Psip1/Ledgf p52 Binds Methylated Histone H3K36 and Splicing Factors and Contributes to the Regulation of Alternative Splicing

Citation for published version:

Digital Object Identifier (DOI):
10.1371/journal.pgen.1002717

Link:
Link to publication record in Edinburgh Research Explorer

Document Version:
Publisher's PDF, also known as Version of record

Published In:
PLoS Genetics

Publisher Rights Statement:
Copyright: © 2012 Pradeepa et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

General rights
Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and/or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.
Psip1/Ledgf p52 Binds Methyalted Histone H3K36 and Splicing Factors and Contributes to the Regulation of Alternative Splicing

Madapura M. Pradeepa¹, Heidi G. Sutherland¹, Jernej Ule², Graeme R. Grimes¹, Wendy A. Bickmore¹*

¹ MRC Human Genetics Unit, MRC Institute of Genetics and Molecular Medicine, University of Edinburgh, Edinburgh, United Kingdom, ² MRC Laboratory of Molecular Biology, Cambridge, United Kingdom

Abstract

Increasing evidence suggests that chromatin modifications have important roles in modulating constitutive or alternative splicing. Here we demonstrate that the PWWP domain of the chromatin-associated protein Psip1/Ledgf can specifically recognize tri-methylated H3K36 and that, like this histone modification, the Psip1 short (p52) isoform is enriched at active genes. We show that the p52, but not the long (p75), isoform of Psip1 co-localizes and interacts with Srsf1 and other proteins involved in mRNA processing. The level of H3K36me3 associated Srsf1 is reduced in Psip1 mutant cells and alternative splicing of specific genes is affected. Moreover, we show altered Srsf1 distribution around the alternatively spliced exons of these genes in Psip1 null cells. We propose that Psip1/p52, through its binding to both chromatin and splicing factors, might act to modulate splicing.


Editor: Wolf Reik, The Babraham Institute, United Kingdom

Received December 29, 2011; Accepted April 2, 2012; Published May 17, 2012

Copyright: © 2012 Pradeepa et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: This work was supported by a project grant from the Wellcome Trust (grant no. WT085767) and by a unit programme grant from the Medical Research Council, UK. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: W.Bickmore@hgu.mrc.ac.uk

Introduction

Pre-mRNA splicing occurs co-transcriptionally [1], whilst the nascent transcript is still associated with the chromatin template. However, until recently there has been little consideration of how chromatin structure might influence the control of splicing. Initial studies indicated a link between promoters and alternative splicing [2–4] and this has been extended to histone modifications enriched at promoters. For example, Gcn5 mediated histone acetylation at promoters in yeast has been shown to facilitate recruitment of splicing factors [5] and mammalian GCN5-containing complexes interact with pre-mRNA splicing factors [6]. The chromatin remodeler CHD1, which recognises a histone mark (H3K4me3) enriched at active promoters, also interacts with spliceosome components and affects the rate of mRNA splicing [7].

A link between the rate of transcriptional elongation and splicing [8–10] has led to a consideration of how chromatin structure within the body of genes might also influence splicing. Increased levels of histone acetylation in gene bodies lead to exon skipping, likely through enhanced RNA polymerase II processivity [11]. Conversely, HP1γ, which binds to H3K9me3, favors inclusion of alternative exons, possibly by decreasing RNA polymerase II elongation rate [12].

Trimethylation of H3 at lysine 36 (H3K36me3) is enriched at exons, particularly those of highly expressed genes [13–17] and its level at alternatively spliced exons is reported to correlate with their inclusion into the spliced transcript [13]. An explanation for this may come from observations that pre-mRNA splicing itself affects the deposition of this histone modification [18,19]. A direct link between H3K36me3 and an effect on mRNA splicing comes from the observation that MRG15, a protein whose chromo-domain can recognise H3K36me3, recruits polypyrimidine tract binding protein (PTB) to alternatively spliced exons [20]. It was not clear whether this is a unique interaction or whether there are other systems that connect H3K36me3 to alternative splicing.

PC4 and SF2 interacting protein 1 (Psip1) has been implicated in transcriptional regulation and mRNA splicing in vitro [21], but its function in vivo is poorly understood. It has been implicated in developmental gene regulation [22] and in guiding the integration of human immunodeficiency virus (HIV) into the host genome [23–26]. Psip1 encodes two protein isoforms - p52 and p75 - generated by alternative splicing within intron 9, and whose relative levels vary between tissues [21,27]. The p75 isoform, also known as lens epithelium derived growth factor (Ledgf), has a C-terminal integrase binding domain (IBD) (Figure 1A) that binds the integrases of HIV-1 and other lentiviruses, preventing their degradation by the proteasome [28] and tethering them to host chromosomes [28–33]. In Psip1 mutant cells, HIV/lentivirus infection is impaired and sites of viral integration into the host genome are altered [24–26]. Though the normal cellular function of Psip1/p75 has not been established, the IBD binds to RAM2/JPO2 - a myc-associated transcriptional regulator [34,35] and p75 is tethered, via Menin and in an IBD-dependant manner, to MLL H3K4 histone methyltransferase [36].

The p52 isoform of Psip1 lacks the IBD (Figure 1A) and does not interact with Menin. Instead, Psip1/p52 has been purified with PC4 transcriptional co-activator [37], and had been shown to immunoprecipitate (IP) with, and to modulate the activity of, the
The regulated processing of mRNAs by splicing of exons and introns has the potential to increase the information content of the genome. Various splicing factors have been identified whose binding to cis-acting sequences can influence whether an alternative exon is included or excluded (skipped) in the mature mRNA. However, increasing evidence suggests that the chromatin template also has an important role in modulating splicing. Here we identify a chromatin-associated protein Psip1/Ledgf that can bind to a histone modification enriched at active genes and that can also interact with other proteins involved in mRNA splicing. Loss of Psip1 reduces the chromatin association of specific splicing proteins and alters the pattern of alternative splicing. We propose that Psip1, through its binding to both chromatin and splicing factors, might act to modulate splicing.

Results

Psip1 PWWP domain can bind to H3K36me3

GFP-tagged full-length, and β-gal tagged gene-trap, versions of Psip1/p75 have been reported on mitotic chromosomes.
H3K36me3 (Spearman’s rank correlation coefficient 0.05) or between H3K36me3 and H3K4me3 (0.013).

To determine if the Psip1 PWWP domain directly interacts with modified histone tails, we used histone tail peptide arrays containing in total 59 different modifications of H3, H4, H2A, and H2B tails in 384 different combinations. In two independent experiments, we observed that GST-tagged Psip1 PWWP domain bound H3K36me3 with high specificity - signal from H3K36me2, H3K36me3 and corresponding unmodified peptide spots were not above background (Figure 1B and 1C, Table S1). Direct binding of p52 with H3K36me3 was confirmed by peptide pulldown (Figure 1D). Immunoblotting with antibodies recognizing different H3 methylation states confirmed a specific enrichment of H3K36me3 in Psip1 IPs from nuclear extracts (Figure 1E).

