroblasts (ATCC PCS-201-012) for DNA-content analysis were cultured in DMEM medium (Wisent Bioproducts 319-005-CL) supplemented with 10% heat-inactivated FBS and 1% penicillin-streptomycin. Cell line authentication was done by means of Short Tandem Repeat (STR), performed by The Hospital for Sick Children The Centre for Applied Genomics (TCAG). Cells were confirmed to be Mycoplasma-free using a PCR Mycoplasma Detection Kit (Applied Biological Materials Inc. G238). Cells were cultured in a 37°C humidified incubator containing 5% CO₂.

DNA-content analysis

DNA content analysis was performed as previously described,⁷⁰ comparing HAP1 cells and diploid fibroblasts. Briefly, cells were washed with 1xPBS (Wisent Bioproducts 311-010-CL), trypsinized, and resuspended in complete media (1×10^6 cells/ μ L), to which a low-toxicity, cell-permeable DNA dye (Vybrant DyeCycle VioletTM; 1:1000, Thermo Fisher Scientific V35003) was added. Cells were incubated at 37°C for 30 minutes. Live/dead cell stain was concurrently performed using propidium iodide (1 μ g/ μ L; Thermo Fisher Scientific P1304MP), added according to manufacturer recommendation. Sample were run on BD LSR-IITM with the BD FACSDivaTM software v9.0, using the services of the Hospital for Sick Children Flow Cytometry Facility. Data were analysed using FlowJo version 10.7.1. Median Fluorescence Intensity (MFI) peaks were compared for HAP1 and fibroblasts at G₀/G₁ and G₂/M.

Cloning

The sgRNA vector, BPK1520_puroR, a plasmid containing a cloning site for sgRNA under control of a U6 promoter, as well as a puromycin resistance cassette, was generated as previously described.³⁵ The following oligonucleotides were used to clone the sgRNA required for mutant generation: E235G, Fwd CACCGATGAACTAGTGGAGTGGGAAG, Rev AAACCTTCCACTCCACTAGTTCATC; K278E, Fwd CACCGCTTAAAAAGTTGGAGGAAT, Rev AAACATTCCTCCAACCTTTTAAAGC; P329L, Fwd CACCGCTGAGGGTGCGTTGGCATGC, Rev: AAACGCATGCCAACGCACCCTCAGC; T385M: Fwd: CACCGGGCACGCACACAAAAGTGA, Rev: AAACCTCACTTTTGTGTGCGTGCCC; D517G, Fwd: CACCGTGTGGACCAGCTGAACATGT, Rev: AAACACATGTTTCAGCTGGTCCACAC; Y701C, Fwd: CACCGTGGATATATCAAGACTGAGT, Rev: AAACACTCAGTCTTGATATATCCAC. Oligonucleotides were annealed using an annealing buffer (10 mM Tris, pH 7.5 - 8.0, 50 mM NaCl, 1 mM EDTA) and phosphorylated using T4 Polynucleotide Kinase (New England BioLabs M0201L)

according to manufacturer recommendations. BPK1520_puroR was linearized with BsmBI (New England BioLabs R0580S) and dephosphorylated using recombinant shrimp alkaline phosphatase (New England BioLabs M0371L). Annealed and phosphorylated oligonucleotides were cloned into linearized and dephosphorylated BPK1520_puroR using T4 DNA Ligase (New England BioLabs M0202L) according to manufacturer recommendations. Subsequently, One Shot™ TOP10 Chemically Competent E. coli (Thermo Fisher Scientific C404003) were transformed with the ligation products and plated on LB-Ampicillin agar plates (50 µg/mL ampicillin). Resultant colonies were inoculated overnight in LB-ampicillin, and plasmids purified using the QIAprep Spin Miniprep Kit (Qiagen 27106) according to manufacturer recommendations. The following plasmids were used for the purpose base-editing: pCMV_ABEmax_P2A_GFP (gift from Dr. David Liu, Addgene plasmid 112101) was used to generate E235G, K278E, P329L, and D517G. Plasmids pCAG-CBE4max-SpG-P2A-EGFP and pCMV-T7-ABEmax(7.10)-SpRY-P2A-EGFP (gifts from Dr. Benjamin Kleinstiver, Addgene plasmids 139998 and 140003) were used to generate T385M and Y701C, respectively.

Transfection and selection

Twenty-four hours prior to transfection, 4×10^5 cells were seeded in a 6-well plate. The following day, cells were transfected with 2500ng of total DNA, containing the Cas9 base-editor expression vector and the sgRNA expression vector containing the PuroR gene, at a 1:1 ratio (w/w). Transfection was performed using Lipofectamine™ 3000 Transfection Reagent (Thermo Fisher Scientific L3000001) according to manufacturer recommendations. To enrich for transfected cells, 24 hours post-transfection cells were subjected to puromycin selection (0.7 µg/mL; Thermo Fisher Scientific A1113803) for 72 hours. Following puromycin selection, estimation of base-editing efficiency in the bulk population was performed as follows: $\sim 1 \times 10^6$ cells were collected from which genomic DNA was isolated using the DNeasy Blood & Tissue Kit (Qiagen 69506). This was followed by PCR amplification of the desired region using DreamTaq Polymerase (Thermo Fisher Scientific EP0705).

