


RESEARCH

Open Access



The nuclear receptor ER β engages AGO2 in regulation of gene transcription, RNA splicing and RISC loading

Roberta Tarallo^{1†}, Giorgio Giurato^{1,2†}, Giuseppina Bruno¹, Maria Ravo^{1,2}, Francesca Rizzo¹, Annamaria Salvati¹, Luca Ricciardi¹, Giovanna Marchese², Angela Cordella³, Teresa Rocco², Valerio Gigantino¹, Biancamaria Pierri¹, Giovanni Cimmino⁴, Luciano Milanesi⁵, Concetta Ambrosino^{6,7}, Tuula A. Nyman⁸, Giovanni Nassa^{1*} and Alessandro Weisz^{1*} 

Abstract

Background: The RNA-binding protein Argonaute 2 (AGO2) is a key effector of RNA-silencing pathways. It exerts a pivotal role in microRNA maturation and activity and can modulate chromatin remodeling, transcriptional gene regulation and RNA splicing. Estrogen receptor beta (ER β) is endowed with oncosuppressive activities, antagonizing hormone-induced carcinogenesis and inhibiting growth and oncogenic functions in luminal-like breast cancers (BCs), where its expression correlates with a better prognosis of the disease.

Results: Applying interaction proteomics coupled to mass spectrometry to characterize nuclear factors cooperating with ER β in gene regulation, we identify AGO2 as a novel partner of ER β in human BC cells. ER β -AGO2 association was confirmed in vitro and in vivo in both the nucleus and cytoplasm and is shown to be RNA-mediated. ChIP-Seq demonstrates AGO2 association with a large number of ER β binding sites, and total and nascent RNA-Seq in ER β + vs ER β - cells, and before and after AGO2 knock-down in ER β + cells, reveals a widespread involvement of this factor in ER β -mediated regulation of gene transcription rate and RNA splicing. Moreover, isolation and sequencing by RIP-Seq of ER β -associated long and small RNAs in the cytoplasm suggests involvement of the nuclear receptor in RISC loading, indicating that it may also be able to directly control mRNA translation efficiency and stability.

Conclusions: These results demonstrate that AGO2 can act as a pleiotropic functional partner of ER β , indicating that both factors are endowed with multiple roles in the control of key cellular functions.

Keywords: Argonaute 2, Estrogen receptor beta, Breast cancer, Interaction proteomics, Transcriptional regulation, RNA splicing

Background

The argonaute protein AGO2 is a RNA-binding protein primarily known for its functions in the cytoplasm, where it is a major component of the RNA-induced silencing complex (RISC). Indeed, this factor controls miRNA maturation and is involved in target recognition by small non-coding RNAs, thereby leading to mRNA degradation or translation inhibition in post-transcriptional gene silencing [1–3]. The

role of AGO2 in the composition of the miRNA machinery and the regulation of miRNA target stability and translation is well documented, among others, in breast cancer (BC) cells [4, 5]. On the other hand, AGO2 also acts in the nucleus, where it has been recently implicated in key events in several species, including mammals, such as transcriptional gene silencing (TGS) mediated by miRNAs [6–9], and it is involved in chromatin remodeling [10] and alternative RNA splicing [11] via RNA Pol II processivity slowdown and/or splicing factor recruitment [12]. Recent results demonstrated that this protein can shuttle between the cytoplasm and nucleus, and that its subcellular distribution is context-dependent [13]. Nucleocytoplasmic shuttling is a

* Correspondence: gnassa@unisa.it; aweisz@unisa.it

†Equal contributors

¹Laboratory of Molecular Medicine and Genomics, Department of Medicine, Surgery and Dentistry "Schola Medica Salernitana", University of Salerno, via S. Allende, 1, 84081 Baronissi, SA, Italy

Full list of author information is available at the end of the article

specific property also of estrogen receptor β (ER β) [14, 15], a member of the nuclear receptor superfamily of transcriptional regulators [16] that shows oncosuppressive activities in BC and other cancers. In BC, where AGO2 has been shown to be associated with tumor progression [17], ER β inhibits cancer cell proliferation and tumor growth and its expression has been found to correlate with a better prognosis of the disease [18]. Furthermore, ER β shows additive effects with anti-estrogens in promotion of apoptotic cell death and cell cycle inhibition [19, 20], and for this reason has been proposed as a marker of tumor responsiveness to endocrine therapy [21, 22]. Although this receptor can bind estrogenic compounds, thereby exerting a modulatory role on the functions of the oncogenic ER α , the other estrogen receptor subtype active in cancer cells, by dimerizing with it and thereby modifying its activity on target genes [16, 23], in the absence of ligand it exhibits significant effects in BC cells [24], including, among others, miRNA-mediated post-transcriptional regulation of the BC cell proteome [25]. Physiologically, the presence of unliganded ER β is a typical condition during specific phases of the menstrual cycle, before puberty, and in post-menopausal women, when this receptor might compensate for the absence of circulating hormones with regard to cell functions.

We show here that expression of unliganded ER β in luminal-like BC MCF-7 cells induces profound effects on the cell transcriptome, represented by changes in both RNA expression and splicing. To elucidate the molecular bases of these actions, we applied interaction proteomics coupled to mass spectrometry (MS) to identify ER β -interacting proteins in BC cell nuclei. AGO2 was among 277 new molecular partners of the receptor identified using this approach. Interestingly, a comparison between this protein dataset and datasets related to AGO2-interacting proteins present in public databases revealed a number of molecular partners in common between the two factors, indicating that they share a sizeable amount of functions in the nucleus, comprising also RNA processing and splicing. Based on these results, we investigated in depth the functional significance of ER β -AGO2 interaction, identifying a dual role of the association between AGO2 and ER β in BC cells in the nucleus and the cytoplasm, for quantitative and qualitative regulation of gene expression at both the transcriptional and post-transcriptional level.

Results

In vivo binding of ER β to the luminal-like BC cell genome and effects on gene expression

ER β is an estrogen receptor that, like many other members of the nuclear receptor superfamily of transcription factors but contrary to ER α , in the absence

of ligands can be found predominantly in the nucleus exerting profound effects on the cell, comprising oncosuppressor activities in BC and other cancer cells [16]. Furthermore, in the absence of ligand ER β does not dimerize with ER α and therefore induces specific effects that are independent of interfering with the activity of the latter [16, 25]. Human BC cell lines expressing endogenous ER β protein to detectable levels are not available, probably due to epigenetic inhibition of this gene promoter by DNA methylation in cancer cells [26], and for this reason activity and functions of this receptor in BC cells can be studied only by exogenous transfer of ER β expression vectors. We previously showed that stable expression of ER β fused to a TAP tag at either the C-terminus (Ct-ER β) or N-terminus (Nt-ER β), suitable for proteomics analyses, causes growth inhibition and re-programming of miRNA expression and the cell proteome in human luminal-like MCF-7 BC cells [25], in line with results obtained in other laboratories [27–29]. These cell lines express ER β to levels comparable to those of endogenous ER α [25], therefore reproducing as much as possible a physiological setting. To investigate the molecular bases of these actions of unliganded ER β , Nt-ER β , Ct-ER β and, as control, Ct-ER α (ER β -) MCF7 cell clones were cultured in steroid-free medium and gene expression profiling was carried out by total RNA extraction and RNA-Seq as described in “Methods”. Reads were aligned to the reference human genome and normalized to “fragments per kilobases of exon per million mapped reads” (FPKM) and genes differentially expressed in the presence of ER β were determined for both clones with respect to TAP-ER α -expressing cells (Ct-ER α , Fig. 1). By setting 0.5 FPKM as an expression level threshold, we identified 15,470 and 16,115 genes differentially expressed in Ct-ER β and Nt-ER β , respectively (Additional file 1: Table S1a, b). Data from these two datasets were compared and 10,399 and 8828 transcripts showed statistically significant (adjusted p value ≤ 0.05 , fold change (FC) $|1.2|$) differences in expression in Ct-ER β and Nt-ER β , respectively (comparing ER β + vs ER β - cells). Among these RNAs, 6739 (3246 upregulated and 3493 downregulated), representing about 65 and 76% of differentially expressed transcripts in Ct-ER β and Nt-ER β , respectively, displayed an identical trend in both cell clones (Fig. 1a; Additional file 1: Table S1c, d). Evaluation of the functional significance of the gene expression changes detected in ER β -expressing cells, performed by IPA comparative analysis, revealed that all the top ten functional annotations identified relate to key cancer cell characteristics, including regulation of cellular movement, cell-to-cell signaling and interactions, cell morphology, growth and proliferation,

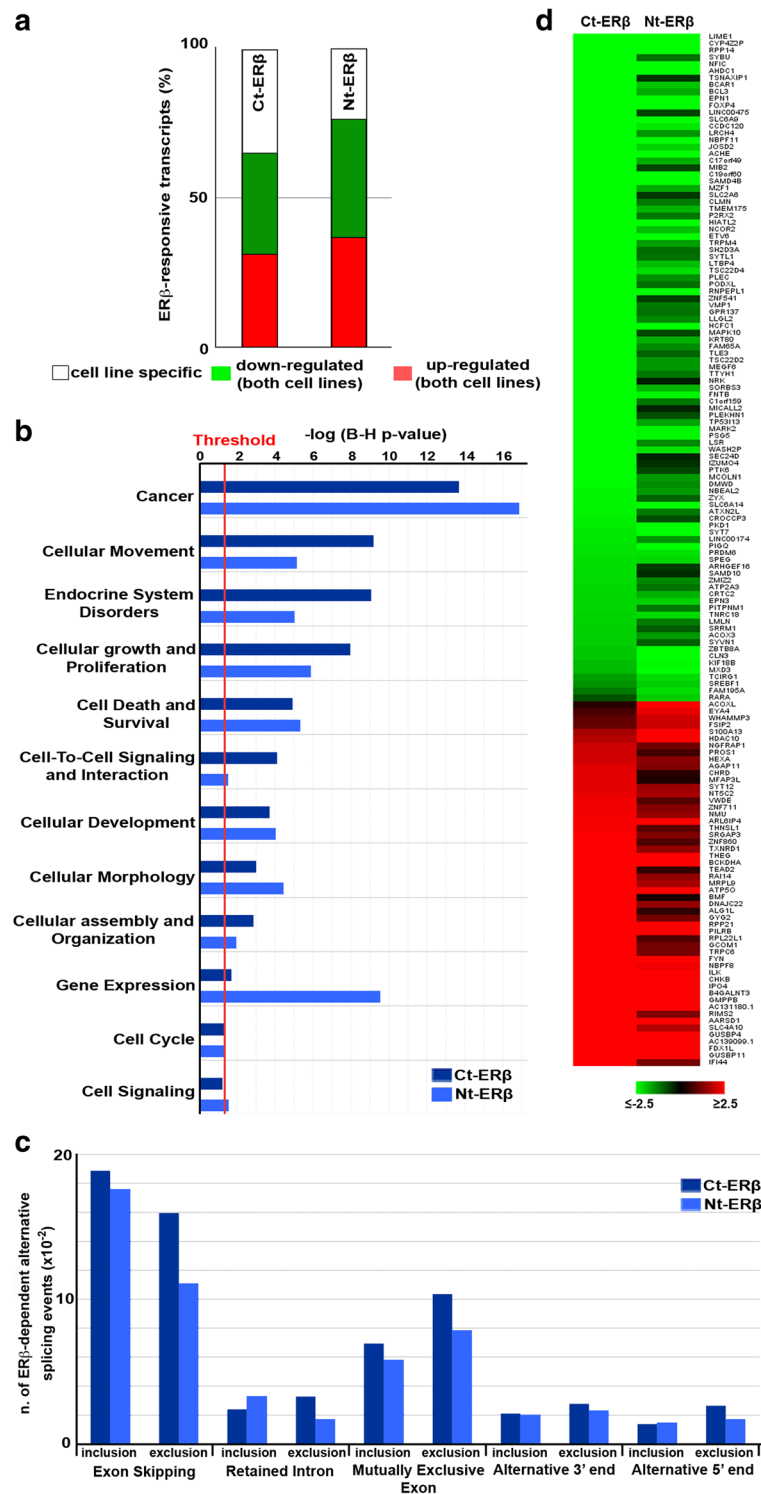


Fig. 1 Effects of unliganded ERβ on the BC cell transcriptome and alternative RNA splicing. **a** The fraction of differentially expressed genes detected in both Ct-ERβ- and Nt-ERβ-expressing cells (red/green) or in only one of the two cell lines (white). **b** Functional annotation by Ingenuity Comparative Analysis on genes differentially expressed in BC cells expressing either Ct-ERβ or Nt-ERβ (fold change cut-off [1.5], false discovery rate (FDR) ≤ 0.05). The red line indicates the p value threshold. **c** Alternative splicing events occurring in the two ERβ-expressing cell lines. Inclusion and exclusion behaviors for each event are shown (FDR ≤ 0.05; inclusion/exclusion cut-off [0.1]). **d** Heatmap showing differentially expressed transcripts subject to alternative splicing events in both ERβ-expressing cell lines

cell cycle or cell death, and survival (Fig. 1b). The fact that all these functions are known to be influenced by ER β in multiple cell types, and that they were similarly affected in both ER β -expressing cell lines, confirms previous observations that the TAP-tag does not significantly influence the receptor activity in vivo [23, 25]. As estrogen-bound ER β has been shown to induce alternative splicing events in this BC cell subtype [30], the effects of unliganded receptor on RNA splicing were also assessed with MATS (Multivariate Analysis of Transcripts Splicing) [31]. Around 900 splicing events were found to be commonly affected in Ct-ER β and Nt-ER β with respect to Ct-ER α cells, considering exon skipping, intron retention, mutually exclusive exons, and alternative 3' and 5' end events. The two clones showed the same splicing patterns, exon skipping being, as expected, the most frequent event, and a comparable percentage of transcripts affected (Fig. 1c; Additional file 2: Table S2a; Additional file 3: Table S2B; Additional file 4: Table S2c). By comparing receptor-mediated differential RNA expression with splicing, it emerged that the 150 ER β -modulated transcripts shown in Fig. 1d also underwent alternative splicing in both cell clones.

To date, the major effects of hormone-bound ER β in BC cells have been shown to result from its binding to the genome. As ligand-free receptor can bind DNA in vitro and is present in the cell nucleus, we mapped its binding to chromatin in vivo by ChIP-Seq, performed in triplicate as described earlier [23, 32]. Triplicate chromatin samples from Ct-ER β and, as negative control, wild-type MCF-7 cells were pulled down with IgG-Sepharose, which binds with high affinity the TAP moiety of the fusion protein (see "Methods"). Purified DNA was used to prepare ChIP-Seq libraries for short-read sequencing on NextSeq 500. Reads obtained were aligned to the human genome sequence and peaks enriched in ER β + libraries with respect to the negative control were identified using MACS2 coupled to MuSERA [33, 34], as described in "Methods". Input DNA was also sequenced as an additional control. This allowed identification of 37,304 ER β -binding sites (Fig. 2a, right). Not surprisingly, a transcription factor binding sequence motif search in the ER β -binding 'peak' sequences revealed a matrix corresponding to the estrogen receptor binding motif ERE (estrogen response element, GGTCAnnnTGACC), indicating that unliganded ER β also binds this element in BC cell chromatin. Indeed, ESR1 and ESR2 binding motifs, together with a number of others, including in particular TP53, TP63, and PPAR γ , were among the transcription factor binding matrixes showing statistically significant enrichment in ER β binding sites (Fig. 2a, left central panel). This last result is particularly interesting when considering that ER β can bind directly to TP53 and TP63 in BC cells and that, in the

presence of mutant TP53, it interferes with its activity on target genes, resulting in inhibition of epithelial-to-mesenchymal transition and cell invasiveness [35]. A full list of these binding sites is available, with relevant information, in Additional file 5: Table S3a. Detailed analysis revealed that 4% of the binding sites identified were located within promoter regions, calculated considering -1000 and +100 bases from the main annotated transcription start site (TSS) according to HOMER guidelines, including 1% in close proximity of TSSs. On the other hand, most of the binding sites were found in intronic (47%) or intergenic (41%) regions (Fig. 2b). Statistical analysis of ER β binding site occurrence within different genomic regions (3' UTR, 5' UTR, intergenic, promoter, etc.), performed with Genome Association Tester (GAT), showed that ER β binds prevalently within 5' UTR and promoter regions (FC > 2 and q-value \leq 0.05) (Fig. 2c), as suggested also by the prevalence of ER β binding around TSSs (Figs. 2a, lower left panel). Alignment of ER β -responsive transcription units (FC cutoff of |1.5|) with the receptor binding sites revealed that, in the absence of ligand, ER β is able to bind the promoter regions of 426 of the genes it regulates, indicating that these are most likely to represent its primary targets (Additional file 5: Table S3b). Functional analysis of these genes by IPA showed that several among them are involved in cellular functions related to known ER β actions in BC and other cancers, in particular cell growth and proliferation, death and survival, and cell cycle (Additional file 6: Figure S1a). Indeed, network representation of cell cycle genes shows that the overall effect of ER β -responsive genes is directed towards inhibition of cell cycle progression (Additional file 6: Figure S1b).

When considering, instead, ER β -responsive genes comprising one or more receptor binding sites within the whole transcription unit, the number increases to 1752, including 476 (27%) whose RNA transcripts undergo alternative splicing in the presence of ER β with patterns identical to those detectable on the whole transcriptome (Additional file 6: Figure S2). Since transcription factors, including nuclear receptors and, in particular, estrogen receptors [36–39], are known to regulate their target genes also through long-range chromatin looping interactions, the high number of ER β binding sites mapped here suggests that a much higher number of these ER β -dependent gene responses and splicing events are likely due to a direct effect of the receptor.