Psip1/p52 is enriched at expressed genes
We assessed the genomic distribution of Psip1 in mouse embryonic fibroblasts (MEFs) by chromatin immunoprecipitation (ChIP) using an antibody A300-847 (see below) and hybridization to a custom tiling array. The hybridization pattern was compared to that from H3K36me3 and H3K4me3 ChIPs.

The large-scale distributions of H3K36me3 and Psip1/p52 were similar to each other and both appeared to be enriched at active genes (Figure 2A). Across the entire array, levels of both Psip1/p52 and H3K36me3 were significantly higher at active genes than inactive genes or intergenic regions, and furthermore were especially enriched at the exons compared to the introns of expressed genes (p<0.05) (Figure 2B). Visual inspection of specific genes revealed a similar distribution of Psip1/p52 and H3K36me3 at some downstream exons (Figure 2C and 2D), distinct from the peak of H3K4me3 at promoters. However, there is also evidence for some enrichment of Psip1/p52 near the transcription start sites (TSSs) suggesting multiple modes of Psip1 association to chromatin. Correlation between the distribution of Psip1/p52 and H3K36me3 (Spearman’s rank correlation coefficient ρ = 0.38, p<0.05) was stronger than that between Psip1/p52 and H3K4me3 (ρ = -0.05) or between H3K36me3 and H3K4me3 (p = 0.013).

Splicing proteins interact with Psip1/p52
To determine whether there are other interacting partners for Psip1 isoforms, apart from H3K36me3, we performed immunoprecipitation with two different antibodies.

Antibody A300-847 was raised against an epitope present in Psip1 (see supplemental information) and detect endogenous wild-type (wt) Psip1/p52 is evidenced by the absence of immunoprecipitation of Srsf1 and other SR proteins in extracts prepared from MEFs homoyzygous for a gene-trap integration into Psip1 (Psip1<sup>+/−</sup>) (Figure S1A and S1B) in which the A300-847 epitope is 3’ to the site of gene trap integration, and so is absent from the resulting fusion protein (Figure 1A) [22].

Antibody A300-848 specifically recognizes the extreme C-terminus - amino acids (a.a.) 400 to 530 - of Psip1/p75 (Figure 1A) and so detects endogenous p75, but not p52, in immunoblots and IPs (Figure 3A and 3B). Only a few transcription related proteins, in addition to p75 itself, were IPed from nuclear extracts by A300-848 (data not shown).

These data indicate a cellular link between Psip1/p52 and the splicing machinery. Immunoblotting of the IP from RNase treated nuclear extracts indicated that Psip1/p52 interacts mainly with the hypophosphorylated form of SRSF1 (Figure 3C). Phosphorylation levels of SR proteins are known to modulate alternative splicing and alter SR protein distribution in relative to splicing-factor enriched nuclear speckles [46–48]. GST-p52 pull down of T7-SRSF1 (over expressed HEK-293T cells), confirmed direct interaction of Psip1/p52 with SRSF1 and that the Psip1 PWWP domain is not sufficient for this (Figure 3D). Furthermore, GST-p52 pulldown of SRSF1 mutants which mimic hypo (RG) and hyper (RD) phosphorylation (serine residues within RS/SR dipeptide repeats of RS domain substituted with Glycine; RG or Aprotic acid: RD) [49] shows higher affinity of Psip1/p52 for hypophosphorylated SRSF1 compared to the hyperphosphorylated form (Figure 3E).

GST pulldown also confirms interaction with SRSF3, but shows that Psip1/p52 does not simply interact non-specifically with all SR proteins, since there is no direct interaction with SRSF2 (SC35) (Figure 3E). Identification of Srsf2 by mass spectrometry in the A300-847 immunoprecipitate presumably is the result of indirect association with other splicing proteins (Table 1).

Association of SRSF1 to H3K36me3 marked chromatin requires p52
Unphosphorylated SRSF1 has been reported to associate with chromatin, especially the H3 tail and to be sensitive to H3 tail post-translational modifications [50]. To investigate whether absence of Psip1 causes any loss of Srsf1 chromatin association in vivo, chromatin purified by ChIP for H3K36me3 was analyzed by immunoblotting. Levels of Srsf1 associated with H3K36me3 modified chromatin were greatly reduced in Psip1<sup>−/−</sup> MEFs cells that do not have detectable Psip1/Ledgf [25], compared to wild type (Figure 3F). As controls, the levels of H3K36me3 associated Ptb and Srsf2 were not changed in the Psip1<sup>−/−</sup> IPed chromatin compared to wild type. These results confirm that Psip1/p52 specifically recruits Srsf1 to H3K36me3 chromatin in vivo, but not Ptb, which has been shown to be recruited to H3K36me3 chromatin through MRG-15 [20].

To investigate whether SRSF1 alone can bind to H3K36me3 in vitro, or whether this occurs via interaction with Psip1, we pulled-down HeLa core histones with T7-SRSF1, with or without addition of Psip1/p52. Immunoblotting with antibodies recognizing different methylated states of H3 revealed a specific enrichment of H3K36me3 in the presence of Psip1/p52 compared to SRSF1 alone (Figure 3G). These results suggest that Psip1/p52 can aid the recruitment of specific splicing factors, including SRSF1, to H3K36me3 modified chromatin.
Figure 2. Genomic distribution of Psip1/p52 and H3K36me3. A) Mean log2 ChIP:input for Psip1/p52 and H3K36me3 in MEFs for an approximately 1.2Mb genomic window from mouse chromosome 5. n = 4 (3 biological and 1 technical replicate). B) Box plots showing the distribution of log2 ChIP:input for Psip1/p52 and H3K36me3 across exons and introns of expressed or non-expressed genes. Data are deposited in NCBI GEO (Accession no. GSM697402-GSM697411). C, D) Mean log2 ChIP:input for Psip1/p52 and H3K36me3 in MEFs at (C) c-Myc and (D) Xist loci. H3K4me3 is also shown for XIST. Filled boxes indicate the positions of exons. n = 4 (3 biological and 1 technical replicate) for H3K36me3 and Psip1. NCBI GEO accession number for array platform is GPL13276. n = 2 biological replicates for H3K4me3.
doi:10.1371/journal.pgen.1002717.g002
Psip1 p52 co-localizes with splicing factors

Given the preponderance of splicing/RNA-binding proteins co-immunoprecipitating with Psip1/p52 but not p75 (Figure 3 and Table 1), we investigated the nuclear localizations of Psip1 isoforms. Antibody A300-848 revealed that, as for Psip1gt/gt [22], endogenous p75 is associated with chromosomes in mitotic cells (Figure 4A) and is generally distributed in the nucleoplasm at interphase.

