CellPress

Sneak Peek

A PREVIEW OF PAPERS UNDER REVIEW

JOURNAL

Cell Stem Cell

ARTICLE TITLE

Wiskott-Aldrich Syndrome Protein Regulates Nuclear Organization, Alternative Splicing and Cell Proliferation

AUTHOR LIST

Mo Li, Keiichiro Suzuki, Xuan Zhou, Takayoshi Yamauchi, James J. Moresco, Sarah Dunn, Reyna Hernandez-Benitez, Tomoaki Hishida, Na Young Kim, Manal M. Andijani, Manching Ku, Yuta Takahashi, Jinna Xu, Jinsong Qiu, Ling Huang, Christopher Benner, Emi Aizawa, Jing Qu, Guang-Hui Liu, Zhongwei Li, Fei Yi, Chongwei Bi, Yanal Ghosheh, Changwei Shao, Maxim Shokhirev, Patrizia Comoli, Francesco Frassoni, John R. Yates III, Xiang-Dong Fu, Concepcion Rodriguez Esteban, Juan Carlos Izpisua Belmonte

This is a paper under consideration at Cell Press and has not been peer-reviewed

Wiskott-Aldrich Syndrome Protein Regulates Nuclear Organization, Alternative Splicing and Cell Proliferation

Authors: Mo Li^{1, 2}, Keiichiro Suzuki¹, Xuan Zhou², Takayoshi Yamauchi¹, James J. Moresco³, Sarah Dunn⁴, Reyna Hernandez-Benitez¹, Tomoaki Hishida¹, Na Young Kim¹, Manal M. Andijani², Manching Ku⁵, Yuta Takahashi^{1, 6}, Jinna Xu^{1, 2}, Jinsong Qiu⁷, Ling Huang⁸, Christopher Benner⁸, Emi Aizawa¹, Jing Qu^{1, *}, Guang-Hui Liu^{1, *}, Zhongwei Li¹, Fei Yi¹, Chongwei Bi², Yanal Ghosheh², Changwei Shao⁷, Maxim Shokhirev⁸, Patrizia Comoli⁹, Francesco Frassoni¹⁰, John R. Yates III³, Xiang-Dong Fu⁷, Concepcion Rodriguez Esteban¹ and Juan Carlos Izpisua Belmonte^{1§}

Affiliations:

1 Gene Expression Laboratory, Salk Institute for Biological Studies, 10010 North Torrey Pines Road, La Jolla, California 92037, USA.

2 4700 King Abdullah University of Science and Technology (KAUST), Thuwal 23955-6900, Kingdom of Saudi Arabia

3 Department of Cell Biology, Scripps Research Institute, La Jolla, California 92037, USA.

4 The Waitt Advanced Biophotonics Core Facility, Salk Institute for Biological Studies, 10010 North Torrey Pines Road, La Jolla, California 92037, USA.

5 Next-generation sequencing core, Salk Institute for Biological Studies, 10010 North Torrey Pines Road, La Jolla, California 92037, USA

6 Life Science Center, Tsukuba Advanced Research Alliance, University of Tsukuba, 1-1-1 Tennoudai, Tsukuba, Ibaraki 305-8577, Japan

7 Department of Cellular & Molecular Medicine, University of California at San Diego, La Jolla, California 92093, USA.

8 Integrative Genomics and Bioinformatics Core, Salk Institute for Biological Studies, 10010 North Torrey Pines Road, La Jolla, California 92037, USA

9 Pediatric Hematology/Oncology and Cell Factory, Fondazione IRCCS Policlinico San Matteo, Pavia, Italy

10 Department of Research Laboratories and Director of Center for Stem Cell and Cell Therapy, Istituto G. Gaslini Children Hospital Scientific Institute, 16147, Genova Italy

* Present address: National Laboratory of Biomacromolecules, Institute of Biophysics, Chinese Academy of Sciences, Beijing 100101, China.

§Correspondence: belmonte@salk.edu, (JCIB)

Abstract

Wiskott-Aldrich syndrome (WAS), caused by mutations in the WASP protein, displays immunological dysfunctions and predisposition to cancer. Despite studies in cell lines and mouse models the molecular mechanisms of WAS remain obscure. We generated induced pluripotent stem cells (iPSCs) from patients with WAS (WAS-iPSCs) and isogenic gene-corrected iPSCs by genome editing. Immune cells derived from WAS-iPSCs, genetically engineered B lymphoblastoid cell lines, and patient primary lymphocytes were subjected to imaging, proteomic and transcriptomic analyses. The WAS-iPSC model not only recapitulated known disease phenotypes but also revealed, for the first time, roles of WASP in the organization of the nucleolus and nuclear speckles and PML nuclear bodies. WASP interacts with components of the nucleolus and nuclear speckles, including chromatin modifiers and splicing factors. Innate and adaptive immune cells from WAS patients display global dysregulation of cell cycle regulation and alternative splicing. WASP mutation is sufficient to drive an accelerated cell cycle and tumor-promoting splicing changes. Our data show that WASP acts as a tumor suppressor and specific WASP mutants behave as oncogenes and cause cell-intrinsic alterations that predispose patients to cancer.

INTRODUCTION

Wiskott-Aldrich syndrome (WAS), an X-linked primary immunodeficiency, results in microthrombocytopenia, eczema, recurrent infections, autoimmunity and predisposition to malignancy (Candotti, 2018). WAS is caused by mutations in the WASP protein, encoded by the *WAS* gene. WASP is best known as an actin nucleation factor (Thrasher and Burns, 2010). WASP also functions in transcriptional regulation via chromatin-mediated mechanisms (Sadhukhan et al., 2014; Sarkar et al., 2014; Taylor et al., 2010). A family member of WASP regulates nuclear organization, especially in the heterochromatin compartment in *Drosophila* (Verboon et al., 2015). However, the pathogenic mechanisms of *WAS* mutations remain obscure. Particularly puzzling is the predisposition of WAS patients to early onset malignancy (Candotti, 2018; Thrasher and Burns, 2010). Research in WAS faces several challenges. Murine models of WAS do not fully recapitulate the disease (Snapper et al., 1998). Primary patient cells are difficult to obtain, owing to the rarity of WAS and the risk in obtaining samples from children who has thrombocytopenia and immunodeficiency. Cell lines can be used to study WAS pathogenesis, but the findings in transformed cell lines must be verified in a normal cellular context.

Mutations in the *WAS* gene affect all hematopoietic lineages except for erythrocytes (Massaad et al., 2013). Patient-derived iPSCs can differentiate into any cell type, therefore constituting an attractive way to model diseases (Inoue et al., 2014; Takahashi et al., 2007; Takahashi and Yamanaka, 2006). Two studies reported WAS disease

modeling using patient-specific iPSCs (Ingrungruenglert et al., 2015; Laskowski et al., 2016). However, these models were not isogenic, nor did they provide new mechanistic insights into the disease. Here we generated isogenic WAS iPSC models, which were combined with efficient hematopoietic differentiation to reveal novel roles of WASP in the organization of nuclear bodies, chromatin structure, gene expression, and alternative splicing (Fig. 1A). Pathogenic mutations of WASP disrupt these processes, cause cell cycle deregulation and promote carcinogenesis.

Generation of isogenic WAS-iPSC model

To avoid potential issues associated with integrated transgenes, we derived integration-free iPSCs from fibroblasts of two patients with WAS using five reprogramming factors (OCT4, SOX2, KLF4, LIN28, and MYCL1) and p53-shRNA encoded in episomal vectors (Fig. S1A) (Li et al., 2011; Okita et al., 2011). Patient 1 had a clinical score 4 (the 1–5 scale increases with severity (Zhang et al., 2010; Zhu et al., 1995)), and a nonsense mutation (c.107_108del, p.F36*, truncation after the PH1 domain). Patient 2 (clinical score unknown) carried a previously unreported mutation (c.1271dupG, p.L425fs, deletion of the c-terminal VCA domain). Fibroblasts from an unaffected family member of Patient 1 were reprogrammed as a control. The mutations did not affect reprogramming efficiency (Patient 1, 0.097%, Patient 2, 0.08%, and normal range, 0.023%~0.2%(Li et al., 2011; Liu et al., 2014)). No ectopic reprogramming factor or episomal vector sequence was detected in the genome of four randomly selected iPSC clones (Fig. S1B&C). The established WAS-iPSC lines expressed markers of

pluripotency, differentiated into all three germ layers (Fig. S1D and S2A), retained the original *WAS* mutation (Fig. S1H), and demonstrated a normal karyotype (Fig. S1I).

To generate an isogenic wild-type control for WAS-iPSC, we performed targeted gene correction by homologous recombination using a helper-dependent adenoviral vector (HDAV)–*WAS*-c-HDAV, which mediates efficient gene targeting with minimal impact on genomic integrity (Liu et al., 2011; Mitani et al., 1995; Suzuki et al., 2008; Suzuki et al., 2014). Gene correction with *WAS*-c-HDAV was confirmed by PCR and sequencing (Fig. S1E&F). The efficiency of gene correction was 5–50% and correlated with the distance between the drug selection marker and the mutation (Fig. S1G). The integrated neomycin-resistance cassette was then excised using the FLP/*FRT* system (data not shown). The corrected WAS-iPSCs (named cWAS-iPSC) had a normal karyotype, retained pluripotency, and were free of random vector integration (Fig. S1I-K).

WAS-iPSC models recapitulate WAS phenotypes

Differentiation of human iPSCs can recapitulate developmental hematopoiesis and produce hematopoietic progenitor and mature cell types useful for investigating WAS pathogenesis (Liu et al., 2014) (Lensch and Daley, 2004). WAS-iPSCs expressed markers of the three germ layers upon spontaneous differentiation (Fig. S2A). The expression of the *WAS* gene was greatly reduced in both patient WAS-iPSCs, which was rescued by gene correction (Fig. S2B&C). Hematopoietic differentiation of WAS-iPSCs showed similar kinetics in the upregulation of genes important for hematopoietic progenitor cells (HPCs) (Fig. S2D). WAS-iPSC derived HPCs gave rise to typical

hematopoietic colonies (Fig. S2F). CD43 expression is defective in some patients (Greer et al., 1989; Piller et al., 1991; Remold-O'Donnell et al., 1984; Siminovitch et al., 1993), which was confirmed at the protein and mRNA level (Fig. S2E, G, H and J). Importantly, the deficit of CD34⁺/CD43⁺ cells was rescued by gene correction (Fig. S2E).

We differentiated WAS-iPSCs into hematopoietic progenitor cells, macrophages and dendritic cells (DCs) to examine if they could recapitulate disease phenotypes observed in patients (Massaad et al., 2013). WAS-iPSC derived macrophages (WAS-iMPs) displayed morphological abnormalities that indicate defects in substrate interaction and spreading, as seen in patient macrophages and DCs (Fig. 1B) (Binks et al., 1998; Linder et al., 1999). WAS-iPSC derived DCs (WAS-iDCs) and WAS-iMPs expressed typical lineage markers, as seen in the clinic (Fig. S2I&K) (Binks et al., 1998; Massaad et al., 2013). Podosome defects of WAS (Linder et al., 1999) were recapitulated in 100% of WAS-iMPs (Fig. 1C). Additionally, phagocytosis and chemotaxis defects (Lorenzi et al., 2000; Massaad et al., 2013) were recapitulated in WAS-iMPs and WAS-iDCs (Fig. 1D and S2L). Importantly, all of these phenotypes were corrected in cWAS-iPSCs-derived cells (Fig. 1B-D and S2). Together, these data show that isogenic WAS-iPSC models faithfully recapitulate known WAS phenotypes and are suitable for studying pathogenic mechanisms of WAS.

WASP regulates nuclear compartments

WASP is expressed by all blood lineages except for erythrocytes (Fig. S3A). We examined WASP localization in major blood lineages and found it to localize to both the

cytoplasm and the nucleus (Fig. S3B&C). Although nuclear roles of WASP in transcription have been described in T_H cells (Sadhukhan et al., 2014; Sarkar et al., 2014; Taylor et al., 2010), the nuclear localization and function of WASP in monocytes/macrophages have not been reported. WASP puncta occupied nucleoplasmic compartments that are associated with active transcription and splicing (Fig. S3B&C) (Bickmore, 2013; Rajapakse and Groudine, 2011; Sutherland and Bickmore, 2009). WASP is known to co-localize with hyperphosphorylated RNA polymerase II (Taylor et al., 2010) at sites of active transcription named “transcription factories” (Bickmore, 2013). However, whether WASP participates in other nuclear compartments remains unknown.

We therefore examined the nuclear localization of WASP in relation to various nuclear bodies. In contrast to T_H1 cells, where WASP is reported to be extranucleolar (Taylor et al., 2010), WASP localized in the nucleolus in macrophages (Fig. 2A). The nucleolar localization was verified at the ultrastructural level (Fig. 2B, S3D&E). WASP puncta partially overlapped with the paraspeckles (Fig. S4B) and were frequently found near but rarely within nuclear speckles and promyelocytic leukemia nuclear bodies (PML-NBs, Fig. 2A).

The nucleolus marked by NMP1 was significantly dimmer and larger in WAS-iMPs than the wild-type iMPs (Fig. 2A&C). Overexpression of the WAS^{F36*} mutant in WAS^{null} B cells led to similar results (Fig. S4C). Transmission electron microscopy (TEM) confirmed the nucleolar hypertrophy in WAS-iMPs and mutant B cells (Fig. 2D-F and

S4D-F). The nucleolar phenotype was rescued by gene correction (Fig. 2E&F and S4D-F). The nucleolus is the site of ribosome biogenesis and a key regulator of cell-cycle progression, chromatin structure, and genome stability (Orsolic et al., 2016). Enlarged nucleoli correlate with cancer initiation and progression and are widely used as a cancer diagnostic marker (O'Day, 2013). Our data thus provide the first link between WASP deficiency and nucleolar dysfunction.

The PML-NBs in WAS-iMP1 were more prominent and numerous than those in normal cells (Fig. 2A&C) (Lallemand-Breitenbach and de The, 2010). A similar increase in PML-NBs was evident in WAS^{F36*}-expressing cells (Fig. S4C). This phenotype was rescued in the isogenic cWAS-iMP1 (Fig. 2C). The morphology and distribution of PML-NBs can change dramatically under various cellular stresses. P53 and interferons—both transcriptional regulators of PML—induce an increase in the number and size of PML-NBs (de Stanchina et al., 2004; Stadler et al., 1995). However, no obvious change in expression levels of p53 or interferons were detected (data not shown). These data suggest that WASP is a novel regulator of PML-NB homeostasis.

Compared to wild-type cells, WAS-iMPs contained larger nuclear speckles with fuzzy edges (Fig. 2A&C). The nuclear speckle marker SRSF2 ectopically aggregated in the cytoplasm of WAS-iMPs (Fig. 2C). Expression of WAS^{F36*} in B cells similarly led to more prominent and irregularly-shaped nuclear speckles (Fig. S4C). Nuclear speckles can be observed as interchromatin granule clusters (IGCs) by negative staining in TEM analysis (Spector and Lamond, 2011). Mutant cells contained numerous prominent

IGCs, whereas IGCs in wild-type cells were sparse and inconspicuous, thus providing an independent confirmation of the immunofluorescence data (Fig. 2D, S4D-F). The nuclear speckle phenotype was rescued by gene correction (Fig. 2E, S4D-F). Nuclear speckles are dynamic structures enriched in splicing factors and involved in pre-mRNA splicing, chromatin structure, transcription initiation and elongation, and RNA modification (Galganski et al., 2017; Spector and Lamond, 2011). Their morphology can be disturbed by carcinogenesis, viral infection, and phosphorylation of serine/arginine-rich splicing factors (Galganski et al., 2017; Spector and Lamond, 2011). For example, inhibition of serine/threonine protein phosphatase 1 is thought to prevent the release of phosphorylated SRSF2 from assembled spliceosomes located in perichromatin fibrils on the periphery of nuclear speckles, thus giving the speckles an enlarged and irregular appearance (Misteli and Spector, 1996; Spector and Lamond, 2011). This is similar to the phenotype of WASP deficient cells (Fig. 2C&D and S4C-F). Thus, these data suggest that WASP regulates the dynamic organization of nuclear speckles.

The fact that the anomalies in nuclear bodies were rectified in isogenic gene-corrected cells suggests that WASP is involved in the organization of multiple nuclear compartments and may regulate transcription, RNA metabolism, and chromatin structure in immune cells.

