Cell Reports
Report

Three Distinct Patterns of Histone H3Y41 Phosphorylation Mark Active Genes

Mark A. Dawson,[1,2,3,4] Samuel D. Foster,[2,4] Andrew J. Bannister,[1,4] Samuel C. Robson,[1] Rebecca Hannah,[2] Xiaonan Wang,[2] Blerta Xhemalce,[1] Andrew D. Wood,[2] Anthony R. Green,[2,3] Berthold Göttgens,[2,5,*] and Tony Kouzarides[1,5,*]

[1]Gurdon Institute and Department of Pathology, Tennis Court Road, Cambridge, CB2 1QN, UK
[2]Department of Haematology, Cambridge Institute for Medical Research and The Wellcome Trust and MRC Stem Cell Institute
[3]Addenbrooke's Hospital
University of Cambridge, Cambridge, CB2 0XY, UK
[4]These authors contributed equally to this work
[5]These authors contributed equally to this work
*Correspondence: bg200@cam.ac.uk (B.G.), t.kouzarides@gurdon.cam.ac.uk (T.K.)
http://dx.doi.org/10.1016/j.celrep.2012.08.016

SUMMARY

The JAK2 tyrosine kinase is a critical mediator of cytokine-induced signaling. It plays a role in the nucleus, where it regulates transcription by phosphorylating histone H3 at tyrosine 41 (H3Y41ph). We used chromatin immunoprecipitation coupled to massively parallel DNA sequencing (ChIP-seq) to define the genome-wide pattern of H3Y41ph in human erythroid leukemia cells. Our results indicate that H3Y41ph is located at three distinct sites: (1) at a subset of active promoters, where it overlaps with H3K4me3, (2) at distal *cis*-regulatory elements, where it coincides with the binding of STAT5, and (3) throughout the transcribed regions of active, tissue-specific hematopoietic genes. Together, these data extend our understanding of this conserved and essential signaling pathway and provide insight into the mechanisms by which extracellular stimuli may lead to the coordinated regulation of transcription.

INTRODUCTION

The canonical JAK2-STAT5 pathway is one of the most widely studied cellular signaling cascades and is critical for normal hematopoiesis (Decker and Müller, 2012). It is also now established that hyperactivation of JAK-STAT signaling is a common event in several hematological neoplasms (Chen et al., 2012). These findings underline the importance of a comprehensive understanding of the function of this signaling pathway.

Until recently, JAKs were largely ascribed the cytoplasmic function of phosphorylating STAT family members to promote STAT dimerization, nuclear translocation, and binding to *cis*-regulatory elements to regulate transcription (Decker and Müller, 2012). However, recent reports have highlighted that members of the JAK family also appear to have specific roles in the nuclei of various tissues (Dawson et al., 2009; Griffiths et al., 2011; Kamakura et al., 2004; Lefrancois-Martinez et al., 2011; Liu et al., 2011; Nilsson et al., 2006; Rui et al., 2010). Indeed, we have demonstrated that JAK2 functions as a histone tyrosine kinase, phosphorylating H3Y41 to regulate chromatin binding of HP1α in hematopoietic cells (Dawson et al., 2009). H3Y41ph represents the only characterized tyrosine phosphorylation in nonvariant mammalian histones. Here we report the genome-wide analysis of this modification by chromatin immunoprecipitation (ChIP) coupled to massively parallel DNA sequencing (ChIP-seq). These data provide insights into the mechanisms of transcription control exerted by this conserved and essential signaling pathway.

RESULTS

H3Y41ph Is Associated with Active Promoters

H3Y41ph is present at the transcriptional start site (TSS) of the *LMO2* gene, overlapping the active H3K4me3 modification (Dawson et al., 2009), and is linked to transcriptional activation. To ascertain whether this relationship is maintained on a genomic scale, we generated a global map of the spatial distribution of H3Y41ph and H3K4me3 by performing ChIP-seq using chromatin from human erythroid leukemia (HEL) cells. When ranked in ascending order of H3Y41ph and H3K4me3 levels, global analysis of gene promoters demonstrates that these two histone modifications are coincident at many loci (Figure 1A, areas I and II).

A detailed analysis then examined the global relationship between H3K4me3 and H3Y41ph across TSSs. 28,996 TSSs were split into 10 sets (the first nine segments have 2,900 and the tenth has 2,896) based on levels of H3K4me3 (Figure 1B). These sets were ranked according to the levels of H3K4me3 at the TSS, as this has previously been shown to quantitatively reflect transcriptional activity; the most highly transcribed genes have the highest levels of H3K4me3 and vice versa (Barski et al., 2007; Min et al., 2011). Once ranked according to H3K4me3 levels, we then looked at the distribution and levels of H3Y41ph across the same promoter cohorts. This analysis revealed that, when present, H3Y41ph is often at the TSS of active genes, where it correlates closely with H3K4me3

