



UNIVERSITÀ DEL PIEMONTE ORIENTALE

School of Medicine

Department of Health Sciences

Master's Degree in Medical Biotechnologies

A comparative analysis of SAP and WAS regulated gene expression

Supervisor:

Prof. Gianluca Baldanzi



Digitally signed by Gianluca Baldanzi
Data: 28.09.2026 21:20:43 CEST
Organizzazione: UNIVERSITA' DEGLI
STUDI DEL PIEMONTE
ORIENTALE/01943490027

Candidate: Dilshan Tharanga Bopitiya
Matricola No: 20063517

Academic Year 2025/2026

INDEX

ABSTRACT	1
1. INTRODUCTION.....	3
1.1. T-cell activation and signalling.	3
1.2. X-linked lymphoproliferative disease and SAP deficiency.	4
1.3. Wiskott-Aldrich syndrome and WASp deficiency.....	6
1.4. Biological relationship between SAP and WASp.....	8
1.5. The clinical and immunological connection between XLP and WAS.....	9
1.6. Transcriptomic and gene expression analysis.	11
2. RATIONALE, AIMS AND OBJECTIVES.....	14
3. MATERIALS AND METHODS	15
3.1. Data sources.....	15
3.2. SAP dataset.....	15
3.3. WAS dataset.	15
3.4. In-house-generated SAP and WAS data set.	16
3.5. Gene Identifier Standardization.....	17
3.6. Comparative (VENN) Analysis.....	18
4. RESULTS	19
4.1. Comparison of SAP- and WAS-regulated gene sets.	19
4.2. Comparison of In-house SAP- and WAS-regulated gene set.....	22
4.3. Comparison of the laboratory datasets with the literature datasets.....	23
5. DISCUSSION	28
6. CONCLUSION	35
7. BIBLIOGRAPHY	36

ABSTRACT

Signalling lymphocytic activation molecule (SLAM)-associated protein (SAP) and Wiskott–Aldrich syndrome protein (WASp) are distinct signalling proteins that contribute to immune-cell function and have been implicated in the pathogenesis of X-linked lymphoproliferative disease (XLP) and Wiskott–Aldrich syndrome (WAS), respectively. Although both proteins are involved in T-cell signalling, it remains unclear whether they regulate a common transcriptional programme. This study aimed to determine SAP- and WAS-regulated gene expression and to assess how experimental conditions influence the resulting gene sets.

Published SAP- and WAS-regulated gene datasets were compared after standardizing the gene identifiers to human Ensembl Gene IDs and mapping mouse WAS genes to their human orthologues. Upregulated and downregulated genes were analysed separately using Venn analysis. The comparison was also extended to an in-house RNA-sequencing dataset from human CD8⁺ lymphocytes following SAP or WAS silencing and CD3 and CD28 restimulation.

Following gene identifier standardization, the literature-derived datasets contained 315 SAP-upregulated and 74 SAP-downregulated genes, compared with 718 WAS-upregulated and 433 WAS-downregulated genes. Only 12 genes were commonly upregulated, and one gene was commonly downregulated across the SAP and WAS literature datasets. In the in-house datasets, SAP silencing resulted in 422 upregulated and 776 downregulated genes, whereas WAS silencing resulted in 184 upregulated and 1,085 downregulated genes. These datasets shared 69 commonly upregulated and 71 commonly downregulated genes. However, overlap between the in-house and literature-derived datasets was limited, with zero common SAP-regulated genes and only four upregulated and 13 downregulated genes shared between the WAS datasets.

Overall, the findings provide no evidence of a single shared transcriptional programme regulated by SAP and WAS. The differences between datasets highlight the influence of experimental conditions and the importance of gene identifier standardization, orthology mapping, and experimental matching when comparing transcriptomic data.

1. INTRODUCTION

1.1. T-cell activation and signalling.

T-cell activation is a coordinated process in which T lymphocytes recognize antigen and develop appropriate immune responses. The process begins when the T-cell receptor (TCR) recognizes antigenic peptide presented by major histocompatibility complex (MHC) molecules on antigen-presenting cells (Smith-garvin, Koretzky and Jordan, 2010). When the TCR interacts with the antigen, it initiates a cascade of intracellular signalling events involving processes such as protein phosphorylation, adaptor proteins, calcium mobilization, cytoskeletal rearrangement, and transcriptional regulation (Smith-garvin, Koretzky and Jordan, 2010). The CD3 complex associated with the TCR contains signalling motifs that become phosphorylated following receptor engagement. During early signalling, kinases such as lymphocyte-specific protein tyrosine kinase (Lck) and zeta-chain-associated protein kinase 70 (ZAP70) play important roles. These proteins contribute to the formation of a larger signalling complex (Rattanasri *et al.*, 2026). These proximal signalling events lead to activation of downstream pathways that regulate cellular activation, proliferation, cytokine production, differentiation, and survival (Smith-garvin, Koretzky and Jordan, 2010). Recent experimental studies show that Wiskott–Aldrich syndrome protein (WASp) contributes not only to actin organisation in T cells but also to proximal stages of TCR signalling. Loss of WASp in T cells reduces phosphorylation of the signalling proteins CD3, Lck, and ZAP70 following TCR stimulation and also associated with reduced expression of T cell activation markers such as CD69 and CD25 (Rattanasri *et al.*, 2026).

T-cell activation also requires changes in the actin cytoskeleton. These changes allow the formation and stabilisation of the immunological synapse between the T cell and antigen-presenting cell (APC). WASp is an important regulator of actin polymerisation and links TCR signalling to cytoskeletal rearrangement (Kumari *et al.*, 2015). WASp-dependent actin structures can also support downstream signalling events, including phospholipase C gamma1 (PLC γ 1) activation and calcium mobilisation. These processes are important for effective T-cell responses (Kumari *et al.*, 2015). The TCR signalling network ultimately activates transcription factors and gene-expression programmes that determine the functional state of the T cell.

Therefore, changes in proteins involved in early signalling can produce broader effects on gene expression. T-cell activation is consequently not restricted to changes at the cell membrane or cytoskeleton but involves coordinated changes extending to the transcriptional level (Kumari *et al.*, 2015; Rattanasri *et al.*, 2026).

SAP is another important signalling protein involved in lymphocyte activation. SAP mainly functions as an intracellular adaptor protein associated with SLAM-family receptors and influences signalling pathways that regulate T-cell and natural killer (NK) cell functions (Li *et al.*, 2003). In addition, recent studies have demonstrated a signalling connection between SAP, WASp, and diacylglycerol kinase alpha (DGK α) during strong TCR activation. WASp can interact with and inhibit DGK α , while NCK1 and CDC42 connect this pathway with the TCR signalling machinery and SAP. This pathway contributes to an appropriate cytokine response, including IL-2 production, and demonstrates that SAP- and WASp-dependent signalling can intersect within activated T cells (Velinati *et al.*, 2023). These findings provide an important biological basis for investigating whether loss of SAP and WASp can produce overlapping downstream transcriptional responses in T cells. However, the two proteins have different molecular functions, and their loss may also produce distinct effects on T-cell gene expression. Therefore, analysing their transcriptional consequences provides an opportunity to determine whether the signalling functions of these two proteins ultimately converge at the level of gene regulation.

1.2. X-linked lymphoproliferative disease and SAP deficiency.

X-linked lymphoproliferative disease (XLP) is a rare primary congenital immunodeficiency characterized by severe immune dysregulation, often following Epstein-Barr virus (EBV) infection (Booth *et al.*, 2011). The disease most commonly results from loss-of-function mutations in the *SH2D1A* gene on chromosome Xq24-25 (11–13), leading to deficiency of the signalling lymphocytic activation molecule (SLAM)- associated protein (SAP) (Snow *et al.*, 2009; Booth *et al.*, 2011).

At the molecular level, SAP is an Src-homology 2 (SH2) domain adaptor protein that is involved in the intracellular signalling of the signalling lymphocyte activation molecule (SLAM) family of immunomodulatory receptors, including SLAM itself and 2B4 (CD244) on NK, T, and B cell antigen (NTB-A) (Parolini *et al.*, 2000;

Tangye *et al.*, 2000). Loss of SAP causes intrinsic defects in lymphocyte function, including cytokine production, cytotoxicity, T-cell-dependent humoral immune responses, and NKT cell development (Snow *et al.*, 2009). NK cells were isolated from an XLP patient, and it was revealed that SAP is specifically required for 2B4-mediated (surface receptor) NK cell activation, since SAP-deficient NK cells failed to kill CD48-expressing target cells through the 2B4 receptor (Tangye *et al.*, 2000). Usually, SAP competes with the inhibitory phosphatase SHP-2 for binding to the cytoplasmic tails of SLAM family receptors, and in the absence of SAP, unopposed SHP-2 recruitment converts these receptors from activating to inhibitory signalling elements (Tangye *et al.*, 2000; Panchal, Houghton and Vassalou, 2022). SHP-2 is hyperphosphorylated in SAP-deficient T cells from XLP patients (Panchal, Houghton and Vassalou, 2022).

SAP also regulates T cell receptor (TCR) signal strength through control of diacylglycerol (DAG) metabolism (Baldanzi *et al.*, 2011). TCR co-stimulation with CD28 or SLAM normally stimulates SAP-dependent inhibition of diacylglycerol kinase alpha (DGK α), and DGK converts DAG into phosphatidic acid, thereby reducing downstream Ras/ERK and PKC θ signalling (Baldanzi *et al.*, 2011). This inhibitory regulation of DGK α failed to occur in SAP-deficient Jurkat T cells, resulting in defective Ras-GTP (the active form of the Ras protein) loading, reduced extracellular signal-regulated kinase (ERK) and protein kinase C theta (PKC θ) activation, and impaired nuclear factor of activated T cells (NFAT) activity and interleukin 2 (IL-2) production (Baldanzi *et al.*, 2011). A study showed that activated T cells from XLP patients are selectively resistant to restimulation-induced cell death (RICD) (Snow *et al.*, 2009). Moreover, RICD, an apoptotic program, limits T cell expansion in the presence of antigen and IL-2, while the same XLP T cells remain fully sensitive to other apoptotic pathways, including Fas-mediated death and Intrinsic Mitochondrial apoptotic pathways (Snow *et al.*, 2009). To identify the genes responsible for this defect, the same study also generated a microarray gene expression dataset comparing SAP-silenced human T cells with T cells transfected with non-specific control siRNA, following TCR restimulation. It also provided the full list of SAP-modulated genes as supplementary data. This dataset forms the SAP-regulated gene list used in the comparative analysis of the present study.