Characterization of the WASP interactome

Few components of nuclear bodies have been reported to interact with WASP (Sadhukhan et al., 2014). To understand how WASP could regulate nuclear bodies we

performed MudPIT (multi-dimensional protein identification technology) analysis using co-immunoprecipitation (co-IP) of endogenous WASP to achieve specific and sensitive detection of WASP interaction partners (see methods). MudPIT analysis in whole cells and purified nuclei (Fig. S5A) identified 78 and 299 high-confidence WASP-interacting polypeptides, respectively (Fig. 3A). Known partners of WASP, including WIPF1 (Ramesh et al., 1997), RAN (Sadhukhan et al., 2014), KPNB1 (Sadhukhan et al., 2014), and VIM (Lee et al., 2004) were identified (Table S1). Novel partners included markers of the nuclear lamina (LMNB1, Fig. 3C), nucleolus (FBL, EIF6, NPM1, and NCL), nuclear speckle (SRSF2, SF3B3, and CDC5L (Grillari et al., 2005)), paraspeckle (NONO, (Fox and Lamond, 2010) SFPQ, (Fox and Lamond, 2010) and RBM14 (Fox and Lamond, 2010)), and PML-NB (PML). Analysis of proteomic datasets (Dellaire and Bazett-Jones, 2004; Galganski et al., 2017; Mohamad and Boden, 2010) showed that the WASP interactome overlapped extensively with components of the nucleolus and nuclear speckle, but minimally with the PML-NB (Fig. 3A, Table S1). The most enriched gene ontology (GO) term was poly(A) RNA binding (Fig. 3B). WASP-interacting proteins were highly enriched in RNA-binding proteins (Fig. S5B) (Castello et al., 2013). Other enriched GO terms included translational initiation, mRNA splicing, nucleolus and nuclear speck (Fig. 3B). Several binding partners of WASP were confirmed using Rapid Immunoprecipitation Mass spectrometry of Endogenous protein (RIME (Mohammed et al., 2016)) in whole cells (RIME is not suitable for nuclear IP) (Table S1, Fig. S5C).

The association of WASP with nucleolar components corroborated our immunofluorescence findings (Fig. 2A&B). Although WASP has never been reported as

a nucleolar protein, another member of the WASP family—WASL—is present in the nucleolus (Mohamad and Boden, 2010). The association between WASP and nuclear speckle components was unexpected, as immunofluorescence foci of WASP and SRSF2 did not have extensively overlap (Fig. 2A). We confirmed the physical interaction between WASP and several splicing factors that are *bona fide* nuclear speckle proteins using co-immunoprecipitation and proximity ligation assay (Fig. 3C&D). Nuclear speckles are sites of storage, assembly, and modification of splicing factors. These splicing factors shuttle between the speckles and nucleoplasm dynamically and transiently interact with nucleoplasmic factors (Spector and Lamond, 2011). Nuclear speckle reorganization upon RNA Pol II inhibition is dependent on actin polymerization (Wang et al., 2006). Therefore, our data suggest that WASP interacts with speckle-associated splicing factors, and that WASP mutations disrupt the organization of nuclear speckles, which could have pathogenic consequences in RNA splicing and transcription.

Epigenetic and transcriptional defects in WAS patients

The nucleolus and nuclear speckle play key roles in chromatin organization and transcription-linked biological processes (Galganski et al., 2017; O'Day, 2013). We therefore examined chromatin structure and gene expression in WASP mutant cells.

The nucleolus organizes a shell of highly condensed and late-replicating heterochromatic DNA around itself (Guetg and Santoro, 2012; O'Day, 2013). Such perinucleolar heterochromatin was abolished in WAS-iMPs (Fig. 3E). Similarly, super-

resolution microscopy showed that the predominant localization of the constitutive heterochromatin around the nucleolus and at the nuclear periphery is lost in WAS-iMPs (Fig. 3F), which is even more pronounced in WAS^{F36*}-expressing cells (Fig. S4C). Furthermore, the contrast of electron density between heterochromatic and euchromatic regions of TEM images was significantly lower in mutant cells. Heterochromatin in WAS mutant cells appeared less condensed and more dispersed (Fig. 2D&E, 3E&F and S4C-F). Together these data suggest that WAS mutations disrupt heterochromatin organization.

We hypothesized that the aberrances in nuclear bodies and chromatin structure in WAS cells could be associated with changes in gene expression. We first performed RNA sequencing (RNA-seq) of WAS-iMPs, cWAS-iMPs, and macrophages derived from normal cord blood (CB-MPs). WAS-iMP1 showed a significant number of differentially expressed genes when compared with CB-MP, while gene correction restored the concordance of gene expression between them (Fig. 4A). There were 652 genes upregulated and 480 downregulated (FDR<0.05). These genes were enriched in chromatin-related processes, extracellular matrix and adhesion, immune function, cell cycle and cancer (Fig. 4B). Similar GO enrichment was seen both patients (Fig. 4C).

Overproduction of pro-inflammatory cytokines has been reported in WAS T_H1 cells and monocytes, and the latter has been suggested to contribute to regulatory T cell dysfunction in autoimmune hepatitis (Sarkar et al., 2015) (Arterbery et al., 2016). WAS-iMPs produced significantly higher levels pro-inflammatory cytokines at both the mRNA and protein level, a phenotype that was corrected in cWAS-iMPs (Fig. 4C&D and S6A).

Thus, our data suggest that overproduction of pro-inflammatory cytokines by WAS patient macrophages also contributes to immune dysfunction and autoimmunity. While misregulation of genes related to cytoskeleton and immune function was unsurprising, given the known function of WASP in these processes (Thrasher and Burns, 2010), the enrichment of differentially expressed genes in chromatin-related processes, cell cycle, and transcriptional misregulation in cancer suggests unappreciated pathogenic mechanisms of WAS mutations.

Unsupervised hierarchical clustering showed that iMPs of both patients clustered together, while their gene-corrected counterpart clustered with CB-MPs (Fig. 4C&E). Two patient WAS-iMPs differed from CB-MPs in distinctive ways in principal component analysis (PCA), yet both resembled normal macrophages after gene correction (Fig. 4F). The two patients shared a common gene signature enriched in GO terms including nucleus, cell cycle, response to virus, and chromosome organization (Fig. S6B). Additionally, p53 pathway was inhibited in both patients (Ingenuity pathway analysis, $p < 2.26 \times 10^{-56}$). Both patient cells activated cancer-associated processes and suppressed processes related to cell death (Fig. S6B&C).

Many nuclear speckle-associated splicing factors were overexpressed in both patient iMPs (Fig. 4G and S6A). We examined alternative splicing by comparing gene expression at the isoform level (see methods). We identified 245 downregulated and 122 upregulated isoforms in WAS-iMPs. Most of these belonged to the differentially expressed gene list. These 367 isoforms were enriched in categories including protein

modification and turnover, alternative splicing, RNA binding and cell cycle (Fig. 4H). We obtained similar results using an alternative method, RASL-seq (Zhou et al., 2012) (Fig. S6D, Table S2). The alternatively spliced genes therein were also enriched in biological processes including acetylation, alternative splicing, protein turnover and cell cycle. Together with the proteomics data, the results suggest a novel role of WASP in RNA splicing through its physical interaction with nuclear factors of the splicing machinery.

WAS patients display dysfunction, clonal proliferation, and malignant transformation of adaptive immune lineages (Candotti, 2018), we therefore analyzed gene expression and splicing in primary patient T and B cells using RNAseq (Fig. 5A). PCA showed that cell type explained 43.6% of the variance (PC1), while disease state explained 27.4% of the variance (PC2, Fig. 5B). Interestingly, clinical severity correlated with deviation from normality (compare P1 and P2, Fig. 5B). Hierarchical clustering of differentially expressed genes showed that P1 and P2 differ from the controls in distinct fashion (Fig. 5B&C). P1, with a worse prognosis, showed much greater changes (Fig. 5C).

We then performed aggregated gene expression analysis of primary cells and iMPs. As expected, lineage difference resulted in different sets of differentially expressed genes in T, B, and macrophage cells (Fig. 5D). Over 4 times as many genes were commonly upregulated (121) than downregulated (29) in at least two cell types (Fig. 5D and S6F). The upregulated genes were enriched in autoimmune disease, chromatin-related processes, and cell proliferation (Fig. 5E). The commonly downregulated genes were not enriched in any pathway. The shared differentially expressed genes included: 1) genes involved in the autoimmune disease systemic lupus erythematosus (e.g.

HIST1H2AL, HIST1H3B, and STAT4), 2) epigenetic regulators (EZH2, LMNB1, UHRF1, and CENPE), 3) cell cycle regulators (MKI67, E2F2, PLK1, MYBL2, CDK1, cyclin B1, TOP2A, MCM3, Aurora kinase A, and p21), 4) DNA damage repair genes (BRCA1, Claspin, EXO1, RAD51, and POLQ (DNA polymerase theta)), and 5) immune function genes (IL2RA, CCR8, ANXA1 and HELLS) (Fig. 5C). Gene set enrichment analysis of RNAseq data of all patients and controls revealed statistically highly significant enrichment in cell cycle and cancer related pathways, suggesting that dysregulation of these pathways is a common feature in WAS regardless of genotype (Fig. 5F).

Patient T and B cells had 310 and 726 genes with differentially expressed isoforms, respectively (Fig. S6F). Most of these belong to the differentially expressed gene set. About 16% of the isoforms were shared by two or more cell types, and were highly enriched in categories including the nucleoplasm, protein binding, RNA binding, and nucleolus (Fig. S6F&G). We confirmed the splicing differences using an alternative analysis pipeline (Fig. S6H).

Patient cells had a strong preference for exon promotion in alternative splicing events (Fig. 5G and S6I). The genes with commonly promoted exons were enriched in protein and nucleic acid turnover (Fig. 5G). De novo motif analysis (see methods) identified multiple motifs in WAS prompted exons that overlap with binding sequences of human splicing factors (e.g., SRSF2, HNRNPA2B1 and etc.) in the SpliceAid-F (Giulietti et al., 2013) database (Fig. S7 and Table S3). This suggests that overexpression of splicing

factors (Fig. 4G and S6A) might contribute to the alterations in splicing regulation, although further study is needed to prove this. Together, the data from iMPs and patient lymphocytes strongly suggest that WAS leads to misregulation of cell cycle and RNA splicing.

WASP mutations derail cell proliferation

WAS patients exhibit both high prevalence (13–22%) and early onset of malignancy (average 9.5 years) (Candotti, 2018; Massaad et al., 2013). WAS patient cells displayed gross abnormality of nuclear bodies that are involved in the regulation of proliferation and tumorigenesis (Fig. 2C&D, S4C-F) (Lallemand-Breitenbach and de The, 2010; O'Day, 2013; Orsolic et al., 2016; Spector and Lamond, 2011). These observations and the dysregulation of cell cycle genes in WAS T, B, and macrophage cells (Fig. 5F) prompted us to examine proliferation of WAS mutant cells.

WAS-iMPs expressed higher levels of the bona fide proliferation marker KI67 (Bullwinkel et al., 2006), which was also shown by RNAseq (Fig. 4C, 6A, and S8A). WAS-iMPs showed a higher proliferation rate than cWAS-iMPs during the first two days of plating immediately after differentiation but proliferated little after maturation (Fig. S8B). WAS^{null} B cells showed accelerated growth that was rescued by lentiviral expression of wild-type WASP (Fig. 6B).

Mechanistically, WASP deficiency was associated with reduced protein levels of the tumor suppressors p53, p16 (CDKN2A) and p21 (CDKN1A), which inhibit proliferation

and promote differentiation (Fig. 6C). The suppression of the p16 and the p53-p21 axis was consistent with our RNA-seq data (Fig. 4C). WAS cells also showed concomitant downregulation of p21 and C/EBP β (a key regulator of immune cell differentiation and function) and activation of STAT3 (Fig. 6C), which is reminiscent of the aberrant proliferative state in acute myeloid leukemias (Gallardo et al., 2015). Additionally, WASP deficiency was associated with elevated levels of c-MYC and hyperphosphorylation of RB, which is consistent with the significant upregulation of MYC and E2F target genes across patient samples (Fig. 5F&6C). Overexpression of c-MYC drives ribosomal biogenesis and protein synthesis, which could explain the nucleolus hypertrophy phenotype in mutant cells (van Riggelen et al., 2010). Additionally, EZH2 levels were elevated in WAS^{null} B cells and WAS-iMPs (Fig. 5C, 6C and S8C). Overexpression of EZH2 has been described in multiple hematologic malignancies, including myelodysplastic syndrome (MDS) and lymphomas (Lund et al., 2014), both observed in WAS patients. Since EZH2 epigenetically represses the cyclin-dependent kinase inhibitor p16, overexpression of EZH2 could contribute to the cell cycle dysregulation in WAS (Cakouros et al., 2012).

Accumulating evidence indicates that dysregulation of alternative splicing promotes tumorigenesis (Dvinge et al., 2016). Although functional consequences of most alternatively spliced isoforms are unclear, specific isoforms have been found to be enriched in cancer due to their ability to promote proliferation and metastasis (David and Manley, 2010; Zhang and Manley, 2013). The full-length isoform of cortactin (CTTN-wt) is overexpressed in B-cell chronic lymphocytic leukemia but not normal B cells, and

correlates with poor prognosis (Gattazzo et al., 2014). RNA-seq and semi-quantitative isoform specific RT-PCR showed that only WAS-iMPs expressed CTTN-wt, whereas wild-type iMPs expressed a shorter isoform lacking exon 11 (CTTN-SV, Fig. 6D, E and I). Moreover, overexpression of WASP mutants resulted in a dramatic increase of CTTN-wt (up to 300-fold) specifically, proving that mutant WASP is responsible for the alternative splicing event (Fig. 6F and S8D). Intriguingly, exon 11 contains a high density of the consensus exonic splicing enhancer motifs of SRSF2 (Fig. 6D). The levels of the transcript containing exon 11 correlated with the levels of SRSF2 (Fig. 6F&G).

FOXM1, a key cell cycle regulator that is overexpressed in many forms of cancer (Myatt and Lam, 2007), was overexpressed in patient cells (Fig. S8E). Consequently, FOXM1 target genes were significantly upregulated in patient cells (Fig. 5F). Interestingly, only the isoforms that are specifically involved in cancer (Kong et al., 2013; Liu et al., 2006) (e.g. FOXM1B and FOXM1C, lacking exon 9 and exons 6 and 9, respectively) were upregulated in patient cells (Fig. 6H&I).

Strikingly, the mutant B cells with very high levels of CTTN-wt were able to form large spherical colonies in the soft-agar colony forming assay, a hallmark of oncogenic transformation (Fig. 6J). They also lost expression of a lineage marker, underwent dramatic changes in morphologic characteristics (Fig. S2J and S8F), and displayed an accelerated cell cycle (a ~2-fold reduction of G1 cells, and a ~2.5-fold increase of S-phase cells, Fig. 6J&S8G). These phenotypes show that the expression of mutant

WASP leads to reprogramming of cellular identity that is often seen in cancers (Wainwright and Scaffidi, 2017). Together, these data suggest that mutant WASP drives multiple oncogenic changes, part of which can be attributed to dysregulation of the alternative splicing pathway (Fig. 6K).

DISCUSSION

WASP is a multifaceted protein important for diverse hematopoietic and immune processes (Thrasher and Burns, 2010). Mutations in WASP cause a complex array of clinical manifestations, the most striking of which is early-onset malignancy (Yoshimi et al., 2013). The pathogenic mechanisms of WAS remain poorly defined, partly due to the lack of an experimental system of human hematopoiesis suitable for modeling this complex disease. Here we take advantage of isogenic WAS-iPSC models to investigate WAS pathogenesis in multiple blood cell types.

The nuclear body phenotypes of WAS mutants have at least two mechanistic explanations. First, it could be an adaptive response to the aberrant gene expression alterations in the mutants. The nucleolus, nuclear speckles, and PML-NBs are known to respond to the transcriptional and proliferation demands of the cell (Lallemand-Breitenbach and de The, 2010; O'Day, 2013; Spector and Lamond, 2011). However, they also play key regulatory roles in transcription, heterochromatin, and splicing, so the direction of causality needs to be established in future studies. Second, WASP directly participates in the formation and maintenance of the nuclear bodies. It could achieve this by facilitating the assembly of nuclear scaffolding structures needed for nuclear

body formation. Nuclear speckles reorganize upon inhibition of transcription by RNA Pol II, which requires actin polymerization (Sinclair and Brasch, 1978). Monomeric actin is found concentrated in nuclear speckles and is thought to stage RNA-processing factors, whereas filamentous actin is proposed to act as a scaffold in inter-chromatin regions (Belin et al., 2013). Mutant WASP defective in actin-related function could conceivably lead to disarray of nuclear organization. It is also possible that WASP facilitates the dynamic movement of factors required for the maintenance of nuclear bodies. The presence of many components of nuclear bodies in WASP-containing complexes supports this hypothesis. WASP interacts with different splicing factors in distinct subcellular locations (Fig. 3D). It will be interesting to study how mutations affect the location of such interaction in the future. WASP can be modified by SUMOylation, which may regulate its targeting to chromatin and various nuclear bodies and interaction with other proteins (Sarkar et al., 2015). Live imaging of the dynamics of WASP partners in the isogenic models will help resolve these possibilities.

