Dear Author

Please use this PDF proof to check the layout of your article. If you would like any changes to be made to the layout, you can leave instructions in the online proofing interface. First, return to the online proofing interface by clicking "Edit" at the top page, then insert a Comment in the relevant location. Making your changes directly in the online proofing interface is the quickest, easiest way to correct and submit your proof.

Please note that changes made to the article in the online proofing interface will be added to the article before publication, but are not reflected in this PDF proof.

If you would prefer to submit your corrections by annotating the PDF proof, please go to the last page and click the link.
Human STAT3 variants underlie autosomal dominant hyper-IgE syndrome by negative dominance

Most patients with autosomal dominant hyper-IgE syndrome (AD-HIES) carry rare heterozygous STAT3 variants. Only six of the 123 in-frame variants reported have been experimentally shown to be dominant negative (DN), and it has been recently suggested that eight out-of-frame variants operate by haploinsufficiency. We experimentally tested these 143 variants, 7 novel out-of-frame variants found in HIES patients, and other STAT3 variants from the general population. Strikingly, all 15 out-of-frame variants were DN via their encoded (1) truncated proteins, (2) neoproteins generated from a translation reinitiation codon, and (3) isoforms from alternative transcripts or a combination thereof. Moreover, 128 of the 135 in-frame variants (95%) were also DN. The patients carrying the seven non-DN STAT3 in-frame variants have not been studied for other genetic etiologies. Finally, none of the variants from the general population tested, including an out-of-frame variant, were DN. Overall, our findings show that heterozygous STAT3 variants, whether in or out of frame, underlie AD-HIES through negative dominance rather than haploinsufficiency.
birth (Ghaffari, 2018; Mogensen, 2013). However, its actual prevalence may be higher due to the incomplete penetrance of most of its clinical manifestations. HIES is typically inherited as an autosomal dominant (AD) trait (Grimbacher et al., 1999). The first genetic etiology to be identified, monoallelic dominant-negative (DN) missense variants of signal transducer and activator of transcription 3 (STAT3), was reported in 2007 and accounts for >90% of sporadic and familial cases (Minegishi et al., 2007). Three other genetic etiologies have since been reported; all three are almost undistinguishable phenocopies of AD STAT3 deficiency: (1) autosomal recessive (AR) ZNF341 deficiency (ZNF341 being a transcription factor required for the expression and activity of STAT3; Béziat et al., 2018; Frey-Jakobs et al., 2018), (2) AR partial deficiency of the common receptor chain gp130 (encoded by IL6ST; Schwerd et al., 2017), and (3) DN IL6ST variants in other kindreds with a typical form of AD-HIES (Béziat et al., 2020). The unifying mechanism underlying most clinical features is the disruption of IL-6- and IL-11–dependent STAT3 activation. Patients with AR complete IL-6R deficiency (Spencer et al., 2019) have a closely related phenotype, without extrahematopoietic manifestations, whereas patients with AR IL-1R deficiency (Keupp et al., 2013) lack the hematological manifestations.

By July 2020, 143 heterozygous variants of STAT3 had been reported to underlie HIES (Abolhassani et al., 2018; Al Khatib et al., 2009; Alcántara-Montiel et al., 2016; Anolik et al., 2009; Avery et al., 2010; Chandesris et al., 2012; Egawa et al., 2019; Eken et al., 2020; Felgentreff et al., 2014; Freeman et al., 2013; Giacomelli et al., 2011; Hagl et al., 2016; He et al., 2012; Heimall et al., 2011; Holland et al., 2007; Jiao et al., 2008; Khourieh et al., 2019; Kim et al., 2009; Kumáновics et al., 2010; Larsen et al., 2019; Ma et al., 2012; Ma et al., 2008; Merli et al., 2014; Minegishi et al., 2009; Minegishi et al., 2007; Miyazaki et al., 2011; Moens et al., 2017; Moens et al., 2014; Mogensen, 2013; Natarajan et al., 2018; Papamatasiou et al., 2010; Pelham et al., 2016; Powers et al., 2009; Renner et al., 2008; Renner et al., 2007; Robinson et al., 2011b; Schimke et al., 2010; Sundin et al., 2014; van de Veerdonk et al., 2010; Vinh et al., 2010; Woellner et al., 2010; Wolach et al., 2014; Wu et al., 2017; Zhang et al., 2013). These 143 variants include 114 (79%) missense mutations, 21 (15%) in-frame indels, and 8 (6%) nonsense or frameshift indels. Only 6 variants, all in-frame, were experimentally shown to be both loss of function (LOF) and DN of STAT3 protein in the patient’s T cells was half that in healthy donors (Natarajan et al., 2018). This report challenged the long-standing paradigm that heterozygous STAT3 variants cause HIES by negative dominance. However, negative dominance was not experimentally excluded (Natarajan et al., 2018). Overall, there is suggestive but inconclusive evidence to suggest that (1) heterozygous loss-of-expression variants of STAT3 can cause HIES and (2) the autosomal dominance underlying HIES can operate both by negative dominance and by HI mechanisms. We addressed both these issues by experimentally testing all known STAT3 variants found in HIES patients, including known and novel variants predicted to be loss of expression (pLOF), and selected variants found in the general population.

Results

Seven families with AD-HIES and STAT3 nonsense or frameshift variants pLOF

In our in-house cohort of patients with HIES, we identified seven new individuals from five families with previously unknown nonsense or frameshift variants of STAT3 (P1–P7; Fig. 1A). These seven patients (P1, P2, P3, P4, P5, P6, and P7) were diagnosed with classical HIES on the basis of signs such as high serum IgE levels, recurrent skin abscesses, and recurrent pneumonia (Table S1; see case reports in Materials and methods). All of these patients carried private heterozygous nonsense or frameshift variants of STAT3 [P1: c.373C>T (p.Q125*), P2: c.1111_1112del(p.D371Lfs*14), P3: c.1552C>T (p.R518*), P4 and P5: c.1685G>A (p.W562*), P6 and P7: c.1897G>T (p.Q633*)]. They had no biallelic rare variants of DOCK8, PGM3, IL6ST, IL6R, and ZNF341 or monoallelic rare variants of CARD11, ERBBIP2, IL6ST, or TGFBR1/2, which underlie HIES, as broadly defined in the 2019 classification of the International Union of Immunological Societies Expert Committee (August, 2018; Béziat et al., 2018; Béziat et al., 2020; Bousfiha et al., 2020; Engelhardt et al., 2009; Felgentreff et al., 2014; Fleisher, 2014; Frey-Jakobs et al., 2018; Lyons et al., 2017; Ma et al., 2017; Sassi et al., 2014; Schwerd et al., 2017; Spencer et al., 2019; Su, 2010; Tange et al., 2020; Zhang et al., 2010b; Zhang et al., 2009; Zhang et al., 2014). These five new nonsense and frameshift STAT3 variants (Fig. 1A) were not found in the Human Gene Mutation Database of pathological variants, the 1,000 Genomes Project, or Genome Aggregation Database (gnomAD). The screening of our whole cohort (>8,000 individuals with severe or recurrent infections) led to the identification of another three individuals (P8, P9, and P10) from two kindreds with two previously unknown frameshift variants [P8 and P9: c.2091delT (p.D698Tfs*9), P10: c.2093ins22 (p.D698Efs*6); Fig. 1A]. P8, P9, and P10 were diagnosed with nonclassical HIES due to their atypical clinical manifestations, with only slightly high serum IgE levels and without CMC or connective tissue disorders, resulting in lower National Institutes of Health (NIH) scores than for P1–P7 (Table S1). They also suffered from tuberculosis (TB), which has rarely been reported in patients with HIES (Ashtekar and Shah, 2016; Grimbacher et al., 2005; Metin et al., 2004). We analyzed the predicted deleteriousness of all variants (P1–P10). All their combined annotation-dependent depletion (CADD) scores were above the 99% confidence threshold for the mutation
significance cutoff (MSC) for STAT3 (16.4; Fig. 1 B; Itan et al., 2016; Rentzsch et al., 2019; Zhang et al., 2018a). These results also suggest that the nonsense or frameshift variants of STAT3 predicted to be LOF may be responsible for AD-HIES.

13 out-of-frame STAT3 variants are DN because they encode truncated proteins
We assessed the potential impact of nonsense or frameshift variants in terms of the molecular mechanism of dominance. We studied the mechanism of dominance not only for the seven nonsense or frameshift variants in our cohort but also for the eight reported variants (Fig. 2 A). Three of the eight reported variants (c.1140-2A>C, c.1140-2A>G, and c.1140-3C>G), all affecting the same splice site, were predicted to create an identical transcript (S381*). We thus assessed the STAT3 activity of the 13 cDNAs corresponding to these 15 variants in a luciferase assay. We used an empty plasmid (empty vector [EV]), WT STAT3, or each variant cDNA to transfect a STAT3-deficient epithelial DLD1
Figure 2. Allele activity and the mechanism of dominance of nonsense or frameshift STAT3 variants. (A) Schematic diagram of nonsense or frameshift STAT3 variants found both in our cohort (labeled in red) and in previous studies (labeled in black). (B and C) Luciferase assay on STAT3−/− A4 cells (B) or HEK293T cells (C) transfected with no vector (mock), EV, WT STAT3 (WT), or STAT3 variants, together with the pGL4.47 luciferase reporter construct and an expression vector for Renilla luciferase. After 24 h, transfected cells were left untreated or treated with 100 ng/ml sIL-6 for 24 h. The y axis represents STAT3 activity (% of WT) or STAT3 transcriptional activity (Fold induction).
transcriptional activity normalized against unstimulated activity in EV-transformed cells. The x axis indicates the alleles used for transfection. R382W, D427ins17, and Y705F, which have been shown to be amorphic and DN, were used as positive controls. In-house variants are shown in red, previously reported variants in black, and positive controls in blue. (B) STAT3 transcriptional activity in STAT3+/− A4 cells transfected with mock, EV, WT, or STAT3 variants (canonical transcripts). STAT3 transcriptional activity in STAT3+/− A4 cells transfected with EV was considered to be 0% (red dashed line) and the level of STAT3 transcriptional activity in STAT3−/− A4 cells transfected with WT was considered to correspond to 100% activity. The variants are classified on the basis of percentage of STAT3 transcriptional activity as follows: <0%, amorphic; between 0% and 25%, severely hypomorphic; between 25% and 75%, mildly hypomorphic; between 75% and 125%, isomorphic; and >125%, hypermorphic. (C) STAT3 transcriptional activity in HEK293T cells expressing endogenous STAT3 transfected with mock, EV, WT STAT3 plasmid or various amounts of the STAT3 variants (described in the upper right box). sIL-6–stimulated EV-transfected HEK293T cells (endogenous STAT3 activity) are indicated by a red dashed line. The black dashed line represents 50% EV activity. The variants are classified on the basis of endogenous STAT3 activity; variants with transcription levels below the endogenous level are considered to be DN (weakly DN if >50% activity and strongly DN if <50%), and those with variants above the endogenous level are considered to be not DN. Each experiment (B and C) was independently performed three times.