1.3. Wiskott-Aldrich syndrome and WASp deficiency.

Wiskott-Aldrich syndrome (WAS) refers to a rare X-linked primary immunodeficiency caused by mutations in the *WAS* gene, which is expressed exclusively in hematopoietic cells. This gene encodes the proline-rich, 502-amino-acid WAS protein (WASp) (Liang, Yang, *et al.*, 2023). Loss-of-function mutations in the *WAS* gene produce the classical WAS phenotype of microthrombocytopenia, eczema, recurrent infections, and a markedly increased incidence of autoimmune disease and malignancy (Liang, Yang, *et al.*, 2023; Velnati *et al.*, 2023). This reflects broad cytoskeleton-dependent defects related to immune cell activation, immune synapse formation, and lymphocyte homeostasis (Liang, Yang, *et al.*, 2023). WASp is an actin nucleation-promoting factor family and contains an N-terminal WH1 domain, a GTPase-binding domain (GBD), a proline-rich region, and a C-terminal verprolin-cofilin-acidic (VCA) domain (Sun *et al.*, 2019). At rest, WASp exists in an auto-inhibited conformation in which the GBD folds back onto the VCA domain, thereby preventing its interaction with the actin-related protein 2/3 (Arp2/3) complex (Sun *et al.*, 2019). Upon TCR engagement, WASp is recruited to the immunological synapse through a ZAP-70–CrkL–WIP–WASp signalling complex. With GTP-bound Cdc42 binding to the GBD and Tyr291 phosphorylation, the VCA domain is released from auto-inhibition, allowing it to bind Arp2/3 and drive branched actin polymerisation at the site of receptor engagement (Sun *et al.*, 2019). This actin-remodelling function underlies WASp's role in T-cell spreading, immune synapse formation and maturation, TCR-induced endocytosis, and T- and B-cell migration and homing (Sun *et al.*, 2019; Velnati *et al.*, 2023). A portion of WASp also translocates to the nucleus, where it associates with histone-modifying activity and regulates transcriptional programmes governing T helper cell differentiation. This includes a contribution to IL-2 gene expression that is independent of its actin-polymerising (VCA/Arp2/3-binding) function (Sun *et al.*, 2019).

WASp binds to diacylglycerol kinase alpha (DGK α) through its WH1 domain, connecting to the recoverin-homology domain of DGK α . Through this binding, during T-cell receptor (TCR) activation, DGK α enzymatic activity is inhibited independently of Arp2/3 activity, while it selectively promotes interleukin-2 (IL-2) production (Velnati *et al.*, 2023). This reveals an actin-independent branch of WASp function in T-cell signal transduction (Velnati *et al.*, 2023). A recent study performed

RNA sequencing on spleen tissue from 10-week-old WASp knockout (WAS-KO) and wild-type (WT) mice to identify the transcriptomic consequences of WASp loss in a major secondary lymphoid organ, and identified 1,964 differentially expressed genes (DEGs) between WAS-KO and WT spleen, of which 996 were upregulated and 968 were downregulated in the knockout mice (Liang, Yang, *et al.*, 2023). Gene Ontology, KEGG pathway, DisGeNET, and Reactome enrichment analyses showed that upregulated genes in WAS-KO spleen were associated with cellular senescence, the p53 signalling pathway, cell cycle, and heme metabolism. In contrast, downregulated genes were associated with platelet activation, antigen processing and presentation, and the MAPK, calcium, and cAMP signal transduction pathways (Liang, Yang, *et al.*, 2023). Hence, WAS pathogenesis has evolved through dysregulated senescence and apoptosis of blood cells in the spleen (Liang, Yang, *et al.*, 2023). This WASp-dependent differentially expressed gene (DEG) set demonstrates the basis of the WAS-regulated gene list analyzed in this study.

Another study conducted transcriptomic analysis at single-cell resolution, performing single-cell RNA sequencing on wild type (WT) and WAS-KO mouse spleen and detecting 30 clusters comprising nine major immune cell types, including B cells, T cells, dendritic cells (DCs), NK cells, monocytes, macrophages, and granulocytes (Liang, Peng, *et al.*, 2023). This analysis illustrated that the proportions of B cells, T cells, and NK cells were diminished in the WAS-KO spleen, while granulocytes and monocytes were elevated, and each cell type displayed a distinct DEG sign linked to chronic inflammation, impaired signal transduction, or apoptosis (Liang, Peng, *et al.*, 2023). Cross-comparison of the DEGs across B cells, T cells, DCs, and NK cells identified eight common genes (*Ubb*, *S100a8*, *S100a9*, *Gm42418*, *Camp*, *Lars2*, *Ngp*, and *Hba-a1*) correlated with chronic inflammatory response, intrinsic apoptosis, and neutrophil chemotaxis. Moreover, DEGs were distributed across macrophages, monocytes, and granulocytes, while four further genes (*Was*, *Isg15*, *Ifi2712a*, and *Irf7*) linked to interferon-beta production and antiviral defence activity (Liang, Peng, *et al.*, 2023). These single-cell findings refine the transcriptome-level analysis of the spleen conducted by Liang and his team by attributing specific DEG reflections to individual immune cell populations rather than the entire spleen tissue (Liang, Peng, *et al.*, 2023).

Furthermore, these studies indicate that WASp deficiency yields broad, cell-based transcriptional dysregulation in the spleen that often extends beyond the classical actin-cytoskeletal defect, incorporating senescence, apoptosis, and inflammatory signalling pathways (Liang, Peng, *et al.*, 2023; Liang, Yang, *et al.*, 2023).

1.4. Biological relationship between SAP and WASp.

SAP and WASp are different proteins with fundamentally different primary molecular functions, but both participate in signalling processes that regulate immune-cell activation. SAP, encoded by the *SH2D1A* gene, is an intracellular adaptor protein that interacts with SLAM-family immune receptors. Through these interactions, SAP can influence the balance between activating and inhibitory signalling molecules associated with SLAM-family receptors (Li *et al.*, 2003). WASp, encoded by the *WAS* gene, is an actin-regulatory protein expressed predominantly in hematopoietic cells and plays a key role in connecting cellular signalling to actin cytoskeletal remodelling (Sun *et al.*, 2019). WASp is recruited and activated following TCR stimulation and contributes to the formation of actin structures required for effective T-cell activation (Kumari *et al.*, 2015). Thus, SAP and WASp can initially be viewed as components of different functional aspects of T-cell signalling. SAP primarily functions as a signalling adaptor, whereas WASp connects signalling pathways with cytoskeletal regulation.

Despite these differences, evidence suggests their signalling pathways can be interconnected. A recent study demonstrated that WASp interacts with DGK α and inhibits its activity following strong TCR stimulation (Velnati *et al.*, 2023). The proteins NCK1 and CDC42 were identified as components that connect WASp-mediated regulation of DGK α with SAP and the TCR signalosome (Velnati *et al.*, 2023). This pathway was shown to contribute to a full IL-2 response in primary human T cells (Velnati *et al.*, 2023). Importantly, the same study also demonstrated that enhanced DAG signalling caused by DGK α inhibition could restore apoptosis sensitivity in SAP-silenced T cells (Velnati *et al.*, 2023). These findings provide direct evidence that SAP and WASp can participate in a connected signalling network rather than functioning in completely isolated pathways.

Therefore, the relationship between these proteins is particularly relevant when considering the transcriptional consequences resulting from their loss. SAP deficiency can alter signalling through SLAM-family receptors and may also affect cellular responses to TCR restimulation (Li *et al.*, 2003; Snow *et al.*, 2009). In contrast, WASp deficiency can interfere with actin organisation, proximal TCR signalling, calcium-dependent signalling, and cellular activation (Kumari *et al.*, 2015; Rattanasri *et al.*, 2026). Experimental studies have shown that WASp deficiency can affect not only cytoskeletal organisation but also gene transcription and broader immune-cell functions (Sun *et al.*, 2019). More recent evidence also indicates that WASp can influence proximal TCR signalling, showing that its function is not limited exclusively to a structural role (Rattanasri *et al.*, 2026). Therefore, although SAP and WASp are not functionally equivalent, disruption of either protein can alter downstream signalling networks and ultimately influence gene expression.

This raises an important biological question: whether the transcriptional changes caused by SAP deficiency and WASp deficiency are partly shared or are mainly independent. Shared transcriptional changes would indicate that different upstream signalling defects may ultimately converge on common downstream cellular responses. Conversely, largely different gene sets would suggest that the two proteins mainly influence separate transcriptional programmes despite their interaction within some signalling pathways. Understanding this distinction is particularly important for determining whether the similarities observed between XLP and WAS can be linked to common molecular responses. Therefore, comparative analysis of SAP- and WASp-regulated genes could help assess the extent of transcriptional convergence between these two signalling systems.

1.5. The clinical and immunological connection between XLP and WAS.

X-linked lymphoproliferative disease (XLP) and Wiskott–Aldrich syndrome (WAS) are two distinct inherited immune disorders caused by defects in different genes and proteins. XLP type 1 is mainly caused by pathogenic variants in *SH2D1A*, which encodes the SAP. In contrast, WAS results from pathogenic variants in the *WAS* gene, which encodes WASp (Li *et al.*, 2003; Snow *et al.*, 2009). Although SAP and WASp have different molecular functions, both disorders can lead to

abnormalities in lymphocyte function and immune regulation. XLP is particularly associated with abnormal immune responses to Epstein–Barr virus (EBV), severe infectious mononucleosis, haemophagocytic lymphohistiocytosis, dysgammaglobulinaemia, and lymphoproliferative disease (Snow *et al.*, 2009).

WAS is characterised classically by thrombocytopenia with small platelets, eczema, and recurrent infections. However, the clinical spectrum of WAS is broader than this classical presentation and can include autoimmunity, chronic inflammation, immune dysregulation, and malignancy (Vallée, Albert and Pai, 2025). Recent studies have emphasised that disease severity can vary considerably between individuals and may be influenced by the underlying WAS gene variant (Vallée, Albert and Pai, 2025). WASp deficiency affects several immune-cell functions because WASp participates in actin-dependent cellular processes and signalling pathways involved in lymphocyte activation (Calixto *et al.*, 2023). More recent work on immunoactinopathies has highlighted the importance of cytoskeletal regulation in immune-cell function and demonstrated that defects in actin-regulatory proteins can result in combined immunodeficiency and immune-dysregulation phenotypes (Hiensch, Dupré and Salzer, 2025).