Immunostaining with A300-847 also showed association with mitotic chromosomes, but at interphase revealed numerous nuclear foci reminiscent of splicing-factor enriched nuclear speckles [51] (Figure 4A). Co-immunostaining for Psip1/p52/
Table 1. Psip1/p52 interacting partners.

<table>
<thead>
<tr>
<th>ID</th>
<th>Protein name</th>
<th>Alias</th>
<th>Protein domains</th>
<th>Functions</th>
<th>Peptides</th>
<th>Mass (kDa)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SR proteins</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gi 5902076</td>
<td>Srsf1 *</td>
<td>SF2/ASF</td>
<td>RRM x2, RS</td>
<td>Constitutive and AS</td>
<td>42</td>
<td>27</td>
</tr>
<tr>
<td>gi 3153821</td>
<td>Srm 1 *</td>
<td>POP101/Srm160</td>
<td>RS, PWI</td>
<td>Constitutive and AS</td>
<td>30</td>
<td>101</td>
</tr>
<tr>
<td>gi 2212585</td>
<td>Srsf7</td>
<td>9G8</td>
<td>RRM, RS</td>
<td>Constitutive and AS</td>
<td>25</td>
<td>27</td>
</tr>
<tr>
<td>gi 6755478</td>
<td>Srsf2</td>
<td>SC-35 *</td>
<td>RRM, RS</td>
<td>Constitutive and AS</td>
<td>25</td>
<td>35</td>
</tr>
<tr>
<td>gi 3929369</td>
<td>Srsf5</td>
<td>Srp40</td>
<td>RRM x2, RS</td>
<td>Constitutive and AS</td>
<td>17</td>
<td>31</td>
</tr>
<tr>
<td>gi 114152161</td>
<td>Srm2</td>
<td>SRm300</td>
<td>RS</td>
<td>Constitutive and AS</td>
<td>17</td>
<td>295</td>
</tr>
<tr>
<td>gi 9622185</td>
<td>Acinus</td>
<td>Acin1</td>
<td>RRM, RS, SAP</td>
<td>mRNA splicing regulation, apoptosis, EJC</td>
<td>16</td>
<td>150</td>
</tr>
<tr>
<td>gi 12844972</td>
<td>Srsf 3</td>
<td>Srp20</td>
<td>RRM, RS</td>
<td>Constitutive and AS</td>
<td>16</td>
<td>19</td>
</tr>
<tr>
<td>gi 27461725</td>
<td>Pinin</td>
<td>Pnn</td>
<td>RS</td>
<td>AS regulator</td>
<td>11</td>
<td>82</td>
</tr>
<tr>
<td>gi 4759098</td>
<td>Srsf10*</td>
<td>Tra2b</td>
<td>RRM, RS x2</td>
<td>Splicing activator</td>
<td>9</td>
<td>33</td>
</tr>
<tr>
<td>gi 13278367</td>
<td>Rbm39</td>
<td>Rpnc2</td>
<td>RRM x3, RS</td>
<td>AS regulation</td>
<td>9</td>
<td>47</td>
</tr>
<tr>
<td>gi 13385016</td>
<td>Srsf9</td>
<td>Srp30</td>
<td>RRM x2, RS</td>
<td>Constitutive and AS</td>
<td>8</td>
<td>25</td>
</tr>
<tr>
<td>gi 124487333</td>
<td>Sppig</td>
<td>SRCyp</td>
<td>RS, PPlase</td>
<td>Regulates localisation of SR proteins</td>
<td>6</td>
<td>88</td>
</tr>
<tr>
<td>gi 1399464</td>
<td>Prp4f4b</td>
<td>Prpk</td>
<td>RS, kinase</td>
<td>mRNA splicing</td>
<td>6</td>
<td>116</td>
</tr>
<tr>
<td>gi 4001720</td>
<td>SRp38</td>
<td>Nssr 1</td>
<td>RRM, RS</td>
<td>General splicing repressor</td>
<td>3</td>
<td>31</td>
</tr>
<tr>
<td>gi 122890231</td>
<td>Srsf4</td>
<td>Srp75</td>
<td>RRM x2, RS</td>
<td>Constitutive and AS</td>
<td>2</td>
<td>56</td>
</tr>
<tr>
<td>DExD/H box helicases</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gi 28076989</td>
<td>DDX54</td>
<td>DP97</td>
<td>DEAD, helicase</td>
<td>Transcription regulation</td>
<td>43</td>
<td>98</td>
</tr>
<tr>
<td>gi 14870489</td>
<td>DHX 9</td>
<td>NDH2</td>
<td>DrRM, DEAD, helicase</td>
<td>Transcription regulation, mRNA splicing</td>
<td>30</td>
<td>132</td>
</tr>
<tr>
<td>gi 2500527</td>
<td>DDX5 *</td>
<td>RNA helicase p68</td>
<td>DEAD, helicase</td>
<td>mRNA splicing, microRNA processing</td>
<td>27</td>
<td>69</td>
</tr>
<tr>
<td>gi 6753620</td>
<td>DDX3X</td>
<td>DAP88</td>
<td>DEAD, helicase</td>
<td>mRNA splicing, RNA export, microRNA processing, translation</td>
<td>27</td>
<td>73</td>
</tr>
<tr>
<td>gi 16716475</td>
<td>DDX 50</td>
<td>RH II/Gu</td>
<td>DEAD, helicase</td>
<td>Ribosome biogenesis, transcription co-factor</td>
<td>25</td>
<td>82</td>
</tr>
<tr>
<td>gi 27371129</td>
<td>DDX10</td>
<td>HRH-J8</td>
<td>DEAD, helicase</td>
<td>Ribosome biogenesis</td>
<td>24</td>
<td>78</td>
</tr>
<tr>
<td>gi 110835726</td>
<td>DHX15</td>
<td>Prp43</td>
<td>DEAD, helicase</td>
<td>mRNA splicing, microRNA processing</td>
<td>27</td>
<td>73</td>
</tr>
<tr>
<td>gi 73620772</td>
<td>DDX41*</td>
<td>Abstrakt</td>
<td>DEAD, helicase, C2HC Zn finger</td>
<td>mRNA splicing</td>
<td>19</td>
<td>70</td>
</tr>
<tr>
<td>gi 25141235</td>
<td>DDX3Y</td>
<td>DBY</td>
<td>DEAD Box</td>
<td>mRNA splicing, microRNA processing, RNA export, translation</td>
<td>14</td>
<td>73</td>
</tr>
<tr>
<td>SnRNPs</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gi 40018610</td>
<td>U5-200KD</td>
<td>BRR2</td>
<td>DEAD, helicase, sec63</td>
<td>mRNA splicing, U5 snRNP complex</td>
<td>13</td>
<td>246</td>
</tr>
<tr>
<td>gi 124430514</td>
<td>U5 100K *</td>
<td>DDX 23/Prp28</td>
<td>RS, DEAD, helicase</td>
<td>mRNA splicing, U5 snRNP complex</td>
<td>12</td>
<td>42</td>
</tr>
<tr>
<td>gi 19527174</td>
<td>SF3B3*</td>
<td>SFib(130)</td>
<td>CPSF, A</td>
<td>mRNA splicing, U2 snRNP complex</td>
<td>10</td>
<td>136</td>
</tr>
<tr>
<td>gi 21539655</td>
<td>U5-102 kDa</td>
<td>Prp6</td>
<td>HAT repeats</td>
<td>mRNA splicing, U5 snRNP complex</td>
<td>6</td>
<td>107</td>
</tr>
<tr>
<td>hnRNPs</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gi 21313308</td>
<td>hnRNpm</td>
<td>KIAA4193</td>
<td>RRM</td>
<td>Pre-mRNA processing</td>
<td>17</td>
<td>78</td>
</tr>
<tr>
<td>gi 3914804</td>
<td>hnRNPg</td>
<td>Rbmx</td>
<td>RRM</td>
<td>Pre-mRNA processing</td>
<td>10</td>
<td>42</td>
</tr>
<tr>
<td>gi 123778087</td>
<td>hnRNPU like</td>
<td>MLF1ANP</td>
<td>RRM, SAR DNA binding</td>
<td>Pre-mRNA processing</td>
<td>4</td>
<td>85</td>
</tr>
<tr>
<td>gi 13384620</td>
<td>hnRNpk</td>
<td>TUNP</td>
<td>KH domains</td>
<td>Pre-mRNA process</td>
<td>5</td>
<td>51</td>
</tr>
<tr>
<td>Other proteins likely to be involved in RNA processing</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gi 54607128</td>
<td>RRPS</td>
<td>PDCD11</td>
<td>S1, HAT</td>
<td>mRNA processing</td>
<td>98</td>
<td>209</td>
</tr>
<tr>
<td>gi 29747798</td>
<td>Rbm 12b</td>
<td>RBM 12ba</td>
<td>RRM x5</td>
<td>Unknown</td>
<td>12</td>
<td>97</td>
</tr>
<tr>
<td>gi 188497644</td>
<td>Noc3</td>
<td>Fad24</td>
<td>mRNA splicing, transcription</td>
<td>11</td>
<td>93</td>
<td></td>
</tr>
<tr>
<td>gi 83921589</td>
<td>Fxr2</td>
<td>KH</td>
<td></td>
<td></td>
<td>10</td>
<td>74</td>
</tr>
<tr>
<td>gi 4835742</td>
<td>FXR1</td>
<td>Fxr1p</td>
<td>KH</td>
<td>translation</td>
<td>8</td>
<td>64</td>
</tr>
<tr>
<td>gi 29747807</td>
<td>Safb2 *</td>
<td>SAF-B2</td>
<td>SAP, RRM</td>
<td>Transcription and splicing</td>
<td>4</td>
<td>114</td>
</tr>
<tr>
<td>gi 34098365</td>
<td>NRF</td>
<td>BTBA4</td>
<td>ds RNA binding, G patch</td>
<td>Transcription regulation</td>
<td>11</td>
<td>78</td>
</tr>
</tbody>
</table>
### Table 1. Cont.

