*Twin Research and Human Genetics* (2026), **29**, 7–19
doi:10.1017/thg.2026.10054

CAMBRIDGE UNIVERSITY PRESS

**Article**

# How Genomic and Structural Context Could Shape JAK-STAT Variant Pathogenicity

Markus Hoffmann[1] and Hye Kyung Lee[2]

[1]Department of Biochemistry and Molecular & Cellular Biology, Georgetown University Medical Center, Washington, DC, USA and [2]National Institute of Diabetes, Digestive and Kidney Diseases (NIDDK), US National Institutes of Health, Bethesda, MD, USA

## Abstract

The Janus kinase (JAK)-Signal Transducer and Activator of Transcription (STAT) pathway is essential for cellular signal transduction, regulating immune responses, hematopoiesis, and cell proliferation. Dysregulation of JAK-STAT signaling due to genetic variations, particularly missense mutations, has been implicated in autoimmune disorders, cancers, and hematological malignancies. This study investigates missense mutations in JAK and STAT genes, focusing on disease-associated single nucleotide polymorphisms (SNPs) and ClinVar benign variants identified in the All of Us and COSMIC databases. We analyzed the distribution of these mutations across functional domains, their structural localization, and biochemical properties. We identified mutation hotspots within specific domains, highlighting their correlation with disease phenotypes. Structural mapping revealed that disease-associated SNPs predominantly localize in linker regions and at the boundaries of secondary structures, suggesting a significant impact on folding, stability, and function of JAK and STAT proteins. Additionally, we examined the genomic context of mutations and identified vulnerable sequences; for example, 'GATC'. Furthermore, our analysis found no predominant association between potential CRISPR-Cas9 target sites and ClinVar benign/disease-associated SNPs. The analysis of amino acid sequence patterns surrounding mutations uncovered an enrichment of hydrophobic residues leucine (Leu), isoleucine (Ile), methionine (Met), and phenylalanine (Phe) in close proximity to disease-associated mutations. Our findings emphasize the importance of structural and biochemical context in determining pathogenicity. In this study, we provide a bioinformatic strategy for refining variant classification and understanding the roles of JAK-STAT pathway mutations in disease.

**Keywords:** JAK-STAT pathway; Missense mutations; SNPs; Protein structure; Disease-associated variants; COSMIC; All of Us database; ClinVar

(Received 17 November 2025; revise received 17 February 2026; accepted 5 March 2026; First Published online 31 March 2026)

The Janus kinase (JAK)-Signal Transducer and Activator of Transcription (STAT) signaling pathway is crucial for gene regulation in common and lineage-specific genetic programs (Suppl. Fig. 1) (Brooks & Putoczki, 2020; X. Hu et al., 2021; Jankowski et al., 2021; Lee et al., 2021; Shillingford, 2002; Xue et al., 2023). Four known JAKs (JAK1, JAK2, JAK3, and TYK2) and seven STATs (STAT1, STAT2, STAT3, STAT4, STAT5A, STAT5B, and STAT6) collectively form the JAK and STAT gene families. The members of these gene families have crucial functions such as immune modulation, cell proliferation, and hematopoiesis (Brooks & Putoczki, 2020; X. Hu et al., 2021; Lee et al., 2021; Xue et al., 2023). The JAK and STAT genes carry mutations, such as single nucleotide polymorphisms (SNPs), representing the most prevalent form of genetic variation among individuals (L. X. Shen et al., 1999). These variations are characterized based on their genomic location and/or their potential influence on gene expression or function (Chu & Wei, 2019). In noncoding regions, such as promoters and enhancers, SNPs can play essential roles in gene regulation (Hecker et al., 2023; Hoffmann et al., 2023, 2025; Lee et al., 2018; Peña-Martínez & Rodríguez-Martínez, 2024). Within coding regions, SNPs are further classified into synonymous variants, which preserve the amino acid sequence and are functionally silent and evolutionarily neutral, and nonsynonymous (missense) variants, which change amino acids and can influence protein structure and/or function (Chu & Wei, 2019; Lio et al., 2025; Tsoy et al., 2024). It is well established that dysregulation of the JAK and STAT genes, through amino acid changes and altered regulatory element activity, can cause diverse pathophysiological outcomes such as autoimmune diseases, cancer, and infectious diseases (Erdogan et al., 2022; Hennighausen & Lee, 2020; Hoffmann, Willruth et al., 2024). Missense mutations drive functional changes of proteins by altering protein stability, disrupting protein-protein interactions, or compromising enzymatic functions, thus serving as potent drivers of disease (Teng et al., 2010).

Some mutations on JAK and STAT genes frequently lead to constitutive activation of the JAK-STAT pathway, a hallmark of various hematological malignancies. A well-known example is the $JAK2^{Val617Phe/V617F}$ mutation, which is highly prevalent in myeloproliferative neoplasms, occurring in approximately 90–95% of polycythemia vera (PV) cases and 50–60% of individuals with

**Corresponding author:** Markus Hoffmann; Email: mh2437@georgetown.edu

**Cite this article:** Hoffmann M and Lee HK. (2026) How Genomic and Structural Context Could Shape JAK-STAT Variant Pathogenicity. *Twin Research and Human Genetics* **29**: 7–19, https://doi.org/10.1017/thg.2026.10054

© The Author(s), 2026. Published by Cambridge University Press on behalf of International Society for Twin Studies. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

Check for updates

essential thrombocythemia and primary myelofibrosis (Perner et al., 2019; Rampal et al., 2014). This gain-of-function mutation enhances JAK-STAT signaling even in the absence of cytokine stimulation, thereby driving uncontrolled cell proliferation and survival. Similarly, activating mutations in JAK1 and JAK3 have been identified in T-cell acute lymphoblastic leukemia (T-ALL), where they contribute to persistent JAK-STAT pathway activation (Girardi et al., 2018; Waldmann, 2017). Beyond hematological malignancies, dysregulation of the JAK-STAT pathway has been implicated in solid tumors and autoimmune diseases, underscoring its broader significance in human pathology (Łączak et al., 2022; O'Shea et al., 2015). Mutations in STAT3, for instance, are linked to increased tumor invasiveness and poor clinical outcomes across multiple cancer types (Deng et al., 2022; Klein et al., 2021).

In our previous work (Hoffmann & Hennighausen, 2025), we conducted a large-scale survey of missense mutations within the JAK and STAT genes across two major repositories — the All of Us database (The All of Us Research Program Genomics Investigators et al., 2024), which captures a broad range of genetic variation from the general but mostly healthy population in the United States, and the COSMIC (Bamford et al., 2004; Sondka et al., 2024) database, a leading resource for somatic mutations in cancer. Our investigation identified hundreds of unique amino acid–changing variants, some of which had been reported as disease-associated. This analysis provided a resource cataloging the breadth and frequency of missense mutations in JAK-STAT genes. Building on these findings, this study aims to focus explicitly on the identified disease-causing (by literature) SNPs and ClinVar (Landrum et al., 2014) benign/likely benign SNPs and their distribution among specific functional domains of JAK and STAT proteins. We assess whether certain functional domains are more frequently affected and how these mutations correlate with disease. By mapping the structural distribution of disease-associated SNPs, we determine whether they predominantly occur on the protein surface or within its core. Additionally, we analyze specific target sequences, CRISPR target proximity, and amino acid composition patterns to identify shared vulnerabilities between disease-causing and ClinVar benign SNPs that may explain why specific regions are more mutation- or disease-prone.

## Materials and Methods

### *All of Us Data Explorer*

The All of Us Research Program gathers health and genomic data from participants residing in the U.S. We accessed the All of Us Controlled Tier Dataset v7 (All of Us Research Program Genomics Investigators et al., 2024) (encompassing 413,000 participants) through the Data Browser to examine SNPs. We focused on missense mutations in the JAK-STAT gene families, specifically SNPs that alter amino acids, and we analyzed their frequency across various demographic groups. All data were anonymized according to program protocols. In accordance with All of Us guidelines, we included only missense mutations (excluding all other SNP types) identified in at least 20 participants and previously associated with a disease in the literature. The cut-off of 20 is dictated by the policy of All of Us due to privacy concerns. We focused only on missense mutations because they change one amino acid to another and therefore could have a significant impact on protein function (Pal & Moult, 2015).

### *COSMIC (Catalogue of Somatic Mutations in Cancer)*

The Catalogue of Somatic Mutations in Cancer (COSMIC) database (https://cancer.sanger.ac.uk/cosmic; Bamford et al., 2004; Sondka et al., 2024) tracks somatic mutations identified in cancer. Using COSMIC v100 (>1,000,000 tumor samples), we examined missense mutations in the JAK-STAT pathway, focusing on those found in cancer samples. COSMIC details the mutational spectrum, tissue distribution, and associated cancers for each SNP, enabling direct comparison with findings from the All of Us cohort. We extracted mutations using COSMIC's online tools, filtering by mutation type. For each SNP, we noted the number of samples carrying the mutation and referenced disease associations from existing literature. To align with the All of Us approach, we included only missense mutations (excluding all other SNP types) present in at least 20 tumor samples and that were previously associated with a disease in the literature. We used the cut-off of 20 as dictated by the policy of All of Us due to privacy concerns to have a consistent filter between the two databases.

### *Obtaining SNPs That are Disease-Associated in the Published Literature and SNPs Classified as Benign in ClinVar*

We used the identified SNPs and literature from Hoffmann and Hennighausen (2025). We used the Athena – OHDSI Vocabularies Repository database (https://athena.ohdsi.org/search-terms/start) to generalize disease terms (Figure 1, Suppl. Table 1) (Reich et al., 2024).

We used the NCBI Entrez E-utilities API (https://eutils.ncbi.nlm.nih.gov/entrez/eutils/esearch.fcgi, with parameter "db": "clinvar") to filter all missense mutations from All of Us and COSMIC for only mutations categorized as benign by ClinVar.

The Python code can be found at GitHub: https://github.com/Firestar93/JAKSTAT_missenseSNPs_properties

### *Enzyme Cut Sites and CRISPR Sites on the Nucleotide Sequence and Amino Acid Composition Analysis on the Amino Acid Sequences*

We downloaded the mRNA sequences from the CCDS database (Farrell et al., 2014; Harte et al., 2012; Pruitt et al., 2009; Pujar et al., 2018) for the following genes: *STAT1* (CCDS2309.1), *STAT3* (CCDS32656.1), *STAT4* (CCDS2310.1), *STAT5B* (CCDS11423.1), *JAK2* (CCDS6457.1), *JAK3* (CCDS12366.1), and *TYK2* (CCDS12236.1).

We used a house-made Python script to detect enzyme cut sites (using the Biopython package, specifically the Bio. Restriction module (Cock et al., 2009)) and the following CRISPR-Cas sites: SpCas9 recognizes the 5"-NGG-3" sequence, while SaCas9 targets 5"-NNGRRT-3". NmCas9 requires the 5"-NNNNGATT-3" PAM, and St1Cas9 recognizes 5"-NNAGAAW-3". Similarly, St3Cas9 targets 5"-NGGNG-3", and CjCas9 recognizes the 5"-NNNNACA-3" sequence. FnCas9 operates with a 5"-YG-3" PAM, whereas TdCas9 exhibits specificity for 5"-NAAAW-3". SpCas9-NG (SpG) has a relaxed PAM requirement of 5"-NG-3", broadening its targeting scope. Lastly, SpRY functions as a near-PAM-less variant, with a preference for NRN sequences. To associate enzyme sites and CRISPR sites to a mutation, we stretch a window of 20 bps in each direction of the mutation since it was reported that 20 bps ensures that any potential restriction sites are adequately captured (Wang, 2018). We used the UpSet Python library for visualizations (Lex et al., 2014).

We used house-made Python scripts to detect amino acid combinations close (3 amino acids upstream and downstream) to ClinVar benign and disease-associated mutations using the Biopython package.

