Cell cycle dependent oscillatory expression of estrogen receptor-α links Pol II elongation to neoplastic transformation

Cristina Vantaggiato[a], Marta Tocchetti[a], Vera Cappelletti[b], Aymone Gurtner[c], Alessandro Villa[a], Maria Grazia Daidone[b], Giulia Piaggio[c], Adriana Maggi[a,1], and Paolo Ciana[a,1,2]

[a]Center of Excellence on Neurodegenerative Diseases and Department of Pharmacological Sciences, University of Milan, 20133 Milan, Italy; [b]Department of Experimental Oncology and Molecular Medicine, Fondazione IRCCS Istituto Nazionale dei Tumori, 20133 Milan, Italy; and [c]Experimental Oncology Department, Istituto Regina Elena, 00158 Rome, Italy

Edited by Jan-Åke Gustafsson, University of Houston, Houston, Texas, and approved May 27, 2014 (received for review November 21, 2013)

Decades of studies provided a detailed view of the mechanism of estrogen receptor-α (ERα) regulated gene transcription and the physio-pathological relevance of the genetic programs controlled by this receptor in a variety of tissues. However, still limited is our knowledge on the regulation of ERα synthesis. Preliminary observations showed that the expression of ERα is cell cycle regulated. Here, we have demonstrated that a well described polymorphic sequence in the first intron of *ERα* (*Pvu*II and *Xba*I) has a key role in regulating the ERα content in cycling cells. We have shown that the RNA Pol II (Pol II) elongation is blocked at the polymorphic site and that the proto-oncogene c-MYB modulates the release of the pausing polymerase. It is well known that the two SNPs are associated to an increased risk, progression, survival and mortality of endocrine-related cancers, here we have demonstrated that the c-MYB-dependent release of Pol II at a specific phase of the cell cycle is facilitated by the px haplotype, thus leading to a higher ERα mitogenic signal. In breast cancer, this mechanism is disrupted when the hormone refractory phenotype is established; therefore, we propose this oscillator as a novel target for the development of therapies aimed at sensitizing breast cancer resistant to hormonal treatments. Because *Pvu*II and XbaI were associated to a broad range physio-pathological conditions beside neoplastic transformation, we expect that the ERα oscillator contributes to the regulation of the estrogen signal in several tissues.

steroid receptors | genetic polymorphisms | oscillatory gene expression | polymerase blockage | transcriptional oscillators

Estrogen receptor-α (ERα) is a member of the intracellular receptor superfamily expressed in a wide variety of tissues where it mediates a multiplicity of estrogen effects on reproductive, skeletal, vascular and metabolic systems. *ERα* human gene is dispersed over a region of more than 250 Kb in the long arm of chromosome 6 where hundreds of single nucleotide polymorphisms (SNPs) were identified. Among these, the restriction fragment length polymorphisms (RFLPs) *Pvu*II (p/P or IVS-397C/T or rs2234693, intron 1 nucleotide position 33,836) and *Xba*I (x/X or IVS-351A/G or rs9340799, intron 1 nucleotide position 33,882) were reported to be genetically linked to several phenotypic traits and disorders influenced by estrogens. For hormone-dependent cancers, association studies demonstrated linkage of these polymorphisms to: (*i*) breast cancer risk (1), survival (2), and mortality associated to hormone treatment (3); (*ii*) insurgence of ductal type carcinoma (4); (*iii*) age of breast cancer onset (5–7); (*iv*) predisposing factors (7), such as mammographic density (8), age of menarche, and menopause (9); and (*v*) endometrial (10) and prostate cancer risks (11). The functional role of *Pvu*II and *Xba*I SNPs in affecting these pathological phenotypes remains to be elucidated. The location of these polymorphisms at the intron 1/exon 2 boundary suggests their potential involvement in the regulation of *ERα* transcription. Indeed, the sequence spanning the *Pvu*II polymorphism was demonstrated in vitro to bind the proto-oncogene c-MYB therein proposed as a potential regulator of ERα expression (12). However, others molecular genetics studies could not conclusively demonstrate an association between the *Pvu*II site and receptor contents in human breast cancers so far (1, 13) or even obtained opposite results (14). However, ERα status is a valuable prognostic and predictive biomarker in the management of breast cancer patients (15). It is current belief that ERα mediates E2-induced growth in endocrine cancers, thus, the receptor synthesis is a plausible checkpoint in neoplastic transformation. The aim of the present study was to assess the extent to which *Pvu*II and *Xba*I control ERα synthesis and how this is critical for neoplastic transformation. Our work, by decrypting the functional role of *Pvu*II and *Xba*I polymorphic sites, provide an unexpected link among control of Pol II elongation, oscillatory gene expression and hormonal tumorigenesis.

Results

Blockage of Transcription Elongation at the First Intron of ERα. To gain insight on *Pvu*II/XbaI involvement in the regulation of *ERα* transcription, we carried out a series of run on assays in MCF-7 breast cancer cells. We used as run on probes either the entire *ERα* coding sequence or five DNA fragments mapping within the receptor genomic locus in sense and antisense orientations. No

Significance

***Pvu*II and *Xba*I polymorphisms at the estrogen receptor-α (*ERα*) intron 1 are associated with several estrogen-dependent pathologies including cancers. Here, we demonstrate that these SNPs map within a pausing site of Pol II elongation, a block that is cyclically released producing an oscillatory expression of ERα during cell division. The protoncogene c-MYB is part of this oscillator modulating the amplitude of ERα fluctuations in the presence of different SNP haplotypes. During breast cancer progression toward a hormone refractory phenotype, Pol II pausing contributes to turn off ERα synthesis highlighting the clinical relevance of the mechanism. Furthermore, the wide range of physio-pathological conditions associated with these SNPs, indicates this oscillator as a novel level of regulation of the hormonal signal throughout the body.**

Author contributions: M.G.D., G.P., A.M., and P.C. designed research; C.V., M.T., V.C., A.G., A.V., and P.C. performed research; C.V., M.T., V.C., A.G., A.V., M.G.D., G.P., A.M., and P.C. analyzed data; and C.V., M.T., A.M., and P.C. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Freely available online through the PNAS open access option.

[1]A.M. and P.C. contributed equally to this work.

[2]To whom correspondence should be addressed. E-mail: paolo.ciana@unimi.it.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1321750111/-/DCSupplemental.

antisense transcript was found. Sense transcripts were detected in exon 1 and intron 1, but, not in exon 2, 4 and 8 (Fig. 1). Similar results were obtained in T47D, another breast cancer cell line (*SI Appendix*, Fig. S1). As control of hybridization capability of the probes, run on experiments were also done on nuclei of SK-ER3 cells constitutively expressing a full length *ERα* cDNA: in this case, all probes were equally able to detect the corresponding transcripts (*SI Appendix*, Fig. S1*C*). These data suggested the presence of an arrest of Pol II transcriptional elongation at intron 1. To map the site responsible for the hypothesized Pol II arrest, we used semiquantitative real-time PCR to measure the relative concentrations of nuclear transcripts containing selected intron 1 sequences in MCF-7 cells (Fig. 2 *A* and *B*). Consistent with the run on data (Fig. 1), transcripts carrying the 5′ intron 1 region (*start* amplicon) were found to be significantly more represented (33-fold) than those carrying the 3′ (*end* amplicon). The levels of nuclear RNA containing the amplicon located immediately upstream the SNPs (*5′-SNP*) were also significantly higher (14-fold) than the transcripts containing the *end* amplicon, pointing to a major block of transcription within the 532-bp sequence containing the two RFLPs (Fig. 2*B*). Similar enrichment of the amplicons 5′ the SNPs were found in human normal and neoplastic endometrium (Fig. 2*C*). In accordance with these observations, Hah and collaborators reported accumulation of nuclear transcripts at the intron 1-exon 2 boundary using a global run on and sequencing approach (16). These results prompted us to analyze the nuclear transcripts from a larger biological sample size including low-grade ($n = 11$) and high-grade ($n = 11$) human breast cancers (*SI Appendix*, Table S1). To obtain an absolute measure of the primary transcripts containing the *5′-SNP* and *end* amplicons, we carried out a quantitative PCR (qPCR) using the standard curve procedure (*SI Appendix*, Fig. S2*A*) to measure the transcriptional blockage in breast cancer cell lines (*SI Appendix*, Fig. S2*B*) and in human breast samples (Fig. 2*D* and *SI Appendix*, Table S1). Pol II pausing was clearly present in the cell lines and in about 80% of the tumor samples. The magnitude of the pausing ratio appeared to be significantly more pronounced in high grade than in low grade tumors (Fig. 2*D*), although quite variable. To normalize the genetic background variability, we measured the pausing ratio within the same patient in the primary tumor and in the subsequent relapse. Seven pairs of ERα-positive primary tumors/ERα-negative relapses and two pairs of ERα-positive primary tumors/ERα-positive relapses were analyzed (*SI Appendix*, Table S2): we observed a major, statistically significant, increase (from 10 to more than 100 times) of the Pol II blockage (Fig. 2*E* and *SI Appendix*, Table S2) during relapse to ERα-negative tumors, whereas only a nonsignificant trend to increase was observed in the ERα-positive recurrences, suggesting that this mechanism may play a role in the development of ER negative, hormone-refractory lesions.