Mapping the nuclear interactome of unliganded ER β identified AGO2 as a novel molecular partner of the receptor in BC cells

To search for leads allowing identification of the molecular mechanisms underlying the ligand-independent activity of ER β on BC cell functions, we applied interaction proteomics [40–43] to map the nuclear

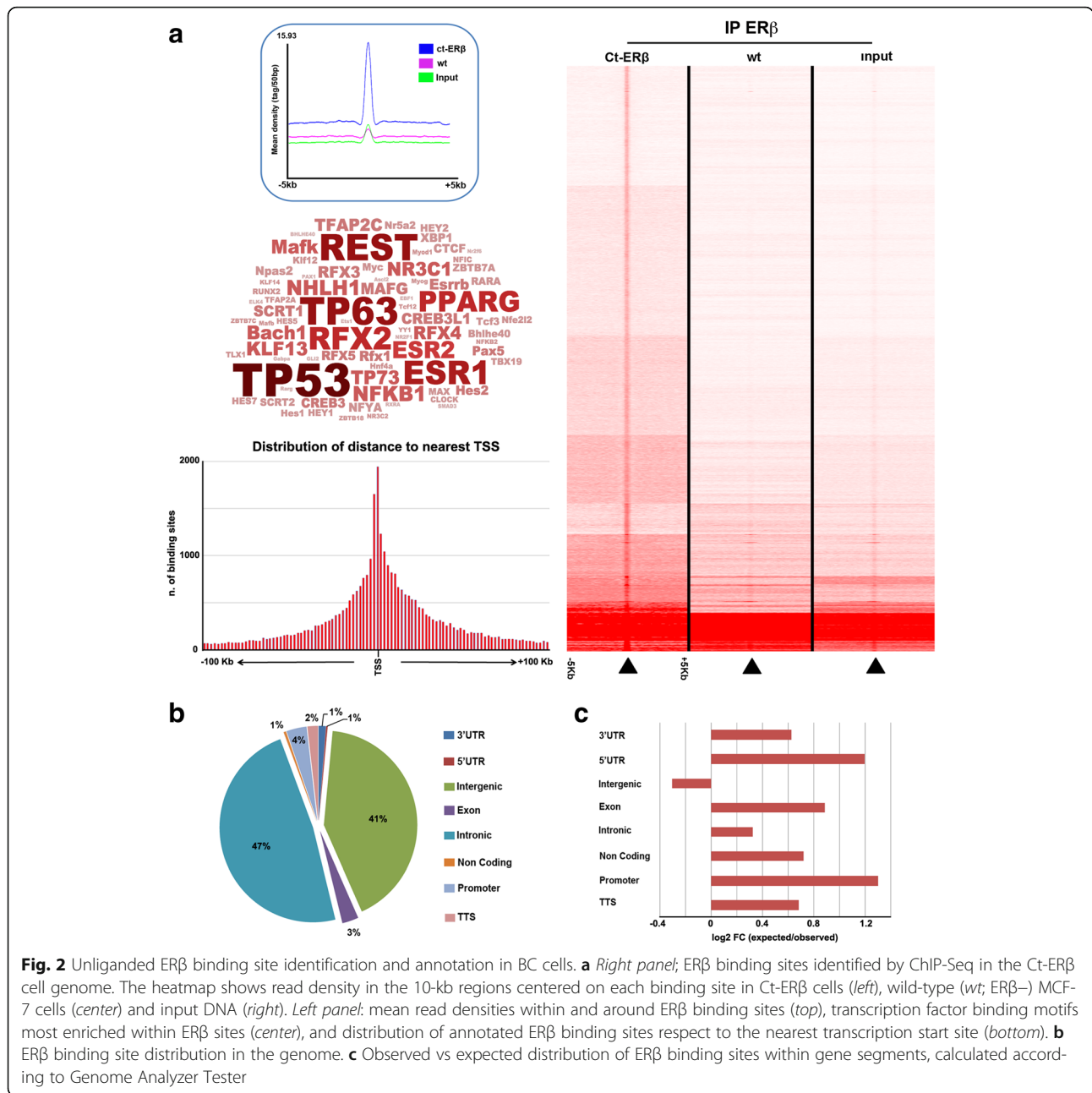


Fig. 2 Unliganded ERβ binding site identification and annotation in BC cells. **a** *Right panel*; ERβ binding sites identified by ChIP-Seq in the Ct-ERβ cell genome. The heatmap shows read density in the 10-kb regions centered on each binding site in Ct-ERβ cells (*left*), wild-type (wt; ERβ⁻) MCF-7 cells (*center*) and input DNA (*right*). *Left panel*: mean read densities within and around ERβ binding sites (*top*), transcription factor binding motifs most enriched within ERβ sites (*center*), and distribution of annotated ERβ binding sites respect to the nearest transcription start site (*bottom*). **b** ERβ binding site distribution in the genome. **c** Observed vs expected distribution of ERβ binding sites within gene segments, calculated according to Genome Analyzer Tester

interactome of the receptor. To this end, nuclear protein extracts from TAP-ERβ-expressing cells were subjected to tandem affinity purification coupled to mass spectrometry for isolation and identification of native protein complexes, as summarized in Fig. 3a. Nuclear extracts from wild-type MCF-7 cells, maintained under the same culture conditions, were processed in parallel as control. As shown in Additional file 6: Figure S3a, b, this procedure led to efficient isolation of the bait protein from Ct-ERβ cell nuclei. Analysis by nano-LC MS/MS of the purified protein mixtures led to unequivocal and robust identification of 277 specific ERβ-interacting proteins (Additional file 7: Table S4), as all those

identified also in ERβ – samples were discarded. Functional annotation analysis by IPA revealed that the ERβ partners identified are involved in multiple molecular functions, either relevant in cancer and/or reflecting known activities of this nuclear receptor, such as RNA post-transcriptional regulation, cell growth and proliferation, cell cycle, DNA replication, gene expression, and cell death and survival (Fig. 3b). Further dissection of the most enriched molecular function related to RNA post-transcriptional modification revealed an involvement of the ERβ interactome in RNA processing, including in particular splicing, alternative splicing, cleavage, polyadenylation, and

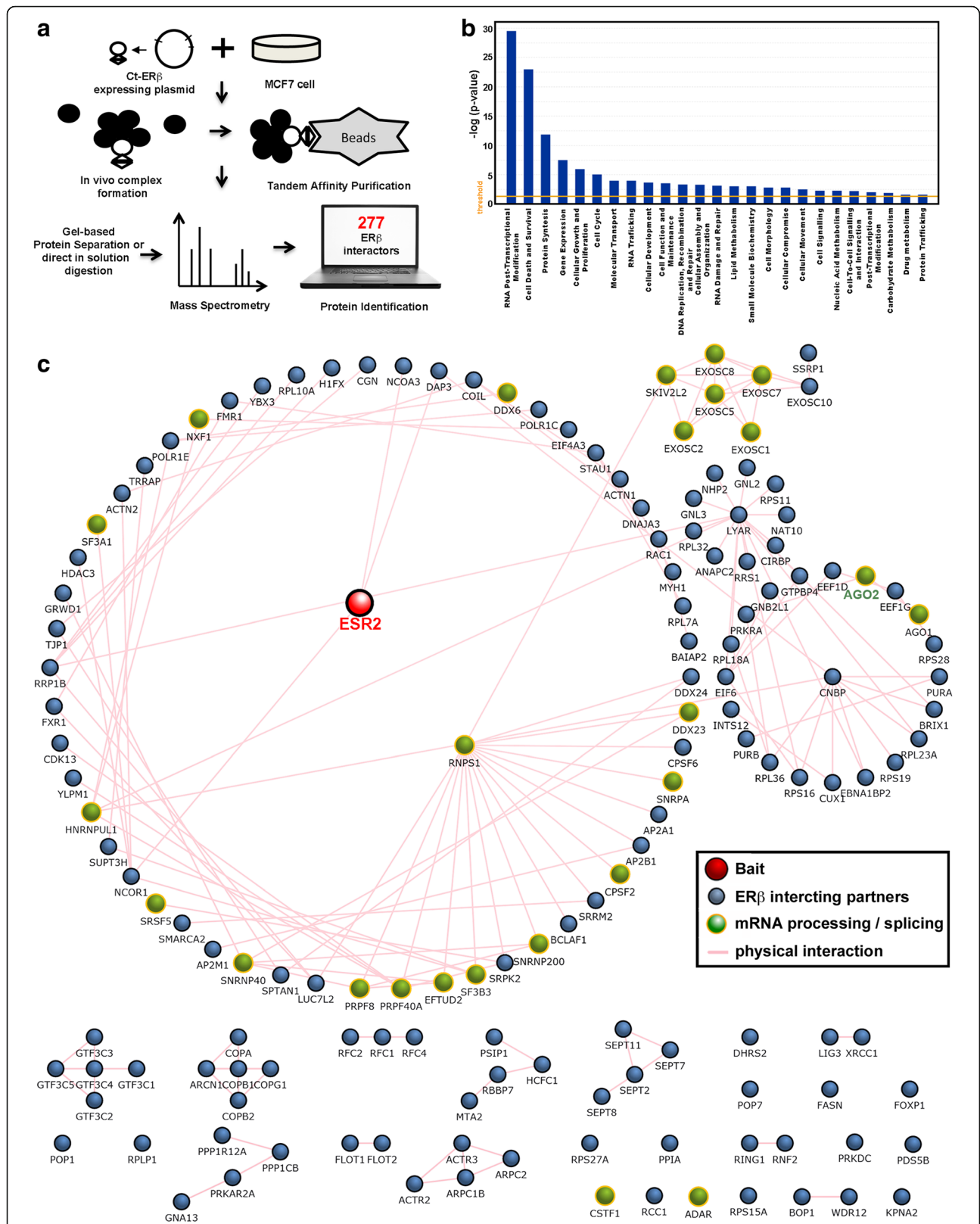


Fig. 3 ER β -binding proteins identified by interaction proteomics. **a** The experimental workflow. **b** Functional enrichment analysis of ER β -associated proteins performed by IPA representation. **c** Interaction network showing known associations reported in protein–protein interaction databases within the identified ER β (ESR2) interactome. Proteins involved in RNA processing and splicing are represented in green

stabilization (Additional file 6: Figure S3c). In Fig. 3c the ER β -interacting proteins identified are shown in an interaction network, with those specifically enrolled in mRNA processing and splicing highlighted (green). Among these, we noticed AGO2, since this factor is a major component of RISC that binds sncRNAs to guide post-transcriptional gene silencing within the cell cytoplasm [44, 45], a process shown to be controlled also by unliganded ER β in BC cells [25]. More recently, AGO2 has been shown to act in the cell nucleus, directly controlling chromatin silencing, transcriptional repression, and pre-mRNA splicing [11, 12], all functions exerted also by ER β and directly related to oncosuppression and regulation of gene transcription. We thus decided to focus our attention on the ER β -AGO2 interaction and, firstly, we compared the set of ER β -associated proteins identified here with that of known AGO2 interactors from protein-protein interaction databases [46, 47] and the literature, identifying 41 partners in common between ER β and AGO2 (Additional file 7: Table S4).

These include, together with several proteins involved in transcription, RNA splicing and maturation and pleiotropic factors controlling key cancer-related cellular functions (see “Discussion” for details), and AGO1, another argonaute protein functionally redundant with AGO2 in the miRNA pathway and known to interact with it. These data suggest that association of AGO2 with ER β in multiprotein nuclear complexes could represent a central hub for regulation of BC cell functions by the nuclear receptors. Co-immunoprecipitation was thus performed in different experimental conditions to confirm the association between the two proteins (Fig. 4a, b). First, Ct-ER β cells were transiently transfected with an expression vector encoding Myc-tagged AGO2 to prevent possible artifacts due to non-specific recognition of endogenous proteins by the anti-AGO2 antibodies. Wild-type MCF-7 cells transfected with a Myc-AGO2 expression vector and untransfected Ct-ER β cells were used as negative controls. ER β pull-down with IgG showed copurification of exogenous AGO2, detected specifically by an anti-Myc tag antibody (Fig. 4a). Secondly, MCF-7 cells expressing a Tet-inducible Myc-Flag-ER β (Tet-On system) were used to exclude the possibility that the interaction could be due to the TAP-tag. Results, reported in Fig. 4b, showed that ER β -AGO2 association is independent of the nature of the tag fused to the receptor. This experimental setting allowed us to confirm also the TAP results relative to ER β association with FXR1 (Fragile X mental retardation syndrome-related protein 1) and EIF6 (Eukaryotic translation initiation factor 6), and with the splicing factor PRPF8 (Pre-mRNA-processing-splicing factor 8; Fig. 4b), all known partners of AGO2. Immunoprecipitation confirmed AGO1 association with ER β , which

was lost upon AGO2 knock-down (*kd*) with shRNA, indicating that it is likely to be mediated by this last (Fig. 4c). To evaluate the role of AGO2 in nuclear ER β interactome composition, TAP/MS analysis was also repeated before and after *in vivo* silencing of AGO2 by shRNA. Results were analyzed by comparing the label-free quantification (LFQ) value of ER β -interacting proteins before and after AGO2 silencing, following normalization with respect to ER β LFQ value, as described in the “Methods” section (Additional file 8: Table S5a, b). This analysis, while providing a confirmation of the identified ER β interactome, revealed that AGO2 is required for the stable association with ER β of a large fraction (70%) of its interacting partners (heatmap in Fig. 4c; Additional file 6: Figure S4; Additional file 8: Table S5c, d), comprising 31 of the 41 AGO2 interactors known to date (75%), a result suggesting that the argonaute protein is likely to play a central role in assembly and/or stability of the nuclear ER β interactome.

ER β -AGO2 association in BC cells was analyzed *in vivo* with two experimental approaches. First, Tet-On MCF-7 cells stably transfected with a tet-inducible vector encoding Myc-Flag-ER β were analyzed by immunofluorescence microscopy with specific antibodies directed against the endogenous AGO2 or the myc epitope of the tagged ER β protein. As shown in Additional file 6: Figure S5a (+DOXY panels), both AGO2 and ER β are detected in the nucleus and the cytoplasm, where they are more abundant, in accordance with the expected intracellular distribution of the two proteins under these conditions, and are colocalized to a significant extent in both compartments. Interestingly, the intracellular distribution of AGO2 is not affected significantly by ER β induction (compare +DOXY with -DOXY panels in Additional file 6: Figure S5a). Then, *in vivo* association of the two proteins was measured by proximity ligation assay (PLA), a method that allows *in situ* detection of the association between two proteins when these are in close proximity (40–50 nm). This test used cells transiently expressing Myc-AGO2 and Flag-ER β . The results, reported in Additional file 6: Figure S5b, confirm *in vivo* association of the two proteins in both the nucleus and the cytoplasm. Under the same experimental conditions the oncogenic ER subtype ER α does not show association with AGO2, indicating that the ability to bind AGO2 is a specific property of ER β . Interestingly, ER β and AGO2 were clearly associated not only in the nucleus but also in the cytoplasmic compartment of the cell. AGO2-ER β interaction was further verified by co-immunoprecipitation following transient expression of Myc-tagged AGO2 in Ct-ER β cells and, as control, wild-type MCF-7 cells. TAP-ER β pull-down from cytoplasmic extracts resulted in co-purification of

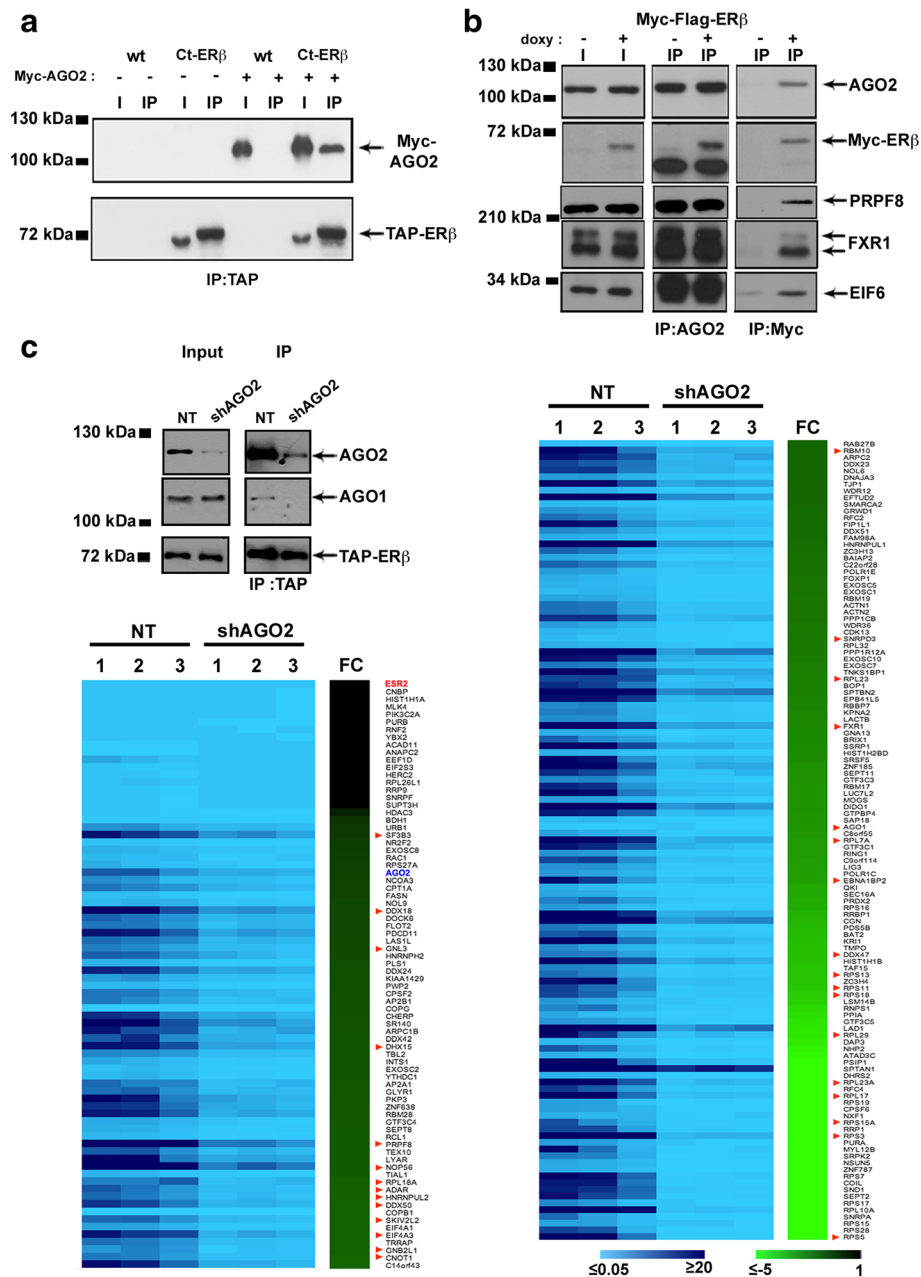
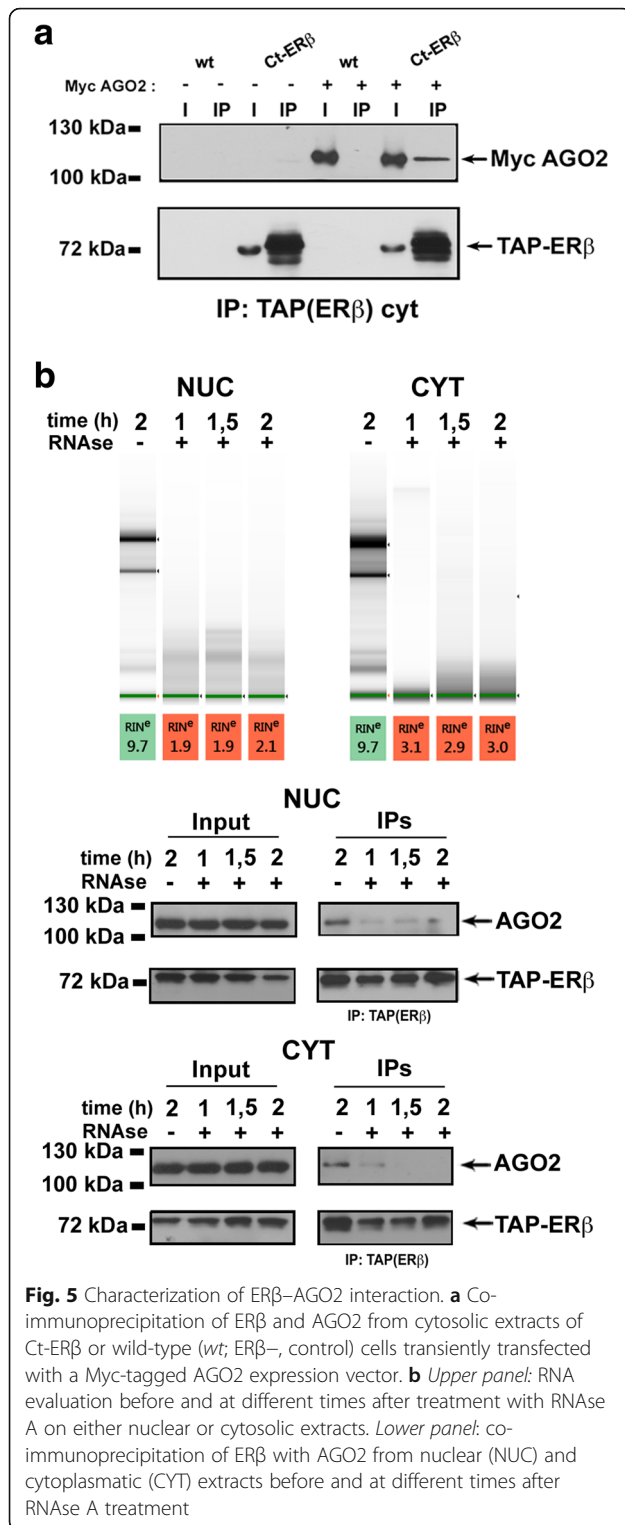


