Citation: Molecular Therapy — Methods & Clinical Development (2016) **3,** 16077; doi:10.1038/mtm.2016.77
Official journal of the American Society of Gene & Cell Therapy

www.nature.com/mtm

REVIEW

The identification of hematopoietic-specific regulatory elements for *WASp* gene expression

Jun Zhan[1], Irudayam Maria Johnson[1], Matthew Wielgosz[1] and Arthur W Nienhuis[1]

Chromosome Conformation Capture (3C) technology was used to identify physical interactions between the proximal Wiskott-Aldrich Syndrome protein (*WASp*) promoter and its distant DNA segments in Jurkat-T cells. We found that two hematopoietic specific DNase I hypersensitive (DHS) sites (proximal DHS-A, and distal DHS-B) which had high interaction frequencies with the proximal WASp promoter indicating potential regulatory activity for these DHS sites. Subsequently, we cloned several DNA fragments around the proximal DHS-A site into a luciferase reporter vector. Interestingly, no fragments showed enhancer activity, but two fragments exhibited strong silencing activity in Jurkat-T cells. After aligning the chromatin state profiling for hematopoietic and nonhematopoietic cells using the human genome browser (UCSC), we found a 5 kb putative hematopoietic specific enhancer region located 250 kb downstream of the *WAS* gene. This putative enhancer region contains two hematopoietic cell specific DHS sites. Subsequently, the hematopoietic specific DHS sites enhanced luciferase expression from the proximal WASp promoter in all hematopoietic cells we tested. Finally, using a lentiviral vector stable expression system, the hematopoietic specific-enhancer(s) increased GFP reporter gene expression in hematopoietic cells, and increased *WASp* gene expression in WASp deficient cells. This enhancer may have the potential to be used in gene therapy for hematological diseases.

Molecular Therapy — Methods & Clinical Development (2016) **3**, 16077; doi:10.1038/mtm.2016.77; published online 14 December 2016

INTRODUCTION

Gene therapy using viral vectors has shown significant benefit in preclinical and clinical trials for the treatment of immune disorders.[1] The Wiskott-Aldrich syndrome (WAS) is an immunological disorders characterized by thrombocytopenia, eczema, immunodeficiency in both humoral and cell-mediated responses, and a tendency toward lymphoproliferative disease.[2] The affected gene encodes a 502 amino acid proline-rich protein with a predicted molecular weight of 53 KD that has been termed the WAS protein (WASp).[3–5]

The encoded WASp consists of a pleckstrin homology (PH) domain at the N-terminus, a CDC42 GTPase binding domain, a proline-rich region or SH3 binding domain and an actin binding C-terminus.[6] The gene for *WASp* is 15 kb, with 12 exons and two alternate promoters.[7] The proximal promoter, which lies immediately adjacent to the translation start site,[8] shows strong activity in hematopoietic cell lines which correlates with its tissue-specific expression pattern.[9,10] The distal promoter resides 6 kb upstream and also shows activity in hematopoietic cell lines.[11]

How *WASp* gene activity is regulated by these two promoters *in vivo* is not yet defined. Gene activity has long been known to be related to chromosome and nuclear organization. The nuclear structure in each living cell is highly organized and dynamically assembled into different cellular states,[8,12,13] which are thought to play essential roles in many aspects of genome regulation, including gene expression, DNA replication, chromosome transmission and maintenance of genome stability.[14] With this dynamic three-dimensional (3D) nuclear organization, gene expression is profoundly dependent upon chromatin folding and looping interactions that facilitate long-range control by distant gene regulatory elements.[15–19] Thus enhancers may influence expression of genes over large genomic distances (up to hundreds of kb) via chromatin folding into 3D structures that enable enhancer-promoter communication.

One way to study spatial organization of long genomic regions in living cells is to use chromosome conformation capture (3C) technologies.[15,20,21] These methods have been successfully used to identify associations between specific sequences involved in gene regulation—including domains of clustered active genes (*e.g.*, the *α-globin* cluster),[22,23] and "looping" between long-range enhancers and their target genes.[24–26] To investigate the potential long-range regulatory elements for *WASp* gene expression, we have done 3C experiments in Jurkat-T cells, a human T-lymphocyte line. We found high interaction frequencies between the proximal *WASp* promoter and its downstream hematopoietic specific DNase I hypersensitive (DHS) sites, DHS-A and DHS-B. Subsequently, we cloned the DNA fragments around the interactive region into a luciferase reporter vector. We have identified two silencers active in Jurkat-T cells from genomic DNA around the DHS-A, and a potential hematopoietic specific enhancer from genomic DNA around the DHS-B site.

[1]Department of Hematology, Division of Experimental Hematology, St. Jude Children's Research Hospital, Memphis, Tennessee, USA. Correspondence: AW Nienhuis (arthur.nienhuis@stjude.org)
Received 21 July 2016; accepted 18 October 2016

RESULTS

Chromatin conformation of the *WASp* gene locus on chromosome X

Active regulatory elements within complex genomes may be identified by pronounced sensitivity to the nonspecific endonuclease, DNase I[27–29] when exposed in the context of intact nuclei. Global whole genome mapping of DHS sites for 125 different cell types was done by John Stamatoyannopoulos' research group.[30] To search for potential *WASp* gene regulatory (enhancer or silencer) elements, we aligned and compared the DHS profiles of hematopoietic versus nonhematopoietic cell types over a 314 kb genomic region covering the *WASp* gene. Two hematopoietic specific DHS sites were identified (Figure 1a). DHS-A is >40 kb, and DHS-B is 250 kb downstream of *WASp* gene. These potential regulatory elements (DHS sites) could influence *WASp* gene expression through long-range interaction.[15,16,18,19,30] One way to evaluate the long-range genomic DNA interactions is with 3C technologies.[13,17,20] We employed 3C technology to identify physical interactions between the *WASp* proximal promoter and distant DNA segments in Jurkat-T cells. 293T cells were used as a negative control, as the *WASp* gene is not expressed in these cells. BAC clones covering the human genomic region of interest were used to generate a 3C control library to normalize polymerase chain reaction (PCR) efficiency. 3C libraries from Jurkat-T and 293T cells and a BAC control library were generated according to the published protocol. 3C primers were designed around the two DHS sites, the restriction enzyme BglII was used to cut DNA and the anchor primer was located at the proximal WASp promoter 3 kb upstream of WASp start site of the coding region. A relative interaction frequency was calculated based on the intensity of the 3C PCR band normalized by that from control library. A potential interaction was represented as a peak for the relative interaction (cross-linking) frequency. As we expected, the calculated cross-linking frequencies (Figure 1b) indicated interactions of the WASp proximal promoter with two hematopoietic specific DHS sites identified above, as proximal and distal DHS. The association (interaction) between DHS-A (proximal) and WASp proximal promoter is stronger than that between DHS-B (distal) and the WASp promoter. The interaction between the WAS promoter and DHS-A site was confirmed by more 3C experiments using different primers around this site (Figure 1c and Supplementary Figure S1; Supplementary Table S1), indicating a potential regulatory function of DHS-A on *WASp* gene activity.

Luciferase assays with reporter vectors driven by a minipromoter

After we identified which genomic regions interact with the *WASp* proximal promoter, we have evaluated what kind of gene regulatory activity these regions exhibited, *e.g.*, enhancer, silencer, or insulator activity. Using the web-based software TRANSFECT, we found several putative enhancers, called lymphocyte enhancer factor1 binding (LEF1) sites, around the DHS-A genomic region. To test the potential enhancer activity of this region, we initially cloned 2.4 kb (D6L+, D6L− in Figure 2a) and 1 kb fragments (D6S+, S+nde in Figure 2a) into a luciferase vector (pGL4.23luc) up stream of a 31 bp synthetic minipromoter that drove firefly luciferase gene expression. Various reporter plasmids in the pGL4.23luc backbone were transfected into Jurkat-T and 293T cells for luciferase reporter gene assays. Renilla luciferase vector pGL4.74hRluc/TK was cotransfected with these reporter vectors and the renilla luciferase activity was used to normalize the transfection efficiencies. After calculating and comparing the relative luciferase activity from different vectors, we found no enhancer activity from these fragments, but we observed silencer activity from the 1 kb fragment in reverse orientation (S+nde) with almost 10-fold reduced gene expression (Figure 2a). Deletion within the 1 kb silencer disrupted its silencing activity (Figure 2b), indicating that the whole 1 kb DNA was required for its function.