Despite their different genetic causes, XLP and WAS can exhibit overlapping features of immune dysfunction. Both disorders may involve abnormal lymphocyte responses, impaired immune homeostasis, increased susceptibility to infection, autoimmune manifestations, and an increased risk of malignancy (Snow *et al.*, 2009; Vallée, Albert and Pai, 2025). However, similar clinical features do not prove that the two diseases share a single molecular mechanism. Instead, similar clinical outcomes can arise through different signalling pathways that affect common biological processes such as lymphocyte activation, survival, proliferation, and immune regulation. This distinction is important when considering the relationship between SAP and WASp.

The biological connection between the two disorders becomes particularly relevant in the context of T-cell signalling. SAP contributes to signalling through SLAM-family receptors and regulates important lymphocyte responses. In contrast, WASp connects receptor signalling to actin remodelling and downstream cellular activation (Li *et al.*, 2003; Kumari *et al.*, 2015). Experimental evidence has

demonstrated that these two proteins can participate in a connected signalling pathway involving DGK α during strong TCR activation (Velnati *et al.*, 2023). Therefore, SAP and WASp deficiency may converge on downstream signalling or gene expression, even though their initiating molecular defects are different.

This possibility provides an important scientific rationale for comparing SAP- and WASp-regulated transcriptional profiles. If the two deficiencies produce similar changes in gene expression, these genes may help identify biological processes that are sensitive to disruption of both signalling systems. Conversely, if most affected genes differ, the clinical similarities between XLP and WAS may instead arise from independent molecular pathways that ultimately disrupt related immune functions. Thus, comparing the transcriptional consequences of SAP and WASp deficiency may help distinguish between shared downstream responses and disease-specific molecular effects.

1.6. Transcriptomic and gene expression analysis.

Transcriptomics is the study of RNA molecules produced by cells or tissues under specific biological conditions. It provides a broad view of gene activity and can identify changes in gene expression associated with cellular activation, differentiation, disease, or genetic perturbation (Antigen and Cells, 2023). RNA sequencing (RNA-seq) is a major approach for transcriptomic analysis because it can measure large numbers of transcripts across the genome and identify differentially expressed genes between experimental conditions. Transcriptomic approaches are increasingly used to investigate T-cell biology because T-cell activation involves rapid and extensive changes in gene expression (Yang *et al.*, 2023). For example, recent research has demonstrated that T-cell activation requires coordinated regulation of RNA processing, RNA stability, protein synthesis, and expression of genes involved in cellular growth and proliferation (Knop *et al.*, 2024).

Transcriptomic analysis can be used to examine the downstream consequences of altering a particular signalling protein. When a gene such as SAP or WAS is silenced or knocked out, comparison with an appropriate control condition can identify genes whose expression increases or decreases following loss of that protein. Genes showing increased expression after loss of a protein may be normally restrained by

that protein, whereas genes showing reduced expression may depend on the protein for normal expression. This approach is particularly useful for signalling proteins because their biological effects may extend beyond their immediate molecular interactions and influence broader transcriptional programmes (Antigen and Cells, 2023; Yang *et al.*, 2023).

Studies using RNA-seq and other high-throughput approaches also demonstrate the usefulness of transcriptomics in T-cell research by characterising activated or genetically modified T cells. Transcriptomic analysis can provide a global view of the cellular state and reveal gene groups associated with particular biological processes or signalling pathways (Antigen and Cells, 2023; Yang *et al.*, 2023). Importantly, transcriptomic data can be analysed at different levels. Researchers can compare individual differentially expressed genes between experimental conditions, and use pathway and functional enrichment analyses to determine whether gene sets are associated with particular biological processes (Yang *et al.*, 2023).

In the context of SAP and WASp, transcriptomic analysis can assess whether loss of these proteins produces similar or distinct gene-expression patterns. Studies have already demonstrated substantial transcriptional changes following WASp deficiency. For example, RNA-seq analysis of WAS-knockout mouse spleen identified extensive differential gene expression involving processes such as cellular senescence, apoptosis, platelet activation, antigen presentation, and signalling pathways. More recent work has also shown that WASp deficiency can affect transcriptional processes in T cells, further supporting the idea that WASp function extends beyond cytoskeletal regulation (Pradeep *et al.*, 2024). The mouse spleen study also used transcriptomic analysis to investigate gene-expression differences between WAS-deficient and wild-type animals, demonstrating the value of transcriptomic approaches for identifying molecular changes associated with WAS deficiency (Liang, Yang, *et al.*, 2023).

However, comparisons of transcriptomic datasets must account for experimental context. Gene-expression profiles can vary according to species, cell type, tissue source, activation state, duration of stimulation, gene-silencing or knockout strategy, sequencing technology, and the statistical criteria used to define differentially expressed genes. Therefore, a lack of overlap between two datasets does not

necessarily mean that the underlying proteins have no biological relationship. Similarly, overlapping genes do not by themselves demonstrate that the proteins directly regulate the same genes. In the present study, used transcriptomic comparison to identify the extent of gene-level convergence between SAP- and WASp-regulated responses. The analysis of published datasets is complemented by comparison of in-house human CD8⁺ lymphocyte datasets generated under a more closely matched experimental setting. This combined approach provides a framework for distinguishing shared transcriptional responses from genes that are specific to SAP or WASp deficiency.

2. RATIONALE, AIMS AND OBJECTIVES

Studies performed in the UPO biochemistry laboratory showed that DGK α , SAP and WASp participate in a signalling pathway in T cells (Snow *et al.*, 2009; Liang, Yang, *et al.*, 2023; Velnati *et al.*, 2023). This suggests they should regulate a common gene set. However, XLP and WAS are two separate genetic diseases with different features, suggesting that SAP and WASp also regulate specific genes. This is the primary rationale for comparing SAP-regulated and WAS-regulated gene expression programs.

This study aims to determine whether the SAP-regulated and WAS-regulated genes in a similar manner, using a standardized gene identifier-based comparison of previously published gene expression data and also data obtained in the UPO biochemistry laboratory.

Workflow:

- 1) Extract and organized the supplementary data lists of upregulated and downregulated genes reported in the published literature (Snow *et al.*, 2009; Liang, Yang, *et al.*, 2023) and in-house gene expression datasets separately.
- 2) To standardize and organize SAP- and WAS-regulated differentially expressed gene identifiers using Ensemble Gene IDs to enable accurate comparison across published and in-house datasets and species, while excluding doubt caused by nomenclature differences, gene symbol aliasing and duplicate identifiers.
- 3) To represent the degree of overlap graphically using VENN diagrams and interpret the biological significance of the unique gene sets and commonly up-regulated and down-regulated genes by both SAP and WAS.
- 4) To evaluate the overlap between the literature and in-house derived datasets and determine whether the observed gene expression patterns are consistent across different experimental systems.
- 5) To interpret the shared and dataset-specific gene sets in the context of the biological functions of SAP and WASp and assess whether the findings support a common transcriptional mechanism or indicate that SAP and WASp predominantly regulate distinct molecular programs despite their involvement in related T-cell signalling pathways.

3. MATERIALS AND METHODS

3.1. Data sources.

This analysis used two published primary literature sources and their associated supplementary gene expression datasets.

3.2. SAP dataset.

The SAP-regulated gene list was captured from a supplementary microarray dataset associated with the study, which was done by Snow *et al.*, “Restimulation induced apoptosis of T cells is impaired in patients with X-linked lymphoproliferative disease caused by SAP deficiency” (Snow *et al.*, 2009). The corresponding author provided this dataset directly through personal communication. The supplementary genes were originally identified by comparing gene expression after 6 hours of T cell receptor (TCR) restimulation. Moreover, the comparison is between normal donor T cells transfected with SAP-specific small interfering RNA (siRNA) relative to cells transfected with non-specific (NS) control siRNA.

In the SAP-regulated dataset, the average gene expression level in the non-silenced control cell (NS-AVG) is much higher than the average gene expression level in which SAP has been silenced (SAP-AVG); these genes are normally re-stimulated, and when SAP is absent or silenced, their expression decreases. Hence, these genes are labelled "SAP-upregulated" because SAP is required for them to be upregulated. When NS-AVG levels are lower than SAP-AVG levels, the gene is normally kept at low expression or suppressed by SAP, and silencing SAP reveals loss of suppression and increased gene expression. Hence, these genes are termed "SAP-downregulated", since they suppress SAP. The raw supplementary dataset contained 357 upregulated and 90 downregulated transcripts.

3.3. WAS dataset.

The WAS-regulated gene list was obtained from the publicly available supplementary dataset in “Experimental and Therapeutic Medicine, 2023” (Liang, Yang, *et al.*, 2023). This dataset was obtained from the article's supplementary materials and lists all 1,964 distinctly expressed genes (DEGs) identified by RNA sequencing of spleen tissue comparing WAS knockout (WAS-KO) versus wild-type

(WT) mice. Moreover, WAS loss in WAS-KO mice increases gene expression relative to WT mice, as indicated by the log2FC threshold. According to this concern, WAS negatively regulates these upregulated genes, since these genes are highly expressed in WAS-KO mice than in WT mice and WAS normally suppresses these genes. These genes are known as “WAS upregulated” genes. Nevertheless, WAS-KO mice show lower gene expression than WT mice due to WAS loss. Hence, these genes are defined as “WAS downregulated,” since their expression is sharply diminished as a result of WAS loss. In this dataset, 996 genes were upregulated and 968 genes were downregulated in WAS-KO relative to WT spleen.

3.4. In-house-generated SAP and WAS data set.

A parallel data set of SAP- and WAS-modulated genes was obtained by our laboratory by RNA sequencing of CD8⁺ human lymphocytes silenced for either SAP or WAS and, after 4 days, re-stimulated with CD3 and CD28 activating antibodies for 4 hours.