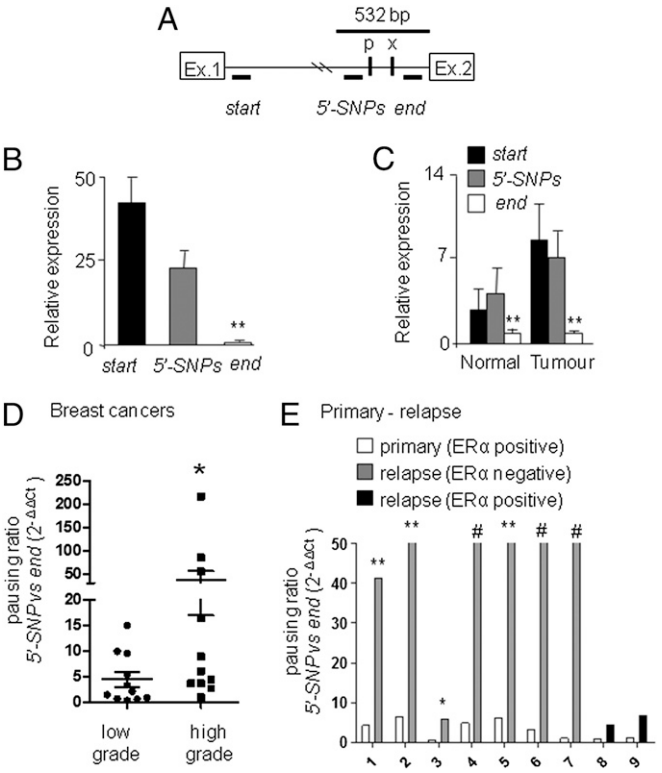

Fig. 2. The Pol II block of transcription involves the *Pvu*II and *Xba*I polymorphisms and is modulated during breast cancer progression. (*A*) Scheme of the real-time PCR amplicons mapping within intron 1; P = *Pvu*II and x = *Xba*1 RFLPs. (*B*) Relative amount of the primary transcripts containing the amplified sequences, bars are the average values ± mean SE (SEM) of the fold enrichment versus the *end* amplicons. $^{**}P < 0.01$ both *start* and the *5′-SNPs versus* the *end* amplicon. (*C*) Same as (*B*) on RNA samples obtained from normal and neoplastic human endometrium. $^{**}P < 0.01$ both *start* or *5′-SNPs versus* the *end* amplicon. (*D*) Pausing is calculated as the ratio between the absolute number of transcripts containing the *5′SNP* and *end* amplicons measured by quantitative real-time PCR from low and high grade breast cancers. Bars are average values ± SEM $^{*}P < 0.05$ calculated on the pausing ratio of high versus low grade tumors with the Mann Whitney test. (*E*) The pausing ratio defined as in (*D*) in primary breast cancers (white bars) and relapses (dark gray or black bars) from n. 9 individuals; bars are averages values ± SEM. #, in these cases, the *end* amplicon could not be detected, the pausing ratio was higher than 50 times; $^{**}P < 0.01$ pausing rate of relapse versus primary breast tumors. All data were collected in at least two independent experiments performed in triplicate. Student *t* test was applied to determine statistical significance.

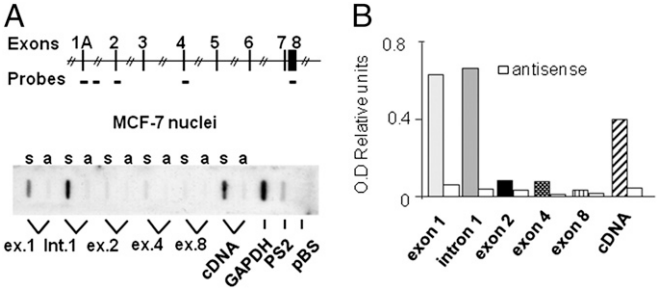

Fig. 1. Block to transcription elongation at *ERα* intron 1. (*A*) The inserted scheme represents the human *ERα* locus with the genomic position of the run on probes (*Upper*). Autoradiography of the run on assay on MCF-7 cells (*Lower*). Probe identity is indicated; s, single strand probes detecting sense transcripts; a, single strand probes antisense transcripts. (*B*) Optical density (OD) analysis of the autoradiographic signals; bars represent the signals derived from sense transcripts normalized on the probe length vs. GAPDH signal; white bars represent the signal generated by antisense transcripts.

Release of Pol II Blockage Is Regulated During Cell Cycle and Correlates with ERα Oscillatory Expression. Oscillations of ERα expression is known to occur during cell division (17–19). In accordance to receptor fluctuation, our immunocytochemistry studies in clones of MCF-7 showed a heterogeneous distribution of the receptor, indicating that ERα expression could be cell cycle-modulated (*SI Appendix*, Fig. S3*A*); interestingly, this heterogeneity is also often found in human endometrial and breast cancer specimens and is used as a parameter for scoring tissues for patient classification (15). We postulated that the mechanism of Pol II block and release involving the intronic region could be responsible of the ERα oscillatory expression during cell cycle. By arresting the cell cycle at different phases, we initially showed that the ERα content significantly decreased at the G1/S transition (*SI Appendix*, Fig. S3*B*). To investigate the kinetics of ERα fluctuation, the experiment was repeated by measuring ERα expression at different times after removal of the synchronizing agent. Upon thymidine removal, 60–70% of the cells synchronously reached the S-phase and returned in G1 20–24 h later as

demonstrated by cytofluorimetry (Fig. 3). ERα mRNA was low during the entire S phase and significantly increased in G2/M and G1 after thymidine removal (Fig. 3*A*); consistent with this observation, ERα protein was low in S-phase, increased in G1 and again decreased in the second S phase as demonstrated by Western blot and immunocytochemistry studies (Fig. 3 *B* and *C* and *SI Appendix*, Fig. S3*C*). Similar oscillatory mRNA and protein expression was observed in synchronized ZR 75.1 cells (*SI Appendix*, Fig. S4). To correlate the Pol II block with the ERα mRNA fluctuations during the cell cycle, we carried run on assays with nuclei of MCF-7 cells at different cell cycle phases (Fig. 4): In G1/S, transcription was barely detectable along the whole *locus*; the transcription blockage observed in S phase was released in late S/G2 phase (13 h after thymidine withdrawal) and resumed 3 h later in G2/M phase. These data showed that the production of a full-length ERα mRNA occurred only for a short time window during the 3–6 h interval in late S/G2 phase. These data suggested that the transient release of the Pol II block from intron 1 initiates the oscillatory wave of ERα expression leading to a peak of receptor levels at G1.

An Evolutionary Conserved 76-bp Sequence Regulates Pol II Elongation During Cell Cycle. Among the 532-bp sequence of intron 1 responsible of the Pol II elongation block (Fig. 2 *A* and *B*), a minimal 76-bp sequence (*76inter1*; i.e., 76 bp intervening sequence 1), containing the two SNPs was found evolutionary conserved (*SI Appendix*, Fig. S5*A*) and therefore was tested for its ability to regulate the Pol II release during cell cycle. To this aim, the *76inter1* sequence was cloned downstream the SV40 promoter driving the luciferase reporter gene (*SI Appendix*, Fig. S5*B*) and tested in transient transfection experiments: the *76inter1* fragment significantly reduced the reporter expression (*SI Appendix*, Fig. S5*B*). Next, we asked whether the *76inter1* was responsible for the regulation of Pol II elongation occurring during cell cycle (Fig. 4), and what was the role of different *Pvu*II/*Xba*I haplotypes in this process. The effects of each of the four haplotypes (Fig. 5*A*) were tested in thymidine synchronized MCF-7 cells transfected with the four constructs. We found that Luciferase transcription was low in S-phase for all constructs, but increased up to sevenfold (for the px containing construct) in late S/G2 phase to decrease at baseline levels at the beginning of G1 phase (Fig. 5*B*). The highest transcriptional

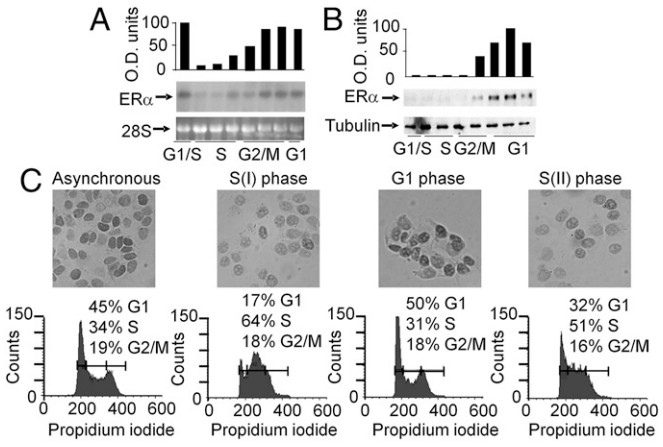

Fig. 3. Intracellular ERα concentration fluctuates during the cell cycle. After thymidine washout cells were harvested at different cell cycle phases. ERα mRNA and protein content were measured by Northern blot (*A*), Western blot (*B*), and immunocytochemistry analysis (*C*). Bars in the graphs represent ERα mRNA levels normalized with 28S ribosomal RNA (*A*) and ERα protein content normalized with tubulin (*B*). Similar data were obtained at least in two independent experiments. (*C*) Immunocytochemistry was carried out as in (*A*) on cells at the indicated phases of the cell cycle: asynchronous cells (As), first S (I), G1, and second S (II). Cell cycle phases were determined by cytofluorometry.

