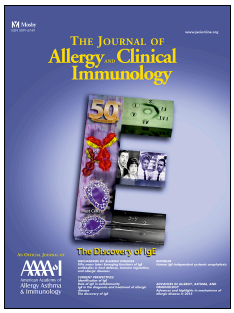
Accepted Manuscript



R-loops cause genomic instability in Wiskott-Aldrich syndrome T_{helper} lymphocytes

Koustav Sarkar, PhD, Seong-Su Han, PhD, Kuo-Kwang Wen, PhD, Hans D. Ochs, MD, Loïc Dupré, PhD, Michael M. Seidman, PhD, Yatin M. Vyas, MD

PII: S0091-6749(17)31899-7

DOI: 10.1016/j.jaci.2017.11.023

Reference: YMAI 13165

To appear in: *Journal of Allergy and Clinical Immunology*

Received Date: 4 July 2017

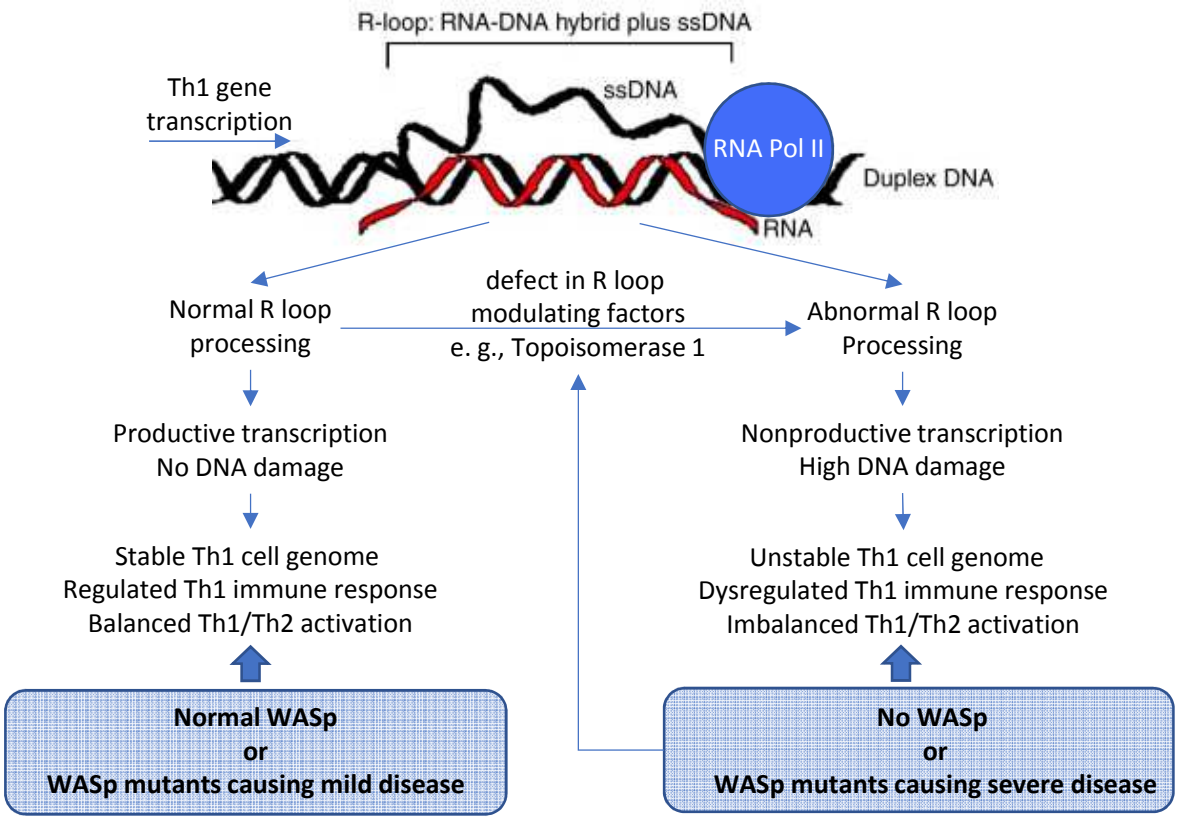
Revised Date: 7 November 2017

Accepted Date: 10 November 2017

Please cite this article as: Sarkar K, Han S-S, Wen K-K, Ochs HD, Dupré L, Seidman MM, Vyas YM, R-loops cause genomic instability in Wiskott-Aldrich syndrome T_{helper} lymphocytes, *Journal of Allergy and Clinical Immunology* (2018), doi: 10.1016/j.jaci.2017.11.023.

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

WASp caught in R loops



WASp: Wiskott-Aldrich syndrome protein

RNA-DNA image adapted from **Ivančić-Baće I, Howard JA, Bolt EL. Tuning in to Interference: R-Loops and Cascade Complexes in CRISPR Immunity. J Mol Biol. 2012 Oct 5;422(5):607-16**

ACCEPTED MANUSCRIPT

R-loops cause genomic instability in Wiskott-Aldrich syndrome T_{helper} lymphocytes

Authors: Koustav Sarkar, PhD[1*§], Seong-Su Han, PhD[1*], Kuo-Kwang Wen, PhD[1*], Hans D. Ochs, MD[2], Loïc Dupré, PhD[3,4,5,6,7], Michael M. Seidman, PhD[8], and Yatin M. Vyas, MD[1]

Affiliations: [1]Division of Pediatric Hematology-Oncology, Carver College of Medicine and the Stead Family University of Iowa Children's Hospital, Iowa City, IA 52242. [2]Center for Immunity and Immunotherapies, Seattle Children's Research Institute, University of Washington, [3]INSERM, UMR1043, Centre de Physiopathologie de Toulouse Purpan, Toulouse, France, [4]Université Toulouse III Paul-Sabatier, Toulouse, France, [5]CNRS, UMR5282, Toulouse, France, [6]Ludwig Boltzmann Institute for Rare and Undiagnosed Diseases, Vienna, Austria, [7]CeMM Research Center for Molecular Medicine of the Austrian Academy of Sciences, Vienna, Austria, [8]Laboratory of Molecular Gerontology, National Institute on Aging, NIH, NIH Biomedical Research Center, 251 Bayview Boulevard, Baltimore, MD 21224, USA.

[*]KS, SSH, and KKW contributed equally.

Corresponding author: Yatin M. Vyas, Division of Pediatric Hematology-Oncology, The Stead Family University of Iowa Children's Hospital, The University of Iowa Carver College of Medicine, Iowa City, IA 52242. E-mail: yatin-vyas@uiowa.edu. Phone: (319) 353 8105.

[§]**Present address**: Department of Biotechnology, SRM University and Research Institute, Kattankulathur, Chennai, Tamil Nadu - 603203, India.

Funding: This work was supported by the NIH grant R01AI084957 (to Y.M.V.), the Intramural Research Program of the NIH, National Institute on Aging grant Z01AG000746-08 (to M.M.S.), and the DeJoria Wiskott-Aldrich Research Fund (to H.D.O).

ACCEPTED MANUSCRIPT

Conflict of interest: All authors declare no conflict of interest with the current study

ACCEPTED MANUSCRIPT

Background: Wiskott-Aldrich syndrome (WAS), X-linked thrombocytopenia (XLT), and X-linked neutropenia (XLN), caused by *WAS* mutations affecting WASp expression or activity, manifest in immunodeficiency, autoimmunity, genomic-instability, and lymphoid-cancer. WASp supports filamentous-actin formation in the cytoplasm and gene-transcription in the nucleus. Although the genetic basis for XLT/WAS has been clarified, the relationships between mutant forms of WASp and the diverse features of these disorders remain ill-defined.

Objective: To define how dysfunctional gene transcription is causally linked to the degree of T_{helper} (Th) cell deficiency and genomic instability in XLT/WAS clinical spectrum.

Methods: In human Th1- or Th2-skewing cell culture systems, co-transcriptional R-loops (RNA:DNA duplex and displaced single-stranded DNA) and DNA double-strand breaks (DSBs) were monitored in multiple XLT and WAS patient samples, and in normal T cells depleted of WASp.

Results: WASp-deficiency provokes increased R-loops and R-loop-mediated DSBs in Th1-cells relative to Th2-cells. Mechanistically, chromatin-occupancy of Serine2-unphosphorylated-RNA Polymerase II is increased and that of topoisomerase-1, a R-loop preventing factor, is decreased at R-loop-enriched regions of *IFNG* and *TBX21* (Th1 genes) in Th1-cells. These aberrations accompany increased unspliced (intron-retained) and decreased spliced mRNA of *IFNG* and *TBX21* but not of *IL13* (Th2-gene). Significantly, increased cellular load of R-loops and DSBs, which are normalized upon RNaseH1-mediated suppression of ectopic R-loops, inversely correlates with disease severity scores.

Conclusion: Transcriptional R-loop imbalance is a novel molecular defect etiologic in Th1-immunodeficiency and genomic-instability in WAS. The study proposes that cellular R-loop load could be used as a potential biomarker for monitoring symptom severity in the XLT-WAS clinical spectrum, and could be targeted therapeutically.

Key Messages:

1. WASp deficiency triggers R-loop-mediated DNA damage in a subset of immune function genes.
2. R-loops accumulate to a greater degree in WASp-deficient Th1-cells as compared to WASp-deficient Th2-cells.
3. The magnitude of R-loop-linked DNA damage correlates with clinical severity scores in the XLT/WAS spectrum for a subset of mutations.

Capsule Summary: Serial monitoring of R-loop and DNA damage load in circulating lymphocytes could identify *WAS* mutations that manifest in severe disease (WAS) at initial presentation or evolve from mild disease (XLT) to severe disease during patient's lifetime.

Key words:

Primary immunodeficiency; Wiskott-Aldrich syndrome; X-linked thrombocytopenia; Th1/Th2-differentiation; R-loops, DNA damage, Genomic instability.

Abbreviations:

WAS: Wiskott-Aldrich syndrome

XLT: X-linked thrombocytopenia

XLN: X-linked neutropenia

WASp: Wiskott-Aldrich syndrome protein

RNAP2: RNA polymerase II

pCTD: Phosphorylated C-terminal domain of RNAP2

S2P or S5P: Phosphorylation of Serine 2 or Serine 5 residues of the CTD.

S2unP: Serine 2 unphosphorylated C-terminal domain of RNAP2.

Th1, Th2, Tfh cells: CD4+ T helper 1, T helper 2, T follicular helper lymphocytes

RNaseH1, H2, I: Ribonuclease H1, H2, I

SWI/SNF: SWItch/Sucrose Non-Fermentable

SRSF1: Serine/Arginine-Rich Protein-Specific Factor 1

mRNP: Messenger ribonucleoprotein

TOP: Topoisomerase

TSS and TTS: transcription start site and transcription termination site.

PBMC: Peripheral blood mononuclear cells

EBV-BLCL: Epstein-Barr virus immortalized B lymphoblastoid cell line

PHA: Phytohemagglutinin

EGFP: Enhanced green fluorescent protein

HTLV-1: Human T-cell lymphotropic virus type 1

LC-MS/MS: Laser capture tandem mass spectrometry

TDRD3: Tudor Domain Containing 3

BRCA1/BRAC2: Breast Cancer Susceptibility Protein Type 1/Type 2

ACCEPTED MANUSCRIPT

Introduction

WAS mutations result in X-linked thrombocytopenia (XLT), X-linked-neutropenia (XLN), or classic WAS.[1-3] Although the genotype-phenotype association is imperfect, most manifest some combination of immunodeficiency, autoimmunity, genomic instability, and/or hematologic malignancy.[4-8] WASp, encoded by *WAS*, is a multi-domain, adaptor protein that regulates actin-polymerization in the cytoplasm[9] and RNA polymerase-II (RNAP2)-dependent transcription in the nucleus of cells of the hematolymphoid lineage.[10-14] WASp-deficient $CD4^{+}$Th-cells show decreased expression of Th1-genes in Th1[10] and Tfh (Th-follicular) cells.[15] WASp is also required for the transcriptional-activation of the *CD19* gene in human B-cells,[16] and nuclear WASp partakes in murine T-cell development.[17] WASp in human myeloid cells,[18] and multiple other WASp-family proteins (Wave1, WASH) in different organisms,[19-22] also support nuclear-functions or modulate transcription. Human WASp by modulating the activity of chromatin-remodeling complexes of the SWI/SNF-family influences promoter "readiness" and gene-activation.[12] Murine WASH modulates chromatin activity of nucleosome-remodeling-factor (NURF) at the *c-Myc* promoter,[20] and *drosophila* Wash influences global nuclear organization by influencing chromatin configuration.[21] Yeast-Las17, a homolog of human WASp, is part of nucleolar rDNA-bound proteins involved in maintaining nucleolar integrity.[23] Xenopus Wave1, like human WASp, facilitates RNAP2 transcription.[19] Accordingly, multiple evidence links WASp-family proteins, well-known for their actin-polymerizing function, also to nuclear/transcriptional function(s) in both higher and lower organisms. As such, the transcriptional function of human WASp and xenopus Wave1, is uncoupled from its actin-polymerizing function,[11,19] and the identification of the nuclear-localizing and nuclear-export sequences (NLS, NES) has allowed linking nuclear WASp, but not cytoplasmic WASp, with transcriptional activities.[11] Together, these findings underscore the evolutionary pressure to preserve the nuclear/transcriptional role(s) of WASp-family

proteins across different species, and for humans, propose WAS also as a disease of dysregulated transcription.

Regulated transcription is a multi-step process that includes safeguarding the newly synthesized, pre-mRNA (intron-containing) strand. Because DNA:RNA association is thermodynamically more stable compared to DNA:DNA association,[24] such nascent RNA transiently hybridize to its template single-strand(ss) DNA, inducing the formation of R-loops, a 3-stranded nucleic-acid structure consisting of RNA:DNA duplex and a displaced non-template ssDNA.[25] R-loops normally occur genome-wide, covering up to 5% of the human genome,[26] where they function in gene transcription, immunoglobulin class-switching, DNA replication, and DNA repair.[25-30] R-loops form both at 5'-TSS (promoter) and 3'-TTS (terminator) of a transcribing gene,[31,32] where they are involved in remodeling the chromatin.[33] These functions represent the physiologic roles of R-loops in transcription-initiation and -elongation. Conversely, unresolved or aberrant R-loops can retard elongation efficiency,[34,35] and are detrimental to the genome by invoking activation of transcription-coupled nucleotide excision repair (TC-NER) factors that cleave R-loops to generate DNA single or double-strand breaks (SSBs or DSBs), which disrupts transcription.[36-38] Accordingly, R-loops both activate or repress transcription, and hence can be beneficial or detrimental.[39-43] To mitigate detrimental effects, cells employ multiple R-loop modulating factors to restrict aberrant R-loops. These are RNaseH1/H2 (degrade RNA moiety of RNA-DNA duplex),[44,45] Aquarius and Senataxin (RNA/DNA-helicases),[46,47] THO-complex and SRSF1 (mRNP-biogenesis factors),[35,48-50] and topoisomerases (TOPs, resolve DNA-supercoiling).[51,52] As such, TOP1 curbs R-loops both by relieving negative DNA-supercoils and facilitating mRNA-splicing.[53,54] These factors are central to preventing R-loop-linked genome instability, since their loss/dysfunction results in human diseases including cancer.[55-57]

147 In this connection, we previously demonstrated that in WASp-deficient Th1-cells, decreased
148 mRNA level of *IFNG*, a signature Th1-cytokine, accompanies reduction in the chromatin-
149 enrichment of RNAP2-pCTD-Ser2 (S2P, a mark of transcription-elongation) but not of
150 RNAP2-pCTD-Ser5 (S5P, a mark of transcription-initiation) in its gene-body.[12] Because
151 RNAP2-complexes that are Ser2-unphosphorylated (hereinafter S2unP) could still produce
152 unspliced (intron-retaining, IR^+) short-transcripts through non-productive elongation[58] and
153 because such IR^+ transcripts could potentially reanneal to their template DNA-strand,[59,60] we
154 tested the possibility that WASp-deficiency might trigger the formation of aberrant R-loops at
155 actively-transcribing genes in T cells. In this study, we demonstrate a link between WASp-
156 deficiency, R-loop accumulation, and DNA-damage, defects that are more pronounced in
157 Th1-cells relative to Th2-cells, which we propose contributes to the development of an aspect
158 of WAS symptomatology. Furthermore, our data of R-loop imbalance rationalize the
159 associated but less severe disorder, XLT, also caused by *WAS* mutations.

ACCEPTED MANUSCRIPT

Methods

Cells, transfection, and pharmacological-modulation

$CD4^+$ T helper (Th) cells were isolated from PBMCs by MACS-apparatus (Meltenyi) from 3 normal-donors: ND1 (hereinafter NIH001-WT), ND3, ND8, and 7 WAS-patients: P1 (p.V75M), P2 (p.L39P), P3 (p.P58R), P4 (del570T-fs260-stop), P5 (del450G-fs167-stop), P6 (c.1509A>C-no stop) and P7 (del23G-fs44-stop, hereinafter NIH002-WAS).[4-6,11,61] NIH001-WT and NIH002-WAS were immortalized by HTLV-1 (gift from Dr. Fabio Candotti).[10] Patient samples were collected under the guidelines of the Institutional Review Board (IRB). NIH002-WAS was stably-transfected to express either Flag/Myc-dual tagged wild-type WASp (hereinafter wt-WASp) or different disease-causing WASp-mutants, as previously described.[11-13] T cells were polarized into Th1 or Th2 cells using Human Th1- or Th2-Differentiation Kit from R&D Systems. Primary, non-immortalized, patient-derived Th cells (P1-P6) were propagated for multiple weeks in vitro in IL2-containing culture by adding irradiated, unrelated donor PBMCs ($1x10^6$/mL) and JY, an EBV-BLCL ($0.1x10^6$/mL), and stimulating with PHA (1μg/mL). For RNaseH1-expression, pEGFP-M27-RNaseH1 plasmid lacking mitochondrial-localizing sequence but containing a nuclear-localization sequence or its vector-only plasmid (gift from Dr. Robert Crouch) was transfected into Th1-cells by Amaxa®Nucleofector® (Lonza). RNaseH1-expressing Th1-cells were enriched by FACS-sorting the GFP^+ subpopulation. For transcription inhibition assays, Th1-cells were treated with 5,6-Dichlorobenzimidazole-1-β-D-Ribofuranoside (DRB) at 25μM for 30min, 50μM for 1h or 2h, and 100μM for 2h, or with 5μg/ml α-amanitin for 16hr. For F-actin depolymerization, normal Th1-cells were treated with 2μM of Cytochalasin-D or Latrunculin-B for 1h or 2h.

CRISPR/Cas9-mediated knock-out and shRNA-mediated knock-down of WASp expression in human Th-cells.

WASp-deficient Th-lines were generated from NIH001 and Jurkat (WT/WKO isogenic pairs) by CRISPR-mediated KO and from ND8 (WT/WKD isogenic pair) by shRNA-mediated KD (and scrambled-vector control KD). NIH001-WT and Jurkat-WT were transfected with 2ug of WASP-CRISPR/CAS9-GFP plasmid (sc-400712-KO-2) using Amaxa-nucleofaction. GFP^+ cells were FACS-sorted, plated into 96-well plates by serial dilution, and screened by PCR to identify clones exhibiting biallelic (total) WAS-KO (NIH001-WKO and Jurkat-WKO). DNA-sequencing of gDNA, RT-PCR of mRNA and Western blot for WASp-expression were employed to verify successful WASp-depletion. For RNAi-mediated KD of *WAS* in primary Th-cells, lentiviral vectors encoding GFP and shRNA-sequences targeting *WAS* or a control (scrambled sequence) were used as described,[62] and WASp expression assessed on sorted GFP^+ Th-cells.

Coimmunoprecipitation and Mass-Spectrometry

Coimmunoprecipitatios/Western blots were performed as we described previously,[10-13] using commercial reagents shown in Supplementary file-1**.** Multiple MS-assays on immunoprecipitated WASp-enriched protein-complexes were performed after pretreating normal donor, primary T cells samples with micrococcal nuclease (MNase),[11] the latter known to mitigate false-positive protein-protein interaction by cleaving single- and double-stranded DNA and RNA.[63]

Confocal Microscopy and Flow cytometry

Immunofluorescence imaging was performed using Zeiss LSM710 integrated with Zen software, and z-stack images acquired at 0.3μM step-size at 63X magnification for analyses, as we described previously.[10,64] FIJI-image analysis software (NIH) was used for all quantifications. Automated counting of S9.6+ foci/nucleus (nucleoplasm+nucleolus) was performed using the "particle analysis" function of Image-J software isolating the analyses to DAPI-stained regions. Signal saturation of γH2A.X in the nucleus was calculated using "3D-Object counter" and "Coloc-2" functions (Image-J software). Intracellular staining of proteins for flow was performed as we described previously.[11-13]

DNA-RNA immunoprecipitation (DRIP), ChIP-qPCR, and RT-PCR

Modified-DRIP was performed as described previously.[65] Briefly, total nucleic-acid isolated from Th1-cells was sonicated to a size of 400-500 bp and pretreated with RNaseI to clear of contaminating single-stranded RNA. 3μg of eluted sample, after phenol/chloroform extraction, was incubated for 15hr with S9.6 antibody and immunoprecipitated RNA:DNA-hybrids were used in PCR-assays. All single-ChIP (1-round) and Re-ChIP (sequential 2-rounds of ChIP) assays were performed with MNase-digested chromatin as previously described.[10-13] ChIP-grade and their isotype-Ig antibodies used to pull-down DNA:protein complexes are shown in Supplementary file-1**.** Non-specific signals obtained with control IgG-ChIP were subtracted from those obtained in the test samples. RT-PCR assays including qPCR analyses on RT² Profiler™ PCR Array Human NF-κB-Signaling-Pathway-Plus (Qiagen) were performed as we described previously.[11-13]

Pre-mRNA estimation

231 Pre-mRNA was enriched using click-it nascent-RNA-capture-kit (Thermo-Fisher). Briefly,
232 $10x10^6$ cells were cultured with 5-ethynyluridine (EU) for 0.5h or 1h. 1ug/µl of RNA was
233 biotinylated with 10mM biotin azide, 50µL of Dynabeads® MyOne™Streptavidin-T1 magnetic
234 beads, and the 5-EU-labelled RNA used for qPCR assays (see Supplementary file-2 for
235 primer/probe sequences).
236
237 **Neutral comet assay**
238 2.5 X 10^5 cells/ml mixed with 50µL of 1% low gelling agarose (Sigma) were pipetted onto
239 CometSlide™ (Trevigen) and placed in the Sub-Cell® GT Agarose Gel Electrophoresis
240 System (Bio-Rad) for 45 min at 30V in neutral electrophoresis buffer.[66] Slides were stained
241 with SYBR-Gold dye (Thermo-Fisher) and z-stack images acquired using a Zeiss-confocal
242 microscope. Comet images were analyzed for tail-moment statistics using OpenComet
243 software.

ACCEPTED MANUSCRIPT

Results

WASp-deficiency triggers R-loop accumulation in human Th1-cells

Using S9.6 antibody (monitors RNA-DNA hybrids), we quantitated nuclear R-loops in NIH002-WAS Th-line (WAS patient-derived) and in 3 isogenic pairs of Th-lines (donor-derived), each either expressing or lacking WASp: 1) NIH001-WT and NIH001-WKO, 2) ND8-scrambled KD and ND8-shRNA WKD, and 3) Jurkat-WT and Jurkat-WKO (See Supplementary Figure 1 for phenotypic characteristics and Supplementary Figure 2 for WASp depletion). Th1-activation triggered high R-loops ($S9.6^+$ foci) in NIH002-WAS and all 3 WASp-deficient Th-lines (NIH001-WKO, ND8-WKD, Jurkat-WKO) relative to their WT-control pairs, captured at a single-cell level by confocal microscopy (Figures 1A, B) and at population level by flow cytometry (Figure 1C). Overexpression of eGFP-tagged RNaseH1 (an endoribonuclease that degrades RNA in a RNA:DNA hybrid) into NIH001-WKO (Supplementary Figure 3), decreased the number of S9.6+foci/nucleus, which establishes $S9.6^+$foci as authentic RNA:DNA hybrids. Moreover, treating NIH001-WKO with DRB or α-amanitin, both inhibitors of RNAP2-driven transcription, also decreased the number of $S9.6^+$foci (Figures 1A, B). Disrupting F-actin assembly in normal (NIH001-WT) Th1-cells, by cytochalasin-D or latrunculin-B (Figure 1D), or stalling transcription-elongation by DRB (Supplementary Figure 4) did not provoke increased R-loop accumulation. The latter suggests that perturbing elongation alone is insufficient to trigger R-loops in human Th1-cells when WASp is present. Interestingly, Th2-skewing did not increase the number of $S9.6^+$ foci in ND8-WKD relative to ND8-WT, but it did increase in NIH001-WKO relative to NIH001-WT (Supplementary Figure 5A). However, the number and size of $S9.6^+$foci were significantly lower in NIH001-WKO Th2-cells compared to NIH001-WKO Th1-cells (Supplementary Figure

5B). Together, these findings demonstrate that R-loops accumulate to a greater degree in WASp-deficient Th1-cells as compared to WASp-deficient Th2-cells.

To clarify the consequences of R-loops on Th1-differentiation, at the chromatin level, we next mapped R-loop frequency, in the presence or absence of WASp, on a subset of Th1-network genes (*IFNG*, *TBX21, RUNX3, IL12RB2, TLR1*). Using ChIP-qPCR, we first show that endogenous-WASp is recruited to multiple regions of the *IFNG*-locus (Figure 1E) in NIH001-WT**,** where its enrichment increases towards the 3'-region of the gene (Figure 1F), as shown previously using the ChIP-on-chip approach.[10] This WASp-binding profile coincides with a similar 3'-biased accumulation of multiple other transcription factors.[67,68] Using S9.6 antibody in DRIP-qPCR, we show in NIH001-WT Th1-cells that R-loops formed at a low level on the *IFNG*-locus, where the R-loop enrichment, like that of WASp, also increased towards the 3'-region of the gene (Figure 1G). This DRIP-profile is consistent with the physiologic role of R-loops in transcription elongation/termination of actively transcribing genes.[27,31] By contrast, R-loops accumulated at a high level in NIH002-WAS Th1 cells, which were normalized upon transfecting either wild-type(wt)-WASp or RNaseH1 (Figures 1H, I), or treating with α-amanitin (Figure-1J), the latter established S9.6 DRIP-signals as representing co-transcriptional RNA:DNA hybrids. High R-loop accumulation was also captured at 3'-ends of multiple other Th1-genes in NIH002-WAS relative to NIH001-WT (Figure 1K). Notably, increased R-loops are not observed in Th1-cells at 3'-ends of *IL5* (Th2-gene), *IL10* (Treg-gene), *FOXP3* (Treg-gene), *IL17A* (Th17-gene) or *ZNF544* (R-loop-resistant locus)[32] (Figure-1K).

WASp-deficiency provokes R-loop-linked DNA-damage in Th1-cells

Because unresolved R-loops cause DSBs,[36,40] we investigated the effect of increased R-loops on genomic stability in WASp-deficient Th1- and Th2-cells. Using confocal microscopy and flow cytometry, we first quantified the signal-saturation of γH2A.X(pSer139)$^{+}$ foci as a surrogate of DNA damage induction.[69] This showed a significantly higher γH2A.X-signal saturations in both NIH002-WAS and NIH001-WKO Th1-cells relative to NIH001-WT Th1-cells by imaging (Figures 2A, B) and flow-cytometry (Figure 2C). Notably, overexpressing RNaseH1 into NIH001-WKO Th1-line or treating with DRB (or α-amanitin) both significantly decreased γH2A.X signal-saturation (Figures 2A, B), as did transfecting wild-type WASp into NIH002-WAS Th1-cells (Figure 2C). Furthermore, neutral comet assay, which reports on DSBs,[66] showed that DSBs accumulated with higher frequency in all 3 WASp-deficient Th1-lines (NIH001-WKO, ND8-WKD, Jurkat-WKO) relative to their WT counterparts (Figure 2D), and the DSB-load is significantly reduced upon degrading R-loops (by RNaseH1) or inhibiting transcription (by DRB) (Figure 2E). Notably, like R-loops, the DSB load (tail-moment value) is also significantly higher in WASp-deficient Th1-cells compared to WASp-deficient Th2-cells (Supplementary Figure 5A).

We next investigated whether DNA-damage occurs on Th1-genes at/or around the site(s) of R-loop formation. By ChIP-qPCR, we show that γH2A.X-enrichment at *IFNG* is significantly higher in NIH002-WAS relative to NIH001-WT Th1-cells (Figure 2F, left). At the 3'-region of *IFNG*, ~80% of histone-H2A is γH2A.X-modified in NIH002-WAS compared to ~20% in NIH001-WT Th1-cells (Figure-2F, right**)**. Moreover, Re-ChIP assays (2-rounds of ChIPs: 1st:γH2A.X>2nd:DNA-PKcs) showed that γH2A.X and DNA-PKcs (targets DNA-breaks)[70] co-enriched at a higher frequency at the same 3'-regions of multiple Th1-network genes in NIH002-WAS relative to NIH001-WT Th1-cells (Figure 2G), which is normalized by transfecting wt-WASp (Figure 2F) or RNaseH1 (Figure 2I). Interestingly, transfecting NIH002-

315 WAS Th1-cells with ΔNES-WASp (nucleus-"only" WASp re-expression)[11] reduced both R-
316 loop- and γH2A.X-load, whereas, transfecting ΔNLS-WASp (cytosol-"only" WASp re-
317 expression)[11] did not (Supplementary Figure 6)**.** We conclude that nuclear deficiency of
318 WASp results in DNA-damage, in part through the accumulation of RNA:DNA-hybrids.
319
320 **WASp-deficiency favors accumulation of intron-retaining, unspliced transcripts of**
321 ***IFNG* and *TBX21* but not *IL13* in Th1-cells**
322 In WASp-deficient Th1-cells, increased R-loops at *IFNG* and *TBX21* loci occurs in the context
323 of decreased mature mRNA-expression of the same genes (Figure-3A).[10-13] Because R-loops
324 are co-transcriptional products, these findings although counterintuitive, suggest that ongoing
325 nascent transcription must be occurring at these Th1-genes to fuel R-loop formation. To test
326 this possibility, we employed nonradioactive Click-iT® assay, in which Th1-cells were labelled
327 with 5-ethynyluridine (EU) for 30min and 1h (Figure-3B). Using multiple qPCR primer-pairs
328 covering all intron-exon boundaries of *IFNG*, *TBX21* (Th1-genes), *IL13* (Th2-gene) *and*
329 *TUBA1A* (control-gene), we show a multifold increase in the expression of pre-mRNA-
330 transcripts of *IFNG* and *TBX21* in NIH001-WKO relative to NIH001-WT Th1-cells (Figure 3B).
331 In contrast, there was a multifold decrease in the expression of mature, processed mRNA-
332 transcripts (reported by primers covering coding-exons) of *IFNG* and *TBX21*, but not *TUBA1B*
333 or *IL13*, in the same RNA-samples (Figure 3A).
334 Because transcriptional defects that generate intron-retaining, pre-mRNAs also result
335 in decreased level of mature mRNA, the latter by its degradation in the nonsense-mediated
336 decay (NMD) pathway,[71] we asked if WASp-deficiency triggers accumulation of pre-mRNA
337 transcripts containing whole introns. Intron-retention was assayed by both quantitative RT-
338 qPCR and semi-quantitative RT-PCR to monitor the relative abundance of unspliced *versus*
339 spliced transcripts of *IFNG*, *TBX21, IL13, and TUBA1B* in 1h EU-labelled total-RNA from

NIH001-WKO and NIH001-WT Th1-cells. MaxENT-scores (maximum-entropy-splicing-strength) were calculated for 5'-donor and 3'-acceptor splice-sites,[72] and an intron associated with high splice-site strength, i.e., one that would be readily spliced-out co-transcriptionally, was selected. Accordingly, intron-2 of *IFNG* (MaxEnt 5':9.85; 3':8.68), intron-4 of *TBX21* (5':8.41; 3':10.5), intron-2 of *TUBA1B* (5':9.27; 3':13.06), and intron-2 of *IL13* (5':9.33; 3':9.88) were analyzed. We found a multi-fold increase in the frequency of intron-retained pre-mRNA transcripts (EU-labeled, 1h) with a high unspliced:spliced transcript ratio for *IFNG* and *TBX21* in NIH001-WKO relative to NIH001-WT Th1-cells (Figure 3C-E). In contrast, *TUBA1* and *IL13* mRNA transcripts were fully-spliced, equally in NIH001-WKO and NIH001-WT Th1-cells, indicating that the splicing defect associated with WASp-deficiency is gene-specific. Together, the data demonstrates that transcription, *albeit* non-productive of mature mRNA, is still ongoing at these Th1-genes even when WASp is absent, and that WASp plays a yet to be defined role in pre-mRNA splicing.

WASp-deficiency favors accumulation of Ser2-unphosphorylated RNAP2-complexes at *IFNG* in Th1-cells

To further define the nature of transcriptional-dysfunction provoked by WASp-deficiency, we investigated the relationship between R-loops and RNAP2. Using ChIP-qPCR, we show decreased accumulation of Ser2-modified-RNAP2 (elongating-S2P) in the gene-body and at 3'-region of *IFNG*, this despite normal accumulation of Ser5-modified-RNAP2 (initiating-S5P) at its 5'-region, in NIH002-WAS Th1-cells compared to NIH001-WT Th1-cells (Figure-4A), which verifies our previous findings.[12] Decreased S2P-occupancy was captured also at the 3'-regions of multiple other Th1-genes, but not at *GAPDH*, in NIH002-WAS Th1-line, in which re-expressing wt-WASp was sufficient to correct S2P-defect (Figures 4B, C).

Because R-loops could destabilize or remodel RNAP2-complexes,[34,73-76] we investigated whether decreased S2P-occupancy is due to decreased occupancy of total-RNAP2 (i.e., destabilized) or from ectopic accumulation of Ser2-*UN*phosphorylated(S2unP)-RNAP2 (i.e., remodeled). ChIP-qPCR using pan-RNAP2-antibody, showed that total-RNAP2-occupancy was high (not low) at the *IFNG*-locus (Figure-4A). Accordingly, WASp-deficiency provokes increased chromatin-accumulation of S2unP-complexes (Figure-4D). Notably, RNaseH1-overexpression into NIH002-WAS Th1-cells resulted in ~4.5-fold increase in S2P-occupancy over its untransfected control (Figure-4E). By comparison, re-expressing wt-WASp into NIH002-WAS Th1-cells increased S2P-occupancy by ~10-fold. Because suppressing R-loops offered a partial S2P rescue, we postulate that loss of WASp impairs S2P-modification by altering multiple mechanisms, at least one of which is linked to R-loop accumulation.

