Argonaute-1 binds transcriptional enhancers and controls constitutive and alternative splicing in human cells

Mariano Alló[a,1,2], Eneritz Agirre[b,1,3], Sergey Bessonov[c,1], Paola Bertucci[a,1], Luciana Gómez Acuña[a,1], Valeria Buggiano[a], Nicolás Bellora[b,4], Babita Singh[b], Ezequiel Petrillo[a], Matías Blaustein[a], Belén Miñana[b,d], Gwendal Dujardin[a], Berta Pozzi[a], Federico Pelisch[a], Elías Bechara[b,d], Dmitry E. Agafonov[c], Anabella Srebrow[a], Reinhard Lührmann[c], Juan Valcárcel[b,d,e], Eduardo Eyras[b,e,5], and Alberto R. Kornblihtt[a,5]

[a]Laboratorio de Fisiología y Biología Molecular, Departamento de Fisiología, Biología Molecular y Celular, Instituto de Fisiología, Biología Molecular y Neurociencias-Consejo Nacional de Investigaciones Científicas y Técnicas, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires, C1428EHA Buenos Aires, Argentina; [b]Computational Genomics, Universitat Pompeu Fabra, E-08003 Barcelona, Spain; [c]Max Planck Institute for Biophysical Chemistry, 37077 Göttingen, Germany; [d]Centre for Genomic Regulation, E-08003 Barcelona, Spain; and [e]Catalan Institution for Research and Advanced Studies, E-08010 Barcelona, Spain

The roles of Argonaute proteins in cytoplasmic microRNA and RNAi pathways are well established. However, their implication in small RNA-mediated transcriptional gene silencing in the mammalian cell nucleus is less understood. We have recently shown that intronic siRNAs cause chromatin modifications that inhibit RNA polymerase II elongation and modulate alternative splicing in an Argonaute-1 (AGO1)-dependent manner. Here we used chromatin immunoprecipitation followed by deep sequencing (ChIP-seq) to investigate the genome-wide distribution of AGO1 nuclear targets. Unexpectedly, we found that about 80% of AGO1 clusters are associated with cell-type-specific transcriptional enhancers, most of them (73%) overlapping active enhancers. This association seems to be mediated by long, rather than short, enhancer RNAs and to be more prominent in intragenic, rather than intergenic, enhancers. Paradoxically, crossing ChIP-seq with RNA-seq data upon AGO1 depletion revealed that enhancer-bound AGO1 is not linked to the global regulation of gene transcription but to the control of constitutive and alternative splicing, which was confirmed by an individual gene analysis explaining how AGO1 controls inclusion levels of the cassette exon 107 in the *SYNE2* gene.

Argonaute proteins | transcriptional enhancers | alternative splicing

Alternative splicing was initially seen as an interesting mechanism to explain protein diversity but affecting a limited number of mammalian genes. The recent development of high-throughput sequencing technologies has dramatically changed this view, generating a renewed interest in alternative splicing. We now know that alternative splicing affects transcripts from more than 90% of human genes (1) and that normal and pathological cell differentiation not only depends on differential gene expression but also on alternative splicing patterns. Mutations in alternative splicing regulatory sequences and factors are involved in the etiology of numerous hereditary diseases, premature aging, and cancer (2).

Recently, amid an avalanche of papers reporting various connections between the chromatin context and splicing (3–9), a relationship between splicing and small RNAs has emerged. The convergence of these previously unrelated areas (RNA interference, chromatin, and splicing) has been studied by our laboratory, showing that siRNAs (20–25 nt long) targeting both intronic and exonic regions near the cassette exon 33 (E33, also known as EDI) of the fibronectin gene were able to regulate its alternative splicing by affecting the chromatin context at the target region, with an increase of histone tail modifications associated with gene silencing (H3K9me2 and H3K27me3, i.e., dimethylation of lysine 9 and trimethylation of lysine 27 of

histone H3 respectively). Moreover, this effect was shown to be dependent on Argonaute proteins (AGO1 and AGO2) and involves a decrease of RNA polymerase II (RNAPII) elongation, which concomitantly up-regulates E33 inclusion into the mature mRNA (3). More recently, a similar effect was found over the variant region of the endogenous *CD44* gene, where both AGO1 and AGO2 were recruited after treatment with phorbol-12-myristate-13-acetate in a Dicer-dependent manner, favoring the methylation of H3K9 and the concomitant recruitment of the

Significance

Argonaute proteins are well characterized factors in post-transcriptional gene silencing, the process by which small RNAs trigger mRNA degradation or inhibit translation in the cytoplasm. We report here that Argonaute proteins also play important roles in the nucleus. Our genome-wide analysis reveals that Argonaute-1 (AGO-1) binds preferentially to active transcriptional enhancers and that this association is mediated by the RNAs that are transcribed from these enhancers (eRNAs). Moreover, the interaction of AGO-1 with enhancers does not seem to regulate transcription of the neighboring genes but of alternative and constitutive splicing. These results contribute to the understanding of the complex regulation of gene expression in eukaryotic cells.

heterochromatin protein HP1 (10). Ameyar-Zazoua et al. also showed that the AGO-mediated mechanism results in a reduction in RNAPII elongation rate that affects *CD44* alternative splicing (10). When first described, with the fibronectin gene as the only example, the control of splicing by nuclear RNAs through chromatin changes was called TGS-AS, for transcriptional gene silencing-regulated alternative splicing (3). However, its endogenous extension and frequency within human cells remained unknown. Consistent with a more general role of Argonaute proteins in the nucleus, two genome-wide surveys in *Drosophila melanogaster* have shown that AGO2 regulates alternative splicing as well as transcription of target genes (11, 12), and a more recent work in human cancer cells suggests that AGO1 interacts with RNAPII and binds to transcriptionally active promoters (13).

In an attempt to investigate the nuclear roles of AGO1 in relation with alternative splicing at a genome-wide level in human cells, here we performed high-throughput DNA and RNA sequencing after AGO1 immunoprecipitation (ChIP-seq) or AGO1 depletion (RNA-seq), respectively. Our analysis uncovers previously unidentified roles for nuclear AGO1, revealing a specific binding to enhancers, bidirectional promoters, and at the 5′region of first introns. Additionally, AGO1 binding to active enhancers seems particularly associated with long rather than small nuclear RNAs.

AGO1 depletion revealed changes in constitutive splicing for about 1,800 internal introns and in the patterns of ~700 alternative splicing events. A detail characterization of one of these events mechanistically illustrates how AGO1 regulates exon skipping when recruited to an enhancer located in the downstream intron of a cassette exon.

Results

To investigate the genome-wide distribution of AGO1 in the nucleus, we performed chromatin immunoprecipitations followed by deep-sequencing (ChIP-seq) in the breast cancer-derived human cell line MCF7, using separately two commercially available antibodies for AGO1: 6D8.2 from Millipore and 4B8 from Sigma. Each antibody recognized a single band in Western blots of MCF7 cell extracts, in agreement with the suppliers' specifications. However, the bands recognized differed in molecular weight, being of around 150 kDa for 6D8.2 and of 90 kDa for 4B8. Surprisingly, ChIP-seq results for both AGO1 antibodies showed important differences in their genomic distribution. For 6D8.2 we observed a widespread enrichment on CpG islands, promoters, 5′ UTRs, exons, and coding sequences (CDS), and exclusion from introns and intergenic regions (Fig. 1*A*). In contrast, using 4B8 we found a more discrete but unexpected enrichment pattern, with significant enrichment at enhancers, 5′ UTRs, bidirectional promoters, and the first 300 nucleotides of intron 1 (Fig. 1*B*). Similar differential patterns were found in the epithelial cell line MCF10 (*SI Appendix*, Fig. S1), often used as a nonmalignant control of MCF7 cells (14). In view of this discrepancy, we decided to investigate in more depth the specificity of both antibodies (*SI Appendix*) and found that only 4B8 was a bona fide antibody against AGO1, which led us to ignore the ChIP-seq results obtained with 6D8.2 (Fig. 1*A*).