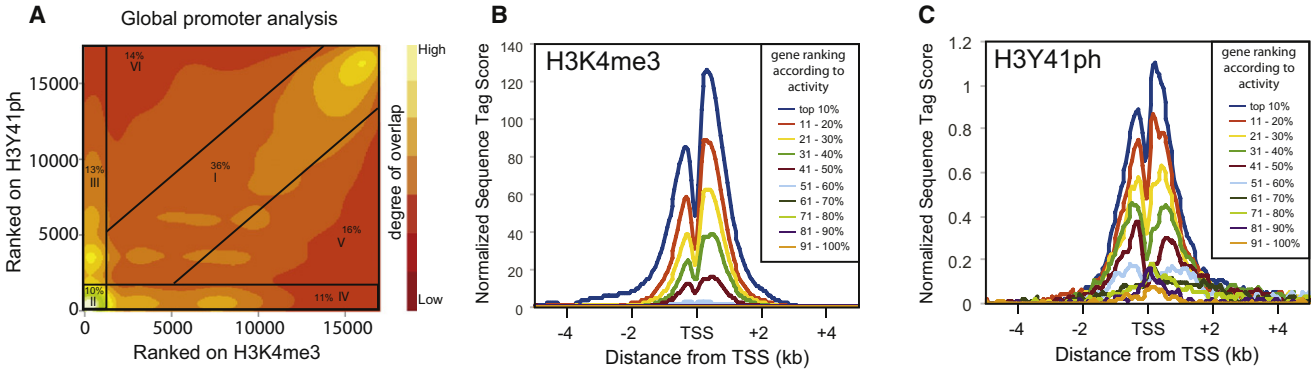


Figure 1. H3Y41ph Is Associated with Active Transcription

(A) Sequence tags for H3K4me3 and H3Y41ph were determined at all TSSs. Density profiles using quantile-normalized ranks of each region for their relative levels of H3K4me3 (*x* axis) and H3Y41ph (*y* axis) were divided into six subdivisions: (I) correlated H3K4me3 and H3Y41ph, (II) low-H3K4me3 and low-H3Y41ph, (III) low-H3K4me3 and mid/high-H3Y41ph, (IV) mid/high-H3K4me3 and low-H3Y41ph, (V) high-H3K4me3 and mid-H3Y41ph, and (VI) high-H3K4me3 and mid-H3Y41ph.

(B) Mean enrichment pattern for H3K4me3 was profiled across all annotated TSSs. Populations used were ranked according to levels of H3K4me3 enrichment and then split into 10 equal sets.

(C) The mean enrichment pattern for H3Y41ph was profiled across the same populations of TSSs as those described in (B).

See also Figure S1 and Table S1.

(Figure 1C). Importantly, selective inhibition of JAK2 activity for 4 hours demonstrated a significant global reduction of H3Y41ph at the TSS of genes marked with this histone modification (Figure S1). Overall, these data indicate that, as predicted from our previous analysis of the *LMO2* gene, H3Y41ph coincides with H3K4me3 surrounding the TSSs of a subset of transcriptionally active genes.

H3Y41ph Marks Intragenic Elements with STAT Binding Sites

Analysis of individual gene loci revealed that H3Y41ph is also enriched at noncoding regions of several JAK-STAT target genes, such as *CISH* and *PIM1* (Figures 2A and 2E). These sites have a high degree of sequence conservation and are also marked by H3K4me1, consistent with a potential function as distal *cis*-regulatory elements (Figure S2A). Indeed, some of these elements have already been defined as enhancer elements, such as those present at *ID1* (Wood et al., 2009) and *CISH* (Nagy et al., 2009). Together these data raise the possibility that H3Y41ph also delineates a subset of conserved elements outside of genes, at least some of which are known functional STAT5 enhancers.

It is firmly established that JAK2 and STAT5 physically and functionally interact (Barahmand-Pour et al., 1998). Thus far, only a few direct target genes bound by STAT5, such as *CISH*, *PIM1, ID1*, and *BCLXL*, have been characterized in the erythroid lineage (Dumon et al., 1999; Matsumoto et al., 1997; Mui et al., 1996; Wood et al., 2009). As HEL cells are a model cell line for malignant erythropoiesis, we investigated the genome-wide distribution of STAT5 by ChIP-seq to see if it coincides with H3Y41ph. Bioinformatic analysis of our newly generated STAT5 ChIP-seq data set identified 676 high-confidence STAT5 binding sites (Table S1). Importantly, the few previously known functional STAT5 binding sites, such as those present at *ID1* and *BCLXL*, were all identified as significant peaks (Figure S2B). Moreover, we now provide several hundred potential targets of STAT5 (Table S1), functional validation of which may uncover new insights into mechanisms of pathogenesis in JAK2-STAT5-driven neoplasms.

The quality of our STAT5 ChIP-seq data set was confirmed by an unbiased de novo motif discovery, which identified the previously reported TTCYNRGAA STAT consensus binding site as the only significantly enriched motif (Figure 2B). Furthermore, we observed an enrichment of STAT5 binding proximal to gene promoters, although the vast majority of STAT5 binding events were seen in nonpromoter regions (Figure 2C). These findings are analogous to those reported from recent genome-wide surveys of STAT5 binding in mouse liver (Zhang et al., 2012). Our ChIP-seq data indicate that both the H3Y41ph modification and STAT5 binding often occur at the same genomic regions (Figure 2E; Table S1). Indeed, when the spatial distribution of the H3Y41ph peaks are determined relative to the STAT5 peaks, we find that the location of H3Y41ph is directly coincident with the STAT5 binding event (Figures 2E and S2A).

H3Y41ph and STAT5 Binding Are Coincident and Concomitantly Regulated by JAK2

Although we observe a significant coincidence of H3Y41ph with STAT5, our data clearly demonstrate that the majority of H3Y41ph is independent of STAT5 (Tables S1 and S2). Nevertheless, the coincidence of H3Y41ph with STAT5 suggests that the interaction of JAK2/STAT5 may extend beyond the cytoplasm and culminate at the chromatin of a key subset of genes. We therefore asked if JAK2 signaling coordinates both the posttranslational modification of histones and the concomitant binding of STAT5 to its target DNA. For this, we performed both an H3Y41ph and STAT5 ChIP-seq analysis with chromatin from HEL cells that had been treated with the selective JAK2 inhibitor, TG101209, for 4 hours. Global analysis demonstrated that H3Y41ph and STAT5 binding events were significantly decreased after JAK2 inhibition (Figures 2D and 2E). These