The full gene list and its characterization are still unpublished and will be described in the thesis of other students. For the present comparative analysis, the finalized Ensembl Gene ID lists were separated into SAP-upregulated, SAP-downregulated, WAS-upregulated, and WAS-downregulated sets. Gene symbols were retained for interpretation, while Ensembl Gene IDs were used as the primary matching identifier.

The published WAS dataset originated from mouse spleen, whereas the laboratory WAS dataset was generated from human CD⁺ lymphocytes. Direct comparison of species-specific identifiers is not appropriate, and the published mouse WAS dataset was represented using the human orthology mapping framework established during the preceding literature analysis. The in-house datasets were already human and were therefore directly comparable at the standardized human Ensembl Gene ID (ENSG) level.

3.5. Gene Identifier Standardization.

Gene symbols from both dataset lists were manually cross-referenced with the gene database, and a unique numeric Ensembl Gene (ENSG) ID was extracted for each gene. The results were then compiled into a single reference spreadsheet containing four parallel gene lists, namely SAP-upregulated, SAP-downregulated, WAS-upregulated, and WAS-downregulated, each paired with the corresponding gene symbols for the Gene ID.

Gene symbols that could not be mapped to a valid, unique ENSG Gene ID were excluded from further analysis, including uncharacterized expressed sequence tags (ESTs), unnamed cDNA clones, and predicted transcripts without an assigned identifier. After this filtering step, the SAP list was reduced from 357 upregulated genes to 315 and from 90 downregulated genes to 74 with resolvable Gene IDs. The WAS list was reduced from 996 upregulated genes to 718 and from 968 downregulated genes to 433 with resolvable Gene IDs.

The published SAP and WAS datasets were derived from different species; an additional orthology-mapping step was required before directly comparing their gene sets. The SAP dataset was obtained from human cells, and the WAS dataset was generated from mouse spleen samples. Therefore, the mouse WAS gene identifiers could not be directly compared with the human SAP gene identifiers, even when the corresponding gene symbols appeared similar.

To address this species difference, the differentially expressed genes identified in the published mouse WAS dataset were mapped to their corresponding human orthologues using g:Orth in g:Profiler. This orthology conversion was performed before the comparative gene set analysis, and the resulting human orthologue identifiers were subsequently used for comparison with the human SAP dataset, as well as with the in-house SAP and WAS datasets.

Gene identifiers with clear orthologous relationships were retained for the analysis, whereas ambiguous or unmapped genes were recorded separately and not forcibly included in the primary comparative analysis. This approach ensured that all four datasets were represented within a common human gene identifier framework,

allowing subsequent gene overlap analyses and Venn diagram comparisons to be performed using biologically comparable gene sets.

3.6. Comparative (VENN) Analysis.

The SAP upregulated and WAS upregulated Gene ID columns were cross-referenced in the newly organized reference spreadsheet using a 'COUNTIF'-based Boolean matching formula. This formula checks whether a value from one list appears anywhere in another list. Each SAP upregulated gene was flagged as 'MATCH' if its Gene ID appeared anywhere in the WAS upregulated column, and as 'NOMATCH' if it did not. An identical procedure was repeated separately for the SAP downregulated and WAS downregulated Gene ID columns. The total number of 'MATCH' flags in each comparison provided the number of genes commonly regulated by both SAP and WASp. These 'MATCH' and 'NOMATCH' results were then used to generate two VENN diagrams, one for the upregulated comparison and the other for the downregulated comparison. Each diagram visualized the number of genes unique to the SAP list, unique to the WAS list, and shared between both lists.

4. RESULTS

4.1. Comparison of SAP- and WAS-regulated gene sets.

We want to understand whether SAP and WAS regulate the same gene set. We searched the literature and found a study in human lymphocytes with SAP silenced (Snow et al., 2009) and one study in WAS-KO mice (Liang, Yang, et al., 2023). The SAP study was conducted only on isolated T cells at a single time point, 6 hours after TCR restimulation, whereas the WAS study was conducted on the whole spleen, an organ containing many different cell types.

After standardizing gene identifiers, the SAP-regulated gene list contained 315 upregulated and 74 downregulated genes with resolvable ENSG Gene IDs; the WAS-regulated gene list contained 718 upregulated and 433 downregulated genes (Table 1). The WAS-regulated dataset was significantly larger than the SAP-regulated dataset in both upregulated and downregulated directions.

Comparison	SAP-regulated genes (n)	WAS-regulated genes (n)	Regulated by both SAP and WAS (n)	% of shared SAP list	% of shared WAS list
Upregulated	315	718	12	3.8%	1.6%
Downregulated	74	433	1	1.4%	0.2%

Table 1. Quantity of the SAP and WAS regulated gene sets and the genes shared between them after standardization to Ensembl Gene IDs.

We manually searched for shared genes; Venn diagram analysis (Figure 1) showed that the two lists share a minimal subset of commonly regulated genes. Therefore, among upregulated genes, 303 of 315 SAP upregulated genes (96.2%) and 706 of 718 WAS upregulated genes (98.3%) were unique to a single dataset, while 12 genes 3.8% of the SAP and 1.6% of the WAS upregulated list regulated by both SAP and WAS. Moreover, among downregulated genes, the overlap was even smaller due to 73 of 74 SAP downregulated genes (98.6%) and 432 of 433 WAS downregulated genes (99.8%) were unique to a single dataset, and 01 gene (*FGF2*) was shared between both SAP and WAS

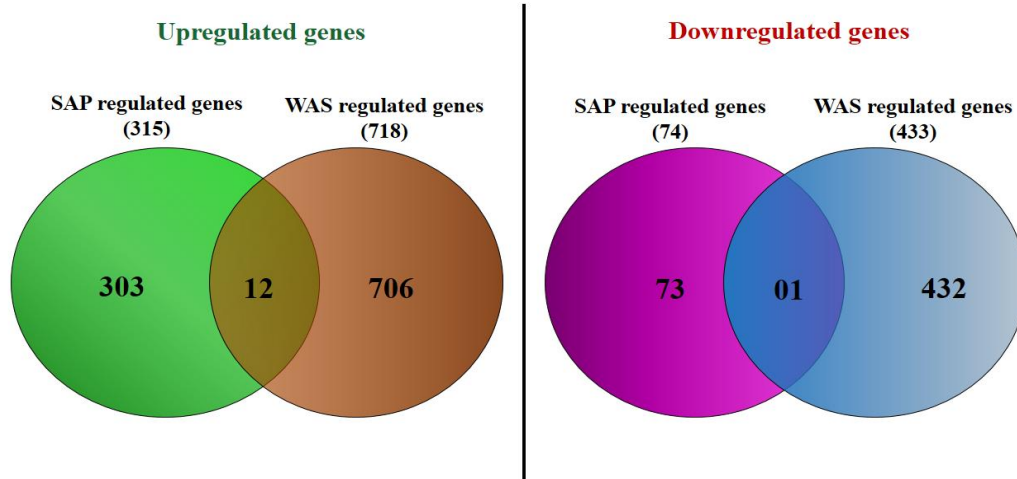


Figure 1. Venn diagrams comparing SAP-regulated and WAS-regulated genes. Left: Upregulated gene sets (315 SAP-regulated vs 718 WAS-regulated genes and 12 genes regulated by both SAP and WAS). Right: Downregulated gene sets (74 SAP-regulated vs 433 WAS-regulated and 01 gene regulated by both SAP and WAS). Data are from the literature (Snow *et al.*, 2009; Liang, Yang, *et al.*, 2023).

The lists of the 12 genes upregulated by both SAP and WAS, and one gene downregulated by both SAP and WAS, are listed in Table 2 and Table 3, respectively together with their shared Ensembl Gene ID and their symbol as reported in the human (SAP) and mouse (WAS) source datasets.

Number	Gene symbol for SAP dataset (human)	ENSG Gene ID	Gene symbol for WAS dataset (mouse)
1	<i>BCL2L1</i>	ENSG00000171552	<i>BCL2L1</i>
2	<i>EPB42</i>	ENSG00000166947	<i>EPB42</i>
3	<i>FABP5</i>	ENSG00000164687	<i>FABP5</i>
4	<i>FNIP2</i>	ENSG00000052795	<i>FNIP2</i>
5	<i>GOLIM4</i>	ENSG00000173905	<i>GOLIM4</i>
6	<i>IL10</i>	ENSG00000136634	<i>IL10</i>
7	<i>MLF1</i>	ENSG00000178053	<i>MLF1</i>
8	<i>NPHP1</i>	ENSG00000144061	<i>NPHP1</i>
9	<i>C1orf198/ANK1</i>	ENSG00000119280	<i>C1orf198</i>
10	<i>SLAMF9</i>	ENSG00000162723	<i>SLAMF9</i>
11	<i>TIMD4</i>	ENSG00000145850	<i>TIMD4</i>
12	<i>TMCC2</i>	ENSG00000133069	<i>TMCC2</i>

Table 2. The 12 genes commonly upregulated in both the SAP-dependent TCR restimulation gene set and the WAS-KO spleen transcriptome.

Number	Gene symbol for SAP dataset (human)	ENSG Gene ID	Gene symbol for WAS dataset (mouse)
01	<i>FGF2</i>	ENSG00000138685	<i>FGF2</i>

Table 3. The single gene commonly downregulated in both the SAP-dependent TCR re-stimulation gene set and the WAS-KO spleen transcriptome.

Among the 12 commonly upregulated genes, *BCL2L1* encodes an anti-apoptotic protein and directly connects to the apoptosis regulatory pathway and cell survival and also common in both XLP (SAP) and WAS disorders (Rawlings *et al.*, 1999). *IL-10* is a well-characterized immune regulatory cytokine gene associated with immune dysregulation in both XLP and WAS (Du *et al.*, 2015). *SLAMF9* and *TIMD4* encode lymphocyte surface receptors from the SLAM and TIM receptor families, respectively (Dragovich and Mor, 2019; Liu *et al.*, 2020). These belong to the same broad category of immunoregulatory surface receptors as NTB-A (SLAMF6), the SLAM family receptor directly linked to SAP-dependent RICD in the source study (Cruz, Davidson and Roncagalli, 2006). The remaining shared genes *EPB42*, *FABP5*, *FNIP2*, *GOLIM4*, *MLF1*, *NPHP1*, *ANK1/C1orf198* and *TMCC2* cover a variety of functions in cytoskeletal and membrane organization, metabolism and vesicle trafficking (Puthenveedu *et al.*, 2003; Satchwell *et al.*, 2009; Lim *et al.*, 2010; Hopkins, 2013; Hasumi *et al.*, 2015; Warren *et al.*, 2023). They do not obviously cluster around a single shared biological pathway. The single common downregulated gene, *FGF2* (fibroblast growth factor 2), is a growth factor involved in cell proliferation and cell survival signalling, and this is the single gene found to be lower in both the SAP-deficient and WAS-deficient conditions (Nickle, Ko and Merrill, 2025).