Fig. 4 Validation of ERβ-AGO2 interaction in the nucleus. **a** Co-immunoprecipitation of ERβ and AGO2 from nuclear extracts of Ct-ERβ or wild-type (wt; ERβ⁻, control) cells transiently transfected with a Myc-tagged AGO2 expression vector. **b** Co-immunoprecipitation of ERβ with AGO2, PRPF8, FXR1, and EIF6 from nuclear extracts of a Tet-inducible MCF-7 cell clone expressing Myc-Flag-ERβ; doxy doxycycline. **c** *Top left*: western blots of AGO2 and AGO1 co-immunoprecipitation with ERβ from nuclear extracts of Ct-ERβ cells before and after AGO2 silencing. *Heatmap*: amount (expressed as LFQ value normalized with respect to ERβ LFQ value) of proteins co-immunoprecipitated with ERβ in control (NT) and AGO2 ‘knock-down’ (shAGO2) cells measured in three biological replicates. *FC* average fold-change in shAGO2 vs control samples (NT). Only statistically significant changes in protein content are reported. ESR2 (bait) and AGO2 are highlighted in red and blue, respectively. Red arrowheads mark known AGO2 interactors from protein-protein interaction databases [46, 47] and the literature [12]

AGO2, detectable here by anti-Myc tag antibodies (Fig. 5a), a result further confirming association of the two proteins also in the extranuclear compartment of the cell. This indicates that either there are different complexes in the two subcellular compartments and/or that at least

some of these complexes undergo nucleo-cytoplasmic shuttling, an interesting possibility given the known ability of ERβ to redistribute within the cell in response to specific stimuli [15, 16], a property recently attributed also to AGO2 [13].



Upon agonist ligand binding, ERβ can dimerize with ERα in the nucleus [23], a condition that could affect AGO2 binding. However, AGO2 is found associated with ERβ also in the presence of 17β-estradiol (E2; Additional file 6: Figure S6a), while under the same condition it did not bind

TAP-tagged ERα (Ct-ERα; Additional file 6: Figure S6b), demonstrating a estrogen receptor subtype-specific ability to associate with AGO2. This result indicates a selective role of the complex(es) comprising the two proteins in ERβ-specific functions in BC cells.

Considering the role exerted by AGO2 in the ERβ interactome, the nature of its interaction with the receptor, i.e., primary or mediated by additional factor(s), was further evaluated. A yeast two-hybrid assay performed using AGO2 fused to the LexA DNA binding domain as ‘bait’ and ERβ fused to the Gal4 activation domain as ‘prey’ failed to demonstrate direct association between the two proteins (data not shown), suggesting that other molecules could be involved in the interaction. For this reason, starting from the assumption that AGO2 is an RNA-binding protein, we investigated whether RNA could represent the bridging factor between the two proteins. To evaluate this possibility, cytoplasmic and nuclear protein extracts from Ct-ERβ cells were treated with RNAse A for different times before ERβ pull-down by IgG binding, followed by immunodetection of the two proteins in the immunoprecipitates, as described by Höck and colleagues [48]. Results showed a strong reduction of ERβ and AGO2 association already 1 h after RNAse treatment, indicating that association between the two factors is indirect and likely to be mediated by one or more RNAs in both the nucleus and cytoplasm (Fig. 5b).

AGO2 binding to the BC cell genome in proximity of ERβ

As AGO2 has been shown to be able to bind chromatin in human cancer cells [49], ChIP-Seq was performed with anti-AGO2 antibodies in Ct-ERβ and, as control, in wild-type MCF-7 cells to investigate the possibility that this protein also binds to BC cell chromatin, alone and in combination with ERβ, where a complex between the two factors could exert specific actions. Results showed the ability of AGO2 to interact with the BC cell genome and how this is greatly influenced by the presence of ERβ. As displayed in Fig. 6a and Additional file 6: Figure S7a, b, ChIP-Seq led to the identification of 3441 and 2552 AGO2 binding sites in ERβ + and ERβ – cells, respectively (Additional file 9: Tables S6a, b); these can be grouped into three clusters, representing, respectively, regions comprising AGO2 binding sites prevalent in Ct-ERβ (blue) or wild-type (red) cells, or similar in both cell lines (grey). The density plots reported in Fig. 6a (boxes to the right) show the signal density profile of the three clusters, highlighting intensity and prevalence of the three sets of AGO2 binding sites detected. Like previously performed for ERβ binding sites, analysis of the annotated AGO2 binding sites prevalent in ERβ + and ERβ – cells (blue and red clusters, respectively) according to their location with

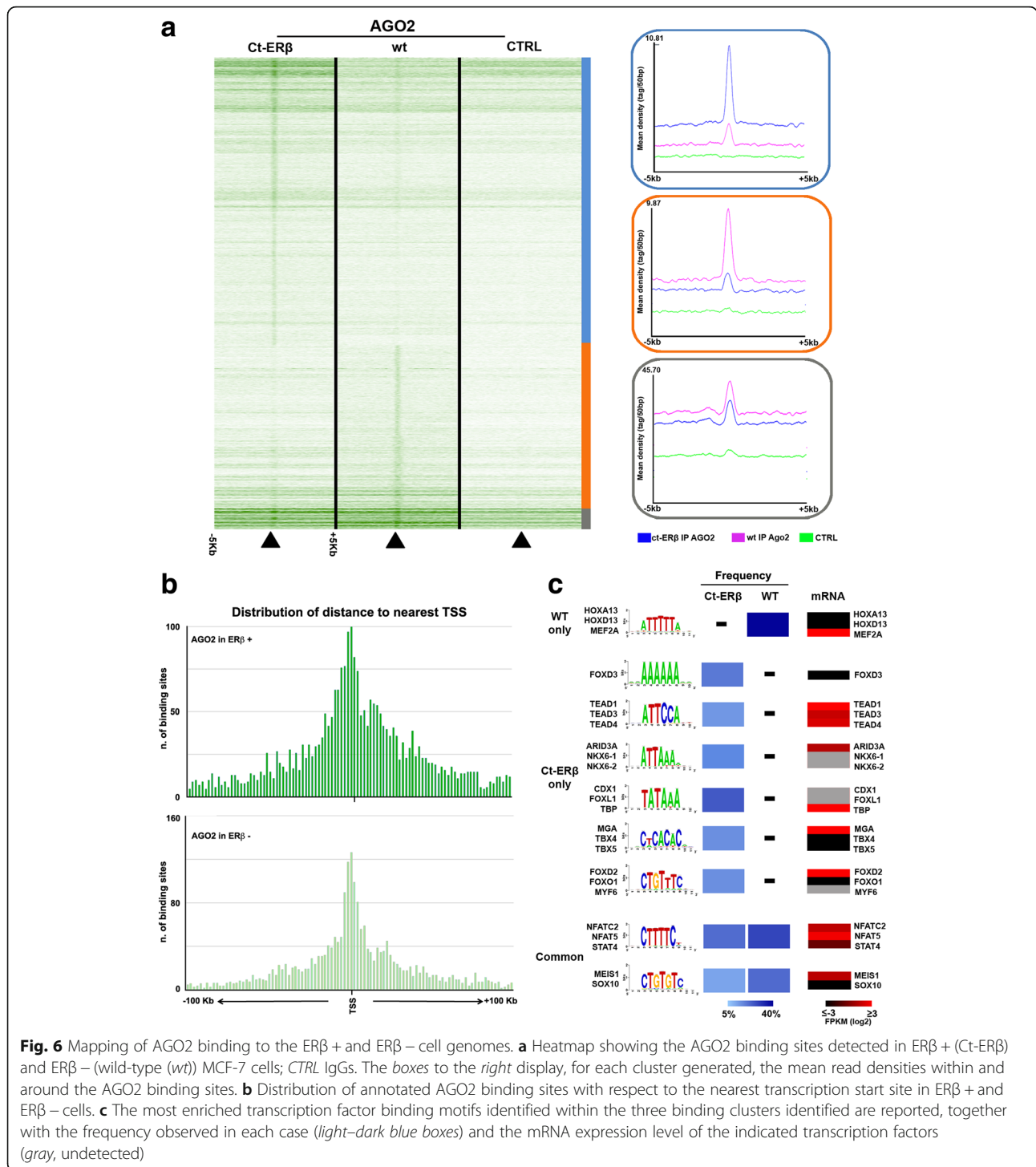


Fig. 6 Mapping of AGO2 binding to the ERβ+ and ERβ- cell genomes. **a** Heatmap showing the AGO2 binding sites detected in ERβ+ (Ct-ERβ) and ERβ- (wild-type (wt)) MCF-7 cells; CTRL IgGs. The boxes to the right display, for each cluster generated, the mean read densities within and around the AGO2 binding sites. **b** Distribution of annotated AGO2 binding sites with respect to the nearest transcription start site in ERβ+ and ERβ- cells. **c** The most enriched transcription factor binding motifs identified within the three binding clusters identified are reported, together with the frequency observed in each case (light–dark blue boxes) and the mRNA expression level of the indicated transcription factors (gray, undetected)

respect to the nearest transcription start site indicates that, in both cases, most of these are positioned in and around gene promoter regions (Fig. 6b). Since AGO2 is not a DNA binding protein, its association with chromatin is likely to be mediated by other factors, which appear to be different in ERβ+ compared to ERβ- cells. To investigate the basis of this difference we performed sequence

analysis of the binding sites within the three AGO2 binding site clusters, searching for enriched binding motifs for other transcription factors that might play a role in mediating AGO2 occupancy of chromatin using a stringent statistical threshold. This revealed that AGO2 binding sites in ERβ+ and ERβ- cells comprise binding motifs for known transcription

factors, some of which are significantly enriched in only one of the two cellular conditions investigated (Fig. 6c; Additional file 10: Table S7a, b). The binding matrixes for homeobox factors (HOXA13 and HOXD13) and MEF2A were overrepresented only within the AGO2 binding sites prevalent in ER β -cells. Interestingly, the mRNAs encoding these three factors are all detectable in MCF-7 cells, with that for MEF2A being expressed at higher levels, indicating that the corresponding genes are active in this cell type. It has been shown that MEF2A controls proliferation of mammary epithelial cells, where its up-regulation coincides with HDAC7 downregulation and promotion of cell cycle exit, and its transcription is repressed in transformed cells with altered morphogenesis [50]. Within the AGO2 binding sites detected in ER β + cells (Ct-ER β only) different binding matrixes were specifically enriched, with several of the corresponding binding factor mRNAs being expressed in the cell, including those for TEAD1-3, known to be involved in tumorigenesis [51], and ARID3A, reported to exert gene regulation activity in BC and other cancers [52]. This result suggests that factors interacting with these genetic elements could influence AGO2 and ER β homing and/or activity in such genomic sites. On the other hand, binding matrixes enriched in AGO2 sites in both cell lines correspond to consensus motifs for NFATC2 and 5 and STAT4, all involved in BC development and metastasis [53–55], and for MEIS1 and SOX10, the first often found dysregulated in BC and the second reported to control stem and mesenchymal cell status in epithelial cells of the mammary gland [56, 57].

By comparing mapping of AGO2 binding in Ct-ER β cells to that of unliganded ER β described in Fig. 2a, results showed that, in a large number of cases, AGO2 binding is located within 1000 bases upstream or downstream of ER β sites, comprising 858 cases showing precise ER β -AGO2 co-localization, demonstrated by overlap analysis performed with GAT [58] and p-overlap (<https://github.com/brentp/poverlap>) (p value < 0.01; Fig. 7a; Additional file 11: Table S8). This result supports the possibility that the two proteins may act together in the genome. Indeed, since AGO2 is not required for ER β binding to BC cell genome, as silencing of this factor did not prevent in a significant way receptor binding to chromatin (data not shown), it is possible to assume that the ER β -AGO2 complex detected in Ct-ER β cell nuclear extracts by interaction proteomics is present also on the chromatin, where it could be conveyed by the receptor to specific sites, which are mainly distributed around TSSs (Fig. 7b), within promoters and 5' UTRs (FC > 2 and q -value \leq 0.05; Fig. 7c). This possibility is further supported by analysis of the enriched motifs bound by both

factors (Fig. 7d; Additional file 10: Table S7C), which showed representation of consensus motifs for ER β binding (ESR2 matrix), as well as TBP/FOXL1 and MEIS1/SOX10 motifs, both also found enriched when considering all AGO2 binding sites mapped in Ct-ER β cells only (TBP) or in both cell lines (MEIS1/SOX10; Fig. 6c).

AGO2 cooperates with unliganded ER β to modulate transcription rate and RNA maturation

To assess the possible cooperation between ER β and AGO2 in the transcriptional modulation of mRNA levels, nascent RNA was isolated from Ct-ER β or wild-type MCF-7 cell nuclei and sequenced (nascent-Seq). The role of AGO2 was analyzed comparing in both cell lines the results obtained with and without AGO2 silencing by shRNA transfection. AGO2 'knock-down' was efficient in both ER β + and ER β - cells, resulting in a ~60% reduction of the corresponding RNA (data not shown) and protein (Fig. 8a). Nascent RNA sequencing revealed a significant influence of both ER β and AGO2 on the transcription rate. In detail, 9273 genes were differentially transcribed in Ct-ER β with respect to wild-type cells, considering a |1.5| FC cut-off and FDR \leq 0.05, including 1301 (14%) that harbored an ER β binding site within the transcription unit (Additional file 12: Table S9A). Interestingly, by comparing the results obtained by expression profiling of mature RNAs by RNA-Seq with those obtained by nascent-Seq in the same cell line, a >60% correlation was observed among two datasets for transcripts modulated by ER β (data not shown), indicating that most changes in transcriptional rate are reflected in the mature transcriptome. On the other hand, the transcription rate of 8163 genes was significantly affected (|1.5| FC cut-off and FDR \leq 0.05) by AGO2 knock-down in ER β + cells (Additional file 12: Table S9b), including 5807 ER β -regulated transcripts (Additional file 12: Table S9c). Interestingly, 77 genes transcriptionally regulated by ER β and harboring both ER β and AGO2 binding sites within the transcription unit underwent an inversion of the ER β -induced transcriptional trend following AGO2 silencing (Additional file 12: Table S9d), demonstrating a role of the functional interaction between these two factors on ER β -mediated transcriptional regulation in BC cells upon physical association of the two proteins within the transcription unit (TU). Functional analysis revealed that these genes encode proteins mainly involved in cell death and survival, movement, growth and proliferation, and morphology. Indeed, they include also LAD1 and BCL9, known to be involved in cancer invasiveness. Indeed, LAD1 (Ladinin 1) has been proposed as a promising new target for 'triple negative' BC treatment [59] and BCL9 (B-cell CLL/lymphoma 9), a co-activator of Wnt-stimulated β -catenin-mediated transcription, is considered a molecular driver of BC early invasion [60] and has been shown to control