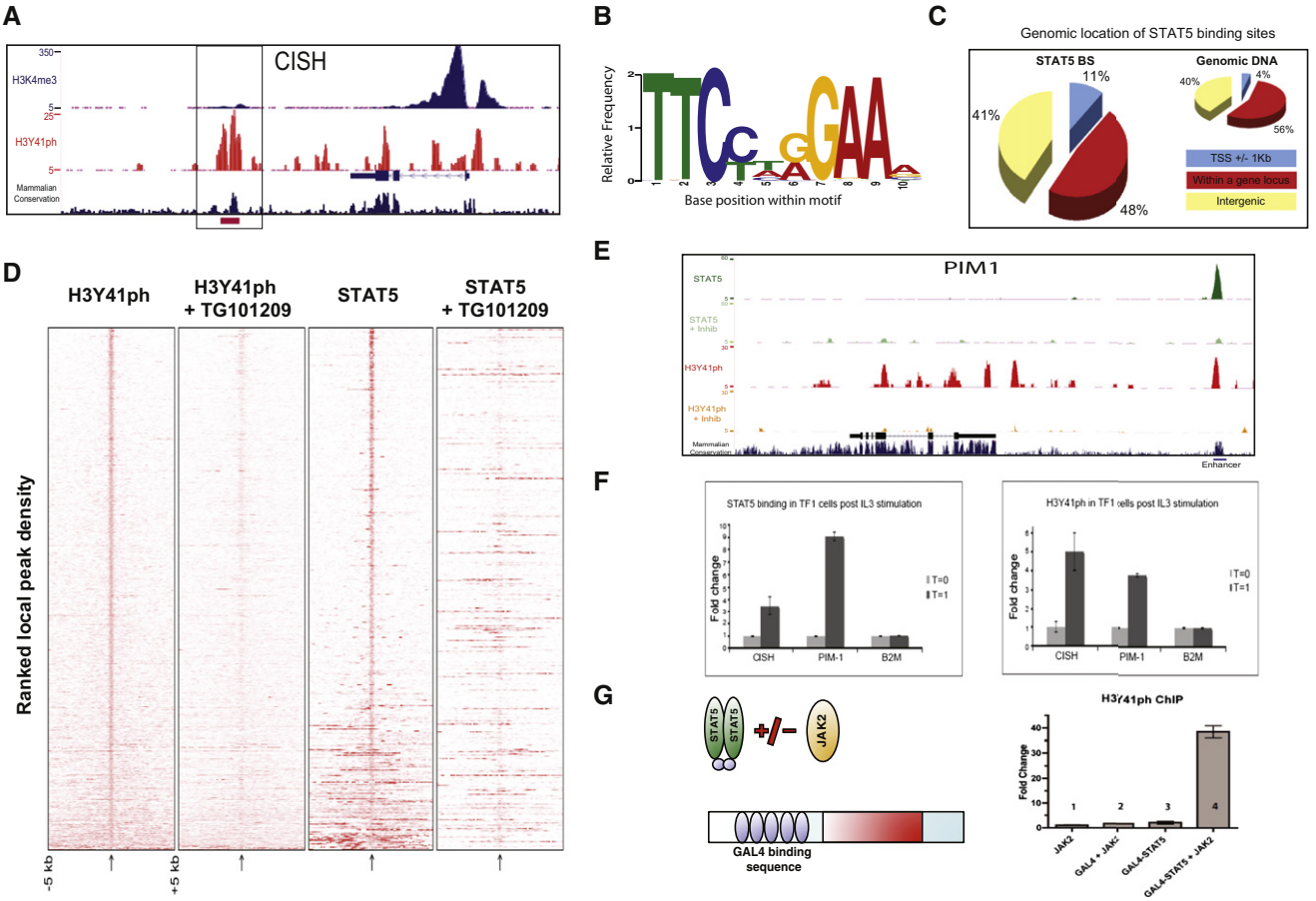


Figure 2. Genome-wide Analysis of STAT5 Binding and H3Y41ph

(A) ChIP-seq density profiles for H3K4me3 and H3Y41ph for the *CISH* gene show H3Y41ph enriched at the TSS and over a regulatory region (boxed) downstream of the gene.

(B) De novo motif discovery of high confidence STAT5 peaks. When displayed as a sequence logo, it is identical to the known STAT5 consensus motif.

(C) Pie chart illustrating the distribution of STAT5 with respect to genes. The number of STAT5 peaks within 1 kb of TSSs was determined. Remaining STAT5 peaks were divided into genic and intergenic. The percentage of STAT5 at each region is shown (large pie chart). For comparison, the percentage of the genome assignable to each region is also shown (small pie chart).

(D) Density of ChIP-seq reads across H3Y41ph and STAT5 peak regions before and after JAK2 inhibition shown as heat maps centered on peak summits with 5 kb of flanking sequence either side. Darker color indicates higher density of reads. H3Y41ph peak regions were rank ordered based on signal intensity at the peak center divided by the average signal across the +/− 5 kb flanking region and were arranged based on relative local enrichment. The same order was used to display the heat maps of H3Y41ph + TG101209. The same procedure was used to generate the analogous heat maps for the STAT5 peaks.

(E) The ChIP-seq density profiles for STAT5 and H3Y41ph, before and after inhibition of JAK2, shows both STAT5 and H3Y41ph enrichment is markedly reduced at an enhancer of *PIM1*.

(F) Cytokine stimulation with IL3 leads to increased STAT5 binding (left panel) and H3Y41ph (right panel) at *cis*-regulatory elements for both *PIM1* and *CISH*.

(G) STAT5 and JAK2 are required for H3Y41 phosphorylation. 293T cells harboring an integrated GAL4 reporter were transfected with the plasmids encoding proteins, as illustrated. H3Y41ph levels at the promoter of the integrated reporter were monitored by ChIP-PCR. Data were normalized for H3 occupancy and are represented relative to H3Y41ph levels in lane 1. Experiments were performed in biological triplicate. Each amplicon was analyzed in duplicate each time and error bars represent the standard deviation for each amplicon.

See also Figure S2 and Table S1.

findings are consistent with our previous results derived from western blot analyses, where we see a global reduction in H3Y41ph following exposure of HEL cells to a 4 hr treatment with TG101209 (Dawson et al., 2009). The dual regulation of H3Y41ph and STAT5 binding by JAK2 signaling is specifically demonstrated at the *PIM1* locus in Figure 2E. Several loci identified in the ChIP-seq data were independently validated by ChIP-PCR at various JAK2-STAT5 target genes in HEL cells (Figure S2C). Moreover, these results were validated using AT9283 (Dawson et al., 2010), a second chemically distinct JAK2 inhibitor (Figure S2D). Finally, we demonstrate the broader application of these data by observing similar findings in SET-2 cells, a megakaryoblastic cell line containing JAK2V617F (Figure S2E). Taken together, these data are consistent with JAK2 being a major histone kinase responsible for H3Y41ph and indicate that signaling via JAK2 concomitantly

regulates both H3Y41ph and STAT5 binding at certain JAK2/STAT5 target genes.