colon carcinoma cell line (A4 STAT3−/−), together with a STAT3-specific luciferase reporter, and we measured luciferase activity. We subsequently activated the cells by incubation with IL-6/IL-6Ra (sIL-6) to assess STAT3 transcriptional activity by measuring luciferase activity. Three of the 13 STAT3 cDNA variants (D698Tfs*9, D698Efs*6, and S701Kfs*17) were amorphic (no activity), five (F493Lfs*16, R518*, W562*, Q633*, and Y657*) were severely hypomorphic (between 0% and 25% activity), and another five (R134Vfs*11, Q125*, D371Lfs*14, S381*, and R729Pfs*12) were mildly hypomorphic (between 25% and 75% activity; Fig. 2 B). We then assessed the molecular mechanism of dominance by using each variant allele to transfected human embryonic kidney (HEK) 293T cells, which endogenously express STAT3. After transfection with the STAT3–dependent luciferase reporter construct and following sIL-6 stimulation, two variant alleles, R134Vfs*11 and Q125*, were found not to have a DN effect. However, a DN effect was observed for the other 11 variant constructs due to the production of a truncated protein, as described for known DN variants (R382W, D427ins17, and Y705F; Fig. 2 C). Interestingly, R518*, S381*, and D371Lfs*14 also behaved in a DN manner, although the DN effect of these variants appeared to be weaker than that of the other variants. Overall, eight variants had a strong DN effect, five variants had a weaker DN effect, and two had no detectable DN effect (Table S2).

### Seven out-of-frame STAT3 variants create alternative transcripts

The results described above suggested that the mechanism of dominance might depend on the variant. We hypothesized that some variants might generate two or more mRNAs: the canonical transcript and at least one alternative transcript generated by the variant and potentially both LOF and DN. This hypothesis has not been tested for any of the previously reported STAT3 nonsense or frameshift variants (Abolhassani et al., 2018; Anolik et al., 2009; Natarajan et al., 2018; Renner et al., 2008; Tavassoli et al., 2019; Woellner et al., 2019; Woellner et al., 2019). We tested the impact of the variants on transcript levels in an exon-trapping assay (Duyk et al., 1990); we tested all HIES-associated STAT3 nonsense or frameshift variants from our in-house cohort and previous studies (Abolhassani et al., 2018; Anolik et al., 2009; Natarajan et al., 2018; Renner et al., 2008; Tavassoli et al., 2019; Woellner et al., 2010). We first generated variant plasmids for seven nonsense or frameshift variants from our cohort and performed exon trapping (Figs. 3 A and S1 and Table S2). The c.1111_1112del (p.D371Lfs*14), c.1552C>T (p.R518*), and c.1685G>A (p.W562*) variants resulted in the generation of new alternative transcripts. The c.1111_1112del variant encoded a D371Lfs*14 due to the retention of intron 12 as an alternative transcript. The c.1552C>T variant encoded a R518fs*14 due to the creation of a new splice donor site. The c.1685G>A variant encoded E552-W562del as an alternative transcript due to the creation of a new splice acceptor site. We then investigated the ratio of these transcripts. For c.1111_1112del, the canonical transcript encoding D371Lfs*14 and the alternative transcript derived from intron retention and encoding D371Lfs*14 were detected at frequencies of 98% and 2%, respectively. For c.1552C>T, the canonical (R518*) and alternative (R518fs*14) transcripts were detected at frequencies of 80% and 20%, respectively. For c.1685G>A, the canonical (W562*) and alternative (E552-W562del) transcripts were present at frequencies of 83% and 17%, respectively. The other variants did not generate any detectable alternative transcripts (Table S2). We then investigated alternative splicing for the eight previously reported nonsense or frameshift variants pLOF (Natarajan et al., 2018; Tavassoli et al., 2019; Figs. 3 A and S1 and Table S2). Next, we performed RNA sequencing (RNA-seq) to detect STAT3 alternative transcripts in the patient’s heterozygous cells. RNA-seq data from P3 EBV-B cells, two Q4 control EBV-B cells, and a patient with a deep intronic mutation creating a new splice site (Khourieh et al., 2019) suggested that the c.1552C>T (p.R518*) mutation led to a new event in the vicinity of the mutation, but we failed to identify alternative transcripts due to low coverage (Fig. S2 A). We thus performed a deep sequencing of preamplified STAT3 cDNA from P3, P4, and two controls’ primary fibroblasts pretreated with emetine dihydrochloride hydrate, an inhibitor of mRNA decay process (Nooensie and Dietz, 2001). These analyses revealed that the patients’ cells (P3, P4) actually display STAT3 alternative transcripts, consistent with the results of exon-trapping (Fig. 3 B; and Fig. S2, A–C). These analyses also revealed that P3 cells harbored a new STAT3 alternative transcript (deletion of exon 17–19, p.N489Vfs*8), which was not detected by exon-trapping (Fig. 3 B). We found that c.1140-2A>C, c.1140-2A>G, and c.1140-3C>G variants resulted in alternative transcripts. Indeed, c.1140-2A>C, c.1140-2A>G, and c.1140-3C>G created S381Pfs*2, and c.2144+1G>A created S701D732delinsN as alternative transcripts. For c.1140-2A>C, S381* had a frequency of 52% and S381Pfs*2 had a frequency of 48%. For c.1140-2A>G, S381* had a frequency of 81% and S381Pfs*2 had a frequency of 19%.
Identification and characterization of abnormal splicing in STAT3 nonsense or frameshift variants in HIES. (A) The ratio of canonical transcripts to alternative transcripts was determined by exon trapping. The percentages of WT (gray) and splice variants (canonical transcript in pink and alternative transcripts in green) of STAT3 transcripts are shown for nonsense or frameshift variants (P3 and P4). Amplified STAT3 cDNA were analyzed by MiSeq to identify the rare splicing events and shown by Sashimi plot representation using Integrative Genomics Viewer. Pink colored lines and legends represent canonical transcripts. (B) Deep RNA analysis for the identification of new alternative splicing in HIES patient primary fibroblasts with STAT3 nonsense variants (P3 and P4). Amplified STAT3 cDNA were analyzed by MiSeq to identify the rare splicing events and shown by Sashimi plot representation using Integrative Genomics Viewer. Pink colored lines and legends represent canonical transcripts. (C) Deep RNA analysis for the identification of new alternative splicing in HIES patient primary fibroblasts with STAT3 nonsense variants (P3 and P4). Amplified STAT3 cDNA were analyzed by MiSeq to identify the rare splicing events and shown by Sashimi plot representation using Integrative Genomics Viewer. Pink colored lines and legends represent canonical transcripts.
splicing. Green colored lines and legends represents alternative splicing already detected in the exon-trapping assay. Blue colored lines and legends represent alternative splicing newly identified in this assay. C, healthy control; P3, R518*, P4, W562*. Red arrow represents the position of each variant. (C and D) STAT3 transcriptional activity in STAT3−/− A4 cells (C) or HEK293T cells (D) transfected with mock, EV, WT, or STAT3 variants, including variants generating alternative transcripts, together with the pGL4.47 reporter construct and an expression vector for Renilla luciferase. After 24 h, the transfected cells were left untreated or treated with 100 ng/ml siL.6 for 24 h. The y axis represents STAT3 transcriptional activity levels normalized against unstimulated activity in EV-transformed cells. The x axis indicates the alleles used for transfection. R382W, D427ins17, and Y705F were used as positive controls. The pink bar represents the canonical transcript and the green bar represents the alternative transcript. Stimulated EV-transfected HEK293T cell activity is indicated by a red dashed line. All variants are classified according to the rules described in Fig. 2, B and C. Each luciferase assay (C and D) was independently performed three times.