For gene therapy applications, we were more interested in finding enhancers than silencers. We wondered if there was any regulatory activity, especially enhancer activity in the DHS-A containing fragments. To this end, we cloned a 1.8 kb fragment covering the DHS-A site, as well as a 1.7 kb fragment immediately upstream of this 1.8 kb fragment, called the DHS-A and D5 fragments, respectively (Figure 2c). Both DNA fragments were amplified from genomic DNA and subsequently cloned into the luciferase reporter vector with the minipromoter. Reporter assays were performed in Jurkat-T and 293T cells. The luciferase activities indicated that both DHS-A and D5 fragments do not have any enhancer activity (Figure 2c). Again we observed silencer activity from the DHS-A fragment in reverse orientation. The deletion of the DHS-A fragment from its 5′ end eliminated the silencer activity (Figure 2d), implying that the 5′ end 500 bp deleted DNA fragment may be essential for the silencer activity. After cloning this 500 bp fragment (named DHS-A-s) into the reporter vector, we found this fragment to have silencer activity (Figure 2e). To find out which element played a role in the silencer activity, we searched for transcriptional repressor binding sites on this 500 bp silencer. Two putative repressor binding elements, Xvent sites[31]were found in DHSA-s. To determine if these Xvent sites are essential for the silencer activity, single or double Xvent sites deletion mutants were generated. Unfortunately, these mutations did not disrupt the silencer activity (Figure 2f), indicating that the Xvent binding sites do not play any role in the silencer activity. Taken together, we identified two silencers around the DHS-A site that acted on the p500 *WASp* proximal promoter.

Luciferase assays in a reporter vector driven by p500 WASp proximal promoter

Using the reporter vectors with the 31 bp synthetic minipromoter for reporter gene expression, we identified two silencers at the genomic locus around the DHS-A site. However, the minipromoter may not reflect the real physiological enhancer or silencer promoter interactions. To determine if these silencers really regulate the *WAS* gene expression, we employed the *WASp* proximal gene promoter to drive reporter gene expression for evaluating silencer or enhancer activity. A 500 bp *WASp* proximal promoter fragment was cloned into a luciferase vector instead of the minipromoter thereby generating a reporter vector driven by the p500 *WAS* proximal promoter. Subsequently, we constructed reporter vectors containing the same silencers previously tested in the minipromoter vectors. Reporter assays were performed in Jurkat-T cells and 293T cells. The relative luciferase activity indicated that both silencers still were functional and acted on the p500 *WASp* proximal promoter (Figure 3a). Other fragments either from the 1 kb silencer or from the 1.8 kb DHSA-Re fragments were also tested. Fragment Silen2 shown weak enhancer activity in the vector with the minipromoter (Figure 2b), and the same in the p500 proximal promoter vector (Figure 3b) while 400/nde shown no activity in vectors driven by minipromoter (Figure 2b), but weak enhancer activity in the context of the p500 proximal promoter (Figure 3b). The combination of two weak enhancers, silen2 and 400/nde, did not yield strong enhancer activity (data not shown).

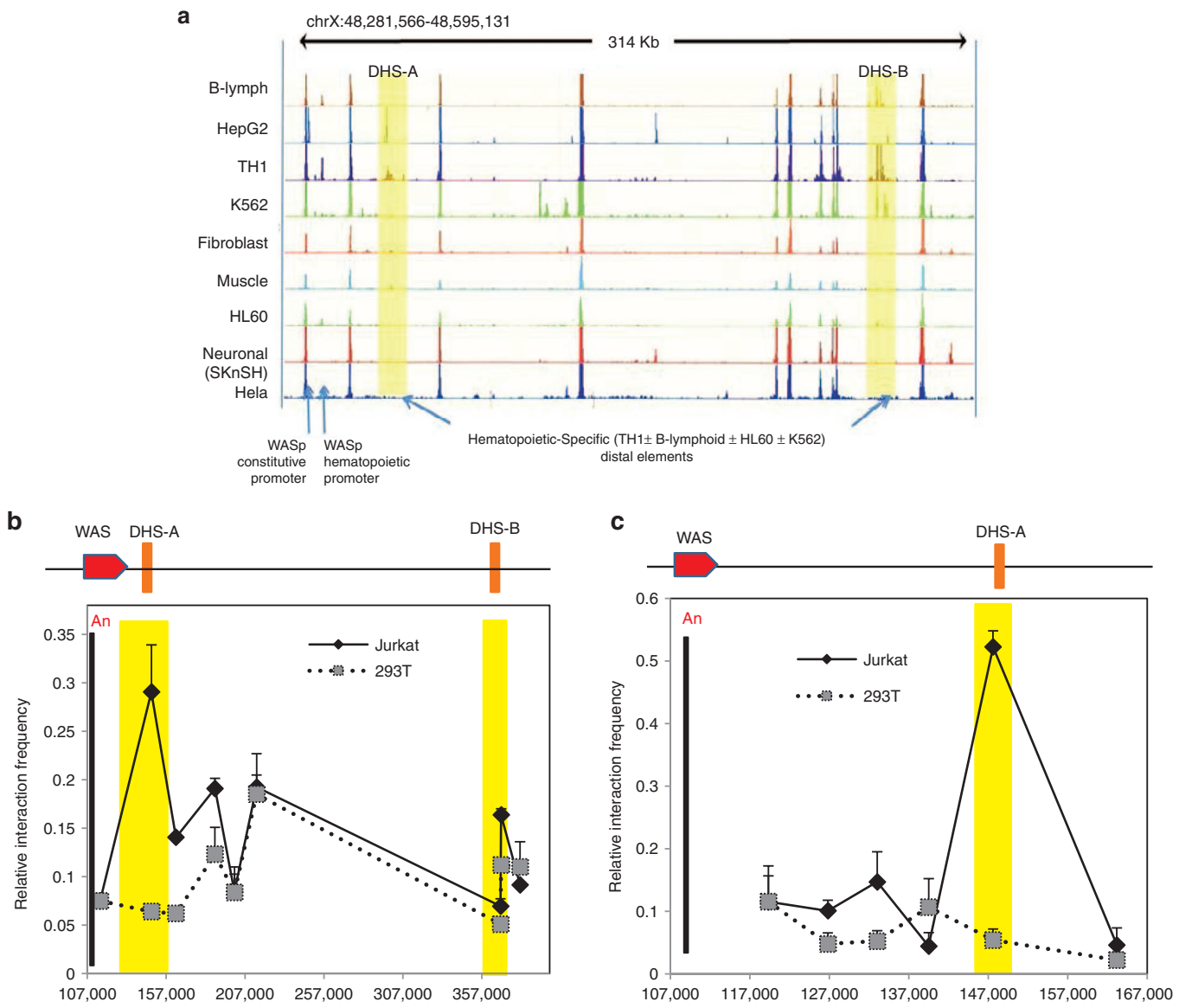

Figure 1 Chromatin confirmation of chromosome X *WAS* gene locus. (**a**) Identification of hematopoietic specific DNase I hypersensitive (DHS) sites. DHS profiles in X chromosome *WAS* gene locus over 300 kb genomic region were aligned for several cells type including B-lymphocyte, HL60 , Th1, k562, Fibroblast, Muscle cell, HepG2, SKnSH (neuronal), and HeLa cells. Each horizontal line represents one cell type, and each peak above the line represents a DHS site at that genomic site. Two hematopoietic specific DHS sites are highlighted, DHS-A is over 40 kb, and DHS-B is 250 kb downstream of *WAS* gene. *WAS* gene constitutive and hematopoietic specific promoter also pointed at the 5′ end. (**b**) 3C-PCR relative interaction frequency for *WAS* gene proximal promoter with its downstream two hematopoietic specific DHS sites. (**c**) 3C-PCR relative interaction frequency for *WAS* gene proximal promoter with its downstream proximal hematopoietic specific DHS-A site (over 40 kb from *WAS* gene proximal promoter). Relative interaction frequency was calculated according the intensity of PCR band from 3C experiments. 293T was a control cell line where *WAS* gene was not expressed. The black bar close to Y-axis is the location for anchor primer. X-axis shows the genomic coordinates and locations of *WAS* gene and DHS sites. The data was an average from six independent experiments. The error bar represents standard error. PCR, polymerase chain reaction; 3C, chromosome conformation capture.