<table>
<thead>
<tr>
<th>ID</th>
<th>Protein name</th>
<th>Alias</th>
<th>Protein domains</th>
<th>Functions</th>
<th>Peptides</th>
<th>Mass (kDa)</th>
</tr>
</thead>
<tbody>
<tr>
<td>gi 23503105</td>
<td>NOC2L</td>
<td></td>
<td>nucleolar</td>
<td>transcription corepressor</td>
<td>30</td>
<td>86</td>
</tr>
<tr>
<td>gi 2645205</td>
<td>Mybbp1A</td>
<td></td>
<td>armadillo</td>
<td>Transcription regulation</td>
<td>28</td>
<td>150</td>
</tr>
<tr>
<td>gi 9790013</td>
<td>AATF</td>
<td></td>
<td></td>
<td>Nucleolus, apoptosis, transcription</td>
<td>25</td>
<td>59</td>
</tr>
<tr>
<td>gi 71979675</td>
<td>LAS1L</td>
<td></td>
<td></td>
<td>Ribosome biogenesis</td>
<td>23</td>
<td>88</td>
</tr>
<tr>
<td>gi 83921589</td>
<td>FMR2</td>
<td></td>
<td></td>
<td></td>
<td>10</td>
<td>74</td>
</tr>
</tbody>
</table>

Proteins identified by mass spectrometry of p52 IPs (200 mM KCl).
*indicates known proteins of the ‘spliceosomal complex C’. Data on protein domains and putative protein functions were taken from http://npd.hgu.mrc.ac.uk/. doi:10.1371/journal.pgen.1002717.t001

---

**Figure 4. Sub-cellular localization of Psip1/p52 and p75.** A) Immunofluorescence and wide-field epifluorescence microscopy on human cells with; (upper row) p75-specific antibody A300-848, (lower row) A300-847 which can recognize both p52 and p75. DNA was counterstained with DAPI. B) Co-immunofluorescence of Psip1/p52 (green/A300-847) and SRSF2 (red) analyzed by confocal microscopy in untreated (upper row), or actinomycin D (ActD) treated cells. C) Co-immunofluorescence of Psip1/p75 (green/A300-848) and SRSF2 (red) in ActD treated cells and analyzed by confocal microscopy. doi:10.1371/journal.pgen.1002717.g004
p75 and SRP2, a marker for the splicing-factor enriched nuclear speckles, confirmed this (Figure 4B). Splicing-factor enriched nuclear speckles become larger and less numerous upon the inhibition of transcription with actinomycin D [40]. Concomitantly, Psip1/p52 also became redistributed to these foci. In contrast, there was no correspondence between the sub-nuclear distribution of Psip1/p75 and splicing-factor enriched nuclear speckles (Figure 4C).

**Loss of Psip1/p52 affects alternative splicing**

To identify whether there are specific exons whose splicing in vivo might be affected by Psip1/p52, we analyzed patterns of alternative splicing in RNA prepared from primary MEFs from three different Psip1<sup>p52<sup>−/−</sup> and corresponding wild type littermate embryos. Psip1<sup>p52<sup>−/−</sup> mutant mice were generated from ES cells with a gene trap integrated between exons 8 and 9 of Psip1. This results in the production of a protein in which only the N-terminal 208 a.a. of Psip1 are present (arrowed in Figure 1A) and are fused to the β-galactosidase reporter [40]. We used a custom Affymetrix microarray containing 40,443 exon junction probe sets derived from 7,175 genes with one or more predicted alternative transcripts and analyzed the data with ASPIRE 3 software [52]. Splicing changes were detected in 95 alternative exons with a score that, in our past experience, can be validated by RT-PCR with high (>90%) success; ΔI rank ≥ 1, or ≤−1 [53,54]. Out of these, 38 exons, from 55 genes, appeared to have decreased inclusion in the mutant MEFs and 37 exons, from 35 genes, had increased inclusion (Table S2).