WASp-deficiency disrupts TOP1 accumulation at Th1-genes

Because TOP1-recruitment to RNAP2 requires S2P-modification,[74] it is possible that S2P-deficiency would disrupt TOP1-occupancy at Th1-genes, which could then provoke R-loop accumulation in WAS Th1-cells, as shown in other systems.[51,52] To test this possibility, we queried our LC-MS/MS-dataset of WASp-interactors previously generated from human Th1-cells.[11] These micrococcal nuclease (MNase)-treated, immunoprecipitation(IP)-based MS-assays[63] captured polypeptides of DNA-topoisomerase-1, -2A, -2B (TOP1, TOP2A, TOP2B) from the nuclear-WASp-complexes (Figure 5A, Supplementary file 3), whereas control IgG-IP/MS did not recover TOP-peptides. The physiologic association between WASp and TOP1 was verified by co-immunoprecipitation/western-blot (Figure 5B) and MNase-treated Re-ChIP (1^{st}ChIP:TOP1$\rightarrow$$2^{nd}$ChIP:WASp), which showed WASp:TOP1-complexes to co-enrich at 3'-

388 regions of multiple Th1-genes in NIH001-WT Th1-cells (Figure 5C). In contrast, WASp-
389 deficiency provoked a significant decrease in TOP1-occupancy, contemporaneously with
390 increased R-loops, in the gene-body and at 3'-region of Th1-genes in NIH002-WAS Th1-line,
391 both of which (TOP1, R-loops) were normalized by re-expressing wt-WASp (Figures 5D, E).
392 We conclude that WASp plays an essential role, direct or indirect, in modulating TOP1-
393 occupancy at certain Th1-genes, which we propose contributes to R-loop formation in WASp-
394 deficient Th1-cells.
395
396 **Suppressing R-loops normalizes mRNA-levels of NF-κB-response genes in WASp-**
397 **deficient Th1-cells**
398 We previously showed,[12,13] and verified here, that WASp-deficiency causes wide-spread
399 decreases in steady-state mRNA-expression of multiple NF-κB-response genes in NIH002-
400 WAS Th1-line (Figure 6A, Supplementary file 4**).** In contrast, expression of *CSF2* and
401 *TNFAIP2,* both pro-inflammatory cytokines, was elevated. (Figures 6A, B). To show a direct
402 connection between R-loops and cell-function, we down-modulated R-loops in NIH002-WAS
403 Th1-cells by RNaseH1, which resulted in normalization of mRNA-levels (e.g., ↑*IFNG, STAT1,*
404 *TBX21, IL12RB2*; ↓*CSF2*, *TNFAIP2*) (Figures 6A-C). These results imply that the level of
405 S2P-upregulation (~4.5-fold) afforded by RNaseH1-mediated R-loop suppression (Figure 4E)
406 is sufficient to restore Th1-cell functionality, at least of the immune-function genes examined.
407 How RNaseH1-mediated R-loop suppression decreases mRNA-levels of *CSF2* and *TNFAIP2*
408 in WASp-deficient Th1-cells is unclear, but we speculate that by decreasing R-loop-mediated
409 DSBs, the expression of pro-inflammatory genes, known to be pathologically increased by
410 chronic DDR-activation,[77,78] will be also be suppressed.
411

R-loop-mediated DNA-damage correlates with disease-severity in XLT/WAS clinical spectrum

Finally, we asked if the R-loop defect in Th1-cells, uncovered here, correlates with the disease-severity scores of XLT/WAS.[4-6,61] In a reconstitution study with engineered WASp-mutants, we show that mutations presenting in a subset of patients initially as XLT but may evolve subsequently into classic WAS with a score of 5 in the same patient (T45M, V75M)[4-6] (Figure 7A) triggered R-loop accumulation and inscribed chromatin marks of DNA breaks at the *IFNG* locus, but not at non-Th1 gene loci in NIH002-WAS Th1-line stably transfected with these pathogenic mutants (Figures 7B, C; Supplementary Figure 7). In contrast, mutations that have been reported to always associate with XLT (A236G, R477K),[4-6] did not manifest these aberrations at Th1 or non-Th1 gene loci. Significantly, in primary Th1-cells (not immortalized) from 6 patients (3 XLT [P1, P2, P3], 3 WAS [P4, P5, P6] and 2 normal-donors [ND8, ND3]), we show that the magnitude of R-loop and DSB-load in the patient Th1-cells with clinical-scores of 4-5 (i.e., WAS) is significantly higher than in the patient Th1-cells with clinical-scores of 0.5-1 (i.e., XLT) (Figure 7D-F). Together, the findings in primary Th cells (Figure 7D) and Th cell lines (Figures 1, 2, and 7A) support the conclusion that the degree of R-loop-linked DNA damage in Th1-cells correlates with the disease-severity scores of XLT-WAS for the pathogenic mutations tested.

Discussion

Our study uncovers R-loop-mediated genomic-instability as a new molecular defect in WAS Th1-cells, where the magnitude of defect correlates with clinical-severity scores in the XLT-WAS disease-spectrum. Accumulation of S2unP-RNAP2, decreased TOP1 chromatin-enrichment, and increased generation of unspliced pre-mRNA transcripts together manifest in increased R-loops, DNA-DSBs, and mature mRNA deficiency in a subset of Th1-genes. These aberrations manifest to a greater degree in WASp-deficient Th1-cells as compared to WASp-deficient Th2-cells, which may explain the observed Th1/Th2-imbalance[10,79] contributing to eczema, hyper-IgE, and other atopic/clinical features of WAS.[4-6] Based on our findings, we propose a model (Figure-8) in which the "platform" function of multi-domain WASp enables targeting of TOP1 and CDK9 to transcriptionally-active Th1-genes, and together with a similar "platform" function described for TDRD3 and BRCA1/BRAC2 in restricting R-loops,[80-82] these observations support "adaptor-signaling" as a means to focally-concentrate anti-R-loop enzymology to transcribing- and replicating-chromatin. Accordingly, by providing new mechanistic clarification of how nuclear-WASp influences gene-transcription, this and previous studies together consolidate the idea that WAS is also a disease of dysregulated transcription.[10-17,64,79]

R-loops and S2P-deficiency: cause or consequence in WAS Th1-cell pathology?

R-loop-'centric' model (R-loop is a primary defect): This model proposes that WASp-deficiency disrupts chromatin recruitment/retention of anti-R-loop factor(s) (e.g., TOP1), as a primary defect. Because TOP1 chromatin-recruitment is regulated by SWI/SNF-activity and H3K4me3,[83] both of which we show are deficient in WAS Th1-cells,[10,12] the model proposes that TOP1-defect could also arise from these chromatin/epigenetic aberrations. Because

TOP1-defect impedes elongation and/or splicing,[51-53,84] the latter causing accumulation of intron-retained (IR+) pre-mRNA, could further fuel R-loop formation (Figure-8), as shown recently in spinal muscular atrophy.[60] Loss of yeast-TOP1 also results in accumulation of pre-rDNA transcripts, R-loops, and impeded RNAP1-transcription.[51] Nonetheless, whether aberrant R-loops are a cause or consequence of splicing-defect in WAS Th-cells remains an open question.

S2P-'centric' model (S2unP is the primary defect): Our findings of increased accumulation of S2unP-complexes present this alternative model, which offers an additional benefit of explaining a 3'-biased gradient for excessive R-loop formation (Figure-8). Our results in WASp-deficient Th1-cells demonstrating impaired chromatin-recruitment/retention of a S2P-kinase CDK9[12] and TOP1, the latter essential also for S2P-modification by BRD4,[74] together support this model. Given NF-κB's role in S2P-elongation,[58] a defect in NF-κB-activity seen in WASp-deficient lymphocytes,[13,64] could also contribute towards S2P-deficiency.

Indeed, it is also possible that WASp exerts parallel, independent effects on both of these phenomena (R-loop and S2P models). Therefore, clarifying whether R-loops are a cause or consequence of defective S2P-elongation, which defect is a primary, disease-driving aberration in WAS, and how global are these aberrations in the WAS Th-cell genomes (Th1, Th2, etc.), will offer better definition of WASp-role in transcriptional R-loop metabolism. Similarly, since WASp-defect renders Th1 gene-promoters "inactive",[10,12-14] whether abnormal gene-body R-loops induce epigenetic-silencing of the 5'-promoter[85] or an abnormal 5'-promoter chromatin-status induces abnormal gene-body R-loop formation[33] in WAS is an open question, clarification of which would provide basis for modulating a specific chromatin-signaling pathway(s), potentially for therapy.

R-loops as a potential metric to monitor disease-severity in the XLT-WAS clinical spectrum

Of immediate clinical relevance are our findings that WAS-mutations that cause classic-WAS associate with high R-loops and DSBs, whereas, mutations that cause XLT do not, at least for mutations/patient-samples tested. Interestingly, for V75M, which has been reported to cause XLT or WAS in different patients, and in the same patient reported to evolve from XLT to WAS, R-loop profiles are instructive. Specifically, WAS^{null} Th1-line (NIH002-WAS) stably-transfected with engineered V75M-WASp mutant shows high R-loop/DSB burden, whereas primary Th1-cells from a V75M-patient obtained at the time of mild-disease (score-1) show low R-loop/DSB burden. Such contrasting findings propose that for the same pathogenic-mutation, R-loop-mediated genomic-instability can be modulated by patient's genetic make-up and/or clinical context, two disease-modifiers that are challenging to test in any inbred, isogenic mouse-model of WAS, which underscores the necessity to conduct experimental studies in patient samples. Accordingly, serial monitoring of R-loop/DSB at different time-points in the same patient is likely to be more informative about disease-scores in XLT/WAS spectrum, especially because some XLT patients are known to subsequently evolve into serious WAS during their lifetime.[4-6] Identification of R-loops as a potential biomarker of disease-severity is of clinical value because complete loss of WASp does not always correlate with severe WAS phenotypes.[4-6,86] Nonetheless, reproducing the R-loop: disease-score correlation, identified here in 10 pathogenic-mutations, in a larger number of mutations, is required to bolster the diagnostic/prognostic value of monitoring R-loops in XLT/WAS spectrum. Certainly, we do not favor the idea that all aspects of WAS complications can be explained by the cell-autonomous consequences of R-loops in Th-cells alone. We anticipate the Th-cell-specific DNA-damage to have systemic consequences, including provoking

503 autoimmunity/autoinflammation as a non-cell autonomous response to chronic DNA
504 damage.[77,78]

ACCEPTED MANUSCRIPT

Author contribution: K.S. performed ChIP, DRIP, coimmunoprecipitation/Western blot, WAS mutant and RNH1 transfections, flow cytometry, RT-qPCR assays, and assisted in comet assay. S.S.H. generated CRISPR/Cas9-mediated WAS KO cellular models, performed all nascent pre-mRNA assays, RNH1 transfection, and Western blot experiments. K.K.W. performed all confocal imaging, comet assays, assisted in RNH1 transfection, and performed statistical analyses for all images. H.D.O. and L.D., provided XLT/WAS patient samples and generated shRNA-mediated WAS KD cellular models. M.M.S. interpreted the data and wrote the paper; Y.M.V. conceived the study, designed the experiments, analyzed and interpreted the data, and wrote the paper.

Acknowledgements

We thank R. Crouch (NIH/NICHD, Bethesda, USA) for providing RNaseH1-eGFP and its control vector, A. Galy (Généthon, Evry, France) for the ShRNA lentiviral vector, A. Selvakumar (Memorial Sloan-Kettering Cancer Center, New York, USA) for assistance with nucleotide blast searches, F. Candotti (University Hospital of Lausanne, Switzerland) and A. Aiuti (San Raffaele Telethon Institute for Gene Therapy, Milan, Italy) for patient samples, and the University of Iowa Dance Marathon (UIDM) for providing laboratory space. This work was supported by the NIH, National Institute of Allergy and Infectious Diseases (NIAID) grant R01AI084957 (to Y.M.V.), the Intramural Research Program of the NIH, National Institute on Aging (NIA) grant Z01AG000746-08 (to M.M.S.), and the DeJoria Wiskott-Aldrich Research Fund (to H.D.O).

References

1. Derry JM, Ochs HD, Francke U. Isolation of a novel gene mutated in Wiskott-Aldrich syndrome. Cell 1994;78:635-44.
2. Villa A, Notarangelo L, Macchi P, Mantuano E, Cavagni G, Brugnoni D, et al. X-linked thrombocytopenia and Wiskott-Aldrich syndrome are allelic diseases with mutations in the WASP gene. Nat Genet 1995;9:414-7.
3. Ancliff PJ, Blundell MP, Cory GO, Calle Y, Worth A, Kempski H, et al. Two novel activating mutations in the Wiskott-Aldrich syndrome protein result in congenital neutropenia. Blood 2006;108:2182-9.
4. Imai K, Morio T, Zhu Y, Jin Y, Itoh S, Kajiwara M, et al. Clinical course of patients with WASP gene mutations. Blood 2004;103:456-64.
5. Albert MH, Bittner TC, Nonoyama S, Notarangelo LD, Burns S, Imai K, et al. X-linked thrombocytopenia (XLT) due to WAS mutations: clinical characteristics, long-term outcome, and treatment options. Blood 2010;115:3231-8.
6. Jin Y, Mazza C, Christie JR, Giliani S, Fiorini M, Mella P, et al. Mutations of the Wiskott-Aldrich Syndrome Protein (WASP): hotspots, effect on transcription, and translation and phenotype/genotype correlation. Blood 2004;104:4010-9.
7. Westerberg LS, Meelu P, Baptista M, Eston MA, Adamovich DA, Cotta-de-Almeida V, et al. Activating WASP mutations associated with X-linked neutropenia result in enhanced actin polymerization, altered cytoskeletal responses, and genomic instability in lymphocytes. J Exp Med 2010;207:1145-52.
8. Moulding DA, Blundell MP, Spiller DG, White MR, Cory GO, Calle Y, et al. Unregulated actin polymerization by WASp causes defects of mitosis and cytokinesis in X-linked neutropenia. J Exp Med 2007;204:2213-24.

9. Symons M, Derry JM, Karlak B, Jiang S, Lemahieu V, McCormick F, et al. Wiskott-Aldrich syndrome protein, a novel effector for the GTPase CDC42Hs, is implicated in actin polymerization. Cell 1996;84:723-34.
10. Taylor MD, Sadhukhan S, Kottangada P, Ramgopal A, Sarkar K, D'Silva S, et al. Nuclear role of WASp in the pathogenesis of dysregulated TH1 immunity in human Wiskott-Aldrich syndrome. Sci Transl Med 2010;2:37ra44.
11. Sadhukhan S, Sarkar K, Taylor M, Candotti F, Vyas YM. Nuclear role of WASp in gene transcription is uncoupled from its ARP2/3-dependent cytoplasmic role in actin polymerization. J Immunol 2014;193:150-60.
12. Sarkar K, Sadhukhan S, Han SS, and Vyas YM. Disruption of hSWI/SNF complexes in T cells by WAS mutations distinguishes X-linked thrombocytopenia from Wiskott-Aldrich syndrome. Blood 2014;124:3409-19.
13. Sarkar K, Sadhukhan S, Han SS, Vyas YM. SUMOylation-disrupting WAS mutation converts WASp from a transcriptional activator to a repressor of NF -kappaB response genes in T cells. Blood 2015;126:1670-82.
14. Teitell MA. Alternative control: what's WASp doing in the nucleus? Sci Transl Med 2010;2:37ps31.
15. Zhang X, Dai R, Li W, Zhao H, Zhang Y, Zhou L, et al. Abnormalities of follicular helper T-cell number and function in Wiskott-Aldrich syndrome. Blood 2016;127:3180-91.
16. Bai X, Zhang Y, Huang L, Wang J, Li W, Niu L, et al. The early activation of memory B cells from Wiskott-Aldrich syndrome patients is suppressed by CD19 downregulation. Blood 2016;128:1723-34.

17. Kuznetsov NV, Almuzzaini B, Kritikou JS, Baptista MAP, Oliveira MMS, Keszei M, et al. Nuclear Wiskott-Aldrich syndrome protein co-regulates T cell factor 1-mediated transcription in T cells. Genome Med 2017;9:91.
18. Looi CY, Sasahara Y, Watanabe Y, Satoh M, Hakozaki I, Uchiyama M, et al. The open conformation of WASP regulates its nuclear localization and gene transcription in myeloid cells. Int Immunol 2014;26:341-52.
19. Miyamoto K, Teperek M, Yusa K, Allen GE, Bradshaw CR, Gurdon JB. Nuclear Wave1 is required for reprogramming transcription in oocytes and for normal development. Science 2013;341:1002-5.
20. Xia P, Wang S, Huang G, Zhu P, Li M, Ye B, et al. WASH is required for the differentiation commitment of hematopoietic stem cells in a c-Myc-dependent manner. J Exp Med 2014;211:2119-34.
21. Verboon JM, Rincon-Arano H, Werwie TR, Delrow JJ, Scalzo D, Nandakumar V, Groudine M, and Parkhurst SM. Wash interacts with lamin and affects global nuclear organization. Curr Biol 2015;25:804-10.
22. Rodriguez-Mesa E, Abreu-Blanco MT, Rosales-Nieves AE, and Parkhurst SM. Developmental expression of Drosophila Wiskott-Aldrich Syndrome family proteins. Dev Dyn 2012;241:608-26.
23. Krol K, Jendrysek J, Debski J, Skoneczny M, Kurlandzka A, Kaminska J, et al. Ribosomal DNA status inferred from DNA cloud assays and mass spectrometry identification of agarose-squeezed proteins interacting with chromatin (ASPIC-MS). Oncotarget 2017;8:24988-5004.
24. Roberts RW, Crothers DM. Stability and properties of double and triple helices: dramatic effects of RNA or DNA backbone composition. Science 1992;258:1463-6.

25. Santos-Pereira JM, Aguilera A. R loops: new modulators of genome dynamics and function. Nat Rev Genet 2015;16:583-97.
26. Sanz LA, Hartono SR, Lim YW, Steyaert S, Rajpurkar A, Ginno PA, et al. Prevalent, Dynamic, and Conserved R-Loop Structures Associate with Specific Epigenomic Signatures in Mammals. Mol Cell 2016;63:167-78.
27. Skourti-Stathaki K, Kamieniarz-Gdula K, Proudfoot NJ. R-loops induce repressive chromatin marks over mammalian gene terminators. Nature 2014;516:436-9.
28. Chaudhuri J, Alt FW. Class-switch recombination: interplay of transcription, DNA deamination and DNA repair. Nat Rev Immunol 2004;4:541-52.
29. Yu K, Chedin F, Hsieh CL, Wilson TE, Lieber MR. R-loops at immunoglobulin class switch regions in the chromosomes of stimulated B cells. Nat Immunol 2003;4:442-51.
30. Hatchi E, Skourti-Stathaki K, Ventz S, Pinello L, Yen A, Kamieniarz-Gdula K, et al. BRCA1 recruitment to transcriptional pause sites is required for R-loop-driven DNA damage repair. Mol Cell 2015;57:636-47.
31. Ginno PA, Lim YW, Lott PL, Korf I, Chedin F. GC skew at the 5' and 3' ends of human genes links R-loop formation to epigenetic regulation and transcription termination. Genome Res 2013;23:1590-600.
32. Ginno PA, Lott PL, Christensen HC, Korf I, Chedin F. R-loop formation is a distinctive characteristic of unmethylated human CpG island promoters. Mol Cell 2012;45:814-25.
33. Chedin F. Nascent Connections: R-Loops and Chromatin Patterning. Trends Genet 2016;32:828-38.
34. Huertas P, Aguilera A. Cotranscriptionally formed DNA:RNA hybrids mediate transcription elongation impairment and transcription-associated recombination. Mol Cell 2003;12:711-21.

623 35. Dominguez-Sanchez MS, Barroso S, Gomez-Gonzalez B, Luna R, Aguilera A.
624 Genome instability and transcription elongation impairment in human cells depleted of
625 THO/TREX. PLoS Genet 2011;7:e1002386.
626 36. Sollier J, Stork CT, Garcia-Rubio ML, Paulsen RD, Aguilera A, Cimprich KA.
627 Transcription-coupled nucleotide excision repair factors promote R-loop-induced
628 genome instability. Mol Cell 2014;56:777-85.
629 37. Schwab RA, Nieminuszczy J, Shah F, Langton J, Lopez Martinez D, Liang CC, et al.
630 The Fanconi Anemia Pathway Maintains Genome Stability by Coordinating Replication
631 and Transcription. Mol Cell 2015;60:351-61.
632 38. Garcia-Rubio ML, Perez-Calero C, Barroso SI, Tumini E, Herrera-Moyano E, Rosado
633 IV, and Aguilera A. The Fanconi Anemia Pathway Protects Genome Integrity from R-
634 loops. PLoS Genet 2015;11:e1005674.
635 39. Aguilera A, Garcia-Muse T. R loops: from transcription byproducts to threats to
636 genome stability. Mol Cell 2012;46:115-24.
637 40. Sollier J, Cimprich KA. Breaking bad: R-loops and genome integrity. Trends Cell Biol
638 2015;25:514-22.
639 41. Costantino L, Koshland D. The Yin and Yang of R-loop biology. Curr Opin Cell Biol
640 2015;34:39-45.
641 42. Wickramasinghe VO, Venkitaraman AR. RNA Processing and Genome Stability:
642 Cause and Consequence. Mol Cell 2016;61:496-505.
643 43. Skourti-Stathaki K, Proudfoot NJ. A double-edged sword: R loops as threats to
644 genome integrity and powerful regulators of gene expression. Genes Dev
645 2014;28:1384-96.

646 44. Wahba L, Amon JD, Koshland D, Vuica-Ross M. RNase H and multiple RNA
647 biogenesis factors cooperate to prevent RNA:DNA hybrids from generating genome
648 instability. Mol Cell 2011;44:978-88.
649 45. Lim YW, Sanz LA, Xu X, Hartono SR, Chedin F. Genome-wide DNA hypomethylation
650 and RNA:DNA hybrid accumulation in Aicardi-Goutieres syndrome. Elife 2015;4.
651 46. De I, Bessonov S, Hofele R, dos Santos K, Will CL, Urlaub H, et al. The RNA helicase
652 Aquarius exhibits structural adaptations mediating its recruitment to spliceosomes. Nat
653 Struct Mol Biol 2015;22:138-44.
654 47. Skourti-Stathaki K, Proudfoot NJ, Gromak N. Human senataxin resolves RNA/DNA
655 hybrids formed at transcriptional pause sites to promote Xrn2-dependent termination.
656 Mol Cell 2011;42:794-805.
657 48. Li X, Manley JL. Inactivation of the SR protein splicing factor ASF/SF2 results in
658 genomic instability. Cell 2005;122:365-78.
659 49. Paulsen RD, Soni DV, Wollman R, Hahn AT, Yee MC, Guan A, et al. A genome-wide
660 siRNA screen reveals diverse cellular processes and pathways that mediate genome
661 stability. Mol Cell 2009;35:228-39.
662 50. Stirling PC, Chan YA, Minaker SW, Aristizabal MJ, Barrett I, Sipahimalani P, et al. R-
663 loop-mediated genome instability in mRNA cleavage and polyadenylation mutants.
664 Genes Dev 2012;26:163-75.
665 51. El Hage A, French SL, Beyer AL, Tollervey D. Loss of Topoisomerase I leads to R-
666 loop-mediated transcriptional blocks during ribosomal RNA synthesis. Genes Dev
667 2010;24:1546-58.
668 52. Drolet M, Phoenix P, Menzel R, Masse E, Liu LF, Crouch RJ. Overexpression of
669 RNase H partially complements the growth defect of an Escherichia coli delta topA

mutant: R-loop formation is a major problem in the absence of DNA topoisomerase I. Proc Natl Acad Sci USA 1995;92:3526-30.

53. Tuduri S, Crabbe L, Conti C, Tourriere H, Holtgreve-Grez H, Jauch A, et al. Topoisomerase I suppresses genomic instability by preventing interference between replication and transcription. Nat Cell Biol 2009;11:1315-24.

54. Rossi F, Labourier E, Forne T, Divita G, Derancourt J, Riou JF, et al. Specific phosphorylation of SR proteins by mammalian DNA topoisomerase I. Nature 1996;381:80-2.

55. Richard P, Manley JL. R Loops and Links to Human Disease. J Mol Biol 2017;429:3168-80.

56. Groh M, and Gromak N. Out of balance: R-loops in human disease. PLoS Genet 2014;10:e1004630.

57. Stork CT, Bocek M, Crossley MP, Sollier J, Sanz LA, Chedin F, et al. Co-transcriptional R-loops are the main cause of estrogen-induced DNA damage. Elife 2016;5.

58. Hargreaves DC, Horng T, Medzhitov R. Control of inducible gene expression by signal-dependent transcriptional elongation. Cell 2009;138:129-45.

59. Tresini M, Warmerdam DO, Kolovos P, Snijder L, Vrouwe MG, Demmers JA, et al. The core spliceosome as target and effector of non-canonical ATM signalling. Nature 2015;523:53-8.

60. Jangi M, Fleet C, Cullen P, Gupta SV, Mekhoubad S, Chiao E, et al. SMN deficiency in severe models of spinal muscular atrophy causes widespread intron retention and DNA damage. Proc Natl Acad Sci USA 2017;114:E2347-E56.

61. De Meester J, Calvez R, Valitutti S, Dupré L. The Wiskott-Aldrich syndrome protein regulates CTL cytotoxicity and is required for efficient killing of B cell lymphoma targets. J Leukoc Biol 2010;88:1031-40.

62. Lafouresse F, Cotta-de-Almeida V, Malet-Engra G, Galy A, Valitutti S, Dupré L. Wiskott-Aldrich syndrome protein controls antigen-presenting cell-driven CD4+ T cell motility by regulating adhesion to intercellular adhesion molecule-1. Immunology. 2012;137:183-96.

63. Nguyen TN, Goodrich JA. Protein-protein interaction assays: eliminating false positive interactions. Nat Methods. 2006; 3:135-9.

64. Huang W, Ochs HD, Dupont B, Vyas YM. The Wiskott-Aldrich syndrome protein regulates nuclear translocation of NFAT2 and NF-kappa B (RelA) independently of its role in filamentous actin polymerization and actin cytoskeletal rearrangement. J Immunol 2005; 174:2602-11.

65. Nadel J, Athanasiadou R, Lemetre C, Wijetunga NA, P OB, Sato H, et al. RNA:DNA hybrids in the human genome have distinctive nucleotide characteristics, chromatin composition, and transcriptional relationships. Epigenetics Chromatin 2015;8:46.

66. Olive PL, Banath JP. The comet assay: a method to measure DNA damage in individual cells. Nat Protoc 2006;1:23-9.

67. Kim H, Erickson B, Luo W, Seward D, Graber JH, Pollock DD, et al. Gene-specific RNA polymerase II phosphorylation and the CTD code. Nat Struct Mol Biol 2010;17:1279-86.

68. Mayer A, Lidschreiber M, Siebert M, Leike K, Soding J, Cramer P. Uniform transitions of the general RNA polymerase II transcription complex. Nat Struct Mol Biol 2010;17:1272-8.

69. Kinner A, Wu W, Staudt C, Iliakis G. Gamma-H2AX in recognition and signaling of DNA double-strand breaks in the context of chromatin. Nucleic Acids Res 2008;36:5678-94.
70. Davis AJ, Chen BP, Chen DJ. DNA-PK: a dynamic enzyme in a versatile DSB repair pathway. DNA Repair (Amst) 2014;17:21-9.
71. Kurosaki T, Maquat LE. Nonsense-mediated mRNA decay in humans at a glance. J Cell Sci 2016;129:461-7.
72. Yeo G, Burge CB. Maximum entropy modeling of short sequence motifs with applications to RNA splicing signals. J Comput Biol 2004;11:377-94.
73. Belotserkovskii BP, Hanawalt PC. Anchoring nascent RNA to the DNA template could interfere with transcription. Biophys J 2011;100:675-84.
74. Baranello L, Wojtowicz D, Cui K, Devaiah BN, Chung HJ, Chan-Salis KY, et al. RNA Polymerase II Regulates Topoisomerase 1 Activity to Favor Efficient Transcription. Cell 2016;165:357-71.
75. Felipe-Abrio I, Lafuente-Barquero J, Garcia-Rubio ML, Aguilera A. RNA polymerase II contributes to preventing transcription-mediated replication fork stalls. EMBO J 2015;34:236-50.
76. Broccoli S, Rallu F, Sanscartier P, Cerritelli SM, Crouch RJ, Drolet M. Effects of RNA polymerase modifications on transcription-induced negative supercoiling and associated R-loop formation. Mol Microbiol 2004;52:1769-79.
77. Bredemeyer AL, Helmink BA, Innes CL, Calderon B, McGinnis LM, Mahowald GK, et al. DNA double-strand breaks activate a multi-functional genetic program in developing lymphocytes. Nature 2008;456:819-23.

78. Kawanishi S, Ohnishi S, Ma N, Hiraku Y, Murata M. Crosstalk between DNA Damage and Inflammation in the Multiple Steps of Carcinogenesis. Int J Mol Sci 2017;18:1808.

79. Trifari S, Sitia G, Aiuti A, Scaramuzza S, Marangoni F, Guidotti LG, Martino S, Saracco P, Notarangelo LD, Roncarolo MG, Dupré L. Defective Th1 cytokine gene transcription in CD4+ and CD8+ T cells from Wiskott-Aldrich syndrome patients. J Immunol. 2006;177:7451-61.

80. Yang Y, McBride KM, Hensley S, Lu Y, Chedin F, Bedford MT. Arginine methylation facilitates the recruitment of TOP3B to chromatin to prevent R loop accumulation. Mol Cell 2014;53:484-97.

81. Hatchi E, Skourti-Stathaki K, Ventz S, Pinello L, Yen A, Kamieniarz-Gdula K, et al. Livingston DM. BRCA1 recruitment to transcriptional pause sites is required for R-loop-driven DNA damage repair. *Mol Cell* 2015;57:636-47.

82. Bhatia V, Barroso SI, García-Rubio ML, Tumini E, Herrera-Moyano E, Aguilera A. BRCA2 prevents R-loop accumulation and associates with TREX-2 mRNA export factor PCID2. Nature 2014;511:362-65.

83. Husain A, Begum NA, Taniguchi T, Taniguchi H, Kobayashi M, Honjo T. Chromatin remodeller SMARCA4 recruits topoisomerase 1 and suppresses transcription-associated genomic instability. Nat Commun 2016;7:10549.

84. Li M, Pokharel S, Wang JT, Xu X, Liu Y. RECQ5-dependent SUMOylation of DNA topoisomerase I prevents transcription-associated genome instability. Nat Commun 2015;6:6720.

85. Chen PB, Chen HV, Acharya D, Rando OJ, Fazzio TG. R loops regulate promoter-proximal chromatin architecture and cellular differentiation. Nat Struct Mol Biol 2015;22:999-1007.

765 86. Ochs HD, Thrasher AJ. The Wiskott-Aldrich syndrome. J Allergy Clin Immunol. 2006
766 Apr;117(4):725-38.
767 87. Herrera-Moyano E, Mergui X, García-Rubio ML, Barroso S, Aguilera A. The yeast and
768 human FACT chromatin-reorganizing complexes solve R-loop-mediated transcription-
769 replication conflicts. Genes Dev. 2014. 28:735-48.
770 88. Kuntziger T, Landsverk HB, Collas P, and Syljuasen RG. Protein phosphatase 1
771 regulators in DNA damage signaling. Cell Cycle. 2011;10(9):1356-62.

ACCEPTED MANUSCRIPT

Figure Legends

Figure 1. R-loops accumulate in WASp-deficient Th1-cells. (**A**) confocal microscopy images of the indicated Th1-skewed cells under indicated experimental conditions dual-labelled with S9.6-antibody (R-loops, green) and DAPI (nucleus, blue). Displayed images are the collapsed composites of 20-30 z-stack images acquired at x63 magnification. (**B**) Box-and-whisker plots (whiskers@10-90%, horizontal bar denotes median, "+" denotes mean) generated from 3D (z-stack) images of at least n=100 cells from n=3 independent experiments. (**C**) Flow cytometry. Histogram plots and bar graphs of their mean fluorescence intensity (MFI) for total histone-H3 (control) and S9.6 (R-loops) in the indicated Th cells. (**D**) confocal images of S9.6+ foci (R-loops) and their abundance (box-and-whisker plot) in normal donor Th1-skewed cells treated with Cytochalasin-D (Cyt-D), Latrunculin-B (LAT-B), or DMSO-control; n=100 cells from n=3 independent experiments. p-values: ****<0.0001; *=0.01; ns, not significant using Mann-Whitney 2-tailed nonparametric analyses. (**E**) Genomic locations of 12 PCR primer/probes across the human *IFNG*-locus (UCSC-Hg18) showing distances (kB) from the 5'TSS and 3'TTS. (**F**) ChIP-qPCR profile of WASp-enrichment across the *IFNG*-locus in indicated Th1-cells/conditions. (**G-J**) DRIP-qPCR profiles of R-loop frequency across the *IFNG*-locus in Th1-skewed, indicated Th-cells/conditions. (**K**) DRIP-qPCR: R-loop frequencies at 3'-ends of the indicated genes in the indicated Th1-cells, transfected with RNaseH1 or control vector. Displayed data is the mean of n=3 independent experiments, +SEM; shown are actual Mann-Whitney p-values that compare all data points for the entire *IFNG*-locus in one condition relative to other. Note: in this study, *SNRPN* is used as a positive-control for R-loops as previously described (primer coordinates, Chr15:25223714-2522381, in the *SNRPN* mRNA-sequence, region1-2).[32] In another study,

796 an upstream region of the *SNRPN-SNURF* locus (Chr 15: 25132924-25132943, not in the
797 *SNRPN* mRNA-sequence) was used as a negative-control for R-loops. [87]
798
799 **Figure 2. Loss of WASp provokes R-loop-mediated DNA damage (A)** Confocal images of
800 indicated Th1-skewed cell lines under indicated conditions, dual-labelled with anti-γH2A.X
801 antibody, an indicator of DNA damage and DNA replication stress (green for NIH002; red for
802 NIH001) and DAPI (blue). **(B)** Box-and-whisker plots display γH2A.X signal
803 saturation/nucleus in 3D images of the cell, n=100-120 cells, ****p<0.0001, *p=0.01 by Mann-
804 Whitney. (**C**) Flow cytometry. Plot showing MFI values for γH2A.X, H2A.Z, and total-H2A
805 (control) immunofluorescence in the indicated T cells/conditions; unTx, untransfected; Tx,
806 transfected. (**D** and **E**) Neutral comet assay. Representative images of comet tails in
807 indicated Th-line pairs/conditions. Tail moment statistics are displayed as box-and-whisker
808 plots, n=100-150 cells from 3 independent experiments for each condition. ****p-
809 value<0.0001, **p-value<0.001 *p-value=0.01 by Mann-Whitney. (**F**) ChIP-qPCR. *Left*,
810 γH2A.X-enrichment at the *IFNG*-locus in indicated Th1-lines. *Right*, relative abundance of
811 γH2A.X-modified histone-H2A at 3'-end of *IFNG* (point#9, Figure 1E). (**G**) Re-ChIP assay. In
812 2 sequential-rounds of ChIPs (1stChIP: γH2A.X or IgG, 2ndChIP: DNA-PKcs), co-enrichment
813 profiles of γH2A.X and DNA-PKcs at the same 3'-ends of indicated genes in indicated Th1-
814 lines. (**H**) ChIP- and DRIP-qPCR profiles of the indicated factors at 3'-end (point #9) of the
815 *IFNG*-locus in indicated Th1-lines transfected with RNaseH1-eGFP (RNH1 tx.) or its empty-
816 vector (vector tx.). The data in panels F-H are mean of n=3, +SEM; actual p-values by Mann-
817 Whitney.
818

Figure 3. Loss of WASp favors accumulation of intron-retained, nascent pre-mRNA transcripts of *IFNG* and *TBX21* in Th1 cells. (**A**) Expression of processed (spliced, intron-removed) mRNA transcripts was quantified by generating multiple RT-qPCR primers that encompass indicated exon-exon junctions of all coding-exons of indicated genes, and is displayed as fold-change in WASp-deficient (WKO) relative to WASp-sufficient (WT) Th1-skewed lines (NIH001, isogenic pair). For *IL13*, only one exon-exon junction was tested. (**B**) *Top*, 5-ethynyluridine(EU)-labelling steps that yield EU-total RNA and EU-immunopurified RNA after 0.5h and 1h of EU-labelling. *Bottom*, expression of unprocessed (unspliced, intron-retained) pre-mRNA transcripts was quantified in both EU-total and EU-immunopurified RNA samples by generating multiple RT-qPCR primers that encompassed indicated intron-exon boundaries of indicated genes. Bar graphs depict mRNA fold-change in WKO Th1-cells relative to WT Th1-cells (NIH001, isogenic pair), after normalizing to their respective *GAPDH* mRNA levels. For fold-change calculations, mRNA-values obtained in WT were normalized as 1.0. Displayed values are mean+s.d., n=3 biological replicates, each with their separate, new EU-labeling. (**C, D)** Intron retention assayed by quantitative RT-qPCR (panel C) and semi-quantitative RT-PCR (panel D) for indicated introns displaying high splice-strengths (5'-acceptor and 3'-donor MaxEnt-scores) in WT and WKO Th1-line pairs. Data average of n=3, +s.d. In panel D, the relative abundance of unspliced (US) over spliced (S) transcripts was calculated using gel-densitometry, n=3 independent experiments. Note, the bar graphs display the abundance of "unspliced" mRNA for *IFNG* and *TBX21*, and "spliced" mRNA for *IL13* and *TUBA1B*. (**E**) nucleotide sequences of the RT-PCR products, aligned to the indicated genomic-sequences (gDNA). Yellow shading denotes aligned sequences.