We measured protein levels with mAbs directed against the STAT3 protein. We assessed the allele activity of each of the alternative transcripts by transfecting A4 cells. All the new alternative transcripts identified by exon-trapping were produced as expected (Fig. S3, A and B; and Table S2). We assessed the allele activity of each of the alternative transcripts by transfecting A4 STAT3−/− cells with each variant allele. We found that one of the variants was amorphic (S701D, N727D, and S701Kfs*17) and a second with the second reinitiation transcript only (Met28 and Met143), and a third with the second reinitiation transcript only (Met99 and Met62; Fig. 4 B). The truncated R13Vfs*11 form with the premature stop codon resulted in a loss of STAT3 activity through negative dominance, whereas both reinitiation constructs were either isomorphic or slightly hypermorphic (Fig. 4 C). The truncated form of the Q125* construct was also LOF and DN. Moreover, the Met143 construct behaved in a hypermorphic, but not DN, manner, whereas the Met62 reinitiation construct was both amorphic and DN (Fig. 4 C). These findings are consistent with previous reports that amino acids 150–162 in the coiled-coil domain (CCD) were indispensable for the nuclear translocation of STAT3 (Liu et al., 2005). Thus, both the R13Vfs*11 and Q125* variants can underlie the reinitiation of translation, with the encoded proteins having a DN effect on STAT3 transcriptional activity. All the out-of-frame variants of STAT3 were DN but by different mechanisms involving truncated proteins (n = 6), isoforms from alternative transcripts (n = 1), proteins generated through translation reinitiation (n = 2), or combinations of these mechanisms (n = 6).

Mutations that lacked a reinitiation tract.

We studied two remaining variant alleles, R13Vfs*11 and Q125*, for which the canonical transcripts were not DN, and which did not express alternative transcripts. We considered the possibility of a reinitiation of translation in R13Vfs*11 and Q125*, because immunoblot revealed the presence of proteins with a lower molecular weight (Fig. S3 A). We first determined which Met after each premature stop codon was used for reinitiation.

Activity and mechanism of dominance in all STAT3 variants reported in AD-HIES patients

In the Human Gene Mutation Database of pathological variants and PubMed, 143 variants are identified as causal variants of AD-HIES (Fig. S3 and Table S3; Abolhassani et al., 2018; Al Khatib et al., 2009; Alcántara-Montiel et al., 2016; Anolik et al., 2009; Avery et al., 2010; Egawa et al., 2019; Eken et al., 2020; Felgentreff et al., 2014; Freeman et al., 2013; Giacomelli et al., 2011; Hagl et al., 2016; He et al., 2012; Heimall et al., 2011; Holland et al., 2007; Jiao et al., 2008; Khourieh et al., 2019; Kim et al., 2009; Kumánovics et al., 2010; Larsen et al., 2019; Ma et al., 2012; Ma et al., 2008; Merli et al., 2014; Minegishi et al., 2007; Miyazaki et al., 2011; Moens et al., 2017; Moens et al., 2014; Mogensen, 2013; Natarajan et al., 2018; Papanastasiou et al., 2010; Pelham et al., 2016; Powers et al., 2012; Qin et al., 2001; Rambotti et al., 2015; Rondou et al., 2010; Schultz et al., 2018; Simmonds et al., 2010; Sokol et al., 2010; Sotiriou et al., 2011; Stolz et al., 2010; Sun et al., 2007; Thomsen et al., 1999; Wang et al., 2008; Zhang et al., 2012; Zhou et al., 2006).
Figure 4. Analysis of translation reinitiation for the R13Vfs*11 and Q125* variants of STAT3. (A) Western blot of extracts from nontransfected STAT3−/− A4 cells (mock), A4 cells transfected with EV, the STAT3 WT allele, or the STAT3 variant allele of interest. All extracts were probed with an antibody against the C terminus (C-ter) of the STAT3 protein. (Left) Each Met codon after the premature stop codon was mutated to give a Leu codon, at Met28, Met99, Met143, Met162, Met185, and Met200. Variants with such mutations at both Met28 and Met99 were also constructed. R13Vfs*11 and each variant with a mutated Met codon were used to transfect A4 cells, and the STAT3 protein profile was then checked. Right: Using the same procedure as for R13Vfs*11, each Met codon after the premature stop codon was mutated to give an Ala codon at Met143, Met162, Met185, and Met200. A construct with a double Met143Ala and
null A4 cells, and STAT3 protein levels were then checked. MW, molecular weight; (B) a short-form construct with a premature stop codon (R13Vfs*11 and Q125*) and a reinitiation construct (translation starts at Met28, Met99, Met143, and Met162) were designed and constructed from each variant plasmid by mutagenesis. CC, coiled-coil domain; DBD, DNA-binding domain; LD, linker domain; TA, transactivation domain. (C) Luciferase assay on STAT3−/− A4 cells (left figure) or HEK293T cells (right figure) transfected with mock, EV, WT, or STAT3 variants, including variants generating alternative transcripts, together with the pGL4.47 reporter construct and an expression vector for Renilla luciferase. After 24 h, the transfected cells were left untreated or treated with 100 ng/ml sIL-6 for 24 h. The y axis represents STAT3 transcriptional activity normalized against unstimulated activity in EV-transformed cells. The x axis shows the alleles used for transfection. R382W was used as a positive control for the DN effect. The black bar represents the results for the R13Vfs*11 variant, and the gray bar represents the results for the Q125* variant. The red dashed line represents activity in stimulated EV-transformed cells. All variants are classified according to the rules described in Fig. 2, B and C. Each experiment (A and C) was independently performed three times.

2009; Renner et al., 2008; Renner et al., 2007; Robinson et al., 2011b; Schimke et al., 2010; Sundin et al., 2014; van de Veerdonk et al., 2010; Vinh et al., 2010; Woellner et al., 2010; Wolach et al., 2014; Wu et al., 2017; Zhang et al., 2013). In summary, 114 of these variants are missense variants, 21 are in-frame indels, and 8 are out-of-frame variants. Among the 143 reported variants, 2% (n = 3) concern the N-terminal domain, 43% (n = 62) the DNA-binding domain, 5% (n = 7) the linker domain, 35% (n = 50) the SH2 domain (SH2), and 15% (n = 21) the transactivation domain (Fig. S4 and Table S4). There were no reported variants of the CCD in patients with HIES (Fig. S4 and Table S4), whereas the remaining seven variants of the CCD have been reported for STAT3 gain-of-function syndrome (Fabre et al., 2019). We assessed the allele activity and mechanism of dominance for all these variants in our luciferase assay system. We found that 107 of the 114 missense variants were amorphic or hypomorphic, 5 were isomorphic, and 2 were hypermorphic (Fig. 5 A and Table S3). Strong DN effects (>50%) were observed for 95 variants, whereas the DN effect was not convincingly strong for 19 variants (Fig. 5 B and Table S3). All 21 in-frame indels were amorphic or hypomorphic, and 20 of them displayed strong DN effects; only one variant (c.1107_1109del (p.D369_L370delinsE)) had a DN effect that was not convincingly strong (Fig. 5, A and B; and Table S3). We then analyzed more precisely the mechanism of dominance for the 20 variants without strong DN effects in luciferase assays (weak DN, 9; not DN, 11). We found that five had dose-dependent DN effects (F384C, D369L370delinsE, R423Q, N472D, and K642E; Fig. S5 A). We considered these variants to be DN, because dose dependence is an element of the definition of negative dominance. The remaining 15 variants had no detectable dose-dependent DN effects, although some did exert mild DN effects (Fig. S4 and Table S5). We investigated whether the variants generated alternative transcripts encoding proteins with DN effects using exon-trapping assays, as in the previous experiment. 13 of the 15 variants generated alternative transcripts (Table S5), and 8 of these alternative transcripts had DN effects (Fig. S5, B and C; and Table S5). The remaining seven variants did not appear to be DN (H58Y, R84Q, H332R, R335W, T341N, N425Y, and A744V). The corresponding patients may benefit from (1) being reanalyzed genetically, as they might carry mutations in other HIES-causing genes (e.g., IL6ST, ZNF341, and IL6R) or (2) being analyzed more thoroughly for STAT3 mRNAs using primary cells, as our exon-trapping may have missed an alternative transcript (Alcántara-Montiel et al., 2016; Freeman et al., 2013; Jabara et al., 2012; Miyazaki et al., 2011; Renner et al., 2008; Woellner et al., 2010). Thus, some missense variants can influence splicing, generating aberrant transcripts and resulting in DN effects. We found that 95% (136 variants) of the 143 variants behaved in a DN manner at the cellular level (Fig. 6; Table S3 and Table S5), either directly or through the production of aberrant transcripts.

Activity and mechanism of dominance of 14 STAT3 variants reported in gnomAD
We investigated a subset of STAT3 variants found in the general population. The gnomAD database includes 1,095 STAT3 variants, of which 141 are missense, 1 is an in-frame deletion, 1 is a nonsense variant, and 167 are synonymous. We selected 14 gnomAD variants, including 12 missense variants (Q25E, R246Q, A376V, G402S, D427G, V461L, I498V, R512Q, A702T, G743V, S763L, and A766T) with either a minor allele frequency above 1 × 10−4 or a CADD score higher than 30, 1 variant (E239del) with an in-frame deletion, and the only variant pLOF (Q247*; minor allele frequency, 4 × 10−6; CADD, 39; Fig. 7 A and Table S6). We assessed their allele activity and the mechanism of dominance in the same luciferase assay system. Only Q247* was severely hypomorphic, three variants (Q25E, D427, and G743V) were mildly hypomorphic, and two variants (A376V and R512Q) were hypermorphic, whereas the remaining eight variants had normal activity (Fig. 7 B). None of the 14 gnomAD variants were found to be DN (Fig. 7 C). The sole LOF variant (Q247*) had no DN effect, even when the amount of the Q247* allele was increased (Fig. 7 D). We also showed that this variant did not create a splice site or reinitiate translation (Fig. 7, E and F). It was not possible to retrieve clinical data for the carrier of this mutation, but this individual belonged to a neurochort from gnomAD and was unlikely to have HIES. Thus, the analysis of STAT3 variants in the general population suggested that most of the common variants are isomorphic and that the only variant predicted to be LOF was indeed LOF but did not display DN behavior, suggesting that STAT3 HI is clinically silent.