The code can be found at GitHub: https://github.com/Firestar93/JAKSTAT_missenseSNPs_properties

https://doi.org/10.1017/thg.2026.10054 Published online by Cambridge University Press

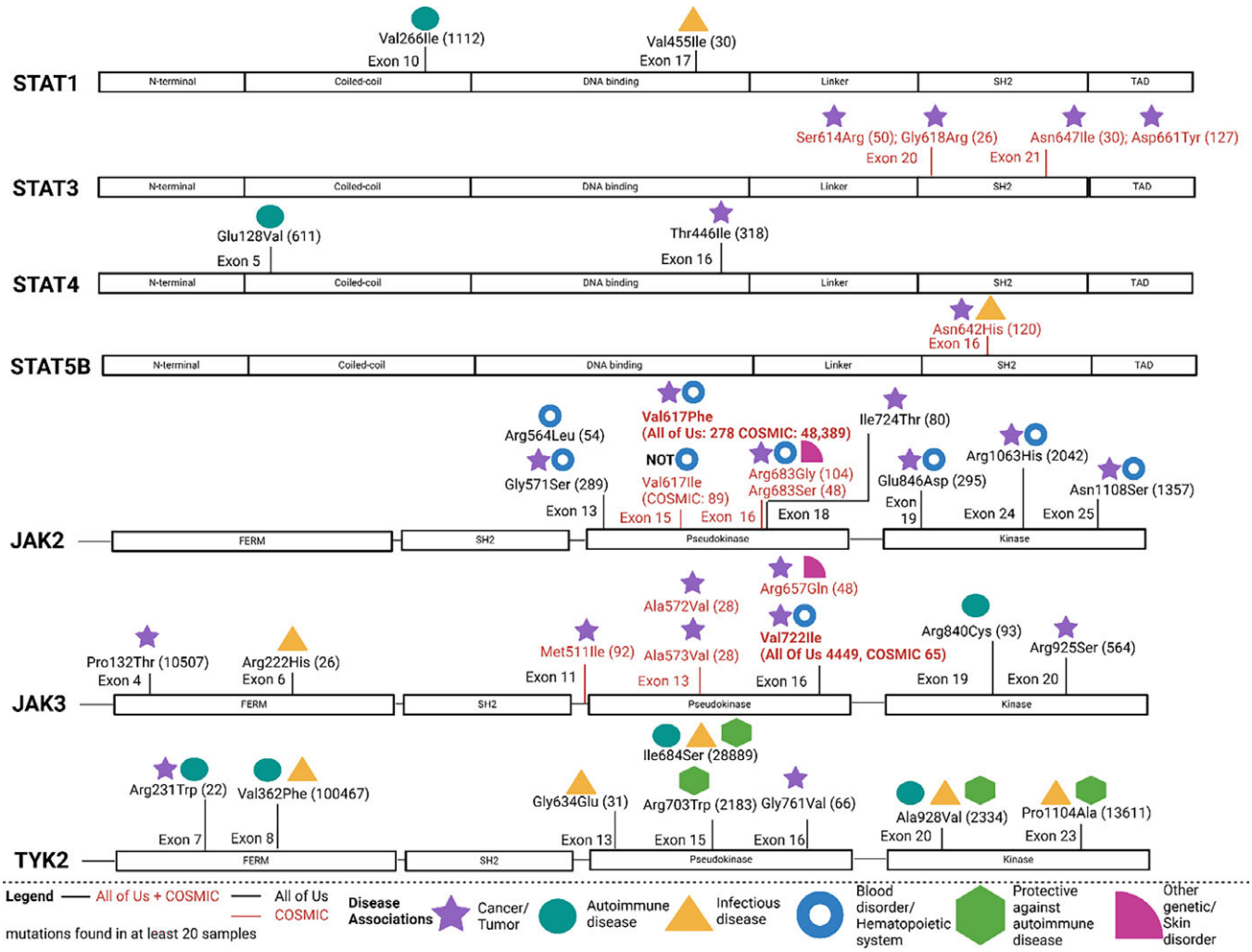



**Figure 1.** Domain-specific distribution of missense mutations in JAK and STAT proteins and their associations with disease. The schematic representation of JAK (JAK2, JAK3, TYK2) and STAT (STAT1, STAT3, STAT4, STAT5B) proteins highlights the locations of missense mutations identified in the All of Us and COSMIC databases. Symbols indicate disease associations, including autoimmune diseases (turquoise circles), cancer/tumor (purple stars), infectious diseases (yellow triangles), blood disorders/hematopoietic system involvement (blue donuts), protective mutations against autoimmunity (green hexagons), and other genetic disorder order skin disorder (dark pink quarter of a circle). Mutations found in at least 20 individuals are labeled, with mutations found in All of Us (black font) or COSMIC (red font) and mutations found in All of Us and COSMIC are highlighted in bold red. This visualization provides insight into mutation clustering within functional domains. Protein domains are annotated as follows: STAT proteins include the N-terminal, coiled-coil, DNA-binding, linker, Src homology 2 (SH2), and transactivation (TAD) domains, while JAK proteins include the FERM (For protein 4.1, Ezrin, Radixin, and Moesin), SH2, pseudokinase, and kinase domains.

### *Visualization of the Protein Structures*

We used the ChimeraX tool for the visualization of the protein structures (Goddard et al., 2018; Meng et al., 2023; Ucsf Chimerax Pettersen et al., 2021). We used ALPHAFOLD MONOMER V2.0 (Jumper et al., 2021) predicted structures for the following proteins: STAT1 (AF-P42224-F1-model_v4), STAT3 (AF-P40763-F1-model_v4), STAT4 (AF-Q14765-F1-model_v4), STAT5B (AF-P51692-F1-model_v4), JAK2 (AF-O60674-F1-model_v4), JAK3 (AF-P52333-F1-model_v4), and TYK2 (AF-P29597-F1-model_v4). ChimeraX files for interactive visualization can be found at figshare: https://doi.org/10.6084/m9.figshare.28597121.

## Results

To better understand the characteristics of JAK and STAT mutations in terms of their location within the nucleotide and amino acid sequence, we focused on their distribution across protein domains and disease associations for STAT1, STAT3, STAT4, STAT5B, JAK2, JAK3, and TYK2. We could not identify any missense variants in STAT2, STAT5A, STAT6, and JAK1 that satisfied the All of Us requirements of at least 20 individuals harboring them. We identified domains particularly susceptible to disease-related alterations. We further assessed where they are located to evaluate their effects on protein structure and stability. Beyond structural localization, we analyzed the genomic context of these mutations, examining their proximity to specific feasible sequences and CRISPR target sites to assess potential regulatory influences and editing feasibility. Finally, we investigated conserved amino acid sequence patterns near disease-associated mutations to identify motifs contributing to mutation susceptibility.

### *Mutation Hotspots in JAK-STAT Proteins and Their Disease Relevance*

We visualized mutations that were disease-associated with autoimmune disease, cancer, chronic disease, etc. (Figure 1, Suppl.

https://doi.org/10.1017/thg.2026.10054 Published online by Cambridge University Press

Table 1) and ClinVar benign (Suppl. Fig. 2) from the All of Us and COSMIC database and investigated distinct patterns in the distribution of disease-associated SNPs across JAK and STAT proteins to highlight specific domains linked to various diseases. Autoimmune disorders were mainly associated with mutations in the coiled-coil domains of STAT1 (Uzel et al., 2013) and STAT4 (Saevarsdottir et al., 2022) critical for dimerization and transcriptional activity. In contrast, cancer-related mutations were predominantly found in the SH2 domains of STAT3 (Cheon et al., 2022; D. Kim et al., 2021; Kristensen et al., 2014; Olson et al., 2020; Ramsey et al., 2023; Rivero et al., 2021; M. Shen, 2023; Yan et al., 2015) and STAT5B (Freiche et al., 2022; Z. Hu et al., 2023; Yin et al., 2023), suggesting that alterations in these key signaling interfaces contribute to tumorigenesis. Hematological malignancies and blood disorders are strongly associated with SNPs in the Pseudokinase (Arai et al., 2023; Delio et al., 2023; Eichstaedt et al., 2020; Haji Paiman et al., 2024; Krah et al., 2023; Panovska-Stavridis et al., 2016) and kinase (Kapralova et al., 2016; Maaziz et al., 2023; Tun et al., 2022) domains of JAK2, regions essential for modulating JAK-STAT signaling. We can also observe a plethora of literature associating the JAK2 Pseudokinase (Bahar et al., 2016; Bourrienne et al., 2024; Carreño-Tarragona et al., 2021; Choi et al., 2024; Gupta, Varma, Kumar et al., 2023; Gupta, Varma, Sreedharanunni et al., 2023; Hassan et al., 2022; Krah et al., 2023; Lin et al., 2019; Mambet et al., 2018; Pace et al., 2023; Patchell et al., 2024; Puli'uvea et al., 2024; Roncero et al., 2016; Schulze et al., 2019; Skoczen et al., 2020; Veitia & Innan, 2022; R. Z. Xu et al., 2022; Yongchao Zhang et al., 2024) and kinase (Benton et al., 2019; Kapralova et al., 2016; Mambet et al., 2018; Schulze et al., 2019) domains to cancer. Similarly, cancer-related mutations were frequently observed in the Pseudokinase domain of JAK3 (Agarwal et al., 2015; Bergmann et al., 2014; Bouchekioua et al., 2014; de Martino et al., 2013; Ehrentraut et al., 2016; Koo et al., 2012; Rivera-Munoz et al., 2018; Sato et al., 2008; Sim et al., 2017; L. Xu et al., 2016), reinforcing its role in malignant transformation. Infectious diseases, on the other hand, appeared to be linked to mutations in the FERM domain of JAK3 (Zhong et al., 2017), the DNA binding domain of STAT1 (Uzel et al., 2013), and all over TYK2 (Kerner et al., 2021, 2019; Ogishi et al., 2022), which play a crucial role in cytokine receptor binding and immune response regulation. Interestingly, mutations were associated with autoimmunity (Li et al., 2013; López-Isac et al., 2016; Motegi et al., 2019) and protective against autoimmunity (Diogo et al., 2015; Enerbäck et al., 2018; Jensen et al., 2023; Motegi et al., 2019) in TYK2, suggesting that variations in these regions may provide resilience against autoimmune diseases, which is an ongoing field of study (Molitor et al., 2025; Syed et al., 2025).

### *Examining the Relationship Between Structural Alterations in JAK-STAT Proteins and the Pathogenic Potential of Mutations*

Next, we investigated the protein structure using the AlphaFold3 AI-predicted model to determine where the mutations impact structural alterations. We observed that disease-associated SNPs were more frequently found in linker regions connecting secondary structural elements, such as between alpha helices and beta sheets. When these mutations occurred within an alpha helix or beta sheet, they were predominantly located at the boundary of the structure, with only rare occurrences in the middle (Figure 2, Figure 3). This suggests that mutations in transition regions may have a more significant impact on protein dynamics and function, potentially altering folding, stability, or interactions with other molecules. In contrast, benign mutations were more often embedded within well-defined secondary structures, but rarely in linker regions. Furthermore, we observed that disease-associated mutations were predominantly located within the interior of the protein's 3D structure, suggesting that these variants may impact structural integrity or disrupt key protein-protein interactions. Conversely, benign mutations were more commonly found on surface-exposed regions of the protein (i.e., an amino acid that faces the outside of the protein) (Suppl. Fig. 3).

### *Conserved Amino Acid Patterns in Proximity to Disease-Associated and ClinVar Benign Variants*

We further compared amino acid patterns surrounding the ClinVar benign and disease-associated mutations (three amino acids upstream and three amino acids downstream of the mutated site (see Materials and Methods, Figure 4). This comparative analysis of amino acid patterns in benign and disease-associated variants revealed distinct compositional differences across disease-associated and ClinVar benign variants. In the single-residue analysis (top panel), certain amino acids, such as valine (Val), glutamic acid (Glu), and methionine (Met), exhibited higher frequencies in proximity to disease-associated variants compared to benign variants, while others, such as serine (Ser), alanine (Ala), tyrosine (Tyr), and arginine (Arg), appeared more frequently in benign variants.

At the dipeptide level (second panel), several combinations were more prominent in proximity to disease-associated mutations such as (1) leucine and glutamic acid (Leu+Glu), (2) leucine and leucine (Leu+Leu), (3) aspartic acid and leucine (Asp, Leu), and (4) arginine and glutamic acid (Arg+Glu). Some combinations were uniquely present close to disease-associated variants: (1) phenylalanine and methionine (Phe+Met), (2) arginine and valine (Arg+Val), and (3) glutamic acid and methionine (Glu+Met).

Expanding to combinations of three amino acids out of six surrounding amino acids (third panel), we observed some combinations that are uniquely in proximity to disease-associated variants: (1) leucine, serine, leucine (Leu+Ser+Leu), (2) aspartic acid, leucine, leucine (Asp+Leu+Leu), (3) aspartic acid, serine, leucine (Asp, Ser, Leu), (4) leucine, leucine, glutamic acid (Leu+Leu+Glu), (5) leucine, isoleucine, glutamic acid (Leu+Ile+Glu), and (6) leucine, glutamic acid, aspartic acid (Leu+Glu+Asp). On the other hand, we detected the following combinations only close to benign variants (1) lysine, proline, glycine (Lys+Pro+Gly), (2) glycine, serine, tyrosine (Gly+Ser+Tyr), (3) arginine, arginine, arginine (Arg+Arg+Arg), and (4) arginine, threonine, arginine (Arg+Thr+Arg).