Amplified DNA was PCR-purified using QIAquick PCR Purification Kit (Qiagen 28106) following the manufacturer's protocol and prepared for Sanger sequencing using the BigDye™ Terminator v3.1 Cycle Sequencing Kit (Thermo Fisher Scientific 4337457). Samples were sequenced on Applied Biosystems SeqStudio Genetic Analyzer (Thermo Fisher Scientific), and sequencing AB1 files were input into the online base-editing analysis tool, editR.⁷¹ This provided an estimated percentage editing of the bulk population, as well as percentage editing (if any) of any adjacent bases.

Single-cell sorting and clone screening

Following estimation of editing efficiency, cells were trypsinized and resuspended in FACS buffer (1xPBS without calcium and magnesium pH 7.4, supplemented 2% FBS and 2.5mM EDTA) at a concentration of 1×10^6 cells/mL. Live/dead cell stain was performed using propidium iodide. Live single cells were sorted on MoFloXDP Cell Sorter (Beckman Coulter), using the services of the Hospital for Sick Children Flow Cytometry Facility. Cells were sorted into a 96-well plate containing full media and allowed to clonally expand for a period of 14 days. Following a 2-week recovery period, single-cell clones underwent genomic DNA isolation, followed by PCR amplification, purification and Sanger sequencing as described above.

Immunoblotting

Immunoblotting was used to determine STAT1 and pSTAT1 protein levels in IFN γ -stimulated or unstimulated cells, over 5 independent experiments. For each experiment, 4×10^5 cells were seeded in a 6-well plate. Twenty-four hours later, Human Recombinant IFN γ (10ng/mL, StemCell Technologies 78020) was added to the media for a period of 60 minutes. Cells were subsequently washed twice with cold 1xPBS, and whole-cell lysates were obtained by lysing cells in RIPA Lysis and Extraction Buffer (Thermo Fisher Scientific 89900), supplemented with Halt™ Protease and Phosphatase Inhibitor Cocktail (Thermo Fisher Scientific 78440), on ice for 30 minutes. Lysates were sonicated and then centrifuged at 12,000xg for a period of 15 minutes at 4°C. Protein concentrations were determined using the Pierce™ BCA Protein Assay Kit (Thermo Fisher Scientific

23552). Samples were then prepared by addition of NuPAGE™ LDS Sample Buffer (4X) (Thermo Fisher Scientific NP0007) followed by boiling at 100°C for 5 minutes. Samples were subjected to SDS-Page separation by running 20µg of total protein on a NuPage 4-12% Bis-Tris gel (Thermo Fisher Scientific NP0336BOX) using NuPAGE™ MOPS SDS Running Buffer (Thermo Fisher Scientific NP000102). For each experiment, samples for pSTAT1 and STAT1 were run in parallel. Protein was subsequently transferred to a nitrocellulose membrane using the iBlot 2 Dry Blotting System (Thermo Fisher Scientific). Following transfer, membranes were blocked in 1xTris-Buffered Saline (50 mM Tris-Cl, pH 7.5. 150 mM NaCl) containing 5% bovine serum albumin (Sigma Aldrich A7906-50G) for 1 hour at room temperature. Membranes were then incubated at 4°C overnight with primary antibodies against pSTAT1 (pY701; clone D4A7; Cell Signaling 7649S), total STAT1 (Clone D1K9Y, Cell Signaling 14994) or alpha-tubulin (Clone DM1A; Sigma Aldrich T6199-100UL). The following day, membranes were washed in tris-buffered saline and incubated for 1 hour at room temperature with one of the following secondary antibodies: Donkey anti-Rabbit IgG (H+L) Highly Cross-Adsorbed Secondary Antibody, Alexa Fluor 647 (Thermo Fisher Scientific A-31573) or Donkey anti-Mouse IgG (H+L) Highly Cross-Adsorbed Secondary Antibody, Alexa Fluor 647 (Thermo Fisher Scientific A-31571). Membranes were imaged using ChemiDoc MP imaging system (Bio-Rad) and analyzed with Image Lab software (©2017 Bio-Rad Laboratories; version 6.0.1).