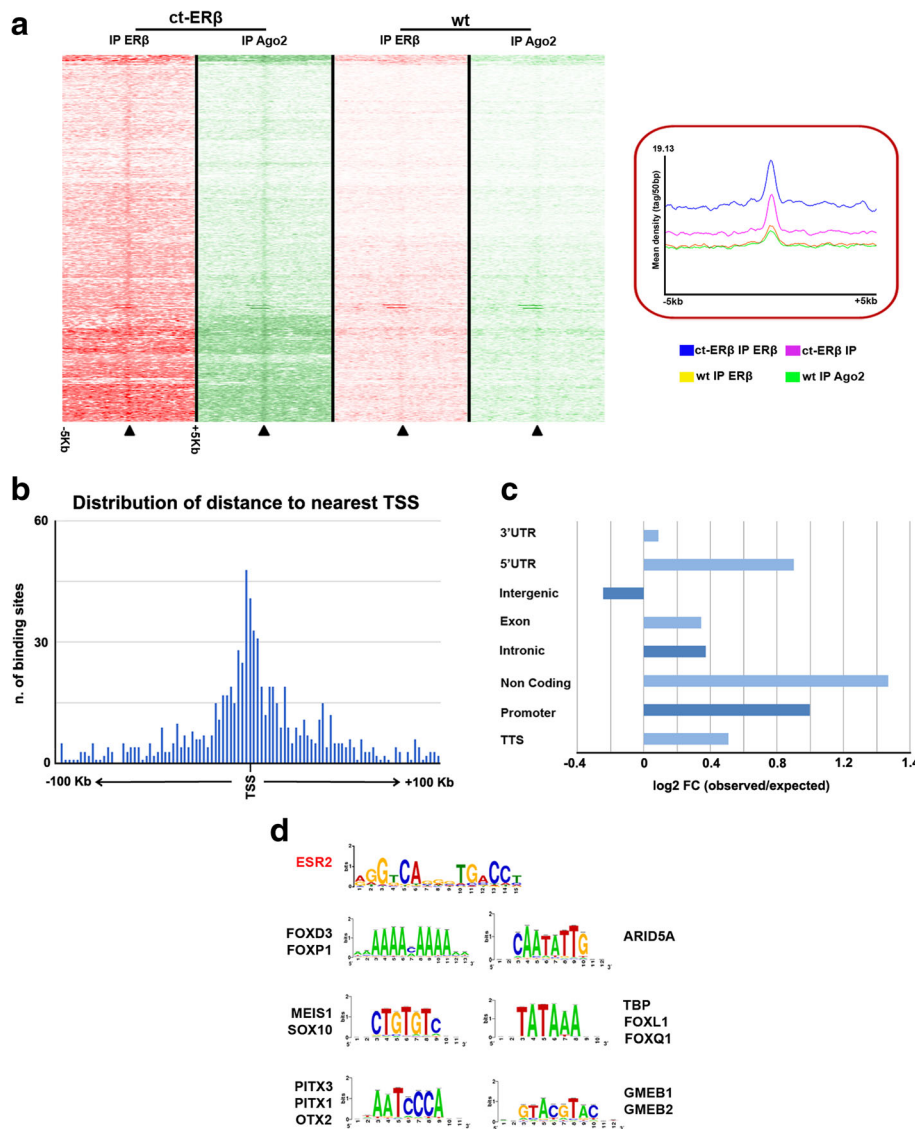


Fig. 7 Co-localization of ERβ and AGO2 binding sites in the BC cell genome. **a** Heatmap summarizing co-localized AGO2 (green) and ERβ (red) binding sites in ERβ + (Ct-ERβ) and ERβ - (wild-type wt) MCF-7 cells, respectively. The right panel shows mean read densities in the four conditions displayed. **b** Distribution of ERβ-AGO2 binding sites with respect to the nearest transcription start site in ERβ + cells. **c** Observed vs expected distribution of ERβ-AGO2 binding sites within gene segments, calculated according to Genome Analyzer Tester. **d** Statistically enriched transcription factor binding motifs most enriched within ERβ-AGO2 binding sites

estrogen signaling and breast carcinogenesis [61]. This evidence further suggests that ERβ association with AGO2 on the genome correlates with the effects of this receptor on BC cell biology.

We then investigated the role of ERβ-AGO2 association in co-transcriptional pre-mRNA splicing. The coupling between transcription and splicing in eukaryotes is well known but the mechanisms that drive it are still unclear, although some evidence points to a kinetic and functional coupling between the two events that determine spliceosome assembly and pre-mRNA loading during transcription [62, 63]. Considering the involvement of ERβ in the

control of the basic events of transcription and the identification, among ERβ-associated proteins, of several splicing factors associated also with nuclear AGO2 [12], we searched for evidence of RNA maturation and intron retention rate in the nascent-Seq datasets. By analyzing the data with the same procedure described for nascent transcript analysis, we investigated the global splicing events, in particular the intron retention coefficient, to verify the existence and nature of RNAs whose maturation may be modulated by ERβ-AGO2 functional cooperation during transcription. To quantify co-transcriptional splicing, we adopted the intron retention metric developed by Khodor

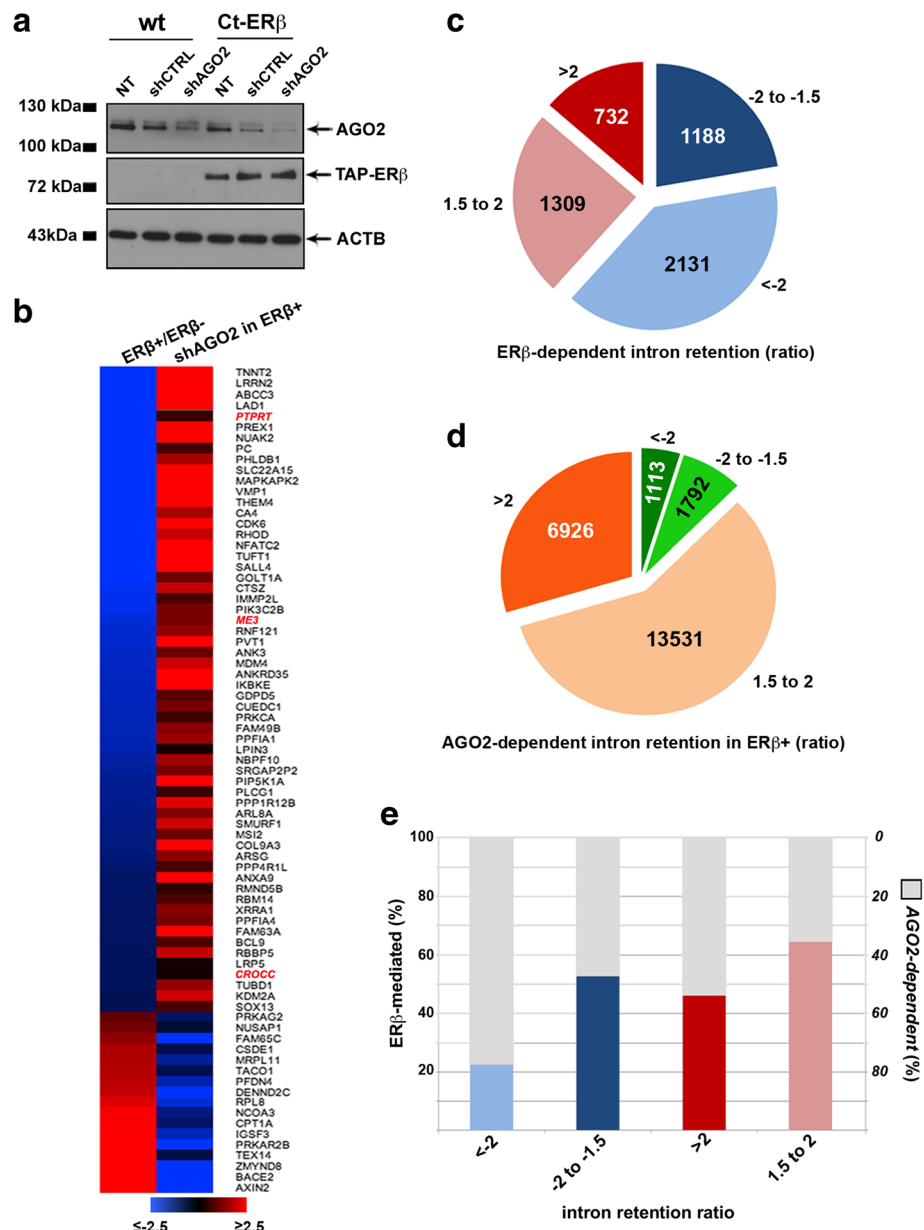


Fig. 8 Effects of ERβ and AGO2 on gene transcription rate and nascent RNA splicing. **a** Western blots showing the extent of AGO2 knock-down by shRNA in Ct-ERβ cells. *ACTB* β-actin, *NT* nontransfected cells, *shCTRL* cells transfected with a non-target shRNA (negative control). **b** Heatmap showing the transcription rate of a subset of genes showing ERβ and AGO2 binding sites within the promoter region or the gene body, expressed as fold change in ERβ + with respect to ERβ - cells (ERβ+/ERβ-) or in ERβ + cells after AGO2 knock-down with respect to control cells (shAGO2 in ERβ+). Genes in *red* and *italics* did not show statistically significant changes in shAGO2 cells. **c** Co-transcriptional pre-mRNA splicing modulation by ERβ. Number of introns showing increased or decreased retention (FDR ≤ 0.05) in Ct-ERβ compared to wild-type ERβ - cells. A positive intron retention ratio indicates reduced splicing efficiency and, conversely, a negative intron retention ratio indicates increased splicing efficiency. **d** Co-transcriptional pre-mRNA splicing modulation by AGO2 in ERβ-expressing cells. Introns showing increased or decreased retention (FDR ≤ 0.05) after AGO2 silencing in Ct-ERβ cells. **e** Bar plot showing, among ERβ-dependent splicing events, those affected by AGO2 knock-down

et al. [64]. In detail, to account for the variability due to imbalance among exons of different length, intron retention was calculated as the ratio between the read number/base pair of a given intron with respect to the same of all exons of the gene. In this way, we identified more than 11,500 splicing events (FDR ≤ 0.05, *t*-test) modulated by ERβ, with

5360 intron retention events being significantly affected in ERβ + cells compared to wild-type MCF-7 (Fig. 8c; Additional file 13: Table S10a), suggesting that ERβ may be directly involved also in this process in BC cells, as already demonstrated for AGO2. Subsequently, we measured 23,362 events (FDR ≤ 0.05, *t*-test) influenced only in ERβ +

cells by AGO2 silencing (Fig. 8d; Additional file 13: Table S10b), a result obtained after filtering out the events observed also in ER β – cells following AGO2 ‘knock-down’. By comparing these two datasets, and considering the 5360 introns influenced by ER β , we highlighted several splicing events modulated by both ER β and AGO2 (Fig. 8e). A stronger inhibition of the splicing efficiency upon AGO2 depletion was observed for ER β -induced intron splicing, where the effect of the receptor was reverted in 78% of the cases showing an intron retention coefficient < -2 and in 52% of the cases where this coefficient was between -2 and -1.5 . A similar, but less pronounced effect of AGO2 silencing was also observed for ER β -mediated intron retention, with increased splicing efficiency in 55% of the cases when the intron retention coefficient was > 2 and of 37% when it was between 1.5 and 2. Taken together, these data indicate that ER β and AGO2 cooperate in modulation of a sizeable amount of co-transcriptional splicing events in luminal-like BC cells, and that their functional association may be important to either promote or reduce the rate of co-transcriptional maturation of their target transcripts. This is further supported by the fact that only 8% of the 513 ER β -dependent intron retention events occurring in the 99 genes showing ER β -AGO2 complex binding were still detectable following AGO2 silencing, indicating that the large majority of them are AGO2-dependent.

Considering that ER β -AGO2 co-occupancy occurs at 858 sites in chromatin, and the fact that the chromatin-bound nuclear receptor can exert transcriptional effects also through long-range chromatin looping, the results reported above strongly suggest a functional role of the cooperation between ER β and AGO2 on direct regulation of gene transcription and co-transcriptional RNA splicing in BC cells. This is further supported by the fact that the ER β interactome of BC cell nuclei includes several transcriptional co-regulators and components of the RNA splicing machinery (Fig. 3) and the evidence that AGO2 plays a central role in assembly and/or stability of the nuclear ER β interactome (Fig. 4).

The AGO2-ER β complex associates with long and small RNAs in BC cells

Given the extent of ER β -AGO2 interaction in the extranuclear compartment observed in vivo by PLA (Additional file 6: Figure S5b) and in vitro by co-immunoprecipitation (Fig. 5) and the role of AGO2 as a major component of RISC [44, 45], we investigated the significance of this interaction in the cytoplasm.

To investigate the involvement of ER β in cytoplasmic mRNA processing, ER β -bound RNAs were searched for by RNA immunoprecipitation followed by sequencing (RIP-Seq). In order to evaluate the possibility of a role of ER β in RNA selection for RISC loading, the purified

samples were first analyzed by western blotting to detect the presence of three key components of the RISC loading complex: Dicer, AGO2, and TRBP. Indeed, cytoplasmic ER β formed complexes with all these proteins, with TRBP being particularly abundant, followed by AGO2 and then Dicer (Fig. 9a). RIP-Seq in Ct-ER β and wild-type MCF-7 cells (negative control) was carried out in triplicate by ER β immunoprecipitation, and input samples were also collected and sequenced for each condition to determine the background (input). Enrichment analysis (see “Methods”) identified a total of 1139 RNA associated with ER β , selected using a cut-off corresponding to the 75th percentile of enrichment-factor distribution (EF > 3.03). To select the enriched RNAs specifically associated with the receptor, we compared the set described above with all RNAs detected also in samples obtained with the same procedure from extracts of wild-type MCF-7 cells used as negative control and sorted those specific for Ct-ER β samples (614/1139). For the remaining 525 RNAs, we considered for further analysis those showing a ratio between the EF determined by RIP-Seq in Ct-ER β cells and the same determined in wild-type (ER β -) cells > 4 , and being above the 75th percentile with respect to the EF calculated in Ct-ER β and, at the same time, showing a negative enrichment (EF < -1) in wild-type cells. In this way we selected an additional 264 RNAs strongly enriched in ER β + samples and obtained a dataset comprising 878 RNAs showing strong evidence for association with ER β (Fig. 9b; Additional file 14: Table S11). By plotting the expression (log2RPKM) of the 878 RNAs in the input sample against the EF after immunoprecipitation, we observed that the majority of the ER β -associated RNAs were expressed in the cell at relatively low levels and all of them were either poorly enriched or not detected in ER β – cells (Fig. 9c). This indicated that the selection procedure was not biased by high concentrations of the RNAs in the starting material and supported the possibility that the RNAs identified were indeed specifically associated with ER β . We then classified the 878 enriched RNAs according to the “Gene_Biotype” term described in the Ensembl annotation file. This revealed that most ER β -bound RNAs were mainly protein coding (38%, p value < 0.01 , hypergeometric test) and antisense lncRNAs (26%, p value $< 2.044 \times 10^{-4}$, hypergeometric test), with the remaining enriched molecules distributed as follows: 13% pseudogenes (p value < 0.01 , hypergeometric test), 8% processed transcripts (p value $< 2.15 \times 10^{-15}$, hypergeometric test), 7% linc RNAs (p value $< 1.93 \times 10^{-7}$, hypergeometric test), 4% sense_intronic lncRNA (p value $< 1.32 \times 10^{-7}$, hypergeometric test) and 4% represented by other classes of RNAs, including pre-miRNAs, snoRNAs, snRNAs, and sense_overlapping (p value < 0.05 , hypergeometric test) (Fig. 9d, pie chart).

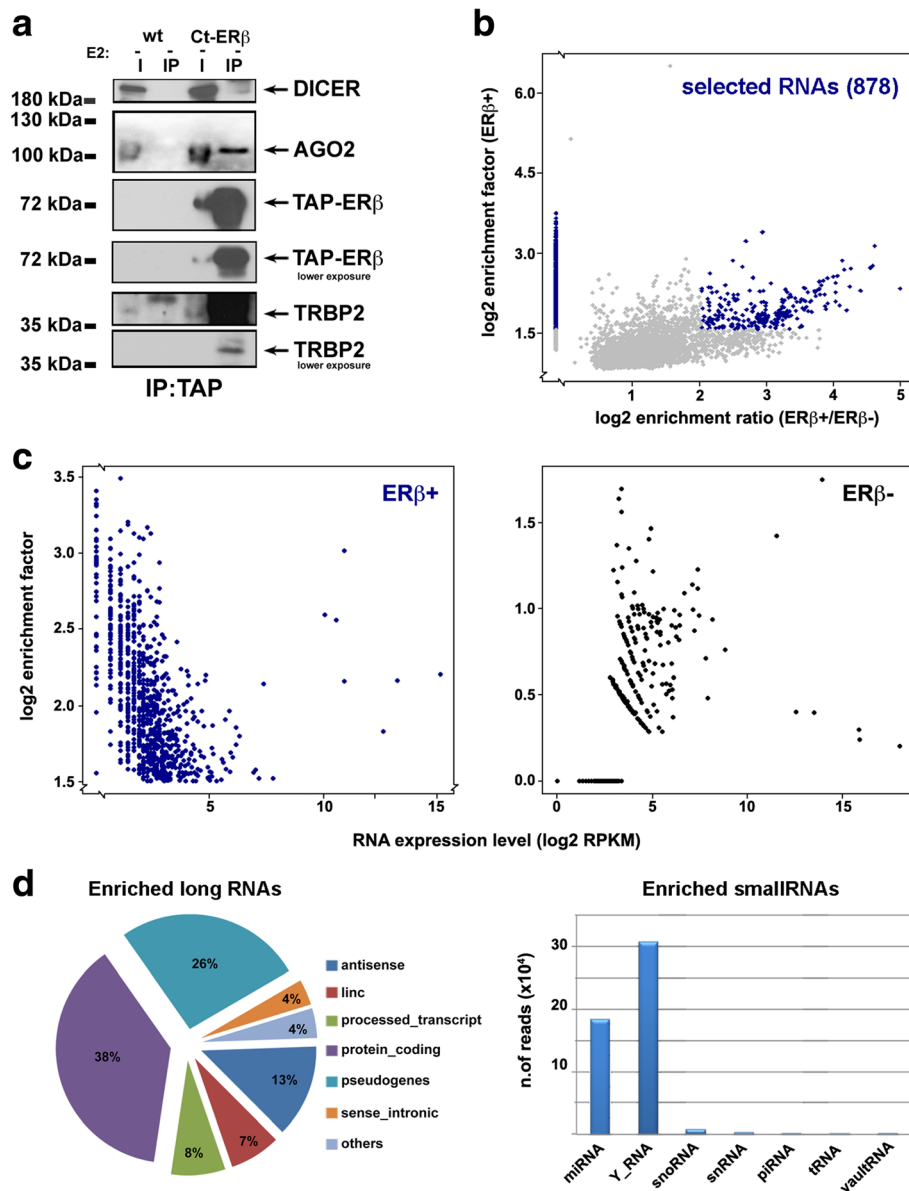


Fig. 9 ERβ association with RISC and a subset of small and long RNAs in BC cell cytoplasm. **a** Association of ERβ with the RISC proteins AGO2, DICER, and TRBP2 analyzed by western blotting following immunoprecipitation of cytosolic extracts with IgG (TAP) in Ct-ERβ and, as control, wild-type (*wt*; ERβ⁻) MCF-7 cells. **b** The enrichment coefficient of all the RNAs (>4500) co-immunoprecipitated with ERβ. *Blue dots* mark the 878 RNAs selected based on the enrichment factor in ERβ + samples and the relative enrichment (ratio between enrichment factors in ERβ + and ERβ⁻, used as negative control). **c** Direct comparison of RIP-Seq results relative to the 878 ERβ-interacting RNAs in ERβ + with respect to ERβ⁻ samples. **d** The relative abundance of different classes of small (*right*) and long (*left*) RNAs, respectively, found associated with ERβ by RIP-Seq