Endogenous AGO1 Is Enriched at Enhancers. The most interesting result from Fig. 1*B* (and also *SI Appendix*, Fig. S1*B*) is that the signal recognized by the 4B8 antibody in both MCF7 and MCF10 cells is selectively enriched in enhancers and bidirectional promoters. Interestingly, AGO1 is enriched around transcription start sites (TSSs) exhibiting two shoulders (*SI Appendix*, Fig. S4*A*). When bidirectional promoters are removed from the analysis the signal is reduced around the TSS, and when enhancers are removed most of the AGO1 signal disappears. There remains, however, some signal downstream of the TSS, which agrees with the observed enrichment of AGO1 on the first 300 bp of intron 1 (Fig. 1). We conclude that AGO1 clusters at promoters could be explained mostly by the contribution from enhancers and, to a lesser extent, by bidirectional promoters (*SI Appendix*, Fig. S4 *A* and *B*).
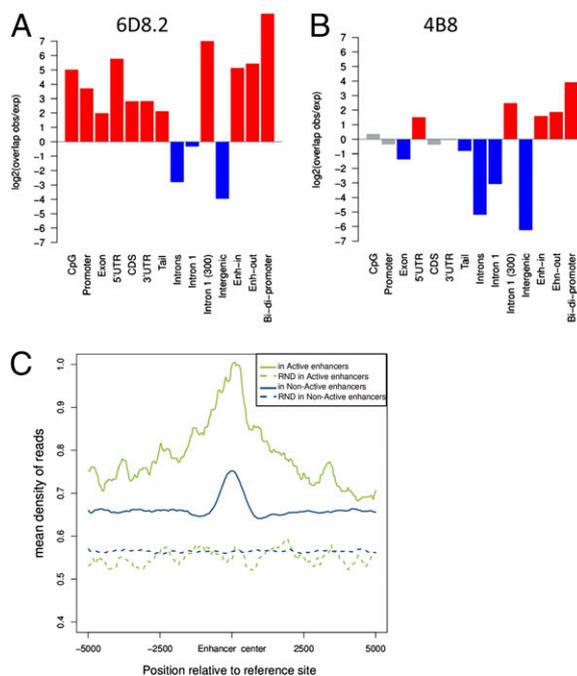
Fig. 1. AGO1 is enriched at enhancers. Enrichment analysis of two ChIP-seq experiments using the commercial antibodies to AGO1 6D8.2 from Millipore (*A*) and 4B8 from Sigma (*B*) in MCF7 cells across various genomic regions: CpG islands (CpG), promoters, exons, 5′ UTRs (5′UTR), coding exons (CDS), 3′ UTRs (3′UTR), downstream of the poly(A) site (Tail), introns, first intron (Intron1), 300 nt downstream of the first 5′ splice site [Intron 1 (300)], intergenic regions (Intergenic), intragenic enhancers (Enh-in), intergenic enhancers (Enh-out), and bidirectional promoters (Bidi-promoters). The *y* axis represents the relative enrichment of observed over expected overlaps as a log₂ rate. Significant enrichment and depletion are indicated as red and blue bars, respectively, with gray indicating no significant difference (see *Methods* for details). (*C*) Density profile of AGO1 reads from significant clusters in active and nonactive enhancers. Dashed lines correspond to the density profile of reads from randomized clusters (*Methods*).

To further investigate the association of AGO1 to enhancers we used mutual information (mi), which is a measure of association that compares the frequency of the two features appearing together to the frequency of the two features appearing independently (*Methods*). We consider enhancer regions predicted according to various histone mark levels measured in a group of nine cell lines (15). We found that AGO1 associates significantly with intragenic enhancers (mi = 0.506, P = 2.2e-16). Moreover, dividing enhancers according to whether they are likely to be active or not in MCF7 cells by analyzing H3K4me3 (trimethylation of lysine 4 of histone H3) and H3K27ac (acetylation of lysine 27 of histone H3) levels (*Methods* and *SI Appendix*, Fig. S5), we observed an even stronger association of AGO1 at intragenic active enhancers (mi = 0.93, P = 2.2e-16). Indeed, from the 1,951 intragenic enhancers classified as active, 128 (6.56%) are bound by AGO1, whereas only 7 (0.76%) of the 917 intragenic nonactive ones are bound by AGO1 (Fisher exact test P = 5.662e-14). Consistent with this result, there is a higher accumulation (larger number of reads) of AGO1 in active enhancers compared with nonactive ones (Fig. 1*C*) and the signal is more prominent in intragenic, compared with intergenic, active enhancers (*SI Appendix*, Fig. S6).

It has been reported that AGO1 is associated with particular histone modifications (16, 17), and we have previously shown that AGO1 is necessary for the TGS-AS mechanism in which siRNAs elicit intragenic H3K9me2 and H3K27me3 histone modifications (3). Interestingly, we observed an enrichment of these histone marks at AGO1 significant clusters (*SI Appendix*, Fig. S7*A*) and at enhancers (Fig. 2*A*). This raised the question of whether this enrichment may be dependent on enhancer activity.

Indeed, we found that whereas AGO1 accumulates at intragenic active enhancers, H3K9me2 and H3K27me3 signals do not accumulate on these sites (*SI Appendix*, Fig. S7B). In contrast, at enhancers classified as nonactive the two silencing marks accumulate similarly to AGO1 (*SI Appendix*, Fig. S8). Thus, the association of silencing histone modifications and AGO1 at enhancers is independent of enhancer localization but dependent on enhancer activity, and H3K9me3 and H3K27me3 mainly overlap with AGO1-bound, nonactive enhancers.

We further studied the association of pairs of different histone marks, including now the typically intragenic H3K36me3 (trimethylation of lysine 36 of histone H3) modification, with AGO1 at enhancers. At the genome-wide level the overlap between pairs of histone marks is lower than 20% (Fig. 2B, *Top Left*). However, at AGO1 target sites the overlap is higher for all possible combinations, reaching up to 55% overlap (Fig. 2B, *Top Right*). Additionally, in active enhancers the overlap between marks seems to be as low as

at the genome-wide level (Fig. 2B, *Middle Left*), whereas at AGO1-bound active enhancers it reaches the 70% for enhancers that have both H3K27me3 and H3K9me2 (Fig. 2B, *Middle Right*) (χ^2 $P = 0.0001743$). Interestingly, the enrichment seems asymmetric, that is, the proportion of enhancers with H3K27me3 that also have H3K36me3 (60%) is higher than for enhancers with H3K36me3 that also have H3K27me3 (16%) (Fig. 2B, *Middle Right*). This could indicate two different activities of AGO1 on enhancers. Similarly, there is an enrichment of overlaps between histone marks on nonactive AGO1-bound enhancers compared with nonactive enhancers with no AGO1 signal (Fig. 2B, *Bottom*). Furthermore, co-occurrence of all pairs of marks in active enhancers with AGO1 is significantly higher compared with their overlap genome-wide and on enhancers with no AGO1 (*SI Appendix*). These results suggest a role for AGO1 either in writing or reading chromatin marks, more prominently in enhancer regions.