4.2. Comparison of In-house SAP- and WAS-regulated gene set.

Comparison	SAP-regulated genes (n) laboratory	WAS-regulated genes (n) laboratory	Commonly regulated by SAP and WAS (n)	SAP unique (n)	WAS unique (n)
Upregulated	422	184	69	353	115
Downregulated	776	1085	71	705	1014

Table 4. Summary of the In-house SAP and WAS gene-set comparison using standardized Ensembl Gene IDs.

The limited overlap between the two lists may reflect the different approaches used in the two reference papers. We repeated the analysis on a series of RNA sequencing experiments conducted in our lab, with CD8 lymphocytes silenced in parallel for SAP or WAS. In our analysis, we identified 422 SAP-upregulated genes, 184 WAS-upregulated genes, 776 SAP-downregulated genes and 1085 WAS-downregulated genes (unpublished data). Comparison of the SAP and WAS gene sets identified 69 genes commonly upregulated and 71 genes commonly downregulated.

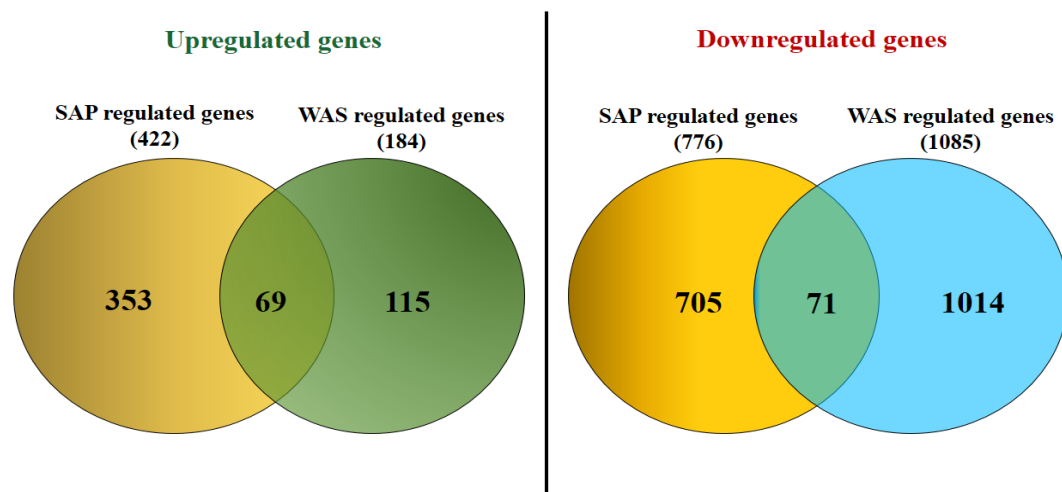


Figure 2. Venn diagrams comparing SAP-regulated and WAS-regulated genes in the In-house datasets. Left: Upregulated gene sets (422 SAP-regulated vs 184 WAS-regulated genes and 69 genes regulated by both SAP and WAS). Right: Downregulated gene sets (776 SAP-regulated vs 1085 WAS-regulated and 71 gene regulated by both SAP and WAS).

As expected, the number of genes regulated by both SAP and WAS is higher but still small, indicating that SAP and WAS do not regulate the same gene set.

To further assess the consistency of SAP- and WAS-regulated genes across datasets, the genes commonly regulated by both SAP and WAS in the literature-derived datasets were compared with the corresponding common SAP and WAS gene sets identified in the in-house dataset. The 12 genes commonly upregulated by SAP and WAS in the literature-derived datasets were not present among the 69 genes commonly upregulated by SAP and WAS in the in-house dataset. Likewise, the single gene commonly downregulated by SAP and WAS in the literature-derived datasets was not present among the 71 genes commonly downregulated by SAP and WAS in the in-house dataset. Thus, the literature-derived and in-house datasets had no overlap for genes commonly upregulated or downregulated by both SAP and WAS.

4.3. Comparison of the laboratory datasets with the literature datasets.

Regulation	Laboratory (n)	Literature (n)	Common gene (n)
SAP upregulated	422	315	0
SAP downregulated	776	74	0
WAS upregulated	184	718	4
WAS downregulated	1085	433	13

Table 5. Comparison of in-house and literature derived SAP and WAS gene sets.

The gene sets produced in our laboratory were next superimposed on the corresponding literature-derived sets to assess reproducibility between the In-house and published analyses. For SAP, zero common Ensembl Gene IDs were identified between the laboratory and literature lists in either the upregulated or downregulated comparison. For WAS, 04 common upregulated genes and 13 common downregulated genes were identified.

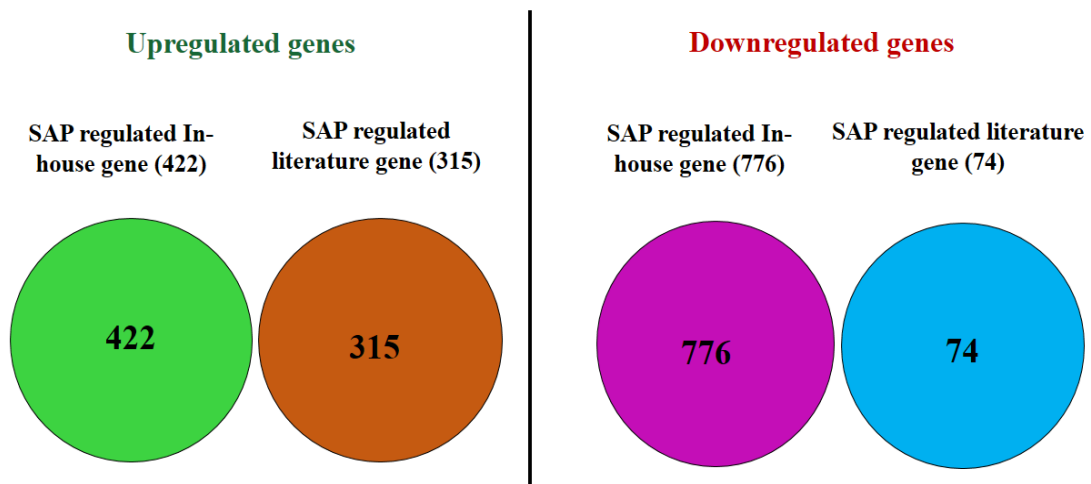


Figure 3. Venn diagram comparing In-house and literature SAP upregulated gene sets and SAP downregulated gene sets. Left: Upregulated gene sets (422 In-house SAP regulated vs 315 literature SAP regulated genes and zero shared genes were identified by SAP upregulated). Right: Downregulated gene sets (776 In-house SAP regulated vs 74 literature SAP regulated genes and zero shared genes were identified by SAP downregulated).

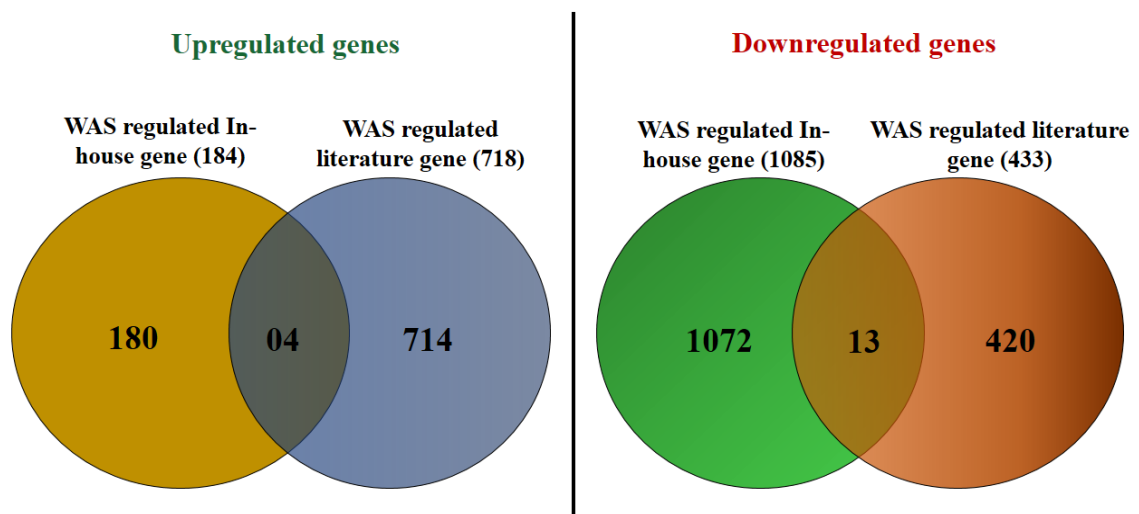


Figure 4. Venn diagram comparing In-house and literature WAS-upregulated gene sets and WAS-downregulated gene sets. Left: Upregulated gene sets (184 in-house WAS-regulated vs 718 literature SAP-regulated genes, and 04 shared genes were identified by WAS upregulated). Right: Downregulated gene sets (1085 in-house WAS-regulated vs 433 literature WAS-regulated genes, and 13 shared genes were identified by WAS-downregulated).

Those data indicate that our dataset does not align with what is reported in the literature, suggesting that the results of those experiments are strongly dependent on the experimental setup and the original materials.

The 04 genes upregulated by both the literature WAS dataset and the in-house WAS dataset, and the 13 genes downregulated by both the literature WAS dataset and the in-house WAS dataset, are listed in Table 6 and Table 7, respectively, together with their shared Ensembl Gene ID and symbol.

Number	Gene symbol for WAS In-house dataset (human)	ENSG Gene ID	Gene symbol for WAS literature dataset (mouse)
1	<i>ATP1B2</i>	ENSG00000129244	<i>ATP1B2</i>
2	<i>CCSER1</i>	ENSG00000184305	<i>CCSER1</i>
3	<i>MAFB</i>	ENSG00000204103	<i>MAFB</i>
4	<i>MT-ND3</i>	ENSG00000198840	<i>MT-ND3</i>

Table 6. The 04 genes commonly upregulated in both the in-house WAS-dependent CD8+ human lymphocyte gene set and the literature WAS-KO spleen transcriptome set.