Figure 4. Accumulation of Ser2-unphosphorylated (S2unP) RNAP2-complexes on the *IFNG* locus in WAS Th1-cells. (A) ChIP-qPCR profiles of indicated RNAP2-modifications

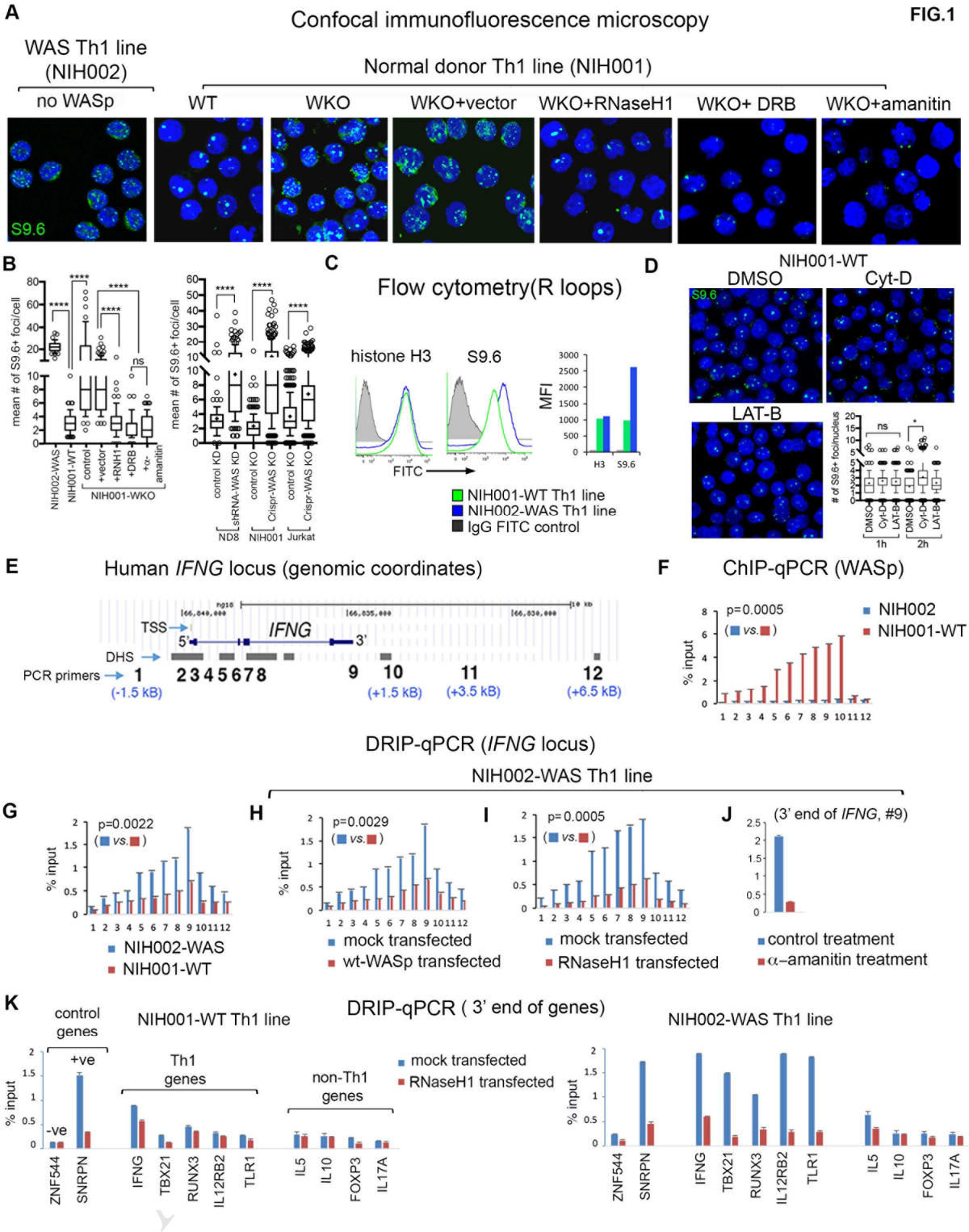
844 at the *IFNG* locus in indicated Th1-lines. S2P or S5P denotes ChIP performed with antibody
845 specific for Ser2- or Ser5-phosphorylated RNAP2. (**B**) ChIP-qPCR profiles of S2P-enrichment
846 at 3'-end of indicated genes in NIH002-WAS Th1-line (untransfected control, unTx.) and the
847 same WAS Th1-line stably-transfected with wt-WASp. p-values are as described in the
848 legend to Figure 1K. (**C**) ChIP-qPCR profiles of S2P- and total-RNAP2 (unmodified) at 3'-end
849 of *IFNG* and *GAPDH* (control locus) in WAS and WT Th1-lines. (**D**) The ChIP-qPCR values
850 for S2P and total-RNAP2 calculated from data in panel A show the relative abundance of
851 S2unP-RNAP2 across the *IFNG*-locus, in WT and WAS Th1-lines. (**E**) Bar graphs show an
852 inverse relationship between chromatin-enrichments of R-loops (S9.6 DRIP) and S2P (ChIP)
853 at the 3'-end of *IFNG*, under indicated conditions (RNH1-transfected, wt-WASp-transfected,
854 or untransfected control); data are an average of n=3 +SEM. ****p<0.0001; ns, nonsignificant,
855 Wilcoxon nonparametric test.
856
857 **Figure 5. Loss of WASp disrupts chromatin-occupancy of TOP1 at same cis-sites**
858 **where R-loops accumulate in Th1 genes.** (**A**) LC-MS/MS. Number of unique polypeptides
859 of indicated topoisomerases associated with WASp isolated from nuclear fraction of MNase-
860 treated Th1-skewed, primary Th-cells (endogenous WASp-proteome) or Jurkat Th-cells
861 (Flag/Myc-tagged, transfected wt-WASp-proteome) by immunoprecipitation (IP) with anti-
862 WASp antibody (Ab) or anti-FLAG/Myc Abs (2-step immunopurification, 1stIP:Flag Ab, 2ndIP:
863 Myc Ab). MultiConsensus reports of peptides/proteins were generated using the SEQUEST-
864 platform from n=4-5 biological replicates after applying the filtering criteria previously
865 described.[11] MS data from IgG-IP conducted in parallel did not recover TOP peptides (not
866 shown). (**B**) Coimmunoprecipitation (coIP)/Western blot. Protein-complexes isolated by coIP
867 from nuclear extracts of WT Th1-cells resolved by sequentially immunoblotting the same gel
868 with TOP1 and WASp. MM, Molecular marker (kDa). Input is 10%. (**C**). Re-ChIP assay. In 2

869 rounds of ChIP (1st ChIP:TOP1/or IgG, 2nd ChIP:WASp), co-enrichment profiles of TOP1 and
870 WASp at the same 3'-ends of indicated genes in WT Th1-cells. (**D**) ChIP-qPCR. Enrichment
871 profiles of TOP1 across the *IFNG*-locus (see Figure 1E) in indicated Th1-lines; unTx,
872 untransfected; wt-WASp Tx, transfected with wt-WASp. (**E**) ChIP- and DRIP-qPCR.
873 Enrichment profiles of indicted reagents at the 3'-ends of indicated genes in the same WAS
874 Th1-line, unTx or wt-WASp Tx. For panels C, D, E, data is from n=3, mean +SEM; p-values
875 in panel D by Wilcoxon nonparametric test.
876
877 **Figure 6. R-loop load modulates the steady-state mRNA expression patterns in human**
878 **Th1-cells**. **(A**) Heat maps reporting on the RT-qPCR-based mRNA expression profiles of NF-
879 κB response genes for WAS Th1-line, untransfected (control), or transfected with either
880 RNaseH1 or wt-WASp using the 96-well plate format from Qiagen. The up/down gene
881 modulation shown is in the context of transitioning from Th0-unstimulated to Th1-skewed
882 state, *in vitro*. Red, gene up-regulation, Green, gene down-regulation, black, unchanged.
883 Also see Supplementary file 4 for actual values. (**B**) Bar graphs display the actual values
884 (fold-change) of the indicated 4 genes derived from the mRNA expression array (heat-maps)
885 shown in panel B. *IFNG* is at C-03 and STAT1 at E-03 in the heat-map layout, and both
886 become upregulated, whereas, *CSF2* (GM-CSF) is at B05, *TNFAIP2* is at G-11, both pro-
887 inflammatory cytokine genes become downregulated in WAS Th1-cells upon expressing
888 RNaseH1 or wt-WASp. (**C**) RT-qPCR. Fold change in the mRNA expression of 3 indicated
889 Th1-genes in WAS Th-line (Th1-skewed) relative to the same WAS Th-line, untransfected
890 (control) or transfected with RNaseH1. n=3, mean +SEM.
891

892 **Figure 7. R-loop and DNA-damage levels in Th1 cells correlate with disease-severity**
893 **scores in WAS-XLT spectrum.** (**A**) The table shows 4 disease-causing *WAS*-mutations
894 engineered *in vitro* for stable transfection into NIH002-WAS-null T-cell line to generate Th-
895 lines representative of mutations that associate with mild or severe clinical phenotypes, as
896 shown. Clinical-severity: score-1, 2: mild-XLT with eczema; score-5A, -5M, severe
897 immunodeficiency with autoimmunity (A) or malignancy (M); 1, 2→5A/5M denotes disease
898 progression to severe WAS after initially presenting as XLT in the same patient.[4-6,86] (**B**) R-
899 loop frequency across the *IFNG*-locus in NIH002-WAS Th1-line stably-expressing indicated
900 WASp-mutants. p-values by Wilcoxon nonparametric test. (**C**) ChIP and DRIP (S9.6)
901 enrichment profiles of indicated factors at 3'-end of *IFNG* (point #9, Figure 1E) in the
902 indicated mutant. n=3, mean +SEM. (**D**) The table shows genotype-phenotype characteristics
903 of pathogenic mutations in patients from whom primary Th cells (P1-P6) and a Th cell line
904 (P7), as well as Th cells from normal donors (ND1, ND3, ND8) were generated.[10,61] **(E**) A
905 representative confocal-image (a collapsed composite of multiple z-stack planes) for each
906 patient and donor T-cell sample shows their respective R-loop load (S9.6+ foci) and DSB
907 load (γH2A.X and neutral comet assay). **(F)** Box-and-whisker plots (whisker@ 10-90%),
908 calculated on z-stacks (n=20-30 stacks/cell) from n=120-190 cells (exception: n=33 cells
909 analyzed for patient-P2 due to small sample volume) for indicated read-outs. S9.6+ foci,
910 γH2A.X signal-saturation in the nucleus, and comet tail moment were processed and
911 quantified using ImageJ 2.0 and Prism7 programs. Entire multi-cell images were optimized
912 equally by Adobe Photoshop to reduce any background fluorescence ("signal-noise").
913 ***p<0.0001, *p=0.01, ns= nonsignificant, p>0.1 by two-tailed non-parametric Mann-Whitney.
914

Figure 8. Model proposing WASp effect on proteins and pathways that modulate R loop balance. Event in blue-box denotes our previously published finding.[12] Events in yellow-box denote our findings of this study. Unboxed events denote predicted consequences of our results. Bold black arrows connect proposed mechanistic pathway for the observed signaling outcomes based on our findings of this study. Hatched grey arrows indicate putative consequences of our findings on other related pathways, not tested in this study, but predicted from published results in other systems. *Model*: Loss of WASp disrupts chromatin-occupancy of CDK9[12] and TOP1. CDK9-deficiency contributes to impaired S2P-modification, the latter further compromising TOP1 recruitment to RNAP2.[74] S2P-deficient RNAP2 can still generate short, unprocessed mRNA transcripts,[58] which could fuel R-loop formation. TOP1 defect will also contribute to R-loop formation through its effects both on DNA-supercoiling and splicing factors.[51-54] R-loops are known to incite DNA-damage, which via ATM-mediated spliceosome displacement,[59] could further fuel R-loop formation. R-loop-mediated DDR-signaling activates S2P-phosphatases, e.g., protein phosphatase-1,[88] which will dephosphorylate S2P contained within the elongation-complex in the gene-body. Together, a feed-forward causative loop of either R-loops⇒elongation-defect⇒R-loops (R-loop model) or elongation-defect⇒R-loops⇒elongation-defect (S2P model) will develop in WASp-deficient Th1-cells, resulting in R-loop-mediated DSBs. DSBs will elicit chronic-DDR, the latter known to perturb the global transcription programs of the affected cell, including provoking pro-inflammation,[77,78] which we propose contributes to immunodeficiency and autoimmunity/autoinflammation in both cell-autonomous and non-cell autonomous fashion.

FIG.1


A
Confocal immunofluorescence microscopy
WAS Th1 line (NIH002)
Normal donor Th1 line (NIH001)
no WASp
WT
WKO
WKO+vector
WKO+RNaseH1
WKO+ DRB
WKO+amanitin
S9.6
B
mean # of S9.6+ foci/cell
NIH002-WAS
NIH001-WT
control
+vector
+RNH1
+DRB
+α-amanitin
NIH001-WKO
control KD
shRNA-WAS KD
control KO
Crispr-WAS KO
control KO
Crispr-WAS KO
ND8
NIH001
Jurkat
C
Flow cytometry(R loops)
histone H3
S9.6
MFI
FITC
NIH001-WT Th1 line
NIH002-WAS Th1 line
IgG FITC control
D
NIH001-WT
DMSO
Cyt-D
LAT-B
# of S9.6+ foci/nucleus
1h
2h
E
Human IFNG locus (genomic coordinates)
TSS
IFNG
DHS
PCR primers
(-1.5 kB)
(+1.5 kB)
(+3.5 kB)
(+6.5 kB)
F
ChIP-qPCR (WASp)
p=0.0005
NIH002
NIH001-WT
% input
DRIP-qPCR (IFNG locus)
NIH002-WAS Th1 line
G
p=0.0022
NIH002-WAS
NIH001-WT
H
p=0.0029
mock transfected
wt-WASp transfected
I
p=0.0005
mock transfected
RNaseH1 transfected
J
(3' end of IFNG, #9)
control treatment
α–amanitin treatment
K
DRIP-qPCR (3' end of genes)
control genes
+ve
-ve
NIH001-WT Th1 line
Th1 genes
non-Th1 genes
mock transfected
RNaseH1 transfected
NIH002-WAS Th1 line
ZNF544
SNRPN
IFNG
TBX21
RUNX3
IL12RB2
TLR1
IL5
IL10
FOXP3
IL17A

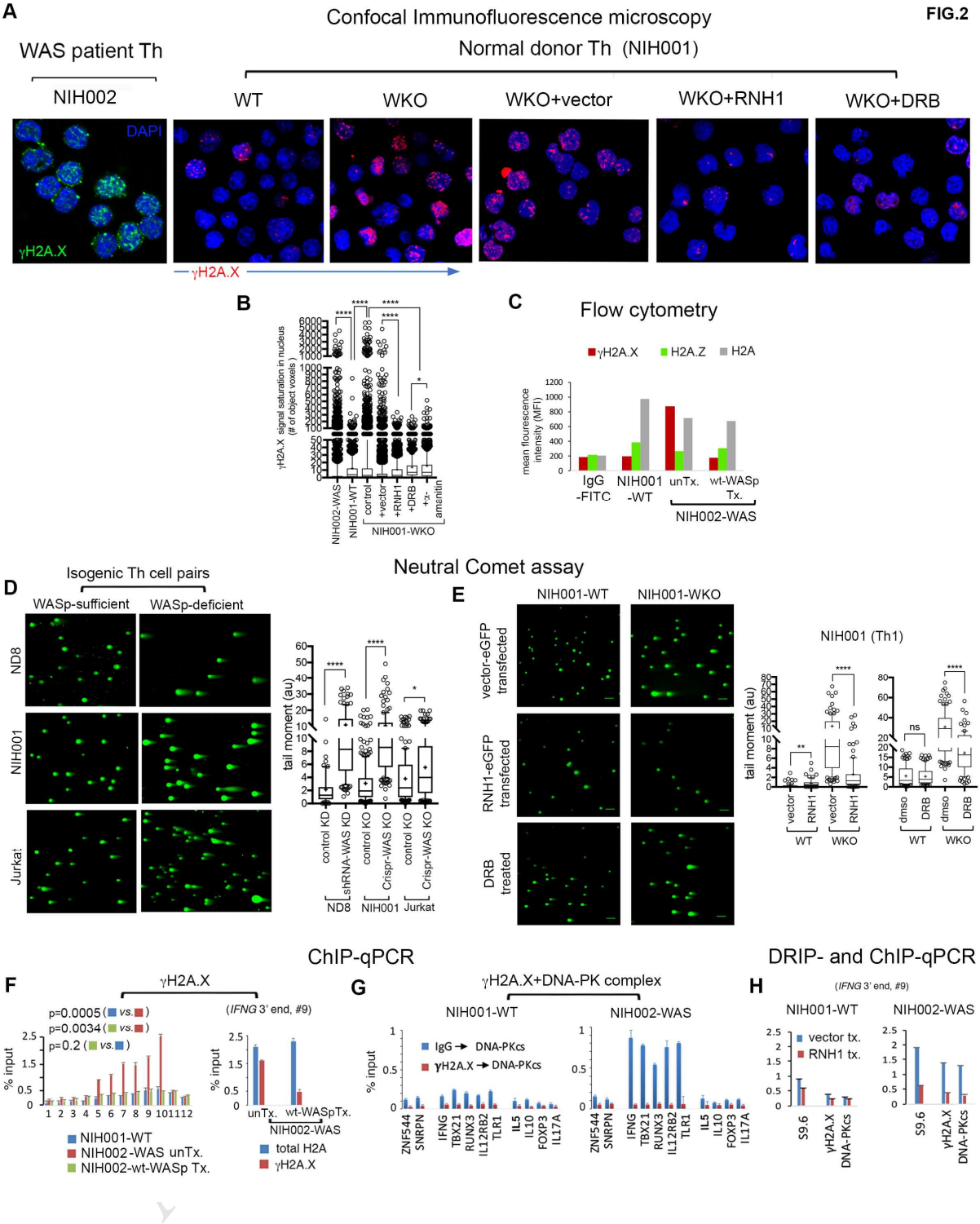

A
Confocal Immunofluorescence microscopy
FIG.2
WAS patient Th
Normal donor Th (NIH001)
NIH002
WT
WKO
WKO+vector
WKO+RNH1
WKO+DRB
DAPI
γH2A.X
B
γH2A.X signal saturation in nucleus (# of object voxels)
NIH002-WAS
NIH001-WT
control
+vector
+RNH1
+DRB
+z-amentin
NIH001-WKO
C
Flow cytometry
γH2A.X
H2A.Z
H2A
mean flourescence intensity (MFI)
IgG -FITC
NIH001 -WT
unTx.
wt-WASp Tx.
NIH002-WAS
D
Isogenic Th cell pairs
WASp-sufficient
WASp-deficient
ND8
NIH001
Jurkat
tail moment (au)
control KD
shRNA-WAS KD
control KO
Crispr-WAS KO
control KO
Crispr-WAS KO
ND8
NIH001
Jurkat
Neutral Comet assay
E
NIH001-WT
NIH001-WKO
vector-eGFP transfected
RNH1-eGFP transfected
DRB treated
NIH001 (Th1)
tail moment (au)
vector
RNH1
vector
RNH1
dmso
DRB
dmso
DRB
WT
WKO
WT
WKO
ChIP-qPCR
F
γH2A.X
p=0.0005 (vs.)
p=0.0034 (vs.)
p=0.2 (vs.)
% input
NIH001-WT
NIH002-WAS unTx.
NIH002-wt-WASp Tx.
(IFNG 3' end, #9)
% input
unTx.
wt-WASpTx.
NIH002-WAS
total H2A
γH2A.X
G
γH2A.X+DNA-PK complex
NIH001-WT
NIH002-WAS
% input
IgG → DNA-PKcs
γH2A.X → DNA-PKcs
ZNF544
SNRPN
IFNG
TBX21
RUNX3
IL12RB2
TLR1
IL5
IL10
FOXP3
IL17A
DRIP- and ChIP-qPCR
H
(IFNG 3' end, #9)
NIH001-WT
NIH002-WAS
vector tx.
RNH1 tx.
S9.6
γH2A.X
DNA-PKcs
S9.6
γH2A.X
DNA-PKcs

FIG.3

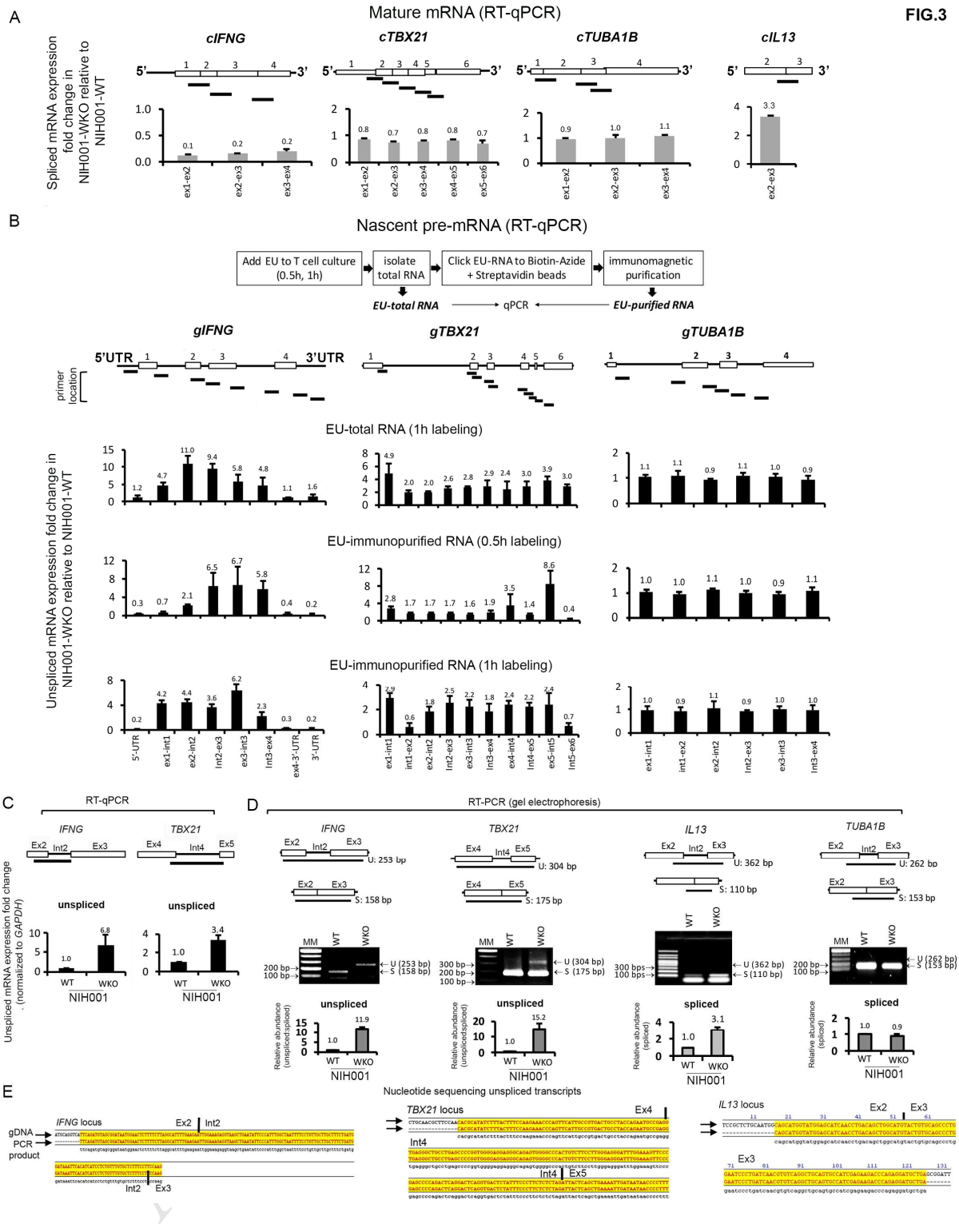
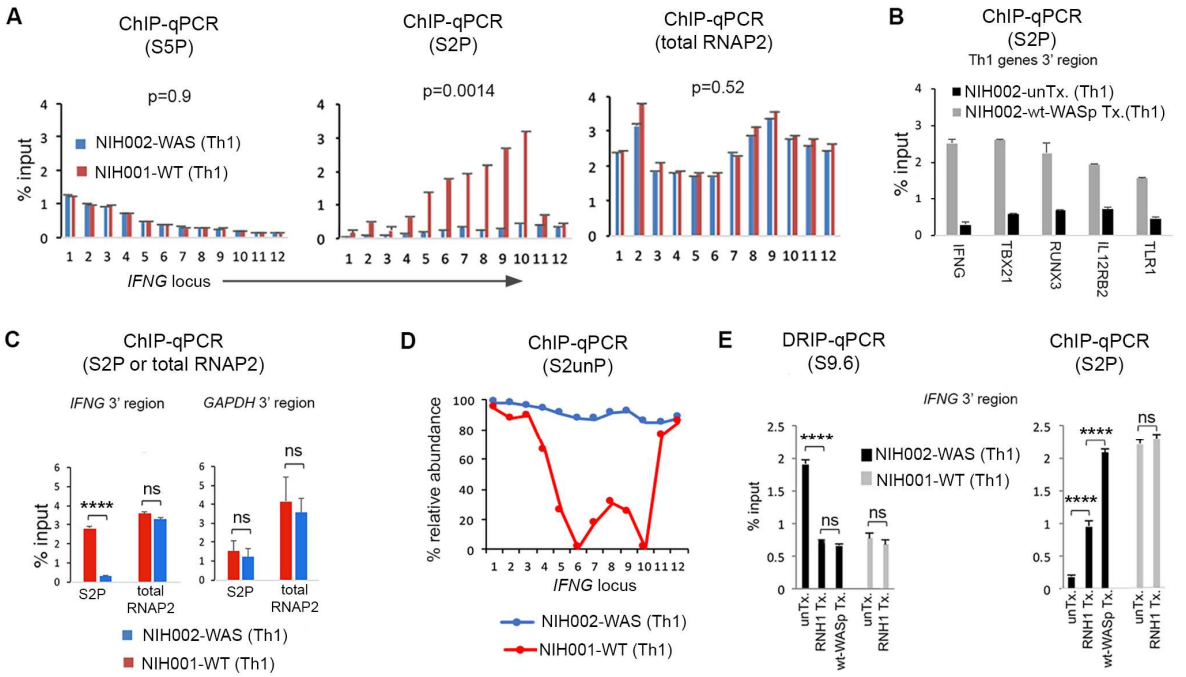

A
Mature mRNA (RT-qPCR)
cIFNG
cTBX21
cTUBA1B
cIL13
Spliced mRNA expression fold change in NIH001-WKO relative to NIH001-WT
ex1-ex2
ex2-ex3
ex3-ex4
ex4-ex5
ex5-ex6
B
Nascent pre-mRNA (RT-qPCR)
Add EU to T cell culture (0.5h, 1h)
isolate total RNA
Click EU-RNA to Biotin-Azide + Streptavidin beads
immunomagnetic purification
EU-total RNA
qPCR
EU-purified RNA
gIFNG
gTBX21
gTUBA1B
5'UTR
3'UTR
primer location
EU-total RNA (1h labeling)
EU-immunopurified RNA (0.5h labeling)
EU-immunopurified RNA (1h labeling)
Unspliced mRNA expression fold change in NIH001-WKO relative to NIH001-WT
5'-UTR
ex1-int1
ex2-int2
Int2-ex3
ex3-int3
Int3-ex4
ex4-3'-UTR
3'-UTR
int1-ex2
ex4-int4
Int4-ex5
ex5-int5
Int5-ex6
C
RT-qPCR
IFNG
TBX21
Ex2
Int2
Ex3
Ex4
Int4
Ex5
unspliced
Unspliced mRNA expression fold change (normalized to GAPDH)
WT
WKO
NIH001
D
RT-PCR (gel electrophoresis)
IL13
TUBA1B
U: 253 bp
S: 158 bp
U: 304 bp
S: 175 bp
U: 362 bp
S: 110 bp
U: 262 bp
S: 153 bp
MM
U (253 bp)
S (158 bp)
U (304 bp)
S (175 bp)
U (362 bp)
S (110 bp)
U (262 bp)
S (153 bp)
Relative abundance (unspliced/spliced)
spliced
Relative abundance (spliced)
E
Nucleotide sequencing unspliced transcripts
gDNA PCR product
IFNG locus
TBX21 locus
IL13 locus
Int4
Int2
Ex5

FIG.4



ACCEPTED MANUSCRIPT

FIG.5

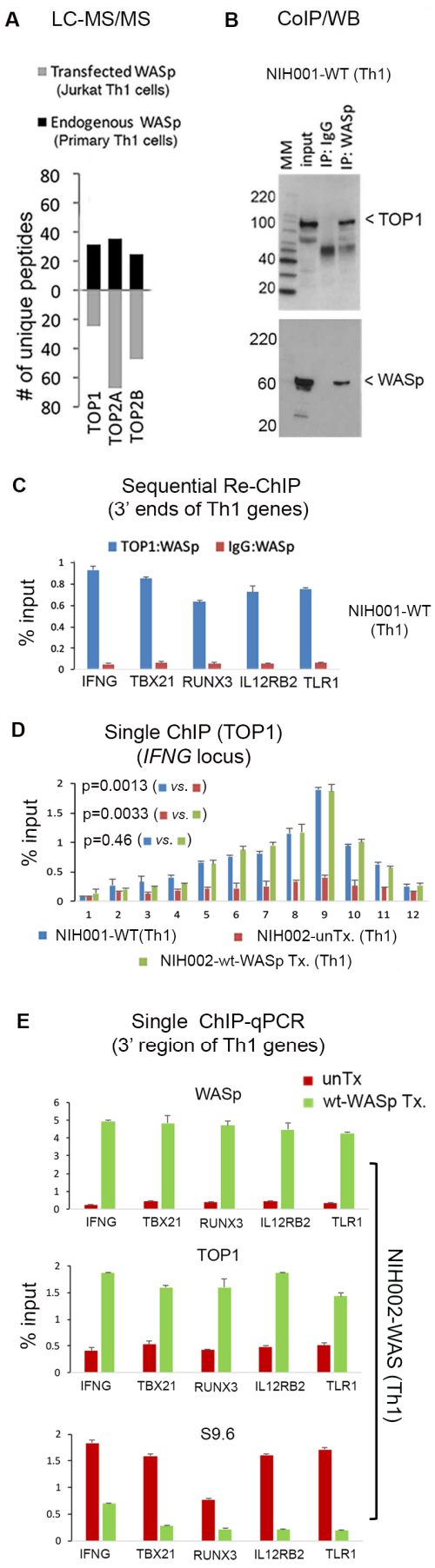
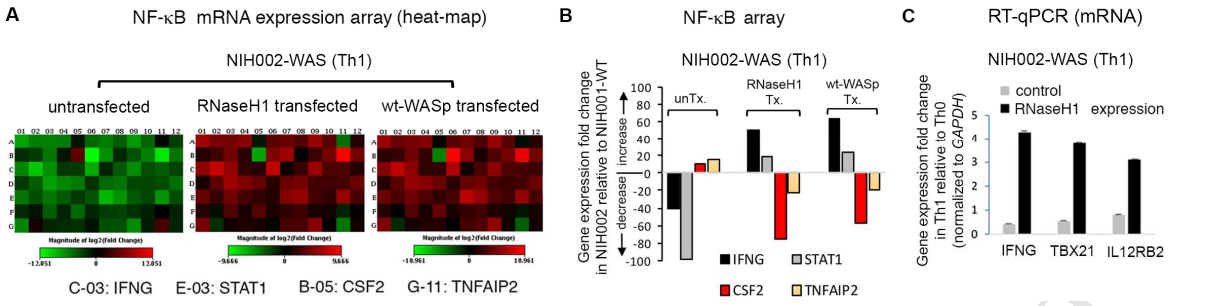