The gene-trap in Psip1<sup>p52<sup>−/−</sup> is between exons 8 and 9 (Figure 1A) [22] so the resulting mRNA lacks exons 9-15. This was evident from the microarray results, which detected Psip1 exons 11 and 12 as those with the most decreased inclusion in the whole analysis (Table S2). At the other extreme, the most increased inclusion of alternative exons in Psip1<sup>p52<sup>−/−</sup> was at Psip1. In mutant cells, increased alternative exon inclusion for Ptp4c, Ptp1shp, Rapsig2, Rapsig3 and Ogfi11, all of which have a ΔI >1, and altered 3′ splice site utilization at alternative exon 4 of Sorb2 (ΔI of <−1), was confirmed by semi-quantitative RT-PCR of RNA from primary MEFs derived from three wild type and three Psip1<sup>p52<sup>−/−</sup> litter mates (Figure 5A). Primer pairs spanned across regions subject to alternative splicing to generate PCR products of different sizes dependent on exon skipping or inclusion (Table S3). A 2−3 fold increase in the ratio of included:skipped exon bands was seen in mutant cells compared to wild-type. The absence of alternative splicing at the alternative exons of Cons1kd, Alg9 and Tpp2 exon 24, which were not detectably altered in the microarray, was also confirmed by RT-PCR (Figure 5B, 5D).

To examine the splicing of specific alternative exons, RT-PCR was also carried out across specific constitutive exon - constitutive exon junctions and across constitutive exon - alternative exon junctions of Vcan, Tpp2 and Duap2 where microarray analysis had indicated increased exon skipping in Psip1<sup>p52<sup>−/−</sup> cells (ΔI = −1) (Table S2). This confirmed the decreased inclusion of alternatively spliced exons in Psip1<sup>p52<sup>−/−</sup> cells (Figure 5C). To rule out the possibility of amplification bias, RT-PCR using primers spanning constitutive exons at either the 5′ or 3′ end of Tpp2, Vcan and Duap2 were tested (Figure 5D).

Although the gene-trapped Psip1 protein produced in Psip1<sup>p52<sup>−/−</sup> cells is truncated and co-localizes with concentrations of chromatin instead of splicing factors [22,40], we wished to confirm a role for Psip1 in the regulation of alternative splicing using an independently derived mutant allele. Therefore, splicing patterns of specific genes were also examined in Psip1<sup>p52<sup>−/−</sup> MEFs in which deletion of Psip1 exon 3 leads to the absence of detectable Psip1/
probably by interacting both with chromatin and proteins involved in pre-mRNA splicing. Despite containing almost all the a.a. residues of p52, the longer (p75) Psip1 isoform neither co-IPs, nor co-localizes, with splicing related proteins (Figure 4). This, together with the inability of the A300-847 antibody to IP p75, even though its epitope is present in the protein sequence and recognized in denatured p75 by immunoblot (Figure 3), suggests that protein folding of Psip1/p75 occludes both the A300-847 epitope and the region capable of interaction with splicing factors. Differential localization and interaction with the transcriptional regulation machinery or with splicing proteins has previously been reported for different isoforms of another protein – WT1 [56–58].

We add Psip1 to the recently identified group of PWPP-containing proteins - Brpf1, Dnm3a, MSH-6, Nsd1, Nsd2 and N-PAC - that have been shown to be able to bind H3K36me3 [43–45] (Figure 1 and Figure 2). This establishes the PWPP members of the ‘royal’ family of protein domains as reader of this histone modification, that has been associated with the exons of active genes [13–16,59] and whose deposition onto chromatin has recently been linked to the process of splicing itself [18,19]. 

MRG15 uses a chromo-domain for methylated histone binding [20]. The chromo-domain, like PWPP, is also a royal family protein domain [42], but the chromo-domain of MRG15 is structurally more similar to the PWPP domain of DNMT3b than to that of more typical chromo-domain proteins that recognize H3K9me3 or H3K27me3 [60].

Psip1/p75 was demonstrated to be important for guiding HIV/ lentiviral integration to the body of genes [24–26]. Our demonstration that the N-terminal PWPP domain, shared by both p52 and p75 Psip1 isoforms, recognizes and binds to H3K36me3 provides a mechanistic explanation for this pattern of HIV integration. There is a growing awareness of the interactions between splicing factors and RNA polymerase II elongation [61] and emerging evidence now highlights the role of histone modifications in this process. At gene promoters, the chromo-domains of CHD1 recognize H3K4me3 [62,63] and CHD1 interacts with the SF3a subcomplex of the U2 snRNP to then facilitate mRNA splicing post-initiation [7]. Similarly, in yeast, the histone acetyltransferase GCN5, found at the promoter regions of active genes, also interacts with components of the U2 snRNP [5]. MRG15 and Psip1/p52 now provide two examples of H3K36me3 binding proteins that can influence the recruitment of splicing components to chromatin.

MRG15 interacts with the RNA-binding protein PTB to regulate alternative splicing [20]. In contrast, we found interactions between Psip1/p52 and; several SR-containing proteins – including Srsf1 (Figure 3 and Table 1), components of the U5 snRNP and other proteins involved in RNA processing. Furthermore, we show that the absence of functional p52 affects alternative splicing of defined endogenous genes in vivo (Figure 5) and alters the pattern of Srsf1 binding across alternatively spliced gene loci (Figure 6).

Differential expression of SR proteins is important for tissue-specific alternative splicing and is abundant in brain and testis [64,65] where, compared to other tissues, mRNA for the p52 isoform of Psip1 is also at high levels compared to that of p75 [21].

Amongst other Psip1 co-immunoprecipitating proteins are many DExD/H box family putative RNA helicases. One of these is DDX10 which, like Psip1/Ledgf, is found as a fusion partner with Nup98 in myeloid leukaemias and myelodysplastic syndromes [66–69], perhaps indicative of their function in a common pathway that is mis-regulated in these malignancies. The presence of the H3K36 methyltransferase NSD1 as another Nup98 fusion partner [70] [71] suggests that the splicing-H3K36me3 connection might be implicated in the aetiology of these myeloid disorders.