The observed differences in amino acid compositions surrounding disease-associated and benign variants suggest underlying structural and functional constraints that contribute to pathogenicity. A key trend is the enrichment of hydrophobic residues, particularly leucine (Leu), isoleucine (Ile), methionine (Met), and phenylalanine (Phe), in proximity to disease-associated variants. The frequent occurrence of Leu+Leu, Leu+Glu, and Leu+Glu+Asp combinations suggests that these mutations often occur within hydrophobic cores, where they may disrupt protein stability or alter packing interactions. Similarly, the presence of methionine (Met) in disease-associated motifs, such as Glu+Met and Phe+Met, points to potential disruptions in hydrophobic regions, particularly in proteins involved in enzymatic activity or membrane function. Conversely, benign variants appear to favor polar and flexible residues, with an overrepresentation of serine

https://doi.org/10.1017/thg.2026.10054 Published online by Cambridge University Press

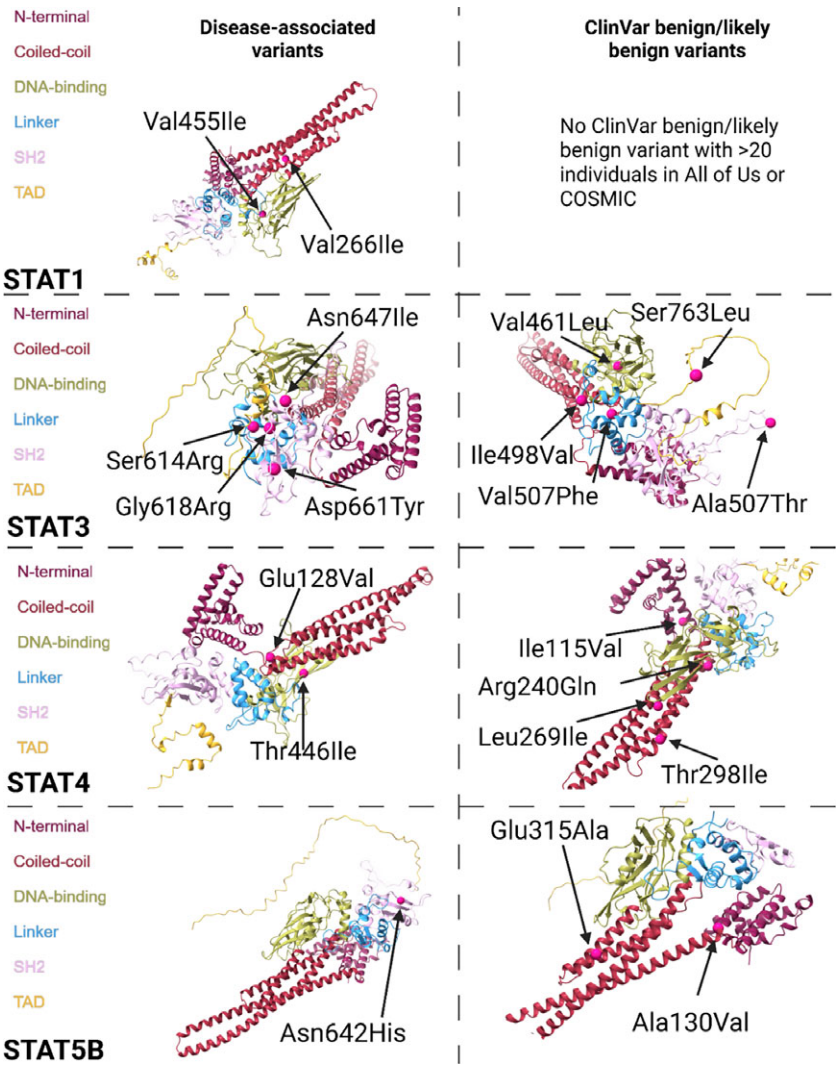



**Figure 2.** Structural analysis of disease-associated and ClinVar benign missense variants in the STAT proteins. The panel illustrates the secondary structure localization of disease-associated (left) and benign/likely benign (right) mutations mapped onto AlphaFold predicted protein structures.

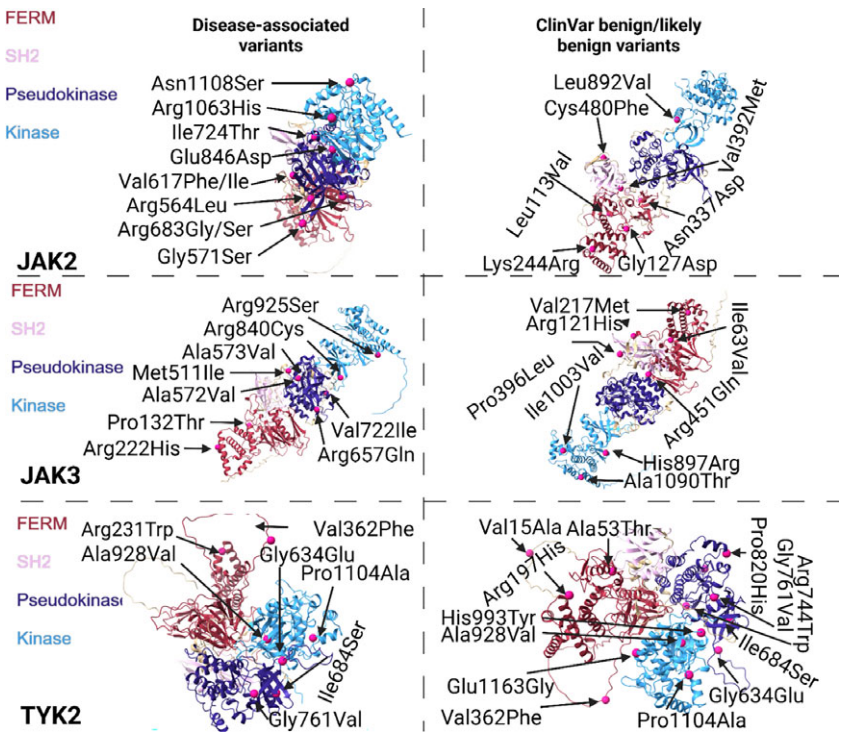



**Figure 3.** Structural analysis of disease-associated and ClinVar benign missense variants in the JAK proteins. The panel illustrates the secondary structure localization of disease-associated (left) and benign/likely benign (right) mutations mapped onto AlphaFold predicted protein structures.

https://doi.org/10.1017/thg.2026.10054 Published online by Cambridge University Press

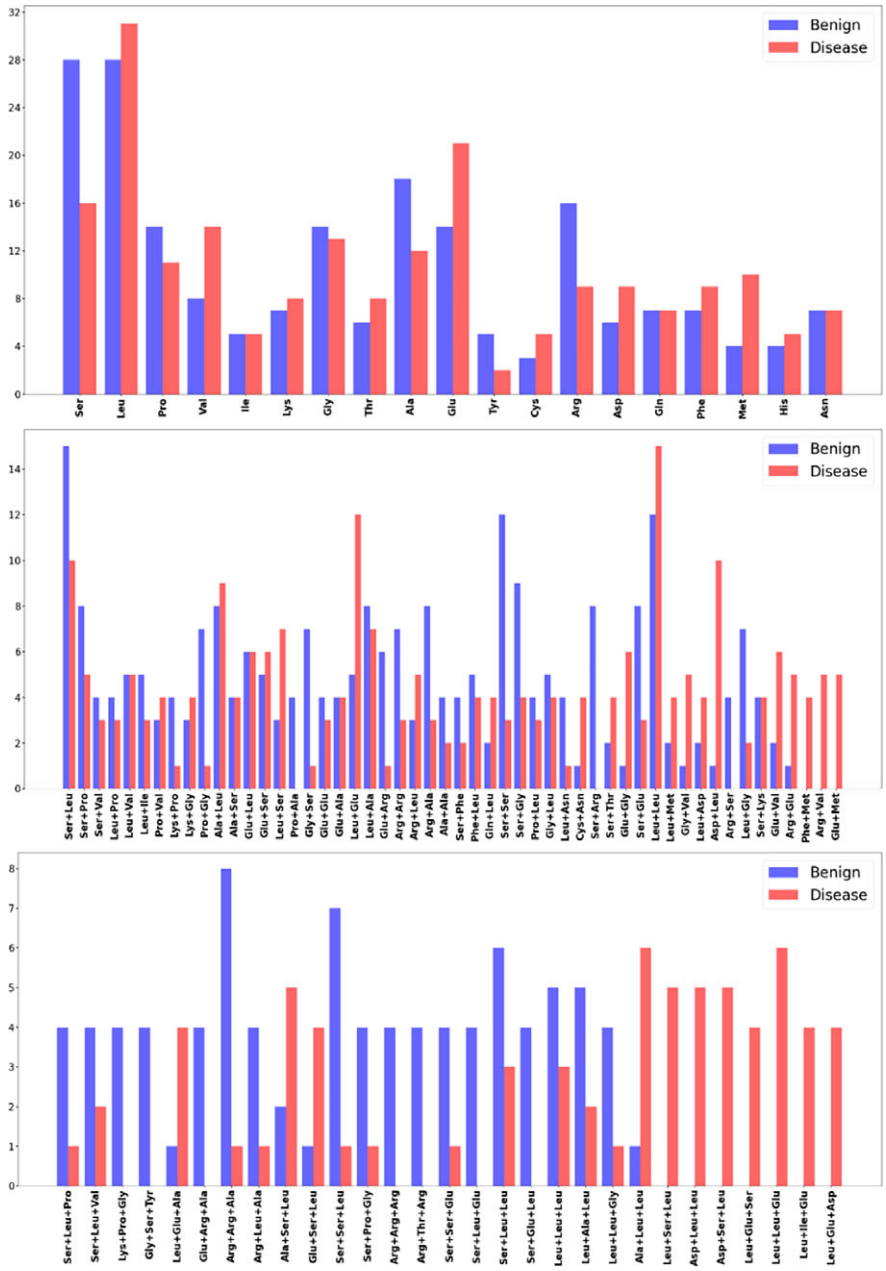



**Figure 4.** Comparative analysis of amino acid patterns (one amino acid, two amino acid combinations, and three amino acid combinations out of three upstream and three downstream of the variant in All of Us or COSMIC) in benign and disease-associated variants. Amino acid compositions for benign variants are blue, and disease variants are red.

(Ser), glycine (Gly), and threonine (Thr). The exclusive presence of Gly+Ser+Tyr and Lys+Pro+Gly in benign variants suggests that these substitutions predominantly occur in solvent-exposed loops or linker regions, where flexibility and structural adaptability mitigate the effects of mutation. Additionally, the recurrent occurrence of arginine-rich motifs (Arg+Arg+Arg and Arg+Thr+Arg) in benign variants indicates that these substitutions are likely involved in electrostatic interactions or protein-protein binding sites that can accommodate mutational changes without significant functional consequences. A second striking pattern is the enrichment of negatively charged residues (glutamic acid, Glu, and aspartic acid, Asp) near disease-associated variants, suggesting potential disruptions in salt-bridge interactions and protein stability. The presence of Arg+Glu, Asp+Leu+Leu, and Leu+Glu+Asp in disease-associated variants indicates that these mutations may destabilize electrostatic interactions or affect protein folding. In contrast, the benign variants tend to retain positively charged arginine (Arg), which is commonly involved in stabilizing protein structures or mediating protein-protein interactions. Taken together, these findings suggest that disease-associated variants frequently occur in structurally constrained regions, where mutations disrupt core hydrophobic interactions, electrostatic balance, or functional interfaces. In contrast, benign variants are more likely to appear in flexible or surface-exposed regions, where mutations are better tolerated due to the preservation of local structural dynamics. The distinct differences in amino acid preferences between disease and benign variants provide insights into the physicochemical constraints that contribute to pathogenicity and may aid in refining predictive models for variant classification.

https://doi.org/10.1017/thg.2026.10054 Published online by Cambridge University Press

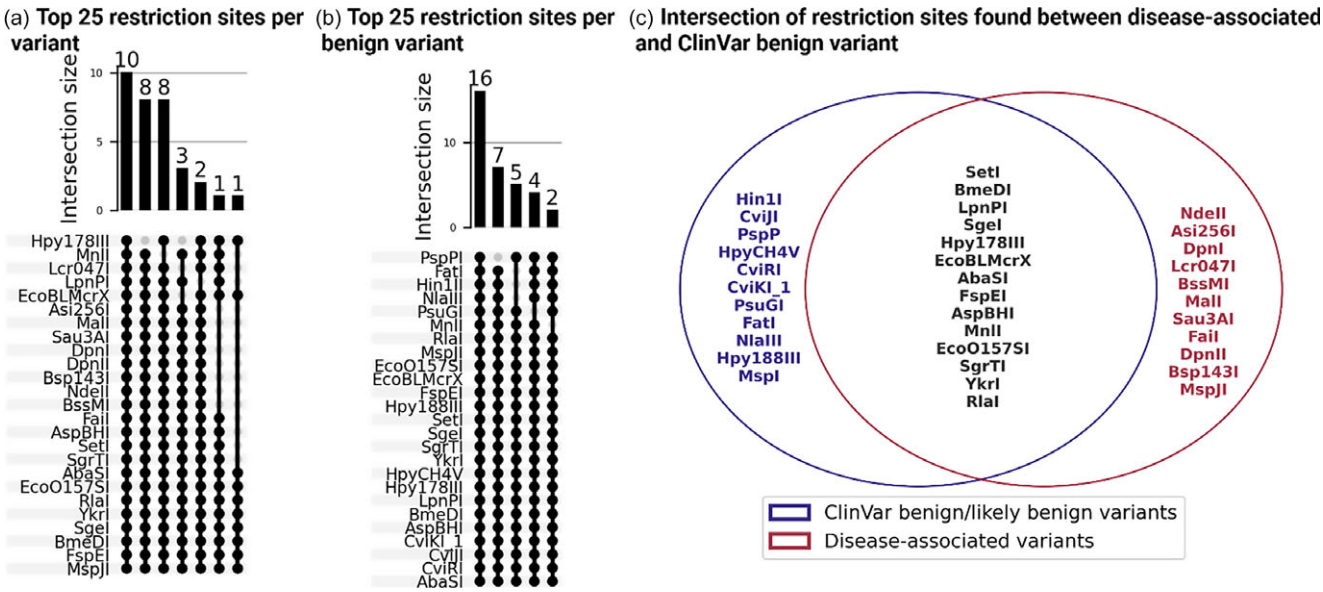



**Figure 5.** Enzyme restriction site analysis in proximity to disease-associated and ClinVar benign variants in JAK and STAT genes. (a,b) The top 25 restriction enzymes identified near disease-associated and benign variants, respectively. (c) Venn diagram illustrating the overlap of restriction sites found near disease-associated variants (red) and ClinVar benign variants (blue).