AGO1 Preferentially Associates with Active Enhancers but Does Not Regulate Transcription. Because enhancers are regulatory elements whose activity is highly variable between cell lines we decided to further investigate the link between AGO1 and the activity of enhancers by comparing our AGO1 cluster dataset with a subset of MCF7 cell enhancers characterized by estrogen dependency (17). The enhancers analyzed are those binding estrogen receptor α (ERα) that have high levels of H3K4me and H3K27ac upon estradiol treatment. Using global run-on sequencing (GRO-seq) data obtained after estradiol treatment (17), these enhancers were separated into those most likely to be active and those most likely to be nonactive (*Methods*). Strikingly, from the 14,205 AGO1 clusters, 73.3% (10,412) were found to overlap active enhancers, whereas only 7.82% (1,111) AGO1 clusters fall in the nonactive ones. Although a small proportion (7.05%) of active enhancers are bound by AGO1, this is significantly larger than the 1.02% of nonactive enhancers with AGO1 (χ^2 test $P = 2.2e-16$). Additionally, active enhancers have significantly higher densities of AGO1 reads than nonactive ones (Kolmogorov–Smirnov test $P < 2.2e-16$) and active enhancers with AGO1 have a significantly higher density of GRO-seq reads than active enhancers without AGO1 (Kolmogorov–Smirnov test $P < 2.2e-16$) (*SI Appendix*, Fig. S9). These and previous results suggest that AGO1 is related to enhancer activity.

A logical consequence of the finding of AGO1 on transcriptional enhancers is to investigate a putative general role of the Argonaute protein on gene transcription. For this we performed RNA-seq of MCF7 cells to assess mRNA levels genome-wide before and after depletion of AGO1 by RNAi. Interestingly, more than 1,000 genes change expression significantly upon AGO1 knockdown (*Methods*): 813 genes were up-regulated whereas 461 showed a reduction in mRNA levels (Dataset S1). However, no relevant correlations seem to exist between the presence of AGO1 clusters on genes and their mRNA levels. In fact, only about 6% (226) of the genes with AGO1 clusters, either intragenically or at their promoters, show expression level changes comparable to genes without AGO1 clusters (~3%, 1,048). Similarly, when genes linked to enhancers are considered using RNAPII ChIA-PET data (18) the difference remains small: from the 1,545 genes with AGO1 at the linked enhancer, 7.31% (113) changed expression, whereas 6.53% (345) of those with no AGO1 at the linked enhancer had a change in expression (Fisher exact test $P > 0.05$). Furthermore, among the 1,274 genes with significant change in expression upon AGO1 knockdown, 594 (46.62%) contain an AGO1 cluster within 100 kbp upstream and 672 (52.74%) between 100 kbp and 300 kbp upstream. However, these percentages are similar when genes that are not regulated by AGO1 knockdown are considered: Of the 13,351 nonregulated genes, 6,638 (49.71%) have an AGO1 cluster within 100 kbp upstream and 7,711 (55.51%) between 100 kbp and 300 kbp upstream (Wilcoxon signed rank test $P > 0.05$).

All together, these analyses indicate that the binding of AGO1 to transcriptional enhancers is not primarily related to the control of transcription. Nevertheless, further investigation is needed
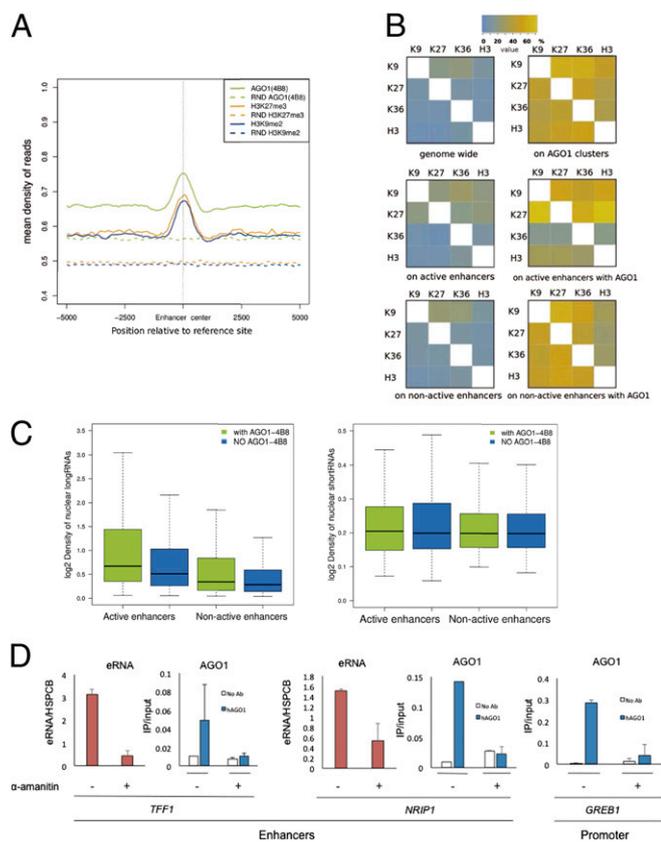
Fig. 2. AGO1 association with silencing marks and RNA at enhancers. (*A*) Density of ChIP-seq reads for AGO1 (4B8 antibody), H3K9me2, and H3K27me3 centered at enhancers. The dashed lines correspond to the read densities from randomized clusters. (*B*) Association of histone marks genome-wide and on enhancers, in relation to the presence of AGO1 binding. The heat maps indicate the proportion (percentage of clusters) of significant sites with a given chromatin signal (rows) that overlap with each chromatin signal (columns). (*Left*) The data for overlaps genome-wide (*Top Left*), on active enhancers (*Middle Left*), and on nonactive enhancers (*Bottom Left*). (*Right*) The proportion of overlap (percentage of clusters) of the same histone marks at sites with AGO1: genome-wide (*Top Right*), on active enhancers (*Middle Right*), and on nonactive enhancers (*Bottom Right*). (C) Density of long (*Left*) and short (*Right*) nuclear RNA-seq reads from MCF7 cells in active and nonactive enhancers with (green) and without (blue) AGO1 binding, respectively. (*D*) Effects of treating MCF7 cells with 2.5 μg/mL alpha-amanitin for 20 h on the abundance of eRNAs (pink bars) and on the AGO1 ChIP signal (blue bars) on the *TFF1* and *NRIP1* transcriptional enhancers (*Left* and *Center*, respectively) and on the *GREB1* promoter (*Right*). White bars correspond to ChIP signals in the absence of antibody. eRNAs and DNA from ChIP were quantified by real-time PCR using specific primers. In all experiments $n = 2$ except for *NRIP* AGO1 signal, where $n = 1$.
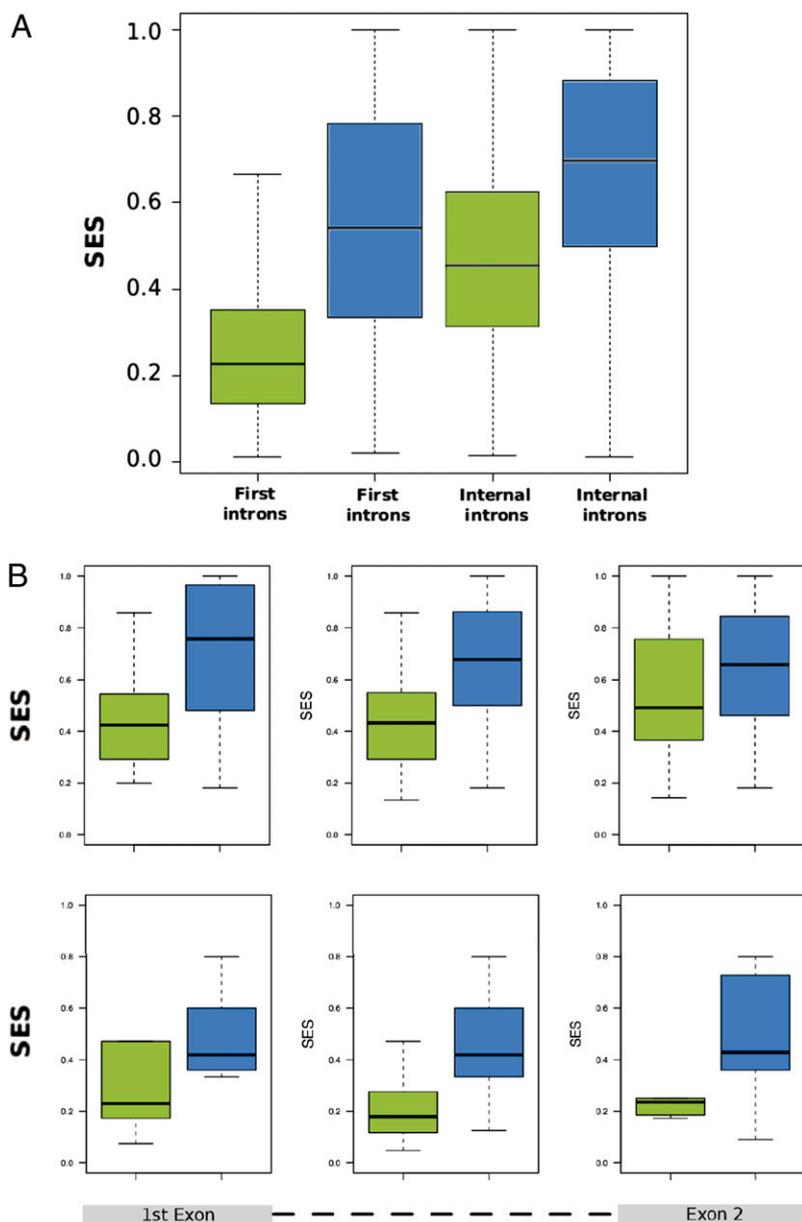
Fig. 3. AGO1 affects constitutive splicing. (*A*) SES (*Methods*) distributions depicted as boxplots in the samples siLuc (green) and siAGO1 (blue) for the introns with significant SES change. Kolmogorov–Smirnov test of the comparison of the distributions yields P = 3.331e-16 for first introns and P < 2.2e-16 for internal introns. (*B*) Distribution of SES scores for introns with AGO1 binding according to the location of the AGO1 clusters in internal introns (*Upper*) and first introns (*Lower*). We considered three possible configurations: AGO1 overlaps with the upstream exon (*Left*), with the downstream exon (*Right*), or it is entirely located in the intron (*Center*).