RNA isolation and Quantitative real-time PCR

RNA analysis was performed to determine gene expression levels across the different genotypes under various conditions. Experiments were repeated a minimum of 5 times for each gene, in at least technical duplicates, and pooled data for each gene was collected. To determine baseline gene expression levels, cells were grown in full media and harvested without stimulation once reaching 70-80% confluence. For determination of baseline gene expression under low-serum conditions, cells were seeded as described above and allowed to adhere in full media for a period of 24 hours. Cells were subsequently thoroughly washed with PBS and placed in low-serum media (IMDM+0.25%

494 FBS) for an additional 24 hours. For stimulation experiments, cells in full media were incubated with
 495 Human Recombinant IFN α -2A (10ng/mL, StemCell Technologies 78076.1) or IFN γ (10ng/mL) for 6
 496 hours prior to harvesting and RNA extraction using RNeasy Mini Kit (Qiagen 74106). Next, 1000ng
 497 of RNA was reverse-transcribed using SuperScript™ III First-Strand Synthesis System (Thermo
 498 Fisher Scientific 18080051) following the manufacturer's protocol. Quantitative real-time PCR (qRT-
 499 PCR) using PowerUp™ SYBR™ Green Master Mix (Thermo Fisher Scientific A25742) was
 500 performed on an Applied Biosystems QuantStudio 3 Real-Time PCR System (Applied Biosystems).
 501 Quantification of the following genes was done: *GBP1*, *IFIT2*, *IRF1*, *APOL6*, *OAS1*, with *GAPDH*
 502 used as housekeeping control. The following primers were used, designed using Primer3 (v.0.4.0):
 503 *GBP1* Fwd AGGAGTTAGCGGCCAGCTAGAAA, Rev
 504 AAAATGACCTGAAGTAAAGCTGAGC; *IFIT2* Fwd GCACTGCAACCATGAGTGAGA, Rev
 505 CAAGTTCCAGGTGAAATGGCA; *IRF1* Fwd TCCTGCAGCAGAGCCAACATGCCCA, Rev
 506 CCGGGATTTGGTTGGAATTAATCTG; *APOL6* Fwd TTGGTTTGCAAAGGGATGAGGATGA,
 507 Rev TCTTTCAATCTGGGAAATTCTCTCA; *OAS1* Fwd
 508 CAAGGTGGTAAAGGGTGGCTCCTCA, Rev TAACTGATCCTGAAAAGTGGTGAGA; *GAPDH*
 509 Fwd CAATGACCCCTTCATTGACCTC, Rev GATCTCGCTCCTGGAAGATG. The relative
 510 expression levels were compared using the $\Delta\Delta C_t$ method.

Statistical analysis

512 Graphical data were represented as means $\pm$ standard error of mean. Statistical analysis was
 513 performed using GraphPad Prism 8 (version 8.4.3). One-way ANOVA with Dunnett's post-hoc test
 514 was used to determine differences among mutants and wildtype. Statistical significance was
 515 represented as: *p < 0.05, **p < 0.01, ***p < 0.001, ****p < 0.0001.

516

Data availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Acknowledgments

Funding this work was provided by Immunodeficiency Canada (OS). Salary support for OS has been provided by the Ontario Ministry of Health Clinician Investigator Program, the Hospital for Sick Children Clinician Scientist Training Program, and the Canadian Child Health Clinician Scientist Program. Figures 1 and 3 were created with [Biorender.com](https://biorender.com).

Author Contributions

Study conception and design were done by OS, CMR, EAI, RDC. Experimental data acquisition was performed by OS, KL. Data analysis was performed by OS, KL, SE. Data interpretation was done by OS, SE, CMR, EAI, RDC. Writing of first draft was performed by OS. Further writing, reviewing and editing was by OS, KL, SE, CMR, EAI, RDC. The work was jointly supervised by EAI and RDC. All authors have approved the submitted version and have agreed both to be personally accountable for their own contributions and to ensure that questions related to the accuracy or integrity of any part of the work, even ones in which the author was not personally involved, are appropriately investigated, resolved, and the resolution documented in the literature.

Competing Interests

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

1. Darnell, J. E., Kerr, I. M., & Stark, G. R. (1994). Jak-STAT pathways and transcriptional activation in response to IFNs and other extracellular signaling proteins. *Science*, 264(5164), 1415-1421.
2. Dale, T. C., Imam, A. M., Kerr, I. M., & Stark, G. R. (1989). Rapid activation by interferon alpha of a latent DNA-binding protein present in the cytoplasm of untreated cells. *Proceedings of the National Academy of Sciences*, 86(4), 1203-1207.
3. Stark, G. R., Kerr, I. M., Williams, B. R., Silverman, R. H., & Schreiber, R. D. (1998). How cells respond to interferons.
4. Levy, D. E., Kessler, D. S., Pine, R., & Darnell, J. E. (1989). Cytoplasmic activation of ISGF3, the positive regulator of interferon-alpha-stimulated transcription, reconstituted in vitro. *Genes & development*, 3(9), 1362-1371.
5. Schindler, C., Levy, D. E., & Decker, T. (2007). JAK-STAT signaling: from interferons to cytokines. *Journal of Biological Chemistry*, 282(28), 20059-20063.
6. Michalska, A., Blaszczyk, K., Wesoly, J., & Bluysen, H. A. (2018). A positive feedback amplifier circuit that regulates interferon (IFN)-stimulated gene expression and controls type I and type II IFN responses. *Frontiers in immunology*, 9, 1135.
7. Dupuis, S., et al. (2003). Impaired response to interferon- α/β and lethal viral disease in human STAT1 deficiency. *Nature genetics*, 33(3), 388-391.
8. Chapgier, A., et al. (2006). 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. *The Journal of Immunology*, 176(8), 5078-5083.