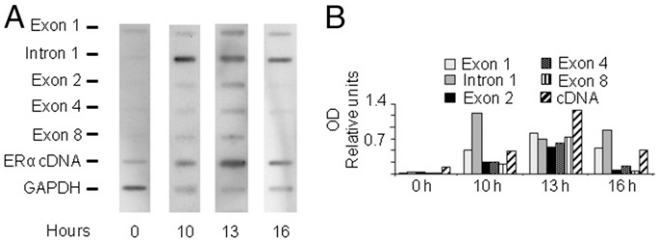

Fig. 4. Pol II block is released in the late S/G2 phase of the cell cycle. (*A*) Representative autoradiography of filters hybridized with ^{32}P-labeled primary transcripts obtained from nuclei of MCF-7 cells taken after release from the thymidine block at the indicated time points; the experiment was repeated twice. (*B*) OD analysis of the autoradiographic signals; bars represent the signals derived from hybridization normalized on the probe length vs. GAPDH signal.

activity was observed for the px containing construct suggesting that this haplotype is coupled with a higher rate of Pol II elongation in late S/G2 phase. These data also indicated that the effect of the haplotype on transcription could be appreciated only in the late S/G2 phase, whereas, in S phase, transcription is prevented for all haplotypes. In agreement with this conclusion, Pol II was found blocked at the polymorphic site even in cells carrying an homozygous px/px genotype, like the T47D (*SI Appendix*, Figs. S1 and S2).

The *Pvu*II and *Xba*I Haplotype Configuration Influences the Amplitude of ERα Fluctuation During Cell Cycle. To demonstrate that px facilitates Pol II read-through also in the chromosomal context, we set up an allele-specific real-time PCR protocol that was applied to the nuclear RNA transcripts obtained from synchronized MCF-7 cells (which carries px and PX haplotypes) in S, late S/G2 and G1 phase. This assay is based on the ability of oligonucleotide probes to discriminate the nuclear transcripts generated by heterozygous alleles at the *Xba*I polymorphic site. The analysis demonstrated a significantly higher amount of transcripts carrying the px haplotype (compared with those carrying the PX) in all phases of the cell cycle (Fig. 5*C*) with a px/PX ratio of transcription of 6.7 fold in S phase, which significantly increased up to 14.3 fold in late S/G2, to decrease at 7.1 fold in G1. We concluded that also in the chromosomal context the px haplotype allowed a more efficient Pol II elongation when the blockade is released in late S/G2. Consistent with this observation, we found a prevalence of the px transcripts in low/high-grade breast cancers (*SI Appendix*, Table S1) and in endometrium (*SI Appendix*, Table S3) human samples; the transcription ratio (px/PX or Px/PX) varied, in some cases was as much as 3,000 times higher and in one case the px was the only allele transcribed (Fig. 5*D*). To gain more insights into the mechanism, we have investigated the recruitment of Pol II, upstream, within and downstream the SNPs site in the context of the px or PX alleles by chromatin immunoprecipitation (ChIP) and allele specific amplification assays. ChIPs were carried out with Pol II specific antibodies including anti-total Pol II and anti-phosphorylated isoforms of Ser5 and Ser2 of the CTD. The results showed that activated Pol II (Ser2 and Ser5 phosphorylated) was significantly accumulated upstream the SNP sites (Fig. 5*E*) and mostly on the px allele (Fig. 5*F*). Activated Pol II is thus enriched at the site of pausing and is ready to be released upon appropriate stimulations occurring during late S/G2 phase of the cell cycle.

The Proto-Oncogene c-MYB Differentially Regulates the Release of the Pol II Block at the Polymorphic Site. Based on the hypothesis of the existence of a cell factor able to recognize the *ERα* intron 1 and facilitate Pol II elongation, we carried out a large series of transfection experiments using the assay reported in *SI Appendix*, Fig. S5*B* to test the effect of several proteins and signaling pathways (*SI Appendix*, Table S4); although some factors displayed a significant activity in strengthening the block of Pol II

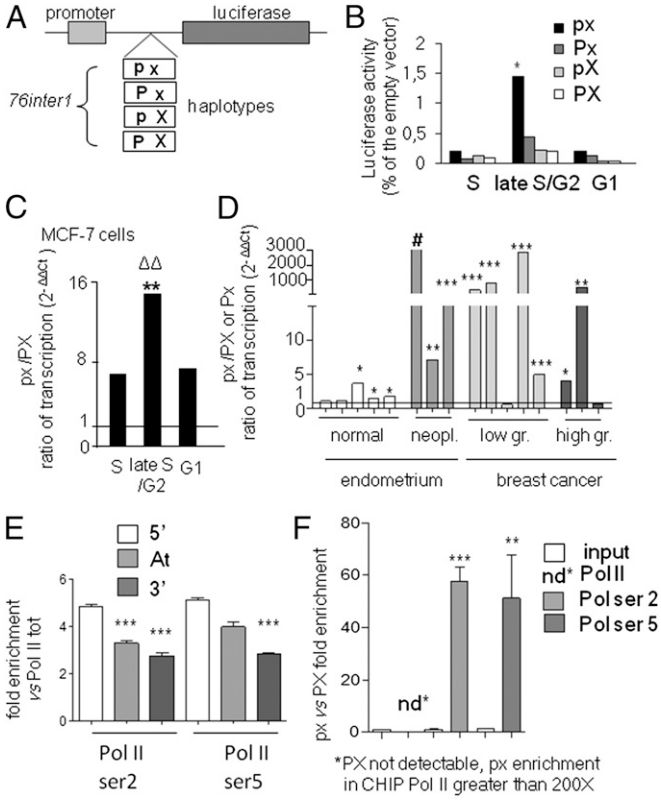

Fig. 5. Higher rate of Pol II elongation is associated with the px haplotype. (*A*) Scheme of the *76inter1* DNA fragments subcloned into the pGL3 vector carrying the four different haplotypes (px, Px, pX, PX). These constructs were transfected in MCF-7 cells. (*B*) Bars are the ratio of the relative light units (RLU) (normalized on β-galactosidase activity) of each haplotype containing construct versus the empty vector measured in transfected cells in S, late S/G2, and G1 phases; $*P < 0.05$, px vs. Px at 13 h. (*C*) Bars are the px/PX ratio ± SEM of allele-specific expression measured by real-time PCR at the indicated cell cycle phases, $**P < 0.005$, late S/G2 vs. S; $\Delta\Delta P < 0.005$, late S/G2 vs. G1. (*D*) The px/PX ratio of allele-specific expression measured in eight human biopsies of normal and neoplastic endometrium and in eight low- and high-grade breast cancers heterozygous for the XbaI polymorphic site; bars are average values ± SEM; # in this case px was the only allele transcribed, $***P < 0.001$, $**P < 0.01$, $*P < 0.05$ px versus PX or Px ratio of transcription. (*E*) ChIPs with anti-total Pol II and anti-phosphorylated isoforms of Ser2 and Ser5 of the CTD; real-time PCRs were done using primers upstream (5′) downstream (3′) or within (At) the attenuation site; bars are average values ± SEM of the fold enrichment relative to the total Pol II, $***P < 0.001$ comparing 3′ and At versus 5′ amplifications. (*F*) The same ChIP samples reported in *E*, but amplifications were carried out with allele-specific real-time PCR; $***P < 0.001$ and $**P < 0.01$ comparing the px/PX ratio from ChIP anti-Pol II Ser2 or Ser5 versus ChIP input, respectively. Data were collected from at least two independent experiments performed in triplicate. *P* values were calculated with one way ANOVA followed by Bonferroni's post test.