Allele activity and prediction of the functional impact of STAT3 variants
An analysis of 150 heterozygous STAT3 variants in AD-STAT3-HIES (143 already reported and 7 new variants reported here), with software for predicting the deleteriousness of mutations (Kircher et al., 2014) identified only one deep intronic variant leading to alternative splicing (Khourieh et al., 2019), which was predicted to be nonpathogenic, with a low CADD score (8.742), below the 99% confidence threshold MSC score for STAT3 (16.4; Itan et al., 2016; Rentzsch et al., 2019; Zhang et al., 2018a). All the
remaining 149 variants were predicted to be deleterious (CADD >16.4; Table S3). By contrast, 99% (141/143) of the 143 heterozygous nonsynonymous coding variants of STAT3 in the gnomAD database (missense, 141; in-frame deletion, 1; nonsense, 1) had CADD scores above the MSC, suggesting that they may be deleterious (Fig. 7 A). We analyzed the correlation between CADD and STAT3 activity for all the variants tested. There was no correlation between STAT3 activity and CADD score ($R^2 = 0.013$). Five variants (3.3%) identified as isomorphic by our assay were predicted to be deleterious by CADD and may therefore be considered false positives (FPs). By contrast, all the hypomorphic or amorphic variants were predicted to be deleterious and are therefore true positives (Fig. 8 A). We also tested the performance of other software for predicting the deleteriousness of mutations (Condel, PPH2, Sift, FATHMM, Mutation Assessor, and Provean; Adzhubei et al., 2010; Choi and Chan, 2015; González-Pérez and López-Bigas, 2011; Reva et al., 2011; Rogers et al., 2018; Vaser et al., 2016). For all six programs, there was no correlation between the predicted deleteriousness score and the activity of the STAT3 variants (Fig. 8 B). Thus, predictions of
Figure 7. Allele activity and the mechanism of dominance for STAT3 variants found in the general population. (A) Graph showing the predicted CADD scores and global allele frequencies of nonsynonymous coding STAT3 variants found in gnomAD (v2.1). 14 variants, including a nonsense variant, were selected and analyzed (shown in red). The CADD-MSC score (99% confidence interval) for STAT3 is indicated by a red dashed line. (B and C) Luciferase assay on STAT3−/− A4 cells (B) or HEK293T cells (C) transfected with mock, EV, WT, or STAT3 variants from gnomAD, together with the pGL4.47 reporter construct and an expression vector for Renilla luciferase. After 24 h, the transfected cells were left untreated or treated with 100 ng/ml sIL-6 for 24 h. The y axis represents STAT3 transcriptional activity normalized against unstimulated activity in EV-transformed cells. The x axis represents the alleles used for transfection. R382W
was used as a positive control, as its functional impact has already been determined. (D) Luciferase assay on HEK293T cells, which have endogenous STAT3, with various amounts of WT or Q247*. (E) Schematic representation of the ratio of canonical to alternative transcripts, according to an exon-trapping assay on the Q247* variant. Exon-trapping analysis revealed that all the transcripts were canonical (represented in pink), with none resulting in alternative transcripts, for the Q247* variant. (F) Western blot of extracts from nontransfected STAT3−/− A4 cells (mock); A4 cells transfected with pCMV6 EV, the STAT3 WT allele, or the STAT3 variant alleles (Q247†). All extracts were probed with mAbs specific for the C-terminal part of STAT3 or the N-terminal part of the STAT3 protein. No protein was detected with an antibody recognizing the C terminus, whereas a faint band at ∼25 kD was detected with an antibody recognizing the N terminus, suggesting that the N-terminal part (1−246) of STAT3 is weakly expressed. Each experiment (B, C, D, and F) was independently performed twice.

deleteriousness do not provide an adequate estimate of the experimentally determined deleteriousness of the variants and, by inference, their pathogenicity.

The population genetics of human STAT3 supports a mechanism of negative dominance

Biologically disruptive heterozygous variants of STAT3 are likely to have clinical effects. Consistently, the gene damage index for STAT3 is low (1.5; Itan et al., 2015). Human genetic variants that cause severe diseases are progressively purged from the population, as their carriers have fewer children than noncarriers. Indeed, STAT3 is under strong purifying selection, with an f value of 0.26 (among the top 2% of human protein genes in terms of constraint; Table S7; Eilertson et al., 2012). This natural process, negative selection, has a stronger effect on disease variants that affects heterozygous individuals (Fuller et al., 2019). We recently developed a new statistical tool, called consensus negative selection (CoNeS), combining interspecies and intraspecies measurements of negative selection, with which we showed that genes underlying AD or X-linked EIE are under stronger negative selection than those underlying AR IEI (Rapaport et al., 2021). In this system, as expected, STAT3 was found to be under very strong negative selection, with a CoNeS value of −1.94 (Fig. 8 C and Table S7). Only three genes underlying EIE have been shown to be under stronger negative selection (ACTB, −2.17; CHD7, −2.11; and KMT2D, −2.75; Table S7). This finding is consistent with the AD inheritance of STAT3-HIES and the low frequency of missense variants in the general population, as shown by our analysis of the gnomAD database. Moreover, there is only one nonsense variant in gnomAD (Q247†) pLOF, but this variant is not DN (Fig. 7, B–D). In our analysis of 150 STAT3 variants, 143 (95%) were predicted to be DN, causing HIES (Table S3). Thus, the population genetics of STAT3 provides no evidence for HI associated with AD disease in STAT3.

Discussion

Following the discovery of the first genetic etiologies of HIES (heterozygous variants of STAT3; Minegishi et al., 2007) in 2007, 143 private STAT3 variants had been reported in HIES patients by July 1, 2020 (and 7 others are reported here). However, STAT3 activity and the mechanism of dominance have been experimentally elucidated for only six variants (Khourieh et al., 2019; Minegishi et al., 2007). We show here that the canonical transcripts of 95.3% (143/150) of these variants encode STAT3 proteins with little or no STAT3 activity. 15 of these variants were pLOF. All but one variant associated with HIES had a CADD score above the MSC (16.4), predictive of deleteriousness, resulting in a very low false-negative (FN) rate, at 0.7% (1/150). The only incorrectly predicted variant is actually a deep intronic variant, with a CADD of 8.742. Misprediction for intronic variants is a well-known source of FNs (Mather et al., 2016). Conversely, five variants from HIES patients predicted to be deleterious actually encoded STAT3 proteins with normal activity, resulting in a FP rate of 3.3%. The remaining two have a lower CADD score. With the benefit of hindsight, the patients carrying the seven isomorphic variants did not have HIES or had not been studied for other genetic etiologies. By contrast, only one of the coding variants from the gnomAD database selected for experimental testing on the basis of their rarity and high CADD score (n = 14), a nonsense variant (pLOF), was experimentally LOF, and none of these variants were DN (including the LOF variant). We did not thoroughly investigate all other 143 nonsynonymous coding variants found in the general population, but we can assume that most public variants are isomorphic, and at least not DN, because they are more common or have lower CADD scores than those tested. Moreover, the gnomAD database includes very few patients with developmental or infectious diseases, the two major characteristic of individuals with HIES. Nevertheless, most of the nonsynonymous coding variants found in the general population were predicted to be deleterious on the basis of CADD, with a FP rate of 98% that questions the usefulness of this score to diagnose HIES patients (141/143; Fig. 8 A). Further experimental studies are required to determine whether these predictions are correct or false. Meanwhile, computational prediction of the functional impact of coding variants of STAT3 remains challenging due to the high FP rate. We therefore suggest that any new variant found in a patient with HIES should be tested functionally for an accurate molecular diagnosis.

The mechanism of dominance underlying AD conditions should, ideally, be reliably established, to improve both our understanding of the pathogenesis of the condition and its clinical management. Before the identification of nonsense variants in HIES patients, most studies referred to the original 2007 paper and did not call the mechanism of negative dominance for HIES into question. By experimentally testing 150 STAT3 variants found in HIES patients, we show that AD STAT3 deficiency invariably underlies HIES by DN. HIES-causing STAT3 variants encode (1) truncated proteins, (2) neo-proteins generated from a translation reinitiation codon, (3) isoforms from alternative transcripts, or a combination thereof, which account for their DN. These findings suggest that variants of uncertain significance in STAT3 or other genes can benefit from being experimentally tested at the transcript level by exon trapping and/or deep RNA-seq (Boisson et al., 2019). In terms of population genetics, STAT3 is under strong purifying selection
Figure 8. Population and evolutionary genetics for human STAT3. (A) Diagram showing the correlation between predicted CADD scores and percentage of STAT3 transcriptional activity for all 150 STAT3 variants (canonical transcripts; red circles, left figure) and 14 gnomAD variants (black circles, right figure) analyzed here. The y axis represents the CADD score, and the x axis represents the percentage of STAT3 transcriptional activity. The STAT3 transcriptional activity of all variants was normalized against WT STAT3 transcriptional activity. The CADD-MSC score (99% confidence interval) for STAT3 is indicated by a red solid line. The gray dashed line is the regression line (R^2). The blue rectangle represents the FN area, and the yellow rectangle represents the FP area. The white rectangle represents the true-positive (TP) area.

(B) We used six programs other than CADD (Condel, FATHMM, Mutation Assessor, Polyphen2 [PPH2], Provean, and Sift) to predict the impact of the STAT3 variants (canonical transcripts), and we compared the scores obtained with transcription levels. In each diagram, the red solid line represents the mutational cutoff value. The blue rectangle represents the FN area, and the gray rectangle represents the FP area. The gray dashed line is the regression line (R^2).