### *Analyzing the Nucleotide Sequence Near Benign and Disease-Associated Variants*

To explore the genomic context of disease-associated and ClinVar benign mutations, we analyzed the nucleotide sequences surrounding these variants, focusing on a 20 bp region around each mutation. The first part of our investigation aimed to identify patterns of nucleotide sequences and assess whether specific sequences are preferentially found near disease-associated or benign mutations (Figure 5, Suppl. Fig. 4). Our analysis identified nucleotide sequences of 625 restriction enzyme sites out of a total of 1,088 within the examined regions. Notably, several nucleotide sequences were exclusively present near disease-associated mutations but absent in ClinVar benign variants. These included the nucleotide sequences of DpnI, Asi256I, DpnII, MalI, Lcr047I, NdeII, Bsp143I, Sau3AI, FaiI, MspJI, and BssMI. We found that the nucleotide sequence 'GATC' was predominantly present near disease-associated mutations but was entirely absent in the vicinity of ClinVar benign mutations.

The second part of this analysis is to assess whether multiple Cas9 enzymes preferentially target sequences near disease-associated or ClinVar benign mutations; we analyzed the presence of Cas9 recognition sites within a 20 bp region surrounding these mutations (Figure 6, Suppl. Fig. 5). We examined a range of Cas9 enzymes, including SpCas9, SaCas9, NmCas9, St1Cas9, St3Cas9, CjCas9, FnCas9, TdCas9, xCas9, SpCas9-NG (SpG), SpRY, HiFi Cas9 (HF1), eSpCas9 (1.1), and HypaCas9 that recognized different PAM sequences (Suppl. Table 2) (Acharya et al., 2024; Du et al., 2023; Guo et al., 2019; Hibshman et al., 2024; Hou et al., 2013; Ikeda et al., 2019; H. K. Kim et al., 2020; Liang et al., 2022; Müller et al., 2016; Schmidheini et al., 2024; Slaymaker et al., 2016; Vakulskas et al., 2018; Wu et al., 2020; Yifei Zhang et al., 2020). Our analysis did not reveal a strong preference for any Cas9 protein's cut site being predominantly located near disease-associated mutations, but absent in ClinVar benign mutations. These findings suggest that while CRISPR/Cas9 target sites are present around the variants, there is no clear enrichment in disease-associated sites that would indicate preferential targetability.

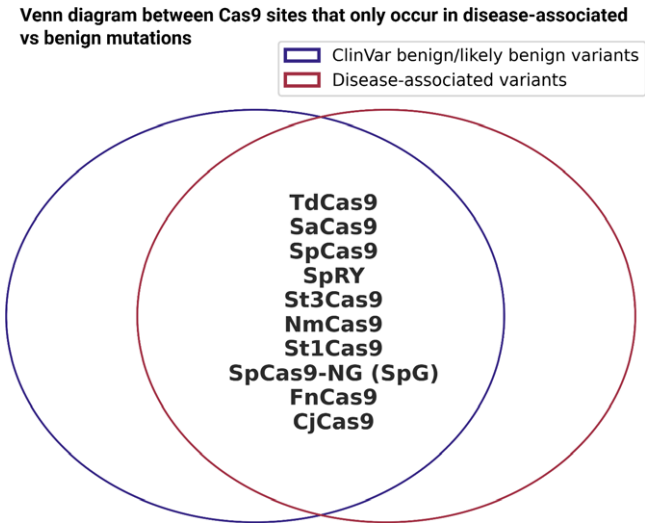



**Figure 6.** CRISPR cut site analysis in proximity to disease-associated and benign mutations. Venn diagram illustrating the overlap of Cas9 cut sites uniquely occurring in either disease-associated (red) or benign (blue) mutations.

## Discussion

This study provides a comprehensive analysis of missense mutations in the JAK-STAT pathway, highlighting differences between disease-associated and ClinVar benign variants regarding structural localization, biochemical properties, and genomic context. By integrating mutation data from the All of Us and COSMIC databases, we identified properties that may help explain why specific variants contribute to disease while others remain functionally neutral. Our findings suggest that disease-associated mutations frequently disrupt core hydrophobic interactions, electrostatic balance, or functional interfaces, whereas benign variants are more commonly found within secondary structures like helices and sheets and also in surface-exposed regions, where

https://doi.org/10.1017/thg.2026.10054 Published online by Cambridge University Press

structural constraints are less restrictive. These observations reinforce the notion that pathogenicity is not merely a function of amino acid substitution alone but is heavily influenced by protein architecture and nearby context on the nucleotide and amino acid level.

Our current understanding of disease-associated mutations is largely based on the analysis of individual SNPs, whereas real-world genetic variation often involves combinations of mutations that may interact in yet unknown ways. The lack of experimental data on epistatic interactions is a major limitation (Suppl. Text 1: Limitations and considerations) in mutation interpretation, as the functional impact of a given variant may depend on the presence of additional mutations within the same gene or pathway (Blumenthal et al., 2020; Hernández-Lorenzo et al., 2022; Hoffmann, Poschenrieder et al., 2024). While our stringent inclusion criteria (minimum 20 occurrences due to All of Us policy) allowed for robust analysis of more prevalent variants, it is important to acknowledge that rare yet highly pathogenic mutations, especially those associated with rare inherited disorders, might not be captured by this approach.

The GATC sequence is well known in bacteria as a recognition site for DNA adenine methyltransferase (Dam) and plays key roles in DNA repair, replication timing, and gene regulation (Flusberg et al., 2010). In *E. coli*, for example, GATC sites guide mismatch repair machinery to correct errors on the newly synthesized DNA strand (Horton et al., 2015). Although human cells do not use the same bacterial repair system, recent work shows that sequence context, including short motifs like GATC, can influence where replication errors occur and how repair systems handle them (Hasenauer et al., 2025). GATC-like sequences can also be recognized by certain transcription factors or occur in open chromatin regions, meaning changes nearby could alter gene expression (Mardenborough et al., 2019).

Future work could explore whether selected missense variants influence STAT dimer formation by combining state-aware structural hypotheses with independent computational estimates of interface perturbation and focused experimental validation, recognizing that dimerization is conformation-dependent and not fully captured by static structural predictions alone.

Understanding genetic variation requires considering not only individual mutations but also their broader structural and functional context. As data and experimental tools improve, distinguishing ClinVar benign from disease-associated variants will become more precise. Addressing gaps such as epistatic interactions and underreported mutations will clarify how JAK-STAT variants affect health. Integrating computational models of protein–protein interactions with high-throughput experiments could systematically investigate multi-variant and epistatic effects. In vivo studies introducing predicted single or combined SNPs into homozygous mouse lines, then crossbreeding and applying RNA-seq, ChIP-seq, and allergenic challenge assays (Gad, 1994), could reveal how complex mutation patterns shape pathway function. Clinical sample analysis will be critical to assess whether variants labeled as benign may contribute to disease in certain contexts, with our curated references offering a starting point for prioritization. Examining noncoding regulatory regions for pathogenic mutations may uncover additional mechanisms, but current datasets (All of Us, COSMIC) lack matched transcriptomic or epigenomic data, and patient samples were unavailable for this study. Future integration of genomic variation with matched expression and chromatin profiles will be essential for linking regulatory mutations to function. Combining large-scale genomics with advanced experimental systems offers a path to further investigate JAK-STAT variation and translate these insights into future therapeutic advances.

**Supplementary material.** To view supplementary material for this article, please visit https://doi.org/10.1017/thg.2026.10054.

**Data availability.** This study used data from the All of Us Research Program's controlled Tier Dataset v.7, available to authorized users on the Researcher Workbench (https://databrowser.researchallofus.org/). Data from COSMIC v100 is available at (https://cancer.sanger.ac.uk/cosmic). The Python code can be found at: https://github.com/Firestar93/JAKSTAT_missenseSNPs_properties. Intermediate results and files can be found at: https://doi.org/10.6084/m9.figshare.28597121.

**Acknowledgments.** The authors want to thank Jakub Jankowski, Lothar Hennighausen, Priscilla A. Furth, and the members of the Laboratory of Cell & Molecular Biology (LCMB), NIDDK, NIH, for their valuable input. The figures were created with Biorender.com. Parts of the figures include icons from Flaticon.com under a paid license. The text was partly rephrased using ChatGPT version 4, Grammarly, and scite.ai under a paid license. Paperpile, under a paid license, was used to collect references in the correct format. We gratefully acknowledge All of Us participants for their contributions, without whom this research would not have been possible. We also thank the National Institutes of Health's All of Us Research Program for making available the participant data, samples, and cohort examined in this study.

**Author contributions.** M.H. planned the project, executed the analysis, and wrote the manuscript. H.K. conceptualized the project and revised the manuscript. All authors read and approved the final version of the manuscript.

**Funding.** This research was supported by the Intramural Research Programs (IRPs) of the National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK) within the National Institutes of Health (NIH). The contributions of the NIH author(s) were made as part of their official duties as NIH federal employees, are in compliance with agency policy requirements, and are considered Works of the United States Government. However, the findings and conclusions presented in this paper are those of the author(s) and do not necessarily reflect the views of the NIH or the U.S. Department of Health and Human Services. Funding support is in part from Georgetown University Medical Center.

**Competing interests.** The authors declare no competing interests.

## References

**Acharya, S., Ansari, A. H., Kumar Das, P., Hirano, S., Aich, M., Rauthan, R., Mahato, S., Maddileti, S., Sarkar, S., Kumar, M., Phutela, R., Gulati, S., Rahman, A., Goel, A., Afzal, C., Paul, D., Agrawal, T., Pulimamidi, V. K., Jalali, S., … Chakraborty, D.** (2024). PAM-flexible engineered FnCas9 variants for robust and ultra-precise genome editing and diagnostics. *Nature Communications*, *15*, 5471. https://doi.org/10.1038/s41467-024-49233-w

**Agarwal, A., MacKenzie, R. J., Eide, C. A., Davare, M. A., Watanabe-Smith, K., Tognon, C. E., Mongoue-Tchokote, S., Park, B., Braziel, R. M., Tyner, J. W., & Druker, B. J.** (2015). Functional RNAi screen targeting cytokine and growth factor receptors reveals oncorequisite role for interleukin-2 gamma receptor in JAK3-mutation-positive leukemia. *Oncogene*, *34*, 2991–2999. https://doi.org/10.1038/onc.2014.243

**All of Us Research Program Genomics Investigators**, **Bick, A. G., Metcalf, G. A., Mayo, K. R., Lichtenstein, L., Rura, S., Carroll, R. J., Musick, A., Linder, J. E., Jordan, I. K., Nagar, S. D., Sharma, S., Meller, R., Basford, M., Boerwinkle, E., Cicek, M. S., Doheny, K. F., Eichler, E. E., Gabriel, S., … NIH All of Us Research Program Staff**. (2024). Genomic data in the All of Us Research Program. *Nature*, *627*, 340–346. https://doi.org/10.1038/s41586-024-07031-8

**Arai, A., Yoshimitsu, M., Otsuka, M., Ito, Y., Miyazono, T., Nakano, N., Obama, K., Nakashima, H., Hanada, S., Owatari, S., Nakamura, D., Tokunaga, M., Kamada, Y., Utsunomiya, A., Haraguchi, K., Hayashida, M., Fujino, S., Odawara, J., Tabuchi, T., … Ishitsuka, K.** (2023).