A
LC-MS/MS
Transfected WASp
(Jurkat Th1 cells)
Endogenous WASp
(Primary Th1 cells)
# of unique peptides
TOP1
TOP2A
TOP2B
B
CoIP/WB
NIH001-WT (Th1)
MM
input
IP: IgG
IP: WASp
< TOP1
< WASp
C
Sequential Re-ChIP
(3' ends of Th1 genes)
TOP1:WASp
IgG:WASp
% input
NIH001-WT
(Th1)
IFNG
TBX21
RUNX3
IL12RB2
TLR1
D
Single ChIP (TOP1)
(IFNG locus)
p=0.0013 (vs.)
p=0.0033 (vs.)
p=0.46 (vs.)
% input
NIH001-WT(Th1)
NIH002-unTx. (Th1)
NIH002-wt-WASp Tx. (Th1)
E
Single ChIP-qPCR
(3' region of Th1 genes)
unTx
wt-WASp Tx.
WASp
TOP1
S9.6
% input
NIH002-WAS (Th1)

FIG.6



ACCEPTED MANUSCRIPT

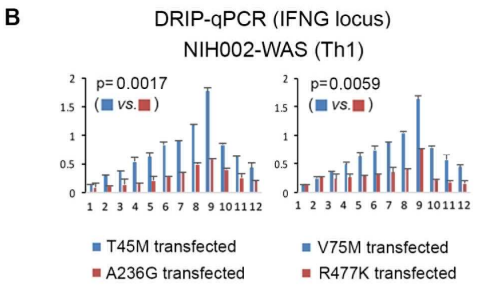
FIG.7

A

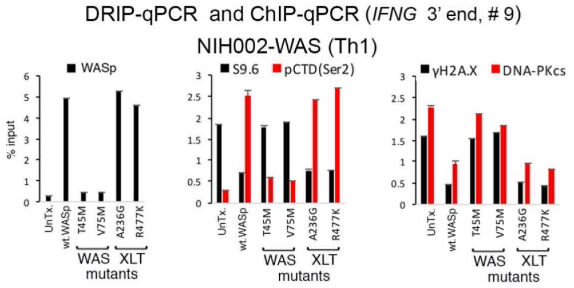
XLT/WAS Th lines

Sample ID	*WAS* Genotype	Clinical scores reported in literature (N=patients)	WASp expression reported in literature	R loop/DSB load
NIH002-T45M	Exon 1, p.T45M	1, 2, 1>5A, 2>5A (n=13)	reduced or absent	high
NIH002-V75M	Exon 2, p.V75M	1, 2, 1>5A, 2>5A (n=22)	normal, reduced, or absent	high
NIH002-A236G	Exon 7, p.A236G	1 (n=1)	absent	low
NIH002-R477K	Exon 11, p.R477K	2 (n=1)	reduced	low

B



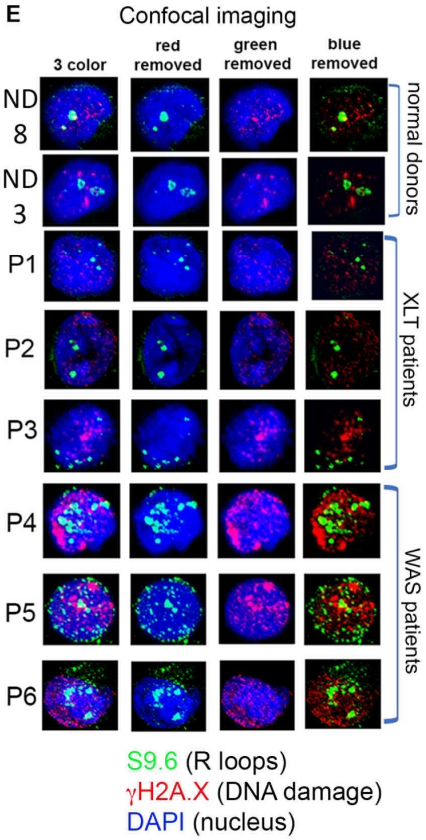
C



D

Sample ID	*WAS* Genotype	clinical score in this patient at sample collection	Clinical scores reported in literature (n=patients)	WASp expression reported in literature	R loop/ DSB load
P1	Exon 2 p.V75M	1	1, 2, 1>5A, 2>5A (n=22)	normal, reduced, or absent	low
P2	Exon 2 p.L39P	1	1, 2, 3, 1>5A/M (n=6)	reduced or absent	low
P3	Exon 2, p.P58R	1	1, 2, 1>5M (n=3)	reduced	low
P4	Exon 6, del570T; fs260stop	5	5 (n=1)	absent	high
P5	Exon 4, del450G; fs167stop	5	5 (n=1)	absent	high
P6	Exon 12 c.1509A>C, no stop	4	4 (n=1)	absent	high
P7 NIH002	Exon1 del23G; fs44stop	5	nd	absent	high
ND1 NIH001	normal donor	-	-	normal	low
ND3	normal donor	-	-	normal	low
ND8	normal donor	-	-	normal	low

E



F

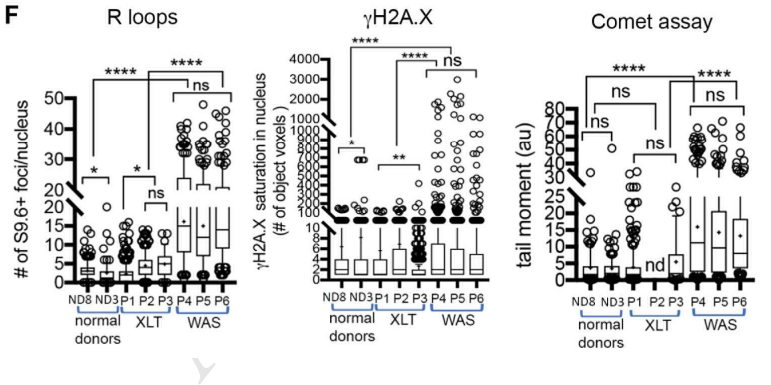
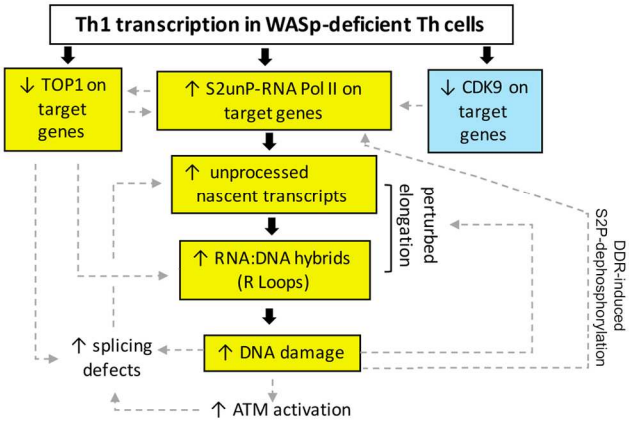


FIG.8



ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

Legends to the Supplementary Figures.

Supplementary Figure 1. Phenotypic characterization of human CD4 Th cells. Flow cytometric analyses of the indicated primary Th cells (ND8) or Th lines (NIH001, Jurkat) stained with a panel of indicated antibodies or their isotype IgG control antibodies are shown as histogram plots. The data tabulated below indicates the Th subtype, i.e., effector, effector memory, naïve etc. based on the combination of cell receptor expression.

Supplementary Figure 2. Experimental depletion of WAS expression in normal (WT) Th cells. (**A**) For CRISPR/Cas9-mediated WAS KO: location of gRNA guides on the *WAS* gene showing primer pairs with the expected amplicon length. (**B**) representative gel images showing clones with PCR amplicon patterns that is consistent with no KO, partial KO, and complete KO of WAS. (**C**) Western blot sequentially reprobed with anti-WASp and α-tubulin loading control antibodies, to demonstrate complete WAS KO clones lacking WASp expression. (**D**) DNA sequencing of complete WAS KO clone show the expected gene region deletion. (**E**) RNA-interference-mediated depletion of WASp expression by shRNA vector. Shown is the western blot image of WASp expression in the indicated primary Th cells.

Supplementary Figure 3. RNaseH1 expression into WT and WKO Th1 cells. Flow cytometry (dot plot) of NIH001-WT and -WKO Th1 lines transfected with RNaseH1-eGFP construct or its empty vector control is shown as pre-sort (transfection efficiency) and post sort (purity) in 3 independent experiments. “X” denotes post-sort sample not tested for purity. The location of the gated population is shown, which was sorted for down-stream assays.

Supplementary Figure 4. R loop modulation in DRB-treated Th1 cells. *Left*, confocal images of WT Th1 cells dual-labelled with S9.6 antibody and DAPI. The image shown is representative of n=100 cells in DRB 100uM for 2 h treatment condition. *Right*, statistics of

S9.6+ foci changes under the indicated DRB dose/duration conditions. The displayed data is from n=100-150 cells analyzed for each condition. p-value: ns= nonsignificant by Mann-Whitney nonparametric test.

Supplementary Figure 5. R-loop and DNA damage load in WASp-deficient Th2-cells. (A) Box-and-whisker plots comparing the magnitude of R-loop (S9.6+ foci) and DSB load (comet assay) in Th1 and Th2 cells, WASp-sufficient (WT) and WASp-deficient (WKO/WKD). The data is representative of analyzing 20-30 z-stack images of at least 100 cells each from 2-3 independent experiments. (**B**) shown are the indicated immunofluorescence images overlaid on the corresponding Nomarski (DIC) images for Th1 and Th2 cells. Right, plot showing the range of individual (single) S9.6+ foci size in Th1 and Th2 cells. ****p<0.0001, Mann-Whitney.

Supplementary Figure 6. Reconstituting nuclear but not cytosolic compartment of WASp-deficient Th1-cells is necessary to normalize R-loop frequency. Change in R-loop frequency in WAS Th1 line stably-transfected with NLS-deleted (ΔNLS) or NES-deleted (ΔNES) WASp mutants. Note, reconstituting only the nuclear compartment with WASp is sufficient to normalize the R-loop frequency (left panel), but reconstituting only the cytosolic compartment does not change the R-loop frequency (right panel). Also shown is the ChIP enrichment profile of γH2A.X. Note, the DNA damage as reported by γH2A.X enrichment is reduced by reconstituting only the nuclear compartment of WASp-deficient Th1 cells. Data is from n=3, mean+SEM.

Supplementary Figure 7. The magnitude of R-loop accumulation and DNA-damage at non-Th1 genes in WASp-deficient Th1 cells. (**A**) R loop frequency at the 3'-ends of the indicated Th1 and non-Th1 genes in WAS Th1 line (NIH002) untransfected (control) or transfected with wt-WASp or the indicated pathogenic mutant. (**B**) DNA damage frequency at the same cis-region in the same set of genes. Data is from n=3, mean+SEM.

ACCEPTED MANUSCRIPT

Supplementary file 1

Antibodies and Reagents

Antibody	**Product number**	**Company name**	**Used for**
AF488 anti-mouse IgG	A21202	Invitrogen	IF, FACS
AF488 anti-rabbit IgG	A21206	Invitrogen	FACS
AF488 anti-mouse IgG	A16119	ThermoFisher	IF
AF568 anti-rabbit IgG	A-11004	ThermoFisher	IF
AF594 anti- rabbit IgG	A11012	Invitrogen	FACS
AF594 anti-mouse IgG	A21203	Invitrogen	FACS
AF647 anti-mouse IgG	A32728	ThermoFisher	IF
Anti-CD28	302902	BioLegend	Cell culture
Anti-CD3	300402	BioLegend	Cell culture
Anti-IL-4	500815	BioLegend	Cell culture
anti-mouse IgG-HRP	sc-2055	Santa Cruz	WB
anti-rabbit IgG-HRP	sc-2054	Santa Cruz	WB
DNAPKcs	Ab18192	Abcam	ChIP
DNAPKcs	sc-5282	Santa Cruz	WB
DNA-RNA Hybrid (S9.6)	ENH001	Kerafast	DRIP, FACS, IF
Flag M2	F1804	Sigma	IP,WB, ChIP
H2A	ab13923	Abcam	FACS
phospho-Histone H2A.X (Ser139)	05-636	Millipore	FACS, ChIP
p-Histone H2A.X (Ser 139)	sc-101696	Santa Cruz	WB, IF
H2A.Z	ab4174	Abcam	FACS
Histone H3	ab1791	Abcam	WB, ChIP, FACS
IgG	sc-66931	Santa Cruz	Gel super-shift
LAMP1	H4A3	DSHB	WB
Normal Mouse IgG	12-371	Upstate	ChIP, FACS
Rabbit Control IgG	Ab46540	Abcam	IP, ChIP, FACS
RNA pol II	05-623	Millipore	ChIP
RNA pol II (Ser2)	Ab5095	Abcam	ChIP
RNA pol II (Ser5)	Ab5131	Abcam	ChIP
Reagent	**Product number**	**Company**	**Used for**
α-Amanitin	A2263	Sigma	Transcription inhibitor
Alexa Fluor™ 568 Phalloidin	A12380	ThermoFisher	IF
Amaxa Cell line nucleofactor kitV	NC9041615	Fisher Scientific	Cell transfection
Agarose, low gelling temperature	A0701	Sigma-Aldrich	Single-Cell Gel Electrophoresis Assay
ChIP-IT Protein G Magnetic Beads	101945	Active Motif	ChIP, IP
CometSlide	4250-200-03	Trevigen	Single-Cell Gel Electrophoresis Assay
Control CRISPR/Cas9 Plasmid	Sc-418922	Santa Cruz	Control for KOs
Cytochalasin D	C8273	Sigma	Actin depolymerization
CytoFix/CytoPerm solutions	554714	BD Pharmingen	For intracellular FACS
DAPI	D3571	Invitrogen	IF
DRB	D1916	Sigma	Transcription inhibitor
Enzymatic shearing cocktail	101197	Active Motif	ChIP
Lysis Solution	4250-050-01	Trevigen	Single-Cell Gel Electrophoresis Assay
High Capacity cDNA Reverse Transcription kit	4368813	Applied Biosystems	cDNA synthesis
Latrunculin B	L5288	Sigma	Actin depolymerization
Myc-DDK-tagged ORF clone of Homo sapiens WAS	RC203457	Origene	Used as mammalian protein expression vector
NE-PER Nuclear & Cytoplasmic Extraction Reagents	78835	Pierce	To extract nuclear and ctyoplasmic proteins

ACCEPTED MANUSCRIPT

Perfecta qPCR SuperMix, LowROX™	95052-500	Quanta Biosciences	RT-PCR
Plasmid Transfection Medium	sc-108062	Santa Cruz	For transfection
Proteinsae K stop solution	Part#101942	Active Motif	ChIP
QuickChange II Site-Directed Mutagenesis (SDM) kit	200519	Stratagene	Site-Directed Mutation

WB: Western Blot
IF: Immunoflourescence
ChIP: Chromatin Immunoprecipitation
IP: Immunoprecipitation
FACS: Fluorescence-activated cell sorting
RT-PCR: Real Time Polymerase Chain Reaction
DRIP: DNA-RNA Immunoprecipitation

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

Supplementary file 2

RT-qPCRPrimers and Probes for mature mRNA and nascent pre-mRNA estimation

Gene (AssayID)	Exon Boundary	Chromosome location
IFNG (Hs99999041_m1)	Ex1-2	Chr.12: 68154770 - 68159741
IFNG (Hs00989290_g1)	Ex2-3	Chr.12: 68154770 - 68159741
IFNG (Hs00174143_m1)	Ex3-4	Chr.12: 68154770 - 68159741
TBX21 (Hs00203436_m1)	Ex1-2	Chr.17: 47733240 - 47746122
TBX21 (Hs00894391_g1)	Ex2-3	Chr.17: 47733240 - 47746122
TBX21 (Hs00894392_m1)	Ex3-4	Chr.17: 47733240 - 47746122
TBX21 (Hs00894393_gH)	Ex4-5	Chr.17: 47733240 - 47746122
TBX21 (Hs00894394_g1)	Ex5-6	Chr.17: 47733240 - 47746122
TUBA1A (Hs04187782_g1)	Ex1-2	Chr.12: 49127782 - 49131521

TUBA1A	Ex2-3	5’-GGCTGTGTTTGTAGACTTGGA-3’
		5’-CTCATCAGGCAAGGAAGAT-3’
		5’-6FAM-TGATGAAGTTCGCACTGGCACCTA-3’
TUBA1A	Ex3-4	5’-CAAGGAAGATGCTGCCAATAAC-3’
		5’-GTCTTCAGGGCTTCTTGGTT-3’
		5’-6FAM-GTCTTCAGGGCTTCTTGGTT-3’
IFNG	5’-UTR	5’-GTTTCCGAGAGAATTAAGCCAAAG-3’
		5’-CATCTGAAGATCAGCTATTAGAAGAGA-3’
		5’-6FAM-TCAAGCTGATCAGGTCCAAAGGACT-3’
IFNG	Ex1-Int1	5’-TTTCAGCTCTGCATCGTTT-3’
		5’-AGCAAGTCGATATTCAGTCATT-3’
		5’-6FAM-ACATATGGGTCCTGGCAGTAACAGC-3’
IFNG	Ex2-Int2	5’-GGTCATTCAGATGTAGCGGATAA-3’
		5’-CCATCAGAAAGCAAGCAACAG-3’
		5’-6FAM-AGAGGTAAGCTGAATATTCCCATTTGGCT-3’
IFNG	Int2-Ex3	5’-TCCTCTGTTTGTGCTCTTTCC-3’
		5’-CTTGATGGTCTCCACACTCTTT-3’
		5’-6FAM-ATGCAGAGCCAAATTGTCTCCTTT-3’

IFNG	Ex3-Int3	5’-AAAGAGTGTGGAGACCATCAAG-3’
		5’-CAAGACCCTCGGCAATGAA-3’
		5’-6FAM-AGCTGACTAATTATTCGGTGAGGCT-3’
IFNG	Int3-Ex4	5’-ACTGACCATCATGACATTAGCA-3’
		5’-CAGTTCAGCCATCACTTGGA-3’
		5’-6FAM-ACTGACTTGAATGTCCAACGCAAAGC-3’
IFNG	Ex4-3’-UTR	5’-TCCAAGTGATGGCTGAACTG-3’
		5’-CAAATATTGCAGGCAGGACAAC-3’
		5’-6FAM-TCGAGGTCGAAGAGCATCCCAGTA-3’
IFNG	3’-UTR-polyA	5’-GTCAGGTGACCCTGATGAAA-3’
		5’-TTGGGTACAGTCACAGTTGTC-3’
		5’-6FAM-TCAGGAGATTTCATGCCTGGTGCT-3’
TBX21	Ex1-Int1	5’-TCCAAGTTTAATCAGCACCAGA-3’
		5’-GACAATGCTTCTGACTCCGT-3’
		5’-6FAM-ATCATCACCAAGCAGGGACGGT-3’
TBX21	Int1-Ex2	5’-TGTCAAGCTGGAGCTGATG-3’
		5’-ACTACAGGATGTTTGTGGACG-3’
		5’-6FAM-TCTCTCTTCTCTCCAGGCGGATGT-3’
TBX21	Ex2-Int2	5’-ATGTTTGTGGACGTGGTCTT-3’
		5’-CTCTGTTTCGCTGGGACTG-3’
		5’-6FAM-GCAAGTGGGTGCAGTGTGGAAAG-3’
TBX21	Int2-Ex3	5’-GCATGTCAAAGAGGTGAACTG-3’
		5’-CATTTGGGAAACTAAAGCTCACA-3’
		5’-6FAM-ACAGGAAACCGCCTGTACGTCC-3’
TBX21	Ex3-Int3	5’-GCGCCAGGAAGTTTCATTTG-3’
		5’-AAAAGGTAGCCTCGCCCT-3’
		5’-6FAM-GCGTCCAACAATGTGACCCAGGTA-3’
TBX21	Int3-Ex4	5’-AACCGTCTTGCTCTGTCTAC-3’

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

		5’-CATATCGTTGAGGTGAACGACG-3’
		5’-6FAM-ATGATTGTGCTCCAGTCCCTCCAT-3’
TBX21	Ex4-Int4	5’-GCTTCCAACACGCATATCTTTAC-3’
		5’-CTTGGGAGGGATTTGGAAA-3’
		5’-6FAM-ATTGCCGTGACTGCCTACCAGAAT-3’
TBX21	Int4-Ex5	5’-GTCTTCCTTGGGAGGGATTTG-3’
		5’-CAAAGGATTCCGGGAGAACTT-3’
		5’-6FAM-CCCAGACTCAGGACTCAGGTGACT-3’
TBX21	Ex5-Int5	5’-AAGGATTCCGGGAGAACTTTG-3’
		5’-CTGAGACACATCCTCTCCCT-3’
		5’-6FAM-GGGTTCAACTCAGCTTTGGTCCCT-3’
TBX21	Int5-Ex6	5’-TCTGCTTGTGACCCGTTT-3’
		5’-CTACCCAACCAGTATCCTGTTC-3’
		5’-6FAM-GCCTGGACCCAACTGTCAATTCCT-3’
TUBA1A	Ex1-Int1	5’-CCTCCTAATCCCTAGCCACTA -3’
		5’-CGTTTGGAGGGCGGAA -3’
		5’-6FAM-TCCCGGCTGCTTTCAGGGAA -3’
TUBA1A	Int1-Ex2	5’-GGCGACAGACCTTATCTCTAAA -3’
		5’-ATCTCCATCCACGTTGGC -3’
		5’-6FAM-TTTCCTTCCCGCAGCGTGAGT -3’
TUBA1A	Ex2-Int2	5’-GGCTGTGTTTGTAGACTTGGAA -3’
		5’-TCACTCACTCACTCCCTCC -3’
		5’-6FAM-CCCACAGTCATTGGTGAGTTGACCT -3’
TUBA1A	Int2-Ex3	5’-CTTCTCTTGTCACTCACTC -3’
		5’-CAAGGAAGATGCTGCCAATAA -3’
		5’-6FAM-CGTCCTTCTCTCCCTCCTCCAGAT -3’
TUBA1A	Ex3-Int3	5’-CTACACCATTGGCAAGGAGAT -3’
		5’-TAGTTCCAGTGCAAGTTGGG -3’
		5’-6FAM-TCGCAAGCTGGTAAGCACCACATA -3’

TUBA1A	Int3-Ex4	5’-CCTTGTGCTTACATCAGTGAGA -3’
		5’-GTCTTCAGGGCTTCTTGGTT -3’
		5’-6FAM-TACAGGCTGACCAGTGCACC -3’

e. *RT-qPCRPrimers and Probes for intron retention*

IFNG	Ex2-Int2	5’-AGATGTAGCGGATAATGGAACTC -3’
		5’-GTTTGTGCTCTTTCCTTCCAAG -3’
		5’-6FAM-TCCTGTTGCTTGCTTTCTGATGGA -3’
TBX21	Ex4-Int3-Ex5	5’-CGTGACTGCCTACCAGAATG -3’
		5’-CTATTTCCCTTCTCTCTAGATTACTCAG -3’
		5’-6FAM-CGTGACTGCCTACCAGAATG -3’
IL13 (Hs01124271_g1)	Ex2-Int2-Ex3-2	Chr.5: 132658173 - 132661109

f. *RT-PCRPrimers for intron retention and sequencing*

IFNG	Ex2-Int2-Ex3 [Forward]	5’-TGCAGGTCATTCAGATGTAGC -3’
	Ex2-Int2-Ex3 [Reverse]	5’-TGACCAGAGCATCCAAAAGA -3’
TBX21	Ex4-Int4-Ex5 [Forward]	5’-GCTGCATATCGTTGAGGTGA -3’
	Ex4-Int4-Ex5 [Reverse]	5’-CAAAGGATTCCGGGAGAACT -3’
TUBA1A	Ex2-Int2-Ex3 [Forward]	5’-CCGGGCTGTGTTTGTAGACT -3’
	Ex2-Int2-Ex3 [Reverse]	5’-ACACCATTGGCAAGGAGATC -3’
IL13	Ex2-Int2-Ex3 [Forward]	5’-CAGCATGGTATGGAGCATCA -3’
	Ex2-Int2-Ex3 [Reverse]	5’-GAAGACCCAGAGGATGCTGA -3’

g.*Site-directed mutagenesis primer sets for generating WASp- mutants*

Primer	**Sequence**
T45M Forward	5'- TGCTTGGACGAAAATGCTTGATGCTGGCCACTG-3'
T45M Reverse	5'- CAGTGGCCAGCATCAAGCATTTTCGTCCAAGCA-3'
V75M Forward	5’-GGGGCTGTGTGCTTCATGAAGGATAACCCCC-3’

ACCEPTED MANUSCRIPT

V75M Reverse	5’-GGGGGTTATCCTTCATGAAGCACACAGCCCC-3’
A236G Forward	5'- AGAAGAAGATCAGCAAAGGTGATATTGGTGCACCC-3'
A236G Reverse	5'- GGGTGCACCAATATCACCTTTGCTGATCTTCTTCT-3'
R477K Forward	5’-CACGTGATGCAGAAGAAAAGCAGAGCCATCCAC-3’
R477K Reverse	5’-GTGGATGGCTCTGCTTTTCTTCTGCATCACGTG-3’

h. *RT-qPCRPrimers and Probes for mRNA estimation*

Gene (AssayID)	**Exon Boundary**	**Chromosome location**
IFNG (Hs00989290_g1)	Ex2-3	Chr.12: 68548550 - 68553521
TBX21 (Hs00894392_m1)	Ex3-4	Chr.17: 45810610 - 45823485
IL12RB2 (Hs02558264_s1)	Ex16-16	Chr.1: 67773047 - 67862583
GAPDH (Hs00266705_g1)	Ex2-3	Chr.12: 6643657 - 6647536

i. *ChIP-qPCRPrimers and Probesbased on NCBI36/hg18(Mar. 2006)*

Genes Names	**Primer/Probes Sequence**
-1kbTSS IFNG (1)	5’-GTTCAAGCCATTCTCCTGCCT-3’
	5’-AAATTAGCCAGGCATGGTGGT-3’
	5’-/56-FAM/CCTCCCTAG/ZEN/TAGCTGAGATTACAGGCA/3IABKFQ/-3’
5’UTR-EXON1 IFNG (2)	5’-CGTAATCCTCAGGAGACTTCAA-3’
	5’-GATCTTTCTCTTCTAATAGCTGATCTTC-3’
	5’-/56-FAM/AAATACCAG/ZEN/CAGCCAGAGGAGGTG/3IABkFQ/-3’
EXON1-INTRON1 IFNG (3)	5’-TTCAGCTCTGCATCGTTT-3’
	5’-GCAAGTCGATATTCAGTCATTT-3’
	5’-/56-FAM/ACATATGGG/ZEN/TCCTGGCAGTAACAGC/3IABkFQ/-3’
Intron1a-intron1a IFNG (4)	5’-ACCTTCTGCTCAGTTTGTATAGAG-3’
	5’-CTGAATGAGTTCCCACCACA-3’
	5’-/56-FAM/AGATGGGCA/ZEN/TAATATGGGTATGAAGCA/3IABkFQ/-3’
Intron1b-intron1b IFNG (5)	5’-GATTTCAACATTCTGCTTCTCTATC-3’
	5’-AAACATTCCTTAGCCAGAATCA-3’
	5’-/56-FAM/CGGAACTTC/ZEN/GTTGCTCACTGGGAT/3IABkFQ/-3’
Intron1c-intron1c IFNG (6)	5’-CACACACACACACACAC-3’
	5’-CATCTACTGTGCCTTCCTGTAG-3’
	5’-/56-FAM/ACACTCGCA/ZEN/CATGTTTGGAACTATCT/3IABkFQ/-3’

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

Intron1c-exon2a IFNG (7)	5’-GGTCATTCAGATGTAGCGGATAA-3’
	5’-CAGAAAGCAAGCAACAGGAAA-3’
	5’-/56-FAM/AGAGGTAAG/ZEN/CTGAATATTCCCATTTGGCT/3IABKFQ/-3’
Intron2a-exon3 IFNG (8)	5’-TCCTCTGTTTGTGCTCTTTCC-3’
	5’-TTTGGATGCTCTGGTCATCTTT-3’
	5’-/56-FAM/ATGCAGAGC/ZEN/CAAATTGTCTCCTTT/3IABkFQ/-3’
Exon4-3’UTR IFNG (9)	5’-GTGTAAGTTCACAACTACTTATGCTG-3’
	5’-TTTCCACAGCTAAGAAGACTCC-3’
	5’-/56-FAM/AGTGAGACC/ZEN/TGGAGTGAAAGAACTACCT/3IABkFQ/-3’
+1.5KB3’UTR IFNG (10)	5’-TCCCTCTCTGAGGTTGAT-3’
	5’-GGATCTTCTCCCTGTCTCATTC-3’
	5’-/56-FAM/TGGCCAGAC/ZEN/TGTGAGACATGAAAGA/3IABkFQ/-3’
+6kb3’UTR IFNG (11)	5’-AGGACAACTAGAAGGGACTCA-3’
	5’-AGAAACAAAGGTCTGCTCTCC-3’
	5’-/56-FAM/AGCCCTCTC/ZEN/ACAGGAGTCCTGGTAA/3IABkFQ/-3’
+3kb3’UTR IFNG (12)	5’-CTTATCATCATGGAGGGCAGAA-3’
	5’-CCTAAAGTCTCCAGAAGGAACAA-3’
	5’-/56-FAM/AAATCAAGG/ZEN/CATCGGCAGGGC/3IABkFQ/-3’
3’TBX21	5’- CTCCAGTCCCTCCATAAGTACCA-3’
	5’- CTCTCCGTCGTTCACCTCAAC-3’
	5’-6FAM- CCCCGGCTGCATAT-3’
3’IL12RB2	5’-GTGGTAGCTCATGCCTGTAATC-3’
	5’-CCATGTTAGCCAGAATGGTCTC-3’
	5’-/56-FAM/AGGCGGATC/Zen/ATGAGGTCAGGAGAT/3IABKFQ/-3’
3’RUNX3	5’-GATGTGTGGAAGGAGTCCAG-3’
	5’-CTGGGCAACAGCAGGAG-3’
	5’-/56-FAM/CCGGTTTCT/Zen/AGGGCTATGCATGGA/3IABkFQ/-3’
3’TLR1	5’-GGATCCACTTGTTTCTTTGCTATAA-3’
	5’-ACTAGAACACACATCACTGAAGAA-3’

ACCEPTED MANUSCRIPT

	5’-/56-FAM/TCATGAAGT/Zen/CAAGGGAAGCACCTGT/3IABkFQ/-3’
3’IL10	5’-CAACAGAAGCTTCCATTCCAAG-3’
	5’-CCAAGCCCAGAGACAAGATAA-3’
	5’-/56-FAM/ACCACGCTT/ZEN/TCTAGCTGTTGAGCT/3IABkFQ/-3’
3’FOXP3	5’-GTCACAATCCTGTCCCTCAC-3’
	5’-GCTGTGCTTGTGTGTGATTG-3’
	5’-/56-FAM/AGCCTGCCT/ZEN/CAGTACACTCAAACA/3IABkFQ/-3’
3’IL17A	5’-GTGCTGTCTTGCCCTACTT-3’
	5’-TGGAGTGAAGAGGAAGGTCT-3’
	5’-/56-FAM/ATATTCTCA/ZEN/GGCCACTCTGCAGGC/3IABKFQ/-3’
3’IL5	5’TTTACTGCAGTGAGAATGAGG-3’
	5’-CTTTCTGGCAAAGTGTCAGTAT-3’
	5’-/56-FAM/ATTAAGGCC/ZEN/TGACTCTTTCTTGGC/3IANkFQ/-3’
3’SNRPN	5’-GAGCAATTAAACTGTGAGGTACTG -3’
	5’-AATCAACTGCTTAATAGGCATCAC-3’
	5’-/56-FAM/TGATGAGAT/Zen/CTTAAGTTACTGTGGATGAGG/3IABkFQ/-3’
3’ ZNF544	5’-GGAGGAGTTGCCATTACACTAC-3’
	5’-CTTGGCAAGATCCCATCCAT-3’
	5’-/56-FAM/AAGATTGGT/Zen/GGACGGGAATGCTGT/3IABkFQ/-3’

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

Human Primary Th1 cells: nuclear fraction (IP with anti-WASp Ab)

Accession	Description	ΣCoverage	Σ# Proteins	Σ# Unique Peptides	Σ# Peptides	Σ# PSMs	# AAs	MW [kDa]
P42768	Wiskott-Aldrich syndrome protein OS=Homo sapiens GN=WAS PE=1 SV=4 - [**WASP_HUMAN**]	10.76	1	3	3	4	502	52.9

Sequence	A3	B3	C3	D3	E3	F3	G3	Modifications	Spectrum File	Charge	m/z [Da]	MH+ [Da]	ΔM [ppm]	RT [min]	PSM Ambiguity	Protein Description	# Proteins	XCorr	# Missed Cleavage	# Protein Groups	tein Group Accessi	earch Engine Rank	q-Value	PEP	Peptides Matched	MS Order	Total Ions
AGISEAQLTDAETSK					High				NP55-75.raw	2	760.87653	1520.74578	1.21	39.41	Unambiguous	Wiskott-Aldrich sy	1	5.41	0	1	P42768	1			565	MS2	28
HVSHVGWDPQNGFDVNNLDPDLR					High				NP55-75.raw	4	658.56573	2631.24111	2.85	59.29	Unambiguous	Wiskott-Aldrich sy	1	4.58	0	1	P42768	1			1967	MS2	132
LIYDFIEDQGGLEAVR					High				NP55-75.raw	3	613.31634	1837.93446	0.72	65.15	Unambiguous	Wiskott-Aldrich sy	1	5.47	0	1	P42768	1			614	MS2	60
LIYDFIEDQGGLEAVR					High				NP55-75.raw	2	919.47296	1837.93865	3.00	65.72	Unambiguous	Wiskott-Aldrich sy	1	6.44	0	1	P42768	1			542	MS2	30

P11387	DNA topoisomerase 1 OS=Homo sapiens GN=TOP1 PE=1 SV=2 - [**TOP1_HUMAN**]	41.70	13	32	32	105	765	90.7