This study uncovers many pro-oncogenic changes caused by WASP mutations, including loss of heterochromatin integrity, nucleolar hypertrophy, dysregulation of cancer related pathways, aberrant splicing, and accelerated proliferation. The prevalence of cancer in WAS has been attributed to oncogenic viruses and immunodeficiency (i.e. Epstein-Barr virus (EBV)-driven lymphoma) (Thrasher and Burns, 2010). However, up-to-date medical reports show that the frequency of other types of malignancy, including MDS and leukemias, in WAS patients is higher than that of EBV-associated malignancy (Albert et al., 2010; Imai et al., 2004; Moratto et al., 2011). The high incidence and early onset of these malignancies among WAS patients are well

outside population norms and cannot be explained by EBV infection alone.

Our data suggest that predisposition to cancer originates from cell intrinsic pro-oncogenic changes. Recent findings that infant WAS patients present clinical features indistinguishable from juvenile myelomonocytic leukemia that are free of viruses support this hypothesis (Sano et al., 2012; Stieglitz et al., 2015; Yoshimi et al., 2013). Our findings thus open up new directions to study the roles of WASP, nuclear bodies, and splicing factors in carcinogenesis. For instance, many epigenetic and splicing regulators are dysregulated in WAS mutants, and interact with WASP. EZH2 and SRSF2 are among the most frequently mutated genes in MDS and myeloproliferative neoplasm (Raza and Galili, 2012). The SRSF2^{P95H} mutation has been shown to contribute to MDS through mutant-specific effects on exon recognition (Kim et al., 2015). Our data suggest a different mechanism, where changes in splicing factor levels contribute to aberrant splicing and oncogenesis. We also showed that WASP interacted with many DNA repair proteins (including TOP1, 2A & 2B, as reported (Sarkar et al., 2017), Table S1). Moreover, the DNA repair pathway was dysregulated in WAS mutant cells, including genes involved in homologous recombination (e.g., RAD51 and BRCA1) and error-prone DNA repair pathways (POLQ and UHRF1) (Fig. 5C).

These findings suggest that the cellular models here presented can be used for elucidating how the recently identified factors contribute to oncogenic transformation. Finally, the signature alterations in gene expression and alternative splicing described in

this study may serve as novel biomarkers for cancer risk in WAS patients, and constitute the basis upon which to develop targeted therapies for WAS-associated malignancies.

ACKNOWLEDGEMENTS

We would like to thank B. Lubin for advice on hematopoiesis and assistance in cord blood unit procurement; C. Cole, R. Walker, J. Lieu, R. Tonai, M. Swearingen for providing cord blood samples; Y. Zheng for advice on hematopoietic differentiation and assistance on flow cytometry; members of the Belmonte lab for helpful discussions; A. T. Gutiérrez, A. Geobl, R. D. Soligalla and Y. Hishida for technical assistance, L. Mack, J. Olvera, C. O'Connor, E. O'Connor and K. Marquez for help with flow cytometry; D. O'Keefe for critical reading of the manuscript; M. Schwarz, P. Schwarz, C. Xia and X. Zhang for administrative help. **Funding:** M.L. was supported by grants from the King Abdullah University of Science and Technology. This work was supported by the Waitt Advanced Biophotonics Core Facility of the Salk Institute with funding from NIH-NCI CCSG: P30 014195, NINDS Neuroscience Core Grant and the Waitt Foundation. J.C.I.B. was supported by grants from The Moxie Foundation, G. Harold and Leila Y. Mathers Charitable Foundation, The Leona M. and Harry B. Helmsley Charitable Trust grant #2012-PG-MED002, The Glenn Foundation and Universidad Católica San Antonio de Murcia (UCAM). Work in the laboratory of M.L. was supported by the KAUST Office of Sponsored Research (OSR) under Awards No. BAS/1/1080-01. **Author contributions:** M.L., K.S. and J.C.I.B conceived the project, designed the experiments and performed data analysis. K.S., T.Y., X.Z, J.J.M., S.D., R.H.B., T.H., N.Y.K, M.M.A., M.K., Y.T., J.X., J.Q., E.A., J.Q., G.L., Z.L., F.Y, C.B., Y.G., C.S., J.R.Y.III., X.D.F., and C.R.E. contributed to project design and conducted the experiments. L.H., C.B., M.S. bioinformatics design, carried out bioinformatics work, and interpreted data. P.C. and

F.F. contributed to the design and experiments related to the primary patient samples.

M.L. and J.C.I.B. wrote the paper. **Competing interests:** Authors declare no competing interests. **Data and materials availability:** All data is available in the main text or the supplementary materials.

REFERENCES:

- Albert, M.H., Bittner, T.C., Nonoyama, S., Notarangelo, L.D., Burns, S., Imai, K., Espanol, T., Fasth, A., Pellier, I., Strauss, G., *et al.* (2010). X-linked thrombocytopenia (XLT) due to WAS mutations: clinical characteristics, long-term outcome, and treatment options. *Blood* 115, 3231-3238.
- Arterbery, A.S., Osafo-Addo, A., Avitzur, Y., Ciarleglio, M., Deng, Y., Lobritto, S.J., Martinez, M., Hafler, D.A., Kleinewietfeld, M., and Ekong, U.D. (2016). Production of Proinflammatory Cytokines by Monocytes in Liver-Transplanted Recipients with De Novo Autoimmune Hepatitis Is Enhanced and Induces TH1-like Regulatory T Cells. *J Immunol* 196, 4040-4051.
- Belin, B.J., Cimini, B.A., Blackburn, E.H., and Mullins, R.D. (2013). Visualization of actin filaments and monomers in somatic cell nuclei. *Mol Biol Cell* 24, 982-994.
- Bickmore, W.A. (2013). The spatial organization of the human genome. *Annual review of genomics and human genetics* 14, 67-84.
- Binks, M., Jones, G.E., Brickell, P.M., Kinnon, C., Katz, D.R., and Thrasher, A.J. (1998). Intrinsic dendritic cell abnormalities in Wiskott-Aldrich syndrome. *European journal of immunology* 28, 3259-3267.
- Bullwinkel, J., Baron-Luhr, B., Ludemann, A., Wohlenberg, C., Gerdes, J., and Scholzen, T. (2006). Ki-67 protein is associated with ribosomal RNA transcription in quiescent and proliferating cells. *Journal of cellular physiology* 206, 624-635.
- Cakouros, D., Isenmann, S., Cooper, L., Zannettino, A., Anderson, P., Glackin, C., and Gronthos, S. (2012). Twist-1 induces Ezh2 recruitment regulating histone methylation along the Ink4A/Arf locus in mesenchymal stem cells. *Molecular and cellular biology* 32, 1433-1441.
- Candotti, F. (2018). Clinical Manifestations and Pathophysiological Mechanisms of the Wiskott-Aldrich Syndrome. *J Clin Immunol* 38, 13-27.
- Castello, A., Fischer, B., Hentze, M.W., and Preiss, T. (2013). RNA-binding proteins in Mendelian disease. *Trends in genetics : TIG* 29, 318-327.
- Choi, K.D., Vodyanik, M., and Slukvin, I. (2011). Hematopoietic differentiation and production of mature myeloid cells from human pluripotent stem cells. *Nature protocols* 6, 296-313.
- Datsenko, K.A., and Wanner, B.L. (2000). One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc Natl Acad Sci U S A* 97, 6640-6645.
- David, C.J., and Manley, J.L. (2010). Alternative pre-mRNA splicing regulation in cancer: pathways and programs unhinged. *Genes & development* 24, 2343-2364.
- de Stanchina, E., Querido, E., Narita, M., Davuluri, R.V., Pandolfi, P.P., Ferbeyre, G., and Lowe, S.W. (2004). PML is a direct p53 target that modulates p53 effector functions. *Molecular cell* 13, 523-535.
- Dellaire, G., and Bazett-Jones, D.P. (2004). PML nuclear bodies: dynamic sensors of DNA damage and cellular stress. *Bioessays* 26, 963-977.
- Dvinge, H., Kim, E., Abdel-Wahab, O., and Bradley, R.K. (2016). RNA splicing factors as oncoproteins and tumour suppressors. *Nature reviews Cancer* 16, 413-430.
- Fox, A.H., and Lamond, A.I. (2010). Paraspeckles. *Cold Spring Harbor perspectives in biology* 2, a000687.

Galganski, L., Urbanek, M.O., and Krzyzosiak, W.J. (2017). Nuclear speckles: molecular organization, biological function and role in disease. *Nucleic acids research* 45, 10350-10368.

Gallardo, M., Lee, H.J., Zhang, X., Bueso-Ramos, C., Paeon, L.R., McArthur, M., Multani, A., Nazha, A., Manshouri, T., Parker-Thornburg, J., *et al.* (2015). hnRNP K Is a Haploinsufficient Tumor Suppressor that Regulates Proliferation and Differentiation Programs in Hematologic Malignancies. *Cancer cell* 28, 486-499.

Gattazzo, C., Martini, V., Frezzato, F., Trimarco, V., Tibaldi, E., Castelli, M., Facco, M., Zonta, F., Brunati, A.M., Zambello, R., *et al.* (2014). Cortactin, another player in the Lyn signaling pathway, is over-expressed and alternatively spliced in leukemic cells from patients with B-cell chronic lymphocytic leukemia. *Haematologica* 99, 1069-1077.

Giulietti, M., Piva, F., D'Antonio, M., D'Onorio De Meo, P., Paoletti, D., Castrignano, T., D'Erchia, A.M., Picardi, E., Zambelli, F., Principato, G., *et al.* (2013). SpliceAid-F: a database of human splicing factors and their RNA-binding sites. *Nucleic acids research* 41, D125-131.

Greer, W.L., Higgins, E., Sutherland, D.R., Novogrodsky, A., Brockhausen, I., Peacocke, M., Rubin, L.A., Baker, M., Dennis, J.W., and Siminovitch, K.A. (1989). Altered expression of leucocyte sialoglycoprotein in Wiskott-Aldrich syndrome is associated with a specific defect in O-glycosylation. *Biochemistry and cell biology = Biochimie et biologie cellulaire* 67, 503-509.

Grillari, J., Ajuh, P., Stadler, G., Loscher, M., Voglauer, R., Ernst, W., Chusainow, J., Eisenhaber, F., Pokar, M., Fortschegger, K., *et al.* (2005). SNEV is an evolutionarily conserved splicing factor whose oligomerization is necessary for spliceosome assembly. *Nucleic acids research* 33, 6868-6883.

Guettg, C., and Santoro, R. (2012). Formation of nuclear heterochromatin: the nucleolar point of view. *Epigenetics* 7, 811-814.

Imai, K., Morio, T., Zhu, Y., Jin, Y., Itoh, S., Kajiwara, M., Yata, J., Mizutani, S., Ochs, H.D., and Nonoyama, S. (2004). Clinical course of patients with WASP gene mutations. *Blood* 103, 456-464.

Ingrungruanglert, P., Amarinthnukrowh, P., Rungsiwiwut, R., Maneesri-le Grand, S., Sosothikul, D., Suphapeetiporn, K., Israsena, N., and Shotelersuk, V. (2015). Wiskott-Aldrich syndrome iPS cells produce megakaryocytes with defects in cytoskeletal rearrangement and proplatelet formation. *Thrombosis and haemostasis* 113, 792-805.

Inoue, H., Nagata, N., Kurokawa, H., and Yamanaka, S. (2014). iPS cells: a game changer for future medicine. *The EMBO journal* 33, 409-417.

Kim, E., Ilagan, J.O., Liang, Y., Daubner, G.M., Lee, S.C., Ramakrishnan, A., Li, Y., Chung, Y.R., Micol, J.B., Murphy, M.E., *et al.* (2015). SRSF2 Mutations Contribute to Myelodysplasia by Mutant-Specific Effects on Exon Recognition. *Cancer cell* 27, 617-630.

Kong, X., Li, L., Li, Z., Le, X., Huang, C., Jia, Z., Cui, J., Huang, S., Wang, L., and Xie, K. (2013). Dysregulated expression of FOXM1 isoforms drives progression of pancreatic cancer. *Cancer research* 73, 3987-3996.

Lallemant-Breitenbach, V., and de The, H. (2010). PML nuclear bodies. *Cold Spring Harbor perspectives in biology* 2, a000661.

Laskowski, T.J., Van Caeneghem, Y., Pourebrahim, R., Ma, C., Ni, Z., Garate, Z., Crane, A.M., Li, X.S., Liao, W., Gonzalez-Garay, M., *et al.* (2016). Gene Correction of

iPSCs from a Wiskott-Aldrich Syndrome Patient Normalizes the Lymphoid Developmental and Functional Defects. *Stem cell reports* 7, 139-148.

Lee, N.P., Mruk, D.D., Conway, A.M., and Cheng, C.Y. (2004). Zyxin, axin, and Wiskott-Aldrich syndrome protein are adaptors that link the cadherin/catenin protein complex to the cytoskeleton at adherens junctions in the seminiferous epithelium of the rat testis. *Journal of andrology* 25, 200-215.

Lensch, M.W., and Daley, G.Q. (2004). Origins of mammalian hematopoiesis: in vivo paradigms and in vitro models. *Current topics in developmental biology* 60, 127-196.

Li, H., Qiu, J., and Fu, X.D. (2012). RASL-seq for massively parallel and quantitative analysis of gene expression. *Current protocols in molecular biology* / edited by Frederick M Ausubel [et al] *Chapter 4*, Unit 4 13 11-19.

Li, M., Suzuki, K., Qu, J., Saini, P., Dubova, I., Yi, F., Lee, J., Sancho-Martinez, I., Liu, G.H., and Izpisua Belmonte, J.C. (2011). Efficient correction of hemoglobinopathy-causing mutations by homologous recombination in integration-free patient iPSCs. *Cell research* 21, 1740-1744.

Linder, S., Nelson, D., Weiss, M., and Aepfelbacher, M. (1999). Wiskott-Aldrich syndrome protein regulates podosomes in primary human macrophages. *Proceedings of the National Academy of Sciences of the United States of America* 96, 9648-9653.

Liu, G.H., Guan, T., Datta, K., Coppinger, J., Yates, J., 3rd, and Gerace, L. (2009). Regulation of myoblast differentiation by the nuclear envelope protein NET39. *Molecular and cellular biology* 29, 5800-5812.

Liu, G.H., Qu, J., Suzuki, K., Nivet, E., Li, M., Montserrat, N., Yi, F., Xu, X., Ruiz, S., Zhang, W., *et al.* (2012). Progressive degeneration of human neural stem cells caused by pathogenic LRRK2. *Nature* 491, 603-607.

Liu, G.H., Suzuki, K., Li, M., Qu, J., Montserrat, N., Tarantino, C., Gu, Y., Yi, F., Xu, X., Zhang, W., *et al.* (2014). Modelling Fanconi anemia pathogenesis and therapeutics using integration-free patient-derived iPSCs. *Nature communications* 5, 4330.

Liu, G.H., Suzuki, K., Qu, J., Sancho-Martinez, I., Yi, F., Li, M., Kumar, S., Nivet, E., Kim, J., Soligalla, R.D., *et al.* (2011). Targeted gene correction of laminopathy-associated LMNA mutations in patient-specific iPSCs. *Cell stem cell* 8, 688-694.

Liu, M., Dai, B., Kang, S.H., Ban, K., Huang, F.J., Lang, F.F., Aldape, K.D., Xie, T.X., Pelloski, C.E., Xie, K., *et al.* (2006). FoxM1B is overexpressed in human glioblastomas and critically regulates the tumorigenicity of glioma cells. *Cancer research* 66, 3593-3602.

Lorenzi, R., Brickell, P.M., Katz, D.R., Kinnon, C., and Thrasher, A.J. (2000). Wiskott-Aldrich syndrome protein is necessary for efficient IgG-mediated phagocytosis. *Blood* 95, 2943-2946.

Lund, K., Adams, P.D., and Copland, M. (2014). EZH2 in normal and malignant hematopoiesis. *Leukemia* 28, 44-49.

Massaad, M.J., Ramesh, N., and Geha, R.S. (2013). Wiskott-Aldrich syndrome: a comprehensive review. *Annals of the New York Academy of Sciences* 1285, 26-43.

Misteli, T., and Spector, D.L. (1996). Serine/threonine phosphatase 1 modulates the subnuclear distribution of pre-mRNA splicing factors. *Mol Biol Cell* 7, 1559-1572.

Mitani, K., Graham, F.L., Caskey, C.T., and Kochanek, S. (1995). Rescue, propagation, and partial purification of a helper virus-dependent adenovirus vector. *Proc Natl Acad Sci U S A* 92, 3854-3858.

Mohamad, N., and Boden, M. (2010). The proteins of intra-nuclear bodies: a data-driven analysis of sequence, interaction and expression. *BMC Syst Biol* 4, 44.