(C) CoNeS analysis. CoNeS is our custom-built score integrating known negative selection scores through a principal-component projection. The mean value is 0, and low values (negative values) correspond to strong negative selection.
(Rapaport et al., 2021), consistent with the reported AD mode of inheritance. Interestingly, in the absence of a pLOF allele in public databases (with one exception), STAT3 clusters with other genes (Table S7) for which AD phenotypes have been shown to be due to HI (Lek et al., 2016). Our work suggests an alternative hypothesis: pLOF variants are absent from public databases because these alleles are expressed and have a DN effect on the WT allele. Our finding that the only nonsense variant (pLOF) in gnomAD (assuming that the person does not have a HIES phenotype) was actually amorphic, but not DN, is consistent with this alternative hypothesis. It is also consistent with the absence of reports of a phenotype similar to HIES in mice heterozygous for a null loss-of-expression STAT3 allele, whereas a DN mouse model (expressing the DN V463del allele) reproduces some of the features of HIES patients (Hillmer et al., 2016; Panopoulos et al., 2006; Siegel et al., 2013; Steward-Tharp et al., 2014; Zhang et al., 2010a; Zhang et al., 2018b). In addition, to our knowledge, incomplete penetrance has not been reported for AD-STAT3-HIES, whereas it is frequently reported in cases of dominance by HI (Bolze et al., 2018; Borghini et al., 2011; Rieux-Laucat and Casanova, 2014). Overall, our work provides an overarching view of STAT3 variants in patients with HIES and the general population and establishes that negative dominance is the only mechanism of dominance underlying AD-HIES in patients heterozygous for pathogenic STAT3 variants.

Materials and methods

Case reports

P1 (kindred A) is a 16-yr-old boy of Syrian origin. He was born at full term to consanguineous parents. He had a history of recurrent sinopulmonary infections during childhood, with failure to thrive (weight and height below the third percentile for age). He developed right lobar pneumonia and bronchiectasis. Bronchoalveolar lavage fluid culture yielded Aspergillus flavus. Mycobacteria were not detected (TB culture, PCR, and cultures of acid-alcohol-resistant bacilli were negative). P1 was diagnosed with invasive pulmonary aspergillosis, leading to the initiation of treatment with itraconazole and methylprednisolone. He was admitted to hospital while on treatment for invasive aspergillosis in the lung. He also had a large abscess in the left frontal lobe of the brain and the left lobe of the cerebellum. Culture of a brain biopsy specimen also yielded A. flavus. Laboratory tests revealed a neutrophil count of 1,700/mm³, a lymphocyte count of 1,300/mm³, a monocyte count of 400/mm³, and an eosinophil count of 1,700/mm³. Serum Ig levels were 2,180 mg/dl for IgG, 130 mg/dl for IgM, 1,119 mg/dl for IgA, and 7,430 IU/ml for IgE. The patient’s IgE level was quite high, and he had an NIH score of 42 but no skin abscesses, primary tooth retention, scoliosis, bone fractures, candidiasis, eczema, rash, or characteristic facial features. HIES was suspected, and a genetic study was conducted. At the age of 14 yr, a heterozygous nonsense variant, c.373C>T, of STAT3 was identified by whole-exome sequencing (WES) and confirmed by Sanger sequencing. The patient was therefore diagnosed with AD-STAT3-HIES. This variant was not identified in his parents, suggesting that it was a de novo variant. The brain and cerebellum lesions regressed on treatment with amphotericin B, voriconazole, caspofungin, and G-CSF. The patient is currently receiving voriconazole prophylaxis and Ig replacement therapy.

P2 (kindred B) is a 37-yr-old man of Australian origin. His mother died at the age of 20 yr from pneumonia (no other details were available). His sister and his two young children are all healthy. From childhood, P2 suffered from an inflammatory skin disease (labeled as psoriasis), but with no obvious history of pyogenic skin eruptions. He also suffered from recurrent episodes of pneumonia, long-standing atypical predominantly cystic bronchiectasis, and asthma. His total IgE levels were high, peaking at 9,000 IU/ml, and eosinophil counts were also high. Scedosporium apiospermum, Cryptococcus gattii, and Aspergillus niger complex were isolated on multiple occasions. P2 is also colonized with Stenotrophomonas maltophilia, and methicillin-resistant Staphylococcus aureus has been identified on occasions. P2 has no connective tissue or skeletal abnormalities. Basic immunophenotyping and proliferation assays were normal. STAT1 phosphorylation was normal, and no autoantibodies against gamma-IFN or GM-CSF were detected. Neutrophil function was also normal. Serological tests for HIV were negative. P2 required multiple arterial embolizations for hemoptysis and requires long-term posaconazole prophylaxis for recurrent fungal infections. WES identified a heterozygous nonsense variant, c.1111_1112del, of STAT3, which was confirmed by Sanger sequencing. P2 was therefore diagnosed with AD-STAT3-HIES.

P3 (kindred C) is a 6-yr-old boy of Indian origin. At the age of 1 yr, he developed watery vesicles all over the body, with fever and itching. Ointments had no effect on the vesicles, which self-resolved, leaving hyperpigmented scars. At the age of 2 yr, P3 suffered pneumonia, with right empyema requiring drainage. The lower lobe of the right lung collapsed. Bronchoalveolar lavage fluid culture yielded S. aureus, and test results for TB were negative. P3 developed right lower chest wall and left calf muscle abscesses and also displayed pallor and multiple macular spots on the lower abdomen. HIV tests were negative, and nitroblue tetrazolium reduction test results were normal, as were the results of the dihydrorhodamine assay, a finding not consistent Q7 with chronic granulomatous disease. P3 had very high absolute eosinophil counts and high IgE levels (9,752 IU/ml). The other Igs were present at normal levels. P3 presented no atopy, coarse facies, or candidiasis. His NIH score reached 40, and he had low T helper cell type 17 levels. HIES was suspected and a genetic study was performed. WES identified a heterozygous nonsense variant, c.1552C>T, of STAT3, which was confirmed by Sanger sequencing when the patient was 3 yr old. This variant was not identified in the parents and siblings of P3, suggesting that it was a de novo variant. The patient was diagnosed with AD-STAT3-HIES.

P4 (kindred D) is a 51-yr-old man with a typical HIES phenotype, with skin abscesses, pneumonia complicated by cavity aspergillosis, and high IgE levels. He also developed vascular abnormalities, with aneurisms of the aorta and coronary artery. NIH score reached 40. At the age of 41 yr, a heterozygous nonsense variant, c.1685G>A, of STAT3 was identified by WES and confirmed by Sanger sequencing. P4 was therefore diagnosed with AD-STAT3-HIES. It was recently shown that he had
transmitted his STAT3 nonsense variant to his son (P5). P5 is a 16-yr-old boy diagnosed with AD-STAT3-HIES at the age of 14 yr. He suffered recurrent skin abscesses, pneumonia, bacterial infections, and osteopenia. He also had high absolute eosinophil counts and high IgE levels (1,133 IU/ml).

P6 and P7 (kindred E) are of New Zealand Maori descent. In terms of the clinical phenotype of HIES, P6 has suffered inflammatory skin disease since childhood, with pyogenic skin infections, bullous lung disease and bronchiectasis, and recurrent staphylococcal pneumonia and bacteremia. He also has high total IgE level (>50,000 IU/ml). He has no connective tissue features. No clinical information is available for his affected daughter (P7) other than her STAT3 genotype.

P8 and P9 (kindred F) are of Moroccan origin. We have limited clinical information for these two patients, both of whom have only a mild elevation of IgE levels (P8, 644 IU/ml; P9, 606 IU/ml) and TB infection. P9, who is an affected daughter of P8, died from infection, and we were unable to obtain any other clinical information in support of suspected HIES. WES identified a heterozygous nonsense variant, c.2093ins22, of STAT3 in both patients, which was confirmed by Sanger sequencing. Our studies implicated this variant in HIES on the basis of functional analysis. These two patients were therefore diagnosed with HIES. This variant was not identified in the father and siblings of these patients by Sanger sequencing, suggesting that it was a de novo variant.

P10 (kindred G) is a 13-yr-old boy of Turkish origin. He had no clinical manifestations indicative of a particular susceptibility to bacterial or viral infections. By contrast, he was susceptible to TB infection. His IgE levels were slightly high (961 IU/ml), and his eosinophil count was 4,180/mm3. We were unable to obtain any information about other clinical manifestations, such as connective tissue signs, that might support a diagnosis of HIES. WES identified a heterozygous nonsense variant, c.2093ins22, of STAT3, which was confirmed by Sanger sequencing. This variant has been implicated in HIES on the basis of functional analysis. The patient was therefore diagnosed with AD-STAT3-HIES. This variant was not identified in his parents and siblings by Sanger sequencing, suggesting that it was a de novo variant.

Ethics statement
The study was approved by and performed in accordance with the requirements of the institutional ethics committees of The Rockefeller University Hospital, New York, NY; Necker-Enfants Malades Hospital, Paris, France; and the Sydney Local Health District Royal Prince Alfred Hospital Zone Human Research Ethics Committee and Research Governance Office, Royal Prince Alfred Hospital, Camperdown, New South Wales, Australia. Informed consent was obtained for all patients and healthy control volunteers reported in the study.