to fully explain the changes in gene expression observed upon AGO1 depletion.

AGO1 and Enhancer RNAs. Because enhancer activity has been associated with transcription of enhancer RNAs (eRNAs) (16, 18–21) we analyzed the density of nuclear long and short RNA on enhancers from RNA-seq experiments in MCF7 cells (*Methods*). We found that AGO1-bound enhancers have significantly higher density of long nuclear RNAs relative to enhancers without significant AGO1 signal (P < 0.005) (Fig. 2*C*, *Left*). This difference becomes in fact higher for active enhancers compared with nonactive ones (P < 2.2×10^{-16}) (Fig. 2*C*, *Left*). In contrast, for small nuclear RNAs we observed no differences between AGO1-bound and AGO1-depleted enhancers (Fig. 2*C*, *Right*). These properties were similar after separating enhancers into intergenic and intragenic (*SI Appendix*, Fig. S10).

To experimentally determine whether the association of AGO1 to enhancers required transcription by RNA polymerase II we assessed the effects of inhibiting this enzyme by treating cells with alpha-amanitin. Fig. 2*D*, *Left* and *Center* shows that in the presence of the inhibitor both eRNA levels and AGO1 ChIP

signals are significantly decreased in two enhancers (*TFF1* and *NRIP1*) of those displaying AGO1 signal in the ChIP-seq experiment. Furthermore, inhibition of Pol II transcription also abolishes the AGO1 signal observed in a promoter (*GREB1*, Fig. 2*D*, *Right*). These results strongly suggest that the association of AGO1 to enhancers and promoters at the chromatin level is mediated by the RNAs transcribed from these elements.

AGO1 Affects Constitutive and Alternative Splicing Genome-Wide. Taking into account the observation that a bigger proportion of AGO1-bound enhancers are intragenic (75.73%), the lack of correlation between AGO1 on enhancers and mRNA expression levels, and the already described relationship between AGO1 and splicing factors (10, 22), we surveyed a possible role of AGO1 in constitutive and/or alternative splicing. Interestingly, AGO1 does not seem to accumulate over exons (*SI Appendix*, Fig. S11); hence, its effect on splicing may be exerted from intronic regions. Accordingly, to assess constitutive splicing we measured the splicing efficiency score (SES) of introns from RNA-seq, defined as the ratio of spliced reads that define an intron over the sum of those spliced reads and the nonspliced
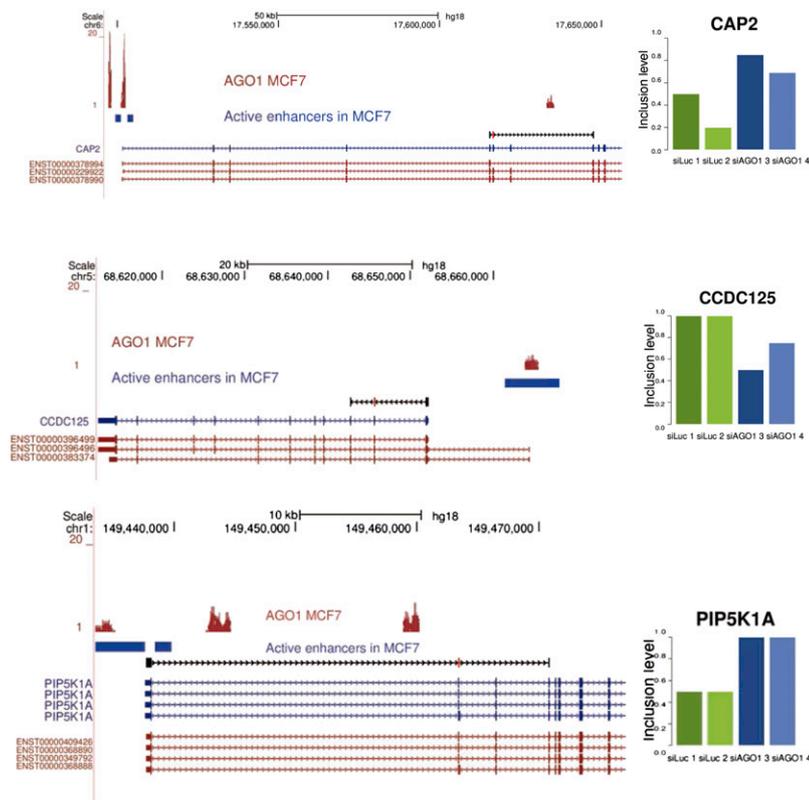
Fig. 4. Alternative splicing upon AGO1 knockdown. (*Left*) The genomic loci for three genes, *CAP2*, *CCDC125*, and *PIP5K1A*, with intragenic active enhancers (blue rectangles) binding AGO1 (red profiles) and containing an alternative splicing event (black track) that changes inclusion upon knockdown of AGO1. (*Right*) The inclusion levels (*y* axis) for the two replicates in control (green bars) and the knockdown of AGO1 (blue bars). Active enhancers and regulated events were determined as specified in *Methods*.