We next asked whether cytokine induction via JAK2 dynamically regulates the appearance of H3Y41ph and the binding of STAT5 on cytokine-regulated genes. TF1 cells (cytokine-dependent human erythroid-leukemia cells diploid for wild-type [WT] JAK2) were stimulated for 1 hr with IL3, a cytokine that signals primarily via JAK2 (Decker and Müller, 2012). The expression of two well-characterized cytokine inducible JAK2-STAT5 target genes, *PIM1* (Mui et al., 1996) and *CISH* (Matsumoto et al., 1997), were markedly increased by IL3, whereas the transcription of the *β2M* housekeeping gene was unchanged (data not shown). Importantly, ChIP analyses from the same cell populations clearly demonstrated that the increased gene expression of *PIM1* and *CISH* was linked to an increase of both H3Y41ph and STAT5 at the respective *PIM1* and *CISH cis*-regulatory elements (Figure 2F). Notably, levels of H3Y41 and STAT5 were unaltered at the promoter of the *β2M* housekeeping gene (Figure 2F). These data provide further evidence for a communication between JAK2 and STAT5, which extends to the chromatin interface in order to regulate transcription of these genes.

To address the possibility that STAT5 plays a direct role in recruiting JAK2 to phosphorylate H3Y41 at a subset of genomic loci, we initially performed small interfering RNA (siRNA)-mediated knockdown of STAT5 in HEL cells and monitored both transcriptional activity and H3Y41ph/STAT5 binding at known JAK2/STAT5 target genes. However, despite a 90% knockdown of STAT5, transcription and H3Y41ph levels were only modestly decreased at some (e.g., *CISH*) but not all (e.g., *PIM1*) JAK2/STAT5 target genes (Figures S1F–S1H). However, since trace amounts of STAT5 were sufficient for essentially normal development and function of the original STAT5 knockout mice (Teglund et al., 1998), the absence of a strong knock down phenotype observed here most likely reflects the fact that the residual 10% of STAT5 is sufficient for near-full activity. An alternative nonmutually exclusive explanation is that other STAT family members may compensate for the reduced levels of STAT5, as previously reported (Lim and Cao, 2006).

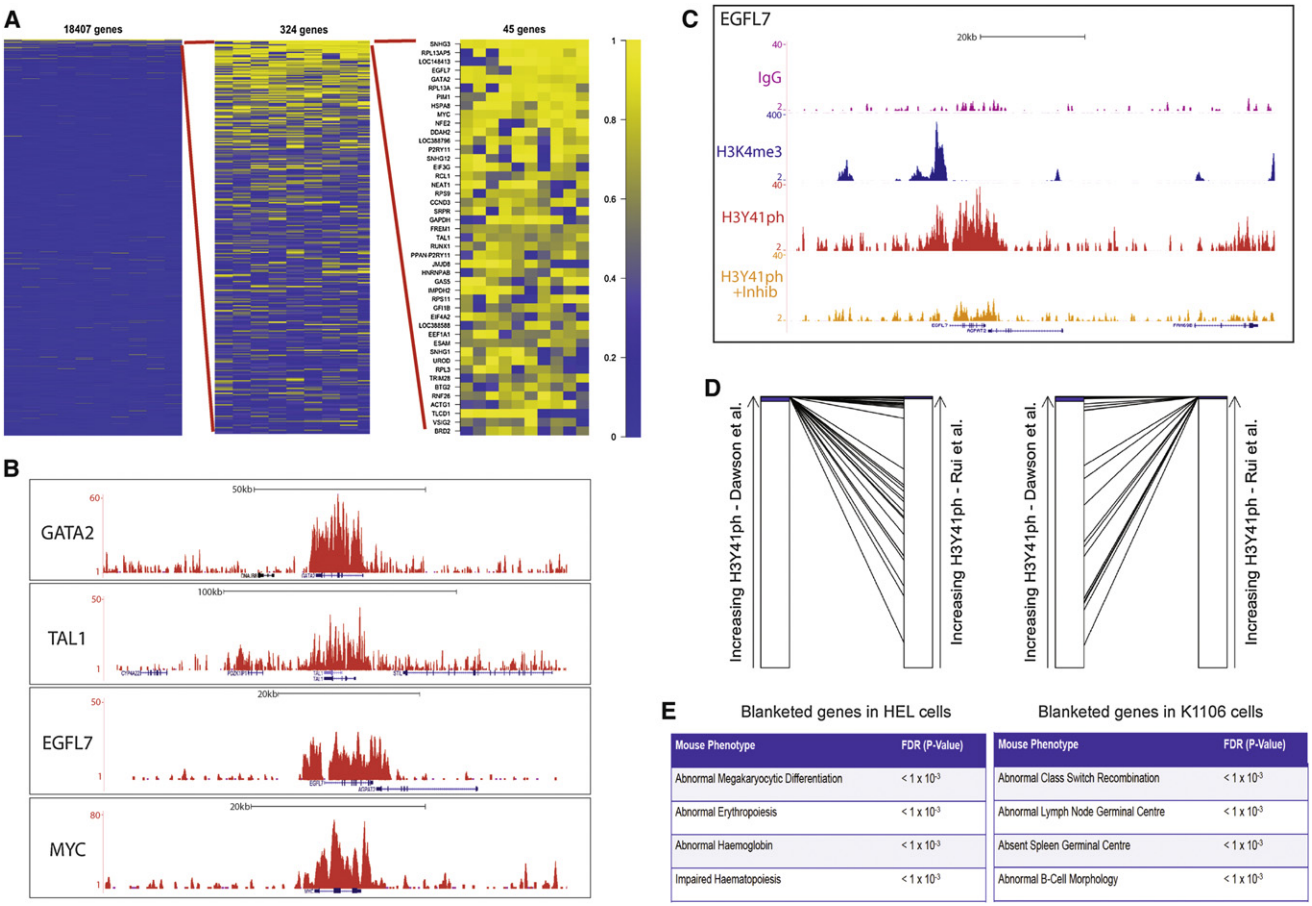
Due to the limitations of the siRNA approach, we chose to investigate the role of STAT5 and JAK2 in mediating H3Y41ph using an alternative strategy. Our method employed a chromosomally integrated Gal4 reporter assay in 293T cells. This approach allowed us to specifically address the issue of whether STAT5 and JAK2 are both required for H3Y41ph at a synthetically engineered chromatinized locus (Figure 2G). In this assay, STAT5 is exogenously expressed in-frame with a Gal4 DNA binding domain (DBD). By using ChIP assays to monitor H3Y41ph levels at the chromatinized Gal4 reporter, we clearly demonstrate that both STAT5 (targeted to chromatin via the Gal4 DBD) and JAK2 are necessary to phosphorylate H3Y41 at this site (Figures 2G, S1I, and S1J). Importantly, GAL4-STAT5 by itself does not lead to an increase in H3Y41ph, as 293T cells contain little, if any JAK2 (Figure 2G). In this assay, we tested human WT STAT5A and STAT5B as well as constitutively active murine STAT5A and STAT5B and observed similar results (Figures 2G, S1G, and S1H). These mechanistic studies are consistent with our genome-wide correlation and together suggest that STAT5 may recruit JAK2 to certain genomic loci to phosphorylate H3Y41.