Identification of putative noncanonical driver mutations in patients with essential thrombocythemia. *European Journal of Haematology*, *110*, 639–647. https://doi.org/10.1111/ejh.13928

**Bahar, B., Barton, K., & Kini, A. R.** (2016). The role of the Exon 13 G571S JAK2 mutation in myeloproliferative neoplasms. *Leukemia Research Reports*, *6*, 27–28. https://doi.org/10.1016/j.lrr.2016.08.001

**Bamford, S., Dawson, E., Forbes, S., Clements, J., Pettett, R., Dogan, A., Flanagan, A., Teague, J., Futreal, P. A., Stratton, M. R., & Wooster, R.** (2004). The COSMIC (Catalogue of Somatic Mutations in Cancer) database and website. *British Journal of Cancer*, *91*, 355–358. https://doi.org/10.1038/sj.bjc.6601894

**Benton, C. B., Boddu, P. C., DiNardo, C. D., Bose, P., Wang, F., Assi, R., Pemmaraju, N., Kc, D., Pierce, S., Patel, K., Konopleva, M., Ravandi, F., Garcia-Manero, G., Kadia, T. M., Cortes, J., Kantarjian, H. M., Andreeff, M., & Verstovsek, S.** (2019). Janus kinase 2 variants associated with the transformation of myeloproliferative neoplasms into acute myeloid leukemia. *Cancer*, *125*, 1855–1866. https://doi.org/10.1002/cncr.32004

**Bergmann, A. K., Schneppenheim, S., Seifert, M., Betts, M. J., Haake, A., Lopez, C., Maria Murga Penas, E., Vater, I., Jayne, S., Dyer, M. J. S., Schrappe, M., Dührsen, U., Ammerpohl, O., Russell, R. B., Küppers, R., Dürig, J., & Siebert, R.** (2014). Recurrent mutation of JAK3 in T-cell prolymphocytic leukemia. *Genes, Chromosomes & Cancer*, *53*, 309–316. https://doi.org/10.1002/gcc.22136

**Blumenthal, D. B., Baumbach, J., Hoffmann, M., Kacprowski, T., & List, M.** (2020). A framework for modeling epistatic interaction. *Bioinformatics*. https://doi.org/10.1093/bioinformatics/btaa990

**Bouchekioua, A., Scourzic, L., de Wever, O., Zhang, Y., Cervera, P., Aline-Fardin, A., Mercher, T., Gaulard, P., Nyga, R., Jeziorowska, D., Douay, L., Vainchenker, W., Louache, F., Gespach, C., Solary, E., & Coppo, P.** (2014). JAK3 deregulation by activating mutations confers invasive growth advantage in extranodal nasal-type natural killer cell lymphoma. *Leukemia*, *28*, 338–348. https://doi.org/10.1038/leu.2013.206

**Bourrienne, M.-C., Loyau, S., Faille, D., Gay, J., Akhenak, S., Farkh, C., Ollivier, V., Solonomenjanahary, M., Dupont, S., Choqueux, C., Villeval, J.-L., Plo, I., Edmond, V., Ho-Tin-Noé, B., Ajzenberg, N., & Mazighi, M.** (2024). Impaired fibrinolysis in JAK2V617F-related myeloproliferative neoplasms. *Journal of Thrombosis and Haemostasis*. https://doi.org/10.1016/j.jtha.2024.07.031

**Brooks, A. J., & Putoczki, T.** (2020). JAK-STAT signalling pathway in cancer. *Cancers*, *12*, 1971. https://doi.org/10.3390/cancers12071971

**Carreño-Tarragona, G., Varghese, L. N., Sebastián, E., Gálvez, E., Marín-Sánchez, A., López-Muñoz, N., Nam-Cha, S., Martínez-López, J., Constantinescu, S. N., Sevilla, J., & Ayala, R.** (2021). A typical acute lymphoblastic leukemia JAK2 variant, R683G, causes an aggressive form of familial thrombocytosis when germline. *Leukemia*, *35*, 3295–3298. https://doi.org/10.1038/s41375-021-01298-0

**Cheon, H., Xing, J. C., Moosic, K. B., Ung, J., Chan, V. W., Chung, D. S., Toro, M. F., Elghawy, O., Wang, J. S., Hamele, C. E., Hardison, R. C., Olson, T. L., Tan, S.-F., Feith, D. J., Ratan, A., & Loughran, T. P., Jr.** (2022). Genomic landscape of TCRαβ and TCRγδ T-large granular lymphocyte leukemia. *Blood*, *139*, 3058–3072. https://doi.org/10.1182/blood.2021013164

**Choi, D. C., Messali, N., Uda, N. R., Abu-Zeinah, G., Kermani, P., Yabut, M. M., Lischer, H. E. L., Castillo Tokumori, F., Erdos, K., Lehmann, T., Sobas, M., Rao, T. N., & Scandura, J. M.** (2024). JAK2V617F impairs lymphoid differentiation in myeloproliferative neoplasms. *Leukemia*. https://doi.org/10.1038/s41375-024-02388-3

**Chu, D., & Wei, L.** (2019). Nonsynonymous, synonymous and nonsense mutations in human cancer-related genes undergo stronger purifying selections than expectation. *BMC Cancer, 19*. https://doi.org/10.1186/s12885-019-5572-x

**Cock, P. J. A., Antao, T., Chang, J. T., Chapman, B. A., Cox, C. J., Dalke, A., Friedberg, I., Hamelryck, T., Kauff, F., Wilczynski, B., & de Hoon, M. J. L.** (2009). Biopython: Freely available Python tools for computational molecular biology and bioinformatics. *Bioinformatics*, *25*, 1422–1423. https://doi.org/10.1093/bioinformatics/btp163

**de Martino, M., Gigante, M., Cormio, L., Prattichizzo, C., Cavalcanti, E., Gigante, M., Ariano, V., Netti, G. S., Montemurno, E., Mancini, V., Battaglia, M., Gesualdo, L., Carrieri, G., & Ranieri, E.** (2013). JAK3 in clear cell renal cell carcinoma: Mutational screening and clinical implications. *Urologic Oncology*, *31*, 930–937. https://doi.org/10.1016/j.urolonc.2011.07.011

**Delio, M., Bryke, C., Mendez, L., Joseph, L., & Jassim, S.** (2023). JAK2 mutations are rare and diverse in myelodysplastic syndromes: Case series and review of the literature. *Hematology Reports*, *15*, 73–87. https://doi.org/10.3390/hematolrep15010008

**Deng, M., Li, Y., Li, Y., Mao, X., Ke, H., Liang, W., Lei, X., Lau, Y.-L., & Mao, H.** (2022). A novel STAT3 gain-of-function mutation in fatal infancy-onset interstitial lung disease. *Frontiers in Immunology*, *13*, 866638. https://doi.org/10.3389/fimmu.2022.866638

**Diogo, D., Bastarache, L., Liao, K. P., Graham, R. R., Fulton, R. S., Greenberg, J. D., Eyre, S., Bowes, J., Cui, J., Lee, A., Pappas, D. A., Kremer, J. M., Barton, A., Coenen, M. J. H., Franke, B., Kiemeney, L. A., Mariette, X., Richard-Miceli, C., Canhão, H., … Plenge, R. M.** (2015). TYK2 protein-coding variants protect against rheumatoid arthritis and autoimmunity, with no evidence of major pleiotropic effects on non-autoimmune complex traits. *PLOS ONE*, *10*, e0122271. https://doi.org/10.1371/journal.pone.0122271

**Du, W., Zhu, H., Qian, J., Xue, D., Zheng, S., & Huang, Q.** (2023). Full-length model of SaCas9-sgRNA-DNA complex in cleavage state. *International Journal of Molecular Sciences*, *24*, 1204. https://doi.org/10.3390/ijms24021204

**Ehrentraut, S., Schneider, B., Nagel, S., Pommerenke, C., Quentmeier, H., Geffers, R., Feist, M., Kaufmann, M., Meyer, C., Kadin, M. E., Drexler, H. G., & MacLeod, R. A. F.** (2016). Th17 cytokine differentiation and loss of plasticity after SOCS1 inactivation in a cutaneous T-cell lymphoma. *Oncotarget*, *7*, 34201–34216. https://doi.org/10.18632/oncotarget.9107

**Eichstaedt, C. A., Verweyen, J., Halank, M., Benjamin, N., Fischer, C., Mayer, E., Guth, S., Wiedenroth, C. B., Egenlauf, B., Harutyunova, S., Xanthouli, P., Marra, A. M., Wilkens, H., Ewert, R., Hinderhofer, K., & Grünig, E.** (2020). Myeloproliferative diseases as possible risk factor for development of chronic thromboembolic pulmonary hypertension — A genetic study. *International Journal of Molecular Sciences*, *21*, 3339. https://doi.org/10.3390/ijms21093339

**Enerbäck, C., Sandin, C., Lambert, S., Zawistowski, M., Stuart, P. E., Verma, D., Tsoi, L. C., Nair, R. P., Johnston, A., & Elder, J. T.** (2018). The psoriasis-protective TYK2 I684S variant impairs IL-12 stimulated pSTAT4 response in skin-homing CD4+ and CD8+ memory T-cells. *Scientific Reports, 8*. https://doi.org/10.1038/s41598-018-25282-2

**Erdogan, F., Qadree, A. K., Radu, T. B., Orlova, A., de Araujo, E. D., Israelian, J., Valent, P., Mustjoki, S. M., Herling, M., Moriggl, R., & Gunning, P. T.** (2022). Structural and mutational analysis of member-specific STAT functions. *Biochimica et Biophysica Acta: General Subjects*, *1866*, 130058. https://doi.org/10.1016/j.bbagen.2021.130058

**Farrell, C. M., O'Leary, N. A., Harte, R. A., Loveland, J. E., Wilming, L. G., Wallin, C., Diekhans, M., Barrell, D., Searle, S. M. J., Aken, B., Hiatt, S. M., Frankish, A., Suner, M.-M., Rajput, B., Steward, C. A., Brown, G. R., Bennett, R., Murphy, M., Wu, W., … Pruitt, K. D.** (2014). Current status and new features of the Consensus Coding Sequence database. *Nucleic Acids Research, 42*, D865–D872. https://doi.org/10.1093/nar/gkt1059

**Flusberg, B. A., Webster, D. R., Lee, J. H., Travers, K. J., Olivares, E. C., Clark, T. A., Korlach, J., & Turner, S. W.** (2010). Direct detection of DNA methylation during single-molecule, real-time sequencing. *Nature Methods*, *7*, 461–465. https://doi.org/10.1038/nmeth.1459

**Freiche, V., Couronné, L., Bruneau, J., & Hermine, O.** (2022). Comment on kieslinger et al. A recurrent STAT5BN642H driver mutation in feline alimentary T cell lymphoma. *Cancers 2021*, 13, 5238. *Cancers, 14*, 4593. https://doi.org/10.3390/cancers14194593

**Gad, S. C.** (1994). The mouse ear swelling test (MEST) in the 1990s. *Toxicology*, *93*, 33–46. https://doi.org/10.1016/0300-483X(94)90163-5

**Girardi, T., Vereecke, S., Sulima, S. O., Khan, Y., Fancello, L., Briggs, J. W., Schwab, C., de Beeck, J. O., Verbeeck, J., Royaert, J., Geerdens, E., Vicente, C., Bornschein, S., Harrison, C. J., Meijerink, J. P., Cools, J., Dinman, J. D., Kampen, K. R., & De Keersmaecker, K.** (2018). The T-cell leukemia-associated ribosomal RPL10 R98S mutation enhances JAK-STAT signaling. *Leukemia*, *32*, 809–819. https://doi.org/10.1038/leu.2017.248

**Goddard, T. D., Huang, C. C., Meng, E. C., Pettersen, E. F., Couch, G. S., Morris, J. H., & Ferrin, T. E.** (2018). UCSF ChimeraX: Meeting modern

challenges in visualization and analysis. *Protein Science*, *27*, 14–25. https://doi.org/10.1002/pro.3235

**Guo, M., Ren, K., Zhu, Y., Tang, Z., Wang, Y., Zhang, B., & Huang, Z.** (2019). Structural insights into a high fidelity variant of SpCas9. *Cell Research*, *29*, 183–192. https://doi.org/10.1038/s41422-018-0128-9

**Gupta, D. G., Varma, N., Kumar, A., Naseem, S., Sachdeva, M. U. S., Sreedharanunni, S., Binota, J., Bose, P., Khadwal, A., Malhotra, P., & Varma, S.** (2023). Genomic and proteomic characterization of Philadelphia-like B-lineage acute lymphoblastic leukemia: A report of Indian patients. *Cancer*, *129*, 1217–1226. https://doi.org/10.1002/cncr.34661

**Gupta, D. G., Varma, N., Sreedharanunni, S., Abdulkadir, S. A., Naseem, S., Sachdeva, M. U. S., Binota, J., Bose, P., Malhotra, P., Khadwal, A., & Varma, S.** (2023). Evaluation of adverse prognostic gene alterations and MRD positivity in BCR::ABL1-like B-lineage acute lymphoblastic leukaemia patients in a resource-constrained setting. *British Journal of Cancer*, *129*, 143–152. https://doi.org/10.1038/s41416-023-02252-2