The RIP samples were also subjected to small RNA-Seq, to evaluate whether mature components of these classes of molecules also associate with the receptor, comprising in particular—among miRNAs—those targeting the co-enriched mRNAs. miRNAs were considered enriched if they showed an EF > 1.5 and an associated adjusted *p* value < 0.1. Applying the same filtering criteria described above for long RNAs, we identified several Y_RNAs, followed by miRNAs (18; Additional file 15:

Table S12), snoRNAs, snRNAs, piRNAs, and vaultRNAs (Fig. 9d, bar graph). The heatmap reported in Fig. 10a displays the most highly enriched (EF > 10; 132 RNAs) long RNAs associated with ERβ. Interestingly, in wild-type MCF-7 cells most of these were expressed below the threshold of ten reads (not detected) or showed a negative EF (ERβ⁻, right lane). The 18 miRNAs identified (Fig. 10b) also had either a negative EF or an expression level below the threshold in ERβ⁻ cells while, in most cases, they were

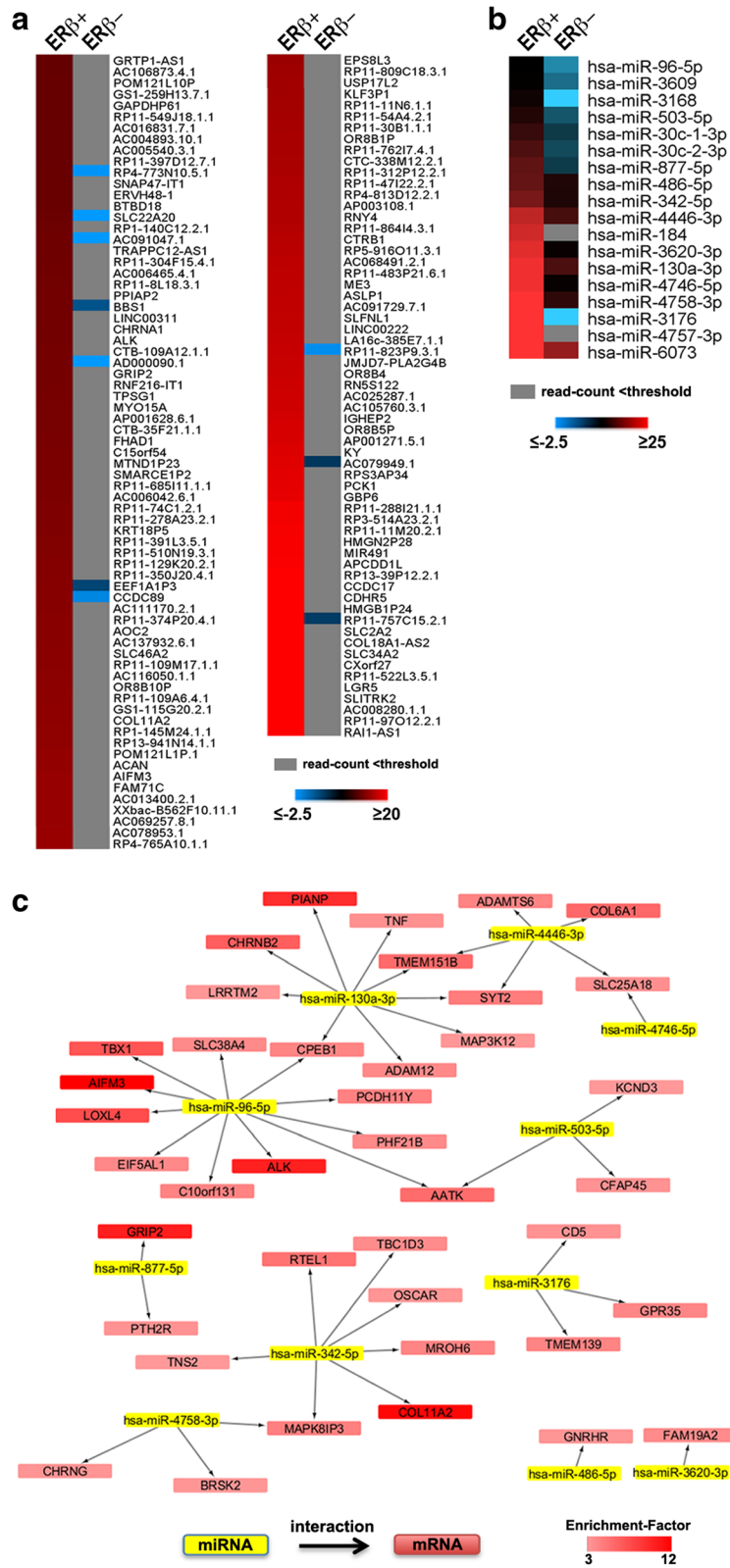


Fig. 10 (See legend on next page.)

(See figure on previous page.)

Fig. 10 Network analysis of the functional interactions between miRNAs and mRNAs found associated with ER β in BC cell cytoplasm by RIP-Seq. **a, b** Heatmaps showing the top enriched miRNAs (**b**) and long RNAs (**a**) associated with ER β . *Gray cells* indicate that the read count of the RNA was below the threshold in ER β – cells. **c** miRNA–mRNA interaction network. Known (experimentally validated) interactions between ER β -associated miRNAs (*yellow*) and mRNAs (*red*)

expressed at very low levels in ER β + cells. Interestingly, most of these miRNAs are known to target mRNAs co-associated with ER β . To highlight the relationships between enriched miRNAs and mRNAs, a network analysis was performed using the miRNA Target Filter Module of IPA, considering only experimentally validated miRNA–RNA interactions (Fig. 10c; Additional file 16: Table S13). Among the most enriched long RNAs represented in the network, ALK mRNA was directly targeted by miR-96-5p. This might represent a very meaningful interaction, since the oncogenic protein encoded by this RNA has been found highly expressed in aggressive BCs, including ‘triple negative’ BC, and the protein it encodes has been proposed as a new drug target in inflammatory BC [65]. Its downregulation through ER β -mediated mRNA decay or translation inhibition could thus be one of the reasons for the better prognosis of ER β -expressing BCs.

When combined with the presence of RISC loading factors in the same ER β complexes, the fact that most of these RNAs are present in low amounts in the cell suggests the involvement of ER β in inducing recruitment of miRNAs and selected target mRNAs by the RISC loading machinery for destabilization of the latter.

Discussion

The estrogen receptors ER α and ER β are directly involved in carcinogenesis and tumor progression in multiple neoplasms of the female genital tract, and ER α was the first molecule amenable to drug targeting in BC, where its presence in cancer cells is still one of the main markers for identification of patients that will benefit from endocrine therapy. For a long time ER α was considered the only estrogen receptor in mammals, but a second one, termed ER β , was subsequently discovered and found to play important roles in breast and other cancers [16]. ER β shows 55% identity with ER α in its ligand-binding domain and approximately 97% similarity in the DNA-binding domain (DBD). Reflecting the high degree of similarity in their DBDs, in chromatin both receptors target predominantly the same conserved estrogen response element (ERE; 5′-GGTCAnnnTGACC-3′) as either homodimers or α/β heterodimers [23]. ER β binds 17 β -estradiol (E2) with relatively low affinity compared to ER α , but, contrary to ER α , shows potent effects also in the absence of ligands [25], like other members of the nuclear receptor superfamily of homeostatic regulators.

ER β is expressed in normal mammary epithelial cells and in a fraction of BCs, showing decreased expression in cancer compared with benign tumors or normal tissues, suggesting that a reduction of this receptor in cancer cells could represent a critical stage in hormone-dependent tumor progression [66]. Interestingly, Förster et al. [67] reported that ER β null mice show impairment of pregnancy-induced terminal differentiation of the mammary gland, suggesting that this receptor subtype is required for normal development of this organ. When combined, these findings led to the hypothesis that ER β might act as oncosuppressor in certain target tissues, including mammary epithelia, by interfering with the tumor promoting actions of estrogen via ER α and of other carcinogenic stimuli and by controlling genetic programs for cell differentiation and proliferation. This was further supported by the observation that mice lacking ER β display multi-focal hyperplasia in prostate and bladder [68]. For all the reasons stated above, understanding the molecular mechanisms of ER β actions is a critical issue in cancer, in particular in BC biology. By interaction proteomics we identified molecular partners of both receptors in the nucleus of cells exposed to agonist and antagonist ligands [40–43, 69], and characterized the effects of unliganded ER β in BC cells, demonstrating its significant effects on cell proliferation, miRNA expression, and the cell proteome [25].

In this study, we demonstrated that unliganded ER β binds to the BC cell genome and induces reprogramming of the cell transcriptome, promoting also alternative splicing of a sizeable number of RNAs transcribed from its target genes. To understand the molecular determinants of these effects, we applied interaction proteomics and identified a large set of ligand-free ER β -interacting proteins in the cell nucleus. The functions of several of the proteins reveal how this receptor can control key processes in BC cells, including gene transcription and RNA splicing and turnover. Among the molecular partners of ER β , our attention was caught by AGO2, for the basic functions this protein exerts not only on miRNA biogenesis and actions but also on gene regulation. We thus investigated in detail the functional significance of the protein complex(es) containing AGO2 and ER β , since we observed that several ER β -interacting proteins were known to be also AGO2 interactors. These include, together with factors involved in RNA biogenesis, splicing, and maturation, also pleiotropic factors controlling key functions in the cell, such as CNOT1 (CCR4–NOT transcription complex subunit 1), a

scaffolding component of the major effector complex of miRNA-mediated gene silencing CCR4-NOT, which associates with the ATP-dependent RNA helicase DDX6 (another interactor in common between ER β and AGO2) to exert this function [70]; the metastasis-associated protein MTA2 (Metastasis associated 1 family member 2), a member of the tumor-associated family of transcriptional regulators and central component of nucleosome remodeling and histone deacetylation complexes [71], shown to be involved in both development and metastasis of a wide spectrum of cancers, including in particular hormone-independent BCs [72]; ADAR (Adenosine deaminase, RNA-specific), an RNA-editing enzyme specifically active in BC, where it has been shown to regulate cell proliferation and apoptosis [73]; COPA (Coatomer protein complex subunit alpha), a component of the coatomer complex of secretory vesicles involved in ER–Golgi transport whose mutation and inactivation have been recently shown to cause growth inhibition and apoptosis in cancer cells [74] and an autosomal dominant immune dysregulatory disease [75]; the human homolog of NOP56 (Nucleolar protein 56/NOP56 ribonucleoprotein), a core component of the C/D box snoRNP complex that controls ribosome biogenesis by regulating pre-rRNA processing and shows dynamic subcellular redistribution in response to growth conditions and nutrient availability [76].

Searching for the biological significance of ER β –AGO2 association, transcriptional co-regulation of genes mediated by the joint action of the two proteins was demonstrated by the identification of several genomic regions occupied by both ER β and AGO2 and by cooperation between the two proteins in modulation of transcription rate and co-transcriptional pre-mRNA splicing. Interestingly, AGO2 binding to the genome appears quite different in ER β + compared to ER β – cells (Fig. 6). This result, which was reproducible in independent experiments, suggests that the nuclear receptor can induce re-positioning of the argonaute protein within chromatin. Since AGO2 is not a DNA-binding protein and, therefore, its association with the genome is mediated by other factors, it is conceivable to assume that ER β might influence the cellular levels of some of these factors, or their ability to bind DNA. Supporting the first possibility, we observed that the mRNAs encoding MGA, FOXP1, and GMEB2 are up-regulated in ER β + cells, while those for ARID5A, GMEB1, NFATC2, NFAT5, TEAD1, and STAT4 are down-regulated. The binding matrixes for all these factors were significantly enriched within the AGO2 sites mapped here (Figs. 6 and 7). Furthermore, it is also possible that AGO2 tethering to the genome is mediated, in some instances, by RNAs whose expression is controlled by the receptor,

which as shown here induces a profound effect on the cell transcriptome (Fig. 1; Additional file 1).

A set of 153 ER β -responsive genes showing co-occupancy by ER β and AGO2 on defined sites in their transcription units was identified, including 77 whose transcription rate was significantly affected ($|FC| > 1.5$) in ER β + cells upon AGO2 silencing. Functional analysis highlighted that these genes are involved in processes such as cell growth and proliferation, death and survival, or motility. Considering the overall involvement of these genes on cellular pathways, G α q and phospholipase C signaling were significantly activated, while protein kinase A and B signaling may be inhibited. Both these pathways are tightly related to cancer progression and apoptosis. In particular, it has been demonstrated that G protein-coupled receptors are involved in BC progression and that Gq signaling promotes cancer cell apoptosis through phospholipase C [77–79]. On the other hand, protein kinase A signaling has been shown to promote mammary tumorigenesis [80] and to determine ER α repositioning at promoters and tamoxifen resistance [81]; its inhibition by ER β –AGO2 cooperation may thus negatively affect cancer cell proliferation and survival.

In addition to the genes for which we could demonstrate both binding of the AGO2–ER β nuclear complex and transcriptional regulation, we detected several others that are influenced in their transcription and/or maturation rate. Since it has been reported that AGO2 association with chromatin induces the formation of heterochromatin mediated by siRNAs in mammalian cells, probably determining slowdown of RNA polymerase II and alternative splicing events [12, 82], its association with ER β may give rise to significant effects on gene activity via different mechanisms. Indeed, considering the capability of ERs to mediate transcription through long-range chromatin interactions [83], and the fact that AGO2 has been shown to co-localize with the insulator factor CTCF [84], known to mediate chromatin looping [85], association between AGO2 and ER β may control gene activity also when occurring at a distance from the targeted transcription units. This could explain, at least in part, their massive effects shown here on gene transcription in BC cells. On the other hand, regulation of genome activity by the combined action of AGO2 and ER β could occur via at least two, independent and not mutually exclusive, events. On one hand, binding to the genome of the complex(es) comprising the transcription factor and the argonaute protein together with other protein(s) and/or RNA(s) determines modulation of target gene expression. On the other, ER β and AGO2 may bind nascent transcripts and modulate pre-mRNA splicing by recruitment and association with splicing factors. Indeed, we observed several

such factors in common between the ER β interactome identified here and the AGO2-associated proteins described by Ameyar-Zazoua et al. [12].

Association between AGO2 and ER β also occurs, both in vivo and in vitro, in the cytoplasm, where isolation of ER β -bound RNAs and miRNAs suggests that the receptor may assist the argonaute protein in the loading of specific miRNA-mRNA molecules in RISC, thus contributing also to post-transcriptional regulation of gene expression. Interestingly, we identified 868 RNAs and 18 miRNAs specifically associated with ER β . Notably, computational analysis revealed that miRNA-mRNA molecules bound to ER β are implicated in Wnt and cadherin signaling pathways. The first has been found dysregulated in BC [86, 87] and associated with metastasis in 'triple negative' tumors [88], while the second is tightly correlated to the Wnt, E-cadherin, and N-cadherin pathway, contributing significantly to epithelial-mesenchymal transition and metastasis [89]. A negative effect of ER β on translation and/or stability of the mRNAs encoding these factors could thus also be part of its activity as an oncosuppressor, contributing to the better prognosis of ER β -expressing tumors. Combined with the relationships between AGO2 and tumorigenesis and cancer progression [90], the results reported here open new avenues for understanding the actions, and resulting effects, of ER β and AGO2 in cancer cells.

Finally, interaction of AGO2 with ER β appears to be indirect, since yeast two-hybrid assays failed to demonstrate direct association between the two proteins and, more important, in vitro RNase A digestion of both nuclear and cytosolic extracts strongly reduced co-immunoprecipitation of the two proteins, suggesting that this interaction may require one or more RNAs. To our knowledge, ER β binding to RNA has not been described previously; however, this is well known for other nuclear receptors, such as the androgen receptor [91] and ER α [92], where a novel RNA binding domain in the N-terminus of the protein has been identified [93]. Attempts to identify the RNA(s) involved in ER β -AGO2 complex formation and/or stability have so far been unsuccessful, but a preliminary computational prediction, performed on RNAs specifically binding to ER β and AGO2 in the cytosolic compartment, suggested that long noncoding RNAs could be likely candidates as bridging molecules (data not shown). Understanding this aspect will need further investigations that go beyond the scope of the present study.

Conclusions

The results of this study demonstrate that AGO2 and ER β can physically and functionally associate, both in the nucleus and the cytoplasm, in complex(es) comprising also several other proteins and RNAs. The final biological outcome of such association appears to depend

upon the sum of different variables, including transcriptional, splicing, and post-transcriptional events and, possibly, the specific cellular context. These findings provide new leads toward understanding the oncosuppressive role of ER β via regulation of gene transcription, RNA maturation, and post-transcriptional control of RNA activity, and the consequences of the loss of this protein in transformed cells. Demonstration of the general importance of these results, obtained here in a cellular model of ER β + BC, for the control of cellular functions and its derangement during carcinogenesis and tumor progression will require, however, further validation in less artificial conditions, in particular in vivo animal models, patient-derived xenografts and tumor biopsies.

Methods

Cell culture

Stable clones expressing ER β tagged with TAP-tag at either the C-terminus (Ct-ER β) or N-terminus (Nt-ER β) and TAP-tagged ER α were obtained from human breast cancer MCF-7 Tet-Off cells (ER-alpha positive; ATCC, catalog number HTB-22) as previously described [40, 69]. For generation of ER β tagged inducible clones, the human full-length cDNA clone pCMV6-ESR2 (RC218519) encoding human ESR2 was purchased from Origene. ESR2 sequence, including the Myc and Flag tags, was subcloned into the BamHI and EcoRI sites of pTRE-Tight expression vector (Clontech). All cell lines were propagated in Dulbecco's modified Eagle medium (DMEM; Sigma-Aldrich) supplemented with 10% FBS (HyClone) and antibiotics: 100 U/ml penicillin, 100 mg/ml streptomycin, 250 ng/ml Amphotericin-B. Steroid deprivation (starvation) was performed by culturing in DMEM without phenol red and 5% dextran coated charcoal stripped serum (DCC-FBS) for 5 days. Cell lines were authenticated by short tandem repeat (STR) profiling and routinely tested for *Mycoplasma* contamination with MycoAlert mycoplasma detection kit (Lonza).