Bioinformatics identification of a strong enhancer downstream of the *WASp* gene

To search for potential hematopoietic specific enhancers, we aligned the chromatin state profiling[32] for hematopoietic versus nonhematopoietic cell lines using the human genome browser (UCSC) and focusing on the *WASp* gene locus on Chromosome X. We found a 5 kb hematopoietic specific putative strong enhancer both in K562 (myeloid leukemia) and GM27878 cells (derived from peripheral blood mononuclear cells transformed by EBV, mostly B cells), but not for nonhematopoietic cells (Figure 4a, top of bottom panel, strong enhancer is color coded as orange). In this putative enhancer region there are four DHS sites and the two on the right side of the region are T cell specific (Figure 4a, red box). The left two are hematopoietic specific DHS sites which show exactly the same pattern with that of *WASp* proximal promoter region (Figure 4a, the yellow boxes). This information implied an important role of this region in *WASp* gene regulation. Furthermore, this 5 kb region partially overlapped with a previously identified hematopoietic specific DHS-B site (Figure 1a). To evaluate the enhancer activity of this region, we cloned the left and right DHS sites into luciferase vector driven with the p500 *WAS* proximal promoter and performed reporter assays in end (Figure 4b, en1.0) encompassing T cell specific DHS sites worked as an enhancer in Jurkat-T cells only, while the left DNA fragment (Figure 4b, en1.3) encompassing hematopoietic specific DHS sites

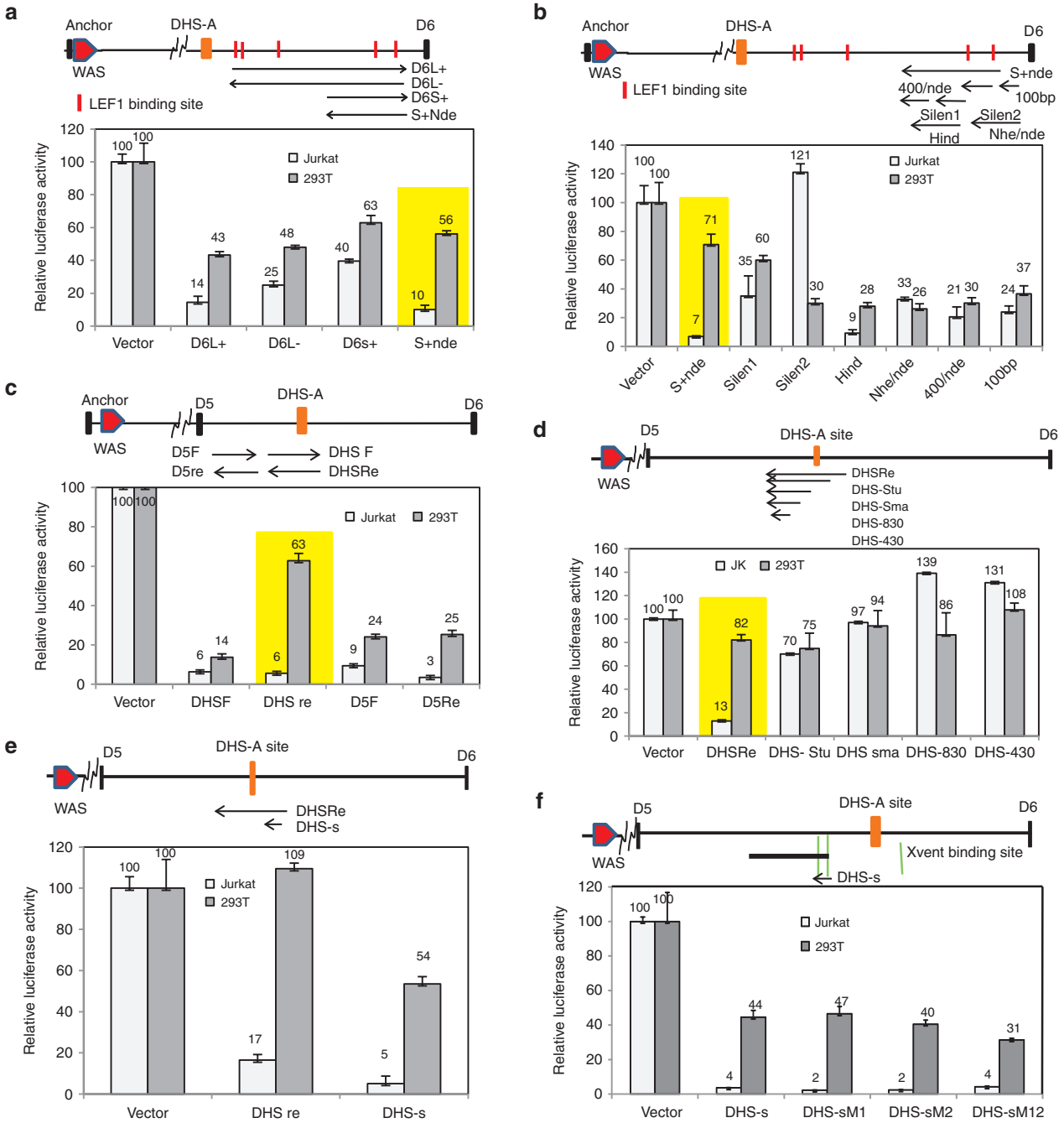

Figure 2 The assessment for potential enhancer or silencer activities up stream of 31 bp minipromoter. Potential gene regulatory elements were inserted into upstream of minipromoter that driven firefly luciferase reporter gene expression. Firefly luciferase reporter plasmids were cotransfected with pGL4.47(hRluc/TK) vector carrying *renilla luciferase* gene as an internal control for transfection efficiency. Jurkat-T cells and 293T cells were transfected and seeded onto 24 well plates. About 42 hours post-transfection cells were harvested and lysed. Both firefly and renilla luciferase were read on moonlight 2010 luminometer. The relative firefly activities were calculated. Results were presented in terms of 100 for control vector and the values represent the mean ± standard error (SE) from three independent transfection experiments. (**a**) Relative luciferase activities for 1 kb and 2.4 kb fragments around DHS-A site. Top panel shows the WAS genomic region and design of reporter vectors containing 2.4 kb and 1 kb DNA fragments in both orientations. Red arrow is *WAS* gene; DHS-A is proximal DHS site; D6 is the primer number 6 for BglII cut downstream of *WAS* gene. On the bottom panel highlighted is the silencer activity of 1 kb fragment (in reverse orientation) in Jurkat-T cells. (**b**) Relative luciferase activities for the 1 kb DNA fragment and its derivatives. Top panel shows the WAS genomic region and design of deletion reporter vectors derived from the 1 kb fragment. Highlighted is the 1 kb fragment in reverse orientation that kept silencer activity. (**c**) Relative luciferase activities for 1.8 kb DNA fragment covering DHS site and a 1.7 kb upstream fragment. Top panel shows the WAS genomic region and design of reporter vectors. Yellow shade highlighted that both DHS and D5 fragments in reverse orientation have silencer activity in Jurkat-T cells comparing to 293T cells. (**d**) Relative luciferase activities for 1.8 kb DHS fragment and its deletions. Top panel shows the WAS genomic region and design of deletion reporter vectors from 1.8 kb DHS. All the deletions disrupted silencer activity; Orange shade highlighted the silencer activity of 1.8 kb fragment. (**e**) Relative luciferase activities for the 500 bp silencer derived from the 1.8 kb DHS fragment. Top panel shows the WAS genomic region and design of reporter vectors. Orange box highlighted the silencer activity of the 500 bp fragment. (**f**) Relative luciferase activities for the 500 bp silencer and its single or double deletion mutation at the Xvent binding sites. Top panel shows the WAS genomic region and design of reporter vectors. DHS, DNase I hypersensitive; WAS; Wiskott-Aldrich Syndrome.

was an enhancer in K562 cells (Figure 4b). The hematopoietic specific enhancer (en1.3) activity was confirmed in HEL92.1.7 and U937 cell lines by luciferase reporter assays (Figure 4c,d for DS-1 cells). To have a complete map of the potential enhancers, we constructed more reporter vectors with insertion of different DNA fragments coving the whole 5 kb DNA and reporter assays again confirmed enhancer activity of en1.3 (Figure 4d).