elongation (NFYA, Phorbol ester, LeptomycinB, LiCl), these experiments enabled us to identify c-MYB as a good candidates for the release of Pol II elongation. Interestingly, c-MYB had been previously proposed as a modulator of *ERα* gene transcription through the *Pvu*II RFLP sequence (12), although the mechanism underlying this activity had never been investigated. To verify whether c-MYB was able to bind the chromatin surrounding the ERα intron 1 polymorphic site, we overexpressed a HA-tagged c-MYB (20) in MCF-7 cells (*SI Appendix*, Fig. S6*A*) and immunoprecipitated the chromatin bound c-MYB using anti-HA, anti-c-MYB and anti IgG antibodies (Fig. 6*A*). We found a specific c-MYB enrichment in the chromosomal region surrounding the two polymorphisms and in the *CCNB2* promoter used as a positive control; no amplification signals were visible when unspecific IgG or primers mapping at – 35 Kb upstream were used (Fig. 6*A* and *SI Appendix*, Fig. S6*B*). Indeed, a cluster of potential c-MYB binding sites is present in the region surrounding the SNPs (the frequency of the site in the remaining sequence of the intron is about one in 1.5 Kb) including the *Pvu*II polymorphic site proposed (12) as a putative MYB regulatory element (*SI Appendix*, Fig. S6*D*). Modulation of c-MYB expression directly influenced Pol II elongation, since increased or decreased c-MYB expression reduced (Fig. 6*B*) or increased the pausing ratio (*SI Appendix*, Fig. S6 *E* and *F*). c-MYB activity was shown highest on the px allele in the transfection assay (*SI Appendix*, Fig. S6*C*), therefore we analyzed the effects of the transcription factor overexpression in the chromosomal context of the endogenous gene; to this aim, we measured the enrichment of the px allele and the ratio of the nuclear transcripts carrying the px and PX allele in ChIP experiments carried out using anti-HA antibody in c-MYB overexpressing cells (Fig. 6 *C* and *D*). A significant increase of the px allele in the immunoprecipitate and an increased level of px transcripts were observed in c-MYB overexpressing cells, confirming that c-MYB preferentially binds and differentially facilitates the release of Pol II elongation depending on the P*vuII*/*XbaI* haplotype in the endogenous *ERα* gene leading to the increased production of mature ERα mRNA (Fig. 6*E*). Interestingly, previous works described an analogous mechanism of regulation by ERα on the Pol II release on the *c-MYB* gene (21), suggesting that a reciprocal regulation between these two transcription factors is occurring at the level of Pol II elongation. In keeping with this reciprocal regulation, c-MYB expression is low during the S-phase both in MCF-7 and ZR 75.1 cell lines (*SI Appendix*, Fig. S7). This low expression is not just a consequence of the reported ERα regulation of c-MYB (21), because the

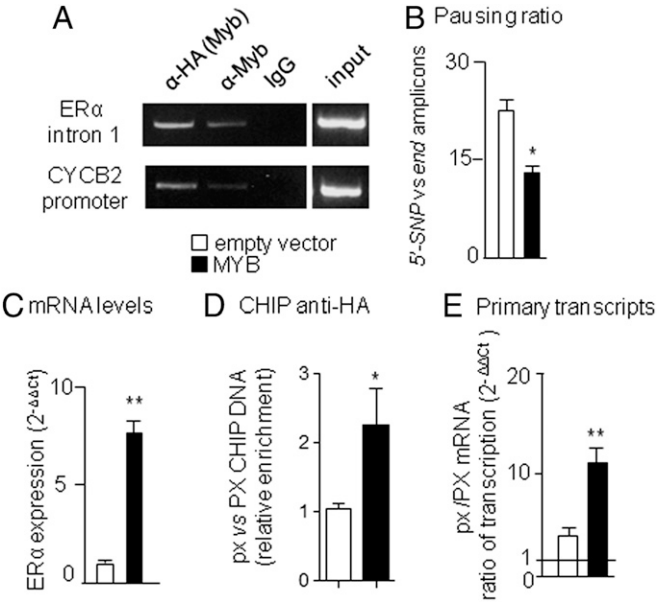

Fig. 6. The effects of c-MYB overexpression on the block of Pol II elongation in MCF7 cells. (*A*) ChIP analysis of the ERα intron 1 site and CYCB2 promoter (as positive control) using anti-HA and anti-cMYB antibodies and mouse IgG. (*B*) Pausing ratio calculated as in Fig. 2*D*; bars represent mean values ± SEM, $*P < 0.05$ MCF-7::HA-cMYB versus MCF-7::empty vector. (*C*) ERα mRNA levels measured by real-time PCR; bars represent means ± SEM, $**P < 0.001$ MCF-7 HA-cMYB versus MCF-7 empty vector. (*D*) ChIPs performed with anti-HA antibody followed by allele-specific real-time PCR; bars represent the ratio of px/PX template, $*P < 0.01$ MCF-7::HA-cMYB versus MCF-7::empty vector. (*E*) Ratio of px/PX mRNA measured by allele-specific real-time PCR; bars represent mean values ± SEM, $**P < 0.01$ MCF-7::HA-cMYB versus MCF-7::empty vector. Data were collected in at least two independent experiments performed in triplicate. *P* values were calculated with one way ANOVA followed by Bonferroni's post test.

ectopical increase of c-MYB expression during S-phase by transient transfection in thymidine synchronized cells, produced a correspondent increased level of ERα expression selectively in c-MYB transfected samples (*SI Appendix*, Fig. S8), again indicating that c-MYB plays a critical role for the release of the Pol II blockage at the intron 1 polymorphic site during the cell cycle. The reciprocal regulation between c-MYB and ERα in humans is also well supported by the significant correlation of their expression found in clinical samples showed by previous studies (21–23) and in our own analysis of public databases (*SI Appendix*, Fig. S9). We propose that the coordinate regulation of these two transcription factors controlling their oscillatory expression during cell cycle represents a novel concept to better understand the hormone action throughout the body.

Discussion

Our study provides the explanation for the genetic association between the *Pvu*II/*Xba*I polymorphisms and several aspects of estrogen physio-pathology including neoplastic transformation of tissues target of estrogen signaling, thus suggesting that the ERα oscillatory expression is involved in these processes. Our finding that the magnitude of pausing at the intron 1 site increases during progression toward a hormone-refractory ERα-negative breast cancer, underlies the clinical relevance of this molecular switch. The ERα oscillatory expression is tightly controlled by an evolutionary conserved site (*SI Appendix*, Fig. S5*A*) where both SNPs map; the mechanism controlling fluctuation involves the c-MYB proto-oncogene, which modulates the Pol II release from the blockage by recognizing, with different affinities, the four haplotypes at the *Pvu*II and *Xba*I sites (Fig. 7).

Decrypting *Pvu*II and *Xba*I Polymorphic Sites in the ERα Intron 1 Links Pol II Elongation to Estrogen-Dependent Carcinogenesis. The present study shows that the association among *Pvu*II and XbaI polymorphisms and hormonal tumorigenesis is likely due to the perturbation of the oscillatory expression of ERα occurring during cell division. Oscillatory expression of the receptor is required for a correct cell division: disruption of the mechanism by ERα constitutive expression was shown to arrest the cell cycle (24) and to induce apoptosis (25) also in the absence of ligand, indicating that hormone-independent pathways are playing a role in the regulation of cell division (26). According to our data, c-MYB, during cell cycle, induces a higher Pol II elongation in the presence of the px haplotype, thus increasing the amplitude and length of the receptor oscillatory expression in proliferating cells; we may speculate that this is causing the activation of overlapping, but not coincident genetic programs during cell cycle, with a differential ability of the receptor to couple its expression with other rhythmic transcriptional programs whose disruption may contribute to the development of hormone-dependent neoplasia (27). It is tempting to speculate that ERα oscillatory expression is part of a cell cycle oscillator (28, 29) involving c-MYB in a positive feed-back loop (21); in agreement with our study, an increasing number of clinical (22, 23) and molecular data (21, 30) suggest a direct involvement of c-MYB in mammary carcinogenesis. Future studies should clarify the relation, if any, of the receptor oscillation in the tumor initiating cell and the tumor suppressive functions of circadian oscillators (27); in line with this idea, oscillation of bHLH transcription factors was recently demonstrated to control multipotency and fate in mouse progenitors (31).

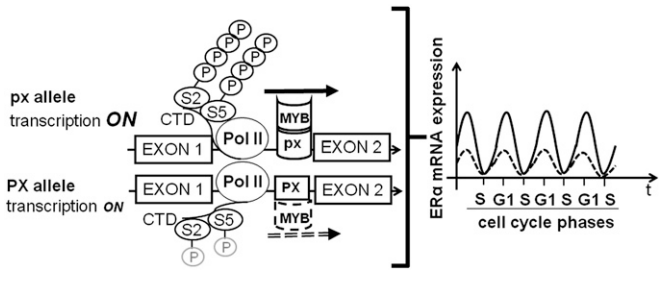

Fig. 7. Model for the control of ERα oscillatory expression during cell cycle by transcriptional pausing. The presence of px in the pausing sequence allows a stronger binding of c-MYB and facilitates the transcription of elongating Pol II (CTD phosphorylated in Ser2 and Ser5) compared with the PX allele, therefore producing an increased amplitude of receptor oscillatory expression during cell cycle.