Genetic analysis
Genomic DNA (gDNA) was prepared from blood samples from patients and controls by the standard phenol-chloroform extraction method. WES was performed with an individual institutional protocol. We performed gDNA extraction, WES data collection, and analysis as previously described (Béziat et al., 2020; Khourieh et al., 2019). Exome capture was performed with the SureSelect Human All Exon 71 Mb kit (Agilent Technologies). Paired-end sequencing was performed on a HiSeq 2500 sequencing system (Illumina) generating 100-base reads. We aligned the sequences with the GRCh37 reference build of the human genome with the Burrows-Wheeler aligner (Li and Durbin, 2010). Downstream processing and variant calling were performed with the Genome Analysis Toolkit (McKenna et al., 2010) and SAMtools (Li et al., 2009). Substitution and InDel calls Q8 were made with the GATK Unified Genotyper. All variants were annotated with annotation software developed in-house (Adzhubei et al., 2010; Kircher et al., 2014; Ng and Henikoff, 2001). For gene-targeting approaches and analysis of familial segregation, exons 2–24 of STAT3 and their flanking intron sequences were amplified by PCR with specific oligonucleotide primers (available on request). PCR products were purified by centrifugation through Sephadex G-50 Superfine resin (Amer sham-Pharmacia-Biotech) and sequenced with the Applied Biosystems Big Dye terminator kit v1.1 (Applied Biosystems) and an ABI Prism 3130xl Analyzer (Applied Biosystems). WES was performed for P1, P2, P3, P4, P6, P8, P9, and P10.

Cell culture, cDNA synthesis, and sequencing
Immortalized SV40 fibroblasts from the patients were grown in complete DMEM supplemented with 10% heat-inactivated FBS. HEK293T cells derived from the human embryonic kidney 293 cell line, which expresses a mutant version of the SV40 large T antigen, were grown in complete DMEM supplemented with 10% FBS. STAT3−/− A4 cells derived from human LDLI colon cancer cells by homologous recombination (Yang et al., 2010) and kindly provided by James E. Darnell, were grown in McCoy’s 5A medium containing 10% heat-inactivated FBS. Cells were incubated at 37°C in the presence of 5% CO2.

mRNA extraction and reverse transcription
Total RNA was extracted from the various cell lines with the RNeasy Plus Mini kit (Qiagen). RNA was then reverse transcribed with SuperScript III reverse transcription (Invitrogen). The cDNA was amplified by PCR with specific oligonucleotide primers (available on request) and sequenced as described above.

Expression vectors and transfection experiments
All STAT3 variants in our analysis except for D427ins17 (Khourieh et al., 2019) were generated by site-directed mutagenesis. The WT or variant alleles were reintroduced into a Myc-DDK-PCMV6 vector ( OriGene). The mutated D427ins17 STAT3 allele was generated by amplifying the full-length cDNA from patient EBV-B cells and inserting it into a TOPO cloning plasmid (pCR2.1-TOPO vector; Invitrogen), as described previously (Khourieh et al., 2019). STAT3−/− A4 cells and HEK293T cells were transfected with the Myc-DDK-PCMV6 vector, empty or containing the WT or a variant allele, in the presence of X-tremeGENE 9 DNA Transfection Reagent (Sigma-Aldrich), according to the manufacturer’s instructions.
Exon-trapping assay
The exon-trapping assay was performed with a previously described method and vector (Buckler et al., 1991; Burn et al., 1995). STAT3 gDNA was amplified from control fibroblasts. Each pair of primers (Table S8) was designed to amplify the previous and following exons and introns surrounding the targeted variant exon. Each STAT3 variant amplified from gDNA was inserted into the pSPL3 plasmid (Life Technologies). Each variant was reintroduced by directed mutagenesis, and the full STAT3 gDNA fragment was Sanger sequenced. Plasmids were used to transfect Cos-7 cells with the X-tremeGENE 9 DNA Transfection Reagent kit (Sigma-Aldrich) according to the manufacturer’s instructions. Cells were harvested 24 h after transfection, and total RNA was extracted as previously described. After cDNA synthesis, PCR was performed with the SD6 (5′-TTCAGTCA CTGGACAACC-3′) and SA2 (5′-ATCTGAATGGTATTTGTGAGC-3′) primers for pSPL3 plasmids. The amplified cDNAs were inserted into the pCR4-TOPO plasmid vector (Invitrogen). M13 forward (5′-GTAACAGCAAGGAGCAG-3′) and M13 reverse (5′-CAGGAAACAGCTATGAC-3′) primers were used for amplifications for the analysis of potential splicing. We screened 100 colonies for each variant to analyze splicing.

Luciferase reporter assay
The luciferase reporter assay was performed as previously described (Khourieh et al., 2019). Briefly, the reporter vector pGL4.47 (E4041; Promega) contains five copies of the STAT3-responsive element, the sis-inducible element, linked to the luciferase reporter gene luc2P. STAT3−/− A4 and HEK293T cells were transfected with the pCMV6 vector bearing WT or variant STAT3 (50 ng), the reporter construct pGL4.47 (200 ng), and an expression vector for Renilla luciferase (20 ng), with the X-tremeGENE™ 9 DNA Transfection Reagent kit (Sigma-Aldrich). After 24 h, the transfected cells were left unstimulated or were stimulated with 100 ng/ml recombinant human sIL-6 (8954-SR; R&D Systems) for 24 h. Relative luciferase activity was then determined by normalizing the values against the firefly Renilla luciferase signal ratio.

Western blotting
STAT3−/− A4 cells mock transfected or transfected with the empty pCMV6 vector, WT STAT3, or a variant STAT3 were stimulated by incubation with 50 ng/ml sIL-6 for 20 min at 37°C, under an atmosphere containing 5% CO2. For whole-cell extracts, the cells were lysed by incubation in the following buffer (30 mM Tris, pH 7.5, 2 mM KCl, 2 mM EDTA, and 1% Triton X-100), supplemented with a mixture of protease inhibitors (Sigma-Aldrich) for 30 min at 4°C. The lysates were then centrifuged at 21,000 g for 20 min at 4°C. The supernatants were processed directly for Western blotting. Western blotting was performed as previously described (Khourieh et al., 2019; Kong et al., 2010) on 20 μg of total extract from transfected STAT3−/− A4 cells, with antibodies against the C terminus (9139; Cell Signaling Technology) or N terminus (sc-8019; Santa Cruz Biotechnology) of the STAT3 protein and Tyr705-phosphorylated STAT3 (9145; Cell Signaling Technology).

Reinitiation analysis
For the R13Vfs* and Q125* alleles, we investigated the site at which translation was reinitiated by mutating the first Met after the premature stop codon in each variant. Directed mutagenesis was performed to change Met residues into Leu or Ala residues for each variant. Each of the reinitiation constructs was used to transfect STAT3−/− A4 cells, and Western blotting was performed as described above. Following confirmation of the reinitiation position, selected constructs (Fig. 4 B) were used to transfect STAT3−/− A4 cells and HEK293T cells, and luciferase assay were performed.

RNA-seq
EBV-B cells from healthy controls (n = 2), D427ins17 patients (n = Q:9 I) who were previously known to have the AD-STAT3-HIES-causing variant with alternative transcript, and AD-STAT3-HIES patients (P3: R518*; n = 1) who we studied herein were analyzed. RNA was extracted with the RNeasy plus micro kit (Qiagen). RNA quality was assessed with the Agilent 2200 TapeStation nucleic acids system. RNA-seq was performed as previously described (Hernandez et al., 2018). In brief, samples were mapped onto the human genome sequence (hg38 assembly) with STAR (Dobin et al., 2013), and counts for the aligned reads for each gene were calculated with HTSeq-count (Anders et al., 2015). RNA-seq datasets for these EBV-B cells have been deposited to the NCBI GEO database under accession numbers Q:10 GSE153886 (GSM4658112, GSM4658118, and GSM4658124) and GSE149602 (GSM4505723). Sashimi plot analysis was done using Integrative Genomics Viewer (Robinson et al., 2017; Robinson et al., 2011a; Thorvaldsdottir et al., 2013).

Identification of rare splicing events
Primary fibroblasts from healthy controls (n = 2) and AD-STAT3-HIES patients (P3: R518*, P4: W562*) that are shown to create alternative transcripts in exon trapping were grown in complete DMEM supplemented with 10% FBS for 24 h. Before mRNA extraction, cells were treated by emetine dihydrochloride hydrate (Sigma-Aldrich) 100 μg/ml for 4 h to prevent mRNA decay. mRNA was extracted from primary fibroblasts with TRIzol (Invitrogen) and RNeasy Plus Mini kit (Qiagen) according to the manufacturer’s protocol. RNA was then reverse transcribed with SuperScript III reverse transcription (Invitrogen). The full-length cDNA for STAT3 was amplified by PCR with specific oligonucleotide primers (5′-ATGGCCCAATGGAATCAG CTACAG-3′ and 5′-CTCACATGGGGGAGGTAGCG-3′). After fragmentation and library prep, sequence was performed using MiSeq nano sequencing (illumina) to detect the rare splicing events. These data have been deposited to the Sequence Read Archive (accession no. PRJNA728716).

Prediction software
We used seven software suites to predict the deleteriousness of STAT3 variants (CADD: https://cadd.gs.washington.edu/score, Condel: http://bbglab.irbbarcelona.org/fannsdb/signin?next=%2Ffannsdb%2Fquery%2Fcondel, Polyphen2 (PPH2): http://genetics.bwh.harvard.edu/pph2/, Sift: http://sift.jcvi.org/, Mutation Assessor: http://mutationassessor.org/r3/, FATHMM:
Online supplemental material

Fig. S1 shows the schematic representation of the exon trapping results for STAT3 nonsense or frameshift variants. Fig. S2 shows STAT3 transcript analysis by RNA-seq in HIES patient cells carrying heterozygous STAT3 nonsense variants. Fig. S3 shows protein expression in STAT3−/− A4 cells and STAT3 activity for nonsense or frameshift STAT3 variants. Fig. S4 shows schematic summary of 150 reported variants, including 8 new STAT3 variants in HIES. Fig. S5 shows deeper analysis for 20 STAT3 variants identified as not DN. Table S1 shows a summary of the infectious and clinical phenotypes of patients with STAT3 nonsense or frameshift variants in our cohort. Table S2 lists canonical and alternative transcripts of STAT3 nonsense or frameshift variants in HIES. Table S3 lists the summary of 150 reported and new STAT3 variants. Table S4 shows the correlation between variants consequences and their localization by domain. Table S5 lists the summary of allele activity of both canonical and alternative transcripts for 15 STAT3 missense variants. Table S6 lists the summary of 14 selected STAT3 variants from gnomAD database. Table S7 lists the summary of parameters correlated with negative selection. Table S8 lists primer pairs for the exon-trapping assay.