H3Y41ph Blankets Key Tissue-Specific Genes

Interrogation of the genomic distribution of H3Y41ph revealed a third pattern unique to this histone modification, i.e., "blanketing" of H3Y41ph over the entire coding region of a gene (Figures 3A–3C and S3A). Notably, H3Y41ph invariably also spans the TSS at these genes (Figure 3B). In order to identify genes most prominently blanketed by H3Y41ph, we analyzed the read density and distribution of H3Y41ph over the entire coding region of all annotated genes. This shows that H3Y41ph heavily blankets only a small fraction of genes (Figures 3A and S3A), many of which are key hematopoietic genes. Indeed, 71% of the identified blanketed genes are enriched in hematopoietic gene sets, and this group includes several genes, such as *GATA2*, *EGFL7*, and *TAL1* (Figure 3B), that have previously been implicated in both normal hematopoiesis and the pathogenesis of acute myeloid leukemia (Table S1). To control for copy number alterations, which could lead to overrepresentation in sequencing data, giving a spurious impression that regions of the genome were "blanketed," we performed an immunoglobulin G (IgG) control from the same chromatin preparation as the H3Y41ph results. When the IgG ChIP-seq data are compared with published genome-wide SNP6 array data in HEL cells, we clearly observe that the "background" of sequencing signal varies, as expected with the copy number (Figure S3B). In contrast, when we look at a locus that is blanketed by H3Y41ph, the IgG trace over the blanketed region does not mirror the pattern of H3Y41ph (Figure S3B). Similar results were also observed for the STAT5 ChIP-seq data (data not shown). Furthermore, a similar "blanketing" analysis for another active histone modification, H3K4me3, does not recapitulate these findings (Table S1).

Interestingly, the H3Y41ph-blanketed genes are not the most highly expressed genes in HEL cells (Figure S3C). Thus, the relatively high level of H3Y41ph spread throughout these genes does not seem to be a consequence of a high rate of transcription. Finally, the vast majority of heavily H3Y41ph-blanketed genes that are responsive to JAK2 inhibition do not contain STAT5 binding events (Table S2). These data serve to demonstrate that, while STAT5 and potentially other STATs may target JAK2 to certain genomic loci, this phenomenon is not ubiquitous, and the majority of genes appear to be marked by H3Y41ph in a STAT-independent manner. Nevertheless, it is clear that JAK2 inhibition is sufficient to substantially reduce H3Y41ph and gene expression, even at these STAT-independent loci (Figure 3C; Table S2).

Together these data underline the fact that blanketing is neither a simple attribute of transcriptional activity nor a common feature of other active histone modifications. This raises the possibility that H3Y41ph may "blanket" genes whose expression is necessary to propagate or maintain a particular cellular state. Consistent with this notion is the fact that unbiased interrogation of the H3Y41ph-blanketed genes in HEL cells for common knock-out mouse phenotypes using the Genomic Regions Enrichment of Annotations Tool (GREAT) (McLean



Mouse Phenotype	FDR (P-Value)
Abnormal Megakaryocytic Differentiation	$< 1 \times 10^{-3}$
Abnormal Erythropoiesis	$< 1 \times 10^{-3}$
Abnormal Haemoglobin	$< 1 \times 10^{-3}$
Impaired Haematopoiesis	$< 1 \times 10^{-3}$

Mouse Phenotype	FDR (P-Value)
Abnormal Class Switch Recombination	$< 1 \times 10^{-3}$
Abnormal Lymph Node Germinal Centre	$< 1 \times 10^{-3}$
Absent Spleen Germinal Centre	$< 1 \times 10^{-3}$
Abnormal B-Cell Morphology	$< 1 \times 10^{-3}$

Figure 3. H3Y41ph Blankets Tissue-Specific Genes

(A) Heat maps showing the coverage of H3Y41ph throughout gene bodies. For each gene, the transcribed region was divided into 10 equally sized sections, and the proportion of bases associated with H3Y41ph in each case was determined. Genes were ranked according to the median of these 10 values. Left panel, H3Y41ph distribution throughout all 18,407 genes; middle panel, H3Y41ph distribution throughout the "top" 324 genes (median H3Y41ph overlap score >0.2); right panel, blanketed genes, H3Y41ph overlap score >0.5 (45 genes).

(B) ChIP-seq screenshots showing four examples of H3Y41ph-blanketed genes.

(C) ChIP-seq screenshot showing the reduction in H3Y41ph blanketing at the *EGFL7* loci following TG101209 treatment.

(D) Genes were ranked based on their median H3Y41ph overlap for both our data (left bar) and that of Rui et al. (2010) (right bar); genes with equal scores were ranked randomly. Genes showing a median overlap >0.2 were classed as being blanketed in the corresponding data set (blue portion of bar). Lines were projected from the rank of the top most blanketed genes (median overlap >0.5; red portion of bar) from our data to the corresponding rank in the Rui et al. data (left panel) and vice versa (right panel).

(E) GREAT database mouse phenotypes associated with the blanketed genes in HEL cells and the lymphoid cell line K1106.