Clinical Implications. Selected studies have tested the predictive value of using the *Pvu*II and *Xba*I polymorphisms as markers of responsiveness to hormone therapy (32, 33), however, a firm evidence that these polymorphisms can be used as genetic markers to identify responders to hormonal therapy or patients likely to develop hormonal resistance is still lacking. A major limitation of the past studies was the fact that they did not take into consideration the haplotype conformation of both SNPs in each patient, whereas we demonstrated this is likely to be most relevant compared with each single polymorphism. The present study may have a significant impact in the generation of novel clinically useful markers. Our results provide the molecular basis for the changes in the receptor status that patients may experience during tumor progression in metastatic disease, changes that significantly influence patient survival (15). The possibility to develop the analysis of nuclear transcripts into clinically useful biomarkers for metastatic patients has to be carefully evaluated through a systematic study involving a sufficiently large number of patients and the normalization of a number of expected confounders, including tumor stage, haplotype, number of proliferating cells and level of c-MYB in the neoplastic tissue. Notwithstanding these confounders, the preliminary inspection made by sorting the patient samples with a decreasing pausing ratio (*SI Appendix*, Table S1), allowed to observe an inverse correlation of the TLI (thymidine labeling index) and pausing in 10 of 13 samples where this index was available (*SI Appendix*, Table S5). This preliminary evidence is certainly encouraging a more in depth investigation of the primary transcripts and pausing in a wider collection of breast cancer cases aimed at establishing their relation with the proliferation of cancer cells as well as their ability to predict a hormone refractory phenotype. Our data also provide a possible explanation of decades of immunohistochemistry observations showing that ERα positive cells in breast tumor specimen are quite variable (1–100%) and that clinical responses to hormonal therapies can be demonstrated in patients with tumors presenting as few as 1–10% of weakly ERα-positive cells. Multiple explanations were proposed to account for the variability of ERα expression observed in breast cancers, including differences in the subclones originated by the tumor originating cells (cancer stem cells) or the clonal selection mechanism intervening during tumor formation (34). Our results indicate that cellular heterogeneity, might be due, at least in part, to the presence of different number of cells in the S-phase of the cell cycle in the histological samples. In favor of this hypothesis, several studies showed a substantial lack of colocalization between S-phase markers and ERα expression in mouse (35), rat (36), bovine (37), and human tissues (38). Because immunohistochemistry analysis offers a snapshot image of the tumor cells, we may conclude that the technique is likely underestimating the number of ERα positive cells limiting the counts to those cells that are not in S-phase. Future studies should re-evaluate ERα immunodetections as a tumor marker in light of the results of the present study. Finally, it is expected that the identification of the key factors underlying the mechanism of ERα instability during breast cancer progression may provide also a new impulse in the development of therapeutics aimed at sensitizing hormone refractory cancers.

Methods

Run-On Assay. Cells (50×10^6) were incubated for 5 min in a lysis buffer (10 mM Tris·HCl, pH 7.4, 3 mM $MgCl_2$, 0.5% Nonidet P-40, and 10 mM NaCl) and nuclei were pelleted by microfuge centrifugation. Nuclear run-on assays were performed as described (39). Four-h prehybridization and 24-h

GENETICS

hybridization were carried on in Church and Gilbert's solution at 72 °C; 2 × 10^6 cpm/mL were used in the hybridization reaction. After two washes of 45 min at 72 °C in a solution containing 40 mM NaHP0, 1% SDS, and 1 mM EDTA, filters were exposed to autoradiographic films. Details concerning the run on probes can be found in *SI Appendix, Methods*.

Real-Time PCR for Amplification and Quantitation of the Primary ERα Transcripts. Templates were amplified using the TAQMAN Universal PCR amplification kit (Applied Biosystems) in a thermocycler (ABI Prism 7000, Applied Biosystems). The level of intron 1 amplicons reported in Fig. 2 *B* and *C* were normalized on the constitutively expressed gene *36B4* and on primer/probe efficiency using the $2^{-\Delta\Delta Ct}$ method. Absolute quantification of the primary transcripts shown in Figs. 2 *D* and *E* and 6*B* and *SI Appendix*, Fig. S2 was carried out using the standard curve procedure according to the manufacturer guidelines (Applied Biosystems). Both 5′-*SNP* and *end* amplicons showed a linear trend down to 10–100 copy of template (*SI Appendix*, Fig. S2). Primers, probes and the plasmid used in the assay are detailed in the *SI Appendix, Methods*.

Allele-Specific Transcription. The protocol for monitoring allele-specific transcription was developed from a preset SNP genotyping assay (Applied Biosystems). Template cDNAs were obtained by random primer retro-transcription of 0.1 μg of total RNA. The ratio of allele-specific cDNAs was determined by using the $2^{-\Delta\Delta Ct}$ method and standardized on efficiency of the two TAQMAN probes used to detect each allele.

Chromatin Immunoprecipitation. ChIP were carried out as described. MCF-7 cells were incubated in 1% of formaldehyde for 10 min at 22 °C, and the reaction was stopped with 0.125 M glycine. Sonicated chromatin was incubated with 4 μg of anti-HA (Roche 1583816), anti-MYB (Millipore 05–175), anti-total Pol II (Santa Cruz Biotechnology), anti-phosphorylated isoforms of Ser5 and Ser2 of the CTD (Abcam), and 4 μg of mouse IgG (Upstate 12–371) as negative control. For PCR analysis, 1 μL of template in 30 μL of total reaction were used. PCR was performed with HOT-MASTER Taq (Eppendorf). Primers used for real-time PCR amplifications are listed in *SI Appendix, Methods*.

Human Tumors. All human tumor specimens were obtained in accordance with the Independent Ethics Committee (CEI) of the "Fondazione IRCCS Istituto Nazionale dei Tumori," Milan, Italy and the main tumor features are listed in *SI Appendix*, Tables S1–S3.

Genotyping. Genomic DNA from cell lines, human endometrial and breast tissues were genotyped with the SNPs genotyping assay (Applied Biosystems) and haplotypes were determined by restriction enzyme analysis of PCR-amplified DNA fragments from genomic DNAs.

Statistical Analysis. Statistical significance of differences between means was determined applying Student *t* test or ANOVA followed by Bonferroni's test.

Additional methods can be found in *SI Appendix, Methods*.

ACKNOWLEDGMENTS. We thank Prof. T. Gonda for kindly providing us with the shMYB vector; F. Zagari, A. Brusadelli, M. Rebecchi, and C. Meda for technical assistance; and E. Tragni for the help with the statistical analysis. This work was supported by Italian Association for Cancer Research IG Grants 11903 (to P.C.), IG-13234 and MFAG-11752 (to G.P.), CARIPLO Foundation Project 2009-2439 (to P.C.), and EPITRON Project LSHC CT 2005 512146 (to A.M.).