Acknowledgments

We would like to thank the patients and their families for participating in this study. We thank James E. Darnell and Claudia Mertens for providing the STAT3-deficient A4 cell lines. We thank members of the St. Giles Laboratory of Human Genetics of Infectious Diseases Laboratory, Masato Ogishi, Paul Bastard, Peng Zhang, and Yuval Itan for highly pertinent intellectual discussions and suggestions and Yelena Nemirovskaya, Dana Liu, Dominick Papandre, Mark Woollett, Cécile Patissier, and Tatjana Kochetkov for administrative and technical assistance.

The Laboratory of Human Genetics of Infectious Diseases is supported by the Job Research Foundation, the National Center for Research Resources, and the National Center for Advancing Sciences of the National Institutes of Health Clinical and Translational Science Award program (UL1TR001866); the National Institute of Allergy and Infectious Diseases (P01AI061093); the French National Research Agency (ANR) under the “Investments for the Future” ANR program (ANR-10-IAHU-01), the Integrative Biology of Emerging Infectious Diseases Laboratory of Excellence (ANR-10-LABX-62-IBEID), and the “PNEUMOPID” project (ANR 14-CE15-0009-01); the French Foundation for Medical Research (EQU201903007998); the Howard Hughes Medical Institute; The Rockefeller University; the St. Giles Foundation; Institut National de la Santé et de la Recherche Médicale; and the Université de Paris. A.N. Spaan was supported by the European Commission (Horizon 2020 Marie Sklodowska-Curie Individual Fellowship 789645), the Dutch Research Council (Rubicon Grant 019.171LW.015), and the European Molecular Biology Organization (Long-Term Fellowship ALT 84-2017, nonstipendia). V. Béziat was supported by the ANR (grant NKIRP-ANR-13-PDOC-0025-01).

Author contributions: T. Asano, J.-L. Casanova, and B. Boisson designed the study and wrote the manuscript. F. Rapaport, V. Béziat, Q. Zhang, A. Puel, and S. Boisson-Dupuis discussed the results and the manuscript. T. Asano, J. Khourieh, P. Zhang, F. Rapaport, S.J. Pelham, J.E. Han, A. Guérin, J. Mackie, D. Hum, T. Habib, and N. Marr performed experiments. T. Asano, J. Khourieh, A. Guérin, J. Mackie, M. Chrabieh, S.G. Tangye, L. Abel, and B. Boisson conducted and analyzed exome analysis. S. Gupta, B. Saikia, J.E.I. Baghdadi, I. Fadil, A. Bousfiha, F. Celmeli, L. Ganeshananadan, J. Peake, L. Droney, A. Williams, N. Hatipoglu, T. Ozcelik, and C. Picard provided material from the patients and conducted clinical explorations. A.N. Spaan, S.J. Pelham, J. Li, and W.-T. Lei provided expertise and feedback. J.-L. Casanova and B. Boisson secured funding. All authors critically reviewed the manuscript.

Disclosures: The authors declare no competing interests exist.

Submitted: 4 December 2020
Revised: 30 March 2021
Accepted: 18 May 2021

References

Abolhassani, H., J. Chow, W. Bainter, C.D. Platt, M. Tavassoli, T. Momen, M. Tavakol, M.H. Eslamian, M. Ghangozouli, M. Movahedi, et al. 2018. Clinical, immunologic, and genetic spectrum of 696 patients with combined immunodeficiency. J. Allergy Clin. Immunol. 141:1450–1458. https://doi.org/10.1016/j.jacl.2017.06.049

Adzhubei, I.A., S. Schmidt, L. Peshkin, V.E. Ramensky, A. Gerasimova, P. Bork, A. Kondrashov, and S.R. Sunyaev. 2010. A method and server for predicting damaging missense mutations. Nat. Methods. 7:248–249. https://doi.org/10.1038/nmeth0410-248

Al Khatib, S., S. Keles, M. Garcia-Lloret, E. Karakoc-Aydiner, I. Reisi, H. Artac, Y. Camcioglu, H. Cokugras, A. Somer, N. Kutukculer, et al. 2009. Defects along the T(H)17 differentiation pathway underlie genetically distinct forms of the hyper IgE syndrome. J. Allergy Clin. Immunol. 124: 342–348: 348.e1–348.e5. https://doi.org/10.1067/mae.2009.05.004

Alcántara-Montiel, J.C., T. Staines-Boone, G. López-Herrera, F. Espinosa-Rosas, S.E. Espinosa-Padilla, R. Hernández-Rivas, and L. Santos-Arugumedo. 2016. Functional characterization of two new STAT3 mutations associated with hyper-IgE syndrome in a Mexican cohort. Clin. Genet. 89:217–221. https://doi.org/10.1111/cge.12658

Anders, S., P.T. Pyl, and W. Huber. 2015. HTSeq—a Python framework to work with high-throughput sequencing data. Bioinformatics. 31:166–169. https://doi.org/10.1093/bioinformatics/btu638

Anolik, R., E. Elmariah, S. Lehrhoff, H.J. Votava, F.T. Martinuk, and W. Levis. 2009. Hyperimmunoglobulin E syndrome with a novel STAT3 mutation. Dermatol. Online J. 15:16.

Ashktor, R., and I. Shah. 2016. Tuberculosis in a case of hyper immuno- globulin E syndrome. J. Exp. Med. Clin. 207:587–597. https://doi.org/10.4103/2249-4863.201167

August, A. 2018. Who regulates whom: ZNF341 is an additional player in the STAT3/T(½)7 song. Sci. Immunol. 3:eaaat9779. https://doi.org/10.1126/sciimmunol.aat9779

Avery, D.T., E.K. Deenick, C.S. Ma, S. Suryani, N. Simpson, G.Y. Chew, T.D. Chan, U. Palendira, J. Bustamante, S. Boisson-Dupuis, et al. 2010. B cell-intrinsinc signaling through IL-21 receptor and STAT3 is required for establishing long-lived antibody responses in humans. J. Exp. Med. 207:155–171. https://doi.org/10.1084/jem.20091706

Béziat, V., J. Li, J.X. Lin, C.S. Ma, P. Li, A. Bousfiha, I. Pelleri, S. Zoghi, S. Baris, S. Keles, et al. 2018. A recessive form of hyper-IgE syndrome by disruption of ZNF341-dependent STAT3 transcription and activity. Sci. Immunol. 3:eaaat956. https://doi.org/10.1126/sciimmunol.aat956
Li, H., and R. Durbin. 2010. Fast and accurate long-read alignment with
Larsen, C.S., M. Christiansen, and T.H. Mogensen. 2019. Autosomal Dominant
Kim, H.J., J.H. Kim, Y.K. Shin, S.I. Lee, and K.M. Ahn. 2009. A novel mutation
Khourieh, J., G. Rao, T. Habib, D.T. Avery, A. Lefèvre-Utile, M.O. Chandesris,
Lyons, J.J., Y. Liu, C.A. Ma, X. Yu, M.P. O’Connell, M.G. Lawrence, Y. Zhang,
Mather, C.A., S.D. Mooney, S.J. Salipante, S. Scroggins, D. Wu, C.C. Pritchard,
Mather, C.A., S.D. Mooney, S.J. Salipante, S. Scroggins, D. Wu, C.C. Pritchard,
Mather, C.A., S.D. Mooney, S.J. Salipante, S. Scroggins, D. Wu, C.C. Pritchard,
Mather, C.A., S.D. Mooney, S.J. Salipante, S. Scroggins, D. Wu, C.C. Pritchard,
Mather, C.A., S.D. Mooney, S.J. Salipante, S. Scroggins, D. Wu, C.C. Pritchard,
Mather, C.A., S.D. Mooney, S.J. Salipante, S. Scroggins, D. Wu, C.C. Pritchard,
Mather, C.A., S.D. Mooney, S.J. Salipante, S. Scroggins, D. Wu, C.C. Pritchard,
Mather, C.A., S.D. Mooney, S.J. Salipante, S. Scroggins, D. Wu, C.C. Pritchard,
Mather, C.A., S.D. Mooney, S.J. Salipante, S. Scroggins, D. Wu, C.C. Pritchard,
Mather, C.A., S.D. Mooney, S.J. Salipante, S. Scroggins, D. Wu, C.C. Pritchard,
Mather, C.A., S.D. Mooney, S.J. Salipante, S. Scroggins, D. Wu, C.C. Pritchard,
Mather, C.A., S.D. Mooney, S.J. Salipante, S. Scroggins, D. Wu, C.C. Pritchard,
Mather, C.A., S.D. Mooney, S.J. Salipante, S. Scroggins, D. Wu, C.C. Pritchard,
Mather, C.A., S.D. Mooney, S.J. Salipante, S. Scroggins, D. Wu, C.C. Pritchard,
Mather, C.A., S.D. Mooney, S.J. Salipante, S. Scroggins, D. Wu, C.C. Pritchard,
Mather, C.A., S.D. Mooney, S.J. Salipante, S. Scroggins, D. Wu, C.C. Pritchard,
Mather, C.A., S.D. Mooney, S.J. Salipante, S. Scroggins, D. Wu, C.C. Pritchard,
Mather, C.A., S.D. Mooney, S.J. Salipante, S. Scroggins, D. Wu, C.C. Pritchard,
Mather, C.A., S.D. Mooney, S.J. Salipante, S. Scroggins, D. Wu, C.C. Pritchard,
Mather, C.A., S.D. Mooney, S.J. Salipante, S. Scroggins, D. Wu, C.C. Pritchard,
Mather, C.A., S.D. Mooney, S.J. Salipante, S. Scroggins, D. Wu, C.C. Pritchard,
Mather, C.A., S.D. Mooney, S.J. Salipante, S. Scroggins, D. Wu, C.C. Pritchard,
Mather, C.A., S.D. Mooney, S.J. Salipante, S. Scroggins, D. Wu, C.C. Pritchard,
Mather, C.A., S.D. Mooney, S.J. Salipante, S. Scroggins, D. Wu, C.C. Pritchard,
Mather, C.A., S.D. Mooney, S.J. Salipante, S. Scroggins, D. Wu, C.C. Pritchard,
Mather, C.A., S.D. Mooney, S.J. Salipante, S. Scroggins, D. Wu, C.C. Pritchard,
Mather, C.A., S.D. Mooney, S.J. Salipante, S. Scroggins, D. Wu, C.C. Pritchard,
Mather, C.A., S.D. Mooney, S.J. Salipante, S. Scroggins, D. Wu, C.C. Pritchard,
Zhang, P., B. Bigio, F. Rapaport, S.Y. Zhang, J.L. Casanova, L. Abel, B. Boisson, and Y. Itan. 2018a. PopViz: a webserver for visualizing minor allele frequencies and damage prediction scores of human genetic variations. Bioinformatics. 34:4307–4309. https://doi.org/10.1093/bioinformatics/bty536
Zhang, Q., B. Boisson, V. Béziat, A. Puel, and J.L. Casanova. 2018b. Human hyper-IgE syndrome: singular or plural? Mamm. Genome. 29:603–617. https://doi.org/10.1007/s00335-018-9767-2
Figure S1. Schematic representation of the exon-trapping assay for nonsense or frameshift variants. Schematic representation of variants with alternative transcripts, based on the results of exon-trapping assays. The WT is shown in gray, and the variants are shown in pink and green (T1 in pink and T2 in green). T1 (pink) is the canonical transcript, and T2 (green) is the alternative transcript.