Materials and Methods

Histone tail peptides arrays and peptide pulldown

A modified histone peptide array (Active motif, #13005) was blocked in TBST buffer (10 mM Tris/ HCl pH 8.3, 0.05% Tween-20, 150 mM NaCl) containing 5% non-fat dried milk at 4°C overnight. The membrane was washed with TBST for 5 min, and incubated with 10 nM purified GST-tagged Psip1 PWPP domain, or GST protein alone, at room temperature (rt) for 1 h in interaction buffer (100 mM KC1, 20 mM HEPEs pH 7.5, 1 mM EDTA, 0.1 mM DTT, 10% glycerol). After washing in TBST, the membrane was incubated with goat α-GST (GE Healthcare #27-4577-01, 1:5000 dilution in TBST) for 1 h at rt. The membrane was then washed 3x with TBST for 10 min each at rt and incubated with horseradish peroxidase conjugated α Goat antibody (Invitrogen #19-1620 1:12000 in TBST) for 1 h at rt. The membrane was submerged in ECL developing solution (Pierce, #32209), imaged (Image-quant, GE Healthcare) and the data quantified using array analyzer software (Active motif).

Biotinylated histone H3 (Ana spec 64440-025) and H3K36me3 (Ana spec 64441-025) peptides coupled to Streptavidin magnetic beads (Invitrogen 656-01), and were used to pull-down GST-p32 as described (http://www.epigenome-noe.net/WWW/ResearchTools/protocol.php?protid = 46).

Cloning, expression, and purification of proteins

Mouse GST-p52 and GST-PWPP (a.a. 1–97), were cloned into pDEST-PGEX6P. Proteins were expressed in BL21 Codonplus E.coli and purified on glutathione sepharose using standard protocols.

Human SRSF1 and Human Psip1/p52 open reading frames were cloned into pCG-T7 and eGFP vector with CMV promoters. pRES2-eGFP-p52-HA and pRES2-eGFP-p75-HA were kindly gifted by Prof. Alan Engelman (Dana-Farber Cancer Institute).
Immunoblotting

Immunoblotting was performed with the following antibody dilutions: A300 847 (1:2000), A300-848 (1:3000), zH3K36me3 (Abcam AB9050, 1:500), zH3K3me2 (Abcam ab7312, 1:500), zH3K4me3 (Millipore 07-473, 1:500), zPan H3 (Abcam Ab 1791), zSRSF1 (1:300), zSRSF1 (Invitrogen 32-4500 1:2000), zPCNA (Santa Cruz, Sc56) zIT7 (Novagen, 65922). Detection was by ECL.

Cell culture and transfection

Mouse embryonic fibroblast (MEF) lines were derived from 13.5 day old Psp1/p52 or embryos and their corresponding wild-type littermates [22]. They were maintained for three passages in DMEM supplemented with 15% Fetal calf serum (FCS), non-essential amino acids, sodium pyruvate, L-glutamine, and Penicillin/Streptomycin. They were transfected with Lipofectamine and GFP

Chromatin immunoprecipitation

MEFs were harvested by trypsinizing and fixed immediately with 1% formaldehyde (25°C, 10 min) in PBS, and stopped with 0.125M Glycine. Chromatin immunoprecipitation (ChiP) was performed as described previously [72]. Nuclei were sonicated using a Diagenode Bioruptor (Liege, full power 30 s, 30 s off, in an icebath for 50 min) to produce fragments of <300 bp. An arbitrary concentration of 200 μg chromatin was incubated with 4 μg rabbit IgG (Santa Cruz, sc-2025), Psp1 antibodies (A300-847), H3K36me3 antibodies (Abcam, Ab 9050-100), zH3K4me3 (Millipore 07-473) or zSRSF1 (Invitrogen 32-4500) and washed, eluted and cross-links reversed.

Histone association assay

To analyze proteins associated with H3K36me3, zH3K36me3 ChIPed chromatin was heated at 95°C in the presence of 1x Laemmli buffer for 10 min, separated on 4–20% SDS-PAGE, transferred onto a PVDF membrane, and probed with zSRSF1, z SRSF2, zSRSF3, zPTB (Invitrogen 32-4800), zPsp1 (A300-847A), and zH3K36me3 antibodies. Instead of species-specific secondary antibodies, HRP coupled Clean-Blot IP Detection Reagent (Thermo Scientific Prod. No. 21230) was used to avoid cross reactivity of HRP coupled antibody to denatured IgGs in the gel.

ChiP on chip for Psp1, H3K36me3, H3K4me3, and Srsf1

For analysis in Figure 2, WGA2 amplified ChIP DNA and input DNA were labeled and hybridized to a 3×720,000 probe custom microarray containing specific tiled regions encompassing 8.2 megabases of the mouse genome (Nimblegen). Array platform number is GPL14175 and the GEO accession numbers for ChIP data are; GSM782590 (Psp1), GSM782591 (H3K36me3), GSM782592 and GSM782593 (Srsf1 in wt), GSM782594 and GSM782595 (Srsf1 in Psp1/p52).

Biological replicates were performed for all the ChIP array experiments and the data were analyzed in R/Bioconductor (http://genomicsbiology.com/2004/5/10/R00) using the Epigenome (PROT14) protocol (http://www.epigenome-noe.net/WWW/researchtools/protocol.php?protid=43) with the following parameters; The mean signal intensity of the 4 replicate probes present on each array was calculated. Loess normalization was used within arrays to correct for dye bias, and scale normalization was used within replicate groups to control inter-array variability. Log enrichment for each group was calculated by subtracting the mean log2 input intensities from the mean of log2 ChIP-enriched intensities. Probes were tested for significant enrichment using the significance analysis of microarrays (SAM) technique [73], and the local false discovery rate based on the SAM statistic was calculated using the Locfdr R package [74]. A false discovery rate of 0.05 was used as the significance cutoff. The spearman rank correlation coefficient was used to assess the correlation between replicate experiments.

The spearman rank correlation coefficient was used on all log enrichment scores between data from Psp1 ChIP and remaining groups to determine, significance and strength of their relationship.

To determine if overlaps between Psp1, H3K36me3 and H3K4me3 enriched probes were significant, 1000 randomized datasets were produced and the 95th percentile of the resultant overlaps was used as a significance cutoff.

To determine the enrichment of probes over genomic features, probes were selected based on the following criteria. Genes were classified as expressed in MEF if they had been detected on an Illumina microarray (unpublished data) with a p value of detection <0.01. Genes classified as non-expressed in MEF cells were defined if they had a p value of detection >0.5 and a signal intensity less than 0. Only those genes that contained significantly enriched Psp1/p52 and H3K36 me3 signal were used for analysis. Exonic probes were defined as those that fall within an exon - probes falling within the 5’UTR and <200 bp from TSS were excluded. Intronic regions were defined as those that fall within an intron and >200 bp from the intron start or end site. Intergenic regions probes were selected from probes that are more than 1 Kbp from either the transcriptional start sites or transcriptional end sites of a gene. The significance of differences between genomic regions was calculated using a Wilcoxon rank sum test, with a p value cutoff <0.05.