Sequence	A3	B3	C3	D3	E3	F3	G3	Modifications	Spectrum File	Charge	m/z [Da]	MH+ [Da]	ΔM [ppm]	RT [min]	PSM Ambiguity	Protein Description	# Proteins	XCorr	# Missed Cleavage	# Protein Groups	tein Group Accessi	earch Engine Rank	q-Value	PEP	Peptides Matched	MS Order	Total Ions
AEEVATTFAK		High							NP10-40-80.raw	2	556.78571	1112.56413	1.64	48.96	Unambiguous	DNA topoisomerase	1	3.52	0	1	P11387	1	0	0.000228087	402	MS2	18
AEEVATTFAK			High						NP10-150-250.raw	2	556.78660	1112.56592	3.24	49.17	Unambiguous	DNA topoisomerase	1	3.22	0	1	P11387	1	0.000188	0.001373608	388	MS2	18
AEEVATTFAK				Medium					NP55.raw	2	556.78485	1112.56243	0.10	51.57	Unambiguous	DNA topoisomerase	1	3.15	0	1	P11387	1			374	MS2	18
AEEVATTFAK					Medium				NP55-75.raw	2	556.78601	1112.56474	2.19	51.57	Unambiguous	DNA topoisomerase	1	3.14	0	1	P11387	1			332	MS2	18
AEEVATTFAK						Medium			NP100-150.raw	2	556.78589	1112.56450	1.97	51.70	Unambiguous	DNA topoisomerase	1	3.01	0	1	P11387	1			344	MS2	18
AEEVATTFAK						High			NP100-150.raw	2	556.78583	1112.56438	1.86	52.28	Unambiguous	DNA topoisomerase	1	3.33	0	1	P11387	1			399	MS2	18
AGNEKEEGETADTVGccSLR		High						C16(Carbamidome	NP10-40-80.raw	3	728.31763	2182.93833	1.96	29.98	Unambiguous	DNA topoisomerase	1	5.19	1	1	P11387	1	0	9.39974E-07	1897	MS2	76
AGNEKEEGETADTVGccSLR			High					C16(Carbamidome	NP10-150-250.raw	3	728.31915	2182.94291	4.06	30.05	Unambiguous	DNA topoisomerase	1	5.78	1	1	P11387	1	0	2.66608E-06	1886	MS2	76
AGNEKEEGETADTVGccSLR					High			C16(Carbamidome	NP55-75.raw	3	728.31738	2182.93760	1.62	32.20	Unambiguous	DNA topoisomerase	1	6.09	1	1	P11387	1			1913	MS2	76
AGNEKEEGETADTVGccSLR						High		C16(Carbamidome	NP100-150.raw	3	728.31787	2182.93906	2.30	32.35	Unambiguous	DNA topoisomerase	1	6.66	1	1	P11387	1			1893	MS2	76
AVAILcNHQR			High					C6(Carbamidomet	NP10-150-250.raw	2	591.31445	1181.62163	0.67	29.72	Unambiguous	DNA topoisomerase	1	3.67	0	1	P11387	1	0	9.61478E-05	321	MS2	18
AVAILcNHQR			Medium					C6(Carbamidomet	NP10-150-250.raw	3	394.54556	1181.62213	1.10	29.78	Unambiguous	DNA topoisomerase	1	1.39	0	1	P11387	1	0.011373	0.1281087	346	MS2	36
AVAILcNHQR					Medium			C6(Carbamidomet	NP55-75.raw	2	591.31433	1181.62139	0.47	31.76	Unambiguous	DNA topoisomerase	1	3.33	0	1	P11387	1			301	MS2	18
AVAILcNHQR						High		C6(Carbamidomet	NP100-150.raw	2	591.31460	1181.62193	0.93	31.98	Unambiguous	DNA topoisomerase	1	3.65	0	1	P11387	1			315	MS2	18
AVALYFIDK		High							NP10-40-80.raw	2	520.29620	1039.58513	2.72	54.46	Unambiguous	DNA topoisomerase	4	2.30	0	1	P11387	1	0.006975	0.08645469	103	MS2	16
AVALYFIDK			High						NP10-150-250.raw	2	520.29608	1039.58489	2.49	54.57	Unambiguous	DNA topoisomerase	4	2.66	0	1	P11387	1	0.001822	0.02354463	101	MS2	16
AVALYFIDK			High						NP10-150-250.raw	2	520.29614	1039.58501	2.60	55.08	Unambiguous	DNA topoisomerase	4	2.48	0	1	P11387	1	0.003331	0.04353877	104	MS2	16
AVALYFIDKLALR			High						NP10-150-250.raw	3	498.30264	1492.89338	3.14	68.00	Unambiguous	DNA topoisomerase	4	1.32	1	1	P11387	1	0.002072	0.02905497	43	MS2	48
cDFTQmSQYFK			High					C1(Carbamidomet	NP10-150-250.raw	2	735.80688	1470.60649	2.47	42.53	Unambiguous	DNA topoisomerase	1	4.05	0	1	P11387	1	0	0.000036739	479	MS2	20
cDFTQmSQYFK						High		C1(Carbamidomet	NP100-150.raw	2	735.80609	1470.60491	1.39	45.01	Unambiguous	DNA topoisomerase	1	3.99	0	1	P11387	1			519	MS2	20
cDFTQMSQYFK			High					C1(Carbamidomet	NP10-150-250.raw	2	727.80988	1454.61247	3.11	55.86	Unambiguous	DNA topoisomerase	1	4.18	0	1	P11387	1	0	3.74162E-05	582	MS2	20
cDFTQMSQYFK						High		C1(Carbamidomet	NP100-150.raw	2	727.80957	1454.61186	2.69	58.34	Unambiguous	DNA topoisomerase	1	3.71	0	1	P11387	1			573	MS2	20
DEDDADYKPK		Medium							NP10-40-80.raw	3	399.17664	1195.51535	3.31	17.85	Unambiguous	DNA topoisomerase	1	1.69	0	1	P11387	1	0.013438	0.1514699	367	MS2	36
DEDDADYKPK		High							NP10-40-80.raw	3	399.17613	1195.51383	2.04	18.36	Unambiguous	DNA topoisomerase	1	1.74	0	1	P11387	1	0.008095	0.1027617	372	MS2	36
DEDDADYKPK			Medium						NP10-150-250.raw	3	399.17645	1195.51480	2.86	17.52	Unambiguous	DNA topoisomerase	1	2.00	0	1	P11387	1	0.022085	0.2316924	358	MS2	36
DEDDADYKPK			High						NP10-150-250.raw	2	598.26062	1195.51397	2.16	17.68	Unambiguous	DNA topoisomerase	1	3.52	0	1	P11387	1	0.000153	0.00024335	355	MS2	18
DEDDADYKPK			High						NP10-150-250.raw	3	399.17697	1195.51636	4.16	18.03	Unambiguous	DNA topoisomerase	1	2.12	0	1	P11387	1	0.000188	0.001237484	353	MS2	36
DEDDADYKPK						High			NP100-150.raw	2	598.26231	1195.51734	4.97	20.16	Unambiguous	DNA topoisomerase	1	3.62	0	1	P11387	1			363	MS2	18
ELTAPDENIPAK		High							NP10-40-80.raw	2	649.33612	1297.66496	1.14	34.77	Unambiguous	DNA topoisomerase	1	2.84	0	1	P11387	1	0.004396	0.05258103	353	MS2	22
ELTAPDENIPAK			High						NP10-150-250.raw	2	649.33690	1297.66652	2.33	34.97	Unambiguous	DNA topoisomerase	1	3.00	0	1	P11387	1	0.007163	0.09027239	360	MS2	22
ELTAPDENIPAK			High						NP10-150-250.raw	2	649.33661	1297.66594	1.89	35.54	Unambiguous	DNA topoisomerase	1	2.88	0	1	P11387	1	0.000153	0.000304114	292	MS2	22
ELTAPDENIPAK							Medium		NP100-150.raw	2	649.33650	1297.66573	1.72	37.29	Unambiguous	DNA topoisomerase	1	3.02	0	1	P11387	1			318	MS2	22
GPVFAPPYEPLPENVK		High							NP10-40-80.raw	2	877.46528	1753.92328	4.13	55.14	Unambiguous	DNA topoisomerase	1	3.52	0	1	P11387	1	0.000148	0.001644673	328	MS2	30
GPVFAPPYEPLPENVK			High						NP10-150-250.raw	2	877.46399	1753.92070	2.66	55.11	Unambiguous	DNA topoisomerase	1	4.34	0	1	P11387	1	0.000153	0.000269315	328	MS2	30
GPVFAPPYEPLPENVK			High						NP10-150-250.raw	2	877.46356	1753.91985	2.18	55.62	Unambiguous	DNA topoisomerase	1	3.41	0	1	P11387	1	0.000188	0.00128767	344	MS2	30
GPVFAPPYEPLPENVK				High					NP55.raw	2	877.46436	1753.92143	3.08	57.74	Unambiguous	DNA topoisomerase	1	4.02	0	1	P11387	1			324	MS2	30
GPVFAPPYEPLPENVK					High				NP55-75.raw	2	877.46295	1753.91863	1.48	57.51	Unambiguous	DNA topoisomerase	1	4.13	0	1	P11387	1			337	MS2	30
GPVFAPPYEPLPENVK						High			NP100-150.raw	2	877.46362	1753.91996	2.24	57.61	Unambiguous	DNA topoisomerase	1	4.13	0	1	P11387	1			341	MS2	30
GPVFAPPYEPLPENVK						Medium			NP100-150.raw	3	585.31116	1753.91892	1.65	57.69	Unambiguous	DNA topoisomerase	1	3.45	0	1	P11387	1			432	MS2	60
HLQDLmEGLTAK			High					M6(Oxidation)	NP10-150-250.raw	2	686.35112	1371.69496	0.88	36.81	Unambiguous	DNA topoisomerase	1	4.01	0	1	P11387	1	0	2.2121E-05	328	MS2	22
HLQDLmEGLTAK			High					M6(Oxidation)	NP10-150-250.raw	3	457.90356	1371.69614	1.75	36.85	Unambiguous	DNA topoisomerase	1	3.24	0	1	P11387	1	0.000188	0.000557672	401	MS2	44
HLQDLmEGLTAK							High	M6(Oxidation)	NP100-150.raw	2	686.35071	1371.69414	0.29	39.19	Unambiguous	DNA topoisomerase	1	3.72	0	1	P11387	1			315	MS2	22
HLQDLmEGLTAK							High	M6(Oxidation)	NP100-150.raw	3	457.90326	1371.69522	1.08	39.22	Unambiguous	DNA topoisomerase	1	3.82	0	1	P11387	1			401	MS2	44
HLQDLMEGLTAK			High						NP10-150-250.raw	2	678.35510	1355.70293	3.02	50.55	Unambiguous	DNA topoisomerase	1	4.18	0	1	P11387	1	0.000153	0.000432702	334	MS2	22
HLQDLMEGLTAK			High						NP10-150-250.raw	3	452.57285	1355.70398	3.80	50.55	Unambiguous	DNA topoisomerase	1	4.05	0	1	P11387	1	0	2.51394E-05	383	MS2	44
HLQDLMEGLTAK					Medium				NP55-75.raw	3	452.57178	1355.70078	1.44	52.80	Unambiguous	DNA topoisomerase	1	3.68	0	1	P11387	1			386	MS2	44
HLQDLMEGLTAK							High		NP100-150.raw	3	452.57196	1355.70133	1.84	53.01	Unambiguous	DNA topoisomerase	1	3.99	0	1	P11387	1			378	MS2	44
HLQDLMEGLTAK							High		NP100-150.raw	2	678.35443	1355.70159	2.03	53.02	Unambiguous	DNA topoisomerase	1	4.11	0	1	P11387	1			298	MS2	22
ILSYNR			Medium						NP10-150-250.raw	2	383.21679	765.42630	1.20	26.81	Unambiguous	DNA topoisomerase	3	1.60	0	1	P11387	1	0.011886	0.1316277	194	MS2	10
ImPEDIIINcSK			High					M2(Oxidation); C1	NP10-150-250.raw	2	724.86192	1448.71657	2.85	47.90	Unambiguous	DNA topoisomerase	1	3.58	0	1	P11387	1	0.000153	0.000446121	458	MS2	22
ImPEDIIINcSK							High	M2(Oxidation); C1	NP100-150.raw	2	724.86140	1448.71552	2.13	50.35	Unambiguous	DNA topoisomerase	1	3.49	0	1	P11387	1			466	MS2	22
ImPEDIIINcSK							High	M2(Oxidation); C1	NP10-150-150.raw	2	724.86194	1448.71660	2.87	51.11	Unambiguous	DNA topoisomerase	1	3.36	0	1	P11387	1			414	MS2	22
ITVAWcK			High					C6(Carbamidomet	NP10-150-250.raw	2	439.23502	877.46276	3.05	40.47	Unambiguous	DNA topoisomerase	1	2.16	0	1	P11387	1	0.002803	0.03745241	260	MS2	12
KLEEEEDGKLK		High							NP10-40-80.raw	3	439.90167	1317.69046	0.57	20.78	Unambiguous	DNA topoisomerase	1	4.81	2	1	P11387	1	0	9.36281E-06	376	MS2	40
KLEEEEDGKLK			High						NP10-150-250.raw	3	439.90256	1317.69312	2.59	20.77	Unambiguous	DNA topoisomerase	1	4.30	2	1	P11387	1	0	2.79841E-05	401	MS2	40
KLEEEEDGKLK			Medium						NP10-150-250.raw	4	330.17899	1317.69411	3.34	20.85	Unambiguous	DNA topoisomerase	1	2.56	2	1	P11387	1	0.017359	0.187114	395	MS2	60
KLEEEEDGKLK			High						NP10-150-250.raw	3	659.35071	1317.69414	3.36	20.97	Unambiguous	DNA topoisomerase	1	4.01	2	1	P11387	1	0.000658	0.006402711	374	MS2	20
KLEEEEDGKLK				Medium					NP55.raw	3	439.90283	1317.69394	3.21	23.15	Unambiguous	DNA topoisomerase	1	3.80	2	1	P11387	1			393	MS2	40
KLEEEEDGKLK						Medium			NP55-75.raw	3	439.90250	1317.69294	2.45	23.01	Unambiguous	DNA topoisomerase	1	3.69	2	1	P11387	1			383	MS2	40
KLEEEEDGKLK							High		NP100-150.raw	3	439.90332	1317.69541	4.32	22.00	Unambiguous	DNA topoisomerase	1	4.17	2	1	P11387	1			376	MS2	40
KLEEEEDGKLK							High		NP100-150.raw	3	439.90353	1317.69605	4.81	23.16	Unambiguous	DNA topoisomerase	1	4.18	2	1	P11387	1			365	MS2	40
KLEEEEDGKLK							High		NP100-150.raw	2	659.35065	1317.69402	3.27	23.35	Unambiguous	DNA topoisomerase	1	3.77	2	1	P11387	1			357	MS2	20
KLEEEEDGKLK							High		NP100-150.raw	3	439.90292	1317.69422	3.42	23.66	Unambiguous	DNA topoisomerase	1	3.92	2	1	P11387	1			389	MS2	40
KWGVPIEK			Medium						NP10-150-250.raw	3	319.52436	956.55854	2.20	37.60	Unambiguous	DNA topoisomerase	1	1.67	1	1	P11387	1	0.017215	0.1836253	206	MS2	28
LEEQLmK			High					M6(Oxidation)	NP10-150-250.raw	2	453.73370	906.46013	-0.02	24.17	Unambiguous	DNA topoisomerase	1	2.97	0	1	P11387	1	0.000705	0.006555981	263	MS2	12
LEEQLMK		High							NP10-40-80.raw	2	445.73706	890.46684	1.81	32.43	Unambiguous	DNA topoisomerase	1	2.42	0	1	P11387	1	0.004396	0.05155847	270	MS2	12
LEEQLMK			High						NP10-150-250.raw	2	445.73746	890.46764	2.70	32.59	Unambiguous	DNA topoisomerase	1	2.57	0	1	P11387	1	0.001468	0.018648	358	MS2	12
LEVQATDREENK		High							NP10-40-80.raw	3	477.90787	1431.70905	1.10	22.44	Unambiguous	DNA topoisomerase	1	2.93	1	1	P11387	1	0.000919	0.01025861	580	MS2	44
LEVQATDREENK			Medium						NP10-150-250.raw	3	477.90796	1431.70932	1.29	22.36	Unambiguous	DNA topoisomerase	1	2.09	1	1	P11387	1	0.019609	0.2063769	586	MS2	44
LEVQATDREENK			High						NP10-150-250.raw	2	716.35883	1431.71037	2.02	22.44	Unambiguous	DNA topoisomerase	1	3.79	1	1	P11387	1	0.000188	0.001146859	583	MS2	22
LEVQATDREENK							High		NP100-150.raw	2	716.35956	1431.71184	3.04	24.69	Unambiguous	DNA topoisomerase	1	4.00	1	1	P11387	1			580	MS2	22
LNTGILNK			High						NP10-150-250.raw	2	436.76462	872.52196	2.21	34.89	Unambiguous	DNA topoisomerase	1	2.93	0	1	P11387	1	0.000771	0.009392723	194	MS2	14
LNTGILNK					Medium				NP55-75.raw	2	436.76440	872.52153	1.72	36.96	Unambiguous	DNA topoisomerase	1	3.01	0	1	P11387	1			184	MS2	14
mLDHEYYTK			High					M1(Oxidation)	NP10-150-250.raw	3	385.17947	1153.52387	3.84	19.27	Unambiguous	DNA topoisomerase	13	1.96	0	1	P11387	1	0.000564	0.005271852	396	MS2	32
MLDHEYYTK			High						NP10-150-250.raw	2	569.26630	1137.52532	0.70	23.91	Unambiguous	DNA topoisomerase	13	2.37	0	1	P11387	1	0.001468	0.01832227	357	MS2	16
NIITNLSK		High							NP10-40-80.raw	2	451.76984	902.53240	2.00	41.15	Unambiguous	DNA topoisomerase	1	2.45	0	1	P11387	1	0.003323	0.03828087	202	MS2	14
NIITNLSK			Medium						NP10-150-250.raw	2	451.77005	902.53282	2.47	41.35	Unambiguous	DNA topoisomerase	1	1.92	0	1	P11387	1	0.049347	0.4786843	271	MS2	14
NLQLFmENK			High					M6(Oxidation)	NP10-150-250.raw	2	576.79144	1152.57561	3.28	45.38	Unambiguous	DNA topoisomerase	4	2.92	0	1	P11387	1	0.00113	0.01607794	442	MS2	16
NLQLFmENK							Medium	M6(Oxidation)	NP100-150.raw	2	576.79089	1152.57451	2.33	47.94	Unambiguous	DNA topoisomerase	4	3.04	0	1	P11387	1			427	MS2	16
NLQLFMENK			High						NP10-150-250.raw	2	568.79364	1136.58000	2.72	51.36	Unambiguous	DNA topoisomerase	4	3.35	0	1	P11387	1	0.000611	0.006196006	333	MS2	16
NLQLFMENK							High		NP100-150.raw	2	568.79333	1136.57939	2.18	53.75	Unambiguous	DNA topoisomerase	4	3.33	0	1	P11387	1			332	MS2	16
NLQLFmENKQPEDDLFDR			High					M6(Oxidation)	NP10-150-250.raw	3	756.69324	2268.06516	2.17	57.12	Unambiguous	DNA topoisomerase	1	4.05	1	1	P11387	1	0	7.50108E-06	1575	MS2	68
NLQLFmENKQPEDDLFDR						Medium		M6(Oxidation)	NP55-75.raw	3	756.69244	2268.06278	1.12	59.53	Unambiguous	DNA topoisomerase	1	3.57	1	1	P11387	1			1594	MS2	68
NLQLFmENKQPEDDLFDR							High	M6(Oxidation)	NP100-150.raw	3	756.69342	2268.06571	2.41	59.51	Unambiguous	DNA topoisomerase	1	4.53	1	1	P11387	1			1602	MS2	68
NLQLFMENKQPEDDLFDR							High		NP100-150.raw	3	751.36182	2252.07090	2.48	63.18	Unambiguous	DNA topoisomerase	1	4.26	1	1	P11387	1			1433	MS2	68
NLQLFmENKQPEDDLFDRLNTGILNK							Medium	M6(Oxidation)	NP100-150.raw	4	781.14862	3121.57265	3.28	68.16	Unambiguous	DNA topoisomerase	1	3.41	2	1	P11387	1			1240	MS2	150
QIALGTSK			Medium						NP10-150-250.raw	2	409.24292	817.47856	0.90	26.26	Unambiguous	DNA topoisomerase	1	2.26	0	1	P11387	1	0.016546	0.1728516	224	MS2	14
QPEDDLFDR			High						NP10-150-250.raw	2	567.75812	1134.50896	2.39	40.85	Unambiguous	DNA topoisomerase	1	2.02	0	1	P11387	1	0.005544	0.07253719	376	MS2	16
SmMNLQTK			Medium					M2(Oxidation)	NP10-150-250.raw	2	484.73016	968.45305	-1.00	24.07	Unambiguous	DNA topoisomerase	1	2.43	0	1	P11387	2	0.027628	0.2844545	398	MS2	14
SMmNLQTK			Medium					M3(Oxidation)	NP10-150-250.raw	2	484.73016	968.45305	-1.00	24.07	Unambiguous	DNA topoisomerase	1	2.45	0	1	P11387	1	0.031223	0.3147461	398	MS2	14
SMMNLQTK			High						NP10-150-250.raw	2	476.73398	952.46068	1.66	33.54	Unambiguous	DNA topoisomerase	1	2.65	0	1	P11387	1	0.00027	0.003632241	260	MS2	14
TYNASITLQQQLK			High						NP10-150-250.raw	2	754.41199	1507.81670	3.43	44.09	Unambiguous	DNA topoisomerase	1	4.67	0	1	P11387	1	0.000238	0.00157097	265	MS2	24
TYNASITLQQQLK				High					NP55.raw	2	754.41101	1507.81474	2.14	46.50	Unambiguous	DNA topoisomerase	1	4.75	0	1	P11387	1			247	MS2	24
TYNASITLQQQLK					High				NP55-75.raw	2	754.41095	1507.81462	2.06	46.48	Unambiguous	DNA topoisomerase	1	4.70	0	1	P11387	1			254	MS2	24
TYNASITLQQQLK							High		NP100-150.raw	2	754.41138	1507.81548	2.62	46.60	Unambiguous	DNA topoisomerase	1	5.16	0	1	P11387	1			270	MS2	24
TYNASITLQQQLK							High		NP100-150.raw	3	503.27631	1507.81437	1.89	46.60	Unambiguous	DNA topoisomerase	1	3.79	0	1	P11387	1			309	MS2	48
VEHINLHPELDGQEYVVEFDFLGK		High							NP10-40-80.raw	4	707.60706	2827.40639	4.44	67.44	Unambiguous	DNA topoisomerase	1	4.28	0	1	P11387	1	0	9.08352E-06	1285	MS2	138
VEHINLHPELDGQEYVVEFDFLGK			High						NP10-150-250.raw	4	707.60712	2827.40664	4.52	67.52	Unambiguous	DNA topoisomerase	1	3.85	0	1	P11387	1	0.000153	0.000408223	1274	MS2	138
VTWLVSWTENIQQSIK			High						NP10-150-250.raw	2	930.99976	1860.99224	3.62	73.09	Unambiguous	DNA topoisomerase	1	3.80	0	1	P11387	1	0.000369	0.004716888	294	MS2	30
WWEEERYPEGIK			Medium						NP10-150-250.raw	3	541.26203	1621.77153	4.27	46.28	Unambiguous	DNA topoisomerase	1	2.76	1	1	P11387	1	0.024261	0.2534447	774	MS2	44
YImLNPSSR			High					M3(Oxidation)	NP10-150-250.raw	2	548.77765	1096.54802	2.21	35.14	Unambiguous	DNA topoisomerase	1	2.98	0	1	P11387	1	0.000771	0.008829159	442	MS2	16