Lenti-vector constructs for evaluating enhancer activity

To evaluate the stable expression regulated by the potential enhancers, we decided to make lentiviral vectors that express either green fluorescence protein (GFP) or *WASp*, as a reporter in the presence of enhancers. We cloned the DNA fragment from hematopoietic specific DHS region en1.3 (~1.3 kb) and its short version en900 (900 bp) into pCL20cw650p500hWASp or pCL20cw650p500hGFP vectors (Figure 5a). Lenti-vectors were packaged in 293T cells and tittered on HeLa cells. The insertion of the enhancers did not reduce the virus titer compared with the control vector.

Subsequently, we transduced different cell lines or hematopoietic stem cells from human or murine sources. At day 6, cells were harvested for mRNA and genomic DNA isolation. WASp mRNA levels were determined by quantitative real-time PCR (qRT-PCR) using woodchuck hepatitis virus post-transcriptional regulatory element (WPRE) as the probe and normalized to vector copy number. The en1.3 and en900 showed enhancer activity in both K562 cells (Figure 5b) and hCD34 (Figure 5c) cells when we analyzed the GFP reporter relative expression which was consistent with the luciferase reporter data. However, both fragments worked as silencers (Figure 5b,c) when we analyzed the WASp mRNA expression as determined with the same probe. This was in contrast to the reporter assays as indicated either by transient reporter luciferase activity or stable GFP expression. That means that the same promoter and its regulatory elements may work differently on each of the reporter genes, GFP, or WASp. We presumed that endogenous WASp expression may influence its own expression from the vectors. Comparing the relative expression of WAS over GFP per copy number from the control vectors driven by p500 *WASp* proximal promoter without any enhancer, we found that the WAS/GFP ratio was remarkably different for different cell types. Jurkat-T cells with strong endogenous WAS expression, have relative little WAS expression from the vector with a very low WAS/GFP ratio (Figure 5d); while K562 cells with little endogenous WAS expression, have relatively more WASp expression from the vector (Figure 5d). To further investigate the enhancer activity that may be affected by the *WAS* protein, we transduced bone marrow cells from WAS deficient mice. The *WASp* gene expression in lineage-marker negative (Lin−) bone marrow cells was very low. In Lin+ cells WASp expression was increased from twofold to threefold by the enhancers (Figure 5e). Taken together, we identified a hematopoietic specific enhancer for WASp expression only in WASp deficient cells.

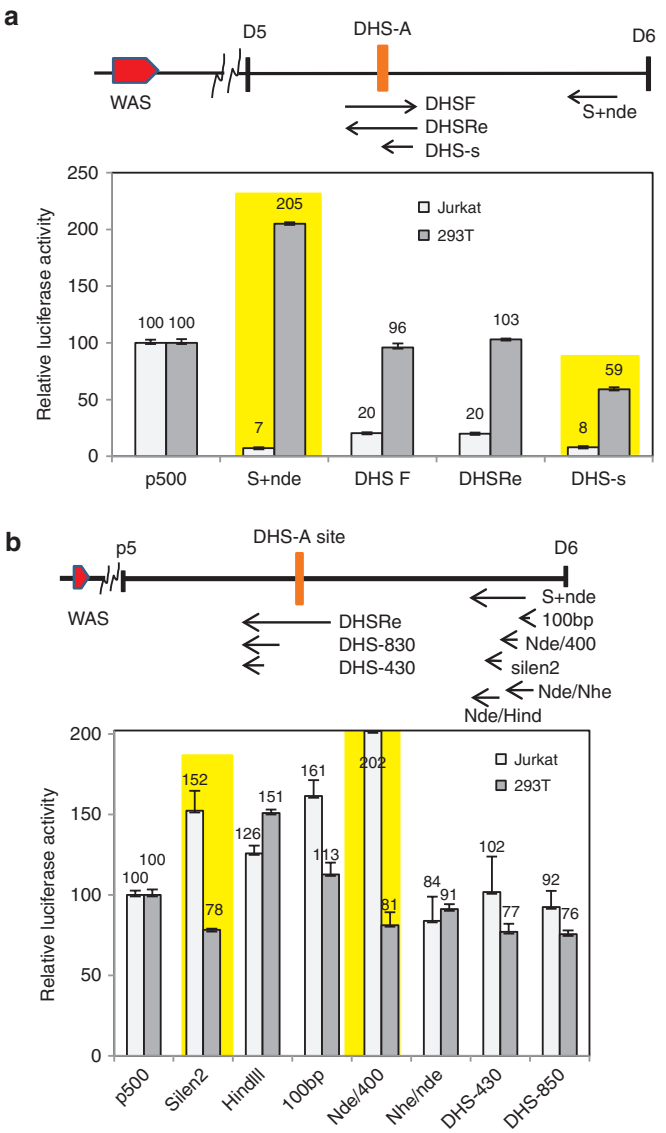

Figure 3 The assessment for potential enhancer or silencer activities up stream of 500 bp *WAS* gene proximal promoter p500. Potential gene regulatory elements were inserted into upstream of WAS p500 promoter that driven firefly luciferase reporter gene expression. Firefly luciferase reporter plasmids were cotransfected with pGL4.47(hRluc/TK) vector carrying *renilla luciferase* gene as internal control for transfection efficiency. Jurkat-T cells and 293T cells were transfected and seeded onto 24 well plates. About 42 hours post-transfection cells were harvested and lysed, both firefly and renilla luciferase were read on moonlight 2010 luminometer. The relative firefly activities were calculated. Results were presented in terms of 100 for control vector and the values represent the mean ± SE from three independent transfection experiments. (**a**) Relative luciferase activities for the silencers identified in the vectors of minimal synthetic promoter. Yellow shade highlighted the silencer activity of s+nde and DHS-s. (**b**) Relative luciferase activities for the deletions of 1 kb and 1.8 DHS fragments. Yellow shade highlighted that two fragments silen2 and 400/nde have shown week enhancer activity. DHS, DNase I hypersensitive; SE, standard error; WASp; Wiskott-Aldrich Syndrome protein.

DISCUSSION

DHS sites are regions of chromatin which are sensitive to cleavage by the DNase I enzyme and have been used as markers of regulatory DNA regions. In these specific regions of the genome, chromatin has lost its condensed structure, exposing the DNA and making it accessible to enzymatic digestion. These accessible chromatin zones are functionally related to transcriptional activity since this remodeled state is necessary for the binding of proteins such as transcription factors. The mapping of DHS sites has led to discovery of all classes of cis-regulatory elements including promoters, enhancers, insulators, silencers, and locus control regions. As part of Encode (Encyclopedia of DNA Elements, a public research project launched by the US National Human Genome Research Institute (NHGRI) in September 2003), John A. Stamatoyannopoulos' research group has done global mapping of DHS sites[30] for about 125 different cell