1. Hill SM, Fuqua SAW, Chamness GC, Greene GL, McGuire WL (1989) Estrogen receptor expression in human breast cancer associated with an estrogen receptor gene restriction fragment length polymorphism. *Cancer Res* 49(1):145–148.
2. Boyapati SM, et al. (2005) Polymorphisms in ER-α gene interact with estrogen receptor status in breast cancer survival. *Clin Cancer Res* 11(3):1093–1098.
3. Ryan J, et al. (2012) Hormone treatment, estrogen receptor polymorphisms and mortality: A prospective cohort study. *PLoS ONE* 7(3):e34112.
4. Wedrén S, et al. (2004) Oestrogen receptor alpha gene haplotype and postmenopausal breast cancer risk: A case control study. *Breast Cancer Res* 6(4):R437–R449.
5. Parl FF, Cavener DR, Dupont WD (1989) Genomic DNA analysis of the estrogen receptor gene in breast cancer. *Breast Cancer Res Treat* 14(1):57–64.
6. Andersen TI, et al. (1994) Oestrogen receptor (ESR) polymorphisms and breast cancer susceptibility. *Hum Genet* 94(6):665–670.
7. Cai Q, et al. (2003) Genetic polymorphisms in the estrogen receptor α gene and risk of breast cancer: Results from the Shanghai Breast Cancer Study. *Cancer Epidemiol Biomarkers Prev* 12(9):853–859.
8. van Duijnhoven FJ, et al. (2005) Polymorphisms in the estrogen receptor α gene and mammographic density. *Cancer Epidemiol Biomarkers Prev* 14(11 Pt 1):2655–2660.
9. Weel AE, et al. (1999) Estrogen receptor polymorphism predicts the onset of natural and surgical menopause. *J Clin Endocrinol Metab* 84(9):3146–3150.
10. Weiderpass E, et al. (2000) Estrogen receptor α gene polymorphisms and endometrial cancer risk. *Carcinogenesis* 21(4):623–627.
11. Suzuki K, et al. (2003) Genetic polymorphisms of estrogen receptor alpha, CYP19, catechol-O-methyltransferase are associated with familial prostate carcinoma risk in a Japanese population. *Cancer* 98(7):1411–1416.
12. Herrington DM, et al. (2002) Common estrogen receptor polymorphism augments effects of hormone replacement therapy on E-selectin but not C-reactive protein. *Circulation* 105(16):1879–1882.
13. Yaich L, Dupont WD, Cavener DR, Parl FF (1992) Analysis of the PvuII restriction fragment-length polymorphism and exon structure of the estrogen receptor gene in breast cancer and peripheral blood. *Cancer Res* 52(1):77–83.
14. Maruyama H, et al. (2000) Lack of an association of estrogen receptor alpha gene polymorphisms and transcriptional activity with Alzheimer disease. *Arch Neurol* 57(2):236–240.
15. Lindström LS, et al. (2012) Clinically used breast cancer markers such as estrogen receptor, progesterone receptor, and human epidermal growth factor receptor 2 are unstable throughout tumor progression. *J Clin Oncol* 30(21):2601–2608.
16. Hah N, et al. (2011) A rapid, extensive, and transient transcriptional response to estrogen signaling in breast cancer cells. *Cell* 145(4):622–634.
17. Jakesz R, et al. (1984) Influence of cell proliferation and cell cycle phase on expression of estrogen receptor in MCF-7 breast cancer cells. *Cancer Res* 44(2):619–625.
18. Dong XF, Berthois Y, Colomb E, Martin PM (1991) Cell cycle phase dependence of estrogen and epidermal growth factor (EGF) receptor expression in MCF-7 cells: Implications in antiestrogen and EGF cell responsiveness. *Endocrinology* 129(5):2719–2728.
19. Rostagno P, Moll JL, Birtwisle-Peyrottes I, Ettore F, Caldani C (1996) Cell cycle expression of estrogen receptors determined by image analysis on human breast cancer cells in vitro and in vivo. *Breast Cancer Res Treat* 39(2):147–154.
20. Tanno B, et al. (2010) Expression of Slug is regulated by c-Myb and is required for invasion and bone marrow homing of cancer cells of different origin. *J Biol Chem* 285(38):29434–29445.
21. Drabsch Y, et al. (2007) Mechanism of and requirement for estrogen-regulated MYB expression in estrogen-receptor-positive breast cancer cells. *Proc Natl Acad Sci USA* 104(34):13762–13767.
22. Guérin M, Sheng ZM, Andrieu N, Riou G (1990) Strong association between c-myb and oestrogen-receptor expression in human breast cancer. *Oncogene* 5(1):131–135.
23. Thorner AR, Parker JS, Hoadley KA, Perou CM (2010) Potential tumor suppressor role for the c-Myb oncogene in luminal breast cancer. *PLoS ONE* 5(10):e13073.
24. Jiang SY, Jordan VC (1992) Growth regulation of estrogen receptor-negative breast cancer cells transfected with complementary DNAs for estrogen receptor. *J Natl Cancer Inst* 84(8):580–591.
25. Lee Y, Renaud RA, Friedrich TC, Gorski J (1998) Estrogen causes cell death of estrogen receptor stably transfected cells via apoptosis. *J Steroid Biochem Mol Biol* 67(4):327–332.
26. Klotz DM, et al. (2002) Requirement of estrogen receptor-α in insulin-like growth factor-1 (IGF-1)-induced uterine responses and in vivo evidence for IGF-1/estrogen receptor cross-talk. *J Biol Chem* 277(10):8531–8537.
27. Fu L, Lee CC (2003) The circadian clock: Pacemaker and tumour suppressor. *Nat Rev Cancer* 3(5):350–361.
28. Guillaumond F, Dardente H, Giguère V, Cermakian N (2005) Differential control of Bmal1 circadian transcription by REV-ERB and ROR nuclear receptors. *J Biol Rhythms* 20(5):391–403.
29. Cai W, et al. (2008) Expression levels of estrogen receptor β are modulated by components of the molecular clock. *Mol Cell Biol* 28(2):784–793.
30. Cesi V, et al. (2011) TGFβ-induced c-Myb affects the expression of EMT-associated genes and promotes invasion of ER+ breast cancer cells. *Cell Cycle* 10(23):4149–4161.
31. Imayoshi I, et al. (2013) Oscillatory control of factors determining multipotency and fate in mouse neural progenitors. *Science* 342(6163):1203–1208.
32. Herrington DM, et al. (2002) Estrogen-receptor polymorphisms and effects of estrogen replacement on high-density lipoprotein cholesterol in women with coronary disease. *N Engl J Med* 346(13):967–974.
33. Jin Y, et al. (2008) Estrogen receptor genotypes influence hot flash prevalence and composite score before and after tamoxifen therapy. *J Clin Oncol* 26(36):5849–5854.
34. Rivenbark AG, O'Connor SM, Coleman WB (2013) Molecular and cellular heterogeneity in breast cancer: Challenges for personalized medicine. *Am J Pathol* 183(4):1113–1124.
35. Zeps N, Bentel JM, Papadimitriou JM, D'Antuono MF, Dawkins HJ (1998) Estrogen receptor-negative epithelial cells in mouse mammary gland development and growth. *Differentiation* 62(5):221–226.
36. Sivaraman L, et al. (2001) Early exposure of the rat mammary gland to estrogen and progesterone blocks co-localization of estrogen receptor expression and proliferation. *J Endocrinol* 171(1):75–83.
37. Capuco AV, Ellis S, Wood DL, Akers RM, Garrett W (2002) Postnatal mammary ductal growth: Three-dimensional imaging of cell proliferation, effects of estrogen treatment, and expression of steroid receptors in prepubertal calves. *Tissue Cell* 34(3):143–154.
38. Clarke RB, Howell A, Potten CS, Anderson E (1997) Dissociation between steroid receptor expression and cell proliferation in the human breast. *Cancer Res* 57(22):4987–4991.
39. Ciana P, et al. (2003) Estrogen receptor α, a molecular switch converting transforming growth factor-α-mediated proliferation into differentiation in neuroblastoma cells. *J Biol Chem* 278(34):31737–31744.

Cell cycle dependent oscillatory expression of estrogen receptor-α links Pol II elongation to neoplastic transformation.

Cristina Vantaggiato, Marta Tocchetti, Vera Cappelletti, Maria Grazia Daidone, Aymone Gurtner, Alessandro Villa, Giulia Piaggio, Adriana Maggi and Paolo Ciana

Supporting Information (SI) Appendix

Contents

SI Appendix Figures

Vantaggiato_ Fig. S1

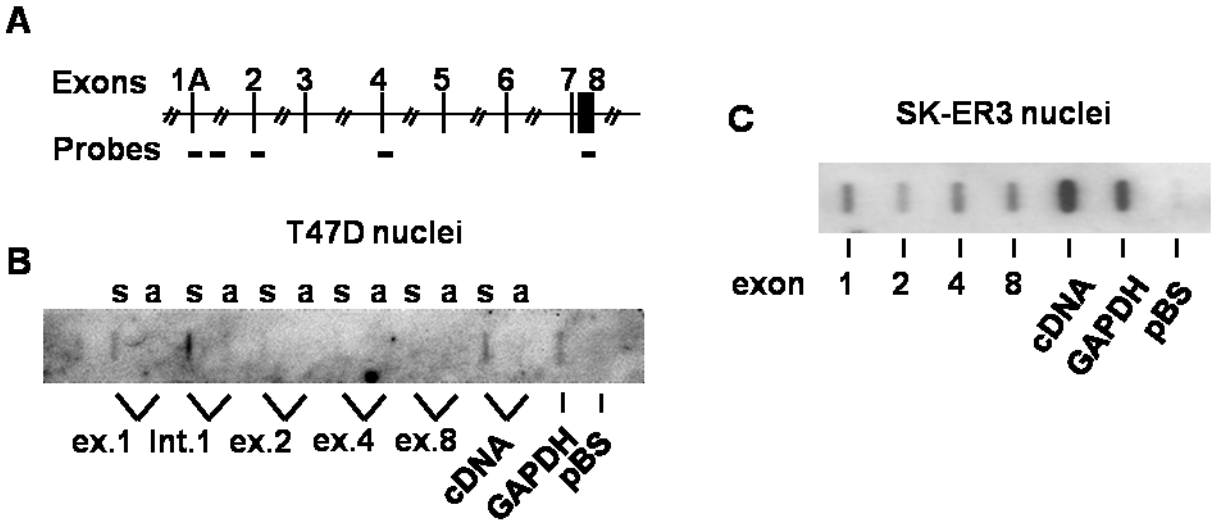

Fig. S1. Pol II attenuation in T47D and SKER3 cell lines. Nuclei generation and *in vitro* transcription assay were carried out as reported in material and methods. (*A*) Schematic representation of the human ERα locus (1) with the genomic position of the run on probes. (*B*) Run on assay on T47D breast cancer cell line; a similar arrest of Pol II elongation at intron 1 found in MCF-7 cells (Fig. 1) is also present in the T47D breast cancer cell line. (*C*) Run on from SK-ER3 a stably transfected neuroblastoma cell line constitutively expressing ERα cDNA (2).