9. Vairo, D., et al. (2011). Severe impairment of IFN- γ and IFN- α responses in cells of a patient with a novel STAT1 splicing mutation. *Blood, The Journal of the American Society of Hematology*, 118(7), 1806-1817.
10. Chapgier, A., et al. (2009). A partial form of recessive STAT1 deficiency in humans. *The Journal of clinical investigation*, 119(6), 1502-1514.
11. Kong, X. F., et al. (2010). A novel form of human STAT1 deficiency impairing early but not late responses to interferons. *Blood*, 116(26), 5895-5906.
12. Kristensen, I. A., Veirum, J. E., Møller, B. K., & Christiansen, M. (2011). Novel STAT1 alleles in a patient with impaired resistance to mycobacteria. *Journal of clinical immunology*, 31(2), 265-271.
13. van de Veerdonk, et al. (2011). STAT1 mutations in autosomal dominant chronic mucocutaneous candidiasis. *New England Journal of Medicine*, 365(1), 54-61.
14. Liu, L., et al. (2011). Gain-of-function human STAT1 mutations impair IL-17 immunity and underlie chronic mucocutaneous candidiasis. *The Journal of experimental medicine*, 208(8), 1635-1648.
15. Toubiana, J., Okada, S., Hiller, J., Oleastro, M., Lagos Gomez, M., Aldave Becerra, J. C., ... & Van Montfrans, J. (2016). Heterozygous STAT1 gain-of-function mutations underlie an unexpectedly broad clinical phenotype. *Blood*, 127(25), 3154-3164.
16. Depner, M., et al. (2016). The extended clinical phenotype of 26 patients with chronic mucocutaneous candidiasis due to gain-of-function mutations in STAT1. *Journal of clinical immunology*, 36(1), 73-84.
17. McGurk, M., & Holmes, M. (1988). Chronic muco-cutaneous candidiasis and oral neoplasia. *The Journal of Laryngology & Otology*, 102(7), 643-645.

18. Marazzi, M. G., Bondi, E., Giannattasio, A., Strozzi, M., & Savioli, C. (2008). Intracranial aneurysm associated with chronic mucocutaneous candidiasis. *European journal of pediatrics*, 167(4), 461.
19. Koo, S., Kejariwal, D., Al-Shehri, T., Dhar, A., & Lilic, D. (2017). Oesophageal candidiasis and squamous cell cancer in patients with gain-of-function STAT1 gene mutation. *United European Gastroenterology Journal*, 5(5), 625-631.
20. Okada, S., Puel, A., Casanova, J. L., & Kobayashi, M. (2016). Chronic mucocutaneous candidiasis disease associated with inborn errors of IL-17 immunity. *Clinical & Translational Immunology*, 5(12), e114.
21. Tamaura, M., et al. (2020). Human gain-of-function STAT1 mutation disturbs IL-17 immunity in mice. *International Immunology*, 32(4), 259-272.
22. Barrat, F. J., Crow, M. K., & Ivashkiv, L. B. (2019). Interferon target-gene expression and epigenomic signatures in health and disease. *Nature immunology*, 20(12), 1574-1583.
23. Reich, N. C. (2019, June). Too much of a good thing: Detrimental effects of interferon. In *Seminars in immunology* (Vol. 43, p. 101282). Academic Press.
24. Wang, B. X., & Fish, E. N. (2012). The yin and yang of viruses and interferons. *Trends in immunology*, 33(4), 190-197.
25. Kaleviste, E., et al. (2019). Interferon signature in patients with STAT1 gain-of-function mutation is epigenetically determined. *European journal of immunology*, 49(5), 790-800.
26. Okada, S., et al. (2020). Human STAT1 Gain-of-Function Heterozygous Mutations: Chronic Mucocutaneous Candidiasis and Type I Interferonopathy. *Journal of Clinical Immunology*, 1-17.
27. Leiding, J. W., et al. (2018). Hematopoietic stem cell transplantation in patients with gain-of-function signal transducer and activator of transcription 1 mutations. *Journal of Allergy and Clinical Immunology*, 141(2), 704-717.