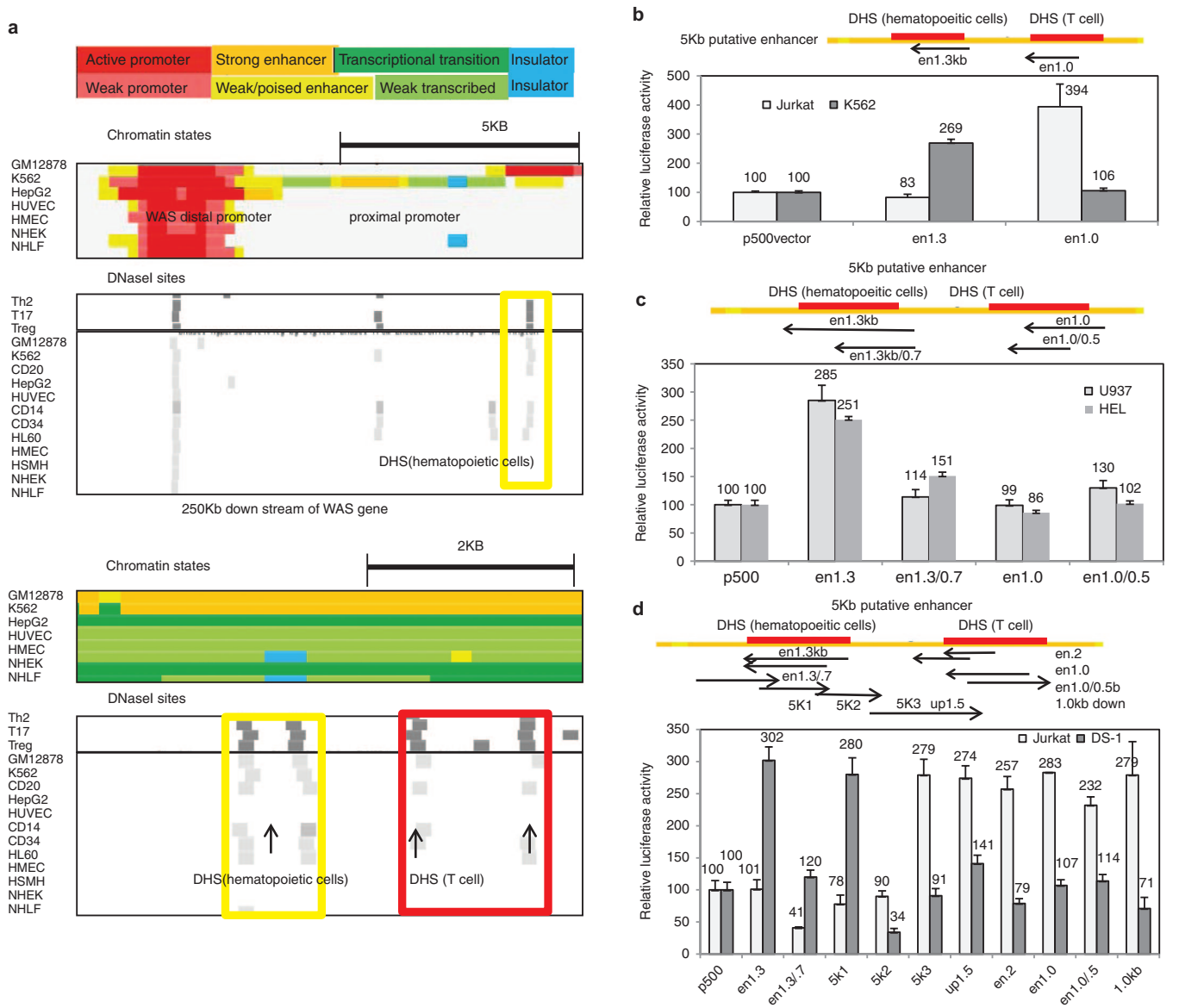

Figure 4 The identification of hematopoietic specific enhancer. (**a**) Bioinformatics identification of hematopoietic specific enhancer. Using human genome browser (UCSC), chromatin states were aligned together for several cell lines on the genomic region of Wiskott-Aldrich Syndrome (*WAS*) gene locus. Different chromatin states were indicated in different colors: red represents promoters, orange indicates a strong enhancer, yellow indicates weak enhancer; and green represents translated regions. The DNase I hypersensitive (DHS) sites were also aligned together for hematopoietic cells and nonhematopoietic cells. Top panel is the chromatin states and DHS sites in the genomic region of *WAS* gene and it's upstream. Bottom pane is the chromatin states and DHS sites in the genomic region about 250 kb downstream of *WAS* gene. The alignment for chromatin states included cells GM12878, K562, HepG2, HUVEC, HMEC, HSMM, NHEK, NHLF among them GM12868 and K562 are hematopoietic cell lines. The alignment for DHS sites included cell lines such as, Th1, Th2, Th17, Treg, GM12878, K562, CD20+, CD14+ CD4+ CD34+, HSMM, and NHEK. A strong enhancer 5 kb region is shown in the top of bottom panel for GM12878 and K562 hematopoietic cells only. Yellow boxes highlighted the hematopoietic specific DNase I site both in the top and bottom panel; red boxes highlighted the T cell specific DNase I sites in the bottom panel. (**b**) Relative luciferase activities for hematopoietic specific DHS sites and T cell specific DHS sites in Jurkat-T and K562 cells. Firefly luciferase reporter plasmids were cotransfected with pGL4.47(hRluc/TK) vector carrying *renilla luciferase* gene as internal control for transfection efficiency. About 42 hours post-transfection cells were harvested and lysed. Both firefly and renilla luciferase were read on moonlight 2010 luminometer. The relative firefly activities were calculated. Results were presented in terms of 100 for control vector and the values represent the mean ± SE (standard error) from three independent transfection experiments. (**c**) Relative luciferase activities in HEL92.1 and U937 hematopoietic cell lines. (**d**) Relative luciferase activities in Jurkat-T and DS-1 (B cells) cell lines.

types by digital footprint. Having aligned and compared the DHS sites in the genomic region containing the *WASp* gene in different cell types, we have identified two hematopoietic specific DHS sites 40 kb and 250 kb downstream of *WASp* gene.

Due to the 3D and dynamic organization of nucleus, distant gene regulatory elements can physically interact through chromatin looping to facilitate long-range control.[15,19,30] Thus, enhancers may influence expression of genes over large genomic distances (up to hundreds of kb) via chromatin folding into 3D structures to enable enhancer-promoter communication. Chromatin Conformation Capture (3C) could provide information on 3D chromatin structures that occur in living cells. Using 3C technology, we found that both DHS sites interact with the *WASp* proximal promoter, indicating a regulatory role of these DHS sites on *WASp* gene expression.