Mohammed, H., Taylor, C., Brown, G.D., Papachristou, E.K., Carroll, J.S., and D'Santos, C.S. (2016). Rapid immunoprecipitation mass spectrometry of endogenous proteins (RIME) for analysis of chromatin complexes. *Nature protocols* 11, 316-326.

Moratto, D., Giliani, S., Bonfim, C., Mazzolari, E., Fischer, A., Ochs, H.D., Cant, A.J., Thrasher, A.J., Cowan, M.J., Albert, M.H., *et al.* (2011). Long-term outcome and lineage-specific chimerism in 194 patients with Wiskott-Aldrich syndrome treated by hematopoietic cell transplantation in the period 1980-2009: an international collaborative study. *Blood* 118, 1675-1684.

Myatt, S.S., and Lam, E.W. (2007). The emerging roles of forkhead box (Fox) proteins in cancer. *Nature reviews Cancer* 7, 847-859.

O'Day, D.H. (2013). Proteins of the nucleolus : regulation, translocation, & biomedical functions (New York: Springer).

Okita, K., Matsumura, Y., Sato, Y., Okada, A., Morizane, A., Okamoto, S., Hong, H., Nakagawa, M., Tanabe, K., Tezuka, K., *et al.* (2011). A more efficient method to generate integration-free human iPS cells. *Nat Methods* 8, 409-412.

Orsolic, I., Jurada, D., Pullen, N., Oren, M., Eliopoulos, A.G., and Volarevic, S. (2016). The relationship between the nucleolus and cancer: Current evidence and emerging paradigms. *Semin Cancer Biol* 37-38, 36-50.

Palmer, D.J., and Ng, P. (2004). Physical and infectious titers of helper-dependent adenoviral vectors: a method of direct comparison to the adenovirus reference material. *Mol Ther* 10, 792-798.

Pandit, S., Zhou, Y., Shiue, L., Coutinho-Mansfield, G., Li, H., Qiu, J., Huang, J., Yeo, G.W., Ares, M., Jr., and Fu, X.D. (2013). Genome-wide analysis reveals SR protein cooperation and competition in regulated splicing. *Molecular cell* 50, 223-235.

Piller, F., Le Deist, F., Weinberg, K.I., Parkman, R., and Fukuda, M. (1991). Altered O-glycan synthesis in lymphocytes from patients with Wiskott-Aldrich syndrome. *The Journal of experimental medicine* 173, 1501-1510.

Rajapakse, I., and Groudine, M. (2011). On emerging nuclear order. *The Journal of cell biology* 192, 711-721.

Ramesh, N., Anton, I.M., Hartwig, J.H., and Geha, R.S. (1997). WIP, a protein associated with wiskott-aldrich syndrome protein, induces actin polymerization and redistribution in lymphoid cells. *Proceedings of the National Academy of Sciences of the United States of America* 94, 14671-14676.

Raza, A., and Galili, N. (2012). The genetic basis of phenotypic heterogeneity in myelodysplastic syndromes. *Nature reviews Cancer* 12, 849-859.

Remold-O'Donnell, E., Kenney, D.M., Parkman, R., Cairns, L., Savage, B., and Rosen, F.S. (1984). Characterization of a human lymphocyte surface sialoglycoprotein that is defective in Wiskott-Aldrich syndrome. *The Journal of experimental medicine* 159, 1705-1723.

Sadhukhan, S., Sarkar, K., Taylor, M., Candotti, F., and Vyas, Y.M. (2014). Nuclear role of WASp in gene transcription is uncoupled from its ARP2/3-dependent cytoplasmic role in actin polymerization. *J Immunol* 193, 150-160.

Sano, H., Kobayashi, R., Suzuki, D., Yasuda, K., Nakanishi, M., Nagashima, T., Yamada, M., and Kobayashi, K. (2012). Wiskott-Aldrich syndrome with unusual clinical features similar to juvenile myelomonocytic leukemia. *Int J Hematol* 96, 279-283.

Sarkar, K., Han, S.S., Wen, K.K., Ochs, H.D., Dupre, L., Seidman, M.M., and Vyas, Y.M. (2017). R-loops cause genomic instability in T helper lymphocytes from patients with Wiskott-Aldrich syndrome. *J Allergy Clin Immunol*.

Sarkar, K., Sadhukhan, S., Han, S.S., and Vyas, Y.M. (2014). Disruption of hSWI/SNF complexes in T cells by WAS mutations distinguishes X-linked thrombocytopenia from Wiskott-Aldrich syndrome. *Blood* 124, 3409-3419.

Sarkar, K., Sadhukhan, S., Han, S.S., and Vyas, Y.M. (2015). SUMOylation-disrupting WAS mutation converts WASp from a transcriptional activator to a repressor of NF-kappaB response genes in T cells. *Blood* 126, 1670-1682.

Senju, S., Haruta, M., Matsumura, K., Matsunaga, Y., Fukushima, S., Ikeda, T., Takamatsu, K., Irie, A., and Nishimura, Y. (2011). Generation of dendritic cells and macrophages from human induced pluripotent stem cells aiming at cell therapy. *Gene therapy* 18, 874-883.

Shu, X., Lev-Ram, V., Deerinck, T.J., Qi, Y., Ramko, E.B., Davidson, M.W., Jin, Y., Ellisman, M.H., and Tsien, R.Y. (2011). A genetically encoded tag for correlated light and electron microscopy of intact cells, tissues, and organisms. *PLoS biology* 9, e1001041.

Siminovitch, K.A., Greer, W.L., Axelsson, B., Rubin, L.A., Novogrodsky, A., and Peacocke, M. (1993). Selective impairment of CD43-mediated T cell activation in the Wiskott-Aldrich syndrome. *Immunodeficiency* 4, 99-108.

Sinclair, G.D., and Brasch, K. (1978). The reversible action of alpha-amanitin on nuclear structure and molecular composition. *Exp Cell Res* 111, 1-14.

Snapper, S.B., Rosen, F.S., Mizoguchi, E., Cohen, P., Khan, W., Liu, C.H., Hagemann, T.L., Kwan, S.P., Ferrini, R., Davidson, L., *et al.* (1998). Wiskott-Aldrich syndrome protein-deficient mice reveal a role for WASP in T but not B cell activation. *Immunity* 9, 81-91.

Spector, D.L., and Lamond, A.I. (2011). Nuclear speckles. *Cold Spring Harbor perspectives in biology* 3.

Stadler, M., Chelbi-Alix, M.K., Koken, M.H., Venturini, L., Lee, C., Saib, A., Quignon, F., Pelicano, L., Guillemin, M.C., Schindler, C., *et al.* (1995). Transcriptional induction of the PML growth suppressor gene by interferons is mediated through an ISRE and a GAS element. *Oncogene* 11, 2565-2573.

Stieglitz, E., Taylor-Weiner, A.N., Chang, T.Y., Gelston, L.C., Wang, Y.D., Mazor, T., Esquivel, E., Yu, A., Seepo, S., Olsen, S., *et al.* (2015). The genomic landscape of juvenile myelomonocytic leukemia. *Nature genetics* 47, 1326-1333.

Strickfaden, H., Xu, Z.Z., and Hendzel, M.J. (2015). Visualization of miniSOG Tagged DNA Repair Proteins in Combination with Electron Spectroscopic Imaging (ESI). *Journal of visualized experiments : JoVE*.

Sutherland, H., and Bickmore, W.A. (2009). Transcription factories: gene expression in unions? *Nature reviews* 10, 457-466.

Suzuki, K., Mitsui, K., Aizawa, E., Hasegawa, K., Kawase, E., Yamagishi, T., Shimizu, Y., Suemori, H., Nakatsuji, N., and Mitani, K. (2008). Highly efficient transient gene expression and gene targeting in primate embryonic stem cells with helper-dependent

adenoviral vectors. *Proceedings of the National Academy of Sciences of the United States of America* 105, 13781-13786.

Suzuki, K., Yu, C., Qu, J., Li, M., Yao, X., Yuan, T., Goebel, A., Tang, S., Ren, R., Aizawa, E., *et al.* (2014). Targeted gene correction minimally impacts whole-genome mutational load in human-disease-specific induced pluripotent stem cell clones. *Cell stem cell* 15, 31-36.

Takahashi, K., Tanabe, K., Ohnuki, M., Narita, M., Ichisaka, T., Tomoda, K., and Yamanaka, S. (2007). Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell* 131, 861-872.

Takahashi, K., and Yamanaka, S. (2006). Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* 126, 663-676.

Taylor, M.D., Sadhukhan, S., Kottangada, P., Ramgopal, A., Sarkar, K., D'Silva, S., Selvakumar, A., Candotti, F., and Vyas, Y.M. (2010). Nuclear role of WASp in the pathogenesis of dysregulated TH1 immunity in human Wiskott-Aldrich syndrome. *Science translational medicine* 2, 37ra44.

Thrasher, A.J., and Burns, S.O. (2010). WASP: a key immunological multitasker. *Nature reviews Immunology* 10, 182-192.

van Riggelen, J., Yetil, A., and Felsher, D.W. (2010). MYC as a regulator of ribosome biogenesis and protein synthesis. *Nature reviews Cancer* 10, 301-309.

Verboon, J.M., Rincon-Arango, H., Werwie, T.R., Delrow, J.J., Scalzo, D., Nandakumar, V., Groudine, M., and Parkhurst, S.M. (2015). Wash interacts with lamin and affects global nuclear organization. *Curr Biol* 25, 804-810.

Vodyanik, M.A., and Slukvin, I. (2007). Hematoendothelial differentiation of human embryonic stem cells. *Curr Protoc Cell Biol Chapter* 23, Unit 23 26.

Wainwright, E.N., and Scaffidi, P. (2017). Epigenetics and Cancer Stem Cells: Unleashing, Hijacking, and Restricting Cellular Plasticity. *Trends Cancer* 3, 372-386.

Wang, I.F., Chang, H.Y., and Shen, C.K. (2006). Actin-based modeling of a transcriptionally competent nuclear substructure induced by transcription inhibition. *Exp Cell Res* 312, 3796-3807.

Yoshimi, A., Kamachi, Y., Imai, K., Watanabe, N., Nakadate, H., Kanazawa, T., Ozono, S., Kobayashi, R., Yoshida, M., Kobayashi, C., *et al.* (2013). Wiskott-Aldrich syndrome presenting with a clinical picture mimicking juvenile myelomonocytic leukaemia. *Pediatric blood & cancer* 60, 836-841.

Zhang, J., and Manley, J.L. (2013). Misregulation of pre-mRNA alternative splicing in cancer. *Cancer Discov* 3, 1228-1237.

Zhang, Z.Y., Xiao, H.Q., Jiang, L.P., Zhou, Y., Zhao, Q., Yu, J., Liu, W., Yang, X.Q., and Zhao, X.D. (2010). Analysis of clinical and molecular characteristics of Wiskott-Aldrich syndrome in 24 patients from 23 unrelated Chinese families. *Pediatr Allergy Immunol* 21, 522-532.

Zhou, Z., Qiu, J., Liu, W., Zhou, Y., Plocinik, R.M., Li, H., Hu, Q., Ghosh, G., Adams, J.A., Rosenfeld, M.G., *et al.* (2012). The Akt-SRPK-SR axis constitutes a major pathway in transducing EGF signaling to regulate alternative splicing in the nucleus. *Molecular cell* 47, 422-433.

Zhu, Q., Zhang, M., Blaese, R.M., Derry, J.M., Junker, A., Francke, U., Chen, S.H., and Ochs, H.D. (1995). The Wiskott-Aldrich syndrome and X-linked congenital thrombocytopenia are caused by mutations of the same gene. *Blood* 86, 3797-3804.

Figure Legend

Figure 1. WAS-iPSC models recapitulate WAS phenotypes. **A**, schematic diagram of the study design. **B**, representative phase-contrast microscopy images of WAS-iMPs, cWAS-iMPs and wild-type macrophages and quantitation of the morphology of WAS and normal macrophages. PB-MP: macrophages derived from peripheral blood; ctrl-MP: macrophages derived from wild-type iPSC. *** $p < 0.001$ (t -test). **C**, immunofluorescence analysis of podosomes in WAS-iMPs and cWAS-iMPs. Podosomes are marked by F-actin and cortactin (encoded by the CTTN gene). **D**, quantitative measurement of phagocytosis using FACS. Left: schematic of the principle of the phagocytosis assay using the pH-sensitive fluorescence dye pHrodo Red-conjugated *E. coli* particles. Right: FACS contour plots of phagocytic activity of WAS-iMP2 and cWAS-iMP2.

Figure 2. WASP regulates the nuclear organization. A, confocal

immunofluorescence analysis of the indicated nuclear body markers in the wild-type control iMPs. Nucleophosmin (NPM1, aka B23) marks the nucleolus. SRSF2 (aka SC35) marks the nuclear speckle. PML marks the PML-NB. Nuclei were stained with DAPI. Bar=5 μ m. **B**, serial block face scanning electron microscopy (SBF-SEM) images of wild-type macrophages expressing a WASP-miniSOG fusion protein. Sub-nuclear localization of WASP was visualized with photo-oxidation of DAB by miniSOG (see methods). White arrowheads indicate WASP signals as high-contrast dark spots. Yellow arrows and outlines indicate nuclear landmarks. 1, nucleolus; 2, nuclear envelope; 3, nuclear pores; 4, heterochromatin. Bar=1 μ m. **C**, confocal immunofluorescence analysis of the indicated nuclear body markers in WAS-iMP1s and cWAS-iMP1s. The rightmost column shows image-based quantitative analysis. For nucleolar area measurement: $p=5.2 \times 10^{-6}$, $n>10$. For PML-NB measurement: $p=1.1 \times 10^{-5}$, $n>10$. For SRSF2 MFI measurement: $p=8.7 \times 10^{-6}$, $n>15$. Insets: magnified view of the nucleolus. MFI: mean fluorescence intensity. *** $p<0.001$. Bar=5 μ m. **D and E**, representative TEM images of 70-nm sections of WAS-iMP and cWAS-iMP collected under identical conditions. Bar=1 μ m in the left and upper right panels; Bar=0.5 μ m in lower right panels and insets. NO: nucleolus, HC: heterochromatin, IGC: interchromatin granule cluster (nuclear speckles), NP, nuclear pore, M: mitochondria, A: autophagosome, G: Golgi. **F**, quantitative analysis of the area of the nucleolus in TEM images of WASP mutant and wild-type cells. ** $p<0.01$.

Figure 3. WASP interacts with nuclear factors and affects chromatin structure. **A**, Venn diagram showing shared and distinct proteins found in the interactome of WASP, the nucleolus, the nuclear speckle and the PML-NB. **B**, GO enrichment of genes identified as part of the WASP protein complex. **C**, co-immunoprecipitation (co-IP) validation of physical interaction between WASP and selected interacting proteins. The interaction between WASP and SRSF2 was verified by reciprocal co-IP. IP, immunoprecipitation. IB, immunoblotting. **D**, left panels: representative immunofluorescence images of PLA signals for WASP-SRSF2, WASP-SF3B3, and WASP-HNRNPA2B1 in GM11518 cells. Nuclei were stained with DAPI. Bar=5 μ m. Upper right graph: quantitative analysis of PLA signals. n=50 for each setting. **** $p<0.0001$. Lower right pie charts: quantitative analysis of PLA signals in different locations of the cell. SRSF2 is known to be nuclear. SF3B3 and HNRNPA2B1 are found in both the cytoplasm and the nucleus. The PLA signals are found in the expected sub-cellular locations. The percentage of signals in a given location is shown in the pie chart. NE: nuclear envelope. NO: nucleolus. OIN: other locations inside the nucleus. **E**, confocal immunofluorescence images of the histone modification marker H3K27me3 and nuclear envelope marker LMNB1 in WAS-iMPs and cWAS-iMPs. Bar=5 μ m. **F**, structured-illumination super-resolution immunofluorescence images of the histone modification H3K9me3 in WAS-iMPs and cWAS-iMPs. Bar=5 μ m. Yellow circles mark nucleolus. **G**, quantitative analysis of the contrast between heterochromatic and euchromatic regions in TEM images of WASP mutant and wild-type cells. ** $p<0.01$.