28. Fujiki, R., et al. (2017). Molecular mechanism and structural basis of gain-of-function of STAT1 caused by pathogenic R274Q mutation. *Journal of Biological Chemistry*, 292(15), 6240-6254.
29. Ovadia, A., Sharfe, N., Hawkins, C., Laughlin, S., & Roifman, C. M. (2018). Two different STAT1 gain-of-function mutations lead to diverse IFN- γ -mediated gene expression. *NPJ genomic medicine*, 3(1), 1-5.
30. Kagawa, R., et al. (2017). Alanine-scanning mutagenesis of human signal transducer and activator of transcription 1 to estimate loss-or gain-of-function variants. *Journal of Allergy and Clinical Immunology*, 140(1), 232-241.
31. Giovannozzi, S., Lemmens, V., Hendrix, J., Gijssbers, R., & Schrijvers, R. (2020). Live Cell Imaging Demonstrates Multiple Routes Toward a STAT1 Gain-of-Function Phenotype. *Frontiers in immunology*, 11, 1114.
32. Bando, H., et al. (2018). Toxoplasma effector TgIST targets host IDO1 to antagonize the IFN- γ -induced anti-parasitic response in human cells. *Frontiers in immunology*, 9, 2073.
33. Ohshima, J., et al. (2014). Role of mouse and human autophagy proteins in IFN- γ -induced cell-autonomous responses against *Toxoplasma gondii*. *The Journal of Immunology*, 192(7), 3328-3335.
34. Gapp, B. V., et al. (2016). Parallel reverse genetic screening in mutant human cells using transcriptomics. *Molecular systems biology*, 12(8), 879.
35. Erwood, S., et al. (2019). Modeling Niemann–Pick disease type C in a human haploid cell line allows for patient variant characterization and clinical interpretation. *Genome Research*, 29(12), 2010-2019.
36. Olbrich, T., et al. (2019). A chemical screen identifies compounds capable of selecting for haploidy in mammalian cells. *Cell reports*, 28(3), 597-604.

37. Romberg, N., et al. (2013). Gain-of-function STAT1 mutations are associated with PD-L1 overexpression and a defect in B-cell survival. *Journal of Allergy and Clinical Immunology*, 131(6), 1691-1693.
38. Yamazaki, Y., et al. (2014). Two novel gain-of-function mutations of STAT1 responsible for chronic mucocutaneous candidiasis disease: impaired production of IL-17A and IL-22, and the presence of anti-IL-17F autoantibody. *The Journal of Immunology*, 193(10), 4880-4887.
39. Mizoguchi, Y., et al. (2014). Simple diagnosis of STAT1 gain-of-function alleles in patients with chronic mucocutaneous candidiasis. *Journal of leukocyte biology*, 95(4), 667-676.
40. Vargas-Hernández, A., et al. (2018). Ruxolitinib partially reverses functional natural killer cell deficiency in patients with signal transducer and activator of transcription 1 (STAT1) gain-of-function mutations. *Journal of Allergy and Clinical Immunology*, 141(6), 2142-2155.
41. Baris, S., et al. (2016). Severe early-onset combined immunodeficiency due to heterozygous gain-of-function mutations in STAT1. *Journal of clinical immunology*, 36(7), 641-648.
42. Chen, X., et al. (2019). Molecular and Phenotypic Characterization of Nine Patients with STAT1 GOF Mutations in China. *Journal of Clinical Immunology*, 1-14.
43. Eren Akarcan, S., et al. (2017). Gain-of-function mutations in STAT1: a recently defined cause for chronic mucocutaneous candidiasis disease mimicking combined immunodeficiencies. *Case reports in immunology*, 2017.
44. Pedraza-Sánchez, S., et al. (2017). Disseminated tuberculosis and chronic mucocutaneous candidiasis in a patient with a gain-of-function mutation in signal transduction and activator of transcription 1. *Frontiers in Immunology*, 8, 1651.
45. Sampaio, E. P., et al. (2013). Signal transducer and activator of transcription 1 (STAT1) gain-of-function mutations and disseminated coccidioidomycosis and histoplasmosis. *Journal of allergy and clinical immunology*, 131(6), 1624-1634.

46. Sharfe, N., et al. (2014). Fatal combined immunodeficiency associated with heterozygous mutation in STAT1. *Journal of Allergy and Clinical Immunology*, 133(3), 807-817.
47. Soltész, B., et al. (2013). New and recurrent gain-of-function STAT1 mutations in patients with chronic mucocutaneous candidiasis from Eastern and Central Europe. *Journal of medical genetics*, 50(9), 567-578.
48. Uzel, G., et al. (2013). Dominant gain-of-function STAT1 mutations in FOXP3 wild-type immune dysregulation–polyendocrinopathy–enteropathy–X-linked–like syndrome. *Journal of allergy and clinical immunology*, 131(6), 1611-1623.
49. Zerbe, C. S., et al. (2016). Progressive multifocal leukoencephalopathy in primary immune deficiencies: Stat1 gain of function and review of the literature. *Clinical Infectious Diseases*, 62(8), 986-994.
50. Hirata, O., et al. (2013). Heterozygosity for the Y701C STAT1 mutation in a multiplex kindred with multifocal osteomyelitis. *Haematologica*, 98(10), 1641-1649.
51. Rusinova, I., et al. (2012). Interferome v2. 0: an updated database of annotated interferon-regulated genes. *Nucleic acids research*, 41(D1), D1040-D1046.
52. Fielhaber, J. A., et al. (2012). Regulation of karyopherin α 1 and nuclear import by mammalian target of rapamycin. *Journal of Biological Chemistry*, 287(18), 14325-14335.
53. Komor, A. C., Kim, Y. B., Packer, M. S., Zuris, J. A., & Liu, D. R. (2016). Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. *Nature*, 533(7603), 420-424.
54. Gaudelli, N. M., et al. (2017). Programmable base editing of A•T to G•C in genomic DNA without DNA cleavage. *Nature*, 551(7681), 464-471.
55. Rees, H. A., & Liu, D. R. (2018). Base editing: precision chemistry on the genome and transcriptome of living cells. *Nature reviews genetics*, 19(12), 770-788.