See also Figure S3 and Tables S1 and S2.

et al., 2010) revealed that these genes are critical for myelopoiesis, especially erythropoiesis and megakaryopoiesis (Figure 3E). To further test this hypothesis, we analyzed a published data set of H3Y41ph in a lymphoid cell line, K1106 (Rui et al., 2010). Interestingly, the genes blanketed by H3Y41ph in this data set were identified by the GREAT database to be germane to B-lymphopoiesis (Figure 3E).

Finally, when we intersected the two data sets, we found that the heavily blanketed gene sets in the two cell types show very little overlap (Figure 3D). These data suggest that H3Y41ph may uniquely delineate genes that define a particular cell lineage. Consistent with this concept is the fact that another independent gene set enrichment analysis of the most highly blanketed genes in HEL cells (using the Database for Annotation, Visualization and Integrated Discovery: DAVID) (Huang da et al., 2009) shows that these genes are primarily expressed in myeloid cells (Figure S3D), and in contrast, the genes most blanketed by H3Y41ph in K1106 cells are those primarily expressed in lymphoid tissues (Figure S3D).

DISCUSSION

Two decades of research have firmly established that a key function of cytoplasmic JAK2 is to mediate cytokine-induced intracellular signaling. The data presented here challenge the view that JAK2 is functionally restricted to the cytoplasm

in this signaling pathway. They indicate that JAK2 can phosphorylate chromatin at H3Y41 and that this modification takes place on a number of genes previously linked to the JAK/STAT pathway. Unfortunately our attempts to localize JAK2 at chromatin with ChIP assays have been unsuccessful. This most likely reflects the transient nature of JAK2′s interaction with chromatin, which is in contrast to the MAPK family of enzymes that have a consensus DNA binding motif and have been successfully localized at chromatin (Hu et al., 2009). Nevertheless, we have demonstrated in various cell types that H3Y41ph is a reliable marker for JAK2′s activity at chromatin (Dawson et al., 2009). The data reported here show that H3Y41ph is coincident with STAT5 at several loci and that both these events are dynamically modulated at the regulatory elements of cytokine-activated genes. Together, these data establish a previously unrecognized nuclear component of the JAK/STAT signaling pathway.

Coincidence of H3Y41ph and STAT5 (Table S1) is seen only at a subset of the total number of genes marked by H3Y41ph or STAT5. Since STAT5 is activated by multiple pathways (Lim and Cao, 2006), genes bound by STAT5 alone may represent interplay with alternative kinases, modifying distinct sites on histones. Alternatively, STAT function at certain sites may not require the cooperation of a kinase, but may depend, for example, upon the recruitment of specific coactivator/repressor complexes. Similarly, the genes marked by H3Y41ph alone may represent genes regulated by other STAT family members or by a STAT-independent pathway (Dawson et al., 2009). In fact, when we cross-referenced our data with the published STAT1 and STAT2 ChIP-seq data sets performed in K562 cells (another erythroid leukemia cell line), we found that a further 10% of the H3Y41ph sites not bound by STAT5 are bound by STAT1/2 (data not shown). A further possible explanation for H3Y41ph peaks that do not overlap STAT5 binding sites is that H3Y41ph is laid down by at least one other kinase (Dawson et al., 2009; Griffiths et al., 2011), which may function independently of STATs.

The data presented here have implications beyond the JAK/STAT pathway; global analysis of H3Y41ph suggests that this modification behaves differently to other well-characterized histone modifications, such as lysine acetylation and methylation, which are often linked to active genes, but their global levels are not regulated by any single signaling pathway. In contrast, H3Y41ph marks a specific set of genes stimulated by a specific signaling pathway and blankets a set of key lineage-specific hematopoietic genes. Moreover, the data presented here provide further evidence that signal-transducing kinases can extend their activity to chromatin. Since many of these pathways have been the focus of intense drug discovery programs, the identification of a nuclear component of kinase cascades increases the scope for future therapeutic intervention.

EXPERIMENTAL PROCEDURES

Cell Culture

HEL, TF-1, and SET-2 cells were grown at 37°C and 5% CO_2 in RPMI-1640 (Sigma-Aldrich) supplemented with 10%–20% fetal calf serum, 1% penicillin/streptomycin, and 10 ng/ml IL3 (TF-1 only).

ChIP-seq and ChIP-PCR Analysis

Chromatin immunoprecipitation was performed as previously described (Dawson et al., 2009). Immunoprecipitated DNA was either amplified for sequencing or analyzed on an ABI7900 RT-PCR machine using TaqMan PCR mastermix according to the manufacturer's instructions (see Extended Experimental Procedures for primers/probes).

Gene Expression Analysis

Messenger RNA (mRNA) was prepared using the QIAGEN RNeasy kit according to the manufacturer's instructions. Complementary DNA (cDNA) was prepared using Superscript III reverse transcriptase (Invitrogen) and analyzed on an ABI7900 RT-PCR machine using power SYBR green PCR mastermix according to the manufacturer's instructions (see Extended Experimental Procedures for primers).

JAK2 Inhibitors

TG101209, TargeGen Inc. (San Diego, CA, USA) and AT9283, Astex Pharmaceuticals, were used at 1 μM and 300 nM, respectively, for 4 hr.

Density Profiling and Genome-wide Feature Analysis

BED files were created from uniquely mapped reads extended to 200 bp and used to create genome-wide density profiles as described (Wilson et al., 2009). For genome-wide feature analysis, regions of 10 kb width were extracted and the mean density plotted normalized relative to the total number of sequence reads in the corresponding data sets. Heat maps were plotted using the same data and ranked based on the ratio of signal at the central 2 kb divided by the average signal across the 10 kb region. To examine global relationships between H3K4me3 and H3Y41ph, TSSs were binned into 10 cohorts of descending H3K4me3 levels. For quantile normalization, TSSs were ranked relative to the quantity of H3K4me3 and H3Y41ph sequence tags within ±500 bp of the TSS and displayed using the kde2d R function for kernel density plots (Venables and Ripley, 2002).