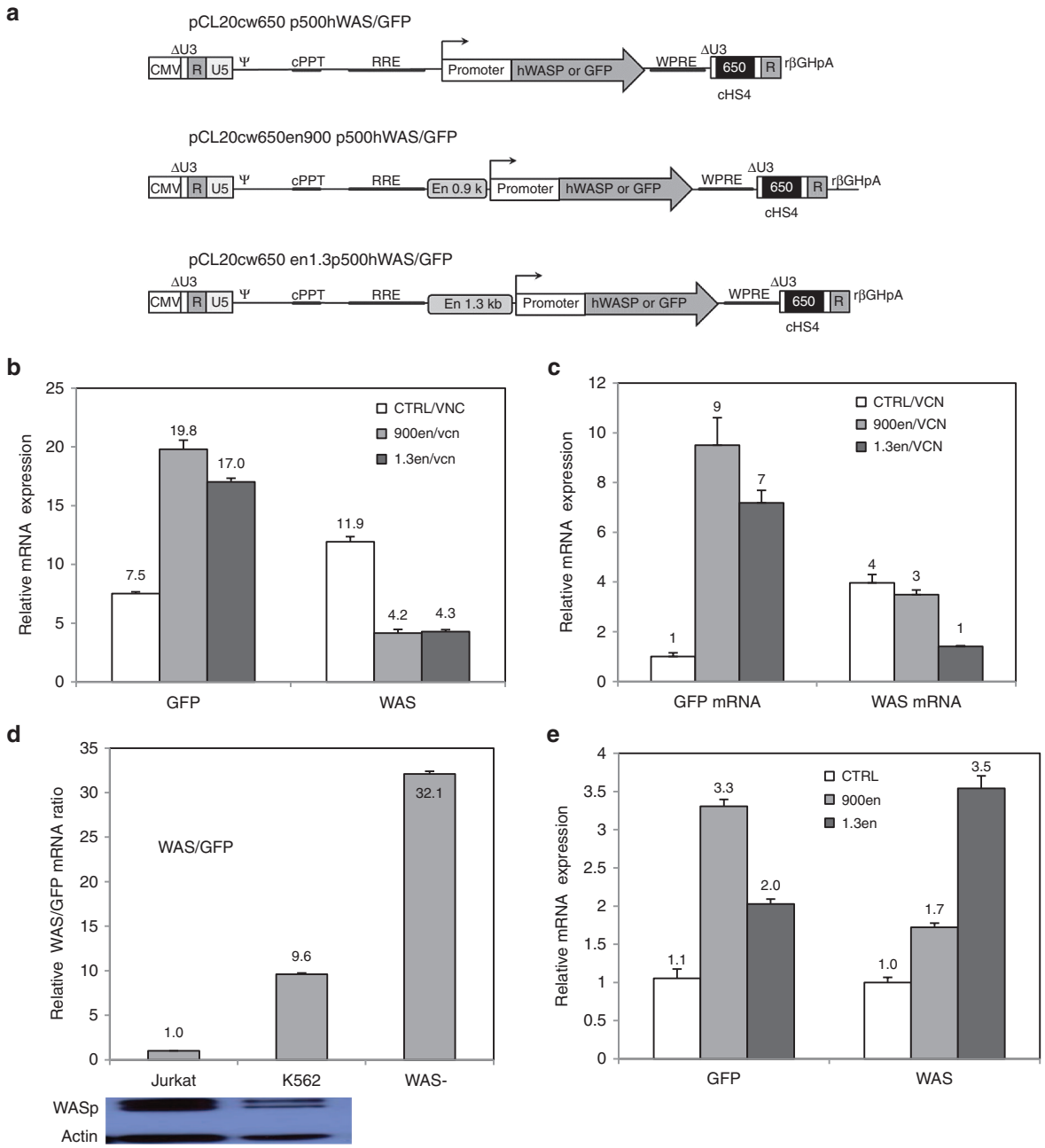

Figure 5 The evaluation of enhancer activity in stable expression system by using lenti vectors. (**a**) Lenti-vector constructions with insertion of potential enhancers. pCL20cw650 p500hWAS/GFP vectors were used as control vectors on top panel. Middle and bottom panels are vectors with 900 bp or 1.3 kb enhancer insertion upstream of p500 WAS promoter. (**b**) Relative GFP and WAS expression in K562 cells. K562 cells were infected by various lentivirus expressing either GFP or WAS at MOI 5. On day 6 cells were harvested for genomic DNA and mRNA. GFP and WAS expressions were determined by q-RT-PCR using WPRE as probe, GAPDH as control. The relative expression was normalized by VCN. The q-RT-PCR was run with triplicated samples . The error bars represent the mean ± SE (standard error). (**c**) Relative GFP or WAS expression in hCD34 cells. Human CD34 cells were transduced by various vectors at MOI 20. On day 6, cells were harvested for gDNA and mRNA. GFP and WAS expressions were determined by q-RT-PCR using WPRE as probe, GAPDH as control. The relative expression was normalized by VCN. (**d**) Relative WAS/GFP mRNA expression ratio from transduced control lenti-vectors. Cells with different level of endogenous WAS expression, Jurkat-T, K562, EBV WAS-, were transduced with pCL20cw650 p500hWAS and pCL20cw650 p500hGFP, On day 6, cells were harvested for genomic DNA and mRNA. GFP and WAS expressions were determined by q-RT-PCR using WPRE as probe, GAPDH as control. The relative expression was normalized by VCN. The western blot of WAS protein was shown on the bottom. (**e**) Relative GFP or WAS expression in murine WAS deficient BM cells. Both Lin– and Lin+ murine BM cells were isolated and transduced by various vectors at MOI 5. On day 6, cells were harvested for genomic DNA and mRNA. GFP and WAS expressions were determined by q-RT-PCR using WPRE as probe, GAPDH as control. The relative expression was normalized by VCN. BM, bone marrow; GFP, green fluorescence protein; GAPDH, glyceraldehyde 3-phosphate dehydrogenase; MOI, multiplicity of infection; qRT-PCR, quantitative real-time polymerase chain reaction; VCN, vector copy number; WAS, Wiskott-Aldrich Syndrome; WPRE, woodchuck hepatitis virus post-transcriptional regulatory element.

The proximal DHS-A site exhibited stronger cell type specific interaction than the DHS-B site which was further confirmed by more 3C experiments with different primers. Therefore we tested the regulatory role of the DNA fragments around this DHS site initially. Unexpectedly, the DNA fragments around proximal DHS-A did not show any enhancer activity in Jurkat-T cells but rather had silencer activity. The silencer activity was confirmed by using two different reporter vectors either driven by the 31 bp synthetic minipromoter or by the p500 WASp proximal promoter.

Chromatin profiling has emerged as a powerful means of detection of regulatory activity. Recently, whole genome chromatin states profiling including 15 different chromatin states such as enhancers and promoters for nine cell types have been mapped and analyzed. All the data are available in the human genome browser (UCSC). Using the human genome browser, we aligned the chromatin profile for several cell lines on the WASp genomic (upstream) region. We found a putative strong enhancer located between the *WASp* distal and proximal promoters and a putative weak enhancer immediately downstream of the WAS distal promoter (Supplementary Figure S2a). The cloning of the putative strong enhancer was not successful due to repeat sequences. We have cloned the putative weak enhancer into a luciferase reporter driven by p500 *WASp* promoter. Only the 1.5 kb fragment had very weak enhancer activity in Jurkat-T cells (Supplementary Figure S2b), indicating that bio-informatics identified DNA regulatory elements correlate with reporter assay results.

To hunt for a potential hematopoietic specific enhancer, we did a genome wide search using the human genome browser. We found a 5 kb enhancer region in K562 and GM12878 (derived from peripheral blood mononuclear cells EBV transformation) but not in nonhematopoietic cells. The region partially overlaps the previously identified DHS-B region. This 5 kb contains four DHS sites, the left two are hematopoietic specific, and the right two are T cell specific. As we expected, the left fragment from hematopoietic specific DHS region was an enhancer in K562 cells, while the right DNA fragment from the T cell specific DHS region was an enhancer in Jurkat-T cells (Figure 4b). Interestingly the left fragment did not enhance gene activity in Jurkat-T cells. As we later tested the DHS-B region we found an enhancer that works in several hematopoietic cells, but not in Jurkat-T cells. Perhaps Jurkat-T cells are a special T cell line which behaves differently from primary cells in regulating the *WASp* gene activity.

Subsequently, we have investigated the enhancer activities in the stable expression system by using lentiviral vectors with GFP or WASp expression as the readout. The enhancer increased GFP expression but decreased WASp mRNA levels in several different cell types, such as K562 and CD34 cells. After comparing the mRNA level for GFP or WASp driven by the same promoter (enhancer), we realized that enhancer activity may not only depend on the promoter but also may depend on what gene it regulates. The coding sequences or introns may have additional regulatory elements and work together with the upstream regulatory elements to control gene activity (*e.g.*, the globin gene expression). We presumed that endogenous WAS protein may negatively regulate its own expression. Finally, in WASp deficient murine cells, we found that the enhancer increased the expression of both GFP and WAS, implying that endogenous WAS may negatively regulate its own expression by an unknown mechanism. Taken together, we identified a hematopoietic specific enhancer that potentially may be used for gene therapy of hematological diseases. Recent gene therapy trials presented by Aiuti *et al.*[35] and Pala *et al.*[36] revealed that physiological levels of WASp can be achieved using a 1.6 kb *WASp* promoter. Aiuti *et al.*[35] used reduced intensity conditioning regimen to treat three patients. All three patients showed stable engraftment of WASP-expressing cells and improvements in platelet counts, immune functions, and clinical score. Pala *et al.*[36] focused on immune functional correction and B cell tolerance. They demonstrated that WASp plays an essential role in the establishment of early B cell tolerance in humans. Given the extensive preclinical data establishing potential efficacy, in our judgment, additional animal studies are not needed. The LV-w1.6W Lentiviral-mediated gene correction of WASp restores both central and peripheral B cell tolerance checkpoints in WAS patients. In our preclinical study,[34] the 1.6 kb promoter was comparable with the p500 promoter and weaker than the MND promoter. We hypothesize that the p500 promoter plus 1.3 kb or 900 bp enhancer will have better WAS expression and thereby more clinical benefits for patients.