ACCEPTED MANUSCRIPT

	YImLNPSSR								Medium		M3(Oxidation)	NP100-150.raw	2	548.77740	1096.54753	1.77	37.58	Unambiguous	DNA topoisomera		1	3.00	0	1 P11387	1			467	MS2	16
	YIMLNPSSR			High								NP1P-150-250.raw	2	540.78021	1080.55315	2.28	40.94	Unambiguous	DNA topoisomera		1	3.25	0	1 P11387	1	0.00027	0.002889659	355	MS2	16
	YIMLNPSSR								Medium			NP100-150.raw	2	540.78040	1080.55351	2.62	43.22	Unambiguous	DNA topoisomera		1	2.97	0	1 P11387	1			293	MS2	16
	YYNKVPVEK			Medium								NP1P-150-250.raw	2	570.30841	1139.60954	-0.01	26.48	Unambiguous	DNA topoisomera		1	2.21	1	1 P11387	1	0.039853	0.3945533	335	MS2	16
P11388	DNA topoisomerase 2-alpha OS=Homo sapiens GN=TOP2A PE=1 SV=3 - [**TOP2A_HUMAN**]	31.81	10	35	50	84	1531	174.3																											
	Sequence	A3	B3	C3	D3	E3	F3	G3	Modifications	Spectrum File	Charge	m/z [Da]	MH+ [Da]	ΔM [ppm]	RT [min]	PSM Ambiguity	rotein Description	# Proteins	XCorr	# Missed Cleavage	# Protein Groups	tein Group Accessi	earch Engine Rani	q-Value	PEP	Peptides Matched	MS Order	Total Ions							
	AYDIAGSTK			High						NP1P-150-250.raw	2	463.23575	925.46422	1.78	26.33	Unambiguous	DNA topoisomera	4	2.85	0	1 P11388	1	0.001007	0.01477411	335	MS2	16							
	DFFELR			Medium						NP1P-150-250.raw	2	413.70972	826.41216	3.31	54.23	Unambiguous	DNA topoisomera	4	1.25	0	1 P11388	5	0.029597	0.30172	368	MS2	10							
	DKYGVFPLR			High						NP1P-150-250.raw	3	365.53870	1094.60154	2.00	46.58	Unambiguous	DNA topoisomera	4	2.37	1	1 P11388	1	0.009594	0.1109833	266	MS2	32							
	DPALNSGVSQKPDPAK			High						NP1P-150-250.raw	3	541.94995	1622.83530	0.95	26.88	Unambiguous	DNA topoisomera	4	4.03	0	1 P11388	1	0	2.28307E-06	518	MS2	60							
	EKQDEQVGLPGK			Medium						NP1P-150-250.raw	3	443.23413	1327.68784	1.92	27.16	Unambiguous	DNA topoisomera	4	2.25	1	1 P11388	1	0.015859	0.1687339	375	MS2	44							
	ELILFSNSDNER			High						NP1P-150-250.raw	2	718.85638	1436.70549	2.67	52.41	Unambiguous	Isoform Beta-1 of	7	4.32	0	2 Q02880-2;P11388	1	0	1.42949E-05	475	MS2	22							
	ELILFSNSDNER						High			NP100-150.raw	2	718.85602	1436.70476	2.16	55.02	Unambiguous	Isoform Beta-1 of	7	3.92	0	2 Q02880-2;P11388	1			469	MS2	22							
	ELILFSNSDNER							High		NP200-250.raw	2	718.85764	1436.70801	4.42	55.22	Unambiguous	Isoform Beta-1 of	7	4.51	0	2 Q02880-2;P11388	1			459	MS2	22							
	ENMTLQPK			Medium						NP1P-150-250.raw	2	480.74554	960.48381	1.95	28.22	Selected	Isoform Beta-1 of	7	2.15	0	2 Q02880-2;P11388	1	0.030736	0.3106677	344	MS2	14							
	EQELDTLK			High						NP1P-150-250.raw	2	488.25461	975.50194	2.64	31.11	Unambiguous	DNA topoisomera	4	2.14	0	1 P11388	1	0.001468	0.01853478	336	MS2	14							
	EVTFVPGLYK			Medium						NP1P-150-250.raw	2	576.81976	1152.63225	1.97	52.82	Unambiguous	Isoform Beta-1 of	8	2.27	0	2 Q02880-2;P11388	1	0.023152	0.2432448	178	MS2	18							
	EWLLGMLGAESAK			High						NP1P-150-250.raw	2	702.86566	1404.72405	3.41	69.66	Unambiguous	DNA topoisomera	4	4.32	0	1 P11388	1	0	2.08268E-06	297	MS2	24							
	EWLTNFmEDR			High					M7(Oxidation)	NP1P-150-250.raw	2	678.79974	1356.59221	2.41	52.21	Selected	Isoform Beta-1 of	7	2.68	0	2 Q02880-2;P11388	2	0.000881	0.0132097	560	MS2	18							
	FLEEFTPIVK			High						NP1P-150-250.raw	2	668.38361	1335.75994	3.00	65.65	Unambiguous	DNA topoisomera	4	3.59	0	1 P11388	1	0.00027	0.002953196	148	MS2	20							
	FLEEFTPIVK						High			NP100-150.raw	2	668.38385	1335.76042	3.36	67.90	Unambiguous	DNA topoisomera	4	3.46	0	1 P11388	1			153	MS2	20							
	FLEEFTPIVK							High		NP200-250.raw	2	668.38519	1335.76311	5.37	68.24	Unambiguous	DNA topoisomera	4	3.48	0	1 P11388	1			140	MS2	20							
	FLYDDNQR			High						NP1P-150-250.raw	2	535.74982	1070.49236	2.04	33.99	Unambiguous	Isoform Beta-1 of	7	2.35	0	2 Q02880-2;P11388	1	0.00997	0.1153091	488	MS2	14							
	GESDDFIHDFDSAVAPR			High						NP1P-150-250.raw	3	632.60272	1895.79361	3.71	50.66	Unambiguous	DNA topoisomera	4	3.22	0	1 P11388	1	0.000238	0.001557641	1372	MS2	64							
	GFQQISFVNSIATSK			High						NP1P-150-250.raw	2	813.93011	1626.85295	2.64	56.73	Unambiguous	DNA topoisomera	4	5.08	0	1 P11388	1	0	5.70246E-06	256	MS2	28							
	GFQQISFVNSIATSK			High						NP1P-150-250.raw	3	542.95544	1626.85178	1.92	56.74	Unambiguous	DNA topoisomera	4	3.83	0	1 P11388	1	0	5.97949E-05	395	MS2	56							
	GFQQISFVNSIATSK						High			NP100-150.raw	3	542.95538	1626.85160	1.80	59.26	Unambiguous	DNA topoisomera	4	3.92	0	1 P11388	1			383	MS2	56							
	GYDSDPVK			High						NP1P-150-250.raw	2	440.70675	880.40621	1.70	22.85	Unambiguous	DNA topoisomera	4	2.48	0	1 P11388	1	0.000771	0.008279213	186	MS2	14							
	HVDYVADQIVTK			High						NP1P-150-250.raw	2	694.36469	1387.72209	0.32	36.81	Unambiguous	DNA topoisomera	4	5.25	0	1 P11388	1	0	2.54164E-06	308	MS2	22							
	HVDYVADQIVTK			High						NP1P-150-250.raw	3	463.24594	1387.72327	1.17	36.93	Unambiguous	DNA topoisomera	4	4.08	0	1 P11388	1	0	5.27814E-05	368	MS2	44							
	IFDEILVNAADNK			High						NP1P-150-250.raw	2	731.38513	1461.76299	3.10	55.16	Unambiguous	Isoform Beta-1 of	8	3.95	0	2 Q02880-2;P11388	1	0.000153	0.000347361	361	MS2	24							
	IFDEILVNAADNK			High						NP1P-150-250.raw	3	487.92593	1461.76325	3.27	55.25	Unambiguous	Isoform Beta-1 of	8	3.79	0	2 Q02880-2;P11388	1	0	4.4581E-07	437	MS2	48							
	IFDEILVNAADNK							High		NP100-150.raw	2	731.38489	1461.76250	2.76	57.65	Unambiguous	Isoform Beta-1 of	8	4.25	0	2 Q02880-2;P11388	1			375	MS2	24							
	IFDEILVNAADNK							High		NP100-150.raw	3	487.92542	1461.76169	2.21	57.77	Unambiguous	Isoform Beta-1 of	8	4.15	0	2 Q02880-2;P11388	1			467	MS2	48							
	IFDEILVNAADNK								High	NP200-250.raw	2	731.38617	1461.76506	4.52	57.87	Unambiguous	Isoform Beta-1 of	8	4.15	0	2 Q02880-2;P11388	1			344	MS2	24							
	IFDEILVNAADNK								Medium	NP200-250.raw	3	487.92667	1461.76545	4.78	57.92	Unambiguous	Isoform Beta-1 of	8	3.41	0	2 Q02880-2;P11388	1			418	MS2	48							
	IIIENKPK		Medium							NP1P-40-80.raw	3	318.87143	954.59974	1.51	23.79	Unambiguous	DNA topoisomera	4	1.68	0	1 P11388	1	0.016662	0.1821939	32	MS2	28							
	IIIENKPK			High						NP1P-150-250.raw	2	477.80347	954.59966	1.43	23.71	Unambiguous	DNA topoisomera	4	2.52	0	1 P11388	1	0.000238	0.001583188	32	MS2	14							
	IIIENKPK			High						NP1P-150-250.raw	3	318.87140	954.59964	1.42	23.72	Unambiguous	DNA topoisomera	4	2.15	0	1 P11388	1	0.008102	0.09612764	32	MS2	28							
	IKNENTEGSPQEDGVELEGLK			High						NP1P-150-250.raw	3	762.70984	2286.11496	2.31	41.30	Unambiguous	DNA topoisomera	4	5.64	1	1 P11388	1	0	3.36329E-09	1252	MS2	80							
	ILNVR			Medium						NP1P-150-250.raw	2	307.70337	614.39946	1.64	27.49	Selected	Isoform Beta-1 of	12	1.32	0	2 Q02880-2;P11388	2	0.029567	0.3010378	103	MS2	8							
	ImImTDQDQDGSHIK			High					M2(Oxidation); M	NP1P-150-250.raw	3	588.60425	1763.79819	2.40	29.97	Unambiguous	Isoform Beta-1 of	7	4.73	0	2 Q02880-2;P11388	1	0	4.58423E-08	1127	MS2	56							
	ImImTDQDQDGSHIK							High	M2(Oxidation); M	NP100-150.raw	3	588.60413	1763.79782	2.20	32.20	Unambiguous	Isoform Beta-1 of	7	4.34	0	2 Q02880-2;P11388	1			1135	MS2	56							
	IPNFDVR			High						NP1P-150-250.raw	2	430.73575	860.46422	1.97	42.06	Unambiguous	DNA topoisomera	4	2.04	0	1 P11388	1	0.001848	0.0247008	339	MS2	12							
	KEWLLGMLGAESAK			High						NP1P-150-250.raw	3	511.61038	1532.81659	1.55	62.68	Unambiguous	DNA topoisomera	4	2.51	1	1 P11388	1	0	3.07411E-05	353	MS2	52							
	KEWLLGMLGAESAK							Medium		NP100-150.raw	3	511.61075	1532.81769	2.26	65.21	Unambiguous	DNA topoisomera	4	3.68	1	1 P11388	1			355	MS2	52							
	KEWLTNFmEDR			High					M8(Oxidation)	NP1P-150-250.raw	3	495.56845	1484.69080	4.64	46.09	Unambiguous	Isoform Beta-1 of	7	3.87	1	2 Q02880-2;P11388	1	0	9.61969E-05	662	MS2	40							
	KEWLTNFMEDR			High						NP1P-150-250.raw	3	490.23584	1468.69297	2.70	55.79	Unambiguous	Isoform Beta-1 of	7	4.11	1	2 Q02880-2;P11388	1	0	9.82995E-05	649	MS2	40							
	KQTTLAFKPIK			High						NP1P-150-250.raw	3	425.60028	1274.78629	2.47	35.68	Unambiguous	DNA topoisomera	4	3.19	1	1 P11388	1	0.000188	0.00054361	27	MS2	40							
	KSPSDLWK			High						NP1P-150-250.raw	3	320.84402	960.51752	2.68	34.48	Unambiguous	Isoform Beta-1 of	7	2.75	1	2 Q02880-2;P11388	1	0.009084	0.1049212	329	MS2	28							
	KYDTVLDILR			High						NP1P-150-250.raw	3	412.57236	1235.70252	2.47	58.08	Unambiguous	DNA topoisomera	4	3.39	1	1 P11388	1	0.000238	0.00194727	170	MS2	36							
	KYDTVLDILR			High						NP1P-150-250.raw	2	618.35516	1235.70305	2.90	58.15	Unambiguous	DNA topoisomera	4	2.99	1	1 P11388	1	0.00027	0.003581974	162	MS2	18							
	KYDTVLDILR							Medium		NP100-150.raw	3	412.57248	1235.70288	2.77	60.54	Unambiguous	DNA topoisomera	4	3.65	1	1 P11388	1			169	MS2	36							
	LcNIFSTK			High					C2(Carbamidomet	NP1P-150-250.raw	2	491.75644	982.50560	2.99	44.41	Unambiguous	Isoform Beta-1 of	7	2.73	0	2 Q02880-2;P11388	1	0.00027	0.003062967	389	MS2	14							
	LQTSLTcNSMVLFDHVGcLK			High					C7(Carbamidomet	NP1P-150-250.raw	3	775.04818	2321.12992	2.37	62.51	Unambiguous	DNA topoisomera	4	4.08	0	1 P11388	1	0.000188	0.000835342	1406	MS2	76							
	mQSLDKDIVALMVR			High					M1(Oxidation)	NP1P-150-250.raw	3	545.62628	1634.86429	2.30	61.28	Unambiguous	DNA topoisomera	4	4.04	1	1 P11388	1	0	0.000102303	372	MS2	52							
	mQSLDKDIVALMVR						High		M1(Oxidation)	NP100-150.raw	3	545.62689	1634.86612	3.42	63.80	Unambiguous	DNA topoisomera	4	4.64	1	1 P11388	1			345	MS2	52							
	MQSLDKDIVALmVR			High					M12(Oxidation)	NP1P-150-250.raw	3	545.62634	1634.86448	2.42	53.18	Unambiguous	DNA topoisomera	4	3.58	1	1 P11388	1	0.000188	0.001221367	378	MS2	52							
	MQSLDKDIVALMVR		High							NP1P-40-80.raw	3	540.29523	1618.87113	3.41	64.50	Unambiguous	DNA topoisomera	4	3.39	1	1 P11388	1	0	0.000139047	328	MS2	52							
	MQSLDKDIVALMVR			High						NP1P-150-250.raw	3	540.29504	1618.87058	3.07	64.53	Unambiguous	DNA topoisomera	4	4.64	1	1 P11388	1	0	4.26375E-07	361	MS2	52							
	MQSLDKDIVALMVR						High			NP100-150.raw	3	540.29480	1618.86985	2.62	66.96	Unambiguous	DNA topoisomera	4	4.34	1	1 P11388	1			366	MS2	52							
	NEKEQELDTLKR			High						NP1P-150-250.raw	3	501.59940	1502.78363	1.77	26.55	Unambiguous	DNA topoisomera	4	3.44	2	1 P11388	1	0.000238	0.001976948	444	MS2	44							
	NHmWIFVNALIENPTFDSQTK			High					M3(Oxidation)	NP1P-150-250.raw	3	841.07825	2521.22019	0.82	68.62	Unambiguous	DNA topoisomera	4	4.17	0	1 P11388	1	0	0.00011786	1378	MS2	80							
	NKQEMAPYSLPFEEIKK			High						NP1P-150-250.raw	3	726.00861	2176.01126	2.58	65.63	Unambiguous	DNA topoisomera	4	4.03	1	1 P11388	1	0.000238	0.001556503	1515	MS2	64							
	NSTEcTLILTEGDSAK			High					C5(Carbamidomet	NP1P-150-250.raw	2	869.91455	1738.82182	3.09	43.34	Unambiguous	DNA topoisomera	4	5.33	0	1 P11388	1	0	6.64661E-07	727	MS2	30							
	NYEDEDSLK			High						NP1P-150-250.raw	2	556.74152	1112.47576	1.34	26.75	Unambiguous	DNA topoisomera	4	2.04	0	1 P11388	1	0.000851	0.01147298	274	MS2	16							
	QImENAEINNIIK			High					M3(Oxidation)	NP1P-150-250.raw	2	773.40283	1545.79839	2.71	46.83	Unambiguous	Isoform Beta-1 of	7	4.41	0	2 Q02880-2;P11388	1	0	7.0743E-05	340	MS2	24							
	QImENAEINNIIK							High	M3(Oxidation)	NP100-150.raw	2	773.40240	1545.79753	2.16	49.31	Unambiguous	Isoform Beta-1 of	7	4.26	0	2 Q02880-2;P11388	1			315	MS2	24							
	QIMENAEINNIIK			High						NP1P-150-250.raw	2	765.40552	1529.80376	2.93	51.92	Unambiguous	Isoform Beta-1 of	7	4.40	0	2 Q02880-2;P11388	1	0	2.87498E-06	337	MS2	24							
	QIMENAEINNIIK							High		NP100-150.raw	2	765.40515	1529.80303	2.45	54.46	Unambiguous	Isoform Beta-1 of	7	4.05	0	2 Q02880-2;P11388	1			319	MS2	24							
	QIMENAEINNIIK								High	NP200-250.raw	2	765.40223	1529.79717	-1.38	54.60	Unambiguous	Isoform Beta-1 of	7	3.50	0	2 Q02880-2;P11388	1			376	MS2	24							
	QTTLAFKPIKK			High						NP1P-150-250.raw	4	319.45157	1274.78444	1.02	33.71	Unambiguous	DNA topoisomera	4	2.01	1	1 P11388	1	0.000771	0.008210947	35	MS2	60							
	RAYDIAGSTK			High						NP1P-150-250.raw	2	541.28632	1081.56536	1.54	23.28	Unambiguous	DNA topoisomera	4	3.47	1	1 P11388	1	0.000851	0.01067797	297	MS2	18							
	RDPALNSGVSQKPDPAK		High							NP1P-10-40.raw	3	593.98468	1779.93949	2.59	24.63	Unambiguous	DNA topoisomera	4	3.75	1	1 P11388	1	0	0.000224664	397	MS2	64							
	RDPALNSGVSQKPDPAK			High						NP1P-150-250.raw	3	593.98395	1779.93729	1.35	24.49	Unambiguous	DNA topoisomera	4	3.82	1	1 P11388	1	0	1.13594E-05	400	MS2	64							
	SFGSTcQLSEK			High					C6(Carbamidomet	NP1P-150-250.raw	2	622.28632	1243.56536	2.40	31.01	Unambiguous	DNA topoisomera	4	3.23	0	1 P11388	1	0.000238	0.001642736	481	MS2	20							
	SPSDLWK			High						NP1P-150-250.raw	2	416.71481	832.42235	2.85	38.48	Unambiguous	Isoform Beta-1 of	7	2.40	0	2 Q02880-2;P11388	1	0.001998	0.02822004	231	MS2	12							
	SVVSDLEADDVK			High						NP1P-150-250.raw	2	638.81873	1276.63017	2.67	44.67	Unambiguous	DNA topoisomera	4	4.21	0	1 P11388	1	0	4.85749E-05	430	MS2	22							
	TLAVSGLGVVGR			High						NP1P-150-250.raw	2	564.84149	1128.67571	1.87	48.07	Unambiguous	DNA topoisomera	4	4.54	0	1 P11388	1	0	6.03892E-07	121	MS2	22							
	TLAVSGLGVVGR							High		NP100-150.raw	2	564.84143	1128.67558	1.76	50.68	Unambiguous	DNA topoisomera	4	4.73	0	1 P11388	1			121	MS2	22							
	TLAVSGLGVVGR								High	NP200-250.raw	2	564.84235	1128.67742	3.38	50.79	Unambiguous	DNA topoisomera	4	4.48	0	1 P11388	1			115	MS2	22							
	TPPLITDYR			High						NP1P-150-250.raw	2	538.29419	1075.58110	2.63	42.51	Unambiguous	DNA topoisomera	4	3.06	0	1 P11388	1	0.000188	0.00070818	196	MS2	16							
	VLFTcFK			Medium					C5(Carbamidomet	NP1P-150-250.raw	2	457.74527	914.48326	3.05	49.86	Unambiguous	Isoform Beta-1 of	7	1.63	0	2 Q02880-2;P11388	2	0.026846	0.2786565	167	MS2	12							
	YDTVLDILR			High						NP1P-150-250.raw	2	554.30707	1107.60686	2.14	63.15	Unambiguous	DNA topoisomera	4	3.04	0	1 P11388	1	0.000188	0.000844845	204	MS2	16							
	YGVFPLR			High						NP1P-150-250.raw	2	426.24405	851.48082	3.98	49.88	Unambiguous	Isoform Beta-1 of	7	2.31	0	2 Q02880-2;P11388	1	0.001589	0.02026924	187	MS2	12							
	YIFTmLSSLAR			High					M5(Oxidation)	NP1P-150-250.raw	2	659.34979	1317.69231	3.89	59.55	Unambiguous	DNA topoisomera	4	4.04	0	1 P11388	1	0	8.07416E-06	382	MS2	20							
	YIFTMLSSLAR			High						NP1P-150-250.raw	2	651.35254	1301.69780	4.25	69.63	Unambiguous	DNA topoisomera	4	3.48	0	1 P11388	1	0.000153	0.000264284	338	MS2	20							
	YIFTMLSSLAR							High		NP100-150.raw	2	651.35193	1301.69658	3.31	72.01	Unambiguous	DNA topoisomera	4	3.86	0	1 P11388	1			321	MS2	20							
	YIFTMLSSLAR								High	NP200-250.raw	2	651.35283	1301.69838	4.69	72.23	Unambiguous	DNA topoisomera	4	3.72	0	1 P11388	1			316	MS2	20							
	YSGPEDDAAISLAFSK			High						NP1P-150-250.raw	2	835.90186	1670.79643	3.32	54.13	Unambiguous	DNA topoisomera	4	4.33	0	1 P11388	1	0	0.000172412	660	MS2	30							
Q02880-2	Isoform Beta-1 of DNA topoisomerase 2-beta OS=Homo sapiens GN=TOP2B - [**TOP2B_HUMAN**]	26.03	11	25	41	82	1621	182.5																											
	Sequence	A3	B3	C3	D3	E3	F3	G3	Modifications	Spectrum File	Charge	m/z [Da]	MH+ [Da]	ΔM [ppm]	RT [min]	PSM Ambiguity	rotein Description	# Proteins	XCorr	# Missed Cleavage	# Protein Groups	tein Group Accessi	earch Engine Rani	q-Value	PEP	Peptides Matched	MS Order	Total Ions							
	AGVSVKPFQVK			High						NP1P-150-250.raw	3	387.23444	1159.68875	4.59	38.68	Unambiguous	Isoform Beta-1 of	3	1.77	0	1 Q02880-2	1	0.003116	0.03976772	88	MS2	40							
	DLIQmLVQR							High	M5(Oxidation)	NP200-250.raw	2	566.31524	1131.62321	3.62	60.73	Unambiguous	Isoform Beta-1 of	3	3.26	0	1 Q02880-2	1			267	MS2	16							
	DLIQMLVQR						High			NP100-150.raw	2	558.31702	1115.62676	2.29	68.14	Unambiguous	Isoform Beta-1 of	3	3.74	0	1 Q02880-2	1			235	MS2	16							
	DLIQMLVQR							High		NP200-250.raw	2	558.31866	1115.63005	5.24	68.08	Unambiguous	Isoform Beta-1 of	3	3.55	0	1 Q02880-2	1			195	MS2	16							
	EDLAAPVEELDKVESQER						Medium			NP100-150.raw	3	703.01337	2107.02555	3.07	71.85	Unambiguous	Isoform Beta-1 of	3	3.58	1	1 Q02880-2	1			1101	MS2	68							
	EFFDLR			Medium						NP1P-150-250.raw	2	413.70901	826.41075	1.61	52.16	Selected	Isoform Beta-1 of	4	1.27	0	1 Q02880-2	2	0.034832	0.3465431	400	MS2	10							
	EIVNNVR			High						NP1P-150-250.raw	2	422.23843	843.46959	1.49	23.55	Unambiguous	Isoform Beta-1 of	3	2.18	0	1 Q02880-2	1	0.006329	0.0786464	204	MS2	12							
	ELILFSNSDNER			High						NP1P-150-250.raw	2	718.85638	1436.70549	2.67	52.41	Unambiguous	Isoform Beta-1 of	7	4.32	0	2 Q02880-2;P11388	1	0	1.42949E-05	475	MS2	22							
	ELILFSNSDNER						High			NP100-150.raw	2	718.85602	1436.70476	2.16	55.02	Unambiguous	Isoform Beta-1 of	7	3.92	0	2 Q02880-2;P11388	1			469	MS2	22							
	ELILFSNSDNER							High		NP200-250.raw	2	718.85764	1436.70801	4.42	55.22	Unambiguous	Isoform Beta-1 of	7	4.51	0	2 Q02880-2;P11388	1			459	MS2	22							
	ENMTLQPK			Medium						NP1P-150-250.raw	2	480.74554	960.48381	1.95	28.22	Selected	Isoform Beta-1 of	7	2.15	0	2 Q02880-2;P11388	1	0.030736	0.3106677	344	MS2	14							
	EVTFVPGLYK			Medium						NP1P-150-250.raw	2	576.81976	1152.63225	1.97	52.82	Unambiguous	Isoform Beta-1 of	8	2.27	0	2 Q02880-2;P11388	1	0.023152	0.2432448	178	MS2	18							
	EWLTNFmEDR			High					M7(Oxidation)	NP1P-150-250.raw	2	678.79974	1356.59221	2.41	52.21	Selected	Isoform Beta-1 of	7	2.68	0	2 Q02880-2;P11388	2	0.000881	0.0132097	560	MS2	18							
	EWLVGMLGAESTK							High		NP100-150.raw	2	710.86615	1420.72502	7.65	65.46	Unambiguous	Isoform Beta-1 of	4	3.66	0	1 Q02880-2	1			313	MS2	24							
	FLYDDNQR			High						NP1P-150-250.raw	2	535.74982	1070.49236	2.04	33.99	Unambiguous	Isoform Beta-1 of	7	2.35	0	2 Q02880-2;P11388	1	0.00997	0.1153091	488	MS2	14							
	GFQQISFVNSIATTK			High						NP1P-150-250.raw	3	547.62787	1640.86905	2.89	57.33	Unambiguous	Isoform Beta-1 of	3	3.96	0	1 Q02880-2	1	0	1.9422E-06	358	MS2	56							
	GFQQISFVNSIATTK							High		NP100-150.raw	3	547.62750	1640.86795	2.22	59.68	Unambiguous	Isoform Beta-1 of	3	5.18	0	1 Q02880-2	1			372	MS2	56							
	GFQQISFVNSIATTK							High		NP100-150.raw	2	820.93817	1640.86907	2.90	59.79	Unambiguous

ACCEPTED MANUSCRIPT

HSLEcTLILTEGDSAK		Medium		C5(Carbamidomet	NP100-150.raw	3	591.96265	1773.87339	2.58	48.97	Unambiguous	Isoform Beta-1 of	3	3.47	0	1 Q02880-2	1			722	MS2	60
HVDYVdQVVGK	High				NPIP-150-250.raw	2	679.36017	1357.71306	1.45	40.51	Unambiguous	Isoform Beta-1 of	3	4.94	0	1 Q02880-2	1	0	1.37336E-06	312	MS2	22
HVDYVdQVVGK		High			NP100-150.raw	3	453.24243	1357.71274	1.22	42.87	Unambiguous	Isoform Beta-1 of	3	3.88	0	1 Q02880-2	1			385	MS2	44
HVDYVdQVVGK			High		NP200-250.raw	2	679.36328	1357.71929	6.04	43.21	Unambiguous	Isoform Beta-1 of	3	4.78	0	1 Q02880-2	1			345	MS2	22
HVDYVdQVVGK			High		NP200-250.raw	3	453.24374	1357.71668	4.12	43.22	Unambiguous	Isoform Beta-1 of	3	3.79	0	1 Q02880-2	1			363	MS2	44
IFDEILVNAADNK	High				NPIP-150-250.raw	2	731.38513	1461.76299	3.10	55.16	Unambiguous	Isoform Beta-1 of	8	3.95	0	2 Q02880-2;P11388	1	0.000153	0.000347361	361	MS2	24
IFDEILVNAADNK	High				NPIP-150-250.raw	3	487.92593	1461.76325	3.27	55.25	Unambiguous	Isoform Beta-1 of	8	3.79	0	2 Q02880-2;P11388	1	0	4.4581E-07	437	MS2	48
IFDEILVNAADNK		High			NP100-150.raw	2	731.38489	1461.76250	2.76	57.65	Unambiguous	Isoform Beta-1 of	8	4.25	0	2 Q02880-2;P11388	1			375	MS2	24
IFDEILVNAADNK		High			NP100-150.raw	3	487.92542	1461.76169	2.21	57.77	Unambiguous	Isoform Beta-1 of	8	4.15	0	2 Q02880-2;P11388	1			467	MS2	48
IFDEILVNAADNK			High		NP200-250.raw	2	731.38617	1461.76506	4.52	57.87	Unambiguous	Isoform Beta-1 of	8	4.15	0	2 Q02880-2;P11388	1			344	MS2	24
IFDEILVNAADNK			Medium		NP200-250.raw	3	487.92667	1461.76545	4.78	57.92	Unambiguous	Isoform Beta-1 of	8	3.41	0	2 Q02880-2;P11388	1			418	MS2	48
ILNVR	Medium				NPIP-150-250.raw	2	307.70337	614.39946	1.64	27.49	Selected	Isoform Beta-1 of	12	1.32	0	2 Q02880-2;P11388	2	0.029567	0.3010378	103	MS2	8
ImImTDQDQDGSHIK	High			M2(Oxidation); M	NPIP-150-250.raw	3	588.60425	1763.79819	2.40	29.97	Unambiguous	Isoform Beta-1 of	7	4.73	0	2 Q02880-2;P11388	1	0	4.58423E-08	1127	MS2	56
ImImTDQDQDGSHIK		High		M2(Oxidation); M	NP100-150.raw	3	588.60413	1763.79782	2.20	32.20	Unambiguous	Isoform Beta-1 of	7	4.34	0	2 Q02880-2;P11388	1			1135	MS2	56
KEWLTNFmEDR	High			M8(Oxidation)	NPIP-150-250.raw	3	495.56845	1484.69080	4.64	46.09	Unambiguous	Isoform Beta-1 of	7	3.87	1	2 Q02880-2;P11388	1	0	9.61969E-05	662	MS2	40
KEWLTNFMEDR	High				NPIP-150-250.raw	3	490.23584	1468.69297	2.70	55.79	Unambiguous	Isoform Beta-1 of	7	4.11	1	2 Q02880-2;P11388	1	0	9.82995E-05	649	MS2	40
KSPSDLWK	High				NPIP-150-250.raw	3	320.84402	960.51752	2.68	34.48	Unambiguous	Isoform Beta-1 of	7	2.75	1	2 Q02880-2;P11388	1	0.009084	0.1049212	329	MS2	28
KYETVQDILK	Medium				NPIP-150-250.raw	2	618.84711	1236.68694	2.80	45.03	Unambiguous	Isoform Beta-1 of	4	2.37	1	1 Q02880-2	1	0.041685	0.4132991	220	MS2	18
KYETVQDILK			Medium		NP200-250.raw	3	412.90094	1236.68827	3.88	47.81	Unambiguous	Isoform Beta-1 of	4	3.25	1	1 Q02880-2	1			223	MS2	36
LAQAEAAGLHK	High				NPIP-150-250.raw	3	370.20929	1108.61332	2.10	23.79	Unambiguous	Isoform Beta-1 of	3	3.16	0	1 Q02880-2	1	0.000188	0.001322709	236	MS2	40
LAQAEAAGLHK	High				NPIP-150-250.raw	2	554.80981	1108.61235	1.23	23.84	Unambiguous	Isoform Beta-1 of	3	3.37	0	1 Q02880-2	1	0.000153	0.000446094	244	MS2	20
LAQAEAAGLHK		High			NP100-150.raw	2	554.81061	1108.61394	2.66	25.98	Unambiguous	Isoform Beta-1 of	3	4.01	0	1 Q02880-2	1			229	MS2	20
LcNIFSTK	High			C2(Carbamidomet	NPIP-150-250.raw	2	491.75644	982.50560	2.99	44.41	Unambiguous	Isoform Beta-1 of	7	2.73	0	2 Q02880-2;P11388	1	0.00027	0.003062967	389	MS2	14
LDKDIVALMTR		Medium			NP100-150.raw	3	425.57724	1274.71717	2.67	54.45	Unambiguous	Isoform Beta-1 of	3	3.63	1	1 Q02880-2	1			178	MS2	40
LDKDIVALMTR			High		NP200-250.raw	3	425.57773	1274.71863	3.82	54.65	Unambiguous	Isoform Beta-1 of	3	3.65	1	1 Q02880-2	1			180	MS2	40
LHGLPEQFLYGTATK	High				NPIP-150-250.raw	3	558.96777	1674.88877	2.22	52.52	Unambiguous	Isoform Beta-1 of	3	2.98	0	1 Q02880-2	1	0.000188	0.000885451	396	MS2	56
LIEVVK	Medium				NPIP-150-250.raw	2	350.73447	700.46166	1.81	34.28	Unambiguous	Cytochrome c oxi	7	1.20	0	2 H3BRM5;Q02880-	1	0.034639	0.3454588	97	MS2	10
LIEVVK	Medium				NPIP-150-250.raw	2	350.73477	700.46227	2.68	41.30	Selected	Cytochrome c oxi	7	1.49	0	2 H3BRM5;Q02880-	1	0.015618	0.1641186	105	MS2	10
LLFPAVDDNLLK	High				NPIP-150-250.raw	2	679.39264	1357.77800	3.94	65.93	Unambiguous	Isoform Beta-1 of	3	2.88	0	1 Q02880-2	1	0.000188	0.00146149	142	MS2	22
LLFPAVDDNLLK		Medium			NP100-150.raw	2	679.39209	1357.77690	3.13	68.26	Unambiguous	Isoform Beta-1 of	3	3.01	0	1 Q02880-2	1			123	MS2	22
LLFPAVDDNLLK			Medium		NP200-250.raw	2	679.39331	1357.77934	4.93	68.44	Unambiguous	Isoform Beta-1 of	3	2.75	0	1 Q02880-2	1			137	MS2	22
NKQELSFYSIPEFDEWKK		High			NP100-150.raw	4	572.78790	2288.12978	2.76	63.59	Unambiguous	Isoform Beta-1 of	3	4.30	2	1 Q02880-2	1			1186	MS2	102
NTVEITELPVR		High			NP100-150.raw	2	635.85522	1270.70317	2.34	50.03	Unambiguous	Isoform Beta-1 of	3	3.82	0	1 Q02880-2	1			181	MS2	20
NTVEITELPVR			High		NP200-250.raw	2	635.85632	1270.70537	4.07	50.09	Unambiguous	Isoform Beta-1 of	3	3.63	0	1 Q02880-2	1			155	MS2	20
QELSFYSIPEFDEWKK	High				NPIP-150-250.raw	3	682.66888	2045.99210	3.20	65.18	Unambiguous	Isoform Beta-1 of	3	2.70	1	1 Q02880-2	1	0.00027	0.002485287	1076	MS2	60
QImENAEINNIIK	High			M3(Oxidation)	NPIP-150-250.raw	2	773.40283	1545.79839	2.71	46.83	Unambiguous	Isoform Beta-1 of	7	4.41	0	2 Q02880-2;P11388	1	0	7.0743E-05	340	MS2	24
QImENAEINNIIK		High		M3(Oxidation)	NP100-150.raw	2	773.40240	1545.79753	2.16	49.31	Unambiguous	Isoform Beta-1 of	7	4.26	0	2 Q02880-2;P11388	1			315	MS2	24
QIMENAEINNIIK	High				NPIP-150-250.raw	2	765.40552	1529.80376	2.93	51.92	Unambiguous	Isoform Beta-1 of	7	4.40	0	2 Q02880-2;P11388	1	0	2.87498E-06	337	MS2	24
QIMENAEINNIIK		High			NP100-150.raw	2	765.40515	1529.80303	2.45	54.46	Unambiguous	Isoform Beta-1 of	7	4.05	0	2 Q02880-2;P11388	1			319	MS2	24
QIMENAEINNIIK			High		NP200-250.raw	2	765.40223	1529.79717	-1.38	54.60	Unambiguous	Isoform Beta-1 of	7	3.50	0	2 Q02880-2;P11388	1			376	MS2	24
RLHGLPEQFLYGTATK		High			NP100-150.raw	3	611.00110	1830.98874	1.41	52.14	Unambiguous	Isoform Beta-1 of	3	4.69	1	1 Q02880-2	1			290	MS2	60
SLAVSGLGVTGR		High			NP100-150.raw	2	564.84167	1128.67607	2.19	54.36	Unambiguous	Isoform Beta-1 of	3	4.86	0	1 Q02880-2	1			123	MS2	22
SLAVSGLGVTGR			High		NP200-250.raw	2	564.84253	1128.67778	3.70	54.67	Unambiguous	Isoform Beta-1 of	3	4.75	0	1 Q02880-2	1			133	MS2	22
SPSDLWK	High				NPIP-150-250.raw	2	416.71481	832.42235	2.85	38.48	Unambiguous	Isoform Beta-1 of	7	2.40	0	2 Q02880-2;P11388	1	0.001998	0.02812004	231	MS2	12
SQDFGNLFSPSYSQK	High				NPIP-150-250.raw	2	926.43304	1851.85881	2.13	65.75	Unambiguous	Isoform Beta-1 of	3	5.62	0	1 Q02880-2	1	0	3.24447E-07	1005	MS2	30
SQDFGNLFSPSYSQK		Medium			NP100-150.raw	3	617.95850	1851.86094	3.28	68.21	Unambiguous	Isoform Beta-1 of	3	3.64	0	1 Q02880-2	1			1168	MS2	60
SQDFGNLFSPSYSQK			High		NP200-250.raw	2	926.43628	1851.86528	5.63	68.42	Unambiguous	Isoform Beta-1 of	3	3.50	0	1 Q02880-2	1			930	MS2	30
TPALISDYKEYHTDTTVK		High			NP100-150.raw	4	521.26617	2082.04287	1.84	43.75	Unambiguous	Isoform Beta-1 of	3	4.21	1	1 Q02880-2	1			906	MS2	102
VDHELANER	High				NPIP-150-250.raw	3	360.86502	1080.58051	0.77	24.11	Unambiguous	Isoform Beta-1 of	3	2.95	0	1 Q02880-2	1	0.000188	0.0005673	289	MS2	32
VLFTcFK	Medium			C5(Carbamidomet	NPIP-150-250.raw	2	457.74527	914.48326	3.05	49.86	Unambiguous	Isoform Beta-1 of	7	1.63	0	2 Q02880-2;P11388	2	0.026846	0.2786565	167	MS2	12
YAGPEDDAAITLAFSK	High				NPIP-150-250.raw	2	834.91272	1668.81816	3.92	56.07	Unambiguous	Isoform Beta-1 of	3	4.39	0	1 Q02880-2	1	0	3.40867E-05	663	MS2	30
YAGPEDDAAITLAFSK		High			NP100-150.raw	2	834.91168	1668.81609	2.68	58.51	Unambiguous	Isoform Beta-1 of	3	5.21	0	1 Q02880-2	1			648	MS2	30
YAGPEDDAAITLAFSK			High		NP200-250.raw	2	834.91358	1668.81988	4.95	58.76	Unambiguous	Isoform Beta-1 of	3	5.22	0	1 Q02880-2	1			656	MS2	30
YGVFPLR	High				NPIP-150-250.raw	2	426.24405	851.48082	3.98	49.88	Unambiguous	Isoform Beta-1 of	7	2.31	0	2 Q02880-2;P11388	1	0.001589	0.02026924	187	MS2	12
YIFTmLSTLAR		High		M5(Oxidation)	NP100-150.raw	2	666.35730	1331.70732	3.37	61.71	Unambiguous	Isoform Beta-1 of	3	3.43	0	1 Q02880-2	1			321	MS2	20
YIFTMLSTLAR	High				NPIP-150-250.raw	2	658.35956	1315.71184	2.98	70.07	Unambiguous	Isoform Beta-1 of	3	3.12	0	1 Q02880-2	1	0.001431	0.01804148	290	MS2	20
YIFTMLSTLAR		High			NP100-150.raw	2	658.35919	1315.71111	2.42	72.15	Unambiguous	Isoform Beta-1 of	3	3.39	0	1 Q02880-2	1			319	MS2	20
YIFTMLSTLAR			High		NP200-250.raw	2	658.36118	1315.71509	5.45	72.32	Unambiguous	Isoform Beta-1 of	3	3.36	0	1 Q02880-2	1			315	MS2	20

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

Human Jurkat Th1 cells: nuclear fraction (2 step IP with anti-Flag and -Myc Abs)