Number	Gene symbol for WAS In-house dataset (human)	ENSG Gene ID	Gene symbol for WAS literature dataset (mouse)
1	<i>APLN</i>	ENSG00000171388	<i>APLN</i>
2	<i>CFAP157</i>	ENSG00000160401	<i>CFAP157</i>
3	<i>CPSF4L</i>	ENSG00000187959	<i>CPSF4L</i>
4	<i>DNAH17</i>	ENSG00000187775	<i>DNAH17</i>
5	<i>GAPDHS</i>	ENSG00000105679	<i>GAPDHS</i>
6	<i>IFITM5</i>	ENSG00000206013	<i>IFITM5</i>
7	<i>IGFBP6</i>	ENSG00000167779	<i>IGFBP6</i>
8	<i>PKDIL3</i>	ENSG00000277481	<i>PKDIL3</i>
9	<i>PMEL</i>	ENSG00000185664	<i>PMEL</i>
10	<i>PRR33</i>	ENSG00000283787	<i>PRR33</i>
11	<i>TRIM7</i>	ENSG00000146054	<i>TRIM7</i>
12	<i>VIL1</i>	ENSG00000127831	<i>VIL1</i>
13	<i>WAS</i>	ENSG00000015285	<i>WAS</i>

Table 7. The 04 genes are commonly downregulated in both the in-house WAS-dependent CD8+ human lymphocyte gene set and the literature WAS-KO spleen transcriptome set.

Among the four commonly upregulated genes identified between the in-house and literature-derived WAS datasets, *ATPIB2* encodes the $\beta 2$ subunit of the Na^+/K^+ -ATPase and contributes to the maintenance of cellular sodium and potassium gradients and plasma membrane function (Pelin *et al.*, 2020). *MAFB* encodes a basic leucine zipper transcription factor involved in lineage-specific hematopoiesis and gene expression regulation (Saiga *et al.*, 2021). Similarly, *MT-ND3* encodes a subunit of mitochondrial respiratory chain complex I and contributes to the mitochondrial electron transport process (Nahar *et al.*, 2024). The *CCSER1* is associated with

cellular regulatory processes. However, its biological function remains less well characterized compared with the other common genes. Therefore, these four upregulated genes represent different biological functions, including membrane and ion homeostasis, transcriptional regulation and mitochondrial function, rather than belonging to a single clearly defined biological pathway.

The 13 commonly downregulated genes also represent a heterogeneous group of genes with diverse biological functions. *APLN* encodes apelin, a peptide involved in vascular signalling, endothelial function and vascular remodelling (Marsault *et al.*, 2020; Park and Adams, 2026). Also, *IGFBP6* encodes an insulin-like growth factor-binding protein involved in IGF-related signalling and cellular processes such as cell migration (Dailly *et al.*, 2001). *VILL1* encodes villin, an actin-associated protein involved in actin filament organization and cytoskeletal remodelling (Epithelial *et al.*, 2026). This is particularly relevant in the context of *WAS* because WASp plays an important role in regulating actin cytoskeletal organization and immune cell function (Blundell *et al.*, 2010; Barda-saad, 2013). The *PMEL* encodes a transmembrane glycoprotein involved in the structural organization of pre-melanosomes (Ma *et al.*, 2025), whereas *TRIM7* encodes an E3 ubiquitin ligase involved in protein regulation and cellular signalling (Fan *et al.*, 2025). The *PKDIL3* encodes a polycystin-family membrane protein that may contribute to cation-channel activity (Yu *et al.*, 2013), while *PRR33* is a relatively poorly characterized protein-coding gene that is predicted to be associated with cellular responses to wounding (Fu *et al.*, 2024). The remaining genes, including *CFAP157*, *CPSF4L*, *DNAH17*, *GAPDHS* and *IFITM5*, have more specialized or incompletely characterized biological functions. Finally, *WAS* itself encodes the Wiskott-Aldrich syndrome protein (WASp), which plays an important role in the regulation of actin polymerization, immune cell signalling and immune synapse formation (Liang, Yang, *et al.*, 2023). Overall, the functional diversity of these 13 downregulated genes and four upregulated genes does not indicate a single common biological pathway.

5. DISCUSSION

This study aimed to determine whether the genes modulated in lymphocytes by SAP deficiency (the molecular basis of XLP) and those regulated in lymphocytes by WASp deficiency (the molecular basis of WAS) are the same. The proteins that cause these two disease conditions (SAP and WASp) are completely different at the molecular level. Still, studies performed at the Biochemistry Laboratory of the UPO University showed that they participate in a common signalling pathway (Velnati *et al.*, 2023). If a single common genetic basis exists, it could provide a unifying transcriptional explanation for the overlapping clinical features, which are commonly detected in both diseases: lymphoproliferation, defective immune homeostasis, and susceptibility to autoimmune and malignant disease (Booth *et al.*, 2011; Buchbinder, Nugent and Fillipovich, 2014)

We retrieved a list of SAP-regulated and WAS-regulated genes from the literature (Snow *et al.*, 2009; Liang, Yang, *et al.*, 2023). The Venn diagram analysis gave a clear solution: SAP and WASp largely regulate minimally overlapping gene expression programs. Indeed, only 12 of 315 genes upregulated by SAP (3.8%) and 1 of 74 genes downregulated by SAP (1.4%) were also found among the genes regulated by WAS. This analysis shows that over 96% of genes regulated by SAP and over 98% of genes regulated by WAS were unique to their respective datasets.

This is consistent with the underlying biological nature of these two proteins. SAP is a signalling adaptor protein that acts early in TCR and SLAM receptor signal transduction in T lymphocytes, specifically during antigen restimulation (Snow *et al.*, 2009; Shah *et al.*, 2021). WASp is a cytoskeletal regulator expressed broadly across hematopoietic cell lineages and sampled here at the level of a heterogeneous whole organ (spleen) under steady-state conditions rather than after a specific antigenic stimulation (Liang, Peng, *et al.*, 2023; Sasson *et al.*, 2026). This suggests that both diseases converge on similar clinical phenotypes through parallel, largely independent molecular routes rather than through a single shared gene expression program. However, the small set of genes distributed between the two conditions may be biologically meaningful rather than coincidental. *BCL2L1* is a core anti-apoptotic regulator of the intrinsic mitochondrial apoptosis pathway (Rawlings *et al.*, 1999).

The distributed upregulation of this gene in both the SAP- and WASp-deficient settings is consistent with the broader theme noted independently in both source publications namely, that dysregulated lymphocyte and blood cell apoptosis contributes to the pathogenesis of both XLP and WAS, even though the two proteins act through different upstream signals (TCR restimulation for SAP and steady-state splenic homeostasis for WASp (Snow *et al.*, 2009; Liang, Yang, *et al.*, 2023; Sasson *et al.*, 2026). The shared upregulation of *IL10*, an immunoregulatory cytokine gene, and of the surface receptor genes *SLAMF9* and *TIMD4* is similarly consistent with the shared theme of immune dysregulation in both diseases (Du *et al.*, 2015; Dragovich and Mor, 2019; Liu *et al.*, 2020). It also echoes the specific mechanistic role of a related SLAM-family receptor, NTB-A, involved in SAP-dependent RICD (Cruz, Davidson and Roncagalli, 2006). The single downregulated gene, *FGF2*, is a growth factor involved in cell proliferation and cell survival signalling (Nickle, Ko and Merrill, 2025), and it is the only gene found to be lower in both SAP-deficient and WAS-deficient conditions. A dedicated, hypothesis-driven follow-up study would be needed to confirm whether it plays a comparable functional role in both disease conditions.

However, we must consider substantial differences between the two studies in cell type, species, stimulation state, and sequencing technology. Therefore, the absence of overlap should not be read as evidence against a shared clinical mechanism, but may reflect different experimental approaches (Ryu *et al.*, 2023; Yu *et al.*, 2024). This comparative analysis has a few limitations that should be mentioned. The two source datasets are not experimentally matched because the SAP dataset was obtained from human T lymphocyte cells, specifically 6 hours after TCR restimulation (Snow *et al.*, 2009), while the WAS dataset was obtained from unstimulated whole mouse spleen (Liang, Yang, *et al.*, 2023). Therefore, this comparison captures cross-species, cross-tissue, and cross-stimulation state overlap rather than a comparison of the same cell type under the same condition. The SAP dataset was generated using an older spotted cDNA microarray technology, whereas the WAS dataset used RNA sequencing, which has substantially greater sensitivity and dynamic range. Nevertheless, the WAS gene list (718 up / 433 down) is much larger relative to the SAP gene list (315 up / 74 down), at least in part be a result of this technological difference and could inflate the apparent asymmetry of the overlap.

Finally, both source studies used rigorous significance limits, and genes falling below each study's threshold are absent from both lists and could not be evaluated for overlap. This means that the true reflection of distributed expression regulation between SAP and WASp pathways may be slightly greater than reported here.

To avoid these problems, the analysis was subsequently extended to the in-house RNA sequencing datasets. These datasets involved human CD8⁺ lymphocytes in which SAP and WAS were silenced under a more closely matched experimental setting than the literature-based comparison. In these experiments, CD8⁺ lymphocytes were silenced for either SAP or WAS and maintained for 4 days before being restimulated with CD3 and CD28 activating antibodies for 4 hours. This experimental design provided a more directly comparable system for investigating the transcriptional consequences of SAP and WAS silencing within the same type of human lymphocyte and following a defined T-cell activation stimulus. The in-house dataset analysis identified 69 genes commonly upregulated by SAP and WAS and 71 genes commonly downregulated by both conditions. In total, 422 genes were upregulated and 776 were downregulated following SAP silencing, while 184 genes were upregulated and 1,085 were downregulated following WAS silencing. Therefore, as expected, the number of genes regulated by both SAP and WAS was higher than that observed in the literature-derived comparison. However, the majority of regulated genes remained specific to either SAP or WAS, indicating that SAP and WASp do not regulate the same gene set. In the upregulated analysis, 353 genes were unique to SAP and 115 were unique to WAS, while in the downregulated analysis, 705 genes were unique to SAP and 1,014 were unique to WAS.