MATERIALS AND METHODS

Tissue culture

The 293T, K562, and HeLa cells were maintained in Dulbecco's Modified Eagle's Media (DMEM) (Cellgro, Manassas, VA) 10% fetal bovine serum (FBS; Thermo Fisher Scientific, Waltham, MA) 1x penicillin/streptomycin (Gibco, Grand Island, NY) referred to hereafter as D10 (Cellgro), 10% FBS (Thermo Fisher Scientific) (mmol/l L-Glutamine (Gibco)), 10% FBS (Thermo Fisher Scientific), Jurkat-T, HEL 92.1.7 and U937 cells were maintained in Roswell Park Memorial Institute (RPMI 1640) medium (Cellgro), 10% FBS (Cellgro) 10% FBS (Thermo Fisher Scientific) 2 mmol/l L-Glutamine (Thermo Fisher Scientific1x penicillin/streptomycin (Thermo Fisher Scientific). DS-1 cells were maintained in RPMI 1640 medium with 2 mmol/l L-Glutamine adjusted to contain 1.5 gm/l, sodium bicarbonate, 4.5 gm/l glucose and 10 mmol/l 2-2 hydroxyethylpiperazine-N9-2-ethanesufonic acid (HEPES), and 1.0 mmol/l sodium pyruvate supplemented with 10 U/ml IL-6 and 10% FBS.

Bioinformatics

Software from Biobase (biological databases company: https://portal.biobase-international.com/cgi-bin/portal/login.cgi) was used for searching for putative transcriptional factor binding sites. Human genome browser (UCSC, http://genome.ucsc.edu/cgi-bin/hgGateway) was used for genome wide searching for enhancer or DHS sites and the alignment of these sites in different cell types.

3C experiments

The construction of 3C libraries and 3C control libraries was based on a published protocol[33] using BglII digestion. Briefly, 5×10^7 cells suspended in fresh medium, were incubated with formaldehyde (1% final concentration) at room temperature for 10 minutes to cross-link DNA-protein complexes, followed by the addition of 2.5 mol/l Gly (0.1 mol/l final concentration) to stop cross-linking. Cells were then spun down for 10 minutes at 400*g* in a clinical centrifuge and resuspended in 500 µl cold lysis buffer (10 mmol/l Tris-HCl pH 8.0, 10 mmol/l sodium chloride, 0.2% Igepal CA-630) on ice for at least 15 minutes. Subsequently cells were disrupted on ice with 15 strokes in a Dounce homogenizer and the cell lysate was transferred to a 1.7-ml microfuge tube and centrifuged for 5 minutes at 2,000*g* at room temperature. The pellet (nuclei with DNA-protein cross-linked complexes) was resuspended in 1× restriction buffer (100 mmol/l NaCl, 50 mmol/l Tris-HCl PH7.9, 10 mmol/l $MgCl_2$ 100 ug/ml bovine serum albumin (BSA)) and subjected to restriction digestion by BglII overnight at 37°C. After heat inactivation of the enzyme at 65°C for 30 minutes, the reactions were transferred to 15 ml conical tubes and incubated in the presence of ligase and 0.1% sodium dodecyl sulfate (SDS) for 2 hours at 16°C. Proteinase K was added to reverse the cross-linked DNA protein complexes. After phenol/chloroform extraction, the DNA (3C library) was precipitated by adding 0.1 volumes of 3 mol/l sodium acetate, pH 5.2 and 2.5 volumes of cold absolute ethanol. The relative interaction frequencies between the proximal WASp promoter and fragments of interest were calculated based on the PCR band intensity normalized to that of a 3C control library.

Reporter plasmids construction

The fragments of interest were PCR amplified from genomic DNA from the Jurkat-T cells and cloned into the pGL4.23Luc vector (Promega, Madison, WI) at HindIII/NheI restriction sites yielding reporter vectors driven by a 31 bp minimal synthetic promoter. To generate the reporter vector driven by p500hWASP proximal promoter, the 500 bp WASp promoter fragment was cloned into pGL4.23Luc at HindIII/NcoI restriction sites after removal of the minipromoter yielding pGLp500. Subsequently, various DNA regulatory fragments were inserted into pGLp500 at NheI /HindIII restriction site, yielding reporter vectors in the pGLp500 backbone. Lentiviral vector construction was based on the pCL20 vector backbone, pCL20cw650p500hWAS or pCL20cw650p500GFP.[34] The 900 bp or 1.3 kb putative enhancer fragments were inserted into these vectors at MluI/BsmBI restriction sites yielding four lenti-vectors, pCl20cw650en900p500hWAS, pCl20cw650en1.3p500hWAS, pCl20cw650en900p500GFP, and pCl20cw650en1.3p500GFP.

Reporter assay

For each transfection, cells were seeded onto 24 well plates at 10^5 cells/well and 0.5 µg reporter DNA and 0.05 µg renilla luciferase plasmid pGL4.74hRluc/TK (Promega) were added. We used Fugene HD (Promega) for 293T and HEL 92.1.7 transfection and Lipofectamine (Life Technologies, Grand Island, NY) for U937 and K562 cells according to the recommended protocols. DS-1 cells and Jurkat-T cells were transfected by neon electroporation (Life Technologies). The electroporation parameters for Jurkat-T cells were 1,700 volts per 20 ms for one pulse and for DS-1 at 1350 volt/40 ms for one pulse. Cells were harvested and lysed ~42 hours post-transfection. Both firefly and renilla luciferase were read on a moonlight 2010 Luminometer (BD Biosciences, 2350Qume Drive, San Jose, CA 95131). The firefly luciferase activity was normalized to the renilla luciferase activity.

Virus production

The 293T cells were plated at $3.5–4.0 \times 10^6$ cells per 10 cm dish in 10 ml of D10 24 hours prior to $CaCl_2$ mediated transfection with 6 µg pCAGK-GP1.1R, 2 µg pCAGA-RtR2, 2 µg pCAG-VSVg and 10 µg of transfer vector. Cells were rinsed 24 hours after transfection with 1× PBS and 10 ml D10 were added. After a further 24 hours, the virus was collected, clarified via filtration through a 0.2 µm cellulose acetate filter, aliquoted, stored at −80°C and subsequently titrated on HeLa cells as previously described.[35]

qPCR and qRT-PCR for VCN and gene expression

Proviral copy number was determined using qRT-PCR on genomic DNA samples prepared using the Gentra genomic DNA Isolation Kit (Qiagen, Valencia, CA). Each sample was assayed for the vector sequence (Fwd 5′CCTCAGACCCTTTTAGTCAGTGT3′, Rev 5′CTTTCGCTTTCAAGTCCCTGTTC3′ Probe 5′6FAMCCACTGCTAGAGATTTTMGBNFQ3 ′) and as a control assay detecting RNase P (cat#4403328, Life technologies). RNA was prepared using a RNA isolation kit (Qiagen). The relative RNA level of GFP or WASp was determined by qRT-PCR using the WPRE probe (WPRE-F 5′ CCCGTTGTCAGGCAACGTGGC 3′, WPRE-R 5′ GAGCTGACAGGTGGTGGCAA 3′, the probe WPRE Probe 5′ FAM TGCTGACGCAACCCCCACTGGT TAMRA3′) and control primers for detecting glyceraldehyde 3-phosphate dehydrogenase (GAPDH) (Life technologies)

KEY POINTS

- Two WAS downstream DHS sites were found to interact with the WASp proximal promoter.
- One of the DHS sites proved to be a hematopoietic specific enhancer.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ACKNOWLEDGMENTS

The authors thank John Stamatoyannopoulos for providing the map shown in Figure 1a. We also acknowledge the excellent administrative assistance provided by Pat Streich. The funds to support this project were provided by the American Lebanese Syrian Associated Charities. J.Z. and A.W.N. wrote the paper. J.Z., I.M.J., and M.W. performed research.