Asano et al. Journal of Experimental Medicine
Defining the mechanism of STAT3 dominance in AD-HIES

https://doi.org/10.1084/jem.20202592
Figure S2. Deep RNA-seq for the identification of novel alternative splicing in HIES patient cells with STAT3 nonsense variants. (A) RNA-seq results using EBV-B cells derived from HIES patient with STAT3 nonsense or frameshift variants (P3: R518*). P3 is already shown to have alternative splicing in our exon-trapping system. D427ins17 was previously confirmed to create an alternative transcript (Khourieh et al., 2019) and was used as a positive control. Red arrows represent mutated position. (B) RNA-seq result (Sashimi plot) for P3 focusing on the mutated exon. Sashimi plot analysis was demonstrated using Integrative Genomics Viewer. (C) Schematic representation of strategy for deep RNA analysis to detect the rare splicing events.

Q:18
Figure S3. Protein production and STAT3 activity for nonsense or frameshift STAT3 variants, taking into account the alternative transcripts. (A and B) Western blot of extracts from nontransfected STAT3−/− A4 cells (mock), A4 cells transfected with pCMV6 EV, the STAT3 WT allele, or the STAT3 variant alleles of interest. Total extracts from nontransfected or transfected STAT3−/− A4 cells after treatment (+) with 50 ng/ml sIL-6 for 20 min. All extracts probed were with mAbs specific for p-STAT3 or the N-terminal (N-ter) part of the STAT3 protein. (C) Luciferase assay on HEK293T cells, which have endogenous STAT3, cotransfected with the canonical transcript and an alternative transcript, according to the ratio estimated in the exon-trapping assay (98% D371Lfs*X14[T1] and 2% D371Lfs*X14 due to IVS12 retention [T2]; 80% R518* [T1] and 20% R518Tfs* [T2]; 83% W562* [T1] and 17% E552_W562del [T2]; 92% S701Kfs*X17 [T1] and 8% S701-D732delinsN [T2]; 52% S381* [T1] and 48% S381Pfs*2 [T2]; and 81% S381* [T1] and 19% S381Pfs*2 [T2]), together with the pGL4.47 reporter construct and an expression vector for Renilla luciferase. T1 represents the canonical transcript, and T2 represents alternative transcripts. The red dashed line represents activity after stimulation in EV-transformed cells, and the blacked dashed line represents 50% of this activity. Following cotransfection in the ratios found in exon-trapping assays, all the variants were found to be DN. Each experiment (A–C) was independently performed twice.
Summary of 150 reported and new STAT3 variants in HIES. Schematic representation of all 150 previously reported and new STAT3 variants (as of July 1, 2020). Variants in black have been reported before, and those in blue are the new nonsense and frameshift variants identified in our cohorts. The variants in red are the six alleles validated experimentally before this work. Missense variants are described in the upper part of the figure, and nonsense, insertion, deletion, and splicing variants are described in the lower part.

Asano et al. Journal of Experimental Medicine
Defining the mechanism of STAT3 dominance in AD-HIES
https://doi.org/10.1084/jem.20202592
Figure S5. Analysis for 20 STAT3 variants identified as not DN in the first luciferase assay. (A) Dose-dependent mechanism of dominance for 20 STAT3 variants. Luciferase assay on HEK293T cells, which have endogenous STAT3, transfected with various amounts of WT plasmid or one of the 20 STAT3 variant plasmids corresponding to variants identified as not DN in Fig. 5B (both not-DN and weak-DN variants were assigned to this category here). The red dashed line represents stimulated activity in EV-transformed cells, and the black dashed line represents 50% of this level of activity. Five variants displayed dose-dependent negative dominance, with a strong or weak DN effect (F384C, D369_L370delinsE, R423Q, N472D, and K642E). The remaining 15 variants (in red) had no DN effect. (B and C) 13 of the 15 variants tested generated alternative transcripts (Table S5). Luciferase assay on STAT3−/− A4 cells (B) or HEK293T cells (C) transfected with the canonical transcript (pink bar) or alternative transcripts (green bar) were performed to measure STAT3 transcriptional activity. Each experiment (A–C) was independently performed twice.
Tables S1–S8 are provided online as separate Excel files. Table S1 shows a summary of the infectious and clinical phenotypes of patients with STAT3 nonsense or frameshift variants in our cohort. Table S2 lists canonical and alternative transcripts of STAT3 nonsense or frameshift variants in HIES. Table S3 is a summary of the 150 previously reported and new STAT3 variants and represents all the mutational and functional features studied. Table S4 lists the correlation between variant consequences and their localization by domain for the 150 previously reported and new STAT3 variants from patients with HIES. Table S5 lists allele activity of canonical and alternative transcripts for 15 STAT3 missense variants. Table S6 is a summary of 14 selected STAT3 variants from the gnomAD database. Table S7 lists the summary of parameters correlated with negative selection (STAT3 and 20 genes displaying HI). Table S8 lists primer pairs for the exon-trapping assay.
Q1 OK as changed from “hyper IgE syndrome” to “hyper-IgE syndrome” in the title and text?

Q2 To make papers more readable and informative, we request that all proper gene names be italicized. Only actual gene names should be italicized, to distinguish them from gene products of the same or similar name; ad hoc designations for genes; gene segments; and gene clusters, families, complexes, or groups. With this in mind, please check that the proper terms have been italicized throughout.

Q3 Supplemental table citations were renumbered; please check carefully.

Q4 Do edits to the sentence beginning with “RNA-seq data from P3 EBV-B cells” preserve your intent?

Q5 Do edits to the sentence beginning with “Serum Ig levels were 2,180 mg/dl” preserve your intent?

Q6 OK to define abbreviation "WES" as "whole-exome sequencing" at first mention?

Q7 OK to spell out abbreviation "DHR" as "dihydrorhodamine" and "Th17" as “T helper cell type 17”?

Q8 Please expand “InDel”, if possible.

Q9 Do edits to the sentence beginning with “EBV-B cells from healthy controls” preserve your intent?

Q10 Accession numbers GSE153886, GSM4658112, GSM4658118, and GSM4658124 appears to be private. Please release these prior to publication, as required by journal policy.

Q11 I could not verify SRA accession no. PRJNA728716. Please verify this number and confirm that the data are accessible (i.e., not private).

Q12 Please provide affiliation information for those who provided materials. Include institution name, city, and state or country location (if not USA).

Q13 Please define the red “M” in Fig. 1.

Q14 For figures that contain error bars (e.g., Figs. 2-5, 7, S3, and S5), please explain what these error bars represent in the figure legends.

Q15 Please define “Mut” in the Fig. 3 legend.

Q16 Please confirm or correct expansions of abbreviations in figure legends.

Q17 Please define NT in the Figure 5 legend.

Q18 In Fig. S2 C, 6, should “pb” be “bp”?

Q19 Should STAT3 in the titles of Tables S4 and S5 be italicized, as it is in other supplemental table titles and in article body?
If you prefer to submit your corrections by annotating the PDF proof, please click the link below. You will be redirected to a system where you can download and then submit your annotated PDF. Please note that, once you click the link below, you will no longer have access to the online editing system.

[Annotate PDF]