P11387	DNA topoisomerase 1 OS=Homo sapiens GN=TOP1 PE=1 SV=2 - [**TOP1**_HUMAN]	31.90	13	25	25	95	765	90.7																								
	Sequence	A3	B3	C3	D3	E3	F3	G3	H3	Modifications	Spectrum File	Charge	m/z [Da]	MH+ [Da]	ΔM [ppm]	RT [min]	PSM Ambiguity	rotein Descriptions	# Proteins	XCorr	r	Missed Cleavage	# Protein Groups	tein Group Accessi	search Engine Rank	q-Value	PEP	Peptides Matched	MS Order	Total Ions		
	AEEVATFFAK			High							Koustav13.raw	2	556.78809	1112.56890	5.92	45.48	Unambiguous	DNA topoisomeras	1	3.73	0	1	P11387	1	0	0.00071099	347	MS2	18		
	AEEVATFFAK					High					Koustav11.raw	2	556.78650	1112.56572	3.07	45.21	Unambiguous	DNA topoisomeras	1	3.36	0	1	P11387	1	0	0.001050806	339	MS2	18		
	AEEVATFFAK						High				Koustav12.raw	2	556.78815	1112.56902	6.03	45.12	Unambiguous	DNA topoisomeras	1	3.22	0	1	P11387	1	0.000346	0.0154077	375	MS2	18		
	AEEVATFFAK							High			Koustav9.raw	2	556.78760	1112.56792	5.04	44.74	Unambiguous	DNA topoisomeras	1	3.14	0	1	P11387	1	0	0.000834045	367	MS2	18		
	AEEVATFFAK								High		Koustav10.raw	2	556.78772	1112.56816	5.26	45.03	Unambiguous	DNA topoisomeras	1	3.28	0	1	P11387	1	0.000241	0.001227148	391	MS2	18		
	AGNEKEEGETADTVGccSLR									C16(Carbamidome	Koustav12.raw	3	728.31796	2182.93933	2.42	25.76	Unambiguous	DNA topoisomeras	1	6.61	1	1	P11387	1	0	1.12309E-07	1904	MS2	76		
	AGNEKEEGETADTVGccSLR									C16(Carbamidome	Koustav9.raw	3	728.31895	2182.94229	3.77	25.37	Unambiguous	DNA topoisomeras	1	5.83	1	1	P11387	1	0	1.44317E-08	1915	MS2	76		
	AGNEKEEGETADTVGccSLR									C16(Carbamidome	Koustav10.raw	3	728.31781	2182.93888	2.21	25.47	Unambiguous	DNA topoisomeras	1	6.79	1	1	P11387	1	0	8.01009E-09	1887	MS2	76		
	AVALYFIDK										Koustav14.raw	2	520.29620	1039.58513	2.72	51.35	Unambiguous	DNA topoisomeras	4	2.40	0	1	P11387	1	0.003911	0.02372156	119	MS2	16		
	AVALYFIDK										Koustav13.raw	2	520.29718	1039.58708	4.60	50.82	Unambiguous	DNA topoisomeras	4	2.36	0	1	P11387	1	0.006325	0.06878856	102	MS2	16		
	AVALYFIDK										Koustav16.raw	2	520.29706	1039.58684	4.37	50.15	Unambiguous	DNA topoisomeras	4	2.57	0	1	P11387	1	0.003308	0.06540293	108	MS2	16		
	AVALYFIDK										Koustav12.raw	2	520.29700	1039.58672	4.25	50.55	Unambiguous	DNA topoisomeras	4	2.36	0	1	P11387	1	0.017113	0.184696	106	MS2	16		
	AVALYFIDK										Koustav9.raw	2	520.29694	1039.58660	4.13	50.35	Unambiguous	DNA topoisomeras	4	2.41	0	1	P11387	1	0.003641	0.03082608	90	MS2	16		
	AVALYFIDK										Koustav10.raw	2	520.29651	1039.58574	3.31	50.49	Unambiguous	DNA topoisomeras	4	2.61	0	1	P11387	1	0.007102	0.08381958	112	MS2	16		
	cDFTQmSQYFK									C1(Carbamidomet	Koustav12.raw	2	735.80872	1470.61016	4.96	38.38	Unambiguous	DNA topoisomeras	1	4.11	0	1	P11387	1	0	0.000880733	487	MS2	20		
	cDFTQmSQYFK									C1(Carbamidomet	Koustav9.raw	2	735.80969	1470.61211	6.29	37.98	Unambiguous	DNA topoisomeras	1	4.32	0	1	P11387	1	0	4.10143E-06	508	MS2	20		
	cDFTQmSQYFK									C1(Carbamidomet	Koustav10.raw	2	735.80865	1470.61003	4.88	38.23	Unambiguous	DNA topoisomeras	1	4.41	0	1	P11387	1	0.000133	0.00049193	466	MS2	20		
	DEDDADYKPK		High								Koustav14.raw	2	598.26239	1195.51750	5.11	14.16	Unambiguous	DNA topoisomeras	1	2.63	0	1	P11387	1	0.000999	0.002927681	357	MS2	18		
	DEDDADYKPK							High			Koustav12.raw	3	399.17657	1195.51517	3.16	13.27	Unambiguous	DNA topoisomeras	1	1.88	0	1	P11387	1	0.001481	0.03934259	325	MS2	36		
	ELTAPDENIPAK		High								Koustav14.raw	2	649.33807	1297.66887	4.15	30.42	Unambiguous	DNA topoisomeras	1	3.17	0	1	P11387	1	0.000185	2.72467E-05	317	MS2	22		
	ELTAPDENIPAK			High							Koustav13.raw	2	649.33691	1297.66655	2.36	31.09	Unambiguous	DNA topoisomeras	1	2.97	0	1	P11387	1	0.000969	0.01558234	349	MS2	22		
	ELTAPDENIPAK							High			Koustav12.raw	2	649.33917	1297.67107	5.84	30.47	Unambiguous	DNA topoisomeras	1	3.22	0	1	P11387	1	0	0.007897976	316	MS2	22		
	ELTAPDENIPAK							High			Koustav12.raw	2	649.33942	1297.67156	6.22	31.06	Unambiguous	DNA topoisomeras	1	3.20	0	1	P11387	1	0.000346	0.01424153	275	MS2	22		
	ELTAPDENIPAK								Medium		Koustav9.raw	2	649.33942	1297.67156	6.22	30.07	Unambiguous	DNA topoisomeras	1	2.87	0	1	P11387	1	0.03046	0.3362047	321	MS2	22		
	ELTAPDENIPAK							High			Koustav9.raw	2	649.33984	1297.67241	6.87	30.57	Unambiguous	DNA topoisomeras	1	3.19	0	1	P11387	1	0	5.0813E-05	294	MS2	22		
	ELTAPDENIPAK								High		Koustav10.raw	2	649.33875	1297.67021	5.18	30.39	Unambiguous	DNA topoisomeras	1	3.33	0	1	P11387	1	0.000292	0.00158149	321	MS2	22		
	FAWAIDmADEDYEF							High		M7(Oxidation)	Koustav12.raw	2	869.85540	1738.70352	5.35	65.98	Unambiguous	DNA topoisomeras	1	3.28	0	1	P11387	1	0.001481	0.04072538	768	MS2	26		
	GPVFAPPYEPLPENVK			High							Koustav13.raw	2	877.46478	1753.92229	3.57	51.63	Unambiguous	DNA topoisomeras	1	3.20	0	1	P11387	1	0.000136	0.002557263	347	MS2	30		
	GPVFAPPYEPLPENVK							High			Koustav12.raw	2	877.46625	1753.92522	5.24	51.23	Unambiguous	DNA topoisomeras	1	3.52	0	1	P11387	1	0	0.003314717	363	MS2	30		
	GPVFAPPYEPLPENVK								High		Koustav9.raw	2	877.46621	1753.92515	5.20	51.23	Unambiguous	DNA topoisomeras	1	3.43	0	1	P11387	1	0	0.000192006	339	MS2	30		
	GPVFAPPYEPLPENVK										Koustav10.raw	2	877.46515	1753.92302	3.99	51.23	Unambiguous	DNA topoisomeras	1	4.15	0	1	P11387	1	0	6.18329E-06	356	MS2	30		
	HLQDLMEGLTAK							High			Koustav12.raw	2	678.35718	1355.70708	6.08	46.71	Unambiguous	DNA topoisomeras	1	4.02	0	1	P11387	1	0	0.000788038	347	MS2	22		
	HLQDLMEGLTAK							High			Koustav12.raw	3	452.57324	1355.70517	4.68	46.73	Unambiguous	DNA topoisomeras	1	2.71	0	1	P11387	1	0	0.001519891	371	MS2	44		
	HLQDLMEGLTAK							High			Koustav9.raw	2	678.35645	1355.70561	5.00	46.37	Unambiguous	DNA topoisomeras	1	3.36	0	1	P11387	1	0	0.000191508	302	MS2	22		
	ImPEDIIINcSK							High		M2(Oxidation); C1	Koustav12.raw	2	724.86353	1448.71977	5.06	43.76	Unambiguous	DNA topoisomeras	1	3.63	0	1	P11387	1	0	0.002011497	366	MS2	22		
	ImPEDIIINcSK							High		M2(Oxidation); C1	Koustav12.raw	2	724.86383	1448.72038	5.48	44.42	Unambiguous	DNA topoisomeras	1	3.44	0	1	P11387	1	0	0.004865129	441	MS2	22		
	ImPEDIIINcSK							High		M2(Oxidation); C1	Koustav9.raw	2	724.86366	1448.72005	5.25	43.46	Unambiguous	DNA topoisomeras	1	3.21	0	1	P11387	1	0	0.000183114	396	MS2	22		
	ImPEDIIINcSK								High	M2(Oxidation); C1	Koustav10.raw	2	724.86340	1448.71953	4.89	43.71	Unambiguous	DNA topoisomeras	1	3.95	0	1	P11387	1	0.000241	0.001193007	382	MS2	22		
	ITVAWcK		Medium							C6(Carbamidomet	Koustav14.raw	2	439.23508	877.46288	3.19	36.21	Unambiguous	DNA topoisomeras	1	2.13	0	1	P11387	1	0.012911	0.2183865	314	MS2	12		
	ITVAWcK			High						C6(Carbamidomet	Koustav13.raw	2	439.23547	877.46367	4.10	36.63	Unambiguous	DNA topoisomeras	1	2.25	0	1	P11387	1	0.004033	0.0491492	292	MS2	12		
	ITVAWcK					Medium				C6(Carbamidomet	Koustav11.raw	2	439.23590	877.46453	5.07	36.46	Unambiguous	DNA topoisomeras	1	2.07	0	1	P11387	1	0.024567	0.2633476	293	MS2	12		
	ITVAWcK								Medium	C6(Carbamidomet	Koustav9.raw	2	439.23437	877.46145	1.57	35.90	Unambiguous	DNA topoisomeras	1	2.10	0	1	P11387	1	0.023194	0.2649287	290	MS2	12		
	ITVAWcK								High	C6(Carbamidomet	Koustav10.raw	2	439.23410	877.46092	0.97	36.08	Unambiguous	DNA topoisomeras	1	2.05	0	1	P11387	1	0.00845	0.1051064	323	MS2	12		
	KLEEEEDGKLK							High			Koustav12.raw	3	439.90356	1317.69614	4.88	15.83	Unambiguous	DNA topoisomeras	1	4.21	2	1	P11387	1	0.009905	0.02255671	375	MS2	40		
	KLEEEEDGKLK								High		Koustav9.raw	3	439.90329	1317.69532	4.25	16.32	Unambiguous	DNA topoisomeras	1	4.28	2	1	P11387	1	0	3.39818E-06	375	MS2	40		
	KLEEEEDGKLK										Koustav10.raw	3	439.90268	1317.69349	2.86	16.27	Unambiguous	DNA topoisomeras	1	4.16	2	1	P11387	1	0	0.000299388	396	MS2	40		
	LEEQLmK		High							M6(Oxidation)	Koustav14.raw	2	453.73523	906.46318	3.35	19.45	Unambiguous	DNA topoisomeras	1	2.64	0	1	P11387	1	0.001301	0.008049621	306	MS2	12		
	LEEQLmK			High						M6(Oxidation)	Koustav13.raw	2	453.73544	906.46361	3.82	19.68	Unambiguous	DNA topoisomeras	1	2.62	0	1	P11387	2	0.006325	0.06853069	317	MS2	12		
	LEEQLmK							High		M6(Oxidation)	Koustav12.raw	2	453.73595	906.46462	4.94	18.84	Unambiguous	DNA topoisomeras	1	2.72	0	1	P11387	1	0.003538	0.06848163	314	MS2	12		
	LEEQLmK								High	M6(Oxidation)	Koustav9.raw	2	453.73574	906.46420	4.47	19.15	Unambiguous	DNA topoisomeras	1	2.76	0	1	P11387	1	0.000579	0.01697035	308	MS2	12		
	LEEQLmK									M6(Oxidation)	Koustav10.raw	2	453.73529	906.46330	3.48	19.28	Unambiguous	DNA topoisomeras	1	2.99	0	1	P11387	1	0.001301	0.0209173	314	MS2	12		
	LEEQLMK								High		Koustav10.raw	2	445.73816	890.46904	4.28	27.98	Unambiguous	DNA topoisomeras	1	2.55	0	1	P11387	1	0.005805	0.06995433	356	MS2	12		
	LEVQATDR							High			Koustav12.raw	2	466.24833	931.48938	5.37	19.17	Unambiguous	DNA topoisomeras	1	2.68	0	1	P11387	1	0.002539	0.05462238	259	MS2	14		
	LEVQATDR								Medium		Koustav9.raw	2	466.24185	931.47643	-8.54	19.39	Unambiguous	DNA topoisomeras	1	2.36	0	1	P11387	1	0.046953	0.4080544	286	MS2	14		
	LEVQATDREENK	Medium									Koustav15.raw	3	477.90857	1431.71115	2.57	17.72	Unambiguous	DNA topoisomeras	1	1.79	1	1	P11387	1	0.012597	0.1565068	606	MS2	44		
	LEVQATDREENK		Medium								Koustav14.raw	3	477.90921	1431.71308	3.91	17.92	Unambiguous	DNA topoisomeras	1	2.00	1	1	P11387	1	0.032689	0.3000182	634	MS2	44		
	LEVQATDREENK							High			Koustav12.raw	2	716.36157	1431.71587	5.86	17.40	Unambiguous	DNA topoisomeras	1	3.42	1	1	P11387	1	0	0.001145393	553	MS2	22		
	LEVQATDREENK								High		Koustav9.raw	3	477.90842	1431.71070	2.25	17.68	Unambiguous	DNA topoisomeras	1	2.21	1	1	P11387	1	0.001121	0.02131889	607	MS2	44		
	LEVQATDREENK								High		Koustav9.raw	2	716.36090	1431.71453	4.92	17.69	Unambiguous	DNA topoisomeras	1	3.11	1	1	P11387	1	0	0.00015314	562	MS2	22		
	LEVQATDREENK									High	Koustav10.raw	3	477.90851	1431.71097	2.44	17.66	Unambiguous	DNA topoisomeras	1	2.11	1	1	P11387	1	0	0.000120766	601	MS2	44		
	LEVQATDREENK									High	Koustav10.raw	2	716.36090	1431.71453	4.92	17.72	Unambiguous	DNA topoisomeras	1	3.91	1	1	P11387	1	0	0.000195796	563	MS2	22		
	LNTGILNK							High			Koustav12.raw	2	436.76584	872.52440	5.01	30.35	Unambiguous	DNA topoisomeras	1	2.79	0	1	P11387	1	0.002279	0.0505782	326	MS2	14		
	LNTGILNK								High		Koustav9.raw	2	436.76553	872.52379	4.31	30.11	Unambiguous	DNA topoisomeras	1	2.67	0	1	P11387	1	0.002713	0.02780445	328	MS2	14		
	LNTGILNK									High	Koustav10.raw	2	436.76532	872.52336	3.82	30.29	Unambiguous	DNA topoisomeras	1	2.63	0	1	P11387	1	0.001301	0.02065668	330	MS2	14		
	mLDHEYTTK							Medium		M1(Oxidation)	Koustav12.raw	3	385.17966	1153.52442	4.32	14.35	Unambiguous	DNA topoisomeras	13	1.54	0	1	P11387	2	0.024386	0.2357938	393	MS2	32		
	mLDHEYTTK							Medium		M1(Oxidation)	Koustav12.raw	2	577.26459	1153.52190	2.13	14.52	Unambiguous	DNA topoisomeras	13	1.92	0	1	P11387	1	0.024022	0.233175	380	MS2	16		
	mLDHEYTTK									High M1(Oxidation)	Koustav10.raw	3	385.17996	1153.52533	5.11	15.05	Unambiguous	DNA topoisomeras	13	1.83	0	1	P11387	1	0.000546	0.006410636	431	MS2	32		
	MLDHEYTTK									High	Koustav10.raw	3	379.84818	1137.52997	4.79	19.20	Unambiguous	DNA topoisomeras	13	1.47	0	1	P11387	1	0.000292	0.004427471	436	MS2	32		
	NIITNLSK				Medium						Koustav16.raw	2	451.77158	902.53588	5.85	36.95	Unambiguous	DNA topoisomeras	1	2.12	0	1	P11387	1	0.03286	0.3527897	243	MS2	14		
	NIITNLSK					High					Koustav11.raw	2	451.77121	902.53514	5.04	37.21	Unambiguous	DNA topoisomeras	1	2.35	0	1	P11387	1	0	0.001455678	246	MS2	14		
	NIITNLSK							High			Koustav12.raw	2	451.77151	902.53575	5.72	37.13	Unambiguous	DNA topoisomeras	1	1.80	0	1	P11387	1	0.001481	0.0410709	232	MS2	14		
	NIITNLSK								High		Koustav10.raw	2	451.77139	902.53551	5.45	36.93	Unambiguous	DNA topoisomeras	1	1.83	0	1	P11387	1	0.001242	0.01867165	229	MS2	14		
	NLQLFmENK					High				M6(Oxidation)	Koustav11.raw	2	576.79266	1152.57805	5.40	41.45	Unambiguous	DNA topoisomeras	4	3.13	0	1	P11387	1	0.00067	0.008325797	419	MS2	16		
	NLQLFmENK							High		M6(Oxidation)	Koustav12.raw	2	576.79291	1152.57854	5.82	41.30	Unambiguous	DNA topoisomeras	4	3.23	0	1	P11387	1	0.000905	0.02847901	400	MS2	16		
	NLQLFmENK								High	M6(Oxidation)	Koustav9.raw	2	576.79321	1152.57915	6.35	40.96	Unambiguous	DNA topoisomeras	4	3.11	0	1	P11387	1	0.000304	0.000834045	390	MS2	16		
	NLQLFmENK									High M6(Oxidation)	Koustav10.raw	2	576.79248	1152.57768	5.08	41.05	Unambiguous	DNA topoisomeras	4	3.23	0	1	P11387	1	0.000292	0.003528611	376	MS2	16		
	NLQLFMENK					High					Koustav11.raw	2	568.79370	1136.58013	2.83	47.32	Unambiguous	DNA topoisomeras	4	2.63	0	1	P11387	1	0.009703	0.1139736	348	MS2	16		
	NLQLFMENK							High			Koustav12.raw	2	568.79486	1136.58245	4.87	47.24	Unambiguous	DNA topoisomeras	4	3.31	0	1	P11387	1	0.000346	0.01236356	269	MS2	16		
	NLQLFMENK								High		Koustav9.raw	2	568.79382	1136.58037	3.04	47.09	Unambiguous	DNA topoisomeras	4	1.87	0	1	P11387	3	0.002468	0.02780445	349	MS2	16		
	NLQLFmENKQPEDDLFDR							High			Koustav12.raw	3	756.69537	2268.07157	5.00	53.36	Unambiguous	DNA topoisomeras	1	3.31	1	1	P11387	1	0	0.00024374	1461	MS2	68		
	NLQLFmENKQPEDDLFDR								High	M6(Oxidation)	Koustav10.raw	3	756.69611	2268.07376	5.96	53.36	Unambiguous	DNA topoisomeras	1	3.67	1	1	P11387	2	0	2.40683E-05	1418	MS2	68		
	QIALGTSK							Medium			Koustav12.raw	2	409.24405	817.48082	3.66	21.81	Unambiguous	DNA topoisomeras	1	2.02	0	1	P11387	1	0.018109	0.190603	187	MS2	14		
	QPEDDLFDR		High								Koustav14.raw	2	567.75934	1134.51140	4.54	36.52	Unambiguous	DNA topoisomeras	4	2.34	0	1	P11387	1	0.001072	0.002927681	369	MS2	16		
	QPEDDLFDR					Medium					Koustav11.raw	2	567.75952	1134.51177	4.86	36.67	Unambiguous	DNA topoisomeras	4	2.47	0	1	P11387	1	0.010289	0.1361433	405	MS2	16		
	QPEDDLFDR							Medium			Koustav12.raw	2	567.75958	1134.51189	4.97	36.84	Unambiguous	DNA topoisomeras	1	1.97	0	1	P11387	1	0.014145	0.1622454	368	MS2	16		
	QPEDDLFDR								High		Koustav9.raw	2	567.75934	1134.51140	4.54	36.32	Unambiguous	DNA topoisomeras	1	2.08	0	1	P11387	1	0.002713	0.02780445	367	MS2	16		
	QPEDDLFDR									High	Koustav10.raw	2	567.75952	1134.51177	4.86	36.51	Unambiguous	DNA topoisomeras	1	1.98	0	1	P11387	1	0.001301	0.0222	399	MS2	16		
	SMMNLQTK							High			Koustav12.raw	2	476.73523	952.46318	4.28	28.98	Unambiguous	DNA topoisomeras	1	2.82	0	1	P11387	1	0.000346	0.01495418	243	MS2	14		
	SMMNLQTK								High		Koustav9.raw	2	476.73561	952.46394	5.08	28.57	Unambiguous	DNA topoisomeras	1	2.69	0	1	P11387	1	0.000304	0.001817355	293	MS2	14		
	TYNASITLQQQLK		High								Koustav14.raw	2	754.41150	1507.81572	2.79	40.22	Unambiguous	DNA topoisomeras	1	4.26	0	1	P11387	1	0.000185	5.09368E-05	278	MS2	24		
	TYNASITLQQQLK							High			Koustav12.raw	2	754.41370	1507.82012	5.70	40.10	Unambiguous	DNA topoisomeras	1	5.08	0	1	P11387	1	0	0.006841181	254	MS2	24		
	YImLNPSSR							High		M3(Oxidation)	Koustav12.raw	2	548.77893	1096.55058	4.55	30.67	Unambiguous	DNA topoisomeras	1	2.59	0	1	P11387	1	0.000346	0.01206872	396	MS2	16		
	YImLNPSSR								High	M3(Oxidation)	Koustav9.raw	2	548.77923	1096.55118	5.09	30.36	Unambiguous	DNA topoisomeras	1	2.72	0	1	P11387	1	0	0.000143792	396	MS2	16		
	YImLNPSSR									High M3(Oxidation)	Koustav10.raw	2	548.77917	1096.55107	4.99	30.58	Unambiguous	DNA topoisomeras	1	2.89	0	1	P11387	1	0.000133	0.000539702	287	MS2	16		
	YIMLNPSSR							High			Koustav12.raw	2	540.77985	1080.55242	1.61	36.41	Unambiguous	DNA topoisomeras	1	2.76	0	1	P11387	1	0	0.001784177	319	MS2	16		

P11388	DNA topoisomerase 2-alpha OS=Homo sapiens GN=TOP2A PE=1 SV=3 - [**TOP2A**_HUMAN]	53.49	11	72	88	235	1531	174.3																								
	Sequence	A3	B3	C3	D3	E3	F3	G3	H3	Modifications	Spectrum File	Charge	m/z [Da]	MH+ [Da]	ΔM [ppm]	RT [min]	PSM Ambiguity	rotein Descriptions	# Proteins	XCorr	r	Missed Cleavage	# Protein Groups	tein Group Accessi	search Engine Rank	q-Value	PEP	Peptides Matched	MS Order	Total Ions		
	AGEmELKPFNGEDYTcITFQPDLSK								High	M4(Oxidation); C1	Koustav10.raw	3	969.45270	2906.34354	7.27	53.99	Unambiguous	DNA topoisomeras	4	5.68	0	1	P11388	1	0	0.000315238	3403	MS2	96		
	AYDIAGSTK			Medium							Koustav13.raw	2	463.23788	925.46849	6.40	22.00	Unambiguous	DNA topoisomeras	4	2.89	0	1	P11388	1	0.021232	0.2513607	481	MS2	16		
	AYDIAGSTK								High		Koustav10.raw	2	463.23724	925.46721	5.01	21.08	Unambiguous	DNA topoisomeras	4	2.68	0	1	P11388	1	0.007993	0.100893	451	MS2	16		
	AYDIAGSTK								High		Koustav10.raw	2	463.23697	925.46666	4.42	21.64	Unambiguous	DNA topoisomeras	4	2.90	0	1	P11388	1	0.005412	0.06749795	403	MS2	16		
	DFFELR								High		Koustav10.raw	2	413.70993	826.41258	3.83	49.95	Unambiguous	DNA topoisomeras	4	1.36	0	1	P11388	1	0.009419	0.1135237	340	MS2	10		
	DFFELR								Medium		Koustav10.raw	2	413.70996	826.41265	3.90	50.49	Unambiguous	DNA topoisomeras	4	1.14	0	1	P11388	1	0.020446	0.2452351	388	MS2	10		
	DIVALMVR								High		Koustav10.raw	2	458.76993	916.53258	4.44	50.55	Unambiguous	DNA topoisomeras	4	3.04	0	1	P11388	1	0.000292	0.003979569	209	MS2	14		
	DKLDETGNSLK								High		Koustav10.raw	3	407.21197	1219.62135	3.95	20.29	Unambiguous	DNA topoisomeras	4	3.24	1	1	P11388	1	0.00858	0.1058416	389	MS2	40		
	DKYGVFPLR				Medium						Koustav16.raw	3	365.53906	1094.60263	3.00	42.34	Unambiguous	DNA topoisomeras	4	2.54	1	1	P11388	1	0.027918	0.3045012	257	MS2	32		
	DKYGVFPLR								Medium		Koustav10.raw	3	365.53976	1094.60474	4.92	42.16	Unambiguous	DNA topoisomeras	4	2.54	1	1	P11388	1	0.021891	0.261396	259	MS2	32		
	DKYGVFPLR								High		Koustav10.raw	2	547.80518	1094.60307	3.40	42.21	Unambiguous	DNA topoisomeras	4	2.85	1	1	P11388	1	0.001963	0.0308754	260	MS2	16		
	DKYGVFPLR								High		Koustav10.raw	3	365.53955	1094.60410	4.34	42.70	Unambiguous	DNA topoisomeras	4	2.76	1	1	P11388	1	0.003985	0.05403274	250	MS2	32		
	DPALNSGVSQKPDPAK	High									Koustav15.raw	3	541.95203	1623.84153	4.78	22.34	Unambiguous	DNA topoisomeras	4	3.94	0	1	P11388	1	0	1.70166E-06	486	MS2	60		
	DPALNSGVSQKPDPAK		High								Koustav14.raw	3	541.95166	1623.84043

ACCEPTED MANUSCRIPT

| | | |

ACCEPTED MANUSCRIPT

| | | |

ACCEPTED MANUSCRIPT

| | | |

ACCEPTED MANUSCRIPT

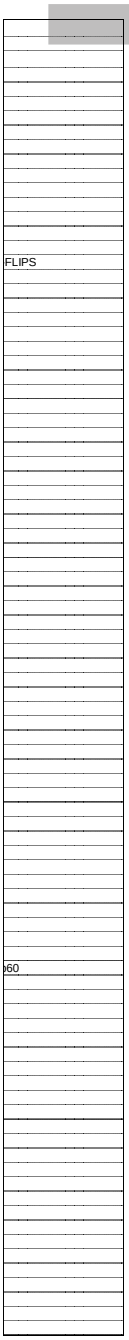
TPALISDYK					High		Koustav10.raw	2	504.27603	1007.54479	3.94	34.95	Unambiguous	Isoform Beta-1 of	3	3.47	0	1	Q02880-2	1	0.000241	0.001125244	224	MS2	16
TPALISDYKEYHTDTTVK					High		Koustav10.raw	3	694.68793	2082.04923	4.90	37.05	Unambiguous	Isoform Beta-1 of	3	4.18	1	1	Q02880-2	1	0	6.01871E-07	817	MS2	68
TWTQVYK					Medium		Koustav10.raw	2	463.24435	925.48142	3.90	30.99	Unambiguous	Isoform Beta-1 of	3	2.02	0	1	Q02880-2	1	0.014272	0.1782489	435	MS2	12
VIHELANER		High					Koustav16.raw	3	360.86707	1080.58664	6.44	19.67	Unambiguous	Isoform Beta-1 of	3	2.96	0	1	Q02880-2	1	0.000529	0.009874636	293	MS2	32
VIHELANER					High		Koustav10.raw	3	360.86624	1080.58417	4.15	19.37	Unambiguous	Isoform Beta-1 of	3	3.00	0	1	Q02880-2	1	0	0.000120842	296	MS2	32
VIHELANER					High		Koustav10.raw	2	540.79523	1080.58318	3.24	19.41	Unambiguous	Isoform Beta-1 of	3	2.81	0	1	Q02880-2	1	0.000292	0.001973292	277	MS2	16
VLFTcFK					Medium	C5(Carbamidomet	Koustav10.raw	2	457.74637	914.48546	5.45	45.53	Unambiguous	Isoform Beta-1 of	7	1.71	0	2	Q02880-2;P11388	1	0.025164	0.2968945	155	MS2	12
VSIDPESNIISIWNNGK					High		Koustav10.raw	2	943.49341	1885.97954	7.44	62.98	Unambiguous	Isoform Beta-1 of	3	5.45	0	1	Q02880-2	1	0	1.4007E-07	356	MS2	32
VSIDPESNIISIWNNGK					High		Koustav10.raw	3	629.33087	1885.97806	6.66	63.10	Unambiguous	Isoform Beta-1 of	3	4.39	0	1	Q02880-2	1	0.000292	0.001459657	464	MS2	64
VYVPALIFGQLLTSSNYDDDEKK					High		Koustav10.raw	3	872.45093	2615.33823	5.43	72.77	Unambiguous	Isoform Beta-1 of	3	1.84	1	1	Q02880-2	1	0	4.78276E-06	567	MS2	88
YAGPEDDAAITLAFSK			High				Koustav11.raw	2	834.91382	1668.82036	5.24	52.37	Unambiguous	Isoform Beta-1 of	3	3.50	0	1	Q02880-2	1	0	5.30864E-08	615	MS2	30
YAGPEDDAAITLAFSK					High		Koustav10.raw	2	834.91357	1668.81987	4.94	51.93	Unambiguous	Isoform Beta-1 of	3	6.04	0	1	Q02880-2	1	0	1.14137E-07	611	MS2	30
YAGPEDDAAITLAFSK					High		Koustav10.raw	3	556.94507	1668.82065	5.41	52.05	Unambiguous	Isoform Beta-1 of	3	3.79	0	1	Q02880-2	1	0	5.88193E-05	810	MS2	60
YETVQDILK					High		Koustav10.raw	2	554.80048	1108.59368	4.67	45.54	Unambiguous	Isoform Beta-1 of	4	3.03	0	1	Q02880-2	1	0.000292	0.004304536	268	MS2	16
YGVFPLR					High		Koustav10.raw	2	426.24435	851.48143	4.70	45.56	Unambiguous	Isoform Beta-1 of	7	2.30	0	2	Q02880-2;P11388	1	0.000936	0.01378671	185	MS2	12
YGVFPLR					High		Koustav10.raw	2	426.24411	851.48094	4.12	46.06	Unambiguous	Isoform Beta-1 of	7	2.28	0	2	Q02880-2;P11388	1	0.001301	0.01925877	168	MS2	12
YIFTmLSTLAR					High	M5(Oxidation)	Koustav10.raw	2	666.35846	1331.70964	5.11	55.37	Unambiguous	Isoform Beta-1 of	3	3.97	0	1	Q02880-2	1	0	0.000285482	327	MS2	20
YIFTmLSTLAR					High	M5(Oxidation)	Koustav10.raw	2	666.35864	1331.71001	5.39	55.94	Unambiguous	Isoform Beta-1 of	3	3.77	0	1	Q02880-2	1	0	1.96081E-05	322	MS2	20
YIFTMLSTLAR				High			Koustav12.raw	2	658.36267	1315.71807	7.71	66.29	Unambiguous	Isoform Beta-1 of	3	3.10	0	1	Q02880-2	1	0.000905	0.01977624	292	MS2	20
YIFTMLSTLAR					High		Koustav10.raw	2	658.36169	1315.71611	6.23	66.83	Unambiguous	Isoform Beta-1 of	3	3.51	0	1	Q02880-2	1	0.000783	0.0104029	296	MS2	20

ACCEPTED MANUSCRIPT

PCR Array Catalog #:	Qiagen	PAHS-025Y			
Position	**Unigene**	**Refseq**	**Symbol**	**Description**	**Gname**
A01	Hs.19383	NM_000029	AGT	Angiotensinogen (serpin peptidase inhibitor, clade A	ANHU/SERPINA8
A02	Hs.525622	NM_005163	AKT1	V-akt murine thymoma viral oncogene homolog 1	AKT/CWS6/PKB/PKB-ALPHA/PRKBA/RAC/RAC-ALPHA
A03	Hs.648565	NM_005171	ATF1	Activating transcription factor 1	EWS-ATF1/FUS/ATF-1/TREB36
A04	Hs.227817	NM_004049	BCL2A1	BCL2-related protein A1	ACC-1/ACC-2/ACC1/ACC2/BCL2L5/BFL1/GRS/HBPA1
A05	Hs.516966	NM_138578	BCL2L1	BCL2-like 1	BCL-XL/S/BCL2L/BCLX/BCLXL/BCLXS/Bcl-X/PPP1R52/bcl-xL/bcl-xS
A06	Hs.31210	NM_005178	BCL3	B-cell CLL/lymphoma 3	BCL4/D19S37
A07	Hs.696238	NM_001166	BIRC2	Baculoviral IAP repeat containing 2	API1/HIAP2/Hiap-2/MIHB/RNF48/c-IAP1/cIAP1
A08	Hs.648101	NM_032415	CARD11	Caspase recruitment domain family, member 11	BENTA/BIMP3/CARMA1/IMD11/PPBL
A09	Hs.2490	NM_033292	CASP1	Caspase 1, apoptosis-related cysteine peptidase (i	ICE/IL1BC/P45
A10	Hs.599762	NM_001228	CASP8	Caspase 8, apoptosis-related cysteine peptidase	ALPS2B/CAP4/Casp-8/FLICE/MACH/MCH5
A11	Hs.303649	NM_002982	CCL2	Chemokine (C-C motif) ligand 2	GDCF-2/HC11/HSMCR30/MCAF/MCP-1/MCP1/SCYA2/SMC-CF
A12	Hs.514821	NM_002985	CCL5	Chemokine (C-C motif) ligand 5	D17S136E/RANTES/SCYA5/SIS-delta/SISd/TCP228/eoCP
B01	Hs.472860	NM_001250	CD40	CD40 molecule, TNF receptor superfamily member	Bp50/CDW40/TNFRSF5/p50
B02	Hs.390736	NM_003879	CFLAR	CASP8 and FADD-like apoptosis regulator	CASH/CASP8AP1/CLARP/Casper/FLAME/FLAME-1/FLAME1/FLIP/I-FLICE/MRIT/c-FLIP/c-FLIPL/c-FLIPR/c-
B03	Hs.198998	NM_001278	CHUK	Conserved helix-loop-helix ubiquitous kinase	IKBKA/IKK-alpha/IKK1/IKKA/NFKBIKA/TCF16
B04	Hs.173894	NM_000757	CSF1	Colony stimulating factor 1 (macrophage)	CSF-1/MCSF
B05	Hs.1349	NM_000758	CSF2	Colony stimulating factor 2 (granulocyte-macrophag	GMCSF
B06	Hs.2233	NM_000759	CSF3	Colony stimulating factor 3 (granulocyte)	C17orf33/CSF3OS/GCSF
B07	Hs.488293	NM_005228	EGFR	Epidermal growth factor receptor	ERBB/ERBB1/HER1/NISBD2/PIG61/mENA
B08	Hs.326035	NM_001964	EGR1	Early growth response 1	AT225/G0S30/KROX-24/NGFI-A/TIS8/ZIF-268/ZNF225
B09	Hs.181128	NM_005229	ELK1	ELK1, member of ETS oncogene family	-
B10	Hs.86131	NM_003824	FADD	Fas (TNFRSF6)-associated via death domain	GIG3/MORT1
B11	Hs.2007	NM_000639	FASLG	Fas ligand (TNF superfamily, member 6)	ALPS1B/APT1LG1/APTL/CD178/CD95-L/CD95L/FASL/TNFSF6
B12	Hs.25647	NM_005252	FOS	FBJ murine osteosarcoma viral oncogene homolog	AP-1/C-FOS/p55
C01	Hs.727017	NM_002133	HMOX1	Heme oxygenase (decycling) 1	HMOX1D/HO-1/HSP32/bK286B10
C02	Hs.37026	NM_024013	IFNA1	Interferon, alpha 1	IFL/IFN/IFN-ALPHA/IFN-alphaD/IFNA13/IFNA@
C03	Hs.856	NM_000619	IFNG	Interferon, gamma	IFG/IFI
C04	Hs.597664	NM_001556	IKBKB	Inhibitor of kappa light polypeptide gene enhancer i	IKK-beta/IKK2/IKKB/IMD15/NFKBIKB
C05	Hs.43505	NM_003639	IKBKG	Inhibitor of kappa light polypeptide gene enhancer i	AMCBX1/FIP-3/FIP3/Fip3p/IKK-gamma/IMD33/IP/IP1/IP2/IPD2/NEMO/ZC2HC9
C06	Hs.193717	NM_000572	IL10	Interleukin 10	CSIF/GVHDS/IL-10/IL10A/TGIF
C07	Hs.1722	NM_000575	IL1A	Interleukin 1, alpha	IL-1A/IL1/IL1-ALPHA/IL1F1
C08	Hs.126256	NM_000576	IL1B	Interleukin 1, beta	IL-1/IL1-BETA/IL1F2
C09	Hs.701982	NM_000877	IL1R1	Interleukin 1 receptor, type I	CD121A/D2S1473/IL-1R-alpha/IL1R/IL1RA/P80
C10	Hs.522819	NM_001569	IRAK1	Interleukin-1 receptor-associated kinase 1	IRAK/pelle
C11	Hs.449207	NM_001570	IRAK2	Interleukin-1 receptor-associated kinase 2	IRAK-2
C12	Hs.696684	NM_002228	JUN	Jun proto-oncogene	AP-1/AP1/c-Jun
D01	Hs.36	NM_000595	LTA	Lymphotoxin alpha (TNF superfamily, member 1)	LT/TNFB/TNFSF1
D02	Hs.1116	NM_002342	LTBR	Lymphotoxin beta receptor (TNFR superfamily, mei	CD18/D12S370/LT-BETA-R/TNF-R-III/TNFCR/TNFR-RP/TNFR2-RP/TNFR3/TNFRSF3
D03	Hs.601217	NM_173844	MALT1	Mucosa associated lymphoid tissue lymphoma tran	IMD12/MLT/MLT1
D04	Hs.653654	NM_005921	MAP3K1	Mitogen-activated protein kinase kinase kinase 1	MAPKKK1/MEKK/MEKK 1/MEKK1/SRXY6
D05	Hs.82116	NM_002468	MYD88	Myeloid differentiation primary response gene (88)	MYD88D
D06	Hs.73090	NM_002502	NFKB2	Nuclear factor of kappa light polypeptide gene enha	CVID10/H2TF1/LYT-10/LYT10/NF-kB2/p100/p52
D07	Hs.738731	NM_006092	NOD1	Nucleotide-binding oligomerization domain containir	CARD4/CLR7.1/NLRC1
D08	Hs.658434	NM_021144	PSIP1	PC4 and SFRS1 interacting protein 1	DFS70/LEDGF/PAIP/PSIP2/p52/p75
D09	Hs.159130	NM_002880	RAF1	V-raf-1 murine leukemia viral oncogene homolog 1	CMD1NN/CRAF/NS5/Raf-1/c-Raf
D10	Hs.633256	NM_002908	REL	V-rel reticuloendotheliosis viral oncogene homolog (	C-Rel
D11	Hs.502875	NM_021975	RELA	V-rel reticuloendotheliosis viral oncogene homolog /	NFKB3/p65
D12	Hs.654402	NM_006509	RELB	V-rel reticuloendotheliosis viral oncogene homolog I	I-REL/IREL/REL-B
E01	Hs.247077	NM_001664	RHOA	Ras homolog gene family, member A	ARH12/ARHA/RHO12/RHOH12
E02	Hs.519842	NM_003804	RIPK1	Receptor (TNFRSF)-interacting serine-threonine kir	RIP/RIP1
E03	Hs.642990	NM_007315	STAT1	Signal transducer and activator of transcription 1, 9	CANDF7/IMD31A/IMD31B/IMD31C/ISGF-3/STAT91
E04	Hs.505874	NM_013254	TBK1	TANK-binding kinase 1	NAK/T2K
E05	Hs.29344	NM_182919	TICAM1	Toll-like receptor adaptor molecule 1	IIAE6/MyD88-3/PRVTIRB/TICAM-1/TRIF
E06	Hs.522632	NM_003254	TIMP1	TIMP metallopeptidase inhibitor 1	CLGI/EPA/EPO/HCI/TIMP
E07	Hs.654532	NM_003263	TLR1	Toll-like receptor 1	CD281/TIL/TIL. LPRS5/rsc786
E08	Hs.519033	NM_003264	TLR2	Toll-like receptor 2	CD282/TIL4
E09	Hs.657724	NM_003265	TLR3	Toll-like receptor 3	CD283/IIAE2
E10	Hs.174312	NM_138554	TLR4	Toll-like receptor 4	ARMD10/CD284/TLR-4/TOLL
E11	Hs.575090	NM_006068	TLR6	Toll-like receptor 6	CD286
E12	Hs.87968	NM_017442	TLR9	Toll-like receptor 9	CD289
F01	Hs.591834	NM_003844	TNFRSF10A	Tumor necrosis factor receptor superfamily, memb	APO2/CD261/DR4/TRAILR-1/TRAILR1
F02	Hs.661668	NM_003842	TNFRSF10B	Tumor necrosis factor receptor superfamily, memb	CD262/DR5/KILLER/KILLER/DR5/TRAIL-R2/TRAILR2/TRICK2/TRICK2A/TRICK2B/TRICKB/ZTNFR9
F03	Hs.713833	NM_001065	TNFRSF1A	Tumor necrosis factor receptor superfamily, memb	CD120a/FPF/MS5/TBP1/TNF-R/TNF-R-I/TNF-R55/TNFAR/TNFR1/TNFR1-d2/TNFR55/TNFR60/p55/p55-R/
F04	Hs.478275	NM_003810	TNFSF10	Tumor necrosis factor (ligand) superfamily, membe	APO2L/Apo-2L/CD253/TL2/TRAIL
F05	Hs.460996	NM_003789	TRADD	TNFRSF1A-associated via death domain	Hs.89862
F06	Hs.522506	NM_021138	TRAF2	TNF receptor-associated factor 2	MGC:45012/TRAP/TRAP3
F07	Hs.510528	NM_003300	TRAF3	TNF receptor-associated factor 3	CAP-1/CAP1/CD40bp/CRAF1/IIAE5/LAP1
F08	Hs.591983	NM_004620	TRAF6	TNF receptor-associated factor 6	MGC:3310/RNF85
F09	Hs.127799	NM_001165	BIRC3	Baculoviral IAP repeat containing 3	AIP1/API2/CIAP2/HAIP1/HIAP1/MALT2/MIHC/RNF49/c-IAP2
F10	Hs.75498	NM_004591	CCL20	Chemokine (C-C motif) ligand 20	CKb4/Exodus/LARC/MIP-3-alpha/MIP-3a/MIP3A/SCYA20/ST38
F11	Hs.595133	NM_004233	CD83	CD83 molecule	BL11/HB15
F12	Hs.75765	NM_002089	CXCL2	Chemokine (C-X-C motif) ligand 2	CINC-2a/GRO2/GROb/MGSA-b/MIP-2a/MIP2/MIP2A/SCYB2
G01	Hs.89690	NM_002090	CXCL3	Chemokine (C-X-C motif) ligand 3	CINC-2b/GRO3/GROg/MIP-2b/MIP2B/SCYB3
G02	Hs.643447	NM_000201	ICAM1	Intercellular adhesion molecule 1	BB2/CD54/P3.58
G03	Hs.624	NM_000584	CXCL8	Interleukin 8	GCP-1/GCP1/IL8/LECT/LUCT/LYNAP/MDNCF/MONAP/NAF/NAP-1/NAP1
G04	Hs.436061	NM_002198	IRF1	Interferon regulatory factor 1	IRF-1/MAR
G05	Hs.618430	NM_003998	NFKB1	Nuclear factor of kappa light polypeptide gene enha	EBP-1/KBF1/NF-kB1/NF-kappa-B/NF-kappaB/NFKB-p105/NFKB-p50/NFkappaB/p105/p50
G06	Hs.81328	NM_020529	NFKBIA	Nuclear factor of kappa light polypeptide gene enha	IKBA/MAD-3/NFKBI
G07	Hs.458276	NM_004556	NFKBIE	Nuclear factor of kappa light polypeptide gene enha	IKBE
G08	Hs.118958	NM_003764	STX11	Syntaxin 11	FHL4/HLH4/HPLH4
G09	Hs.310640	NM_052864	TIFA	TRAF-interacting protein with forkhead-associated d	T2BP/T6BP/TIFAA
G10	Hs.241570	NM_000594	TNF	Tumor necrosis factor	DIF/TNF-alpha/TNFA/TNFSF2
G11	Hs.525607	NM_006291	TNFAIP2	Tumor necrosis factor, alpha-induced protein 2	B94/EXOC3L3
G12	Hs.211600	NM_006290	TNFAIP3	Tumor necrosis factor, alpha-induced protein 3	A20/OTUD7C/TNFA1P2
control genes					
H01	Hs.520640	NM_001101	ACTB	Actin, beta	BRWS1/PS1TP5BP1
H02	Hs.546285	NM_001002	RPLP0	Ribosomal protein, large, P0	L10E/LP0/P0/PRLP0/RPP0