Figure 4. WASP mutations disrupt normal gene expression and alternative splicing programs in iMPs. **A**, quantitation of gene expression (in FPKM) of WAS-iMP (red) or cWAS-iMP (blue) verses those of wild-type CB-MP. Correlation coefficient $r=0.62$ before gene correction, compared to $r=0.88$ after. **B**, GO enrichment of genes differentially expressed between WAS-iMP1 and cWAS-iMP1. **C**, hierarchical clustering of WAS-iMP1 (in triplicate), cWAS-iMP1 (in triplicate), WAS-iMP2, cWAS-iMP2 and CB-MP using genes in the enriched GO terms in **B**. **D**, ELISA quantitation of cytokine secretion by macrophages upon stimulation with lipopolysaccharide (LPS). Values shown are mean $\pm$ SEM, $n=3$. * $p<0.05$, ** $p<0.01$, *** $p<0.001$. **E**, hierarchical clustering of all differentially expressed genes among indicated samples. Note that patient T and B cells formed separate clusters outside of the T or B cell clusters, indicating that disease-associated gene expression signature is dominant over the cell-type specific transcriptional program. **F**, PCA of RNA-seq data of WAS-iMPs and cWAS-iMPs derived from two patients, and the wild-type CB-MP. **G**, hierarchical clustering of samples shown in **C** based on differentially expressed splicing factors. **H**, GO enrichment of genes with differentially expressed isoforms between WAS-iMPs and cWAS-iMPs.

Figure 5. WASP pathogenic mutations disrupt normal gene expression and alternative splicing programs in primary patient lymphocytes. **A**, schematic diagram of the experimental design. **B**, PCA of RNA-seq data of patients (P1 and P2, clinical score of 4 and 2, respectively) and normal donors (D1 and D2). **C**, hierarchical clustering of genes that are differentially expressed in at least two cell types among indicated samples. **D**, Venn diagrams showing overlap of genes that are upregulated (DE up) in the three cell types. FDR <0.05. **E**, heatmap shows GO enrichment of genes differentially expressed in the three cell types. **F**, top enriched pathways identified by Gene Set Enrichment Analysis (GSEA) using RNAseq data of all patient and normal samples. NES: normalized enrichment score. NOM p: nominal p value. FDR: false discovery rate. **G**, Venn diagram showing the number of exons that are promoted in B and T cells from WAS patients compared to those from normal individuals. Bottom: GO Biological Process and KEGG pathway enrichment in genes containing commonly promoted exons in B and T cells from patients with WAS.

Figure 6. WASP deficiency causes dysregulation of pathways controlling cell proliferation. **A**, FACS analysis of Ki67 expression by intracellular staining in WAS-iMPs and cWAS-iMPs. **B**, cell proliferation curve of WAS mutant B cells (ID00003), wild-type B cells (GM11518) and mutant B cells rescued with wild-type WASP (ID00003/WASP). Data are shown as mean $\pm$ SEM. $n\geq 3$. **C**, Western blot analysis of the expression of tumor suppressors (p53, p16, p21 and RB), oncogenic factors (MYC, pSTAT3 and EZH2) and differentiation related transcription factor (C/EBP β) in WASP mutant and wild-type cells. HSP90 was used as a loading control. **D**, graph shows RNA-seq reads mapped to exon 10, 11 and 12 of the CTTN gene (transcript ID: NM_005231). The DNA sequence of the cassette exon 11 is shown below the genome browser view. The green boxes represent the distribution of the consensus exonic splicing enhancer binding motifs of SRSF2 (GGNG or CCNG) in exon 11. **E&F**, semi-quantitative isoform specific RT-PCR shows expression of two CTTN variants in macrophages (**E**) and B cell lines (**F**). Rep1 and Rep2: two biological replicates. **G**, Western blot analysis of the expression of SRSF2 and EZH2 in the wild-type, WAS^{null}, and WAS^{F36*}-expressing cells. GAPDH and PPIB were loading controls. **H**, graph shows RNA-seq reads mapped to exon 8, 9 and 10 of the FOXM1 gene (transcript ID: NM_202002). The DNA sequence of the cassette exon 9 is shown below the genome browser view. The green boxes represent the distribution of the consensus exonic splicing enhancer binding motifs of SRSF2 (GGNG or CCNG) in exon 9. **I**, isoform-level analysis of RNAseq data of WAS patient and control samples. Please note that CTTN is not expressed in T and B cells. **J**, representative images of soft-agar colony forming assay in the 6-well plate format (top row, $n=3$). Middle images show magnified view of the colonies. Bottom panels show cell

cycle analysis of the corresponding cells using propidium iodide. Right bar graph shows the number of colonies per well (n=3). **K**, schematic of a model showing how overexpression of splicing factors could result in aberrant splicing of cancer related genes and promote tumorigenesis. The black box represents a cassette exon. The green bars represent the distribution of binding motifs of splicing factors.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Mouse monoclonal anti-beta Actin antibody	Abcam	ab6276
Mouse monoclonal anti-Histone H3 (tri methyl K27) antibody	Abcam	ab6002
Mouse monoclonal anti-Histone H3 (tri methyl K9) antibody	Abcam	ab8898; RRID: AB_306848
Mouse monoclonal anti-HSP90	BD Biosciences	Cat# 610419
Mouse monoclonal anti-human CD43-APC (1:50 dilution)	BD Biosciences	Cat# 560198
Mouse monoclonal anti-CD45-APC (1:50 dilution)	BD Biosciences	Cat# 555485
Mouse monoclonal anti-CD34-PE (1:50 dilution)	BD Biosciences	Cat# 555822
Mouse monoclonal anti-human CD3 PE (1:50)	BD Biosciences	Cat# 555333
Mouse monoclonal anti-CD11c PE (1:50 dilution)	BD Biosciences	Cat# 560999
Mouse monoclonal anti-CD20 FITC (1:50 dilution)	BD Biosciences	Cat# 555622
Mouse monoclonal anti-CD8 FITC	BD Biosciences	Cat# 557085
Mouse monoclonal anti-CD4 PE	BD Biosciences	Cat# 560649
Mouse monoclonal anti-CD1a Alexa Fluor 700	BD Biosciences	Cat# 300120
Mouse monoclonal isotype control APC (1:50)	BD Biosciences	Cat# 555751
Mouse monoclonal isotype control PE (1:50)	BD Biosciences	Cat# 555749
Mouse monoclonal isotype control FITC (1:50)	BD Biosciences	Cat# 555742
Mouse monoclonal anti-WASP antibody (5A5)	BD Biosciences	Cat# 557773
Rabbit polyclonal anti-NPM1 antibody	Bethyl laboratories	Cat# A302-402A-M
Mouse monoclonal Alexa Fluor® 647 anti-human Ki-67 (1:20)	Biolegend	Cat# 350509
Mouse monoclonal anti-CD14 PE	Biolegend	Cat# 301805
Mouse monoclonal Alexa Fluor® 647 Mouse IgG1, κ Isotype control (1:20)	Biolegend	Cat# 400130
Mouse monoclonal anti-Ezh2 antibody	Cell Signaling Technology	Cat# 3147
Rabbit anti-c-Myc antibody	Cell Signaling Technology	Cat# 9402
Mouse monoclonal anti-Rb antibody	Cell Signaling Technology	Cat# 9309

Rabbit monoclonal anti-phospho-Rb (Ser780) antibody	Cell Signaling Technology	Cat# 8180
Rabbit monoclonal anti-phospho-Rb (Ser807/811) antibody	Cell Signaling Technology	Cat# 8516
Rabbit monoclonal anti-phospho-Rb (Ser795) antibody	Cell Signaling Technology	Cat# 9301
Mouse monoclonal anti-p53 antibody	Cell Signaling Technology	Cat# 2524
Rabbit anti-Stat3 antibody	Cell Signaling Technology	Cat# 9132
Rabbit anti-Phospho-Stat3 (Tyr705) antibody	Cell Signaling Technology	Cat# 9131
Mouse monoclonal anti-human CD235a-PE	eBioscience	Cat# 12-9987-80
Mouse monoclonal anti-human CD11b FITC	eBioscience	Cat# 11-0118-41
Mouse monoclonal anti-human CD209 (DC-SIGN)	eBioscience	Cat# 17-2099-41
Mouse monoclonal anti-human CD115 PE	eBioscience	Cat# 12-1159-41
Mouse monoclonal anti-human CD38 eFluor 450	eBioscience	Cat# 48-0388-42
Mouse monoclonal anti-SC35 (SRSF2)	Genetex	Cat# GTX11826
Rabbit polyclonal anti-SF3B3	Genetex	Cat# GTX106450
	Genetex	Cat#
	Genetex	Cat#
Mouse monoclonal anti-human CD34-APC (1:50)	Miltenyi Biotec	Cat#130-090-954
Mouse monoclonal anti-human CD34-PerCP Vio700 (1:50)	Miltenyi Biotec	Cat# 130-097-915
Mouse monoclonal anti-human CD163 FITC	Miltenyi Biotec	Cat# 130-099-969
Mouse monoclonal anti-human HLA-DR APC	Miltenyi Biotec	Cat# 130-095-297
Rabbit polyclonal anti-NONO antibody	Novus Biologicals	Cat# NB100-1556
Mouse monoclonal anti-human Tra-1-85 APC (1:50)	R&D Systems	Cat# FAB3195A
Rabbit polyclonal anti-cortactin	Santa Cruz Biotechnology	Cat# sc-11408
Mouse monoclonal anti-WASP antibody (B9)	Santa Cruz Biotechnology	Cat# sc-13139
Mouse monoclonal anti-WASP antibody (D1)	Santa Cruz Biotechnology	Cat# SC-5300
Rabbit polyclonal anti-WASP antibody (H250)	Santa Cruz Biotechnology	Cat# sc-8353
Goat polyclonal anti-PML antibody (N-19)	Santa Cruz Biotechnology	Cat# SC9862 X
Goat polyclonal anti-Lamin B (M-20)	Santa Cruz Biotechnology	Cat# sc-6217
C/EBP β antibody (C-19)	Santa Cruz Biotechnology	Cat# sc-150
Rabbit p21 antibody (C-19)	Santa Cruz Biotechnology	Cat# sc-397
Rabbit p21 antibody (N-20)	Santa Cruz Biotechnology	Cat# sc-469
Mouse monoclonal anti-p16 antibody (F-12)	Santa Cruz Biotechnology	Cat# sc-1661
Rabbit polyclonal anti-p16 antibody (N-20)	Santa Cruz Biotechnology	Cat# sc-467
Mouse monoclonal anti-FLAG® M2 antibody	Sigma-Aldrich	Cat# F1804
Bacterial and Virus Strains		

pCIHDAAdGT8-4	Dr. Kohnosuke Mitani, Saitama Medical University	PNAS 2008 September, 105 (37) 13781-13786.
pLenti-CAG-PL9	Drs. Inder M. Verma And Ronald Evans, the Salk Institute for Biological Studies	PNAS 2002 February, 99 (4) 2140-2145
pLKO.1-puro	Mission shRNA library	https://www.sigmaaldrich.com/catalog/product/sigma/shcng?lang=en&region=SA
Bacterial stock of shRNA lentiviral vector targeting the WAS gene	Mission shRNA library bacterial stock	TRCN0000029819
Bacterial stock of shRNA lentiviral vector targeting the WAS gene	Mission shRNA library bacterial stock	TRCN0000029822
Bacterial stock of shRNA lentiviral vector targeting the WAS gene	Mission shRNA library bacterial stock	TRCN0000029823
Biological Samples		
Human fibroblasts from a male patient with WAS	Coriell institute	https://www.coriell.org/
Human fibroblasts from a male patient with WAS	Cell Line and DNA Bank from Patients Affected by Genetic Diseases (Genova, Italy)	http://dppm.gaslini.org/biobank/
Human peripheral blood mononuclear cells from male patients of WAS	This paper	N/A
Human fibroblasts from a normal individual	Cell Line and DNA Bank from Patients Affected by Genetic Diseases (Genova, Italy)	http://dppm.gaslini.org/biobank/
Human buffy coat	San Diego Blood Bank (San Diego, CA)	https://www.sandiegobloodbank.org/
Human umbilical cord blood	San Diego Blood Bank (San Diego, CA) and StemCyt (Covina, CA)	https://www.sandiegobloodbank.org/ ; https://www.stemcyte.com/
Chemicals, Peptides, and Recombinant Proteins		
Polybrene	Millipore	Cat# TR-1003
Sodium butyrate	Stemgent	Cat# 04-0005
Ascorbic acid	Sigma-Aldrich	Cat# A4544
monothioglycerol	Sigma-Aldrich	Cat# M6145
Hyclone defined FBS	GE life sciences	SH30070.03
pHrodo™ Red E. coli BioParticles® Conjugate for Phagocytosis	ThermoFisher Scientific	Cat# P35361
NuPAGE Novex 4%–12% Bis-Tris protein gels	ThermoFisher Scientific	Cat# NP0321
Immobilon-P Membrane, PVDF	Millipore	Cat# IPVH00010
ECL SuperSignal	GE Healthcare	Cat# RPN2232

iScript Reverse Transcription Supermix	Bio-Rad	Cat# 1708840
SsoAdvanced SYBR Green Supermix	Bio-Rad	Cat# 1725270
LPS-Lipopolysaccharides from Escherichia coli 0111:B4	Sigma-Aldrich	Cat# L3024-5MG
Lipofectamine 3000	ThermoFisher Scientific	Cat# L3000-015
Critical Commercial Assays		
CytoSelect™ 24-Well Cell Migration Assay, 5 µm	Cell Biolabs Inc.	Cat# CBA-102
CellTiter 96 AQueous One cell proliferation assay kit	Promega	Cat# G3582
The Gibson Assembly® HiFi 1-Step Kit	Synthetic Genomics	Cat# GA1100-10
TruSeq Stranded mRNA Library Prep Kit	Illumina	Cat# RS-122-2101
TruSeq RNA Access Library Prep Kit	Illumina	Cat# RS-301-2001
Nuclei EZ Prep kit	Sigma-Aldrich	Cat# NUC101
Peprotech Mini ELISA development kit in conjunction with TMB substrate	Peprotech	Cat# 52-00-02
Hematopoietic CFU assay with MethoCult GF + H4435 semisolid medium	Stem Cell Technologies	Cat# 04435
Hematopoietic CFU assay with StemMACS human HSC-CFU complete medium with Epo	Miltenyi Biotec	Cat# 130-091-280
DUO-LINK in situ red starter kit mouse/rabbit	Sigma Aldrich	Cat# DUO92101-1KT
Deposited Data		
The raw sequencing data and the processed gene expression tables.	NCBI GEO database	GSE107963
Experimental Models: Cell Lines		
293t	ATCC	Cat# CRL-3216
Lymphoblastoid cell line GM11528	Coriell institute	https://www.coriell.org/
Lymphoblastoid cell line ID00003	Coriell institute	https://www.coriell.org/
Lymphoblastoid cell line GM16749	Coriell institute	https://www.coriell.org/
Lymphoblastoid cell line GM13022	Coriell institute	https://www.coriell.org/
Oligonucleotides		
See Table S4	This paper	N/A
Recombinant DNA		
pMDLg/pRRE	Addgene	Cat# 12251
pRSV-Rev	Addgene	Cat# 12253
pMD2.G	Addgene	Cat# 12259
mEos3.2-C1	Addgene	Cat# 54550
miniSOG-N1	Obtained under MTA from Roger Y Tsien lab at UCSD	PLoS biology 9, e1001041
Software and Algorithms		
STAR	https://github.com/alexdobin/STAR	https://academic.oup.com/bioinformatics/article/29/1/15/272537

DAVID (v6.8)	https://david.ncifcrf.gov	Nature Protocols 2009; 4(1):44 & Nucleic Acids Res. 2009;37(1):1
Cuffdiff (v2.2.1)	cole-trapnell-lab.github.io/cufflinks/releases/v2.2.1/	https://www.nature.com/articles/nbt.2450
IGV	software.broadinstitute.org/software/igv/	Nature Biotechnology 29, 24–26 (2011)
Java TreeView	jtreeview.sourceforge.net/	http://jtreeview.sourceforge.net/docs/AlokThesisChap2.pdf
Cluster 3.0	https://www.encodeproject.org/software/cluster/	PMID:14871861
Ingenuity pathway analysis	Qiagen bioinformatics	https://www.qiagenbioinformatics.com/products/ingenuity-pathway-analysis/
JunctionSeq	https://bioconductor.org/packages/release/bioc/html/JunctionSeq.html	https://www.ncbi.nlm.nih.gov/pubmed/27257077
Cytoscape	http://www.cytoscape.org	Genome Research 2003 Nov; 13(11):2498-504

CONTACT FOR REAGENTS AND RESOURCE SHARING

Further information and requests for resources and reagents should be directed to the Lead Contact Juan Carlos Izpisua Belmonte (belmonte@salk.edu).

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Cells Used and Culture Conditions

Human fibroblasts from a male Wiskott-Aldrich syndrome patient with the c.1271dupG mutation (in the WAS gene) were purchased from the Coriell institute (Camden, NJ).