reads that overlap with the exon–intron boundary (see *Methods* for details) in cells transfected with siLuc (control) or siAGO1. Interestingly, there is a significant difference in SES comparing cells depleted of AGO1 with controls, for first introns (Wilcoxon sign rank test $P < 2.2$e-16) and for internal introns (Wilcoxon sign rank test $P < 2.2$e-16) (Fig. 3*A*). Moreover, the change in SES upon AGO1 knockdown is significant for 150 first introns and 1,767 internal introns (*Methods*), from which 91.97% and 96.87%, respectively, overlap with AGO1 clusters. In contrast, only 9.73% and 3.54% of the 2,609 first and 14,250 internal introns, respectively, that do not change significantly have overlap with AGO1. Most of the significant changes imply an increase in splicing efficiency: 87.3% for first introns and 74.2% for internal introns. Additionally, this effect seems to be independent of the position of the AGO1 cluster for first introns (Fig. 3*B*), whereas for internal introns the effect seems to be stronger when AGO1 overlaps the upstream exon (Fig. 3*B*). This indicates that AGO1 inhibits intron excision.

In view of these results, we wondered whether intragenic enhancers, upon activation or inactivation and in combination with AGO1, might affect the alternative splicing of nearby exons, possibly through the association to local chromatin changes or by affecting RNAPII elongation. Upon AGO1 knockdown in MCF7 cells, 354 and 305 alternative splicing cassette exon events undergo significant up-regulation and down-regulation of exon inclusion, respectively ([Dataset S2](#)). More interestingly, we found that 23% of the genes that modify their alternative splicing pattern upon AGO1 knockdown contain an AGO1 cluster at an intragenic enhancer that maps at either the downstream or the upstream intron of the alternatively spliced exon. This figure falls to 12% for genes whose splicing is not regulated by AGO1 knockdown, which is significantly lower (Fisher exact test $P = 2.972$e-07). Some examples of the alternative splicing events regulated by AGO1 are shown in Fig. 4. This evidence, together with our previous findings (3), suggests that AGO1 binding to transcriptional enhancers serves to regulate neighboring alternative splicing events.

SYNE2, a Convergent Example. To mechanistically explore how AGO1 regulates alternative splicing through its binding to an enhancer, among the several alternative splicing events regulated by the knockdown of AGO1 we chose the *SYNE2* gene (Fig. 5*A*) to illustrate the cross-talk between these processes. *SYNE2* is a giant gene of 373 kbp and 116 exons. In MCF7 cells, its alternative exon 107 (E107) is characterized by an AGO1 cluster spanning the 3′ half of intron 107 (I107) and part of E108 and the histone marks H3K27ac and H3K4me3, characteristic of enhancers (Fig. 5 *B* and *C*). All these features prompted us to investigate whether *SYNE2* E107 inclusion is regulated in an AGO1-dependent manner, in particular, through the TGS-AS mechanism (3, 23). Because an essential part of the TGS-AS mechanism is inhibition of RNAPII elongation, we first evaluated how *SYNE2* E107 responds to activators and inhibitors of elongation. The chromatin-relaxing drugs trichostatin A (TSA) and 5 aza deoxycytidine (5azadC) are known to affect alternative splicing by promoting elongation (3, 9, 24), whereas the topoisomerase inhibitor camptothecin (CPT) was shown to inhibit elongation (24–27). Unlike the exon E33 of fibronectin gene (*FN1*), where reduction in elongation promotes inclusion (Fig. 5*D*, *Lower*) (3, 24), SYNE E107 behaves in the opposite way: Promotion of elongation increases E107 inclusion, whereas inhibition of RNAPII elongation promotes E107 skipping (Fig. 5*D*, *Upper*), similarly to what was recently reported for the CFTR alternative exon 9 (24, 27, 29), for which it was demonstrated that inhibition of elongation provides more time for the recruitment of the negative factor ETR-3 to inhibit exon 9 inclusion into the nascent pre-mRNA (24). Owing to the presence of an AGO1 cluster downstream of E107, we speculated that endogenous AGO1 over the enhancer could locally favor E107 skipping by affecting RNAPII elongation. In agreement with this, RNA-seq data show that AGO1 knockdown increases E107 inclusion in MCF7 cells (Fig. 5*E*). To further test this mechanism, we applied the TGS-AS strategy by using intronic siRNAs (3) targeting I107 DNA sequences located near the AGO1 endogenous cluster. We transfected three siRNAs targeting different sequences in
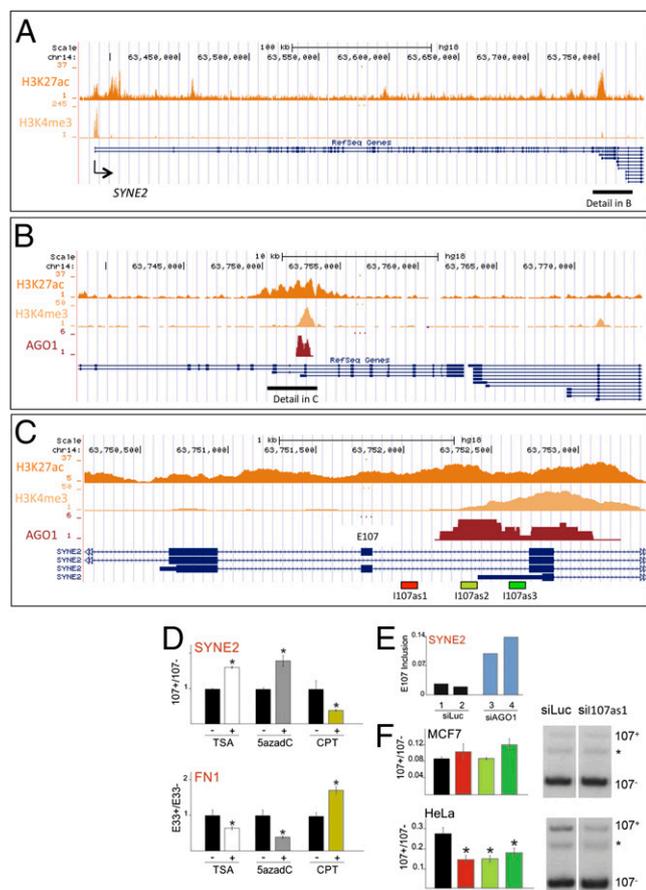
Fig. 5. *SYNE2* alternative splicing is regulated by AGO1. (*A*) Genomic locus of the SYNE 2 human gene indicating the distribution of the H3K27ac and H3K4me3 marks. (*B*) Higher resolution of the *SYNE2* gene corresponding to the region encompassed by the black horizontal line in *A*, showing overlap of the H3K27ac and H3K4me3 marks with a conspicuous AGO1 cluster. (*C*) Detail of the *SYNE2* gene around exon E107 corresponding to the region encompassed by the black horizontal line in *B*. (*D*) Effect of different drugs affecting RNAPII elongation on alternative splicing of *SYNE2* exon E107 and fibronectin exon E33. The elongation-promoting reagents TSA and 5azadC increase E107 inclusion (*Upper*), whereas they cause E33 skipping (*Lower*). Conversely, the inhibitor of elongation CPT promotes E107 skipping but increases E33 inclusion. (*E*) Transfection of MCF7 cells with siAGO1, using siLuc as control, increases *SYNE2* E107 inclusion. (*F*) Effects of intronic siRNAs in MCF7 (*Upper*) and HeLa (*Lower*) cells on E107 alternative splicing. E107+ /E107− and E33+/E33− ratios were determined using reverse transcriptase-radioactive PCR (RT-rPCR). Gel images for the RT-PCRs of siLUC and siI107as1 experiments are shown on the right of each quantification. The asterisk indicates a spurious band. Note that in HeLa cells siI107as1 not only decreases the intensity of the inclusion band (E107+) but also decreases the intensity of the exclusion band (E107−).