RNA extraction sequencing and data analysis

Total RNA was extracted from ER β + and ER β - (Ct-ER α and/or wild-type) MCF-7 cells using the standard RNA extraction method with TRIzol (Life Technologies). Before use, the RNA concentration in each sample was assayed with a ND-2000c spectrophotometer (Thermo-scientific) and its quality and integrity assessed with the Agilent 2100 Bioanalyzer with Agilent RNA 6000 nano kit (Agilent Technologies). For RNA sequencing experiments, indexed libraries were prepared using 1 μ g of total RNA as starting material, with a TruSeq Stranded Total RNA Sample Prep Kit (Illumina Inc.). Libraries were sequenced (paired-end, 2 \times 100 cycles) at a concentration of 8 pM/lane on the HiSeq 2500 platform (Illumina Inc.). The raw sequence files generated (.fastq files)

underwent quality control analysis using FASTQC (<http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>) and quality-checked reads were then aligned to the human genome (assembly hg19) using TopHat version 2.0.10 [94] with the standard parameters. The expression value of each mRNA was normalized to FPKM (fragments per kilobase of exon model per million of sequenced reads) as computed by Cufflink [95]. Differentially expressed mRNAs were identified using DESeq2 [96]. Firstly, gene annotation was obtained for all known genes in the human genome, as provided by Ensemble (GRCh37; https://support.illumina.com/sequencing/sequencing_software/igenome.html). Using the reads mapped to the genome, we computed the number of reads mapping to each transcript with HTSeq-count [97]. A given mRNA was considered expressed when detected by at least ≥ 10 reads. The raw read counts were then used as input to DESeq for calculation of normalized signal for each transcript in the samples. Differential expression was reported as fold change [1.5] along with associated adjusted p values ($FDR \leq 0.05$), computed according to Benjamini–Hochberg. Alternative splicing data analysis was performed as described previously [30]. Raw RNA sequencing data are deposited in the EBI ArrayExpress database (<http://www.ebi.ac.uk/arrayexpress>) with accession number E-MTAB-4363.

Chromatin immunoprecipitation, sequencing, and data analysis

C-TAP-ER β and MCF-7 control cells were hormone-deprived for 5 days. For each assay, a total of about 15×10^6 cells were fixed, lysed to isolate nuclei, sonicated, and diluted as described by Schmidt et al. [98], with minor modifications. An aliquot of nuclear extract was taken as input. For ER β pull-down, chromatin samples were incubated, as described earlier [69], at 4 °C for 3 h with 40 μ l of IgG Sepharose 6 fast Flow (GE Healthcare Bio-Science AB) properly equilibrated in Poly-Prep chromatography columns (0.84 cm, Bio-Rad), according to the manufacturer's instructions. For AGO2 immunoprecipitation, chromatin samples were incubated at 4 °C overnight with 40 μ l of pre-blocked magnetic beads (Dynabeads, ThermoFisher) conjugated with 1 μ g of mouse monoclonal anti-AGO2/eIF2C2 (ab57113, Abcam). As negative control for these experiments, chromatin samples were also incubated overnight with 1 μ g of mouse monoclonal anti-Flag M2 affinity purified (F1804, Sigma-Aldrich). Bead washing, elution, reverse crosslinking and DNA extraction were then performed as described [69]. The size distribution of each ChIP DNA sample was assessed by running a 1 μ l aliquot on an Agilent High Sensitivity DNA chip using an Agilent Technologies 2100 Bioanalyzer (Agilent Technologies). The concentration of each DNA sample was determined by using a Quant-IT DNA Assay Kit-High Sensitivity and a Qubit Fluorometer (Life Technologies). Purified ChIP

DNA (10 ng) was used as the starting material for sequencing library preparation from three independent ChIP experiments. Indexed triplicate libraries were prepared with a TruSeq ChIP Sample Prep Kit (Illumina Inc.) and were sequenced (single read, 1×50 cycles) on a NextSeq 500 (Illumina Inc.).

Read alignment and quality control of ChIP-seq data

The raw sequence files generated (.fastq) underwent quality control analysis using FASTQC (<http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>). Reads were aligned to the reference human genome assembly (hg19) using bowtie [99], allowing up to one mismatch and considering uniquely mappable reads. Duplicated reads were removed using Picard tools v 2.9.0 (MarkDuplicates; <https://broadinstitute.github.io/picard>).

Peak calling

For each biological replicate and corresponding control samples, peak calling was performed using MACS2 [33] with p value set to 0.05. The peaks obtained for each biological replicate were combined using MuSERA [34] with the following parameters: replicate type, biological; Ts, 1E-08; Tw, 1E-04; γ , 1E-08; Benjamini–Hochberg false discovery rate (α), 0.005, using the lowest p value when multiple regions from a sample intersected with the region of another sample and considering peaks common to at least two replicates (C:2). The annotation of peaks to the nearest gene was performed combining the information obtained using the annotatePeaks.pl function of HOMER [100] and the Annotation and Statistics of Genomatix Software suite. Comparison, integration, and quantification were performed using seqMINER [101]. Over-represented sequence motifs for known transcription factors, according to motif descriptors in the JASPAR database, were determined using PScan-ChIP [102]. Only over-represented motifs with p value $\leq 1E-10$ were considered.

De novo motif discovery

The predicted sequences of ER β and Ago2 binding sites were extracted and used for de novo motif discovery using the RSAT peak motifs method with default parameters [103] and Meme-ChIP [104]. For ER β binding sites, the ERE binding motif was searched with MatInspector application [105], using a core similarity threshold of 0.75 and a matrix similarity threshold of Optimal -0.02.

Binding site statistics

The overlap between ER β and Ago2 binding sites was calculated using bedtools intersect [106]. The significance of the overlaps was assessed using the poverlap tool (<https://github.com/brentp/poverlap>) by performing 100,000 simulations and allowing shuffling of both

datasets. The significance of overlaps of ER β with different genomic regions (3' UTR, 5' UTR, intergenic, exonic, intronic, promoter, and TSS) was assessed using Genomic Association Test (GAT) [58] with 10,000 simulations. In each case considered, the distance to the TSS was computed using ChIPseek [107]. Raw ChIP-Seq data have been deposited in the EBI ArrayExpress database (<http://www.ebi.ac.uk/arrayexpress>) with accession number E-MTAB-4359.

Tandem affinity purification

Nuclear extraction and tandem affinity purification from C-TAP-ER β and control (wild-type MCF-7) cells were performed as previously described [40, 41, 69]. Partially purified samples, coming from the first purification step, were then subjected to mass spectrometry analysis for protein identification.

Nano LC-MS/MS and data analysis

Three biological replicates of partially purified samples from Ct-ER β and control MCF-7 cells were separated on SDS-PAGE and visualized with silver-staining. After separation, SDS-PAGE lanes were sliced into six pieces, and the proteins were in-gel digested with trypsin into peptides and analyzed by LC-MS/MS as previously described [69]. MS data were acquired using Analyst QS 2.0 software. The information-dependent acquisition method consisted of a 0.5 s TOF-MS survey scan of m/z 400–1400. From every survey scan two most abundant ions with charge states +2 to +4 were selected for product ion scans. Once an ion was selected for MS/MS fragmentation, it was put on an exclusion list for 60 s. LC-MS/MS data from the biological replicates were combined and searched against SwissProt 2010 (517,802 sequences, 182,492,287 residues; human, 20,283 sequences) for control (TAP-only) and SwissProt 2010 (523,151 sequences, 184,678,199 residues; human, 20,259 sequences) for Ct-ER β samples. The search criteria for Mascot searches were: trypsin digestion with one missed cleavage allowed, carbamidomethyl (C) as fixed modification and oxidation (M), phospho (ST), phospho (Y) as variable modifications. For the LC-MS/MS spectra the maximum precursor ion mass tolerance was 50 ppm and MS/MS fragment ion mass tolerance 0.2 Da, and a peptide charge state of +1, +2, or +3 was used. All of the reported protein identifications were statistically significant ($p < 0.05$). To eliminate the redundancy of proteins that appear in the database under different names and accession numbers, the single protein member with the highest protein score (top rank) was selected from multiprotein families for the identification results. Protein reported as Ct-ER β molecular interacting partners were selected by filtering them against the proteins identified in negative control after quality assessment of the identification peptides. The Mascot search results,

including peptide sequences identifying Ct-ER β interacting proteins, are reported in Additional file 7: Table S4 (Mascot search results sheets).

For experiments performed in the presence or absence of AGO2 silencing, three biological replicates of partially purified samples from Ct-ER β for each of the two conditions were analyzed. The proteins were precipitated with 10% TCA in acetone solution and dissolved in 40 μ L 0.2% ProteaseMAX™ Surfactant, Trypsin Enhancer (Promega) in 50 mM NH₄HCO₃ followed by protein reduction, alkylation, and in-solution digestion with trypsin (Promega), performed overnight at 37 °C. Peptides were desalted and concentrated before mass spectrometry by the STAGE-TIP method, using a C18 resin disk (3 M Empore). The peptides were eluted twice with 0.1% TFA/50% ACN, dried, and solubilized in 7 μ L 0.1% TFA for mass spectrometry analysis. Each peptide mixture was analyzed on an Easy nLC1000 nano-LC system connected to a quadrupole Orbitrap mass spectrometer (QExactive Plus, ThermoElectron) equipped with a nano-electrospray ion source (EasySpray/Thermo). For the liquid chromatography separation of the peptides, an EasySpray column capillary of 50 cm bed length (C18, 2 μ m beads, 100 Å, 75 μ m inner diameter, Thermo) was employed. The flow rate was 300 nL/min, and the peptides were eluted with a 2–30% gradient of solvent B in 120 min. Solvent A was aqueous 0.1% formic acid and solvent B 100% acetonitrile/0.1% formic acid. The data-dependent acquisition automatically switched between MS and MS/MS mode. Survey full scan MS spectra were acquired from a mass-to-charge ratio (m/z) of 400 to 1200 with the resolution $R = 70,000$ at m/z 200 after accumulation to a target of 3,000,000 ions in the quadrupole. For MS/MS, the ten most abundant multiple-charged ions were selected for fragmentation on the high-energy collision dissociation (HCD) cell at a target value of 100,000 charges or maximum acquisition time of 100 ms. The MS/MS scans were collected at a resolution of 17,500. Target ions already selected for MS/MS were dynamically excluded for 30 s. The resulting MS raw files were submitted to MaxQuant software version 1.5.7.4 for protein identification using the Andromeda search engine. Carbamidomethyl (C) was set as a fixed modification and protein N-acetylation and methionine oxidation were set as variable modifications. First search peptide tolerance of 20 ppm and main search error 4.5 ppm were used. Trypsin without the proline restriction enzyme option was used, with two allowed miscleavages. The minimal unique + razor peptide number was set to 1, and the allowed FDR was 0.01 (1%) for peptide and protein identification. Label-free quantification (LFQ) was employed with default settings. The SwissProt human database (August 2016 release, with 154,660 entries) was used for the database searches. Known contaminants as provided by MaxQuant and identified in the samples were excluded from further analysis. LFQ

intensities were used for differential expression analysis. Protein LFQ values were further normalized by ESR2 LFQ value in each replicate for each dataset. Then, to identify statistically modulated proteins a two-sample *t*-test statistical analysis with a permutation based FDR cut-off of 0.01 was performed. All the protein identification and quantification data are reported in Additional file 8: Table S5a–d.

The mass spectrometry proteomics data have been deposited in the ProteomeXchange Consortium via the PRIDE [108] partner repository with the dataset identifier PXD006280.

Nuclear protein extraction and co-immunoprecipitation

Nuclear protein extracts were prepared, as described [69], from inducible MCF-7 Tet-On cells expressing Myc-Flag-ER β treated or not with doxycycline (2 μ g/ml) for 24 h. To immunoprecipitate AGO2, 1 mg of nuclear proteins was incubated overnight at 4 °C with 2 μ g of mouse monoclonal anti-AGO2/eIF2C2 (ab57113, Abcam) and then at 4 °C for 1 h with 35 μ l of equilibrated slurry Protein A/G Plus-Agarose (sc-2003, Santa Cruz Biotechnology). To immunoprecipitate myc-flag-tagged ER β , the same amount of nuclear proteins was incubated for 2 h at 4 °C with 35 μ l of equilibrated slurry EZview Red Anti-c-Myc Affinity Gel (E6654, Sigma Aldrich). After binding, the beads were sequentially washed with IPP150 buffer (7.14 mM HEPES pH 7.5, 8.92% glycerol, 150 mM NaCl, 0.54 mM MgCl₂, 0.07 mM EDTA pH 8, 1 \times protease inhibitors) and wash buffer (50 mM Tris-HCl pH 7.6, 150 mM NaCl, 1 \times protease inhibitors). To elute ER β -immunoprecipitated samples from the beads, an elution at 4 °C for 30 min was performed using c-Myc Peptide (M 2435, Sigma Aldrich).

Immunofluorescence assays

MCF-7 Tet-On cells stably expressing tet-inducible Myc-Flag-ER β were seeded on microscope glass slides and starved for 3 days before treatment with doxycycline for 24 h. Cells were then fixed with 4% paraformaldehyde for 15 min, washed with PBS-Tween three times and permeabilized with 0.1% Triton X-100 in PBS. After washing with PBS and blocking with 0.5% BSA, slides were incubated with mouse anti-Myc (clone 4A6, Merck Millipore, 1:200) and rabbit anti-AGO2 (07-590, Millipore, 1:66), incubated and washed for three times each with 0.5% BSA and then PBS, before incubation with Alexa Fluor 488 goat anti-mouse IgG (ThermoFisher) and Cy3 anti-rabbit (Jackson Immuno Research) secondary antibodies. BSA and PBS washes were repeated and cells were covered with mounting medium containing 4',6-diamidino-2-phenylindole (DAPI 1:20,000) and imaged with a confocal microscope (Leica DM6000 B). Images were processed with ImageJ software (<https://imagej.net>).

Proximity ligation assay

MCF-7 cells were plated on microscope glass slides and after 5 days of starvation, transiently transfected with plasmids expressing either Myc-tagged AGO2, Flag-tagged ER β , or Flag-tagged ER α . Non-transfected cells were used as control. Cells were washed three times in ice-cold PBS and fixed by incubating them in 4% paraformaldehyde for 20 min under gentle agitation in the dark. After three washes with PBS, cells were permeabilized with 0.2% Triton X100 for 5 min under gentle agitation and then washed again. A proximity ligation assay was performed following the manufacturer's instructions. In detail, fixed and permeabilized cells were blocked in a pre-heated humidity chamber for 30 min at 37 °C with one drop of blocking solution per 1 cm². Then, primary antibodies were added (rabbit anti-Flag Tag, F7425 and mouse anti-Myc Tag: clone 4A6, Merck Millipore) and incubated for 1 h at 37 °C in a pre-heated humidity chamber. Slides were washed twice for 5 min in wash buffer A in a staining jar with gentle orbital shaking and then incubated with PLA probes (Mouse \pm for the detection of exogenous AGO2, Rabbit \pm for the detection of ER β or ER α and Mouse – and Rabbit + for the detection of AGO2/ER interactions) in a pre-heated humidity chamber for 1 h at 37 °C. After two other washes with wash buffer A, a ligation reaction was performed by adding the ligase to the slides (1:40 dilution of the stock) and incubating them in a humidity chamber for 30 min at 37 °C. Slides were washed twice with wash buffer A for 2 min under gentle agitation and then the amplification-polymerase solution was added to the cells and left to act in a pre-heated humidity chamber for 100 min at 37 °C. Two last final wash steps were performed, submerging slides twice in wash buffer B for 10 min and then in 0.01 \times wash buffer B for 1 min. The slides were then dried in the dark, prepared for imaging by adding Duolink II Mounting Medium with DAPI, and visualized using a confocal microscope (Leica DM6000 B).

Western blotting

SDS-PAGE and western blot analyses were performed using standard protocols. The following primary antibodies were used: rabbit anti-TAP (CAB1001, Thermo Scientific-Pierce), anti-Myc Tag clone 4A6 (05-724, Millipore), rabbit anti-ER α (sc-543, Santa Cruz Biotechnology), mouse anti-AGO2/eIF2C2 (ab57113, Abcam), rabbit polyclonal to FXR1 (ab50841, Abcam), rabbit polyclonal to integrin beta 4 binding protein (EIF6; ab77298, Abcam), anti-PRPF8 antibody (ab79237, Abcam), mouse monoclonal anti-AGO1 clone 4G7-E12 (MABE143, Millipore), mouse anti- β -actin (A1978, Sigma Aldrich), mouse monoclonal to Dicer (ab14601, Abcam), anti-TRBP2 (H-57; sc-292550, Santa Cruz Biotechnology).

AGO2 knock-down

For nascent-Seq experiments, C-TAP-ER β and MCF-7 control cells were starved for 5 days and then AGO2 knock-down was performed using a combination of three pLKO.1 plasmid vectors expressing shRNAs (Sigma Aldrich: TRCN0000007864; TRCN0000007867; TRCN0000011203) targeting the AGO2 transcript (GenBankTM accession number NM_012154) in different regions. AGO2 silencing was conducted by co-transfecting C-TAP-ER β and control cells with shRNA vectors, using Lipofectamine 2000 (Life Technologies), for 48 h. The transfection medium was replaced with fresh culturing medium 6 h after treatment. Non-transfected and transfected cells with pLKO.1-puro Non-Target shRNA Control Plasmid DNA (Sigma-Aldrich) were used as control. For TAP/MS after AGO2 silencing, hormone-deprived Ct-ER β cells were transfected with SMARTvector human lentiviral shRNA pooled libraries (Dharmacon) for 72 h. Western blotting was performed to verify the level of 'knock-down' of the target protein.

Nascent RNA isolation, sequencing, and data analysis

Nascent RNA was extracted from each sample as described by Khodor et al. [64]. In brief, following TRIzol (Life Technologies) addition, samples were incubated at 65 °C to dissolve DNA-Histone-Pol II-RNA pellets and RNA was extracted following the manufacturer's protocol. For sequencing, indexed libraries were prepared using 1 μ g of Nascent RNA as starting material, with TruSeq Stranded Total RNA Sample Prep Kit (Illumina Inc.). Libraries were sequenced (paired-end, 2 \times 100 cycles) at a concentration of 8 pM/lane on the HiSeq 2500 platform (Illumina Inc.) [30].