**Haji Paiman, N. S., Mat Nasir, N., Miptah, H. N., Saidon, N., & Abdul Monir, M.** (2024). Challenges in diagnosing polycythemia Vera in primary care: A 55-year-old Malaysian woman with atypical presentation. *The American Journal of Case Reports*, *25*, e944202. https://doi.org/10.12659/AJCR.944202

**Harte, R. A., Farrell, C. M., Loveland, J. E., Suner, M.-M., Wilming, L., Aken, B., Barrell, D., Frankish, A., Wallin, C., Searle, S., Diekhans, M., Harrow, J., & Pruitt, K. D.** (2012). Tracking and coordinating an international curation effort for the CCDS Project. *Database: The Journal of Biological Databases and Curation, 2012*, bas008. https://doi.org/10.1093/database/bas008

**Hasenauer, F. C., Barreto, H. C., Lotton, C., & Matic, I.** (2025). Genome-wide mapping of spontaneous DNA replication error-hotspots using mismatch repair proteins in rapidly proliferating *Escherichia coli*. *Nucleic Acids Research, 53*. https://doi.org/10.1093/nar/gkae1196

**Hassan, N. M., Abdellateif, M. S., Radwan, E. M., Hameed, S. A., Desouky, E. D. E., Kamel, M. M., & Gameel, A. M.** (2022). Prognostic significance of CRLF2 overexpression and JAK2 mutation in Egyptian pediatric patients with B-precursor acute lymphoblastic leukemia. *Clinical Lymphoma, Myeloma & Leukemia*, *22*, e376–e385. https://doi.org/10.1016/j.clml.2022.01.006

**Hecker, D., Lauber, M., Behjati Ardakani, F., Ashrafiyan, S., Manz, Q., Kersting, J., Hoffmann, M., Schulz, M. H., & List, M.** (2023). Computational tools for inferring transcription factor activity. *Proteomics*, e2200462. https://doi.org/10.1002/pmic.202200462

**Hennighausen, L., & Lee, H. K.** (2020). Activation of the SARS-CoV-2 receptor Ace2 through JAK/STAT-dependent enhancers during pregnancy. *Cell Reports*, *32*, 108199. https://doi.org/10.1016/j.celrep.2020.108199

**Hernández-Lorenzo, L., Hoffmann, M., Scheibling, E., List, M., Matías-Guiu, J. A., & Ayala, J. L.** (2022). On the limits of graph neural networks for the early diagnosis of Alzheimer's disease. *Scientific Reports*, *12*, 17632. https://doi.org/10.1038/s41598-022-21895-8

**Hibshman, G. N., Bravo, J. P. K., Hooper, M. M., Dangerfield, T. L., Zhang, H., Finkelstein, I. J., Johnson, K. A., & Taylor, D. W.** (2024). Unraveling the mechanisms of PAMless DNA interrogation by SpRY-Cas9. *Nature Communications*, *15*. https://doi.org/10.1038/s41467-024-45858-6

**Hoffmann, M., & Hennighausen, L.** (2025). Spotlight on amino acid changing mutations in the JAK-STAT pathway: From disease-specific mutation to general mutation databases. *Scientific Reports*, *15*. https://doi.org/10.1038/s41598-025-xxxx-x

**Hoffmann, M., Poschenrieder, J. M., Incudini, M., Baier, S., Fritz, A., Maier, A., Hartung, M., Hoffmann, C., Trummer, N., Adamowicz, K., Picciani, M., Scheibling, E., Harl, M. V., Lesch, I., Frey, H., Kayser, S., Wissenberg, P., Schwartz, L., Hafner, L., … Blumenthal, D. B.** (2024). Network medicine-based epistasis detection in complex diseases: Ready for quantum computing. *Nucleic Acids Research*, *52*, 10144–10160. https://doi.org/10.1093/nar/gkae620

**Hoffmann, M., Trummer, N., Schwartz, L., Jankowski, J., Lee, H. K., Willruth, L.-L., Lazareva, O., Yuan, K., Baumgarten, N., Schmidt, F., Baumbach, J., Schulz, M. H., Blumenthal, D. B., Hennighausen, L., & List, M.** (2023). TF-Prioritizer: A Java pipeline to prioritize condition-specific transcription factors. *GigaScience*, *12*, giad026. https://doi.org/10.1093/gigascience/giad026

**Hoffmann, M., Vaz, T., Chhatrala, S., & Hennighausen, L.** (2025). Data-driven projections of candidate enhancer-activating SNPs in immune regulation. *BMC Genomics*, *26*. https://doi.org/10.1186/s12864-025-11374-7

**Hoffmann, M., Willruth, L.-L., Dietrich, A., Lee, H. K., Knabl, L., Trummer, N., Baumbach, J., Furth, P. A., Hennighausen, L., & List, M.** (2024). Blood transcriptomics analysis offers insights into variant-specific immune response to SARS-CoV-2. *Scientific Reports*, *14*. https://doi.org/10.1038/s41598-024-55686-2

**Horton, J. R., Zhang, X., Blumenthal, R. M., & Cheng, X.** (2015). Structures of *Escherichia coli* DNA adenine methyltransferase (Dam) in complex with a non-GATC sequence: Potential implications for methylation-independent transcriptional repression. *Nucleic Acids Research*, *43*, 4296–4308. https://doi.org/10.1093/nar/gkv207

**Hou, Z., Zhang, Y., Propson, N. E., Howden, S. E., Chu, L.-F., Sontheimer, E. J., & Thomson, J. A.** (2013). Efficient genome engineering in human pluripotent stem cells using Cas9 from *Neisseria meningitidis*. *Proceedings of the National Academy of Sciences*, *110*, 15644–15649. https://doi.org/10.1073/pnas.1313587110

**Hu, X., Li, J., Fu, M., Zhao, X., & Wang, W.** (2021). The JAK/STAT signaling pathway: From bench to clinic. *Signal Transduction and Targeted Therapy*, *6*, 402. https://doi.org/10.1038/s41392-021-00791-1

**Hu, Z., Medeiros, L. J., Xu, M., Yuan, J., Peker, D., Shao, L., Tang, Z., Mai, B., Thakral, B., Rios, A., Hu, S., & Wang, W.** (2023). T-cell prolymphocytic leukemia with t(X;14)(q28;q11.2): A clinicopathologic study of 15 cases. *American Journal of Clinical Pathology*, *159*, 325–336. https://doi.org/10.1093/ajcp/aqac163

**Ikeda, A., Fujii, W., Sugiura, K., & Naito, K.** (2019). High-fidelity endonuclease variant HypaCas9 facilitates accurate allele-specific gene modification in mouse zygotes. *Communications Biology*, *2*, 371. https://doi.org/10.1038/s42003-019-0627-8

**Jankowski, J., Lee, H. K., Wilflingseder, J., & Hennighausen, L.** (2021). JAK inhibitors dampen activation of interferon-activated transcriptomes and the SARS-CoV-2 receptor ACE2 in human renal proximal tubules. *iScience*, *24*, 102928. https://doi.org/10.1016/j.isci.2021.102928

**Jensen, L. T., Attfield, K. E., Feldmann, M., & Fugger, L.** (2023). Allosteric TYK2 inhibition: Redefining autoimmune disease therapy beyond JAK1–3 inhibitors. *EBioMedicine*, *97*, 104840. https://doi.org/10.1016/j.ebiom.2023.104840

**Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A., Bridgland, A., Meyer, C., Kohl, S. A. A., Ballard, A. J., Cowie, A., Romera-Paredes, B., Nikolov, S., Jain, R., Adler, J., … Hassabis, D.** (2021). Highly accurate protein structure prediction with AlphaFold. *Nature*, *596*, 583–589. https://doi.org/10.1038/s41586-021-03819-2

**Kapralova, K., Horvathova, M., Pecquet, C., Fialova Kucerova, J., Pospisilova, D., Leroy, E., Kralova, B., Milosevic Feenstra, J. D., Schischlik, F., Kralovics, R., Constantinescu, S. N., & Divoky, V.** (2016). Cooperation of germ line JAK2 mutations E846D and R1063H in hereditary erythrocytosis with megakaryocytic atypia. *Blood*, *128*, 1418–1423. https://doi.org/10.1182/blood-2016-03-706093

**Kerner, G., Laval, G., Patin, E., Boisson-Dupuis, S., Abel, L., Casanova, J.-L., & Quintana-Murci, L.** (2021). Human ancient DNA analyses reveal the high burden of tuberculosis in Europeans over the last 2,000 years. *The American Journal of Human Genetics*, *108*, 517–524. https://doi.org/10.1016/j.ajhg.2021.02.009

**Kerner, G., Ramirez-Alejo, N., Seeleuthner, Y., Yang, R., Ogishi, M., Cobat, A., Patin, E., Quintana-Murci, L., Boisson-Dupuis, S., Casanova, J.-L., & Abel, L.** (2019). Homozygosity for TYK2 P1104A underlies tuberculosis in about 1% of patients in a cohort of European ancestry. *Proceedings of the National Academy of Sciences of the United States of America*, *116*, 10430–10434. https://doi.org/10.1073/pnas.1903561116

**Kim, D., Park, G., Huuhtanen, J., Ghimire, B., Rajala, H., Moriggl, R., Chan, W. C., Kankainen, M., Myllymäki, M., & Mustjoki, S.** (2021). STAT3 activation in large granular lymphocyte leukemia is associated with cytokine signaling and DNA hypermethylation. *Leukemia*, *35*, 3430–3443. https://doi.org/10.1038/s41375-021-01208-0

**Kim, H. K., Lee, S., Kim, Y., Park, J., Min, S., Choi, J. W., Huang, T. P., Yoon, S., Liu, D. R., & Kim, H. H.** (2020). High-throughput analysis of the activities of xCas9, SpCas9-NG and SpCas9 at matched and mismatched target sequences in human cells. *Nature Biomedical Engineering*, *4*, 111–124. https://doi.org/10.1038/s41551-019-0505-0

https://doi.org/10.1017/thg.2026.10054 Published online by Cambridge University Press

**Klein, K., Stoiber, D., Sexl, V., & Witalisz-Siepracka, A.** (2021). Untwining anti-tumor and immunosuppressive effects of JAK inhibitors—A strategy for hematological malignancies? *Cancers*, *13*, 2611. https://doi.org/10.3390/cancers13112611

**Koo, G. C., Tan, S. Y., Tang, T., Poon, S. L., Allen, G. E., Tan, L., Chong, S. C., Ong, W. S., Tay, K., Tao, M., Quek, R., Loong, S., Yeoh, K.-W., Yap, S. P., Lee, K. A., Lim, L. C., Tan, D., Goh, C., Cutcutache, I., … Lim, S. T.** (2012). Janus kinase 3–activating mutations identified in natural killer/T-cell lymphoma. *Cancer Discovery*, *2*, 591–597. https://doi.org/10.1158/2159-8290.CD-12-0028

**Krah, N. M., Miotke, L., Li, P., Patel, J. L., Bowen, A. R., Pomicter, A. D., & Patel, A. B.** (2023). JAK2 R683S mutation resulting in dual diagnoses of chronic eosinophilic leukemia and myelodysplastic/myeloproliferative overlap syndrome. *Journal of the National Comprehensive Cancer Network*, *21*, 1218–1223. https://doi.org/10.6004/jnccn.2023.7068

**Kristensen, T., Larsen, M., Rewes, A., Frederiksen, H., Thomassen, M., & Møller, M. B.** (2014). Clinical relevance of sensitive and quantitative STAT3 mutation analysis using next-generation sequencing in T-cell large granular lymphocytic leukemia. *The Journal of Molecular Diagnostics*, *16*, 382–392. https://doi.org/10.1016/j.jmoldx.2014.03.005

**Łączak, M., Kuczyńska, M., Grygier, J., Andrzejewska, D., Grochowska, W., Gulaczyk, H., & Lewandowski, K.** (2022). JAK and STAT gene mutations and JAK-STAT pathway activation in lympho- and myeloproliferative neoplasms. *Hematology in Clinical Practice*, *12*, 89–104.