I107 (Fig. 5*C*) in both MCF7 and HeLa cells. These siRNAs (I107as1, I107as2, and I107as3) were designed with chemical modifications (Stealth technology from Invitrogen) in such a way that the anti-sense strand acts as the guide strand (i.e., enters preferentially in the silencing complex). A significant promotion of E107 skipping was observed in HeLa cells with all three siRNAs (Fig. 5*F*, *Lower*), but not in MCF7 cells, where the inclusion/exclusion ratio is already too low to detect further decrease (Fig. 5F, *Upper*). These results show the complementary effect of AGO1 knockdown favoring exon recognition and targeting AGO1 to the endogenous enhancer location promoting exon skipping, demonstrating that AGO1 recruitment to intragenic enhancers is able to modify the alternative splicing process.

Discussion

Argonaute proteins have been profusely studied and characterized for their cytoplasmic roles as mediators of posttranscriptional gene silencing mediated by siRNAs and microRNAs (30–33). The existence of specific nuclear roles for these proteins in mammalian cells was demonstrated at the individual gene level for the mechanism of transcriptional gene silencing either acting at the promoter level (18, 34, 35) or intragenically in the regulation of alternative splicing (3, 10, 23). In particular, the fact that AGO1 was shown to interact with chromatin-embedded proteins such as chromatin modifiers, splicing factors (10), and RNA polymerase II (13) prompted us to define the DNA targets of AGO1 at a genome-wide level using ChIP-seq in human cells. During our ChIP-seq experiments using two different commercial antibodies to AGO1 we revealed that one of them, 6D8.2, supplied by Millipore, does not recognize AGO1 as its main target, but rather the splicing factor SF3b145, previously syndicated as the main partner in AGO1 immunoprecipitations (10, 22). Considering that SF3b145 and AGO1 bear undetectable sequence identity, that the sequence of the synthetic peptide claimed to be used to generate 6D8.2 (04-083; Millipore) is not present in the SF3b145 amino acid sequence, and that 6D8.2 immunofluorescence signal is exclusively nuclear, unlike bona fide anti-AGO1 antibodies such as 4B8, whose signal is both nuclear and cytoplasmic, we conclude that the use of 6D8.2 is a potential source of dramatic artifacts and indicates that some experiments in the current literature may have to be revised.

The first relevant conclusion of the ChIP-seq experiment is that AGO1 target sites do not bear a particular DNA consensus sequence. This suggests that AGO1 is not binding directly to DNA and, if binding is mediated by base pairing with DNA of an AGO1-associated nuclear RNA, this RNA is not a particular species defined by its sequence. Most surprisingly, we found that AGO1 distinctively binds to active transcriptional enhancers, this association being more prevalent at intragenic enhancers. This conclusion is reinforced by the finding that AGO1-bound enhancers have significantly higher density of long nuclear RNAs compared with enhancers without significant AGO1 signal, because eRNAs have been associated with enhancers that are active in regulating the transcription of distant genes (18, 36, 37). In parallel, we observed that AGO1 associates genome-wide with silencing chromatin marks such as H3K9me2 and H3K27me3. Consistently, this occurs mostly at nonactive enhancers.

A logical consequence of AGO1 binding to enhancers would be that AGO1 participates in the regulation of transcription. However, our AGO1 depletion experiments clearly indicate that there is no correlation between AGO1 genome localization and those genes whose mRNA levels are controlled by AGO1. On the contrary, and confirming previous evidence produced at the individual gene level (3, 10, 23), nuclear AGO1 regulates alternative splicing. Additionally, we show here that AGO1 also affects constitutive splicing, thereby likely playing a role in stimulating the excision of internal rather than first introns. However, alternative splicing of a number of exons is affected by the depletion of AGO1, and we found that 23% of the genes that modify their alternative splicing pattern upon AGO1 knockdown contain an AGO1 cluster at an intragenic enhancer, which is significantly higher than the 12% of the genes that do not modify their alternative splicing pattern upon AGO1 knockdown containing an AGO1 cluster at an intragenic enhancer. Additionally, there are no regulated alternative splicing events close to nonactive enhancers with AGO1 signal. Although we cannot rule out an indirect effect owing to changes in expression levels of constitutive splicing factors upon AGO1 knockdown, the fact that the effect is not global points at a direct mechanism. We provide further support for the AGO1-enhancer mediated control of alternative splicing by studying in detail the E107 alternative splicing event of the gene *SYNE2*, which contains an AGO1-bound active enhancer in intron 107. Indeed, the AGO1 cluster observed in MCF7 cells overlaps with features that are frequently used to define enhancers (i.e., conspicuous peaks of H3K27ac and H3K4me3) (Fig. 5 *A–C*). We showed that E107

inclusion is up-regulated upon AGO1 depletion and by treatments that increase RNAPII elongation (TSA and 5azadC). Conversely, E107 skipping is promoted by inhibition of RNAPII elongation with CPT (Fig. 5D). We have previously demonstrated that transfection with siRNAs targeting sequences in the respective downstream intron promotes higher inclusion of the alternative cassette exons number 33 of fibronectin (3) and number 18 of the neural cell adhesion molecule (23) through the AGO1-dependent TGS-AS mechanism (3). In these two exons slow elongation promotes higher inclusion. Because SYNE E107 behaves in an opposite way with respect to elongation, we predicted that transfection with siRNAs targeting I107 should cause E107 skipping. Experiments in Fig. 5 confirm this prediction and constitute an additional example for the TGS-AS mechanism, which could trigger either inclusion or skipping of alternative exons depending on the architecture of each particular event.

AGO1 was recently shown to interact with RNAPII at gene promoters, thereby regulating their transcription (13). It is known that RNAPII accumulates at enhancers (19), and it was proposed that intragenic enhancers could act as internal alternative promoters (38). One possibility to explain why AGO1 association to chromatin occurs mostly at active enhancers is that it is due to its interaction with RNAPII. Another nonexclusive possibility is that AGO1 associates with the RNA molecules produced at enhancers, possibly recruiting them to a nearby target, similarly to what was proposed recently for anti-sense RNAs (10), but specifically at enhancers. Additionally, AGO1 may be involved in the stabilization of the activity of these enhancers by eRNAs, as proposed before (18).

In view of the unforeseen nature of our findings, which reveal an association of an Argonaute protein with enhancers, it becomes crucial to rule out possible artifacts owing to the ChIP-seq technique. Indeed, there have been reports of pervasive artifacts in which highly expressed genes are prone to misleading ChIP localization of multiple unrelated proteins (39–42). These reports largely relate to yeast and mainly when grown under rich media in the absence of stress. More importantly, two additional reports in mouse ES cells (43), a closer system to our results, and in Drosophila (44) have described a similar phenomenon but actually provided evidence for the function of some of these regions as ES cell-specific enhancers. These two studies have shown that, at least for multicellular organisms, the presence of "hotspots" for transcription factors and other related proteins at certain specific genomic sites is not due to technical artifacts. On the contrary, in both cases they represent biologically relevant binding sites for transcriptional regulatory proteins with specific functions.