56. Walton, R. T., Christie, K. A., Whittaker, M. N., & Kleinstiver, B. P. (2020). Unconstrained genome targeting with near-PAMless engineered CRISPR-Cas9 variants. *Science*, 368(6488), 290-296.
57. Molla, K. A., Qi, Y., Karmakar, S., & Baig, M. J. (2020). Base Editing Landscape Extends to Perform Transversion Mutation. *Trends in Genetics*.
58. Zimmerman, O., et al. (2019). STAT1 gain-of-function mutations cause high total STAT1 levels with normal dephosphorylation. *Frontiers in immunology*, 10, 1433.
59. Bernasconi, A. R., et al. (2018). Increased STAT1 amounts correlate with the phospho-STAT1 level in STAT1 gain-of-function defects. *Journal of Clinical Immunology*, 38(7), 745-747.
60. Chatterjee Kishore, M., Wright, K. L., Ting, J. P. Y., & Stark, G. R. (2000). How Stat1 mediates constitutive gene expression: a complex of unphosphorylated Stat1 and IRF1 supports transcription of the LMP2 gene. *The EMBO journal*, 19(15), 4111-4122.
61. Cheon, H., & Stark, G. R. (2009). Unphosphorylated STAT1 prolongs the expression of interferon-induced immune regulatory genes. *Proceedings of the National Academy of Sciences*, 106(23), 9373-9378.
62. Majoros, A., et al. (2016). Response to interferons and antibacterial innate immunity in the absence of tyrosine-phosphorylated STAT 1. *EMBO reports*, 17(3), 367-382.
63. Meyer, T., Begitt, A., Lödige, I., van Rossum, M., & Vinkemeier, U. (2002). Constitutive and IFN- γ -induced nuclear import of STAT1 proceed through independent pathways. *The EMBO journal*, 21(3), 344-354.
64. Wang, W., et al. (2017). Unphosphorylated ISGF3 drives constitutive expression of interferon-stimulated genes to protect against viral infections. *Science Signaling*, 10(476).

65. Kobbe, R., et al. (2016). Common variable immunodeficiency, impaired neurological development and reduced numbers of T regulatory cells in a 10-year-old boy with a STAT1 gain-of-function mutation. *Gene*, 586(2), 234-238.
66. Meesilpavikkai, K., et al. (2017). A novel heterozygous mutation in the STAT1 SH2 domain causes chronic mucocutaneous candidiasis, atypically diverse infections, autoimmunity, and impaired cytokine regulation. *Frontiers in immunology*, 8, 274.
67. Zheng, J., et al. (2015). Gain of function STAT1 mutations impair STAT3 activity in patients with chronic mucocutaneous candidiasis (CMC). *European journal of immunology*, 45(10), 2834-2846.
68. Zhang, Y., et al. 2017). PD-L1 up-regulation restrains Th17 cell differentiation in STAT3 loss-and STAT1 gain-of-function patients. *Journal of Experimental Medicine*, 214(9), 2523-2533.
69. Qian, W., Miner, C. A., Ingle, H., Platt, D. J., Baldrige, M. T., & Miner, J. J. (2019). A human STAT1 gain-of-function mutation impairs CD8+ T cell responses against gammaherpesvirus 68. *Journal of virology*, 93(19), e00307-19.
70. Darzynkiewicz, Z., Halicka, H. D., & Zhao, H. (2010). Analysis of cellular DNA content by flow and laser scanning cytometry. In *Polyploidization and cancer* (pp. 137-147). Springer, New York, NY.
71. Kluesner, M. G., Nedveck, D. A., Lahr, W. S., Garbe, J. R., Abrahante, J. E., Webber, B. R., & Moriarity, B. S. (2018). EditR: a method to quantify base editing from Sanger sequencing. *The CRISPR journal*, 1(3), 239-250.

Figure Legends

Figure 1. *STAT1* mutant generation. | (a) Schematic representation of *STAT1* mutations selected for modelling and their respectively-affected protein domains. Mutations designated in blue (E235G, K278E, O329L, T385M, D517G) are “gain-of-function” mutations, while orange (Y701C) designates a “loss-of-function” mutation. (b) Overview of the process of heterozygous cell model generation using CRISPR/Cas9 base editing. (c) Sanger sequencing confirmation of generated cell models, compared with their respective wildtype counterparts. Highlighted nucleotides denote the edited base. Note that for E235G, editing of an adjacent “A” on both alleles took place, resulting in a silent bystander mutation. This yielded the trinucleotide change: GAA/GAA → GAG/GGG, leading to the amino acid substitution: Glutamine/Glutamine → Glutamine/Glycine. This corresponds with the same amino acid change found in patients affected by the E235G mutation.