**Landrum, M. J., Lee, J. M., Riley, G. R., Jang, W., Rubinstein, W. S., Church, D. M., & Maglott, D. R.** (2014). ClinVar: Public archive of relationships among sequence variation and human phenotype. *Nucleic Acids Research, 42*, D980–D985. https://doi.org/10.1093/nar/gkt1113

**Lee, H. K., Jung, O., & Hennighausen, L.** (2021). JAK inhibitors dampen activation of interferon-stimulated transcription of ACE2 isoforms in human airway epithelial cells. *Communications Biology*, *4*, 654. https://doi.org/10.1038/s42003-021-02222-w

**Lee, H. K., Willi, M., Shin, H. Y., Liu, C., & Hennighausen, L.** (2018). Progressing super-enhancer landscape during mammary differentiation controls tissue-specific gene regulation. *Nucleic Acids Research*, *46*, 10796–10809. https://doi.org/10.1093/nar/gky955

**Lex, A., Gehlenborg, N., Strobelt, H., Vuillemot, R., & Pfister, H.** (2014). UpSet: Visualization of intersecting sets. *IEEE Transactions on Visualization and Computer Graphics*, *20*, 1983–1992. https://doi.org/10.1109/TVCG.2014.2346248

**Li, Z., Gakovic, M., Ragimbeau, J., Eloranta, M.-L., Rönnblom, L., Michel, F., & Pellegrini, S.** (2013). Two rare disease-associated Tyk2 variants are catalytically impaired but signaling competent. *The Journal of Immunology*, *190*, 2335–2344. https://doi.org/10.4049/jimmunol.1202368

**Liang, F., Zhang, Y., Li, L., Yang, Y., Fei, J.-F., Liu, Y., & Qin, W.** (2022). SpG and SpRY variants expand the CRISPR toolbox for genome editing in zebrafish. *Nature Communications*, *13*, 1–10. https://doi.org/10.1038/s41467-022-29750-2

**Lin, M., Nebral, K., Gertzen, C. G. W., Ganmore, I., Haas, O. A., Bhatia, S., Fischer, U., Kuhlen, M., Gohlke, H., Izraeli, S., Trka, J., Hu, J., Borkhardt, A., Hauer, J., & Auer, F.** (2019). JAK2 p.G571S in B-cell precursor acute lymphoblastic leukemia: a synergizing germline susceptibility. *Leukemia*, *33*, 2331–2335. https://doi.org/10.1038/s41375-019-0423-1

**Lio, C. T., Düz, T., Hoffmann, M., Willruth, L.-L., Baumbach, J., List, M., & Tsoy, O.** (2025). Comprehensive benchmark of differential transcript usage analysis for bulk and single-cell RNA sequencing. *NAR Genomics and Bioinformatics*, *7*, lqaf117. https://doi.org/10.1093/nargab/lqaf117

**López-Isac, E., Campillo-Davo, D., Bossini-Castillo, L., Guerra, S. G., Assassi, S., Simeón, C. P., Carreira, P., Ortego-Centeno, N., García de la Peña, P., the Spanish Scleroderma Group, Beretta, L., Santaniello, A., Bellocchi, C., Lunardi, C., Moroncini, G., Gabrielli, A., Riemekasten, G., Witte, T., Hunzelmann, N., … Martín, J.** (2016). Influence of TYK2 in systemic sclerosis susceptibility: a new locus in the IL-12 pathway. *Annals of the Rheumatic Diseases*, *75*, 1521–1526. https://doi.org/10.1136/annrheumdis-2015-207570

**Maaziz, N., Garrec, C., Airaud, F., Bobée, V., Contentin, N., Cayssials, E., Rimbert, A., Aral, B., Bézieau, S., Gardie, B., & Girodon, F.** (2023). Germline JAK2 E846D substitution as the cause of erythrocytosis? *Genes*, *14*, 1066. https://doi.org/10.3390/genes14051066

**Mambet, C., Babosova, O., Defour, J.-P., Leroy, E., Necula, L., Stanca, O., Tatic, A., Berbec, N., Coriu, D., Belickova, M., Kralova, B., Lanikova, L., Vesela, J., Pecquet, C., Saussoy, P., Havelange, V., Diaconu, C. C., Divoky, V., & Constantinescu, S. N.** (2018). Cooccurring JAK2 V617F and R1063H mutations increase JAK2 signaling and neutrophilia in myeloproliferative neoplasms. *Blood*, *132*, 2695–2699. https://doi.org/10.1182/blood-2018-05-848002

**Mardenborough, Y. S. N., Nitsenko, K., Laffeber, C., Duboc, C., Sahin, E., Quessada-Vial, A., Winterwerp, H. H. K., Sixma, T. K., Kanaar, R., Friedhoff, P., Strick, T. R., & Lebbink, J. H. G.** (2019). The unstructured linker arms of MutL enable GATC site incision beyond roadblocks during initiation of DNA mismatch repair. *Nucleic Acids Research*, *47*, 11667–11680. https://doi.org/10.1093/nar/gkz1020

**Meng, E. C., Goddard, T. D., Pettersen, E. F., Couch, G. S., Pearson, Z. J., Morris, J. H., & Ferrin, T. E.** (2023). UCSF ChimeraX: Tools for structure building and analysis. *Protein Science*, *32*, e4792. https://doi.org/10.1002/pro.4792

**Molitor, T. P., Hayashi, G., Lin, M.-Y., Dunn, C. J., Peterson, N. G., Poston, R. G., Kurnellas, M. P., Traver, D. A., Patel, S., Akgungor, Z., Leonardi, V., Lewis, C., Segales, J. S., Bennett, D. S., Truong, A. P., Dani, M., Naphade, S., Wong, J. K., McDermott, A. E., … Rassoulpour, A.** (2025). Central TYK2 inhibition identifies TYK2 as a key neuroimmune modulator. *Proceedings of the National Academy of Sciences of the United States of America*, *122*, e2422172122. https://doi.org/10.1073/pnas.2422172122

**Motegi, T., Kochi, Y., Matsuda, K., Kubo, M., Yamamoto, K., & Momozawa, Y.** (2019). Identification of rare coding variants in TYK2 protective for rheumatoid arthritis in the Japanese population and their effects on cytokine signalling. *Annals of the Rheumatic Diseases*, *78*, 1062–1069. https://doi.org/10.1136/annrheumdis-2018-214639

**Müller, M., Lee, C. M., Gasiunas, G., Davis, T. H., Cradick, T. J., Siksnys, V., Bao, G., Cathomen, T., & Mussolino, C.** (2016). Streptococcus thermophilus CRISPR-Cas9 systems enable specific editing of the human genome. *Molecular Therapy*, *24*, 636–644. https://doi.org/10.1038/mt.2015.218

**Ogishi, M., Arias, A. A., Yang, R., Han, J. E., Zhang, P., Rinchai, D., Halpern, J., Mulwa, J., Keating, N., Chrabieh, M., Lainé, C., Seeleuthner, Y., Ramírez-Alejo, N., Nekooie-Marnany, N., Guennoun, A., Muller-Fleckenstein, I., Fleckenstein, B., Kilic, S. S., Minegishi, Y., … Boisson-Dupuis, S.** (2022). Impaired IL-23–dependent induction of IFN-γ underlies mycobacterial disease in patients with inherited TYK2 deficiency. *The Journal of Experimental Medicine*, *219*. https://doi.org/10.1084/jem.20220094

**Olson, K. C., Moosic, K. B., Jones, M. K., Larkin, P. M. K., Olson, T. L., Toro, M. F., Fox, T. E., Feith, D. J., & Loughran, T. P., Jr.** (2020). Large granular lymphocyte leukemia serum and corresponding hematological parameters reveal unique cytokine and sphingolipid biomarkers and associations with STAT3 mutations. *Cancer Medicine*, *9*, 6533–6549. https://doi.org/10.1002/cam4.3294

**O'Shea, J. J., Schwartz, D. M., Villarino, A. V., Gadina, M., McInnes, I. B., & Laurence, A.** (2015). The JAK-STAT pathway: Impact on human disease and therapeutic intervention. *Annual Review of Medicine*, *66*, 311–328. https://doi.org/10.1146/annurev-med-051113-024537

**Pace, M., Guadagno, E., Russo, D., Gencarelli, A., Carlea, A., Di Spiezio, A., Bertuzzi, C., Mascolo, M., Grimaldi, F., & Insabato, L.** (2023). Myeloid sarcoma of the breast as blast phase of JAK2-mutated (Val617Phe Exon 14p) essential thrombocythemia: A case report and a systematic literature review. *Pathobiology*, *90*, 123–130. https://doi.org/10.1159/000528607

**Pal, L. R., & Moult, J.** (2015). Genetic basis of common human disease: Insight into the role of missense SNPs from genome-wide association studies. *Journal of Molecular Biology*, *427*, 2271–2289. https://doi.org/10.1016/j.jmb.2015.04.014

**Panovska-Stavridis, I., Eftimov, A., Ivanovski, M., Pivkova-Veljanovska, A., Cevreska, L., Hermouet, S., & Dimovski, A. J.** (2016). Essential thrombocythemia associated with germline JAK2 G571S variant and somatic CALR type 1 mutation. *Clinical Lymphoma, Myeloma & Leukemia*, *16*, e55–e57. https://doi.org/10.1016/j.clml.2016.02.025

**Patchell, D., Keohane, C., O'Shea, S., & Langabeer, S. E.** (2024). Incidence and impact of non-canonical JAK2 p.(Val617Phe) mutations in

myeloproliferative neoplasm molecular diagnostics. *Journal of Clinical Pathology*. https://doi.org/10.1136/jcp-2023-209276

**Peña-Martínez, E. G., & Rodríguez-Martínez, J. A.** (2024). Decoding non-coding variants: Recent approaches to studying their role in gene regulation and human diseases. *Frontiers in Bioscience*, *16*, 4. https://doi.org/10.31083/j.fbs1601004

**Perner, F., Perner, C., Ernst, T., & Heidel, F. H.** (2019). Roles of JAK2 in aging, inflammation, hematopoiesis and malignant transformation. *Cells*, *8*, 854. https://doi.org/10.3390/cells8080854

**Pruitt, K. D., Harrow, J., Harte, R. A., Wallin, C., Diekhans, M., Maglott, D. R., Searle, S., Farrell, C. M., Loveland, J. E., Ruef, B. J., Hart, E., Suner, M.-M., Landrum, M. J., Aken, B., Ayling, S., Baertsch, R., Fernandez-Banet, J., Cherry, J. L., Curwen, V., … Lipman, D.** (2009). The consensus coding sequence (CCDS) project: Identifying a common protein-coding gene set for the human and mouse genomes. *Genome Research*, *19*, 1316–1323. https://doi.org/10.1101/gr.080531.108

**Pujar, S., O'Leary, N. A., Farrell, C. M., Loveland, J. E., Mudge, J. M., Wallin, C., Girón, C. G., Diekhans, M., Barnes, I., Bennett, R., Berry, A. E., Cox, E., Davidson, C., Goldfarb, T., Gonzalez, J. M., Hunt, T., Jackson, J., Joardar, V., Kay, M. P., … Pruitt, K. D.** (2018). Consensus coding sequence (CCDS) database: A standardized set of human and mouse protein-coding regions supported by expert curation. *Nucleic Acids Research*, *46*(D1), D221–D228. https://doi.org/10.1093/nar/gkx1031

**Puli'uvea, C., Immanuel, T., Green, T. N., Tsai, P., Shepherd, P. R., & Kalev-Zylinska, M. L.** (2024). Insights into the role of JAK2-I724T variant in myeloproliferative neoplasms from a unique cohort of New Zealand patients. *Hematology, 29*. https://doi.org/10.1080/16078454.2023.2297597

**Rampal, R., Al-Shahrour, F., Abdel-Wahab, O., Patel, J. P., Brunel, J.-P., Mermel, C. H., Bass, A. J., Pretz, J., Ahn, J., Hricik, T., Kilpivaara, O., Wadleigh, M., Busque, L., Gilliland, D. G., Golub, T. R., Ebert, B. L., & Levine, R. L.** (2014). Integrated genomic analysis illustrates the central role of JAK-STAT pathway activation in myeloproliferative neoplasm pathogenesis. *Blood*, *123*, e123–e133. https://doi.org/10.1182/blood-2013-11-536003

**Ramsey, M. C., Sabatini, P. J. B., Watson, G., Chawla, T., Ko, M., & Sakhdari, A.** (2023). Case report: Identification of a novel STAT3 mutation in EBV-positive inflammatory follicular dendritic cell sarcoma. *Frontiers in Oncology*, *13*. https://doi.org/10.3389/fonc.2023.1266897

**Reich, C., Ostropolets, A., Ryan, P., Rijnbeek, P., Schuemie, M., Davydov, A., Dymshyts, D., & Hripcsak, G.** (2024). OHDSI standardized vocabularies—a large-scale centralized reference ontology for international data harmonization. *Journal of the American Medical Informatics Association*, *31*, 583–590. https://doi.org/10.1093/jamia/ocad247

**Rivera-Munoz, P., Laurent, A. P., Siret, A., Lopez, C. K., Ignacimouttou, C., Cornejo, M. G., Bawa, O., Rameau, P., Bernard, O. A., Dessen, P., Gilliland, D. G., Mercher, T., & Malinge, S.** (2018). Partial trisomy 21 contributes to T-cell malignancies induced by JAK3-activating mutations in murine models. *Blood Advances*, *2*, 1616–1627. https://doi.org/10.1182/bloodadvances.2018018161

**Rivero, A., Mozas, P., Jiménez, L., López-Guerra, M., Colomer, D., Bataller, A., Correa, J., Rivas-Delgado, A., Bastidas, G., Baumann, T., Martínez-Trillos, A., Delgado, J., Giné, E., Campo, E., López-Guillermo, A., Villamor, N., Magnano, L., & Matutes, E.** (2021). Clinicobiological characteristics and outcomes of patients with T-cell large granular lymphocytic leukemia and chronic lymphoproliferative disorder of natural killer cells from a single institution. *Cancers*, *13*, 3900. https://doi.org/10.3390/cancers13153900