For data in Figure 6, WGA2 amplified ChIP DNA and input DNA were labeled and hybridized to a 3×720,000 probe custom microarray containing specific tiled regions encompassing 8.2 megabases of the mouse genome (Nimblegen). Array platform number is GPL14175 and the GEO accession numbers for ChIP data are; Psp1: GSM782590, H3K36me3: GSM782591, Srsf1 (Wt MEFs): GSM782592, GSM782593, Srsf1 (Psp1/p52 MEFs): GSM782594, GSM782595. The median signal of replicate probes was taken prior to normalization. Data was normalized as above. Because levels of Srsf1 binding were generally quite low we used quantized correlation coefficients (QCC), which are less effected by the amount of binding signal present in the data, to determine the correlation between replicate experiments [75]. Across the entire

Biological replicates were performed for all the ChIP array experiments and the data were analyzed in R/Bioconductor (http://genomicsbiology.com/2004/5/10/R00) using the Epigenome (PROT14) protocol (http://www.epigenome-noe.net/WWW/researchtools/protocol.php?protid=43) with the following parameters; The mean signal intensity of the 4 replicate probes present on each array was calculated. Loess normalization was used within arrays to correct for dye bias, and scale normalization was used within replicate groups to control inter-array variability. Log enrichment for each group was calculated by subtracting the mean log2 input intensities from the mean of log2 ChIP-enriched intensities. Probes were tested for significant enrichment using the significance analysis of microarrays (SAM) technique [73], and the local false discovery rate based on the SAM statistic was calculated using the Locfdr R package [74]. A false discovery rate of 0.05 was used as the significance cutoff. The spearman rank correlation coefficient was used to assess the correlation between replicate experiments. The spearman rank correlation coefficient was used on all log enrichment scores between data from Psp1 ChIP and remaining groups to determine, significance and strength of their relationship.

To determine if overlaps between Psp1, H3K36me3 and H3K4me3 enriched probes were significant, 1000 randomized datasets were produced and the 95th percentile of the resultant overlaps was used as a significance cutoff. To determine the enrichment of probes over genomic features, probes were selected based on the following criteria. Genes were classified as expressed in MEF if they had been detected on an Illumina microarray (unpublished data) with a p value of detection <0.01. Genes classified as non-expressed in MEF cells were defined if they had a p value of detection >0.5 and a signal intensity less than 0. Only those genes that contained significantly enriched Psp1/p52 and H3K36 me3 signal were used for analysis. Exonic probes were defined as those that fall within an exon - probes falling within the 5’UTR and <200 bp from TSS were excluded. Intronic regions were defined as those that fall within an intron and >200 bp from the intron start or end site. Intergenic regions probes were selected from probes that are more than 1 Kbp from either the transcriptional start sites or transcriptional end sites of a gene. The significance of differences between genomic regions was calculated using a Wilcoxon rank sum test, with a p value cutoff <0.05.

For data in Figure 6, WGA2 amplified ChIP DNA and input DNA were labeled and hybridized to a 3×720,000 probe custom microarray containing specific tiled regions encompassing 8.2 megabases of the mouse genome (Nimblegen). Array platform number is GPL14175 and the GEO accession numbers for ChIP data are; Psp1: GSM782590, H3K36me3: GSM782591, Srsf1 (Wt MEFs): GSM782592, GSM782593, Srsf1 (Psp1/p52 MEFs): GSM782594, GSM782595. The median signal of replicate probes was taken prior to normalization. Data was normalized as above. Because levels of Srsf1 binding were generally quite low we used quantized correlation coefficients (QCC), which are less effected by the amount of binding signal present in the data, to determine the correlation between replicate experiments [75]. Across the entire
array the QCC between Srsf1 replicates was 0.37 in wild-type cells and 0.18 in Psip1/p52 cells likely reflecting a loss of overall Srsf1 binding captured in the mutant cells. However, considering only the regions on the array around exons, where most Srsf1 binding is likely to be located, the QCC in wild-type cells rises to 0.5 and to 0.23 in mutant cells. Enriched probes were identified as those above a threshold defined using the upperBoundNull method from Ringo Bioconductor Package [76]. Probes above the threshold must also be located within 300 bp of 2 or more probes to be called enriched. A hypergeometric test was applied to determine significant overlap between enriched probe groups.

Nuclear extract preparation and immunoprecipitation from NIH 3T3 cells

Nuclear extract was prepared from NIH 3T3 cells according to [77] with the following modifications: after precipitation with 1/10th vol of 4 M (NH₄)₂SO₄ and mixing for 20 min, the lysate was cleared by centrifugation at 11600g in a TL-100 ultracentrifuge (Beckman, Mountain View, CA). The supernatant was dialyzed against 3 changes of buffer C (25 mM Hepes pH 7.6, 150 mM KCl, 12.5 mM MgCl₂, 0.1 mM EDTA, 10% (v/v) glycerol, 1 mM DTT, 0.2 mM PMSF and complete protease inhibitors (Roche)) and flash frozen in liquid nitrogen. The extracts were quantified by Bradford assay (Bio-Rad). A total of 200 μg nuclear extract were immunoprecipitated by incubation for 45 minutes at 4°C with 5 μg rabbit IgG, 5 μg rabbit A300-847(1:200 dilution, Bethyl laboratories,) and 10 μl Protein A Dynal beads. After washing three times with buffer C, but containing 200 mM KCl, for 10 min each, the bound proteins were boiled in SDS sample buffer, separated on a 4–20% tris glycine polyacrylamide gel and either stained with colloidal coomassie (Invitrogen) to identify the proteins, or transferred to nitrocellulose membrane for western blotting. Individual protein bands or 1 cm² gel pieces were cut and subjected to mass spectrometry analysis.

MS/MS analysis

Excised gel pieces were treated with trypsin at 37°C and the peptides extracted with 10% formic acid. Peptides were separated using an Ultimate nanoLC (LC Packings, Amsterdam) equipped with a PepMap C18 trap & column. The eluent was sprayed into a Q-Star tandem mass spectrometer (Applied Biosystems, Foster City, CA) and analyzed in Information Dependent Acquisition (IDA) mode, performing 1 s of MS followed by 3 s MSMS analyses of the 2 most intense peaks seen by MS. The MS/MS data file generated was analyzed using the Mascot 2.1 search engine (Matrix Science, London, UK) against UniProt April 2009 (7966092 sequences) or NCBInr March 2010 (10530540 sequences) databases with no species restriction. The data was searched with tolerances of 0.2 Da for the precursor and fragment ions. The Mascot search results were accepted if a protein hit included at least 2 peptides with a score above the homology threshold.

In vitro pulldowns

For p52 pulldown, T7 tagged SRSF1, SRSF3 and RG and RD mutants of SRSF1 and GFP- SRSF2, were overexpressed in 293T cells [49,78], and the cell lysates incubated with Glutathione beads coupled with p52 in GST lysis buffer. Unbound proteins were washed 5 times with the same buffer. Bound proteins were separated on 12% SDS PAGE. After transferring to nitrocellulose membrane, the proteins were probed with α77 monoclonal antibody (Novagen) and imaged.