Figure 2. Immunoblot analysis of pSTAT1 (Tyr701) and total STAT1 levels among *STAT1* mutants following IFN γ stimulation. | Levels of pSTAT1 (Tyr701; a) and total STAT1 (b) were measured in whole cell protein lysates at baseline and following a 60-minute stimulation with IFN γ (10ng/mL), normalized to a loading control (β -Tubulin). Densitometry analysis results from 5 independent experiments were plotted for pSTAT1 (c) and STAT1 levels following IFN γ stimulation (d) and compared with WT. Results are presented as mean $\pm$ standard error of mean. Statistical analysis: One-way ANOVA with Dunnett’s post-hoc test (* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$).

Figure 3. *STAT1*-associated gene transcription under various experimental conditions. | (a) at baseline, constitutive Interferon-Stimulated Gene (ISG) expression is maintained at a basal level in an interferon (IFN)-independent manner via two unphosphorylated transcription complexes: unphosphorylated gamma activating factor (U-GAF) made up of two *STAT1* units, and unphosphorylated interferon-stimulated gene factor 3 (U-ISGF3), made up of *STAT1*, *STAT2* and IRF9. At baseline, mammalian target of rapamycin (mTOR) serves as an inhibitor of *STAT1*

migration to the nucleus, preventing excessive ISG transcription; (b) following nutrient depletion, mTOR is inhibited, allowing for higher nuclear accumulation unphosphorylated-STAT1 containing transcriptional complexes. This results in an ISG transcriptional response which is IFN-independent; (c) Stimulation with IFN α begins with ligation of the type I IFN receptor, made up of two subunits: IFNAR1 and IFNAR2. Receptor ligation leads to recruitment of the tyrosine kinases TYK2 and JAK1, which cross-phosphorylate each other as well as the intracellular receptor domains. This creates a docking site for STAT1 and STAT2 which bind, are phosphorylated, and along with IRF9 form the transcription complex ISGF3. ISGF3 migrates to the nucleus where it binds to interferon-stimulated response element (ISRE) in gene promoters, resulting in a Type I IFN response; (d) Stimulation with IFN γ begins with ligation of the type II IFN receptor, made up of two subunits each of IFNGR1 and IFNGR2. Receptor ligation leads to recruitment of the tyrosine kinases JAK1 and JAK2, which cross-phosphorylate each other as well as the intracellular receptor domains. This creates a docking site for STAT1 which bind, are phosphorylated, and form the homodimer GAF. GAF migrates to the nucleus where it binds to gamma-activating sequence (GAS) in gene promoters, resulting in a Type II IFN response.

Figure 4. Relative gene expression of Interferon-Stimulated Genes (ISG) at baseline and following serum starvation. | (a) mRNA expression levels were measured for five ISG (*GBP1*, *IFIT2*, *IRF1*, *APOL6*, *OAS1*) at baseline in cells grown under normal serum conditions; (b) mRNA expression levels for five ISG measured in cells following 24-hours of serum starvation. Expression levels were normalized to housekeeping *GAPDH* expression and plotted relatively to wildtype for each experiment. Pooled data from at least 5 experiments are presented. Data are represented as mean $\pm$ standard error of mean. Statistical analysis: One-way ANOVA with Dunnett's post-hoc test (* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$)

781 **Figure 5.** Relative gene expression of Interferon-Stimulated Genes (ISG) following IFN α or IFN γ
782 stimulation. | mRNA expression levels were measured (*GBP1*, *IFIT2*, *IRF1*, *APOL6*, *OAS1*) at
783 baseline and following 6 hours of stimulation with IFN α (a) or IFN γ (b). Expression levels were
784 normalized to housekeeping *GAPDH* and plotted as fold increase from baseline. Pooled data from at
785 least 5 experiments are presented. Data are represented as mean $\pm$ standard error of mean. Statistical
786 analysis: One-way ANOVA with Dunnett's post-hoc test (* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, ****
787 $p < 0.0001$)

Table I: Clinical manifestations of *STAT1* mutations modelled in this work

Mutation	Designation	Fungal infections	Bacterial Infections	Viral Infections	Mycobacterial infections	Autoimmunity / Endocrinopathy	Vascular abnormalities	Malignancy	Other
E235G (c.704A>G) ³⁷	AD GOF	CMCC	Sinopulmonary infections, bacterial skin infections with abscess formation	Recurrent HSV labialis	Mycobacterial pneumonia	Alopecia	Carotid and celiac/ splenic artery dissection	HPV+ squamous cell carcinoma, basal cell carcinoma	N.D.
K278E (c.832A>G) ³⁸	AD GOF	CMCC	Bacterial skin infections	Recurrent HSV stomatitis, recurrent Herpes zoster	N.D.	Positive auto-antibodies (ANA, microsomal, anti-TSH, anti-IL-17F)	N.D.	N.D.	N.D.
P329L (c.986T>C) ^{39,40}	AD GOF	CMCC	Yes, no details provided	Recurrent HSV stomatitis	N.D.	Autoimmune hemolytic anemia, red blood cell aplasia	N.D.	N.D.	N.D.
T385M (c.1154C>T) ^{16,41-49}	AD GOF	CMCC, invasive fungal infections	Sinopulmonary infections	EBV, CMV, recurrent severe VZV	Variety of Mycobacterial infections reported	Enteropathy, T1DM, thyroiditis, autoimmune cytopenias, vitiligo, growth hormone deficiency, autoimmune hepatitis	Intracranial aneurysms	N.D.	Recurrent fractures, eczema, chronic adenopathy, multifocal leuko-encephalopathy
D517G (c.1550A>G) ¹⁵	AD GOF	CMCC, invasive fungal infections	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.
Y701C (c.2102A>G) ⁵⁰	AD LOF	N.D.	N.D.	N.D.	Invasive Mycobacterial infections	N.D.	N.D.	N.D.	Multifocal osteomyelitis