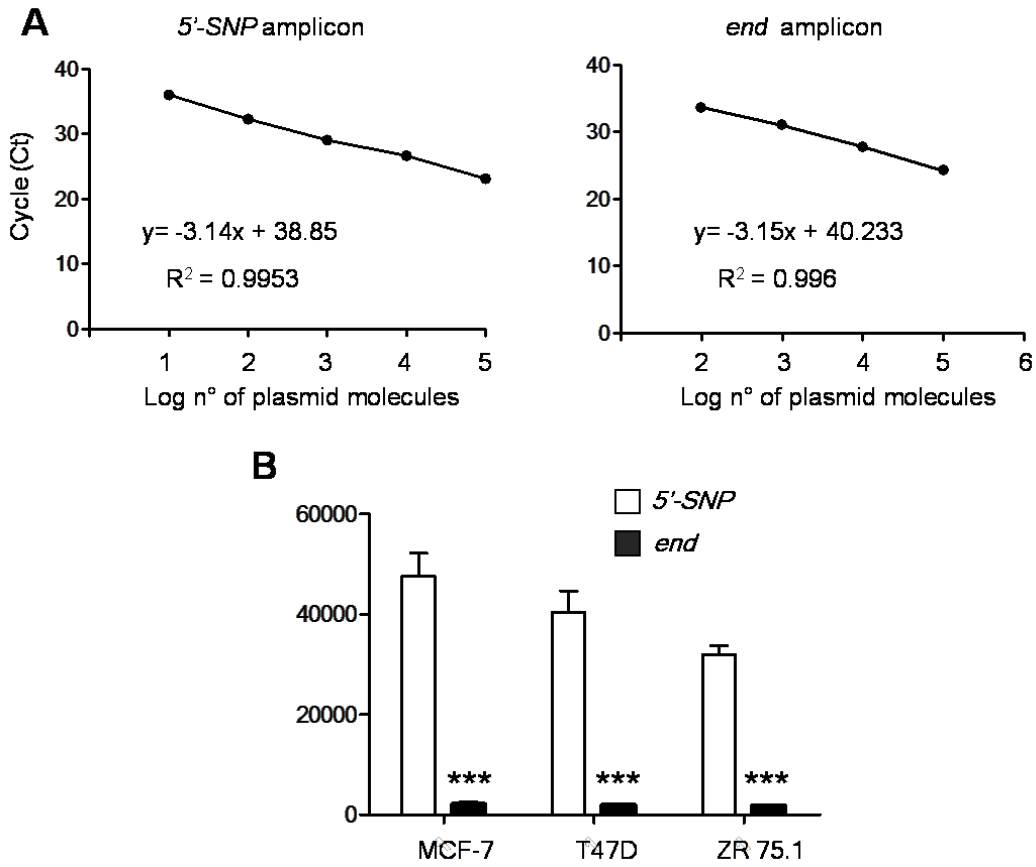

Fig. S2. Absolute quantification of the pausing mechanism by quantitative RT-PCR. (*A*) The standard curve was generated by real time PCR amplification using as template increasing dilutions of a plasmid containing the 532 bp region spanning the two amplicons (5'*SNP* and *end*); the real time PCR showed a linear amplification down to 10 (5'-*SNP* amplicon) or 100 (*end* amplicon) molecules of template. (*B*) The standard curve was used to measure the absolute number of template molecules amplified by the two amplicons in 0.1 µg of retrotranscribed total RNA obtained from MCF-7, T47D and ZR 75.1 breast cancer cell lines. The same procedure was applied to estimate the number of nuclear RNA transcripts obtained from breast tumors (see Fig. 2, Supplementary *SI Appendix* Table S1 and Table S2).

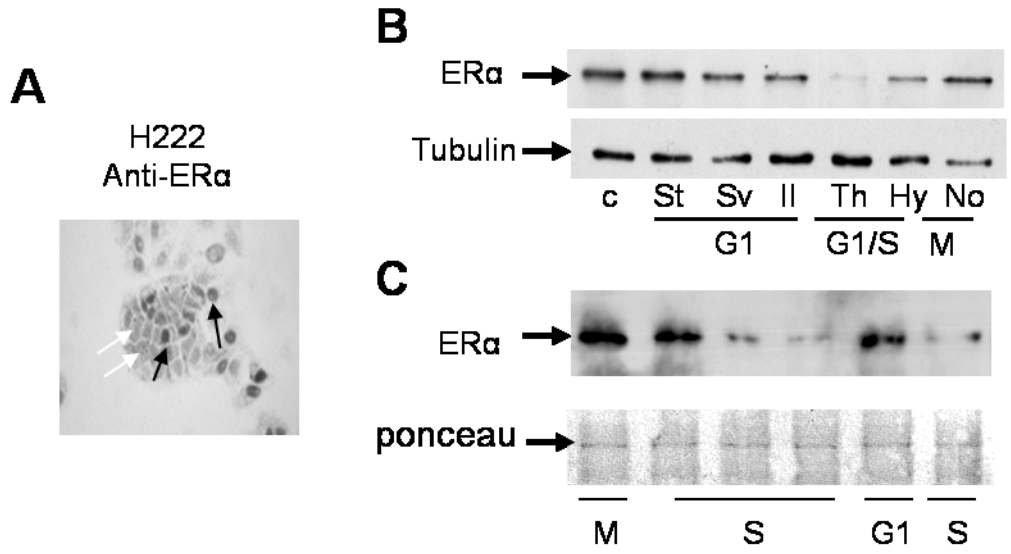

Fig. S3. Supplementary results supporting the oscillatory expression of ERα during cell cycle. (*A*) Immunocytochemistry of non-synchronized MCF-7 cells carried out with the H222 anti-ERα antibody; white and black arrows indicate clonal cells not expressing or expressing ERα, respectively. *B*) Western blot analysis of ERα expression in protein extracts from MCF-7 cells treated with vehicle (c) or different synchronizing agents (*3*) including 2 mM thymidine (Th), 10 nM interleukin-1 α (Il), 4 nM staurosporine (St), 20 μM simvastatin (Sv) and 1 mM hydroxyurea (Hy), 1.7 nM nocodazole (No). *C*) Dynamic of ERα fluctuation in MCF-7 cells synchronized with a nocodazole treatment. Briefly, cells were treated with 1.7 nM nocodazole; mitotic cells loosely attached to the Petri's dish were recovered by gentle pipetting; the drug was washed away before seeding the recovered cells in fresh medium and the cells were then collected at different phase of the cell cycle for Western blot analysis. Filters were immunostained with the antibody directed against ERα (H222); anti-tubulin or ponceau staining are shown as a relative measure of the proteins loaded in the gels.

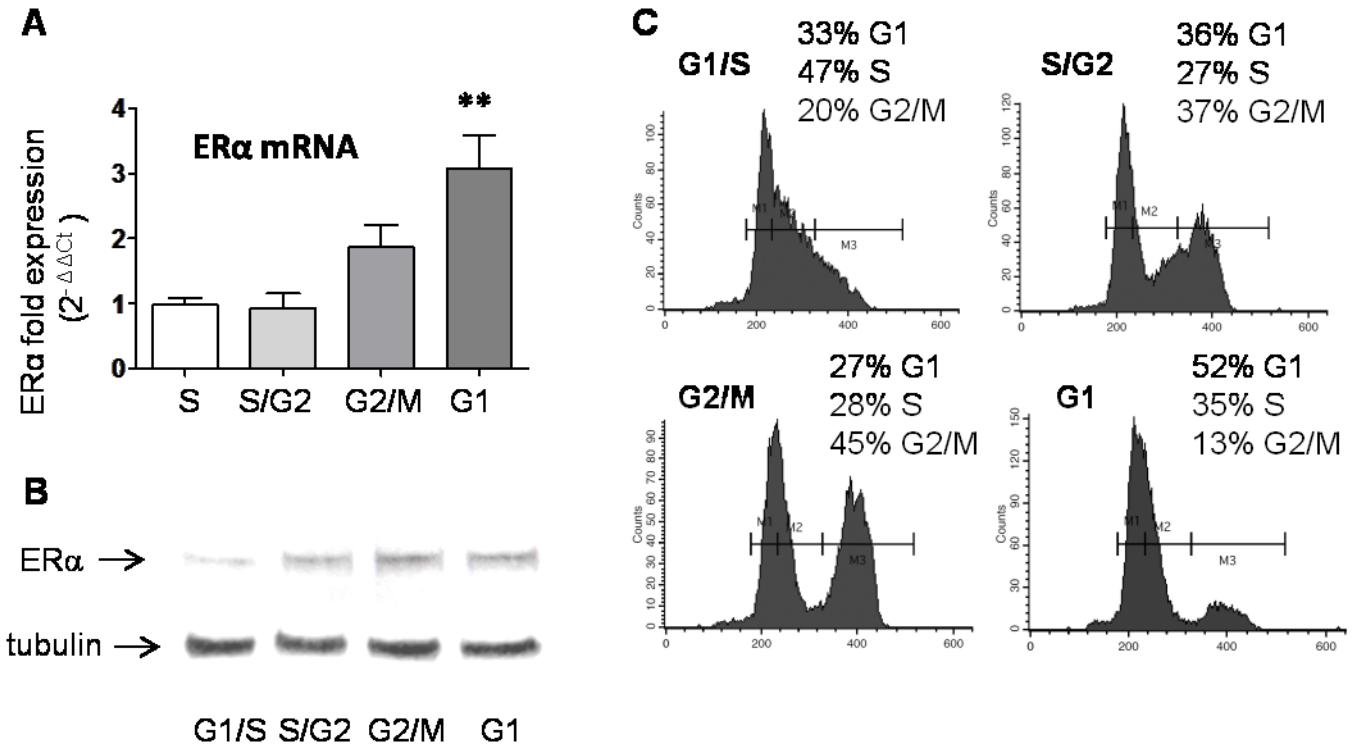

Fig. S4. Oscillatory expression of ERα in ZR 75.1 breast cancer cell line. Similarly to what observed in MCF-7 cells, ERα mRNA (A) and protein (B) is low during the S phase and increase at the end of S/beginning of the G2 phase, reaching the peak of expression in G1. C) Cell cycle phases were determined by cytofluorimetry**.**