**Roncero, A. M., López-Nieva, P., Cobos-Fernández, M. A., Villa-Morales, M., González-Sánchez, L., López-Lorenzo, J. L., Llamas, P., Ayuso, C., Rodríguez-Pinilla, S. M., Arriba, M. C., Piris, M. A., Fernández-Navarro, P., Fernández, A. F., Fraga, M. F., Santos, J., & Fernández-Piqueras, J.** (2016). Contribution of JAK2 mutations to T-cell lymphoblastic lymphoma development. *Leukemia*, *30*, 94–103. https://doi.org/10.1038/leu.2015.224

**Saevarsdottir, S., Stefansdottir, L., Sulem, P., Thorleifsson, G., Ferkingstad, E., Rutsdottir, G., Glintborg, B., Westerlind, H., Grondal, G., Loft, I. C., Sorensen, S. B., Lie, B. A., Brink, M., Ärlestig, L., Arnthorsson, A. O., Baecklund, E., Banasik, K., Bank, S., Bjorkman, L. I.**, … **The Danish RA Genetics Working Group**. (2022). Multiomics analysis of rheumatoid arthritis yields sequence variants that have large effects on risk of the seropositive subset. *Annals of the Rheumatic Diseases*, *81*, 1085–1095. https://doi.org/10.1136/annrheumdis-2021-221527

**Sato, T., Toki, T., Kanezaki, R., Xu, G., Terui, K., Kanegane, H., Miura, M., Adachi, S., Migita, M., Morinaga, S., Nakano, T., Endo, M., Kojima, S., Kiyoi, H., Mano, H., & Ito, E.** (2008). Functional analysis of JAK3 mutations in transient myeloproliferative disorder and acute megakaryoblastic leukaemia accompanying Down syndrome. *British Journal of Haematology*, *141*, 681–688. https://doi.org/10.1111/j.1365-2141.2008.07079.x

**Schmidheini, L., Mathis, N., Marquart, K. F., Rothgangl, T., Kissling, L., Böck, D., Chanez, C., Wang, J. P., Jinek, M., & Schwank, G.** (2024). Continuous directed evolution of a compact CjCas9 variant with broad PAM compatibility. *Nature Chemical Biology*, *20*, 333–343. https://doi.org/10.1038/s41589-023-01448-7

**Schulze, S., Stengel, R., Jaekel, N., Wang, S.-Y., Franke, G.-N., Roskos, M., Schneider, M., Niederwieser, D., & Al-Ali, H. K.** (2019). Concomitant and noncanonical JAK2 and MPL mutations in JAK2V617F- and MPLW515L-positive myelofibrosis. *Genes, Chromosomes & Cancer*, *58*, 747–755. https://doi.org/10.1002/gcc.22763

**Shen, L. X., Basilion, J. P., & Stanton, V. P., Jr.** (1999). Single-nucleotide polymorphisms can cause different structural folds of mRNA. *Proceedings of the National Academy of Sciences of the United States of America*, *96*, 7871–7876. https://doi.org/10.1073/pnas.96.14.7871

**Shen, M.** (2023). A case report of T-LGL leukemia-associated pure red cell aplasia harboring STAT3, TNFAIP3, and KMT2D mutation. *Translational Cancer Research*, *12*, 1054–1059. https://doi.org/10.21037/tcr-22-2590

**Shillingford, J. M.** (2002). Jak2 is an essential tyrosine kinase involved in pregnancy-mediated development of mammary secretory epithelium. *Molecular Endocrinology*, *16*, 563–570. https://doi.org/10.1210/mend.16.3.0785

**Sim, S. H., Kim, S., Kim, T. M., Jeon, Y. K., Nam, S. J., Ahn, Y.-O., Keam, B., Park, H. H., Kim, D.-W., Kim, C. W., & Heo, D. S.** (2017). Novel JAK3-activating mutations in extranodal NK/T-cell lymphoma, nasal type. *The American Journal of Pathology*, *187*, 980–986. https://doi.org/10.1016/j.ajpath.2016.12.008

**Skoczen, S., Stepien, K., Mlynarski, W., Centkowski, P., Kwiecinska, K., Korostynski, M., Piechota, M., Wyrobek, E., Moryl-Bujakowska, A., Strojny, W., Rej, M., Kowalczyk, J., & Balwierz, W.** (2020). Genetic signature of acute lymphoblastic leukemia and netherton syndrome co-incidence—first report in the literature. *Frontiers in Oncology*, *9*, 1477. https://doi.org/10.3389/fonc.2019.01477

**Slaymaker, I. M., Gao, L., Zetsche, B., Scott, D. A., Yan, W. X., & Zhang, F.** (2016). Rationally engineered Cas9 nucleases with improved specificity. *Science*, *351*, 84–88. https://doi.org/10.1126/science.aad5227

**Sondka, Z., Dhir, N. B., Carvalho-Silva, D., Jupe, S., Madhumita, McLaren, K., Starkey, M., Ward, S., Wilding, J., Ahmed, M., Argasinska, J., Beare, D., Chawla, M. S., Duke, S., Fasanella, I., Neogi, A. G., Haller, S., Hetenyi, B., Hodges, L., … Teague, J.** (2024). COSMIC: a curated database of somatic variants and clinical data for cancer. *Nucleic Acids Research*, *52*, D1210–D1217. https://doi.org/10.1093/nar/gkad984

**Steri, M., Orrù, V., Sidore, C., Mulas, A., Pitzalis, M., Busonero, F., Maschio, A., Serra, V., Dei, M., Lai, S., Virdis, F., Lobina, M., Loizedda, A., Marongiu, M., Masala, M., Floris, M., Curreli, N., Balaci, L., Loi, F., … Zoledziewska, M.** (2025). TYK2 :P.Pro1104Ala variant protects against autoimmunity by modulating immune cell levels. *Immunology*, *174*, 462–469. https://doi.org/10.1111/imm.13902

**Syed, F., Ballew, O., Lee, C.-C., Rana, J., Krishnan, P., Castela, A., Weaver, S. A., Chalasani, N. S., Thomaidou, S. F., Demine, S., Chang, G., Coomans de Brachène, A., Alvelos, M. I., Vazquez, E. M., Marselli, L., Orr, K., Felton, J. L., Liu, J., Kaddis, J. S., … Evans-Molina, C.** (2025). Pharmacological inhibition of tyrosine protein-kinase 2 reduces islet inflammation and delays type 1 diabetes onset in mice. *EBioMedicine*, *117*, 105734. https://doi.org/10.1016/j.ebiom.2025.105734

**Teng, S., Srivastava, A. K., Schwartz, C. E., Alexov, E., & Wang, L.** (2010). Structural assessment of the effects of amino acid substitutions on protein stability and protein-protein interaction. *International Journal of Computational Biology and Drug Design*, *3*, 334–349. https://doi.org/10.1504/IJCBDD.2010.038396

**Tsoy, O., Ameling, S., Franzenburg, S., Hoffmann, M. D., Liv-Willuth, L., Lee, H. K., Knabl, L., Furth, P. A., Voelker, U., Hennighausen, L.,**

**Baumbach, J., Kacprowski, T., & List, M.** (2024). RNA sequencing depth guidelines for the study of alternative splicing. *bioRxiv*. https://doi.org/10.1101/2024.10.09.617406

**Tun, P. W. W., Buka, R. J., Graham, J., & Dyer, P.** (2022). Heterozygous, germline JAK2 E846D substitution as the cause of familial erythrocytosis. *British Journal of Haematology*, *198*, 923–926. https://doi.org/10.1111/bjh.18304

**Pettersen, E. F., Goddard, T. D., Huang, C. C., Meng, E. C., Couch, G. S., Croll, T. I., Morris, J. H., & Ferrin, T. E.** (2021). UCSF ChimeraX: Structure visualization for researchers, educators, and developers. *Protein Science*, *30*, 70–82. https://doi.org/10.1002/pro.3943

**Uzel, G., Sampaio, E. P., Lawrence, M. G., Hsu, A. P., Hackett, M., Dorsey, M. J., Noel, R. J., Verbsky, J. W., Freeman, A. F., Janssen, E., Bonilla, F. A., Pechacek, J., Chandrasekaran, P., Browne, S. K., Agharahimi, A., Gharib, A. M., Mannurita, S. C., Yim, J. J., Gambineri, E., … Holland, S. M.** (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*, 1611–1623. https://doi.org/10.1016/j.jaci.2012.11.015

**Vakulskas, C. A., Dever, D. P., Rettig, G. R., Turk, R., Jacobi, A. M., Collingwood, M. A., Bode, N. M., McNeill, M. S., Yan, S., Camarena, J., Lee, C. M., Park, S. H., Wiebking, V., Bak, R. O., Gomez-Ospina, N., Pavel-Dinu, M., Sun, W., Bao, G., Porteus, M. H., & Behlke, M. A.** (2018). A high-fidelity Cas9 mutant delivered as a ribonucleoprotein complex enables efficient gene editing in human hematopoietic stem and progenitor cells. *Nature Medicine*, *24*, 1216–1224. https://doi.org/10.1038/s41591-018-0137-0

**Veitia, R. A., & Innan, H.** (2022). Pathogenic 'germline' variants associated with myeloproliferative disorders in apparently normal individuals: Inherited or acquired genetic alterations? *Clinical Genetics*, *101*, 371–374. https://doi.org/10.1111/cge.14121

**Waldmann, T. A.** (2017). JAK/STAT pathway directed therapy of T-cell leukemia/lymphoma: Inspired by functional and structural genomics. *Molecular and Cellular Endocrinology*, *451*, 66–70. https://doi.org/10.1016/j.mce.2017.02.040

**Wang, Y.** (2018). Identification of SNPs and their association with disease traits in a large population. *Genetics Research*, *100*. https://doi.org/10.1017/S0016672318000012

**Wu, J., Tang, B., & Tang, Y.** (2020). Allele-specific genome targeting in the development of precision medicine. *Theranostics*, *10*, 3118–3137. https://doi.org/10.7150/thno.42234

**Xu, L., Wilson, R. A., Laetsch, T. W., Oliver, D., Spunt, S. L., Hawkins, D. S., & Skapek, S. X.** (2016). Potential pitfalls of mass spectrometry to uncover mutations in childhood soft tissue sarcoma: A report from the Children's Oncology Group. *Scientific Reports*, *6*, 33429. https://doi.org/10.1038/srep33429

**Xu, R. Z., Karsan, A., Xu, Z., & Berry, B. R.** (2022). A rare de novo pure erythroid leukemia with JAK2 R683S mutation. *Annals of Hematology*, *101*, 921–922. https://doi.org/10.1007/s00277-021-04727-5

**Xue, C., Yao, Q., Gu, X., Shi, Q., Yuan, X., Chu, Q., Bao, Z., Lu, J., & Li, L.** (2023). Evolving cognition of the JAK-STAT signaling pathway: autoimmune disorders and cancer. *Signal Transduction and Targeted Therapy*, *8*, 1–24. https://doi.org/10.1038/s41392-023-01334-6

**Yan, Y., Olson, T. L., Nyland, S. B., Feith, D. J., & Loughran, T. P., Jr.** (2015). Emergence of a STAT3 mutated NK clone in LGL leukemia. *Leukemia Research Reports*, *4*, 4–7. https://doi.org/10.1016/j.lrr.2014.12.002

**Yin, C. C., Tam, W., Walker, S. M., Kaur, A., Ouseph, M. M., Xie, W., Weinberg, O. K., Li, P., Zuo, Z., Routbort, M. J., Chen, S., Medeiros, L. J., George, T. I., Orazi, A., Arber, D. A., Bagg, A., Hasserjian, R. P., & Wang, S. A.** (2023). STAT5B mutations in myeloid neoplasms differ by disease subtypes but characterize a subset of chronic myeloid neoplasms with eosinophilia and/or basophilia. *Haematologica, 109*. https://doi.org/10.3324/haematol.2023.284311

**Zhang, Y., Zhang, H., Xu, X., Wang, Y., Chen, W., Wang, Y., Wu, Z., Tang, N., Wang, Y., Zhao, S., Gan, J., & Ji, Q.** (2020). Catalytic-state structure and engineering of *Streptococcus thermophilus* Cas9. *Nature Catalysis*, *3*, 813–823. https://doi.org/10.1038/s41929-020-00500-9

**Zhang, Y., Zhao, Y., Liu, Y., Zhang, M., & Zhang, J.** (2024). New advances in the role of JAK2 V617F mutation in myeloproliferative neoplasms. *Cancer*. https://doi.org/10.1002/cncr.35559

**Zhong, L., Wang, W., Ma, M., Gou, L., Tang, X., & Song, H.** (2017). Chronic active Epstein–Barr virus infection as the initial symptom in a Janus kinase 3 deficiency child. *Medicine*, *96*, e7989. https://doi.org/10.1097/MD.0000000000007989

https://doi.org/10.1017/thg.2026.10054 Published online by Cambridge University Press