Human fibroblasts from a male Wiskott-Aldrich syndrome patient with the c.107_108del mutation and from the mother of the patient (unaffected) were purchased from the Cell Line and DNA Bank from Patients Affected by Genetic Diseases (Genova, Italy). All human fibroblasts were cultured at 37°C in DMEM containing GlutaMAX, non-essential amino acids, sodium pyruvate, and 15% fetal bovine serum (FBS). The B-lymphocyte

cell lines ID00003 (WAS: c.431G>A, p.E133K, clinical score 4–5) and GM11518 (wild type) were purchased from the Coriell institute (Camden, NJ). All lymphoblastoid cell lines were cultured at 37C in RPMI1640 containing 2mM L-glutamine and 15% FBS. H1 and H9 human ESCs were purchased from WiCell Research. Human ESCs and generated iPSC lines were cultured on Matrigel or mitotically inactivated MEF cells at 37C in DMEM/F12 (Invitrogen) supplemented with 0.1 mM non-essential amino acids (Invitrogen), 1mM GlutaMAX (Invitrogen), 20% Knockout Serum Replacement (Invitrogen), 55 mM beta-mercaptoethanol (Invitrogen) and 10 ng ml²¹ bFGF (Joint Protein Central) as described (Liu et al., 2012). Peripheral blood mononuclear cells from two WAS patients were obtained under approved IRB at IRCCS Fondazione Policlinico San Matteo. Human buffy coats were purchased from San Diego Blood Bank (San Diego, CA). Human umbilical cord blood units were gifts from StemCyte (Covina, CA) and San Diego Blood Bank (San Diego, CA). The cord blood units were anonymized and donated to research by individuals who gave informed consent. Cord blood CD34-positive cells were isolated by immune-selection using human CD34 magnetic microbeads and MiniMACS separator (Miltenyi Biotec, San Diego, CA) according to manufacturer's instructions. Purity of the cells was checked by staining with anti-CD34-PE antibody on a BD LSRFortessa cytometer. Cryopreserved CB CD34⁺ cells were FACS sorted again before being used in experiments.

METHOD DETAILS

iPSCs Generation

The episomal vector-based reprogramming was performed as previously described(Li et al., 2011; Liu et al., 2014). Briefly, fibroblasts were nucleofected with the episomal vectors using a Nucleofector II (Lonza) with program U20. Five days after nucleofection, fibroblasts were reseeded onto mitotically inactivated mouse embryonic fibroblasts (MEFs). The reprogrammed colonies were allowed to grow until they were ready to be mechanically picked and transferred onto MEFs. Upon successful expansion of the iPSC lines on MEFs, they were then cultured on Matrigel (BD Biosciences) in mTeSR medium (StemCell Technologies). Analyses of pluripotency of iPSCs, including detection of transgene insertion, transgene expression, teratoma formation, immunostaining of pluripotency markers, karyotyping were carried out as previously described(Li et al., 2011; Liu et al., 2014). The genotype of the WAS-iPSCs was periodically verified during the course of this study. The identity of the iPSCs was checked by Sanger sequencing and DNA fingerprinting.

Construction and Preparation of HDAdV

Construction and preparation of WAS-c-HDAdV for gene correction was generated using a BAC clone containing the human *WAS* locus (RP11-1148L6, BACPAC Resources) (Datsenko and Wanner, 2000). In brief, an *FRT*-PGK-EM7-neo-bpA-*FRT* fragment was recombined into intron 2 of *WRN* in the BAC clone. A total of 20.4 kb of *WAS* homology sequence and *FRT*-PGK-EM7-neo-bpA-*FRT* cassette were subcloned into the HDAdV plasmid pCIHDAdGT8-4 (kindly provided by Dr. Kohnosuke Mitani) (Suzuki et al 2008 PNAS). The WAS-c-HDAdV virus was generated according to a

previously described method (Palmer and Ng, 2004; Suzuki et al., 2008). β gal-transducing units (btu) were determined in 293 cells to define infectious vector titers.

Isolation of Gene-corrected Human iPSCs

For generation of gene-corrected iPSCs, $2.5\text{--}3.0 \times 10^6$ WAS patient iPSCs were infected with WAS-c-HDAdV at multiplicity of infection (MOI) of 3-30 btu/cell. Two days after infection, G418 (50 $\mu\text{g/ml}$; Invitrogen) was added to the medium to start positive selection. After 9 days, 2 μM Ganciclovir (GANC; Invitrogen) in addition to G418 was added to the medium to start negative selection. After an additional 8 days, G418/GANC double-resistant clones were transferred to 96-well plates and expanded for further characterization. Gene-targeted clones were determined by PCR of genomic DNA from drug-resistant clones with the following primers (P1, 5'-CATCCTGCCCCAGCCGACCAGACCTTAATGCTC-3'; P2, 5'-CCCCAAAGGCCTACCCGCTTCCATTGCTCA-3'; P3, 5'-CTACCTGCCCATTGACCAAGCGAAACATC-3'; P4, 5'-AAATTTTCCGCCATCTTTCCCCACGGCTAACGAC-3'; see Figure 1E) using PrimeSTAR GXL DNA Polymerase (TAKARA). To determine gene-corrected clones for exon1, exon 1 of WAS was PCR-amplified with the following primers: 5'-CCACCCAGGCCCATGACTACTCCTTGCCACA -3' and 5'-AGGCAACCATCCCGGCTGAGAGAATACCCA -3' with PrimeSTAR GXL DNA Polymerase. To determine gene-corrected clones for exon10, exon 10 of WAS was PCR-amplified with the following primers: 5'-ACTGCTTCAGTCAGGAGTTG-3' and 5'-GAGGCTGACACAAGATTCAT-3' with PrimeSTAR GXL DNA Polymerase. Amplicons

were sequenced using an ABI 3730 sequencer (Applied Biosystems). The inserted neomycin-resistance cassettes were removed according to a previously described method(Liu et al., 2012).

Plasmid Construction

The FLAG-tagged WASP, WASP^{F36*}, WASP^{ΔEX2}, and WASP^{L425fs} were cloned into the pLenti-CAG-PL9 (kind gift from the Drs. Inder Verma and Ronald Evans) using the The Gibson Assembly® HiFi 1-Step Kit (Synthetic Genomics, San Diego, CA). All the constructs generated were subjected to DNA sequencing to confirm accurate sequences.

Lentivirus Production

Lentiviruses were generated by co-transfecting the 293T cells (ATCC) with the lentiviral vectors together with the packaging plasmids (pMDLg/pRRE, pRSV-Rev, and pMD2.G, from Addgene, 12251, 12253, and 12259, respectively) into 293T cells using Lipofectamine 3000 (Invitrogen). Lentiviruses were collected 24 h after transfection, filtered through a 0.45-μm filter, and concentrated using the PEG-it virus precipitation solution (SBI).

Fluorescence in situ Hybridization

The WAS-c-HDAdV probe was custom labeled with 5-TAMRA by Empire Genomics (Buffalo, NY). 5-Fluorescein labeled chromosome 11 probe was purchased from Empire Genomics. Cells were fixed in 4% formaldehyde for 10 min at room temperature,

permeabilized with Triton X-100 and repeated freezing in liquid nitrogen, and deproteinized by treating with 0.1 M HCl. Hybridization was carried out at 37°C overnight. Samples were counterstained with DAPI, mounted and examined using a Zeiss LSM 780 Laser Scanning Confocal Microscope.

Hematopoietic Differentiation

Hematopoietic differentiation efficiency and hematopoietic colony formation activity were assayed as previously described (Vodyanik and Slukvin, 2007) with some modifications. iPSCs and ESCs cultured on MEF feeders in human ESC medium were manually dissociated and plated on OP9 feeders to start differentiation. Cells were co-cultured with the OP9 feeders for 12-14 days in OP9 differentiation medium (Alpha MEM containing 10% FBS, 100 µM monothioglycerol and 50 µg/ml ascorbic acid) as described (Vodyanik and Slukvin, 2007). Differentiated cells were dissociated with collagenase and Accutase (Innovative Cell Technologies) and subjected to flow cytometry analyses, cell sorting, qPCR analyses and clonogenic progenitor cell assays. Macrophage and dendritic cell differentiation were performed as described (Senju et al., 2011). Briefly, 2-week differentiated hiPSC were harvested and the non-adherent fraction (or FACS sorted $+/+^{iPSCs}$) were cultured under myeloid progenitor expansion condition for 10 days and then induced to maturity by lineage specific cytokines.

Characterization of iPSC Derived Cells

Characterization of the identity of iMPs and iDCs using lineage specific cell surface markers were performed as previously described (Choi et al., 2011). Macrophage

morphology was quantitated using the circularity plugin of ImageJ. Macrophage phagocytosis assay was performed with pHrodo™ Red *E. coli* BioParticles® Conjugate for Phagocytosis (Life Technologies, Carlsbad, CA). The pHrodo red particles were prepared according to the manual and opsonized *Escherichia coli* BioParticles® opsonizing Reagent (Life Technologies, Carlsbad, CA) as instructed. Adherent macrophages were incubated with pHrodo red *E. coli* particles for 1 hour at 37°C, dissociated, washed and analyzed by FACS. Migration assay was performed using the CytoSelect™ 24-Well Cell Migration Assay, 5 µm kit (Cell Biolabs, inc., San Diego, CA).

Clonogenic Progenitor Cell Assay

Hematopoietic clonogenic assays were performed in 35-mm low adherent plastic dishes (Stem Cell Technologies) in triplicate using 1.1 ml/dish of MethoCult GF + H4435 semisolid medium (Stem Cell Technologies) or StemMACS human HSC-CFU complete medium with Epo (Miltenyi Biotec, 130-091-280). Colony-forming cells (CFCs) were scored after 15 days of incubation. Cytospins were stained with the Diff stain kit (IMEB K7128) according to manufacturer's instructions.

qRT-PCR

Total RNA extraction and cDNA synthesis were performed with the RNeasy micro kit (Qiagen) and iScript™ Reverse Transcription Supermix for RT-qPCR (Biorad). Quantitative RT-PCR was carried out with SsoAdvanced™ Universal SYBR® Green Supermix (Biorad) on a CFX384 real-time PCR detection system (Biorad). Primer sequences are given in the **primer list**.

Immunofluorescence Analysis

Cells were fixed with 4% formaldehyde in PBS for 20min at room temperature, and then permeabilized with 0.4% Triton X-100. After a blocking step with 10% normal donkey serum in PBS, cells were incubated with the primary antibody at 4 °C overnight, followed by incubation at room temperature with the corresponding secondary antibody for 1 h. Nuclei were counterstained with 4',6-diamidino-2-phenylindole (DAPI) and mounted with Vectashield (Vector Laboratories). Images were acquired with a Zeiss LSM 780 laser-scanning inverted confocal microscope.

ELISA

IL1b, IL6 and TNF α ELISA assays were performed using Peprotech Mini ELISA development kit in conjunction with TMB substrate (KPL product code 52-00-02). Cells were treated with DMSO or 100ng/ml LPS. Cell culture supernatants were collected and assayed according to manufacturer instructions.

Cell Proliferation Assay

Cell proliferation assay were performed using the CellTiter 96 AQueous One cell proliferation assay kit (Promega, Madison, WI) according to manufacturer instructions. Macrophages were seeded onto 96-well plates at 10000 cells per well. B lymphocyte lines were seeded onto 96-well plates at 25000 cells per well. At each assay time point, assays are performed by adding 20 μ l of the CellTiter 96® AQueous One Solution Reagent directly to culture wells containing 100 μ l media, incubating for 1–4 hours and

then recording the absorbance at 490nm with a Bio-Tek Synergy H1 microplate reader. Background absorbance was corrected at 650 nm.

Immunoprecipitation and Immunoblot Analysis

The ID00003 and GM11518 cell lines described above or ID00003 cells expressing FLAG-tagged wild-type WASP were lysed with a buffer comprising 50 mM Tris-HCl (pH 7.5), 150 mM NaCl, 0.5% Triton X-100, aprotinin (10 µg/ml), leupeptin (10 µg/ml), 1 mM PMSF, 400 µM Na₃VO₄, 400 µM EDTA, 10 mM NaF, and 10 mM sodium pyrophosphate. The lysates were centrifuged at 20,000 × g for 10 min at 4°C to remove debris, and the resulting supernatants were then incubated for 30 min at 4°C with protein A–Sepharose 4 Fast Flow (P9424, Sigma) conjugated with antibodies as indicated. The resulting immunoprecipitates were washed three times with ice-cold lysis buffer, fractionated by SDS-PAGE, and subjected to immunoblot analysis with primary antibodies and HRP-conjugated secondary antibodies to either mouse (7076, Cell Signaling Technology), or rabbit (7074, Cell Signaling Technology) IgG. Immune complexes were detected with an ECL system (ThermoFisher Scientific).

Serial Block Face Scanning Electron Microscopy

miniSOG(Shu et al., 2011) (generously provided by R.Y. Tsien) was fused to the carboxyl terminus of WASP in a lentivirus expression vector. For a negative control, we used WASP fused to mEOS3, a fluorescence tag incapable of photo-oxidation of DAB. Macrophages derived from normal cord blood adhered to 35mm glass bottom grid-500 dishes (Ibidi, 81168) were fixed in 1% glutaraldehyde:1% paraformaldehyde in 0.15M

sodium cacodylate buffer, pH 7.4 containing 2mM CaCl_2 for 30 minutes at room temperature. miniSOG mediated photo-oxidation of diaminobenzidine (DAB) was done as previously described (Strickfaden et al., 2015). Samples were prepared for serial block face scanning electron microscopy as described in “NCMIR methods for 3D EM” (<http://www.gatan.com/products/sem-imaging-spectroscopy/3view-system#resources>). Briefly, they were then stained with 2% OsO_4 , 1.5% potassium ferrocyanide in 0.15M cacodylate buffer, pH 7.4 containing 2mM CaCl_2 for 1 hour at 4°C. Samples were washed with deionized water (ddH_2O), treated with thiocarbohydrazide for 20 minutes at room temperature, washed with ddH_2O and stained again in 2% OsO_4 for 1 hour at room temperature. Cell samples were counterstained with 1% uranyl acetate overnight at 4°C and subsequently washed with ddH_2O before incubation in *en bloc* Walton’s lead nitrate at 60 degrees for 30 minutes. Samples were serially dehydrated with cold ethanol (35%, 50%, 70%, 90%, 100%, 100%) for 5 minutes for each step at 4°C, followed by cold absolute ethanol for 10 minutes on ice, and an additional incubation in absolute ethanol for 10 minutes at room temperature. Samples were serially infiltrated into Durcupan resin (11.4g part A, 10g part B, 0.3g part C, 0.075g part D) over the course of 6 hours (25:75 resin:ethanol, 50:50 resin:ethanol, 75:25 resin:ethanol) and left in 100% resin overnight. The next day, samples were incubated again with fresh Durcupan resin for 2 hours. To reduce potential sample charging while imaging, an additional half-batch of Durcupan resin (5.7g part A, 5g part B, 0.15g part C, 0.0375g part D) was made that contained 10mg of Polyaniline emeraldine base powder (Alfa Aesar, 32038). The samples were cured in the aforementioned resin at 60°C for 48 hours. Glass from the 35mm dish was etched off with hydrofluoric acid (Sigma Aldrich,

30107). The resin embedded ROIs that were marked during the photo-oxidation steps were excised using a jeweler's saw, and then mounted (Electron Microscopy Sciences, 12642-14) onto a specimen pin (Gatan, 3VMRT12) using two-part conductive silver epoxy. The ROI was trimmed to yield a block face with dimensions $\sim 0.5 \times 0.5$ mm and a thin layer of platinum was evaporated onto the block face using the Leica EM SCD 500 prior to imaging. Serial block face scanning electron microscopy was performed on a Zeiss Sigma VP (Cambridge, UK) equipped with a Gatan 3View XP2 (Pleasanton, CA, USA). The SEM was operated at 2.5 KeV with a 30 μ m aperture in variable pressure mode set to 10 Pa pressure with a slice thickness of 50-70 nm. Images were taken at ~ 10 nm/pixel with either a 6 μ s dwell time using the 3View's integrated backscatter detector with image dimensions ranging between 2848 x 2848 pixels and 3538 x 3538 pixels. Sectioning and imaging continued until the entire cell(s) of interest was imaged.

Flow Cytometry

For cell-cycle analysis of HPCs, differentiated cells were stained with antibodies against indicated cell surface markers, fixed and permeabilized using the BD Cytofix/Cytoperm™ kit, stained with Alexa Fluor® 647 anti-human Ki-67 and DAPI, and then analyzed on a BD LSRFortessa cytometer. Cell sorting was performed by the Flow cytometry core facilities of the Salk Institute and UCSD on a FACS Aria II cytometer or a BD Influx cytometer.

Small Molecules

The following small molecules were used: Sodium butyrate (catalog number 04-0005, Stemgent, Lexington, MA), working concentration: 0.25 mM. All drugs were dissolved in DMSO and used at concentrations as indicated.