For histone pulldowns, 1 μg of T7 tagged SRSF1, purified from 293T cells, was incubated with T7 beads in GST lysis buffer for 1 hr at 4°C. After washing unbound proteins in same buffer, 1 μg of GST-p52 and 1 μg HeLa core histones (Active motif, cat. 55301) were added and incubated for 3 hrs. Unbound proteins were washed off 5 times with the same buffer and bound proteins were separated on 17% SDS-PAGE. After transferring to nitrocellulose membrane, the proteins were probed with αH3K36me3 antibodies and imaged. The membrane was then stripped and reprobed with αH3K14me2 and αH3K18me3 antibodies.

Immunofluorescence

Cells grown on slides were fixed in 3% paraformaldehyde (pFA) as previously described [79] and incubated with primary antibodies; rabbit A300-847(1:200 dilution, Bethyl laboratories,) which recognizes an epitope (a.a. 225–275) present in both p52 and p75, A300-848(1:200, Bethyl laboratories) which recognizes only p75 (a.a. 480–530), mouse monoclonal αSc35 (1:50, Sigma S4045), Secondary antibodies, and image capture by wide-field epifluorescence microscopy were as previously described [79]. Confocal analysis was performed using a Zeiss LSM510 confocal microscope.

Alternative splicing microarray

Microarray analysis of alternative splicing was performed as described [53]. Five hundred ng total RNA, isolated from primary MEFs derived from three littermates of E13.5 wild-type or Pspip1/p52 embryos [22], were used to generate sense-strand cDNA (Ambion WT expression kit #411974). Purified cDNA was fragmented and labelled with biotin-conjugated nucleotides using terminal transferase (Affymetrix, #900670). Arrays were hybridized with labelled cDNA for 16 h at 50°C in 7% dimethylsulfoxide. Washing and detection were performed in an Affymetrix Fluidics Station using standard protocols for eukaryotic targets [53]. Scanned microarrays were analyzed using ASPIRE3 (Analysis of SPliCing Isoform Reciprocity, version 3) [52], which predicts splicing changes from reciprocal sets of microarray probes that recognize either inclusion or skipping of an alternative exon. Data were quantified as the change in the fraction of exon inclusion (AI), where a value of 1.0 indicates a 100% increase, and −1.0 a 100% decrease in exon inclusion.

RT–PCR

Primers corresponding to exons flanking the alternate spliced exons were designed (Table S2). 5 μg of RNA was reverse transcribed with superscript reverse transcriptase II (Invitrogen) using random primers, and each of the forward primers were labeled with 32P γ-ATP. PCR was performed for 24–30 cycles, and the products were separated on a 5% denaturing polyacrylamide gel and analyzed by autoradiography for 3–16 h, or separated on 1.5% agarose gel.

Supporting Information

Figure S1 A300-847 antibody immunoprecipitates from wild-type and mutant cells. A) Silver-stained gel of immunoprecipitates with IgG, and A300-847 antibodies (anti Psip1-p52/p75) from nuclear extracts prepared from wild-type and Pspip1/p52 MEFs, 5% of the nuclear extract was loaded as input. Duplicate gel was stained with colloidal coomassie (Invitrogen), and 1 cm² of the lanes corresponding to molecular weight of 25–40 KDa (indicated by boxed area) were subjected to mass spectrometry. Srsf1, Srsf5 and hnRNPa were identified from I.P with wild-type nuclear extracts prepared from NIH 3T3 WT expression kit [22], were used to generate sense-strand cDNA (Ambion WT expression kit #411974). Purified cDNA was fragmented and labelled with biotin-conjugated nucleotides using terminal transferase (Affymetrix, #900670). Arrays were hybridized with labelled cDNA for 16 h at 50°C in 7% dimethylsulfoxide. Washing and detection were performed in an Affymetrix Fluidics Station using standard protocols for eukaryotic targets [53]. Scanned microarrays were analyzed using ASPIRE3 (Analysis of SPliCing Isoform Reciprocity, version 3) [52], which predicts splicing changes from reciprocal sets of microarray probes that recognize either inclusion or skipping of an alternative exon. Data were quantified as the change in the fraction of exon inclusion (AI), where a value of 1.0 indicates a 100% increase, and −1.0 a 100% decrease in exon inclusion.

RT–PCR

Primers corresponding to exons flanking the alternate spliced exons were designed (Table S2). 5 μg of RNA was reverse transcribed with superscript reverse transcriptase II (Invitrogen) using random primers, and each of the forward primers were labeled with 32P γ-ATP. PCR was performed for 24–30 cycles, and the products were separated on a 5% denaturing polyacrylamide gel and analyzed by autoradiography for 3–16 h, or separated on 1.5% agarose gel.

Supporting Information

Figure S1 A300-847 antibody immunoprecipitates from wild-type and mutant cells. A) Silver-stained gel of immunoprecipitates with IgG, and A300-847 antibodies (anti Psip1-p52/p75) from nuclear extracts prepared from wild-type and Pspip1/p52 MEFs, 5% of the nuclear extract was loaded as input. Duplicate gel was stained with colloidal coomassie (Invitrogen), and 1 cm² of the lanes corresponding to molecular weight of 25–40 KDa (indicated by boxed area) were subjected to mass spectrometry. Srsf1, Srsf5 and hnRNPa were identified from I.P with wild-type nuclear extract, but not from I.P with Pspip1/p52 nuclear extract. B) Western blot of A300-847 IPs from wild-type and Pspip1/p52 (A) with anti Srsf1.

(PDF)
Table S1  Peptide array quantification values of all 59 histone modifications. Cells highlighted with yellow are with specificity factor >2, and were used to generate Figure 1C.

Table S2  Alternative splicing array results in Excel spread sheets (Sheet 1 includes data from all the exons on the array, Sheet 2 includes only exons which shows significant changes in alternative splicing between wild-type and Psip1/−/− cells, Sheet 3 includes the annotation for the data. Related to Figure 3.

Table S3  Sequence of PCR primers used for RT-PCR validation of alternative splicing events in wild-type, Psip1+/+/*, and Psip1/−/− cells.

References


Acknowledgments

We acknowledge Melis Kayikci (MRC LMB) for analyzing the splicing microarray data. We thank Javier Caceres (MRC HGU) for key discussions about mRNA splicing and SRSF1 antibodies and the St. Andrews Biomedical Research Support centre (BRSC) for mass spectrometry. We are very grateful to Alan Engelman (Dana-Farber Cancer Institute) for Psip1−/− MEFs and Psip1 plasmids.

Author Contributions

Conceived and designed the experiments: MMP WAB. Performed the experiments: MMP HG5 HU. Analyzed the data: MMP JU GRG. Wrote the paper: MMP WAB.