Consequently, the 69 commonly upregulated and 71 commonly downregulated genes may provide stronger evidence for a degree of transcriptional convergence between SAP and WAS than the very small overlap observed across the published datasets. However, the shared genes still represent only a small proportion of the total genes regulated in either condition. Therefore, the in-house results do not support the conclusion that SAP and WAS regulate the same overall transcriptional programme. Instead, these findings suggest that the two proteins may influence a limited common transcriptional response while retaining largely distinct downstream effects.

Another important observation was the difference in the sizes of the SAP and WAS gene sets. In the in-house analysis, SAP silencing was associated with 422 upregulated and 776 downregulated genes, whereas WAS silencing was associated with 184 upregulated and 1,085 downregulated genes. This asymmetry suggests that the two deficiencies may affect the transcriptional state of the cells in different ways.

However, the number of differentially expressed genes (DEGs) alone should not be used to conclude that one protein is more biologically important than the other. The size of a DEG list is influenced by factors such as the magnitude and variability of gene expression changes, experimental thresholds, sequencing depth, biological variation, and the specific molecular consequences of each perturbation. Therefore, differences in gene set size are best interpreted as evidence that SAP and WAS induce distinct transcriptional responses rather than as a direct measure of the relative importance of SAP or WAS.

The comparison of the in-house datasets with the corresponding literature-derived datasets yielded a further important finding regarding reproducibility. No common Ensembl Gene IDs were identified between the in-house SAP-upregulated genes and the published SAP-upregulated genes, and similarly, no common genes were detected between the SAP-downregulated datasets. In contrast, the WAS comparison identified 4 commonly upregulated genes and 13 commonly downregulated genes. These findings indicate that the in-house datasets did not closely reproduce the published SAP transcriptional signature, whereas a small degree of reproducibility was observed for the WAS dataset.

The absence of shared SAP genes between the in-house and literature datasets should not necessarily be interpreted as evidence that the in-house SAP experiment is biologically unrelated to the published SAP study. The published SAP experiment (Snow et al., 2009) and the in-house experiment may differ in several experimental parameters, including cell preparation, stimulation conditions, duration of stimulation, gene expression measurement platform, biological variability, and the criteria used to define differential expression. Even when the same biological protein is studied, relatively small differences in experimental design can alter the set of genes that pass a statistical significance threshold. Furthermore, genes with biologically meaningful but relatively small changes in gene expression may fall below the selected threshold

in one experiment. In such cases, these genes may not appear in the corresponding DEG list.

The comparison between the in-house and literature-derived WAS datasets showed a small but measurable overlap. Four genes were identified as commonly upregulated across both datasets. These are *ATP1B2*, *CCSER1*, *MAFB* and *MT-ND3*. Similarly, 13 genes were identified as commonly downregulated in both datasets. These are *APLN*, *CFAP157*, *CPSF4L*, *DNAH17*, *GAPDHS*, *IFITM5*, *IGFBP6*, *PKDIL3*, *PMEL*, *PRR33*, *TRIM7*, *VIL1* and *WAS*. Among the four commonly upregulated genes identified between the in-house and literature-derived WAS datasets, *ATP1B2* encodes the $\beta 2$ subunit of the Na^+/K^+ -ATPase and helps maintain cellular sodium and potassium gradients and plasma membrane function (Pelin et al., 2020). *MAFB* encodes a basic leucine zipper transcription factor involved in lineage-specific haematopoiesis and gene expression regulation (Saiga et al., 2021). Similarly, *MT-ND3* encodes a subunit of mitochondrial respiratory chain complex I and contributes to mitochondrial electron transport (Nahar et al., 2024). *CCSER1* is associated with cellular regulatory processes. However, its biological function remains less well characterised than those of the other common genes. Therefore, these four upregulated genes represent distinct biological functions, including membrane and ion homeostasis, transcriptional regulation and mitochondrial function, rather than belonging to a single clearly defined biological pathway.

The 13 commonly downregulated genes also form a heterogeneous group with diverse biological functions. *APLN* encodes apelin, a peptide involved in vascular signalling, endothelial function and vascular remodelling (Marsault et al., 2020; Park and Adams, 2026). *IGFBP6* encodes an insulin-like growth factor-binding protein involved in IGF-related signalling and cellular processes such as cell migration (Dailly et al., 2001). *VIL1* encodes villin, an actin-associated protein involved in actin filament organization and cytoskeletal remodelling (Epithelial et al., 2026). This is particularly relevant in the context of *WAS*, as WASp plays an important role in regulating actin cytoskeletal organization and immune cell function (Blundell et al., 2010; Barda-saad, 2013). *PMEL* encodes a transmembrane glycoprotein involved in the structural organization of pre-melanosomes (Ma et al., 2025), whereas *TRIM7* encodes an E3 ubiquitin ligase involved in protein regulation and cellular signalling (Fan et al., 2025). *PKDIL3* encodes a polycystin-family membrane protein that may

contribute to cation-channel activity (Yu et al., 2013), while *PRR33* is a relatively poorly characterized protein-coding gene that is predicted to be associated with cellular responses to wounding (Fu et al., 2024). The remaining genes, including *CFAP157*, *CPSF4L*, *DNAH17*, *GAPDHS* and *IFITM5*, have more specialized or incompletely characterized biological functions. Finally, *WAS* itself encodes the Wiskott-Aldrich syndrome protein (WASp), which plays an important role in the regulation of actin polymerization, immune cell signalling and immune synapse formation (Liang, Yang, et al., 2023). Overall, the functional diversity of these 13 downregulated genes and four upregulated genes does not indicate a single common biological pathway.

The presence of these genes in both experimental settings suggests that a small subset of the transcriptional response associated with *WAS* deficiency may be relatively reproducible across datasets. However, because most genes were not shared between the two datasets, these common genes should be considered candidate reproducible markers rather than evidence of complete replication of the published WAS transcriptional programme.

The interpretation of the overlap between the WAS in-house and literature datasets also requires consideration of the species conversion performed during the analysis. The published WAS dataset originated from mouse spleen, whereas the laboratory WAS dataset was generated from human CD8⁺ lymphocytes. The published mouse genes were therefore mapped to their corresponding human orthologues before comparison with the human datasets. This enabled comparison using a common human Ensembl Gene ID framework rather than comparing mouse and human gene identifiers directly. This step was essential because species-specific gene identifiers cannot be treated as directly comparable. Nevertheless, orthology mapping cannot eliminate the biological differences between mouse spleen and human CD8⁺ lymphocytes. Therefore, the remaining overlap should be interpreted cautiously, with these limitations in mind.

The in-house and literature dataset comparison further suggests that transcriptomic findings may depend on the original biological material and experimental setup. The absence of overlap for SAP and the limited overlap for WAS highlight the importance of experimental matching when using published gene lists to

infer shared molecular mechanisms. These results also underscore the need for standardised gene identifiers and orthology mapping for cross-study comparisons.

The 13 commonly downregulated WAS genes encompass a range of biological functions, indicating that the shared transcriptional response is not confined to a single molecular pathway. The presence of genes with diverse functional roles suggests that WASp deficiency may affect multiple cellular processes simultaneously. However, the present Venn analysis alone cannot determine whether these genes are functionally connected or whether they participate in a single common pathway. Therefore, additional functional enrichment and pathway analyses are required to determine whether these shared genes form biologically meaningful networks.

The combined analysis provides a clear understanding of the relationship between SAP- and WASp-regulated transcription. The literature datasets indicate very limited gene overlap under significantly different experimental conditions, whereas the in-house human CD8⁺ lymphocyte datasets identify a relatively large but still minority shared transcriptional response. Therefore, these findings do not support the presence of a single large common SAP and WAS gene programme.

Further analysis of the shared and unique gene sets using Gene Ontology (GO), Kyoto Encyclopedia of Genes and Genomes (KEGG), Reactome, or other pathway enrichment methods would be highly valuable. In addition, network analysis could help identify potential functional relationships among the shared genes and determine whether they are associated with T-cell activation, apoptosis, cytoskeletal regulation, inflammatory signalling, metabolism, or other cellular processes. These approaches would provide a pathway-level complement to the present gene-level Venn analysis and help determine whether the observed gene overlap represents a biologically meaningful common response.

6. CONCLUSION

We can obtain two main conclusions from this work:

- 1) There is no evidence of a shared genetic programme controlled by SAP and WASp, as only a small proportion of genes were commonly regulated by both SAP and WASp in both the literature-derived and in-house datasets.
- 2) We have evidence that the gene lists are strongly dependent on the experimental settings, as substantial differences were observed between datasets obtained using different experimental conditions, including cell and tissue type, species, stimulation conditions, and experimental approaches.

7. BIBLIOGRAPHY

Antigen, C. and Cells, R. T. (2023) ‘Transcriptomic Approaches in Studies on and Applications of Chimeric Antigen Receptor T Cells’, pp. 1–15.

Baldanzi, G. *et al.* (2011) ‘SAP-mediated inhibition of diacylglycerol kinase α regulates TCR-induced diacylglycerol signaling’, 187(11), pp. 5941–5951. doi: 10.4049/jimmunol.1002476.SAP-MEDIATED.

Barda-saad, M. (2013) ‘Wiskott – Aldrich syndrome protein – dynamic regulation of actin homeostasis : from activation through function and signal termination in T lymphocytes’, pp. 10–29.

Blundell, M. P. *et al.* (2010) ‘The Wiskott-Aldrich syndrome : The actin cytoskeleton and immune cell function’, 29, pp. 157–175. doi: 10.3233/DMA-2010-0735.

Booth, C. *et al.* (2011) ‘X-linked lymphoproliferative disease due to SAP / SH2D1A deficiency : a multicenter study on the manifestations , management and outcome of the disease’, 117(1), pp. 53–62. doi: 10.1182/blood-2010-06-284935.The.

Buchbinder, D., Nugent, D. J. and Fillipovich, A. H. (2014) ‘Wiskott – Aldrich syndrome : diagnosis , current management , and emerging treatments’, pp. 55–66.

Calixto, R. *et al.* (2023) ‘Understanding immunoactinopathies : A decade of research on WAS gene defects’, (March), pp. 1–19. doi: 10.1111/pai.13951.