Proteomic Analyses

MudPIT We suspect that the sensitivity and specificity of MS analysis might be confounded by the close association between WASP and the ubiquitous cytoskeletal components. To improve confidence in specificity, we performed WAS co-IP in two independent B-lymphocyte lines, each with two well-characterized monoclonal antibodies. We then filtered out peptides identified in the IgG co-IP or in the WAS^{null} B cell line by either of the two antibodies. Seventy-eight high-confidence polypeptides were thus identified as WASP partners (Table S1). To improve the detection of nuclear factors, we repeated the MudPIT analysis with highly purified nuclei and identified 299 high-confidence WASP-interacting polypeptides. Proteins associated with endogenous WASP were co-immunoprecipitated with two different monoclonal antibodies (sc-13139, Santa Cruz Biotechnology; 557773, BD Biosciences) against WASP from wild-type B lymphoblastoid cell lines using the Active Motif Universal Magnetic Co-IP Kit. Co-IP using mouse IgG in wild type cells and co-IP using WASP antibodies in WASP mutant cells were used to discriminate against nonspecific proteins. To specifically immunoprecipitate WASP-interacting proteins in the nucleus, nuclei were isolated using the Nuclei EZ Prep kit (Sigma-Aldrich, St. Louis, MO) and processed as other co-IP experiments. The co-IP samples were subjected to MudPIT analysis as previously described (Liu et al., 2009; Liu et al., 2012)

RIME Assay

Chromatin Immunoprecipitation

Cells were fixed with 1% methanol-free formaldehyde for 8 min and quenched with 0.125 M glycine. Chromatin was isolated by the addition of lysis buffer, followed by disruption with a Dounce homogenizer. Lysates were sonicated and the DNA sheared to an average length of 300-500 bp. Genomic DNA (Input) was prepared by treating aliquots of chromatin with RNase, proteinase K and heat for de-crosslinking, followed by ethanol precipitation. Pellets were resuspended and the resulting DNA was quantified on a NanoDrop spectrophotometer. Extrapolation to the original chromatin volume allowed quantitation of the total chromatin yield.

An aliquot of chromatin (150 ug) was precleared with protein G agarose beads (Invitrogen). Proteins of interest were immunoprecipitated using 15 ug of antibody against WASP and protein G magnetic beads. Protein complexes were washed then trypsin was used to remove the immunoprecipitate from beads and digest the protein sample. Protein digests were separated from the beads and purified using a C18 spin column (Harvard Apparatus). The peptides were vacuum dried using a speedvac.

Mass Spectrometry

Digested peptides were analyzed by LC-MS/MS on a Thermo Scientific Q Exactive Orbitrap Mass spectrometer in conjunction with a Proxeon Easy-nLC II HPLC (Thermo Scientific) and Proxeon nanospray source. The digested peptides were loaded on a 100

micron x 25 mm Magic C18 100Å 5U reverse phase trap where they were desalted online before being separated using a 75 micron x 150 mm Magic C18 200Å 3U reverse phase column. Peptides were eluted using a 90 minute gradient with a flow rate of 300nl/min. An MS survey scan was obtained for the m/z range 300-1600, MS/MS spectra were acquired using a top 15 method, where the top 15 ions in the MS spectra were subjected to HCD (High Energy Collisional Dissociation). An isolation mass window of 1.6 m/z was for the precursor ion selection, and normalized collision energy of 27% was used for fragmentation. A five second duration was used for the dynamic exclusion.

Database Searching

DATABASE SEARCHING-- Tandem mass spectra were extracted by [unknown] version [unknown]. Charge state deconvolution and deisotoping were not performed. All MS/MS samples were analyzed using X! Tandem (The GPM, thegpm.org; version X! Tandem Alanine (2017.2.1.4)). X! Tandem was set up to search the Activ_motiv_human database (unknown version, 143818 entries) assuming the digestion enzyme trypsin and the decoy.uniprot-proteome-3AUP000005640-absUnique database (unknown version, 143206 entries) the cRAP database of common laboratory contaminants (www.thegpm.org/crap; 114 entries) plus an equal number of reverse protein sequences also assuming the digestion of enzyme trypsin. X! Tandem was searched with a fragment ion mass tolerance of 20 PPM and a parent ion tolerance of 20 PPM. Carbamidomethyl of cysteine was specified in X! Tandem as a fixed modification. Glu->pyro-Glu of the n-terminus, ammonia-loss of the n-terminus, deamidated of asparagine

and glutamine, oxidation of methionine and tryptophan and dioxidation of methionine and tryptophan were specified in X! Tandem as variable modifications.

Criteria for Protein Identification

Scaffold (version Scaffold_4.8.4, Proteome Software Inc., Portland, OR) was used to validate MS/MS based peptide and protein identifications. Peptide identifications were accepted if they exceeded specific database search engine thresholds. Peptide identifications were accepted if they could be established at greater than 95.0% probability by the Scaffold Local FDR algorithm. Protein identifications were accepted if they could be established at greater than 99.0% probability and contained at least 2 identified peptides. Protein probabilities were assigned by the Protein Prophet algorithm (Nesvizhskii, Al et al Anal. Chem. 2003;75(17):4646-58). Proteins that contained similar peptides and could not be differentiated based on MS/MS analysis alone were grouped to satisfy the principles of parsimony. Proteins sharing significant peptide evidence were grouped into clusters.

List Filtering

Protein coverage percentage is observed in the raw data file. Final list generation was done by taking all proteins with a spectral count of five and above from each replicate reaction and comparing them in a Venn-diagram against IgG control replicates. Proteins unique to both experimental replicates were then applied to the PANTHER database for protein ontology results.

RASL-seq Profiling of Alternative Splicing

RASL-seq analysis was carried out as previously described (Li et al., 2012; Pandit et al., 2013). A pool of oligonucleotides was designed to detect 5530 alternative splicing events conserved between humans and mice. RASL-seq primers were annealed to total RNA isolated from WAS mutant and wild-type cells. After ligation, 5 µl eluted ligated oligos was used for PCR amplification and the bar-coded PCR products were sequenced on HiSeq2000. Data analysis was as described (Zhou et al., 2012).

High-throughput Whole Transcriptome Sequencing

RNA was isolated by Qiagen RNeasy kit with in column DNase digestion. The Agilent TapeStation or Bioanalyzer was used to determine RNA integrity (RIN) numbers prior to library preparation. Stranded mRNA-Seq libraries were prepared using the TruSeq Stranded mRNA Library Prep Kit or TruSeq RNA Access Library Prep Kit according to the manufacturer's instructions (Illumina). Briefly, RNA with polyA tail was isolated using poly-T oligos conjugated to magnetic beads. mRNA was then fragmented and reverse-transcribed into cDNA. dUTPs were incorporated, followed by second strand cDNA synthesis. dUTP-incorporated second strand was not amplified. cDNA was then end-repaired, index adapter-ligated and PCR amplified. Purification of nucleic acids after each enzymatic steps was achieved by using AMPure XP beads (Beckman Coulter).

Next-gen Sequencing

mRNA Libraries were quantified, pooled and sequenced at single-end 50 base-pair (bp) using the Illumina HiSeq 2500 or paired-end 75 bp using the Illumina NextSeq 500

platform at the Salk Next Generation Sequencing Core and the KAUST Bioscience Core Lab. Raw sequencing data was demultiplexed and converted into FASTQ files using CASAVA (v1.8.2). Libraries were sequenced at an average depth of 50 million reads per library.

Bioinformatics Analysis

Sequencing reads were aligned to the hg19 human genome using STAR (v2.5.1) (<https://academic.oup.com/bioinformatics/article/29/1/15/272537>) with default parameters. Cuffdiff (v2.2.1) (<https://www.nature.com/articles/nbt.2450>) was used to analyze the differentially expressed (DE) genes and isoforms (-b -u --library-type fr-firststrand) on a pair-wise experiment design. A default cutoff of q-value < 0.05 was used to identify the DE genes. The spliced ratio was defined as the expression of one isoform over all isoforms of a specific gene under one condition. Isoforms with a difference in spliced ratio > 20% between two conditions and FPKM (Fragments Per Kilobase of transcript per Million mapped reads) > 5 in both conditions were identified as differentially spliced (DS) isoforms. Pathway analysis was performed using DAVID (v6.8) (Nature Protocols 2009; 4(1):44 & Nucleic Acids Res. 2009;37(1):1) where the list of DE genes/DS isoforms was used as input and the whole genome was used as background. As an alternative isoform expression analysis method, we used STAR v2.5.3a to build a genome index from ENSEMBL GRCh38 human primary assembly. We augmented the index with the ENSEMBL annotation file release 90 through the --sjdbGTFfile command line parameter. The paired-end reads were then aligned using STAR v2.5.3a to the previously built human genome. Afterwards, we used the QC command of QoRT

(<https://www.ncbi.nlm.nih.gov/pubmed/26187896>) to generate counts for novel and annotated splice junctions based on the alignment files using the arguments (--stranded --runFunctions writeKnownSplices writeNovelSplices writeSpliceExon). Next, we used the mergeNovelSplices command to aggregate the splice junctions and filter out low coverage (<10 normalized read coverage) novel splice junctions resulting in flat exon and splice junction count files for each sample. We used the previously generated flat count files as input for JunctionSeq (<https://www.ncbi.nlm.nih.gov/pubmed/27257077>) to determine exon and splice junction differential usage. We followed the user manual to obtain the differentially used junctions and exons. Using the previously generated alignment files, we obtained gene-level counts through htseq-count v0.6.0, part of the HTSeq framework (<https://www.ncbi.nlm.nih.gov/pubmed/25260700>), based on ENSEMBL annotation file release 90 (<https://www.ncbi.nlm.nih.gov/pubmed/27899575>). We then used DESeq2 (<https://www.ncbi.nlm.nih.gov/pubmed/25516281>) to obtain the differentially expressed genes following the standard workflow. The splicing pattern was visualized using IGV (Nature Biotechnology 29, 24–26 (2011)). Gene expression cluster analysis was conducted using Cluster 3.0 and Java TreeView. Dendrogram on the heatmap represented the hierarchical clustering of the gene expression based on their correlation using the average linkage method in Cluster 3.0. GO term enrichment analysis was performed using DAVID Bioinformatics Resources 6.8 and visualized with Cytoscape (<http://www.cytoscape.org>). Ingenuity pathway analysis of the disease and function categories affected using RNA-seq data was performed according to manufacturer instructions.

Skipped exon events were detected using rMATS (v4.0.1; doi: 10.1073/pnas.1419161111) by evaluating reads that were mapped to the splicing junctions and the “on target” exons. Each condition was tested using two replicates of paired-end 75 bp reads. To get the reads with fixed length, reads were pre-processed by cutadapt (v1.14; DOI:10.14806/ej.17.1.200) to trim the first nucleotide (nt) off whenever necessary and aligned by STAR (v2.5.1) with default parameters and the soft-clipping option off. The resulted skipped exon events with > 10% difference in splicing ratio and FDR < 0.05 were identified as significantly alternatively spliced (AS) exons.

Motif enrichment tests on the exons were performed using Homer (v4.8; PMID: 20513432). The WAS promoted exons were tested for *de novo* enriched motifs on the same strand against the WAS repressed exons as background, and the reverse analysis was repeated on the WAS repressed exons using the promoted exon sequences as background. The motif lengths from 4 nt to 6 nt were tested for enrichment with respect to background. Significantly enriched motifs were defined as motifs with hypergeometric $p < 1e-3$, greater than 5% enrichment in target sequences, and fold-enrichment > 1.5.

R (<http://www.R-project.org/>) was used to generate graphs, unless mentioned specifically. Venn diagrams were plotted using the R package VennDiagram (<https://doi.org/10.1186/1471-2105-12-35>).

Statistical Analysis

Results are shown as mean $\pm$ sem unless indicated otherwise. Comparisons were performed with student's *t*-test.

Data Availability

The raw sequencing data and the processed gene expression tables were submitted to the NCBI GEO database with the accession number GSE107963.

Fig. 1

Fig. 2

Fig. 3

A

B

C

D

E

F

G

Fig. 4

A

B

C

D

F

E

G

H

Fig. 5

Fig. 6

SUPPLEMENTARY MATERIALS

SUPPLEMENTAL Fig. S1-S8

Table S1. List of Proteins Identified by MudPIT and Their Relationship with Nuclear Compartments

Table S2. RASL-seq Read Count Ratios of Macrophages

Table S3. List of Known Splicing Factor Binding Sites in the SpliceAid-F Database that Overlap with Motifs in Figure S7.

Table S4. List of Oligonucleotides

Figure S1. Generation of WAS-specific iPSCs and isogenic gene-corrected iPSCs, related to Figure 1. **A**, alkaline phosphatase (AP) and NANOG immunostaining of colonies of reprogrammed WAS patient fibroblasts. Scale bar, 0.5 mm. **B**, quantitative RT-PCR (qRT-PCR) analysis of the expression of endogenous pluripotency transcription factors (Oct4, Sox2, Klf4, L-Myc and Lin28) or transgenes (Oct4-tg, Sox2-tg, Klf4-tg, L-Myc-tg and Lin28-tg) in the two patient WAS-iPSCs. Mean $\pm$ SEM, n=3. **C**, quantification of the copy number of episomal vectors (EBNA1) and the reprogramming factor transgenes that may integrate in the genome. An iPSC line generated with integrative retroviruses were included as a control. Mean $\pm$ SEM, n=3. **D**, immunofluorescence analysis of pluripotency markers OCT4, Tra-1-60, NANOG and SOX2 in WAS-iPSC. DNA was stained with DAPI. Bar, 100 μ m. **E**, schematic molecular representation of WAS gene-correction with gene-correction vector (WAS-c-HDAV). The primers for PCR verification of gene targeting are shown as arrows (P1, P2, P3 and P4). HSV tk stands for herpes simplex virus thymidine kinase gene cassette used for negative selection; neo stands for neomycin-resistant gene cassette used for positive selection; CMV- β gal: β -gal expression cassette for determination of HDAV titer; Red crosses mark the mutations c.107_108del in exon1 and c.1271dupG in exon10 present in Patient1 and Patient2, respectively. **F**, DNA Sequencing of the mutation sites in WAS-iPSCs and cWAS-iPSCs. **G**, gene-targeting and gene-correction efficiencies at the WAS locus in WAS-iPSCs achieved in multiple clones. **H**, Verification of patient specific mutations by DNA sequencing in WAS-iPSC clones. Wild-type sequence is shown on top of the chromatographs. **I**, Karyotyping analysis revealed normal karyotypes in randomly selected clones of WAS-iPSC lines, either before or after gene

correction. **J**, Hematoxylin and eosin (H&E) histological staining and immunostaining in teratomas derived from a randomly selected gene-corrected WAS-iPSC clone show *in vivo* differentiation towards ectodermal, mesodermal and endodermal tissues. **K**, fluorescence in situ hybridization analysis of random integration of WAS-c-HDAdV. Interphase iPSC nuclei were hybridized with a TAMRA-labeled probe specific for WAS-c-HDAdV. Gene-corrected WAS-iPSCs showed one hybridization signal (red puncta) per nucleus. A chromosome 11 probe (green puncta) was used as control. Scale bar: 10 μ m. Blue: DAPI.

Figure S2. Differentiation and characterization of WAS-iPSC-derived

hematopoietic cells, related to Figure 1 and 2. A, qRT-PCR analysis showing

upregulation of endodermal (GATA4 and GATA6), mesodermal (CTnT and MSX1), and ectodermal (FGF5) markers in embryoid bodies generated from H9 ESC, WAS-iPSC and cWAS-iPSC. Data are shown as mean $\pm$ SEM. **B,** qRT-PCR analysis of the upregulation

of the *WAS* gene in embryoid bodies generated from H9 ESCs, WAS-iPSCs and cWAS-iPSCs. Data are shown as mean $\pm$ SEM, n=3. **C,** qRT-PCR analysis of the expression of

the *WAS* gene in ESCs and iPSCs following hematopoietic differentiation. Expression of *NWASP*, a constitutively expressed WAS family gene is shown as a control. Error bars: SEM, n=3. **D,** qRT-PCR analysis of the kinetics of the upregulation of hematopoietic

lineage genes during differentiation of iPSCs. CB: cord blood mononuclear cells. Values shown are normalized to GAPDH (mean $\pm$ SEM, n=3). **E,** representative FACS analysis

of HPC populations in differentiated iPSCs and ESCs. Cells are in the Tra-1-85⁺ gate. Numbers represent percentages. **F,** images of typical WAS-iPSC-derived CFU-GM,

BFU-E, and CFU-GEMM. Bar=0.1mm. **G,** quantitation of knockdown efficiency of three shRNAs against WASP. Lentiviruses expressing shRNA against WASP or a scramble shRNA were introduced to 293XT cells expressing a WASP-mEOS fusion protein. The level of knockdown of the shRNAs was quantitated by FACS using fluorescence of

mEOS as a proxy. shRNA1 was selected for further experiments. **H,** representative FACS analysis of HPC populations in differentiated iPSCs treated with indicated

shRNAs. Cells shown are in the Tra-1-85⁺ gate. Numbers represent percentage. The bar graph shows quantitation of the percentage of CD34⁺/CD43⁺ cells. ***, p<0.001 (*t*-

test). **I,** FACS analysis of iMPs. Histograms show expression of typical macrophage cell

surface markers (solid colored lines). Grey lines: isotype controls. **J**, left: expression values of CD43 measured in RPKM (Reads Per Kilobase of transcript per Million mapped reads) in macrophages of indicated genotypes; right: qRT-PCR analysis of the expression of the CD43 gene in B cells of indicated genotype. The expression of the WAS^{F36*} mutant in WAS^{null} B cells caused a further 44-fold reduction in CD43 expression, suggesting that WAS^{F36*} has a gain of function phenotype. Data are shown as mean $\pm$ SEM, n=3. **K**, FACS analysis of dendritic cells differentiated from iPSCs (iDCs). Histograms show expression of typical dendritic cell surface markers (solid colored lines). Grey lines: isotype controls. **L**, quantitative measurement of the migration of iDCs and iMPs toward CCL3. RFU: relative fluorescence unit.