Nevertheless, we wish to summarize a series of observations and controls presented here that, in our view, greatly mitigate the possibility of an artifact. (i) The presence of AGO1 in the nucleus was confirmed by a method independent from ChIP-seq. *SI Appendix*, Fig. S2 (lanes 4–6) clearly shows Western blot evidence for the presence of AGO1 both in nuclear and cytoplasmic extracts. (ii) The whole characterization of the wrong antibody (Millipore 6D8.2, supposed to recognize AGO1, but recognizing SF3b) served as a nice negative control for the specific antibody to AGO1 (Sigma 4B8). ChIP-seq patterns for these two antibodies are completely distinct (Fig. 1 *A* and *B* and *SI Appendix*, Fig. S1 *A* and *B*), which reinforces the specificity of bona fide AGO1 signals. (iii) AGO1 peaks were not only detected through ChIP-seq but also through conventional ChIP followed by real-time PCR as the quantification method (Fig. 2D). (iv) The AGO1 signal depends on transcription by RNAPII. Fig. 2D shows that the signal completely disappears from two enhancers and one promoter when cells are treated with alpha-amanitin. If the signal were produced by highly expressed target sequences, one would expect to have higher signals on exons than on enhancers or promoters. Fig. 1B clearly shows that exons, exonic sequences (5′ and 3′ UTRs), and introns have lower observed signals than expected by random distribution (blue negative bars). On the contrary, signals are higher than expected by chance on enhancers and promoters. (v) Particular features of nonspecific ChIP enrichments (42) include association with highly expressed

genes (including ribosomal protein genes and noncoding RNAs) and with regions of low nucleosome densities. These features are not common to the AGO1 signals detected in our study. Indeed, the percentage of gene bodies [considered from 1 kb upstream of the transcription initiation site to 1 kb downstream of the poly(A) site] containing AGO1 clusters is similar for genes with high, medium, or low expression levels (18.5, 16.0, and 14.26%, respectively).

Taken together, our results strongly argue for specific and functional association of AGO1 with genomic loci correlating with transcriptional and posttranscriptional regulatory events.

We would like to propose that, apart from generating long poly(A)+ eRNAs, intragenic enhancers may regulate alternative splicing of neighboring exons by contributing to modulate RNAPII elongation intragenically, in connection with AGO1. Moreover, this association could happen in both directions, depending on the chromatin context of either active or nonactive enhancers.

In summary, we provide here genome-wide evidence for an association of AGO1 to active transcriptional enhancers through the RNAs transcribed from them in combination with a control of alternative splicing.

Methods

RNA and ChIP Sequencing. ChIP was performed as published (5). Library preparation and sequencing were carried out following standard protocols. More details about these experiments and the antibody characterization can be found in *SI Appendix*.

ChIP-Seq and Enhancer Analyses. Details about the processing of ChIP-seq reads and calculation of significant clusters can be found in *SI Appendix*. Random clusters were calculated by relocating each cluster in an arbitrary new position in the same chromosome, avoiding satellites, gaps, pericentromeric regions, and the overlap with any other random cluster. The co-occurrence of significant ChIP-seq clusters in specific regions was calculated with the block bootstrap and segmentation method (version 0.8.1) (encodestatistics.org/) with parameters -r 0.1 -n 10000, where r is the minimum overlapping fraction of each cluster to a region and n is the number of bootstrap samples used. The observed vs. expected overlap ratios and corresponding z-scores were calculated for each sample-region pair. Only cases with |z-score| > 3.3, corresponding to a two-tail *P* value < 0.001, were considered significant. The association of AGO1 and enhancers in genes was also analyzed using a 2 × 2 contingency table

	Genes with enhancer	Genes with no enhancer
Genes with AGO1	n11	n12
Genes with no AGO1	n21	n22

and calculating the mutual information

$$\text{mi} = \log_2\left(\frac{n_{11}}{(n_{11}+n_{12})(n_{11}+n_{21})}(n_{11}+n_{12}+n_{21}+n_{22})\right)$$

and performing a χ^2 test. For all enhancers: (n11, n12, n21, n22) = (3,084, 66, 16,890, 8,935), mi = 0.506, chi = 1,383.534, *P* = 2.2e-16. For active enhancers: (n11, n12, n21, n22) = (2,451, 699, 9,334, 16,491), mi = 0.935, chi = 2,018.21, *P* = 2.2e-16. For active enhancers restricted to lengths 400–2,600 bp: (n11, n12, n21, n22) = (1,898, 1,252, 6,744, 19,081), mi = 1.014, chi = 1,561.801, *P* < 2.2e-16.

Regions predicted by chromHMM (15) as candidate enhancers were labeled as active if they had significant enrichment (z-score >3) of H3K4me3 and H3K27ac ChIP-seq over input signal, calculated with Pyicos (45) using ENCODE ChIP-seq data for MCF7 cells and those nonenriched were defined as nonactive, yielding 2,622 active and 2,219 nonactive enhancers. Estrogen-dependent enhancers were defined as the regions occupied by an ERα peak and an H3K27ac peak whose centers lie at ≤1 kb (17). Selected enhancers were ranked according to coverage of GRO-seq reads independently for each replicate. Active and nonactive enhancers were defined according to whether they were in the top 25% or bottom 25%, respectively, in any of the two replicates, producing 157,957 active and 166,453 nonactive estrogen-dependent enhancers.

Expression and Splicing Analysis. Details about the processing of RNA-seq reads can be found in *SI Appendix*. Differential expression was calculated with DEGSEq (46) using the MATR method. Using as threshold a Benjamini–Hochberg corrected *P* value of 0.05 yielded 813 up-regulated and 461 down-regulated genes in the AGO1 knockdown compared with control cells.

Splice junctions were generated for potential exon cassettes (E1, E2, and E3) in the annotation. RNA-seq reads were mapped to the junctions (see *SI Appendix* for details) and the inclusion level (*I*) of the middle exon E2 was calculated as the fraction of reads that include the exon (n_{12}, n_{23}) over the total number of reads that include or skip the exon (n_{12}, n_{23}, n_{13}):

$$I = \frac{n_{12} + n_{23}}{n_{12} + n_{23} + 2n_{13}}.$$

The inclusion change between the AGO1 knockdown (siAGO1) and control cells (siLuc) was computed for each event as

$$M = \log_2\left(\frac{I_{siAGO1}}{I_{siLuc}}\right).$$

Significant inclusion changes were calculated with Pyicos (45) by comparing the *M* values between the two conditions to those between replicas as a function of the number of reads in the junctions:

$$A = \log_2\left((n_{12} + n_{23} + n_{13})|_{siAGO1} + (n_{12} + n_{23} + n_{13})|_{siLuc}\right).$$

Keeping only those events with a significant change in all comparisons (Benjamini–Hochberg corrected $P < 0.01$) resulted in 354 up- and 305 down-regulated events. The SES of an intron was defined as

$$SES = \frac{s}{s + u},$$

where *s* is the number of spliced reads defining the intron and *u* the number of reads over exon–intron boundaries. All reads that fall entirely inside any of the annotated exons were first discarded. Significant changes in the SES score were calculated analogously to the splicing analysis described above, using a Benjamini–Hochberg corrected $P < 0.05$.

Other Methods. See *SI Appendix* for additional discussion of methods.