REFERENCES

1. O'Reilly, M, Kohn, DB, Bartlett, J, Benson, J, Brooks, PJ, Byrne, BJ *et al.* (2013). Gene therapy for rare diseases: summary of a National Institutes of Health workshop, September 13, 2012. *Hum Gene Ther* **24**: 355–362.
2. Rosen, FS, Cooper, MD and Wedgwood, RJ (1995). The primary immunodeficiencies. *N Engl J Med* **333**: 431–440.
3. Derry, JM, Ochs, HD and Francke, U (1994). Isolation of a novel gene mutated in Wiskott-Aldrich syndrome. *Cell* **79**: following 922.
4. Derry, JM, Ochs, HD and Francke, U (1994). Isolation of a novel gene mutated in Wiskott-Aldrich syndrome. *Cell* **78**: 635–644.
5. Kwan, SP, Hagemann, TL, Radtke, BE, Blaese, RM, and Rosen, FS (1995). Identification of mutations in the Wiskott-Aldrich syndrome gene and characterization of a polymorphic dinucleotide repeat at DXS6940, adjacent to the disease gene. *Proc Natl Acad Sci USA* **92**: 4706–4710.
6. Snapper, SB and Rosen, FS (1999). The Wiskott-Aldrich syndrome protein (WASP): roles in signaling and cytoskeletal organization. *Annu Rev Immunol* **17**: 905–929.
7. Hagemann, TL and Kwan, SP (1999). The identification and characterization of two promoters and the complete genomic sequence for the Wiskott-Aldrich syndrome gene. *Biochem Biophys Res Commun* **256**: 104–109.
8. Petrella, A, Doti, I, Agosti, V, Giarrusso, PC, Vitale, D, Bond, HM *et al.* (1998). A 5′ regulatory sequence containing two Ets motifs controls the expression of the Wiskott-Aldrich syndrome protein (*WASP*) gene in human hematopoietic cells. *Blood* **91**: 4554–4560.
9. Parolini, O, Berardelli, S, Riedl, E, Bello-Fernandez, C, Strobl, H, Majdic, O *et al.* (1997). Expression of Wiskott-Aldrich syndrome protein (*WASP*) gene during hematopoietic differentiation. *Blood* **90**: 70–75.
10. Stewart, DM, Treiber-Held, S, Kurman, CC, Facchetti, F, Notarangelo, LD and Nelson, DL (1996). Studies of the expression of the Wiskott-Aldrich syndrome protein. *J Clin Invest* **97**: 2627–2634.
11. Cavalli, G and Misteli, T (2013). Functional implications of genome topology. *Nat Struct Mol Biol* **20**: 290–299.
12. Gibcus, JH and Dekker, J (2013). The hierarchy of the 3D genome. *Mol Cell* **49**: 773–782.
13. Misteli, T (2007). Beyond the sequence: cellular organization of genome function. *Cell* **128**: 787–800.
14. Misteli, T and Soutoglou, E (2009). The emerging role of nuclear architecture in DNA repair and genome maintenance. *Nat Rev Mol Cell Biol* **10**: 243–254.
15. Cremer, T and Cremer, C (2001). Chromosome territories, nuclear architecture and gene regulation in mammalian cells. *Nat Rev Genet* **2**: 292–301.
16. Dean, A (2011). In the loop: long range chromatin interactions and gene regulation. *Brief Funct Genomics* **10**: 3–10.
17. Dekker, J (2008). Gene regulation in the third dimension. *Science* **319**: 1793–1794.
18. Dekker, J, Rippe, K, Dekker, M and Kleckner, N (2002). Capturing chromosome conformation. *Science* **295**: 1306–1311.
19. Miele, A and Dekker, J (2008). Long-range chromosomal interactions and gene regulation. *Mol Biosyst* **4**: 1046–1057.
20. Miele, A, Gheldof, N, Tabuchi, TM, Dostie, J and Dekker, J (2006). Mapping chromatin interactions by chromosome conformation capture. *Curr Protoc Mol Biol* **Chapter 21**: Unit 21.11.
21. Splinter, E, Grosveld, F and de Laat, W (2004). 3C technology: analyzing the spatial organization of genomic loci *in vivo*. *Methods Enzymol* **375**: 493–507.
22. Baù, D, Sanyal, A, Lajoie, BR, Capriotti, E, Byron, M, Lawrence, JB *et al.* (2011). The three-dimensional folding of the α-globin gene domain reveals formation of chromatin globules. *Nat Struct Mol Biol* **18**: 107–114.
23. Wang, KC, Yang, YW, Liu, B, Sanyal, A, Corces-Zimmerman, R, Chen, Y *et al.* (2011). A long noncoding RNA maintains active chromatin to coordinate homeotic gene expression. *Nature* **472**: 120–124.
24. Andrey, G, Montavon, T, Mascrez, B, Gonzalez, F, Noordermeer, D, Leleu, M *et al.* (2013). A switch between topological domains underlies HoxD genes collinearity in mouse limbs. *Science* **340**: 1234167.
25. Berlivet, S, Paquette, D, Dumouchel, A, Langlais, D, Dostie, J and Kmita, M (2013). Clustering of tissue-specific sub-TADs accompanies the regulation of HoxA genes in developing limbs. *PLoS Genet* **9**: e1004018.

26. Hughes, JR, Roberts, N, McGowan, S, Hay, D, Giannoulatou, E, Lynch, M *et al.* (2014). Analysis of hundreds of cis-regulatory landscapes at high resolution in a single, high-throughput experiment. *Nat Genet* **46**: 205–212.
27. Gross, DS and Garrard, WT (1988). Nuclease hypersensitive sites in chromatin. *Annu Rev Biochem* **57**: 159–197.
28. Elgin, SC (1984). Anatomy of hypersensitive sites. *Nature* **309**: 213–214.
29. Wu, C (1980). The 5′ ends of Drosophila heat shock genes in chromatin are hypersensitive to DNase I. *Nature* **286**: 854–860.
30. Hesselberth, JR, Chen, X, Zhang, Z, Sabo, PJ, Sandstrom, R, Reynolds, AP *et al.* (2009). Global mapping of protein-DNA interactions *in vivo* by digital genomic footprinting. *Nat Methods* **6**: 283–289.
31. Friedle, H, Rastegar, S, Paul, H, Kaufmann, E and Knöchel, W (1998). Xvent-1 mediates BMP-4-induced suppression of the dorsal-lip-specific early response gene XFD-1′ in Xenopus embryos. *EMBO J* **17**: 2298–2307.
32. Ernst, J, Kheradpour, P, Mikkelsen, TS, Shoresh, N, Ward, LD, Epstein, CB *et al.* (2011). Mapping and analysis of chromatin state dynamics in nine human cell types. *Nature* **473**: 43–49.
33. Dostie, J and Dekker, J (2007). Mapping networks of physical interactions between genomic elements using 5C technology. *Nat Protoc* **2**: 988–1002.
34. Koldej, RM, Carney, G, Wielgosz, MM, Zhou, S, Zhan, J, Sorrentino, BP *et al.* (2013). Comparison of insulators and promoters for expression of the Wiskott-Aldrich syndrome protein using lentiviral vectors. *Hum Gene Ther Clin Dev* **24**: 77–85.
35. Dewey, RA, Avedillo Díez, I, Ballmaier, M, Filipovich, A, Greil, J, Güngör, T *et al.* (2006). Retroviral *WASP* gene transfer into human hematopoietic stem cells reconstitutes the actin cytoskeleton in myeloid progeny cells differentiated *in vitro*. *Exp Hematol* **34**: 1161–1169.
36. Aiuti, A, Biasco, L, Scaramuzza, S, Ferrua, F, Cicalese, MP *et al.* (2013). Lentivirus-based gene therapy of hematopoietic stem cells in Wiskott-Aldrich syndrome. *Science* **341**: 1233151–1233179.
37. Pala, F, Morbach, H, Castiello, MC, Schickel, JN, Scaramuzza, S, Chamberlain, N *et al.* (2015). Lentiviral-mediated gene therapy restores B cell tolerance in Wiskott-Aldrich syndrome patients. *J Clin Invest* **125**: 3941–3951.

Supplementary Information accompanies this paper on the *Molecular Therapy—Methods & Clinical Development* website (http://www.nature.com/mtm)