Cruz, M., Davidson, D. and Roncagalli, R. (2006) ‘NK cell regulation by SLAM family receptors and SAP-related adapters’, 214, pp. 22–34.

Dailly, Y. P. *et al.* (2001) ‘Structure and characterization of the human insulin-like growth factor binding protein (IGFBP) -6 promoter : identi cation of a functional retinoid response element 1’, 1518, pp. 145–151.

Dragovich, M. and Mor, A. (2019) ‘inflammatory and autoimmune diseases’, 17(7), pp. 674–682. doi: 10.1016/j.autrev.2018.01.018.The.

Du, H. *et al.* (2015) ‘Effects of Wiskott – Aldrich Syndrome Protein Deficiency on IL-10-Producing Regulatory B Cells in Humans and Mice’, pp. 483–493. doi: 10.1111/sji.12282.

Epithelial, T. *et al.* (2026) ‘The Epithelial Cell Cytoskeleton and Intracellular Trafficking III. How is villin involved in the actin cytoskeleton dynamics in intestinal cells?’, pp. 496–502.

Fan, J. *et al.* (2025) ‘A brief overview of the E3 ubiquitin ligase : TRIM7’, *Cellular Signalling*, 134(April), p. 111886. doi: 10.1016/j.cellsig.2025.111886.

Fu, X. *et al.* (2024) ‘Adapting cytoskeleton-mitochondria patterning with myocyte differentiation by promyogenic PRR33’, *Cell Death & Differentiation*, (August 2023). doi: 10.1038/s41418-024-01363-w.

Hasumi, H. *et al.* (2015) ‘syndrome : Clinical and molecular aspects of Birt – Hogg – Dub e recently identified kidney cancer syndrome’, pp. 11–13. doi: 10.1111/iju.13015.

Hiensch, F., Dupré, L. and Salzer, E. (2025) ‘Review Article Immunoactinopathies revisited : understanding clinical manifestations and biological pathways’, *BLOOD*, 145(23), pp. 2709–2732. doi: 10.1182/blood.2024026763.

Hopkins, P. C. R. (2013) ‘Neurodegeneration in a Drosophila Model for the Function of TMCC2 , an Amyloid Protein Precursor- Interacting and Apolipoprotein E-Binding Protein’, 8(2). doi: 10.1371/journal.pone.0055810.

Knop, K. *et al.* (2024) ‘RAM is upregulated during T cell activation and is required for RNA cap formation and gene expression’, (September 2023), pp. 1–11.

Kumari, S. *et al.* (2015) ‘Actin foci facilitate activation of the phospholipase C- γ in primary T lymphocytes via the WASP pathway’, pp. 1–31. doi: 10.7554/eLife.04953.

Li, C. *et al.* (2003) ‘Dual Functional Roles for the X-linked Lymphoproliferative Syndrome Gene Product SAP / SH2D1A in Signaling through the Signaling Lymphocyte Activation Molecule (SLAM) Family of Immune Receptors *’, *Journal of Biological Chemistry*, 278(6), pp. 3852–3859. doi: 10.1074/jbc.M206649200.

Liang, F., Peng, C., *et al.* (2023) ‘A single-cell atlas of immunocytes in the spleen of a mouse model of Wiskott-Aldrich syndrome’, *Cellular Immunology*, 393–394(November), p. 104783. doi: 10.1016/j.cellimm.2023.104783.

Liang, F., Yang, J. U. N., *et al.* (2023) ‘Transcriptomic insights into the role of the

spleen in a mouse model of Wiskott - Aldrich syndrome', pp. 1–9. doi: 10.3892/etm.2022.11763.

Lim, G. *et al.* (2010) 'Detection of t (3 ; 5) and NPM1 / MLF1 rearrangement in an elderly patient with acute myeloid leukemia : clinical and laboratory study with review of the literature', *Cancer Genetics and Cytogenetics*, 199(2), pp. 101–109. doi: 10.1016/j.cancergencyto.2010.02.009.

Liu, W. *et al.* (2020) 'Tim-4 in Health and Disease : Friend or Foe ?', 11(April), pp. 1–10. doi: 10.3389/fimmu.2020.00537.

Ma, B. *et al.* (2025) 'Atomic structure and in situ visualization of native PMEL lamellae in melanosomes', (April).

Marsault, E. *et al.* (2020) 'The apelinergic system: a perspective on challenges and opportunities in cardiovascular and metabolic disorders.', 1455(1), pp. 12–33. doi: 10.1111/nyas.14123.The.

Nahar, N. *et al.* (2024) 'Mitochondrion Identification of a novel MT-ND3 variant and restoring mitochondrial function by allotopic expression of MT-ND3 gene', *Mitochondrion*, 76(March), p. 101858. doi: 10.1016/j.mito.2024.101858.

Nickle, A., Ko, S. and Merrill, A. E. (2025) 'Fibroblast Growth Factor 2', pp. 1–20. doi: 10.1016/j.diff.2023.10.001.Fibroblast.

Panchal, N., Houghton, C. and Vassalou, E. (2022) 'Allosteric inhibition of SHP2 rescues functional T-cell abnormalities in SAP deficiency', *Journal of Allergy and Clinical Immunology*, 150(6), pp. 1507-1516.e7. doi: 10.1016/j.jaci.2022.06.021.

Park, H. and Adams, R. H. (2026) 'Apelin – APLNR pathway across development , in fl ammation , and vascular remodeling : an endothelial perspective', *Experimental & Molecular Medicine*, (April). doi: 10.1038/s12276-026-01760-w.

Parolini, B. S. *et al.* (2000) 'X-linked Lymphoproliferative Disease : 2B4 Molecules Displaying Inhibitory Rather Than Activating Function Are Responsible for the Inability of Natural Killer Cells to Kill Epstein-Barr Virus – infected Cells', 192(3).

Pelin, M. *et al.* (2020) 'In Vitro Cell Sensitivity to Palytoxin Correlates with High Gene Expression of the Na + / K + -ATPase β 2 Subunit Isoform'.

Pradeep, R. *et al.* (2024) 'Wiskott Aldrich syndrome protein (WASp) -deficient Th1 cells promote R-loop-driven transcriptional insufficiency and transcription-coupled nucleotide excision repair factor (TC-NER) -driven genome-instability in the pathogenesis of T cell acute lymphoblastic leukemia', 263(January).

Puthenveedu, M. A. *et al.* (2003) 'Basolateral Cycling Mediated by a Luminal Domain Targeting Determinant', 14(May), pp. 1801–1807. doi: 10.1091/mbc.E02.

Rattanasri, A. *et al.* (2026) 'Regulation of Proximal TCR Signalling Requires Wiskott – Aldrich Syndrome Protein', pp. 704–716. doi: 10.1111/imm.70149.

Rawlings, B. S. L. *et al.* (1999) 'Spontaneous Apoptosis in Lymphocytes From Patients With Wiskott-Aldrich Syndrome : Correlation of Accelerated Cell Death and Attenuated Bcl-2 Expression', *BLOOD*, 94(11), pp. 3872–3882. doi: 10.1182/blood.V94.11.3872.

Ryu, Y. *et al.* (2023) 'Molecules and Cells Integration of Single-Cell RNA-Seq Datasets : A Review of Computational Methods', *Molecules and Cells*, 46(2), pp. 106–119. doi: 10.14348/molcells.2023.0009.

Saiga, H. *et al.* (2021) 'Transcription factor MafB-mediated inhibition of type I interferons in plasmacytoid dendritic cells', 34(3), pp. 159–172.

Sasson, I. *et al.* (2026) 'Regulation and dysregulation at a critical hematopoietic junction : The human WASP-WIP complex', *Journal of Structural Biology*, 218(3), p. 108354. doi: 10.1016/j.jsb.2026.108354.

Satchwell, T. J. *et al.* (2009) 'Blood Cells , Molecules , and Diseases Protein 4 . 2 : A complex linker', *Blood Cells, Molecules, and Diseases*, 42(3), pp. 201–210. doi: 10.1016/j.bcmd.2009.01.005.

Shah, K. *et al.* (2021) 'T cell receptor (TCR) signaling in health and disease', (February). doi: 10.1038/s41392-021-00823-w.

Smith-garvin, J. E., Koretzky, G. A. and Jordan, M. S. (2010) 'T Cell Activation', pp. 591–619. doi: 10.1146/annurev.immunol.021908.132706.T.

Snow, A. L. *et al.* (2009) 'patients with X-linked lymphoproliferative disease caused by Restimulation-induced apoptosis of T cells is impaired in patients with X-linked

lymphoproliferative disease caused by SAP deficiency'. doi: 10.1172/JCI39518.2976.

Sun, X. *et al.* (2019) 'The role of WASp in T cells and B cells', *Cellular Immunology*, 341(January), p. 103919. doi: 10.1016/j.cellimm.2019.04.007.

Tangye, S. G. *et al.* (2000) 'Cutting Edge: Functional Requirement for SAP in 2B4-Mediated Activation of Human Natural Killer Cells as Revealed by the X-Linked Lymphoproliferative Syndrome1'.

Vallée, T. C., Albert, M. H. and Pai, S. (2025) 'How I Treat How I treat Wiskott-Aldrich syndrome', *BLOOD*, 146(1), pp. 41–51. doi: 10.1182/blood.2024026288.

Velnati, S. *et al.* (2023) 'Wiskott-Aldrich syndrome protein interacts and inhibits diacylglycerol kinase alpha promoting IL-2 induction', (April), pp. 1–21. doi: 10.3389/fimmu.2023.1043603.

Warren, W. G. *et al.* (2023) 'The emerging role of fatty acid binding protein 5 (FABP5) in cancers', *Drug Discovery Today*, 28(7), p. 103628. doi: 10.1016/j.drudis.2023.103628.

Yang, J. *et al.* (2023) 'Advancing CAR T cell therapy through the use of multidimensional omics data', 20(4), pp. 211–228. doi: 10.1038/s41571-023-00729-2.Advancing.

Yu, Y. *et al.* (2013) 'Molecular mechanism of the assembly of an acid-sensing receptor ion channel complex'. doi: 10.1038/ncomms2257.Molecular.

Yu, Y. *et al.* (2024) 'Assessing and mitigating batch effects in large - scale omics studies', *Genome Biology*. doi: 10.1186/s13059-024-03401-9.