1. Wang ET, et al. (2008) Alternative isoform regulation in human tissue transcriptomes. *Nature* 456(7221):470–476.
2. Cáceres JF, Kornblihtt AR (2002) Alternative splicing: Multiple control mechanisms and involvement in human disease. *Trends Genet* 18(4):186–193.
3. Alló M, et al. (2009) Control of alternative splicing through siRNA-mediated transcriptional gene silencing. *Nat Struct Mol Biol* 16(7):717–724.
4. Alló M, et al. (2010) Chromatin and alternative splicing. *Cold Spring Harb Symp Quant Biol* 75:103–111.
5. Bertucci PY, et al. (2013) Progesterone receptor induces bcl-x expression through intragenic binding sites favoring RNA polymerase II elongation. *Nucleic Acids Res* 41(12):6072–6086.
6. de Almeida SF, et al. (2011) Splicing enhances recruitment of methyltransferase HYPB/Setd2 and methylation of histone H3 Lys36. *Nat Struct Mol Biol* 18(9):977–983.
7. Luco RF, Alló M, Schor IE, Misteli T (2011) Epigenetics in alternative pre-mRNA splicing. *Cell* 144(1):16–26.
8. Luco RF, et al. (2010) Regulation of alternative splicing by histone modifications. *Science* 327(5968):996–1000.
9. Schor IE, Rascovan N, Pelisch F, Alló M, Kornblihtt AR (2009) Neuronal cell depolarization induces intragenic chromatin modifications affecting NCAM alternative splicing. *Proc Natl Acad Sci USA* 106(11):4325–4330.
10. Ameyar-Zazoua M, et al. (2012) Argonaute proteins couple chromatin silencing to alternative splicing. *Nat Struct Mol Biol* 19(10):998–1004.
11. Cernilogar FM, et al. (2011) Chromatin-associated RNA interference components contribute to transcriptional regulation in Drosophila. *Nature* 480(7377):391–395.
12. Taliaferro JM, et al. (2013) Two new and distinct roles for Drosophila Argonaute-2 in the nucleus: Alternative pre-mRNA splicing and transcriptional repression. *Genes Dev* 27(4):378–389.
13. Huang V, et al. (2013) Ago1 Interacts with RNA polymerase II and binds to the promoters of actively transcribed genes in human cancer cells. *PLoS Genet* 9(9):e1003821.
14. Nagaraja GM, et al. (2006) Gene expression signatures and biomarkers of noninvasive and invasive breast cancer cells: Comprehensive profiles by representational difference analysis, microarrays and proteomics. *Oncogene* 25(16):2328–2338.
15. Ernst J, et al. (2011) Mapping and analysis of chromatin state dynamics in nine human cell types. *Nature* 473(7345):43–49.
16. Hawkins PG, Santoso S, Adams C, Anest V, Morris KV (2009) Promoter targeted small RNAs induce long-term transcriptional gene silencing in human cells. *Nucleic Acids Res* 37(9):2984–2995.
17. Kim DH, Villeneuve LM, Morris KV, Rossi JJ (2006) Argonaute-1 directs siRNA-mediated transcriptional gene silencing in human cells. *Nat Struct Mol Biol* 13(9):793–797.
18. Li W, et al. (2013) Functional roles of enhancer RNAs for oestrogen-dependent transcriptional activation. *Nature* 498(7455):516–520.
19. Bonn S, et al. (2012) Tissue-specific analysis of chromatin state identifies temporal signatures of enhancer activity during embryonic development. *Nat Genet* 44(2):148–156.
20. Rada-Iglesias A, et al. (2011) A unique chromatin signature uncovers early developmental enhancers in humans. *Nature* 470(7333):279–283.
21. Creyghton MP, et al. (2010) Histone H3K27ac separates active from poised enhancers and predicts developmental state. *Proc Natl Acad Sci USA* 107(50):21931–21936.
22. Höck J, et al. (2007) Proteomic and functional analysis of Argonaute-containing mRNA-protein complexes in human cells. *EMBO Rep* 8(11):1052–1060.
23. Schor IE, Fiszbein A, Petrillo E, Kornblihtt AR (2013) Intragenic epigenetic changes modulate NCAM alternative splicing in neuronal differentiation. *EMBO J* 32(16):2264–2274.

24. Dujardin G, et al. (2014) How slow RNA polymerase II elongation favors alternative exon skipping. *Mol Cell* 54(4):683–690.
25. Boireau S, et al. (2007) The transcriptional cycle of HIV-1 in real-time and live cells. *J Cell Biol* 179(2):291–304.
26. Darzacq X, et al. (2007) In vivo dynamics of RNA polymerase II transcription. *Nat Struct Mol Biol* 14(9):796–806.
27. Ip JY, et al. (2011) Global impact of RNA polymerase II elongation inhibition on alternative splicing regulation. *Genome Res* 21(3):390–401.
28. de la Mata M, et al. (2003) A slow RNA polymerase II affects alternative splicing in vivo. *Mol Cell* 12(2):525–532.
29. Dutertre M, et al. (2010) Cotranscriptional exon skipping in the genotoxic stress response. *Nat Struct Mol Biol* 17(11):1358–1366.
30. Jacquier A (2009) The complex eukaryotic transcriptome: Unexpected pervasive transcription and novel small RNAs. *Nat Rev Genet* 10(12):833–844.
31. Huntzinger E, Izaurralde E (2011) Gene silencing by microRNAs: Contributions of translational repression and mRNA decay. *Nat Rev Genet* 12(2):99–110.
32. Castel SE, Martienssen RA (2013) RNA interference in the nucleus: Roles for small RNAs in transcription, epigenetics and beyond. *Nat Rev Genet* 14(2):100–112.
33. Bartel DP (2004) MicroRNAs: Genomics, biogenesis, mechanism, and function. *Cell* 116(2):281–297.
34. Morris KV, Chan SW, Jacobsen SE, Looney DJ (2004) Small interfering RNA-induced transcriptional gene silencing in human cells. *Science* 305(5688):1289–1292.
35. Suzuki K, et al. (2008) Closed chromatin architecture is induced by an RNA duplex targeting the HIV-1 promoter region. *J Biol Chem* 283(34):23353–23363.
36. Kim T-K, et al. (2010) Widespread transcription at neuronal activity-regulated enhancers. *Nature* 465(7295):182–187.
37. Lai F, et al. (2013) Activating RNAs associate with Mediator to enhance chromatin architecture and transcription. *Nature* 494(7438):497–501.
38. Kowalczyk MS, et al. (2012) Intragenic enhancers act as alternative promoters. *Mol Cell* 45(4):447–458.
39. Park D, Lee Y, Bhupindersingh G, Iyer VR (2013) Widespread misinterpretable ChIP-seq bias in yeast. *PLoS ONE* 8(12):e83506.
40. Teytelman L, Thurtle DM, Rine J, van Oudenaarden A (2013) Highly expressed loci are vulnerable to misleading ChIP localization of multiple unrelated proteins. *Proc Natl Acad Sci USA* 110(46):18602–18607.
41. Kasinathan S, Orsi GA, Zentner GE, Ahmad K, Henikoff S (2014) High-resolution mapping of transcription factor binding sites on native chromatin. *Nat Methods* 11(2):203–209.
42. Ward LD, Wang J, Bussemaker HJ (2014) Characterizing a collective and dynamic component of chromatin immunoprecipitation enrichment profiles in yeast. *BMC Genomics* 15:494–510.
43. Chen X, et al. (2008) Integration of external signaling pathways with the core transcriptional network in embryonic stem cells. *Cell* 133(6):1106–1117.
44. Moorman C, et al. (2006) Hotspots of transcription factor colocalization in the genome of Drosophila melanogaster. *Proc Natl Acad Sci USA* 103(32):12027–12032.
45. Althammer S, González-Vallinas J, Ballaré C, Beato M, Eyras E (2011) Pyicos: A versatile toolkit for the analysis of high-throughput sequencing data. *Bioinformatics* 27(24):3333–3340.
46. Wang L, Feng Z, Wang X, Wang X, Zhang X (2010) DEGseq: An R package for identifying differentially expressed genes from RNA-seq data. *Bioinformatics* 26(1):136–138.