AD: Autosomal dominant; ANA: anti-nuclear antibody; CMCC: chronic mucocutaneous Candidiasis; CMV: Cytomegalovirus; EBV: Epstein-Barr virus; GOF: gain-of-function; HPV: Human papillomavirus; HSV: Human Herpesvirus; LOF: loss-of-function; N.D.: not described; T1DM: Type 1 diabetes mellitus; TSH: thyroid stimulating hormone; VZV: Varicella zoster virus

Table II: Summary of base-editing features for *STAT1* mutation generation

Mutation	Desired trinucleotide change	Single-guide RNA sequence*	Base-editor used	Protospacer adjacent motif (PAM) sequence	On-target base-editing (%)	Off-target changes within the editing window (%)	Number of clones screened to obtain the desired mutation in a heterozygous state
E235G (c.704A>G)	GAA>GGA	gATGAACTAGTGGAGTGGAAG	SpCas9 ABEmax	CGG	50	38	39**
K278E (c.832A>G)	AAA>GAA	GCTTAAAAAGTTGGAGGAAT	SpCas9 ABEmax	TGG	67	7	17
P329L (c.986T>C)	CCT>CCC (Editing on opposite strand: AGG>GGG)	gCTGAGGGTGC GTTGGCATGC	SpCas9 ABEmax	AGG	20	1	13
T385M (c.1154C>T)	ACG>ATG	GGGCACGCACACAAAAGTGA	SpG CBE4max	TGA	21	4	14
D517G (c.1550A>G)	GAC>GGC	gTGTGGACCAGCTGAACATGT	SpCas9 ABEmax	TGG	77	2	12
Y701C (c.2102A>G)	TAT>TGT	gTGGATATATCAAGACTGAGT	SpRY ABEmax	TGA	36	15	24

*Non-template “g” appended to single-guide RNA to facilitate transcription downstream of the U6 promoter designated in lowercase

**Resultant clone contains a silent bystander nucleotide change within the editing window

Table III: Summary of differences in relative expression or fold increase across *STAT1* mutants compared with wildtype under various conditions

Mutation	Condition/Stimulation	GBP1	IFIT2	IRF1	APOL6	OAS1
E235G	Baseline (full serum) *	↑	NS	NS	↑	NS
	Serum starvation **	↑	NS	NS	↑	NS
	IFN α †	NS	↑	↑	NS	NS
	IFN γ ‡	↓	NS	NS	↓	↑
K278E	Baseline (full serum) *	NS	NS	NS	NS	NS
	Serum starvation **	NS	NS	NS	NS	NS
	IFN α †	↓	NS	NS	↓	↓
	IFN γ ‡	NS	↑	NS	NS	NS
P329L	Baseline (full serum) *	NS	NS	NS	↑	↑
	Serum starvation **	NS	NS	↑	↑	↑
	IFN α †	↓	NS	NS	↓	↓
	IFN γ ‡	↓	NS	NS	↓	NS
T385M	Baseline (full serum) *	NS	NS	NS	NS	NS
	Serum starvation **	NS	NS	NS	NS	NS
	IFN α †	↑	↑	↑	↑	↑
	IFN γ ‡	NS	NS	NS	↓	NS
D517G	Baseline (full serum) *	NS	↑	NS	NS	NS
	Serum starvation **	NS	NS	NS	NS	NS
	IFN α †	NS	NS	NS	↓	↓
	IFN γ ‡	NS	NS	NS	↓	NS
Y701C	Baseline (full serum) *	NS	NS	NS	NS	↓
	Serum starvation **	NS	NS	NS	NS	NS
	IFN α †	↓	↓	↓	↓	NS
	IFN γ ‡	↓	↓	↓	↓	NS

* Relative gene expression compared with wildtype under full-serum culture conditions without cytokine stimulation

** Relative gene expression compared with wildtype following 24 hours of low-serum conditions without cytokine stimulation

† Fold increase in gene expression from baseline compared with wildtype following 6 hours of stimulation with IFN α (10ng/mL)‡ Fold increase in gene expression from baseline compared with wildtype following 6 hours of stimulation with IFN γ (10ng/mL)

NS: not significant

a

**No Stimulation - Baseline
Constitutive Expression**

b

**Nutrient Depletion - Stress
Response**

c

**IFN α Stimulation - Type I IFN
Response**

d

**IFN γ Stimulation - Type II IFN
Response**