Identification of Blanketed Genes

Refseq genes were split into 10 equally sized regions and two measures were calculated for each: the proportion of bases showing significant normalized H3Y41ph sequence tag score >10 ("overlap") and sum of normalized H3Y41ph read counts per base ("density"). Both measures were highly correlated (Spearman correlation = 1.00, $p < 2 \times 10^{-16}$). Genes within 50 kb of the centromere or less than 2.5 kb were removed. Of the 19,067 unique genes analyzed, 1,009 showed H3Y41ph coverage. A histogram of the median H3Y41ph overlap of each gene identified a subset of genes with high levels of protracted gene coverage. Genes were classified as "blanketed" or "highly blanketed" for H3Y41ph if the overlap was greater than 20% (324 genes) or 50% (45 genes), respectively. Three control gene sets were created of a similar size to the blanketed gene list: active genes (based on H3K4me3 read density at the promoter), active genes with no H3Y41ph, and genes with H3Y41ph in the promoter but not across the gene body. Functional gene set analysis was conducted using the DAVID tool (Huang da et al., 2009).

Peak Analysis

Peak calling for the STAT5 and H3Y41ph data sets was performed using Findpeaks 3.1 (Fejes et al., 2008) and MACs (Zhang et al., 2008). In addition, STAT5 peaks were called using PeakSeq (Rozowsky et al., 2009). Peaks called by all programs were retained and filtered further based on Findpeaks 3.1 peaks in the control IgG ChIP-seq data set, distance of more than 100 kb of any RefSeq or University of California Santa Cruz (UCSC) annotated gene, and more than 70% of repetitive sequence. De novo motif discovery was performed using Multiple Expectation Maximization for Motif Elicitation (MEME) with the central 100 bp of STAT5 peaks (Bailey et al., 2009) with default settings.

Overlap between STAT5- and H3Y41ph-enriched regions was determined based on peaks that overlapped by at least one base pair and also determined for STAT1/STAT2 peaks in K562 cells from the ENCODE project (Zhang et al., 2007). Statistical significance of overlaps was calculated by performing 1,000 random permutations and a nonparametric Wilcoxon rank-sum test under the

NULL hypothesis that STAT and H3Y41ph enrichments are independent of one another.

ACCESSION NUMBERS

The ArrayExpress accession number for the ChIP-seq data reported in this paper is E-MTAB-1096.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Extended Experimental Procedures, three figures, and two tables and can be found with this article online at http://dx.doi.org/10.1016/j.celrep.2012.08.016.

LICENSING INFORMATION

ACKNOWLEDGMENTS

We would like to thank Yongjun Zhao for overseeing the sequencing and Sarah Teichmann for helpful discussions. Work in the authors' laboratories is supported by grants from Cancer Research UK, Leukaemia & Lymphoma Research UK, the UK Medical Research Council, the Wellcome Trust, Leukemia & Lymphoma Society of America, NIHR Cambridge Biomedical Research Centre, and the 6th Research Framework Programme of the European Union (Epitron and SMARTER). M.A.D. is a Wellcome-Beit Intermediate Clinical Fellow. S.D.F. was funded by a PhD fellowship grant from the UK Medical Research Council to the Cambridge Institute for Medical Research. T.K. is a director of Abcam Ltd and he is on the scientific advisory board of GlaxoSmithKline. A.R.G. is on the clinical advisory board of Astex Therapeutics, Cambridge, UK.

Received: April 23, 2012
Revised: July 16, 2012
Accepted: August 16, 2012
Published online: September 20, 2012

REFERENCES

Bailey, T.L., Boden, M., Buske, F.A., Frith, M., Grant, C.E., Clementi, L., Ren, J., Li, W.W., and Noble, W.S. (2009). MEME SUITE: tools for motif discovery and searching. Nucleic Acids Res. *37*(Web Server issue), W202–W208.

Barahmand-Pour, F., Meinke, A., Groner, B., and Decker, T. (1998). Jak2-Stat5 interactions analyzed in yeast. J. Biol. Chem. *273*, 12567–12575.

Barski, A., Cuddapah, S., Cui, K., Roh, T.Y., Schones, D.E., Wang, Z., Wei, G., Chepelev, I., and Zhao, K. (2007). High-resolution profiling of histone methylations in the human genome. Cell *129*, 823–837.

Chen, E., Staudt, L.M., and Green, A.R. (2012). Janus kinase deregulation in leukemia and lymphoma. Immunity *36*, 529–541.

Dawson, M.A., Bannister, A.J., Göttgens, B., Foster, S.D., Bartke, T., Green, A.R., and Kouzarides, T. (2009). JAK2 phosphorylates histone H3Y41 and excludes HP1alpha from chromatin. Nature *461*, 819–822.

Dawson, M.A., Curry, J.E., Barber, K., Beer, P.A., Graham, B., Lyons, J.F., Richardson, C.J., Scott, M.A., Smyth, T., Squires, M.S., et al. (2010). AT9283, a potent inhibitor of the Aurora kinases and Jak2, has therapeutic potential in myeloproliferative disorders. Br. J. Haematol. *150*, 46–57.

Decker, T., and Müller, M. (2012). Jak-stat signaling: from basics to disease, First Edition (New York: Springer).

Dennis, G., Jr., Sherman, B.T., Hosack, D.A., Yang, J., Gao, W., Lane, H.C., and Lempicki, R.A. (2003). DAVID: Database for Annotation, Visualization, and Integrated Discovery. Genome Biol. *4*, 3.

Dumon, S., Santos, S.C., Debierre-Grockiego, F., Gouilleux-Gruart, V., Cocault, L., Boucheron, C., Mollat, P., Gisselbrecht, S., and Gouilleux, F. (1999). IL-3 dependent regulation of Bcl-xL gene expression by STAT5 in a bone marrow derived cell line. Oncogene *18*, 4191–4199.

Fejes, A.P., Robertson, G., Bilenky, M., Varhol, R., Bainbridge, M., and Jones, S.J. (2008). FindPeaks 3.1: a tool for identifying areas of enrichment from massively parallel short-read sequencing technology. Bioinformatics *24*, 1729–1730.

Griffiths, D.S., Li, J., Dawson, M.A., Trotter, M.W., Cheng, Y.H., Smith, A.M., Mansfield, W., Liu, P., Kouzarides, T., Nichols, J., et al. (2011). LIF-independent JAK signalling to chromatin in embryonic stem cells uncovered from an adult stem cell disease. Nat. Cell Biol. *13*, 13–21.