Alignment to the human genome

Raw sequence files (.fastq files) underwent quality control analysis using FASTQC (<http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>) and the quality checked reads were then aligned to the human genome (assembly hg19) using TopHat version 2.0.10 [94], according to the criteria used by Menet et al. [109].

Quantification of gene signal

Quantification of nascent RNA was done as in Menet et al. [109]. Differentially expressed nascent RNAs were identified using DESeq2 [96]. The differential expression was reported as fold change [1.5] along with associated adjusted p values ($FDR \leq 0.05$) computed according to Benjamini–Hochberg.

Intron retention determination

Before proceeding with intronic quantification, we extracted intronic intervals as described by St Laurent et al. [110], while intron retention was

computed as described by Khodor et al. [64]. The statistical significance of intron retention events observed between the several conditions was assessed using t -test ($FDR < 0.05$). Raw data are deposited in the EBI ArrayExpress database (<http://www.ebi.ac.uk/arrayexpress>) with accession number EMTAB-4368.

RNA immunoprecipitation, sequencing, and data analysis

Cells were lysed with polysome lysis buffer, as described by Keene et al. [111]. An aliquot of whole-cell extract (10% of total) was taken as input. For ER β immunoprecipitation, samples were incubated at 4 °C for 3 h with 50 μ l of IgG Sepharose 6 fast Flow (GE Healthcare Bio-Science AB) pre-treated with NT2 buffer supplemented with 5% BSA. After binding, the isolation of RNA co-precipitated with ER β was carried out by adding TRIzol (Life Technologies) directly to the washed beads, following the manufacturer's instructions. For RNA-Seq analyses, indexed sequencing libraries were prepared starting from 1 μ g of RNA input and 300 ng of RNA immunoprecipitated, pooling three independent experiments (biological replicates) and using TruSeq Stranded Total RNA. For miRNA-Seq experiments, libraries were generated from 120 ng of the same pooled RNA using TruSeq Small RNA Sample Prep Kits (Illumina Inc.). Libraries were sequenced (single read 1 \times 50 cycles and 2 \times 100 cycles for miRNA- and RNA-Seq experiments, respectively) on a HiSeq 2500 (Illumina Inc.). Data analysis was performed as follows.

Alignment

Raw sequence files (.fastq files) underwent quality control analysis using FASTQC (<http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>) and the quality checked reads were then aligned to the human genome (assembly hg19) using TopHat version 2.0.10 [94]. HTSeq-count [112] was used to compute gene-level read counts.

Enrichment analysis

The read counts obtained were used as input to DESeq2 [96] to perform enrichment analysis. RNAs showing enrichment factor (EF) > 1 and adjusted p value ≤ 0.05 were considered for further analysis. To define enriched RNAs in Ct-ER β IP versus input RNAs, we applied a more stringent analysis: firstly, we selected the RNAs showing an EF more than 75th percentile of its distribution, and subsequently we compared these RNAs with those identified comparing wild-type IP β versus input RNAs. Hence, we selected those RNAs specific for the Ct-ER β IP β group, and those that, when compared to wild type, showed a ratio (Ct-ER β EF/wild-type EF) ≥ 4 , or showing a negative EF in wild-type IP β vs input. Small RNA-Seq data were analyzed using iSmaRT [113] with standard parameters, using miRBase v20 as reference track.

miRNAs showing $EF > 1$ and p value ≤ 0.05 were considered for further analysis. To select ER β -specific enriched miRNAs, those with $EF > 1.5$ were considered and the EF in ER β + cells was compared with the same in wild-type cells. miRNAs showing a ratio between the two conditions (i.e., Ct-ER β EF/wild-type EF) ≥ 2 or showing a negative EF in wild-type IP β vs input were selected. The different classes of small RNAs obtained in IP β in Ct-ER β cells were assessed using sRNA-Bench [114]. Classification of enriched RNAs was performed using the "Gene biotype" term in ENSEMBL using a.gtf file downloaded from Genecode (<http://www.genecodegenes.org/#>).

Functional and pathway analyses

Functional and interaction network analysis of ER β -associated proteins was performed with the FunRich tool [115] according to the user manual. The lists of transcripts were analyzed using Ingenuity Pathway Analysis Software (IPA, Ingenuity® Systems, www.ingenuity.com). It refers to a proprietary knowledge base (Ingenuity Pathways Knowledge Base) in which cellular molecules, biological interactions, and functional properties are annotated. IPA Functional Analysis on "molecular and cellular functions" category and Canonical Pathway investigation were carried out, calculating the likelihood that the association between our RNA dataset and a specific function or pathway is due to random choice, and it is expressed as a p value calculated using the right-tailed Fisher exact test. The activation z-score is used to infer likely activation states of enriched pathways, based on comparison with a model that assigns random regulation directions. Finally, the "microRNA Target Filter" IPA module was used to provide insights into the biological effects of microRNAs, using miRNA-mRNA interactions from TarBase and miRecords, as well as predicted miRNA-mRNA interactions from TargetScan examining miRNA-mRNA pairings in the pathways of interest. Finally, a network representing miRNA-RNA target interaction was created using Cytoscape [116].

Additional files

Additional file 1: Table S1. Transcriptome profiling of ER β expressing cells. **Table S1a** Expressed transcripts in Ct-ER β expressing cells. **Table S1b.** Expressed transcripts in Nt-ER β expressing cells. **Table S1c** Differentially expressed transcripts in Ct- and Nt-ER β expressing cells. **Table S1d.** Commonly differentially expressed transcripts in Ct- and Nt-ER β expressing cells. (XLSX 4816 kb)

Additional file 2: Table S2a. Alternative splicing events in Ct-ER β cells. (XLSX 850 kb)

Additional file 3: Table S2b. Alternative splicing events in Nt-ER β cells. (XLSX 706 kb)

Additional file 4: Table S2c. Differentially expressed transcripts coupled with splicing events. (XLSX 140 kb)

Additional file 5: Table S3. Mapping of ER β binding sites to the BC cell genome. **Table S3a.** ER β binding sites. **Table S3b.** Differentially expressed transcripts with ER β binding sites in promoter regions. **Table S3c** Differentially expressed transcripts harboring ER β binding sites in the transcriptional unit. (XLSX 1334 kb)

Additional file 6: Figures S1–S7. Supplementary figures with legends. (DOCX 3897 kb)

Additional file 7: Table S4. Proteins interacting with ER β in MCF-7 cell nuclei in the absence of estrogen stimuli (including Mascot search files). (XLSX 1010 kb)

Additional file 8: Table S5. Proteomics analysis of ER β interactome following AGO2 silencing. **Table S5a.** Raw data. **Table S5b.** Normalized data. **Table S5c.** Statistically significant changes. **Table S5d.** Not statistically significant changes. (XLSX 201 kb)

Additional file 9: Table S6. Mapping of AGO2 binding sites to the BC cell genome. **Table S6a** AGO binding sites in ER β -positive cells. **Table S6b.** AGO binding sites in ER β negative cells. (XLSX 232 kb)

Additional file 10: Table S7. AGO2 binding matrices. **Table S7a.** Motifs discovered among AGO2 binding sites in ER β -expressing cells.

Table S7b. Motifs discovered among AGO2 binding sites in wild-type cells. **Table S7c.** Motifs discovered among AGO2-ER β shared binding sites. (XLSX 16 kb)

Additional file 11: Table S8. ER β and AGO2 shared binding sites. (XLSX 39 kb)

Additional file 12: Table S9. Genes whose transcription rate is modulated by ER β and AGO2. **Table S9a.** Genes showing transcriptional regulation by ER β (Ct-ER β vs wild type). **Table S9b.** Genes responding to AGO2 silencing in ER β + cells (shAGO2 vs Ct-ER β). **Table S9c.** Genes showing transcriptional regulation by both ER β (Ct-ER β vs wild type) and AGO2 (shAGO2 vs Ct-ER β). **Table S9d.** Genes differentially expressed in Ct-ER β vs wild-type cells harboring both ER β and AGO2 binding sites and showing an inversion of the ER β -induced transcriptional trend after AGO2 silencing. (XLSX 1259 kb)

Additional file 13: Table S10. Nascent transcripts whose maturation is modulated by ER β and AGO2. **Table S10a.** Intron retention modulated by ER β (FDR ≤ 0.05 , t -test). **Table S10b.** Intron retention modulated by AGO2 (FDR ≤ 0.05 , t -test). (XLSX 3193 kb)

Additional file 14: Table S11. ER β -bound mRNAs. (XLSX 50 kb)

Additional file 15: Table S12. ER β -bound miRNAs. (XLSX 10 kb)

Additional file 16: Table S13. ER β -bound miRNAs targeting enriched mRNAs. (XLSX 17 kb)

Acknowledgements

GN was supported by a 'Mario e Valeria Rindi' fellowship of the Italian Foundation for Cancer Research.

Funding

Work supported by: Italian Association for Cancer Research (grants IG-17426), Italian Ministry for Education, University and Research (grant FIRB RBF12W5V5_003 to RT), Italian Ministry of Health (Young Researcher grants GR-2011-02347781 to GN and GR-2011-02350476 to MR), University of Salerno (Fondi FARB 2015-2016) and CNR (Flagship Project InterOmics). We also acknowledge ELIXIR-ItB (<http://elixir-italy.org/>), the Italian Node of the European ELIXIR infrastructure (<https://elixir-europe.org/>), for the computational power support provided.

Availability of data and materials

The sequencing datasets generated and analyzed during the current study are available in the EBI ArrayExpress database repository (<http://www.ebi.ac.uk/arrayexpress>) with accession numbers E-MTAB-4363, E-MTAB-4359, and E-MTAB-4368. The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier PXD006280.

Authors' contributions

All authors participated in conception and design of the study. RT, GB, GN, MR, FR, GM, AC, TR, GC, VG, and BP performed in vitro experimental work and RNA sequencing. AS and LR performed in vivo experimental work. TN and GN performed the proteomics analyses. GG performed the statistical and bioinformatics analyses. RT, GG, GN, CA, LM, and AW coordinated and finalized figure preparation, manuscript drafting, and revision. All authors read and approved the final manuscript.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Publisher's note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Author details

¹Laboratory of Molecular Medicine and Genomics, Department of Medicine, Surgery and Dentistry "Schola Medica Salernitana", University of Salerno, via S. Allende, 1, 84081 Baronissi, SA, Italy. ²Genomix4Life srl, Department of Medicine, Surgery and Dentistry "Schola Medica Salernitana", University of Salerno, Baronissi, SA, Italy. ³IRCCS SDN, Napoli, Italy. ⁴Department of Cardiothoracic and Respiratory Sciences, University of Campania "L. Vanvitelli", Naples, Italy. ⁵Institute of Biomedical Technologies, National Research Council, Segrate, MI, Italy. ⁶Department of Science and Technology, University of Sannio, Benevento, Italy. ⁷IRGS Biogem, Ariano Irpino, AV, Italy. ⁸Department of Immunology, Institute of Clinical Medicine, University of Oslo and Rikshospitalet Oslo, Oslo, Norway.

Received: 9 July 2017 Accepted: 20 September 2017

Published online: 06 October 2017

References

- Meister G, Landthaler M, Patkaniowska A, Dorsett Y, Teng G, Tuschl T. Human Argonaute2 mediates RNA cleavage targeted by miRNAs and siRNAs. *Mol Cell*. 2004;15:185–97.
- Rand TA, Ginalski K, Grishin NV, Wang X. Biochemical identification of Argonaute 2 as the sole protein required for RNA-induced silencing complex activity. *Proc Natl Acad Sci U S A*. 2004;101:14385–9.
- Liu J, Carmell MA, Rivas FV, Marsden CG, Thomson JM, Song JJ, et al. Argonaute2 is the catalytic engine of mammalian RNAi. *Science*. 2004;305:1437–41.
- Kwon SY, Lee JH, Kim B, Park JW, Kwon TK, Kang SH, Kim S. Complexity in regulation of microRNA machinery components in invasive breast carcinoma. *Pathol Oncol Res*. 2014;20:697–705.
- Fan M, Krutilina R, Sun J, Sethuraman A, Yang CH, Wu ZH, et al. Comprehensive analysis of microRNA (miRNA) targets in breast cancer cells. *J Biol Chem*. 2013;288:27480–93.
- Morris KV. RNA-mediated transcriptional gene silencing in human cells. *Curr Top Microbiol Immunol*. 2008;320:211–24.
- Janowski BA, Huffman KE, Schwartz JC, Ram R, Nordsell R, Shames DS, et al. Involvement of AGO1 and AGO2 in mammalian transcriptional silencing. *Nat Struct Mol Biol*. 2006;13:787–92.
- Kim DH, Saetrom P, Snove Jr O, Rossi JJ. MicroRNA-directed transcriptional gene silencing in mammalian cells. *Proc Natl Acad Sci U S A*. 2008;105:16230–5.
- Benhamed M, Herbig U, Ye T, Dejean A, Bischof O. Senescence is an endogenous trigger for microRNA-directed transcriptional gene silencing in human cells. *Nat Cell Biol*. 2012;14:266–75.
- Carissimi C, Laudadio I, Cipolletta E, Gioiosa S, Mihailovich M, Bonaldi T, et al. ARGONAUTE2 cooperates with SWI/SNF complex to determine nucleosome occupancy at human Transcription Start Sites. *Nucleic Acids Res*. 2015;43:1498–512.
- Taliaferro JM, Aspden JL, Bradley T, Marwha D, Blanchette M, Rio DC. Two new and distinct roles for Drosophila Argonaute-2 in the nucleus: alternative pre-mRNA splicing and transcriptional repression. *Genes Dev*. 2013;27:378–89.
- Ameyar-Zazoua M, Rachez C, Souidi M, Robin P, Fritsch L, Young R, et al. Argonaute proteins couple chromatin silencing to alternative splicing. *Nat Struct Mol Biol*. 2012;19:998–1004.
- Sharma NR, Wang X, Majerciak V, Ajiro M, Kruhlak M, Meyers C, Zheng ZM. Cell-type and tissue context-dependent nuclear distribution of human Ago2. *J Biol Chem*. 2016;291:2302–9.
- Damdimopoulos AE, Spyrou G, Gustafsson JA. Ligands differentially modify the nuclear mobility of estrogen receptors alpha and beta. *Endocrinology*. 2008;149:339–45.
- Maruvada P, Baumann CT, Hager GL, Yen PM. Dynamic shuttling and intranuclear mobility of nuclear hormone receptors. *J Biol Chem*. 2003;278:12425–32.
- Thomas C, Gustafsson JA. The different roles of ER subtypes in cancer biology and therapy. *Nat Rev Cancer*. 2011;11:597–608.
- Adams BD, Claffey KP, White BA. Argonaute-2 expression is regulated by epidermal growth factor receptor and mitogen-activated protein kinase signaling and correlates with a transformed phenotype in breast cancer cells. *Endocrinology*. 2009;150:14–23.
- Leung YK, Lee MT, Lam HM, Tarapore P, Ho SM. Estrogen receptor-beta and breast cancer: translating biology into clinical practice. *Steroids*. 2012;77:727–37.
- Hodges-Gallagher L, Valentine CD, El Bader S, Kushner PJ. Estrogen receptor beta increases the efficacy of antiestrogens by effects on apoptosis and cell cycling in breast cancer cells. *Breast Cancer Res Treat*. 2008;109:241–50.
- Murphy LC, Peng B, Lewis A, Davie JR, Leygue E, Kemp A, et al. Inducible upregulation of oestrogen receptor-beta1 affects oestrogen and tamoxifen responsiveness in MCF7 human breast cancer cells. *J Mol Endocrinol*. 2005;34:553–66.
- Huang B, Warner M, Gustafsson JA. Estrogen receptors in breast carcinogenesis and endocrine therapy. *Mol Cell Endocrinol*. 2015;418(Pt 3):240–4.
- Madeira M, Mattar A, Logullo AF, Soares FA, Gebirim LH. Estrogen receptor alpha/beta ratio and estrogen receptor beta as predictors of endocrine therapy responsiveness—a randomized neoadjuvant trial comparison between anastrozole and tamoxifen for the treatment of postmenopausal breast cancer. *BMC Cancer*. 2013;13:425.
- Grober OM, Mutarelli M, Giurato G, Ravo M, Cicatiello L, De Filippo MR, et al. Global analysis of estrogen receptor beta binding to breast cancer cell genome reveals an extensive interplay with estrogen receptor alpha for target gene regulation. *BMC Genomics*. 2011;12:36.
- Vivar OI, Zhao X, Saunier EF, Griffin C, Mayba OS, Tagliaferri M, et al. Estrogen receptor beta binds to and regulates three distinct classes of target genes. *J Biol Chem*. 2010;285:22059–66.
- Nassa G, Tarallo R, Giurato G, De Filippo MR, Ravo M, Rizzo F, et al. Post-transcriptional regulation of human breast cancer cell proteome by unliganded estrogen receptor beta via microRNAs. *Mol Cell Proteomics*. 2014;13:1076–90.
- Zhao C, Lam EW, Sunters A, Enmark E, De Bella MT, Coombes RC, et al. Expression of estrogen receptor beta isoforms in normal breast epithelial cells and breast cancer: regulation by methylation. *Oncogene*. 2003;22:7600–6.
- Chang EC, Frasar J, Komm B, Katzenellenbogen BS. Impact of estrogen receptor beta on gene networks regulated by estrogen receptor alpha in breast cancer cells. *Endocrinology*. 2006;147:4831–42.
- Lin CY, Strom A, Li Kong S, Kietz S, Thomsen JS, Tee JB, et al. Inhibitory effects of estrogen receptor beta on specific hormone-responsive gene expression and association with disease outcome in primary breast cancer. *Breast Cancer Res*. 2007;9:R25.
- Williams C, Edvardsson K, Lewandowski SA, Strom A, Gustafsson JA. A genome-wide study of the repressive effects of estrogen receptor beta on estrogen receptor alpha signaling in breast cancer cells. *Oncogene*. 2008;27:1019–32.
- Dago DN, Scafoglio C, Rinaldi A, Memoli D, Giurato G, Nassa G, et al. Estrogen receptor beta impacts hormone-induced alternative mRNA splicing in breast cancer cells. *BMC Genomics*. 2015;16:367.
- Shen S, Park JW, Huang J, Dittmar KA, Lu ZX, Zhou Q, et al. MATS: a Bayesian framework for flexible detection of differential alternative splicing from RNA-Seq data. *Nucleic Acids Res*. 2016;291:2302–9.
- Paris O, Ferraro L, Grober OM, Ravo M, De Filippo MR, Giurato G, et al. Direct regulation of microRNA biogenesis and expression by estrogen receptor beta in hormone-responsive breast cancer. *Oncogene*. 2012;31:4196–206.