Figure S3. Localization of wild type WASP in hematopoietic lineages, related to Figure 2. **A**, FACS analysis of WASP expression by intracellular staining in indicated hematopoietic lineages from peripheral blood of a normal individual. **B**, immunofluorescence analysis of the nuclear and cytoplasmic localization of WASP in wild-type macrophages. Lamin B1 and phalloidin were used as nuclear and cytoplasmic landmarks, respectively. DNA is stained with DAPI. Bar=10 μ m. **C**, Immunofluorescence analysis of the nuclear and cytoplasmic localization of WASP in wild-type hematopoietic stem/progenitor cells (Lin⁻CD34⁺CD38⁻CD49f⁺), B cells (CD20⁺), T cells (CD3⁺) and dendritic cells (CD11c⁺ DC-SIGN⁺). Green: WASP; magenta: LMNB; Blue: DNA. Bar=5 μ m. **D, E**, SBF-SEM images of wild-type macrophages expressing a WASP-miniSOG fusion protein. **D**, sub-nuclear localization of WASP was visualized with photo-oxidation of DAB by miniSOG (see methods). White arrowheads indicate WASP signals (high-contrast dark spots) at the nuclear envelop, nuclear periphery and nuclear interior regions. **E**, SBF-SEM images of wild-type macrophages expressing a WASP-mEOS fusion protein, which is incapable of photo-oxidation of DAB and serves as a negative control for the specificity of the positive signals in Fig. 3F and **D**. Bar=1 μ m.

Figure S4. WASP is involved in the organization of nuclear bodies, related to

Figure 2 and 3. **A**, confocal microscopy image showing colocalization of WASP and lamin B at the nuclear envelop in peripheral blood monocytes. Bar=5 μ m. **B**, confocal immunofluorescence analysis of the nuclear paraspeckle marker NONO in the wild-type control iMPs. Nuclei were stained with DAPI. Bar=5 μ m. **C**, confocal immunofluorescence analysis of the nuclear body markers and the histone modification H3K9me3. Overexpression of the *WASP*^{F36*} in *WASP*^{null} B cells led to an increase in total nucleolar area—composed of numerous dimly stained nucleoli—compared to the control. Nuclei were stained with DAPI. Bar=5 μ m. **D-F**, representative TEM images of *WASP*^{null} B cell lines overexpressing wild-type WASP (**D**), *WASP*^{F36*} (**E**) and *WASP*^{L425fs} (**F**). Bar=1 μ m in the left and upper right panels; Bar=0.5 μ m in lower right panels and insets. NO: nucleolus, HC: heterochromatin, IGC: interchromatin granule cluster, NP, nuclear pore, M: mitochondria, A: autophagosome, G: Golgi.

Figure S5. A, Western blot analysis of the purity of isolated nuclei using Lysosomal-associated membrane protein 1 (LAMP1) as a marker of cytoplasm and RNA polymerase II (RNAP2) as a nuclear marker. **B,** quantitation of known RNA-binding proteins in the WASP interactome. *** $P < 0.0001$, χ^2 test. **C,** examples of the sequence and spectral count of peptides in WASP binding partners identified by mass spectrometry. Detected peptides are highlighted in yellow.

Figure S6. WASP mutations disrupt normal gene expression and alternative splicing programs, related to Figure 4 and 5. **A**, qRT-PCR of selected differentially expressed genes identified in RNA-seq. $n=3$. *** $p<0.001$, Student's t-test. **B**, top cellular component and biological process GO terms that are enriched among common differentially expressed genes in both WAS-iMPs visualized as a network. Unsupervised hierarchical clustering of samples based on gene expression of genes in indicated GO categories is shown as heatmap. **C**, Ingenuity pathway analysis (IPA) of enriched disease and function annotations in the WAS-iMP RNA-seq data. The x-axis indicates statistical significance of the enrichment of the disease and function category (larger is more significant). The y-axis indicates IPA activation Z-score, which predicts the activation status of the disease/function process based on observed expression values and knowledge in the literature. Higher positive Z-scores are more consistent with activation, while lower negative Z-scores are more consistent with inhibition. **D**, unsupervised hierarchical clustering of short to long isoform ratio of the 513 alternative splicing events that passed the quality filter (Table S2) in WAS-iMPs, cWAS-iMPs, CB-MPs and control-iMPs. The data showed that WAS-iMPs differed significantly from macrophages derived from cord blood or wild type iPSCs, and that gene correction restored the normal splicing pattern. Color scale represents $\log_2$ (fold change). **E**, Venn diagrams showing shared and distinct genes that are downregulated (DE down) in the three cell types. FDR <0.05 . **F**, Venn diagram shows shared and distinct differentially expressed isoforms in the three cell types. **G**, heatmap shows GO enrichment of differentially expressed isoforms in the three cell types. **H**, GO enrichment of genes with differentially used junctions detected by Junction-seq between T cells from WAS

patients and normal donors. I, Venn diagram showing the number of exons that are repressed in B and T cells from WAS patients compared to those from normal individuals.

Figure S7. Motif analysis of promoted vs. repressed exons in WAS. Logograms of most enriched motif in WAS promoted exons using WAS repressed exons as background. The cutoff criteria were 1) hypergeometric $p < 10^{-3}$; 2) fold enrichment > 1.5 ; 3) target% $> 5\%$.

Figure S8. WASP deficiency promotes deregulation of cell proliferation, related to

Figure 6. A, FACS analysis of Ki67 expression by intracellular staining in WAS knockdown wild type iMPs. Cells shown in the analysis were in the CD11b⁺/CD14⁺ gate. **B**, cell proliferation curve of WAS-iMPs and cWAS-iMPs. Data are shown as mean $\pm$ SEM. n $\geq$ 3. **C**, Western blot analysis of the expression of WASP and EZH2 in WAS-iMPs and cWAS-iMPs. HSP90 was used as a loading control. IB, immunoblotting. **D**, qRT-PCR of CTTN isoforms. Values are normalized against wild-type B cells. n=3. **E**, expression values of FOXM1 measured in RPKM (Reads Per Kilobase of transcript per Million mapped reads) in indicated samples. **F**, phase contrast images of B cell lines of indicated genotypes. Note that the wild-type, *Was*^{null}, and *Was*^{null}+WASP cDNA cells grow as suspension cells, while the *Was*^{null}+WASP mutant cDNAs cells change their morphology and grow as adherent cells. Bar=50 μ m. **G**, FACS analysis of cell cycle of mutant B cells. The percentage of cells in the G1, S and G2 phase is shown.

Fig. S1

Fig. S2

Fig. S3

Fig. S4

Fig. S5

A

B

C

WIPF1

WIPF1_HUMAN (100%), 51,260.2 Da
WAS/WASL-interacting protein family member 1 OS=Homo sapiens GN=WIPF1 PE=1 SV=3
8 exclusive unique peptides, 8 exclusive unique spectra, 16 total spectra, 117/503 amino acids (23% coverage)

MPVPPPPAPP	PPPTFALANT	EKPTLNKTEO	AGRNALLSDI	SKGKKLKKT V
TNDRSAPI LD	KPKGAGAGGG	GGGFGGGGGF	GGGGGGGGGG	SFGGGGPPGL
GGLFOAGMPK	LRSTANRDND	SGGSRPPLL P	PGGRSTSAKP	FSPPSGPGRF
PVPSPGHRSG	PPEPORNRMP	PPRPDVGSKP	DSIPPPVPST	PRPIOSSPHN
RGSPVPVGGP	ROPSPGPTPP	PFPGNRGTAL	GGGSIROSPL	SSSSPFSNRP
PLPPTPSRAL	DDKPPPPPPP	VGNRPSTHRE	AVPPPPPPONN	KPPVPSTPRP
SASSQAAPPPP	PPPSRPGPPP	LPPSSSGNDE	TPRLPORNLS	LSSSTPPLPS
PGRSGPLPPP	PSERPPPPVR	DPPGRSGPLP	PPPPVSRNGS	TSRALPATPO
LPSRSGVDS P	RSGRPPLLP	DRPSAGAPPP	PPPSTSI RNG	FODSPCEDEW
ESRFFYFHPIS	DLPPPEPYVQ	TTKSYPSKLA	RNESRSGSNR	REGAPPLPP
IPR				

HNRNPA2B1

ROA2_HUMAN (100%), 37,430.3 Da
Heterogeneous nuclear ribonucleoproteins A2/B1 OS=Homo sapiens GN=HNRNPA2B1 PE=1 SV=2
13 exclusive unique peptides, 15 exclusive unique spectra, 22 total spectra, 160/353 amino acids (45% coverage)

MEKTLETVPL	ERKKREKEOF	RKLFIGGLSF	ETTEESLRNY	YEQWGKLTDC
VVMRDPASKR	SRGFGFVTFS	SMAEVDAA MA	ARPHSIDGRV	VEPKRAVARE
ESGKPGAHV T	VKKLFVGGIK	EDTEEHLRD	YFEEYGKIDT	IEIITDROSG
KKRGGFGFVT F	DDHDPVDKIV	LOKYHTINGH	NAEVRKALSR	OEMOEVOSSR
SGRGGNFGGF G	DSRGGGGNFG	PGPGSNFRGG	SDGYGSGRGF	GDGYNGYGGG
PGGGNFGGSP	GYGGGRGGYG	GGGPGYGNQG	GGYGGGYDNY	GGGNYGSGNY
NDFGNYNQQP	SNYGPMKSGN	FGGSRNMGGP	YGGGNYGPGG	SGGSGGYGGR
SRY				

MEF2D

MEF2D_HUMAN (100%), 55,938.3 Da
Myocyte-specific enhancer factor 2D OS=Homo sapiens GN=MEF2D PE=1 SV=1
3 exclusive unique peptides, 3 exclusive unique spectra, 8 total spectra, 58/521 amino acids (11% coverage)

MGRKKIQIOR	ITDERNROVT	FTKRKEGLMK	KAYELSVLCD	CEIALIIFNH
SNKLFOYAST	DMDKVLLKYT	EYNEPHESRT	NADI IETLRK	KGFNGCDSPE
PDGEDSLEOS	PLLEDKYRRA	SEELDGLFRR	YGSTVPAPNF	AMPVTVPVSN
OSSLQFSNPS	GSLVTPSLVT	SSLTDPRLLS	POOPALORNS	VSPGLPQORPA
SAGAMLGGDL	NSANGACPS P	VGNGYVSARA	SPGLLPVANG	NSLNKVI PAK
SPPPPPTHSTO	LGAPSRKPD L	RVITSOAGKG	LMHHLTEDHL	DLNNAORLGV
SOSTHSLTTP	VVS VATPSLL	SOGLPFSSMP	TAYNTDYOLT	SAELSSLP AF
SSPGGLSLGN	VTAWOOPOP	OQPOOPOPPO	OQPOPOPOPO	POQPOQPOQ
POOQSHLVPV	SLSNLI PGSP	LPHVGAALTV	TTHPHISIKS	EPVSPSRERS
PAPPPPAVFP	AARPEPGDGL	SSPAGGSYET	GDRDDGRGDF	GPTLGLLRPA
PEPEAEGS AV	KRMRLDTWTL	K		

TMPO

GSE972_HUMAN (100%), 46,307.5 Da
Lamina-associated polypeptide 2, isoforms beta/gamma OS=Homo sapiens GN=TMPO PE=1 SV=1
2 exclusive unique peptides, 2 exclusive unique spectra, 3 total spectra, 43/414 amino acids (10% coverage)

MPEFLEDPSV	LTKDKLKSEL	VANNVTLPAG	EORKDVYVOL	YLOHLTARNR
PPLPAGTNSK	GPPDFSSDEE	REPTPVLGSG	AAAAGRSRAA	VGRKATKKTD
KPROEDKDDL	DVTELTNEDL	LDOLVKYGVN	PGPIVGTTRK	LYEKKLLKLR
EOGTESRSST	PLPTISSSAE	NTRONGSND S	DRYSDNEEDS	KIELKLEKRE
PLKGRAKTPV	TLKORRVEHN	OVETSEHFR I	DGPVISESTP	IAETIMASSN
ESLVVNRVTG	NFKHASPILP	ITEFSDIPRR	APKKPLTRAE	VGEKTEERRV
ERDILKEMFP	YEASTPTGIS	ASCRRPIKGA	AGRPLELSDF	RMEESFSSKY
VPKYVPLADV	KSEKTKKGRS	IPVWIKILLF	VVAVVFLFLV	YQAMETNQVN
PFSNFLHVDP	RKSN			

Fig. S6

Fig. S7

WAS-iMPs vs cWAS-iMPs

Rank	Motif	Hypergeometric p-value	Target %	Background %	Fold enrichment
1		1e-3	10.42	4.07	2.56
2		1e-3	10.42	4.02	2.59
3		1e-3	8.85	3.31	2.67
4		1e-3	9.38	3.64	2.58
5		1e-3	13.28	6.55	2.03

WAS B cells vs wild-type B cells

Rank	Motif	Hypergeometric p-value	Target %	Background %	Fold enrichment
1		1e-8	5.42	1.34	4.04
2		1e-5	18.46	12.27	1.50
3		1e-3	10.98	6.59	1.67

WAS T cells vs wild-type T cells

Rank	Motif	Hypergeometric p-value	Target %	Background %	Fold enrichment
1		1e-6	11.19	6.35	1.76
2		1e-4	9.39	5.66	1.66
3		1e-4	9.78	6.11	1.60

Fig. S8

Table S3 Oligonucleotides

Oligonucleotides		
h-EZH2-left: CTGATTTTACACGCTTCCGC	This paper	N/A
h-EZH2-right: GGAACAACGCGAGTCGG	This paper	N/A
h-TNF-left: AGATGATCTGACTGCCTGGG	This paper	N/A
h-TNF-right: CTGCTGCACTTTGGAGTGAT	This paper	N/A
h-KRAS-left: TGACCTGCTGTGTCGAGAAT	This paper	N/A
h-KRAS-right: TTGTGGACGAATATGATCCAA	This paper	N/A
h-CDKN2A left: GTGAGAGTGGCGGGGTC	This paper	N/A
h-CDKN2A right: GTTACGGTCGGAGGCCG	This paper	N/A
H-SFPQ right: CCCATTCCTCTAGGACCCTG	This paper	N/A
h-SFPQ left: GTGGTTCCATGATGGGAAGT	This paper	N/A
h-HNRNPA3 left: CAACTGCTCTGGTTCCTTTG	This paper	N/A
h-HNRNPA3 right: ATGGAGGTAAAACCGCCG	This paper	N/A
h-GATA2-F: GGCCCACTCTCTGTGTACC	This paper	N/A
h-GATA2-R: CATCTTCATGCTCTCCGTCAG	This paper	N/A
h-RUNX1-F: TCGAAGACATCGGCAGAACT	This paper	N/A
h-RUNX1-R: GAGGCTGAGGGTTAAAGGCA	This paper	N/A
h-GAPDH-F: GCACCGTCAAGGCTGAGAAC	This paper	N/A
h-GAPDH-R: AGGGATCTCGCTCCTGGAA	This paper	N/A
h-HoxB4-F: GTGAGCACGGTAAACCCCAAT	This paper	N/A
h-HoxB4-R: CGAGCGGATCTTGGTGTTG	This paper	N/A
h-CD45-F: ACAGCCAGCACCTTTCCTAC	This paper	N/A
h-CD45-R: GTGCAGGTAAGGCAGCAGA	This paper	N/A
h-CD34-F: GCGCTTTGCTTGCTGAGTTT	This paper	N/A
h-CD34-R: GCCATGTTGAGACACAGGGT	This paper	N/A
h-CD31-F: AACAGTGTTGACATGAAGAGCC	This paper	N/A

h-CD31-R: TGATAAACAGCACGTCATCCTT	This paper	N/A
----------------------------------	------------	-----