Hu, S., Xie, Z., Onishi, A., Yu, X., Jiang, L., Lin, J., Rho, H.S., Woodard, C., Wang, H., Jeong, J.S., et al. (2009). Profiling the human protein-DNA interactome reveals ERK2 as a transcriptional repressor of interferon signaling. Cell *139*, 610–622.

Huang da, W., Sherman, B.T., and Lempicki, R.A. (2009). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat. Protoc. *4*, 44–57.

Kamakura, S., Oishi, K., Yoshimatsu, T., Nakafuku, M., Masuyama, N., and Gotoh, Y. (2004). Hes binding to STAT3 mediates crosstalk between Notch and JAK-STAT signalling. Nat. Cell Biol. *6*, 547–554.

Lefrancois-Martinez, A.M., Blondet-Trichard, A., Binart, N., Val, P., Chambon, C., Sahut-Barnola, I., Pointud, J.C., and Martinez, A. (2011). Transcriptional control of adrenal steroidogenesis: novel connection between Janus kinase (JAK) 2 protein and protein kinase A (PKA) through stabilization of cAMP response element-binding protein (CREB) transcription factor. J. Biol. Chem. *286*, 32976–32985.

Lim, C.P., and Cao, X. (2006). Structure, function, and regulation of STAT proteins. Mol. Biosyst. *2*, 536–550.

Liu, F., Zhao, X., Perna, F., Wang, L., Koppikar, P., Abdel-Wahab, O., Harr, M.W., Levine, R.L., Xu, H., Tefferi, A., et al. (2011). JAK2V617F-mediated phosphorylation of PRMT5 downregulates its methyltransferase activity and promotes myeloproliferation. Cancer Cell *19*, 283–294.

Matsumoto, A., Masuhara, M., Mitsui, K., Yokouchi, M., Ohtsubo, M., Misawa, H., Miyajima, A., and Yoshimura, A. (1997). CIS, a cytokine inducible SH2 protein, is a target of the JAK-STAT5 pathway and modulates STAT5 activation. Blood *89*, 3148–3154.

McLean, C.Y., Bristor, D., Hiller, M., Clarke, S.L., Schaar, B.T., Lowe, C.B., Wenger, A.M., and Bejerano, G. (2010). GREAT improves functional interpretation of cis-regulatory regions. Nat. Biotechnol. *28*, 495–501.

Min, I.M., Waterfall, J.J., Core, L.J., Munroe, R.J., Schimenti, J., and Lis, J.T. (2011). Regulating RNA polymerase pausing and transcription elongation in embryonic stem cells. Genes Dev. *25*, 742–754.

Mui, A.L., Wakao, H., Kinoshita, T., Kitamura, T., and Miyajima, A. (1996). Suppression of interleukin-3-induced gene expression by a C-terminal truncated Stat5: role of Stat5 in proliferation. EMBO J. *15*, 2425–2433.

Nagy, Z.S., LeBaron, M.J., Ross, J.A., Mitra, A., Rui, H., and Kirken, R.A. (2009). STAT5 regulation of BCL10 parallels constitutive NFkappaB activation in lymphoid tumor cells. Mol. Cancer *8*, 67.

Nilsson, J., Bjursell, G., and Kannius-Janson, M. (2006). Nuclear Jak2 and transcription factor NF1-C2: a novel mechanism of prolactin signaling in mammary epithelial cells. Mol. Cell. Biol. *26*, 5663–5674.

Rozowsky, J., Euskirchen, G., Auerbach, R.K., Zhang, Z.D., Gibson, T., Bjornson, R., Carriero, N., Snyder, M., and Gerstein, M.B. (2009). PeakSeq enables systematic scoring of ChIP-seq experiments relative to controls. Nat. Biotechnol. *27*, 66–75.

Rui, L., Emre, N.C., Kruhlak, M.J., Chung, H.J., Steidl, C., Slack, G., Wright, G.W., Lenz, G., Ngo, V.N., Shaffer, A.L., et al. (2010). Cooperative epigenetic modulation by cancer amplicon genes. Cancer Cell *18*, 590–605.

Teglund, S., McKay, C., Schuetz, E., van Deursen, J.M., Stravopodis, D., Wang, D., Brown, M., Bodner, S., Grosveld, G., and Ihle, J.N. (1998). Stat5a

and Stat5b proteins have essential and nonessential, or redundant, roles in cytokine responses. Cell *93*, 841–850.

Venables, W.N., and Ripley, B.D. (2002). Modern applied statistics with S, Fourth Edition (New York: Springer).

Wilson, N.K., Miranda-Saavedra, D., Kinston, S., Bonadies, N., Foster, S.D., Calero-Nieto, F., Dawson, M.A., Donaldson, I.J., Dumon, S., Frampton, J., et al. (2009). The transcriptional program controlled by the stem cell leukemia gene Scl/Tal1 during early embryonic hematopoietic development. Blood *113*, 5456–5465.

Wood, A.D., Chen, E., Donaldson, I.J., Hattangadi, S., Burke, K.A., Dawson, M.A., Miranda-Saavedra, D., Lodish, H.F., Green, A.R., and Göttgens, B. (2009). ID1 promotes expansion and survival of primary erythroid cells and is a target of JAK2V617F-STAT5 signaling. Blood *114*, 1820–1830.

Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nusbaum, C., Myers, R.M., Brown, M., Li, W., and Liu, X.S. (2008). Model-based analysis of ChIP-Seq (MACS). Genome Biol. *9*, R137.

Zhang, Y., Laz, E.V., and Waxman, D.J. (2012). Dynamic, sex-differential STAT5 and BCL6 binding to sex-biased, growth hormone-regulated genes in adult mouse liver. Mol. Cell. Biol. *32*, 880–896.

Zhang, Z.D., Paccanaro, A., Fu, Y., Weissman, S., Weng, Z., Chang, J., Snyder, M., and Gerstein, M.B. (2007). Statistical analysis of the genomic distribution and correlation of regulatory elements in the ENCODE regions. Genome Res. *17*, 787–797.