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT



ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

		Up-Down Regulation (comparing to control group: WAS Th1 cells; untransfected)								
		Group 1 (WAS Th1; wt-WASp transfected		Group 2 (WAS Th1; RNaseH1 transfected						
		Fold Regulation	Comments	Fold Regulation	Comments					
A01	AGT	96.687	A	70.2779	A					
A02	AKT1	23.9798		22.8227						
A03	ATF1	34.2148		21.8885						
A04	BCL2A1	47.5393		38.245						
A05	BCL2L1	-1.1013		-1.7589						
A06	BCL3	14.0194		10.8521						
A07	BIRC2	26.2264		14.998						
A08	CARD11	9.8468		8.6231						
A09	CASP1	92.8835		57.6081						
A10	CASP8	13.0697		9.8485						
A11	CCL2	-1.4907		-17.5366						
A12	CCL5	3.2151		2.5394						
B01	CD40	3.2523		2.3953						
B02	CFLAR	4.9814		5.0312						
B03	CHUK	26.241		17.8704						
B04	CSF1	4.3962		2.9506						
B05	CSF2	-51.9824		-77.1396						
B06	CSF3	1993.8622	A	30.19	A					
B07	EGFR	63.2218	A	64.2712	A					
B08	EGR1	2.2583		-1.0898						
B09	ELK1	10.1877		4.9157						
B10	FADD	34.5724		22.4229						
B11	FASLG	1157.6232	A	812.1486	A					
B12	FOS	29.0779	A	20.9576	A					
C01	HMOX1	55.5668	A	43.0965	A					
C02	IFNA1	319.318	A	200.0477	A					
C03	IFNG	66.6831		50.0258						
C04	IKBKB	4.1654		3.6933						
C05	IKBKG	4.1409		2.7123						
C06	IL10	176.5643	A	158.9798	A					
C07	IL1A	1.6785		1.4464						
C08	IL1B	9.2404		4.9509						
C09	IL1R1	66.8729	A	56.9293	A					
C10	IRAK1	9.3544		7.3773						
C11	IRAK2	8.6984		7.4071						
C12	JUN	5.257		2.7566						
D01	LTA	3.0023		1.9592						
D02	LTBR	13.6285		11.0678						
D03	MALT1	91.8845		66.6947						
D04	MAP3K1	61.0809	A	29.3645	A					
D05	MYD88	23.8538		13.6876						
D06	NFKB2	3.2566		2.8499						
D07	NOD1	21.4789		16.2864						
D08	PSIP1	28.6358		21.9067						
D09	RAF1	6.6587		5.0037						
D10	REL	5.6889		5.0169						
D11	RELA	7.8112		5.8417						
D12	RELB	25.7099		15.1737						
E01	RHOA	24.3078		14.2126						
E02	RIPK1	9.0678		4.2416						

ACCEPTED MANUSCRIPT

E03	STAT1	25.0207		18.561						
E04	TBK1	10.2188		6.3577						
E05	TICAM1	3.8929		3.5751						
E06	TIMP1	9.2352		2.5971						
E07	TLR1	274.2884	A	37.0192	A					
E08	TLR2	90.1247	A	52.0238	A					
E09	TLR3	11.604		5.8434						
E10	TLR4	6.8227		3.7239						
E11	TLR6	80.4963	A	51.5679	A					
E12	TLR9	8.7201		6.3392						
F01	TNFRSF10A	2.8429		2.7031						
F02	TNFRSF10B	1.1033		-1.4955						
F03	TNFRSF1A	12.4343		8.8864						
F04	TNFSF10	2.226		2.1007						
F05	TRADD	19.4628		8.5203						
F06	TRAF2	2.3713		2.2138						
F07	TRAF3	5.0861		3.6095						
F08	TRAF6	5.8834		4.8544						
F09	BIRC3	2.5369		1.8274						
F10	CCL20	1.7494		1.7066						
F11	CD83	-1.8208		-2.6029						
F12	CXCL2	2.0126		1.1872						
G01	CXCL3	43.2537	A	22.231	A					
G02	ICAM1	-3.5046		-3.5643						
G03	CXCL8	1.9341		1.6475						
G04	IRF1	2.5918		2.2856						
G05	NFKB1	2.9954		2.0922						
G06	NFKBIA	1.0739		-1.1787						
G07	NFKBIE	7.7761		6.5428						
G08	STX11	10.1869		6.7885						
G09	TIFA	76.3126		34.7779						
G10	TNF	1.5032		1.1242						
G11	TNFAIP2	-17.9716		-21.769						
G12	TNFAIP3	1.9326		1.4901						
control genes										
H01	ACTB	1.6923		1.7671						
H02	RPLP0	-1.6923		-1.7671						

Comments:

A: This geneï¿½s average threshold cycle is relatively high (> 30) in either the control or the test sample, and is reasonably low in the other sample (< 30).

These data mean that the geneï¿½s expression is relatively low in one sample and reasonably detected in

ACCEPTED MANUSCRIPT

the other sample suggesting that the actual fold-change value is at least as large as the calculated and reported fold-change result.						
This fold-change result may also have greater variations if p value > 0.05; therefore, it is important to have a sufficient number of biological replicates to validate the result for this gene.						
B: This geneï¿½s average threshold cycle is relatively high (> 30), meaning that its relative expression level is low, in both control and test samples, and the p-value for the fold-change is either unavailable or relatively high (p > 0.05).						
This fold-change result may also have greater variations; therefore, it is important to have a sufficient number of biological replicates to validate the result for this gene.						
C: This geneï¿½s average threshold cycle is either not determined or greater than the defined cut-off value (default 35), in both samples meaning that its expression was undetected, making this fold-change result erroneous and un-interpretable.						
Fold Change & Fold Regulation:						
Fold-Change (2^(- Delta Delta Ct)) is the normalized gene expression (2^(- Delta Ct)) in the Test Sample divided the normalized gene expression (2^(- Delta Ct)) in the Control Sample.						
Fold-Regulation represents fold-change results in a biologically meaningful way. Fold-change values greater than one indicate a positive- or an up-regulation, and the fold-regulation is equal to the fold-change.						
Fold-change values less than one indicate a negative or down-regulation, and the fold-regulation is the negative inverse of the fold-change.						
Fold-change and fold-regulation values greater than 2 are indicated in red; fold-change values less than 0.5 and fold-regulation values less than -2 are indicated in blue.						

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

		Up-Down Regulation (comparing to control group: Normal Th1 cells)					
		Group 1 (WAS Th1, untransfected)		Group 2 (WAS Th1, wt-WASp transfected		Group 3 (WAS Th1, RNaseH1 transfected	
		Fold Regulation	Comments	Fold Regulation	Comments	Fold Regulation	Comments
A01	AGT	-211.6079	A	-2.1886		-3.011	
A02	AKT1	-60.286		-2.514		-2.6415	
A03	ATF1	-69.8675		-2.042		-3.192	
A04	BCL2A1	-59.7576		-1.257		-1.5625	
A05	BCL2L1	-1.9463		-2.1435		-3.4234	
A06	BCL3	-31.1109		-2.2191		-2.8668	
A07	BIRC2	-56.6086		-2.1585		-3.7744	
A08	CARD11	-21.2538		-2.1585		-2.4648	
A09	CASP1	-161.7195		-1.7411		-2.8072	
A10	CASP8	-26.3211		-2.0139		-2.6726	
A11	CCL2	-1.2605		-1.879		-22.1058	
A12	CCL5	-7.5415		-2.3457		-2.9698	
B01	CD40	-4.4427		-1.366		-1.8548	
B02	CFLAR	-7.709		-1.5476		-1.5322	
B03	CHUK	-49.9963		-1.9053		-2.7977	
B04	CSF1	-2.094		2.0994		1.4091	
B05	CSF2	23.9168		-2.1735		-3.2253	
B06	CSF3	-4244.415	A	-2.1287		-140.59	
B07	EGFR	-112.3885	A	-1.7777		-1.7487	
B08	EGR1	-3.7199		-1.6472		-4.054	
B09	ELK1	-16.665		-1.6358		-3.3902	
B10	FADD	-66.7895		-1.9319		-2.9786	
B11	FASLG	-1987.7905	A	-1.7171		-2.4476	
B12	FOS	-44.6891	A	-1.5369		-2.1324	
C01	HMOX1	-84.2234	A	-1.5157		-1.9543	
C02	IFNA1	-738.7027	A	-2.3134		-3.6926	
C03	IFNG	-103.1965		-1.5476		-2.0629	
C04	IKBKB	-6.3575		-1.5263		-1.7214	
C05	IKBKG	-6.9161		-1.6702		-2.5499	
C06	IL10	-286.8296	A	-1.6245		-1.8042	
C07	IL1A	-3.3338		-1.9862		-2.3048	
C08	IL1B	-19.0003		-2.0562		-3.8377	
C09	IL1R1	-103.4902	A	-1.5476		-1.8179	
C10	IRAK1	-12.7785		-1.366		-1.7321	
C11	IRAK2	-19.9837		-2.2974		-2.6979	
C12	JUN	-8.9647		-1.7053		-3.2521	
D01	LTA	-6.2595		-2.0849		-3.1949	
D02	LTBR	-20.5142		-1.5052		-1.8535	
D03	MALT1	-195.5983		-2.1287		-2.9327	
D04	MAP3K1	-92.5814	A	-1.5157		-3.1528	
D05	MYD88	-54.4232		-2.2815		-3.9761	
D06	NFKB2	-5.0748		-1.5583		-1.7807	
D07	NOD1	-30.7998		-1.434		-1.8911	
D08	PSIP1	-48.1595		-1.6818		-2.1984	
D09	RAF1	-10.0928		-1.5157		-2.0171	
D10	REL	-9.7011		-1.7053		-1.9337	
D11	RELA	-11.7578		-1.5052		-2.0127	
D12	RELB	-35.8587		-1.3947		-2.3632	
E01	RHOA	-48.9538		-2.0139		-3.4444	
E02	RIPK1	-13.0932		-1.4439		-3.0869	
E03	STAT1	-48.673		-1.9453		-2.6223	
E04	TBK1	-18.6765		-1.8277		-2.9376	
E05	TICAM1	-6.4123		-1.6472		-1.7936	
E06	TIMP1	-15.2121		-1.6472		-5.8574	
E07	TLR1	-418.6353	A	-1.5263		-11.3086	
E08	TLR2	-126.5751	A	-1.4044		-2.433	
E09	TLR3	-18.2086		-1.5692		-3.1161	
E10	TLR4	-10.4857		-1.5369		-2.8158	

ACCEPTED MANUSCRIPT

E11	TLR6	-123.7127	A	-1.5369		-2.399	
E12	TLR9	-12.2469		-1.4044		-1.9319	
F01	TNFRSF10A	-3.83		-1.3472		-1.4169	
F02	TNFRSF10B	-1.8174		-1.6472		-2.7178	
F03	TNFRSF1A	-22.1042		-1.7777		-2.4874	
F04	TNFSF10	-2.9576		-1.3287		-1.4079	
F05	TRADD	-29.2964		-1.5052		-3.4384	
F06	TRAF2	-3.5693		-1.5052		-1.6123	
F07	TRAF3	-7.4464		-1.4641		-2.063	
F08	TRAF6	-8.9176		-1.5157		-1.837	
F09	BIRC3	-4.8001		-1.8921		-2.6267	
F10	CCL20	-2.8028		-1.6021		-1.6423	
F11	CD83	1.0385		-1.7532		-2.5063	
F12	CXCL2	-3.1146		-1.5476		-2.6234	
G01	CXCL3	-76.8914	A	-1.7777		-3.4587	
G02	ICAM1	1.9175		-1.8277		-1.8588	
G03	CXCL8	-3.5842		-1.8532		-2.1755	
G04	IRF1	-4.2692		-1.6472		-1.8678	
G05	NFKB1	-5.0028		-1.6702		-2.3912	
G06	NFKBIA	-1.7936		-1.6702		-2.1141	
G07	NFKBIE	-12.8978		-1.6586		-1.9713	
G08	STX11	-14.5067		-1.4241		-2.1369	
G09	TIFA	-98.6228		-1.2924		-2.8358	
G10	TNF	-2.4761		-1.6472		-2.2025	
G11	TNFAIP2	9.9704		-1.8025		-2.1834	
G12	TNFAIP3	-3.2055		-1.6586		-2.1512	

control genes

H01	ACTB	-1.0203		1.6586		1.732	
H02	RPLP0	1.0203		-1.6586		-1.732	

Comments:

A: This geneï¿½s average threshold cycle is relatively high (> 30) in either the control or the test sample, and is reasonably low in the other sample (< 30).
These data mean that the geneï¿½s expression is relatively low in one sample and reasonably detected in the other sample suggesting that the actual fold-change value is at least as large as the calculated and reported fold-change result.
This fold-change result may also have greater variations if p value > 0.05; therefore, it is important to have a sufficient number of biological replicates to validate the result for this gene.

B: This geneï¿½s average threshold cycle is relatively high (> 30), meaning that its relative expression level is low, in both control and test samples, and the p-value for the fold-change is either unavailable or relatively high (p > 0.05).
This fold-change result may also have greater variations; therefore, it is important to have a sufficient number of biological replicates to validate the result for this gene.

C: This geneï¿½s average threshold cycle is either not determined or greater than the defined cut-off value (default 35), in both samples meaning that its expression was undetected, making this fold-change result erroneous and un-interpretable.

ACCEPTED MANUSCRIPT

Fold Change & Fold Regulation:

Fold-Change (2^(- Delta Delta Ct)) is the normalized gene expression (2^(- Delta Ct)) in the Test Sample divided the normalized gene expression (2^(- Delta Ct)) in the Control Sample.

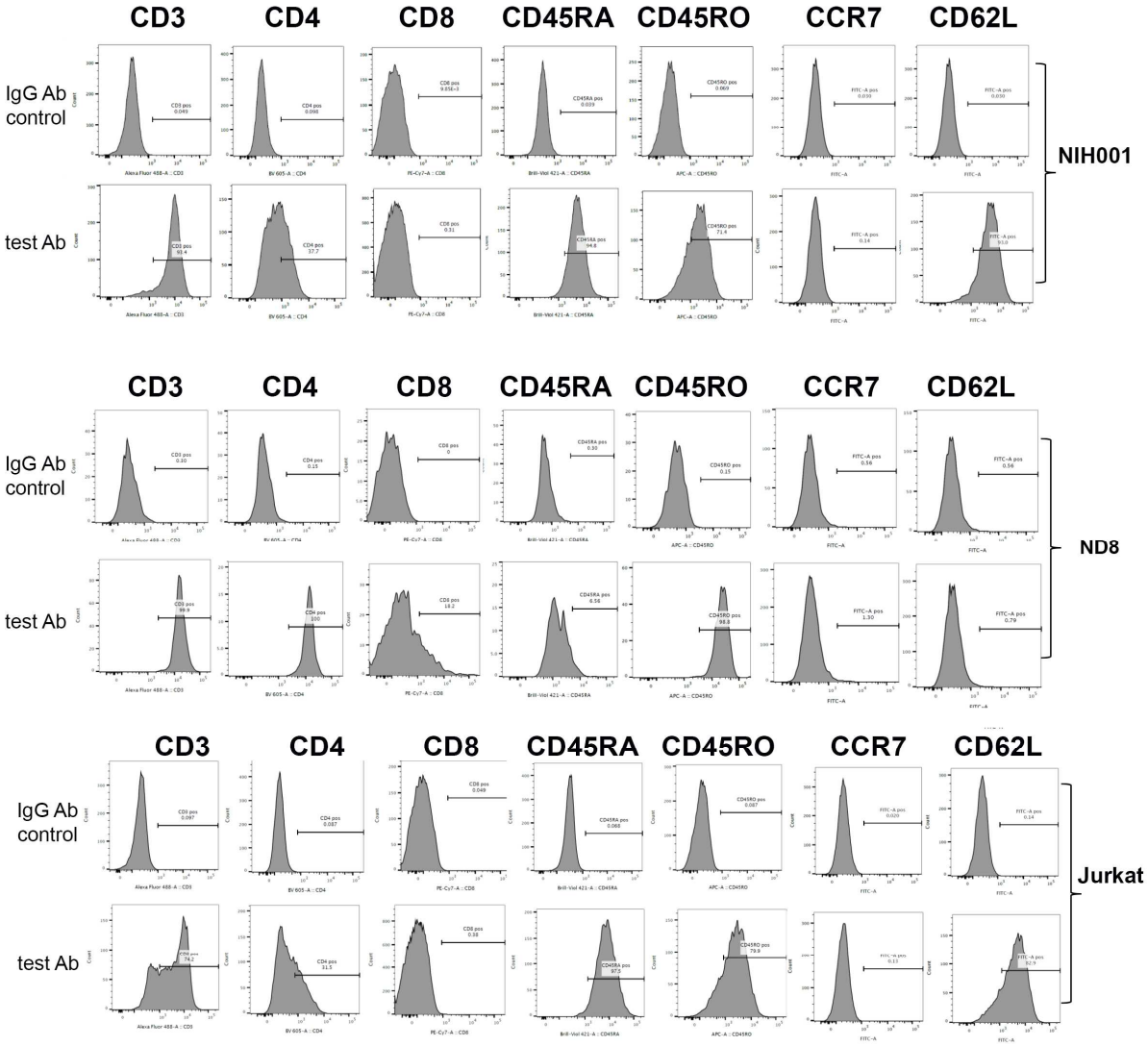
Fold-Regulation represents fold-change results in a biologically meaningful way. Fold-change values greater than one indicate a positive- or an up-regulation, and the fold-regulation is equal to the fold-change.

Fold-change values less than one indicate a negative or down-regulation, and the fold-regulation is the negative inverse of the fold-change.

Fold-change and fold-regulation values greater than 2 are indicated in red; fold-change values less than 0.5 and fold-regulation values less than -2 are indicated in blue.

ACCEPTED MANUSCRIPT

Supplementary Figure 1



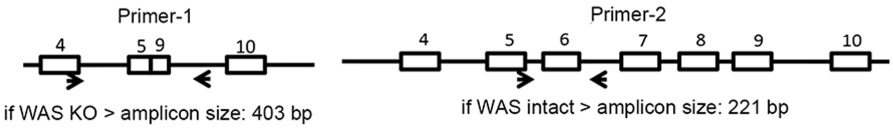
	CD3	CD4	CD8	CD45RA	CD45RO	CCR7	CD62L	*Th subtype*
NIH001	+	+	-	+	+	-	+	**Effector cell** (not naïve or central memory cell)
Milan008	+	+	-	-	+	-	-	**Effector memory cell** (not naïve or central memory cell)
Jurkat	+	+	-	+	+	-	+	**Effector cell** (not naïve or central memory cell)

ACCEPTED MANUSCRIPT

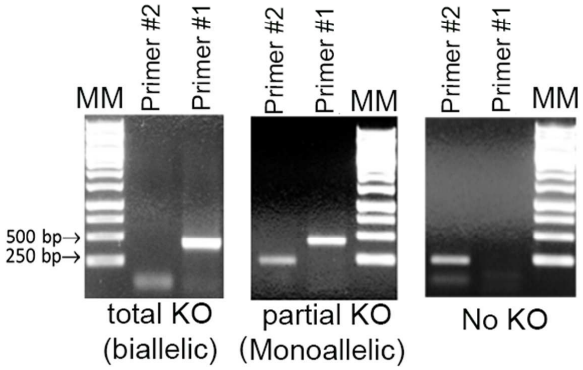
Supplementary Figure 2

Generation of CRISPR/Cas9-mediated WAS KO Th cell line (NIH001)

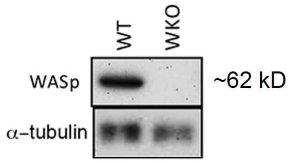
A Designing guide RNAs



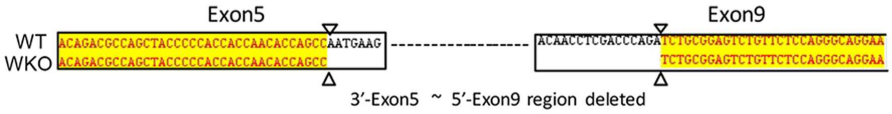
B Screening WAS KO clones by PCR



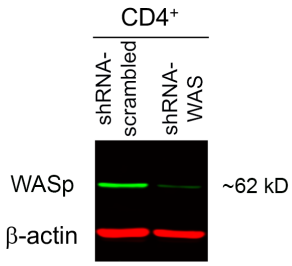
C Verifying WASp depletion by Western blot



D Confirming WAS KO by gDNA sequencing

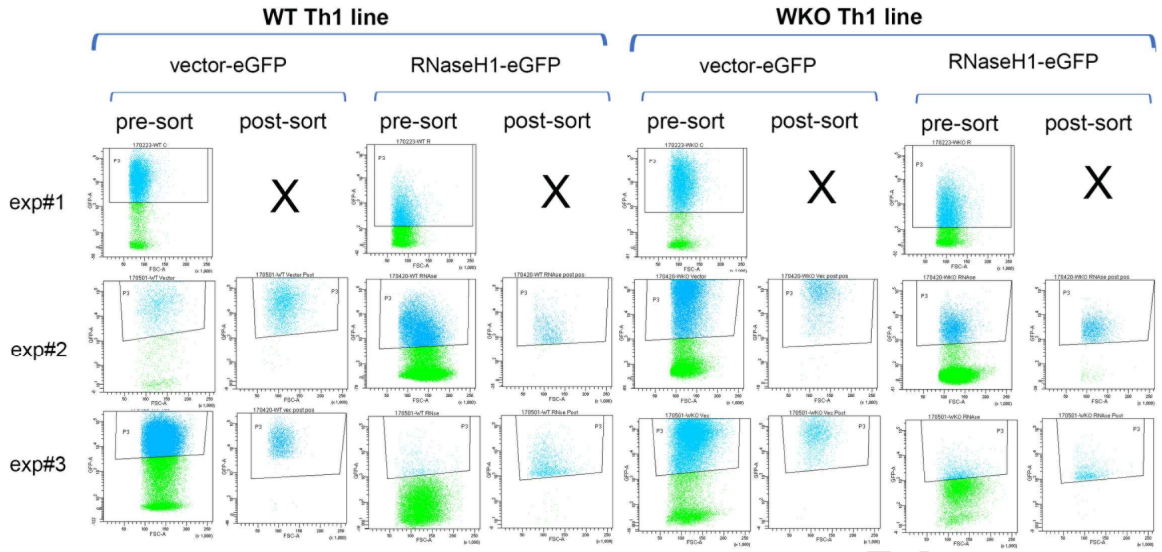


E **Generation of RNA-interefence mediated WAS KD primary Th cells (ND8)**



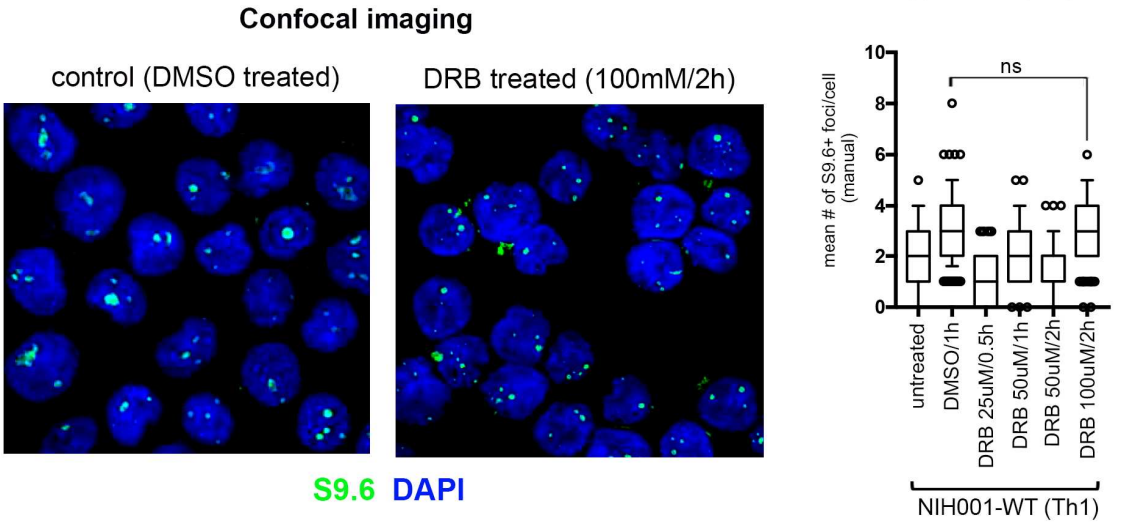
ACCEPTED MANUSCRIPT

Supplementary Figure-3



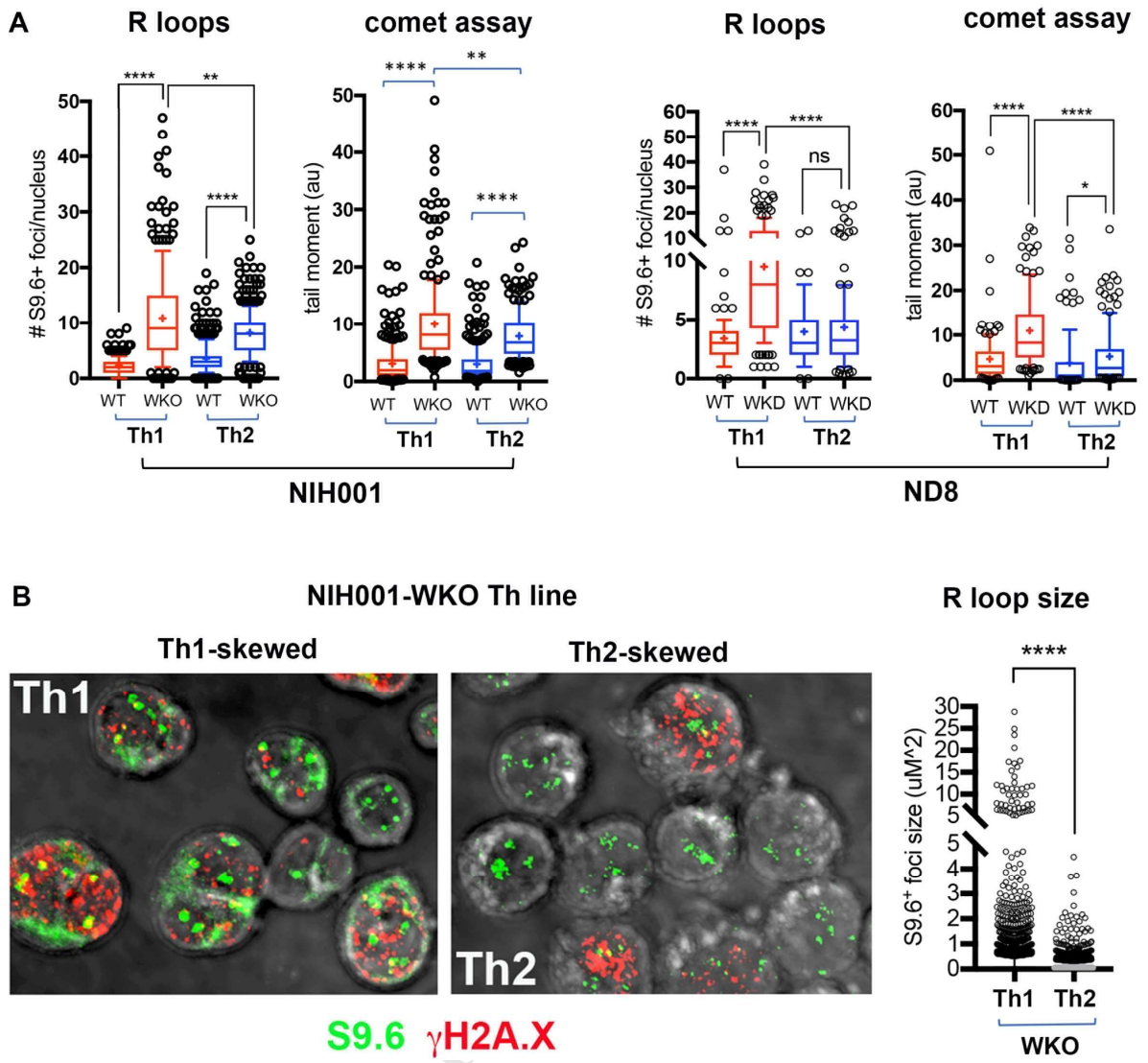
ACCEPTED MANUSCRIPT

Supplementary Figure 4



ACCEPTED MANUSCRIPT

Supplementary Figure-5

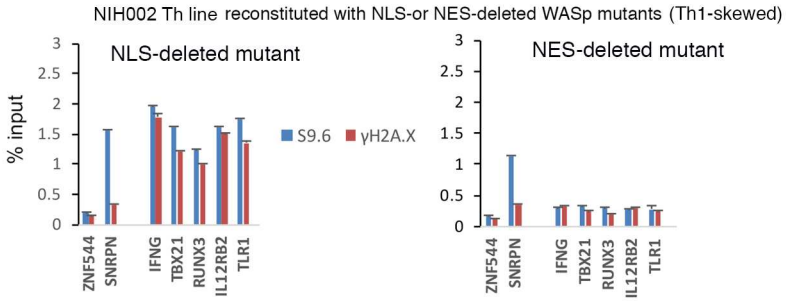


ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

Supplementary Figure-6

DRIP/ChIP-qPCR: 3' end of genes



ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

Supplementary Figure-7