33. Zhang Y, Liu T, Meyer CA, Eeckhoutte J, Johnson DS, Bernstein BE, et al. Model-based analysis of ChIP-Seq (MACS). *Genome Biol.* 2008;9:R137.
34. Jalili V, Matteucci M, Morelli MJ, Masseroli M. MuSERA: Multiple Sample Enriched Region Assessment. *Brief Bioinform.* 2017;18:367–81.
35. Bado I, Nikolos F, Rajapaksa G, Gustafsson JA, Thomas C. ERbeta decreases the invasiveness of triple-negative breast cancer cells by regulating mutant p53 oncogenic function. *Oncotarget.* 2016;7:13599–611.
36. Carter D, Chakalova L, Osborne CS, Dai YF, Fraser P. Long-range chromatin regulatory interactions in vivo. *Nat Genet.* 2002;32:623–6.
37. Fullwood MJ, Liu MH, Pan YF, Liu J, Xu H, Mohamed YB, et al. An oestrogen-receptor-alpha-bound human chromatin interactome. *Nature.* 2009;462:58–64.
38. Korkmaz G, Lopes R, Ugalde AP, Nevedomskaya E, Han R, Myacheva K, et al. Functional genetic screens for enhancer elements in the human genome using CRISPR-Cas9. *Nat Biotechnol.* 2016;34:192–8.
39. Bourgo RJ, Singhal H, Greene GL. Capture of associated targets on chromatin links long-distance chromatin looping to transcriptional coordination. *Nat Commun.* 2016;7:12893.
40. Nassa G, Tarallo R, Ambrosino C, Bamundo A, Ferraro L, Paris O, et al. A large set of estrogen receptor beta-interacting proteins identified by tandem affinity purification in hormone-responsive human breast cancer cell nuclei. *Proteomics.* 2011;11:159–65.
41. Tarallo R, Bamundo A, Nassa G, Nola E, Paris O, Ambrosino C, et al. Identification of proteins associated with ligand-activated estrogen receptor alpha in human breast cancer cell nuclei by tandem affinity purification and nano LC-MS/MS. *Proteomics.* 2011;11:172–9.
42. Cirillo F, Nassa G, Tarallo R, Stellato C, De Filippo MR, Ambrosino C, et al. Molecular mechanisms of selective estrogen receptor modulator activity in human breast cancer cells: identification of novel nuclear cofactors of antiestrogen-ERalpha complexes by interaction proteomics. *J Proteome Res.* 2013;12:421–31.
43. Stellato C, Nassa G, Tarallo R, Giurato G, Ravo M, Rizzo F, et al. Identification of cytoplasmic proteins interacting with unliganded estrogen receptor alpha and beta in human breast cancer cells. *Proteomics.* 2015;15:1801–7.
44. Hutvagner G, Simard MJ. Argonaute proteins: key players in RNA silencing. *Nat Rev Mol Cell Biol.* 2008;9:22–32.
45. Hock J, Meister G. The Argonaute protein family. *Genome Biol.* 2008;9:210.
46. Kalathur RK, Pinto J, Hernandez-Prieto MA, Machado RS, Almeida D, Chaurasia G, Futschik ME. UniHI 7: an enhanced database for retrieval and interactive analysis of human molecular interaction networks. *Nucleic Acids Res.* 2014;42:D408–14.
47. Cowley MJ, Pinese M, Kassahn KS, Waddell N, Pearson JV, Grimmond SM, et al. PINA v2.0: mining interactome modules. *Nucleic Acids Res.* 2012;40:D862–5.
48. Hock J, Weinmann L, Ender C, Rudel S, Kremmer E, Raabe M, et al. Proteomic and functional analysis of Argonaute-containing mRNA-protein complexes in human cells. *EMBO Rep.* 2007;8:1052–60.
49. Cho S, Park JS, Kang YK. AGO2 and SETDB1 cooperate in promoter-targeted transcriptional silencing of the androgen receptor gene. *Nucleic Acids Res.* 2014;42:13545–56.
50. Clocchiatti A, Di Giorgio E, Viviani G, Streuli C, Sgorbissa A, Picco R, et al. The MEF2-HDAC axis controls proliferation of mammary epithelial cells and acini formation in vitro. *J Cell Sci.* 2015;128:3961–76.
51. Zhou Y, Huang T, Cheng AS, Yu J, Kang W, To KF. The TEAD family and its oncogenic role in promoting tumorigenesis. *Int J Mol Sci.* 2016;17(1):138.
52. Zang H, Li N, Pan Y, Hao J. Identification of upstream transcription factors (TFs) for expression signature genes in breast cancer. *Gynecol Endocrinol.* 2017;33:193–98.
53. Quang CT, Leboucher S, Passaro D, Fuhrmann L, Nourieh M, Vincent-Salomon A, Ghysdael J. The calcineurin/NFAT pathway is activated in diagnostic breast cancer cases and is essential to survival and metastasis of mammary cancer cells. *Cell Death Dis.* 2015;6:e1658.
54. Jauliac S, Lopez-Rodriguez C, Shaw LM, Brown LF, Rao A, Toker A. The role of NFAT transcription factors in integrin-mediated carcinoma invasion. *Nat Cell Biol.* 2002;4:540–4.
55. Watson CJ. Stat transcription factors in mammary gland development and tumorigenesis. *J Mammary Gland Biol Neoplasia.* 2001;6:115–27.
56. Swingen T, Tonissen KF. Altered HOX gene expression in human skin and breast cancer cells. *Cancer Biol Ther.* 2003;2:518–23.
57. Dravis C, Spike BT, Harrell JC, Johns C, Trejo CL, Southard-Smith EM, et al. Sox10 regulates stem/progenitor and mesenchymal cell states in mammary epithelial cells. *Cell Rep.* 2015;12:2035–48.
58. Heger A, Webber C, Goodson M, Ponting CP, Lunter G. GAT: a simulation framework for testing the association of genomic intervals. *Bioinformatics.* 2013;29:2046–8.
59. Wang X, Guda C. Integrative exploration of genomic profiles for triple negative breast cancer identifies potential drug targets. *Medicine (Baltimore).* 2016;95:e4321.
60. Elsarraj HS, Hong Y, Valdez KE, Michaels W, Hook M, Smith WP, et al. Expression profiling of in vivo ductal carcinoma in situ progression models identified B cell lymphoma-9 as a molecular driver of breast cancer invasion. *Breast Cancer Res.* 2015;17:128.
61. Zatulá N, Wiese M, Bunzendorf J, Birchmeier W, Perske C, Bleckmann A, Brembeck FH. The BCL9-2 proto-oncogene governs estrogen receptor alpha expression in breast tumorigenesis. *Oncotarget.* 2014;5:6770–87.
62. Ip JY, Schmidt D, Pan Q, Ramani AK, Fraser AG, Odom DT, Blencowe BJ. Global impact of RNA polymerase II elongation inhibition on alternative splicing regulation. *Genome Res.* 2011;21:390–401.
63. Das R, Dufu K, Romney B, Feldt M, Elenko M, Reed R. Functional coupling of RNAP II transcription to spliceosome assembly. *Genes Dev.* 2006;20:1100–9.
64. Khodor YL, Rodriguez J, Abruzzi KC, Tang CH, Marr 2nd MT, Rosbash M. Nascent-seq indicates widespread cotranscriptional pre-mRNA splicing in *Drosophila*. *Genes Dev.* 2011;25:2502–12.
65. Siraj AK, Beg S, Jehan Z, Prabhakaran S, Ahmed M, RH A, et al. ALK alteration is a frequent event in aggressive breast cancers. *Breast Cancer Res.* 2015;17:127.
66. Bardin A, Boule N, Lazennec G, Vignon F, Pujol P. Loss of ERbeta expression as a common step in estrogen-dependent tumor progression. *Endocr Relat Cancer.* 2004;11:537–51.
67. Forster C, Makela S, Warri A, Kietz S, Becker D, Hultenby K, et al. Involvement of estrogen receptor beta in terminal differentiation of mammary gland epithelium. *Proc Natl Acad Sci U S A.* 2002;99:15578–83.
68. Krege JH, Hodgins JB, Couse JF, Enmark E, Warner M, Mahler JF, et al. Generation and reproductive phenotypes of mice lacking estrogen receptor beta. *Proc Natl Acad Sci U S A.* 1998;95:15677–82.
69. Ambrosino C, Tarallo R, Bamundo A, Cuomo D, Franci G, Nassa G, et al. Identification of a hormone-regulated dynamic nuclear actin network associated with estrogen receptor alpha in human breast cancer cell nuclei. *Mol Cell Proteomics.* 2010;9:1352–67.
70. Chen Y, Boland A, Kuzuoglu-Ozturk D, Bawankar P, Loh B, Chang CT, et al. A DDX6-CNOT1 complex and W-binding pockets in CNOT9 reveal direct links between miRNA target recognition and silencing. *Mol Cell.* 2014;54:737–50.
71. Covington KR, Fuqua SA. Role of MTA2 in human cancer. *Cancer Metastasis Rev.* 2014;33:921–8.
72. Ning Z, Gan J, Chen C, Zhang D, Zhang H. Molecular functions and significance of the MTA family in hormone-independent cancer. *Cancer Metastasis Rev.* 2014;33:901–19.
73. Fumagalli D, Gacquer D, Rothe F, Lefort A, Libert F, Brown D, et al. Principles governing A-to-I RNA editing in the breast cancer transcriptome. *Cell Rep.* 2015;13:277–89.
74. Sudo H, Tsuji AB, Sugyo A, Kohda M, Sogawa C, Yoshida C, et al. Knockdown of COPA, identified by loss-of-function screen, induces apoptosis and suppresses tumor growth in mesothelioma mouse model. *Genomics.* 2010;95:210–6.
75. Watkin LB, Jessen B, Wiszniewski W, Vece TJ, Jan M, Sha Y, et al. COPA mutations impair ER-Golgi transport and cause hereditary autoimmune-mediated lung disease and arthritis. *Nat Genet.* 2015;47:654–60.
76. Kakiyama Y, Makhnevych T, Zhao L, Tang W, Houry WA. Nutritional status modulates box C/D snoRNP biogenesis by regulated subcellular relocalization of the R2TP complex. *Genome Biol.* 2014;15:404.
77. Zuo H, Wong YH. Association of activated Galphaq to the tumor suppressor Fhit is enhanced by phospholipase Cbeta. *BMC Cancer.* 2015;15:775.
78. Dorsam RT, Gutkind JS. G-protein-coupled receptors and cancer. *Nat Rev Cancer.* 2007;7:79–94.
79. Golebiewska U, Guo Y, Khalikapradas N, Zurawsky C, Yerramilli VS, Scarlata S. gamma-Synuclein interacts with phospholipase Cbeta2 to modulate G protein activation. *PLoS One.* 2012;7:e41067.
80. Beristain AG, Molyneux SD, Joshi PA, Pomroy NC, Di Grappa MA, Chang MC, et al. PKA signaling drives mammary tumorigenesis through Src. *Oncogene.* 2015;34:1160–73.
81. de Leeuw R, Flach K, Bentin Toaldo C, Alexi X, Canisius S, Neeffes J, et al. PKA phosphorylation redirects ERalpha to promoters of a unique gene set to induce tamoxifen resistance. *Oncogene.* 2013;32:3543–51.

82. Allo M, Buggiano V, Fededa JP, Petrillo E, Schor I, de la Mata M, et al. Control of alternative splicing through siRNA-mediated transcriptional gene silencing. *Nat Struct Mol Biol.* 2009;16:717–24.
83. Liu MH, Cheung E. Estrogen receptor-mediated long-range chromatin interactions and transcription in breast cancer. *Mol Cell Endocrinol.* 2014;382:624–32.
84. Moshkovich N, Nisha P, Boyle PJ, Thompson BA, Dale RK, Lei EP. RNAi-independent role for Argonaute2 in CTCF/CP190 chromatin insulator function. *Genes Dev.* 2011;25:1686–701.
85. Holwerda SJ, de Laat W. CTCF: the protein, the binding partners, the binding sites and their chromatin loops. *Philos Trans R Soc Lond B Biol Sci.* 2013;368:20120369.
86. Jang GB, Kim JY, Cho SD, Park KS, Jung JY, Lee HY, et al. Blockade of Wnt/beta-catenin signaling suppresses breast cancer metastasis by inhibiting CSC-like phenotype. *Sci Rep.* 2015;5:12465.
87. Zardawi SJ, O'Toole SA, Sutherland RL, Musgrove EA. Dysregulation of Hedgehog, Wnt and Notch signalling pathways in breast cancer. *Histol Histopathol.* 2009;24:385–98.
88. Dey N, Barwick BG, Moreno CS, Ordanic-Kodani M, Chen Z, Oprea-Ilie G, et al. Wnt signaling in triple negative breast cancer is associated with metastasis. *BMC Cancer.* 2013;13:537.
89. Ashaie MA, Chowdhury EH. Cadherins: the superfamily critically involved in breast cancer. *Curr Pharm Des.* 2016;22:616–38.
90. Ye Z, Jin H, Qian Q. Argonaute 2: a novel rising star in cancer research. *J Cancer.* 2015;6:877–82.
91. Yang L, Lin C, Jin C, Yang JC, Tanasa B, Li W, et al. lncRNA-dependent mechanisms of androgen-receptor-regulated gene activation programs. *Nature.* 2013;500:598–602.
92. Lanz RB, McKenna NJ, Onate SA, Albrecht U, Wong J, Tsai SY, et al. A steroid receptor coactivator, SRA, functions as an RNA and is present in an SRC-1 complex. *Cell.* 1999;97:17–27.
93. Ghosh SK, Patton JR, Spanjaard RA. A small RNA derived from RNA coactivator SRA blocks steroid receptor signaling via inhibition of Pusp-mediated pseudouridylation of SRA: evidence of a novel RNA binding domain in the N-terminus of steroid receptors. *Biochemistry.* 2012;51:8163–72.
94. Trapnell C, Pachter L, Salzberg SL. TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics.* 2009;25:1105–11.
95. Trapnell C, Roberts A, Goff L, Pertea G, Kim D, Kelley DR, et al. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nat Protoc.* 2012;7:562–78.
96. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 2014;15:550.
97. Anders S, Huber W. Differential expression analysis for sequence count data. *Genome Biol.* 2010;11:R106.
98. Schmidt D, Wilson MD, Spyrou C, Brown GD, Hadfield J, Odom DT. ChIP-seq: using high-throughput sequencing to discover protein-DNA interactions. *Methods.* 2009;48:240–8.
99. Langmead B, Trapnell C, Pop M, Salzberg SL. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* 2009;10:R25.
100. Heinz S, Benner C, Spann N, Bertolino E, Lin YC, Laslo P, et al. Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol Cell.* 2010;38:576–89.
101. Ye T, Krebs AR, Choukallah MA, Keime C, Plewniak F, Davidson I, Tora L. seqMINER: an integrated ChIP-seq data interpretation platform. *Nucleic Acids Res.* 2011;39:e35.
102. Zambelli F, Pesole G, Pavesi G. ScanChIP: Finding over-represented transcription factor-binding site motifs and their correlations in sequences from ChIP-Seq experiments. *Nucleic Acids Res.* 2013;41:W535–43.
103. Medina-Rivera A, Defrance M, Sand O, Herrmann C, Castro-Mondragon JA, Delerice J, et al. RSAT 2015: regulatory sequence analysis tools. *Nucleic Acids Res.* 2015;43:W50–6.
104. Machanick P, Bailey TL. MEME-ChIP: motif analysis of large DNA datasets. *Bioinformatics.* 2011;27:1696–7.
105. Cartharius K, Frech K, Grote K, Klocke B, Haltmeier M, Klingenhoff A, et al. MatInspector and beyond: promoter analysis based on transcription factor binding sites. *Bioinformatics.* 2005;21:2933–42.
106. Quinlan AR. BEDTools: the Swiss-Army tool for genome feature analysis. *Curr Protoc Bioinformatics.* 2014;47:11. 12 1–34.
107. Chen TW, Li HP, Lee CC, Gan RC, Huang PJ, Wu TH, et al. ChIPseeker, a web-based analysis tool for ChIP data. *BMC Genomics.* 2014;15:539.
108. Vizzaino JA, Csordas A, Del-Toro N, Dianes JA, Griss J, Lavidas I, et al. 2016 update of the PRIDE database and its related tools. *Nucleic Acids Res.* 2016;44:11033.
109. Menet JS, Rodriguez J, Abruzzi KC, Rosbash M. Nascent-Seq reveals novel features of mouse circadian transcriptional regulation. *Elife.* 2012;1:e00011.
110. St Laurent G, Shtokalo D, Tackett MR, Yang Z, Eremina T, Wahlestedt C, et al. Intronic RNAs constitute the major fraction of the non-coding RNA in mammalian cells. *BMC Genomics.* 2012;13:504.
111. Keene JD, Komisarow JM, Friedersdorf MB. RIP-Chip: the isolation and identification of mRNAs, microRNAs and protein components of ribonucleoprotein complexes from cell extracts. *Nat Protoc.* 2006;1:302–7.
112. Anders S, Pyl PT, Huber W. HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics.* 2015;31:166–9.
113. Panero R, Rinaldi A, Memoli D, Nassa G, Ravo M, Rizzo F, et al. iSmaRT: a toolkit for a comprehensive analysis of small RNA-Seq data. *Bioinformatics.* 2017;33:938–40.
114. Rueda A, Barturen G, Lebron R, Gomez-Martin C, Alganza A, Oliver JL, Hackenberg M. sRNAtoolbox: an integrated collection of small RNA research tools. *Nucleic Acids Res.* 2015;43:W467–73.
115. Pathan M, Keerthikumar S, Ang CS, Gangoda L, Quek CY, Williamson NA, et al. FunRich: An open access standalone functional enrichment and interaction network analysis tool. *Proteomics.* 2015;15:2597–601.
116. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* 2003;13:2498–504.

Submit your next manuscript to BioMed Central and we will help you at every step:

- We accept pre-submission inquiries
- Our selector tool helps you to find the most relevant journal
- We provide round the clock customer support
- Convenient online submission
- Thorough peer review
- Inclusion in PubMed and all major indexing services
- Maximum visibility for your research

Submit your manuscript at
www.biomedcentral.com/submit