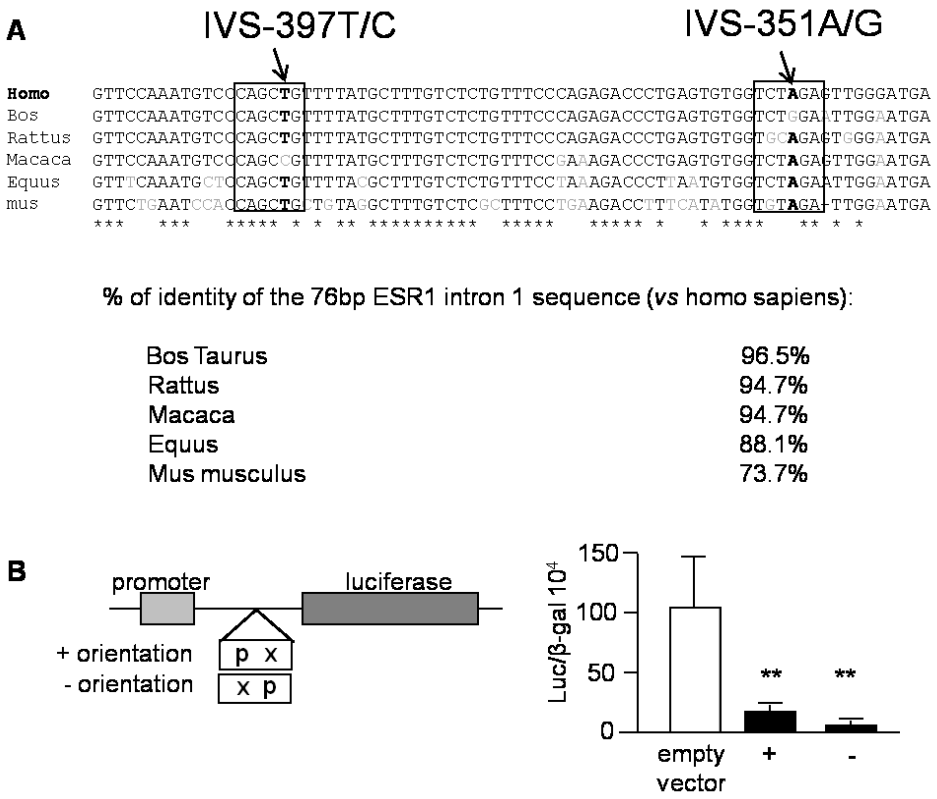

Fig. S5. Evolutionary conservation and functional significance of the 76 bp DNA sequence of ERα intron 1 containing the *Pvu*II and *Xba*I polymorphic sites. (*A*) The 76 bp intron 1 sequence of ERα spanning the polymorphic sites is highly conserved in mammalian. Boxes indicate the *Puv*II (IVS-397T/C) and *Xba*I (IVS-351A/G) restriction sites, arrows indicate the RFLPs which are also highlighted in bold; single nucleotide differences among species are reported in grey, while conserved nucleotide are highlighted with asterisks on the bottom part of the alignment. Alignments were performed with CLUSTAL 2.0.11 Multiple Sequence Alignments software. (*B*) Scheme of the pGL3 vector carrying the *76inter1* DNA fragment in sense (+) or antisense (-) orientations. Bars are average values ± SEM of the relative light units (RLU) normalized on β-galactosidase activity, ** $P<0.01$ cells transfected with the *76inter1* sequence *versus* the empty vector.

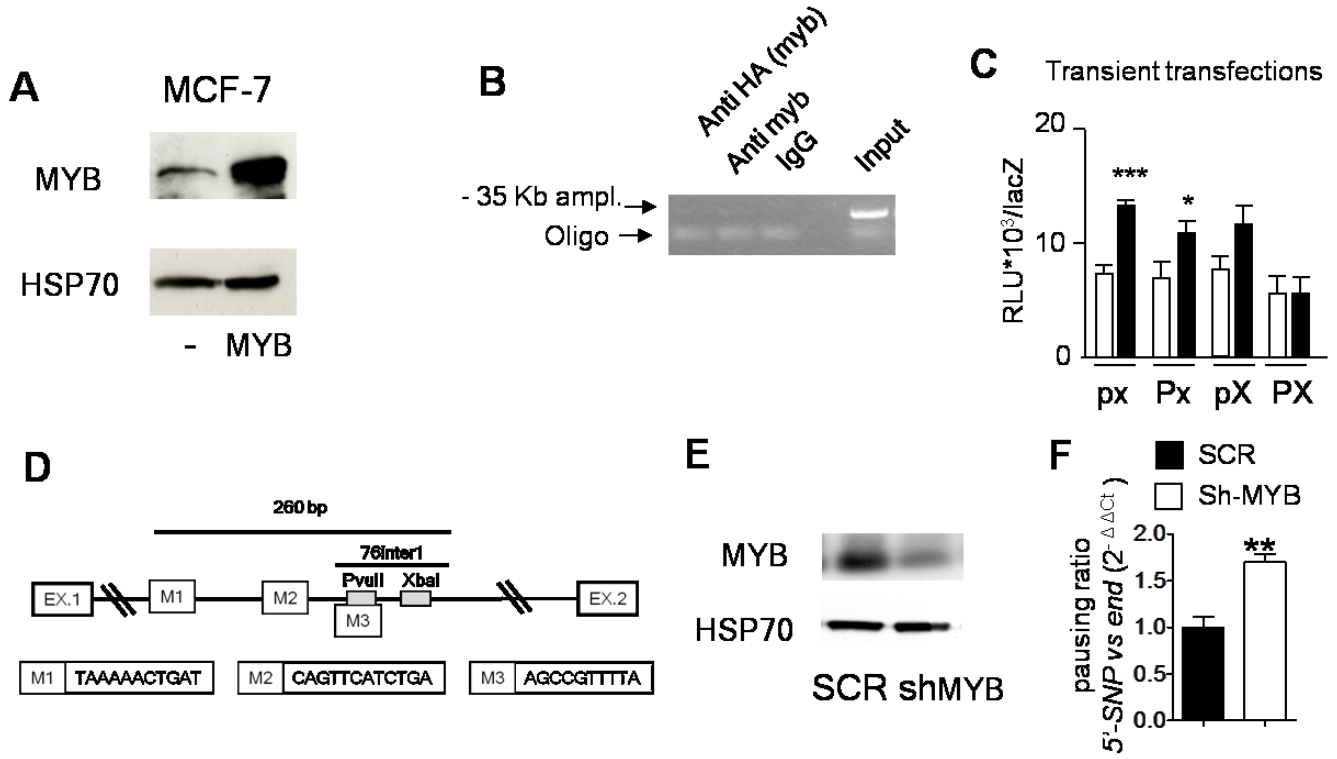

Fig. S6. Supplementary results of the c-Myb effects on the control of ERα expression. (*A*) Western blot analysis of transiently transfected MCF-7 cells with pHA-MYB (*4*) or with the empty vector of the experiments reported in Fig. 6. (*B*) Control of the ChIP experiment of Fig. 6; the immunoprecipitated chromatin was PCR amplified with an amplicon mapping 35 Kb upstream the pausing site. (*C*) Luciferase activity measured in MCF-7 24h after transfection with the reporter constructs carrying the four different haplotypes (see also Fig. 5*A* and *B*). Bars represent mean values ± SEM, ***P<0.001 and * P<0.05 MCF-7::HA-cMYB *versus* MCF-7::empty vector. (*D*) A cluster of high confident (>96%) c-MYB binding sites (M1, M2, M3) were mapped using the algorithm of Farré and colleagues (*5*); the "M3" site was previously characterized (*6*). (*E*) western blot (E) and real time PCR (F) analysis of c-MYB expression and pausing ration in MCF-7 infected with a shMYB expressing vector kindly provided by Prof. T. Gonda (University of Queensland, Australia) (7); pausing ratio was set to 1 for the cells infected with the control scramble vector (SCR).

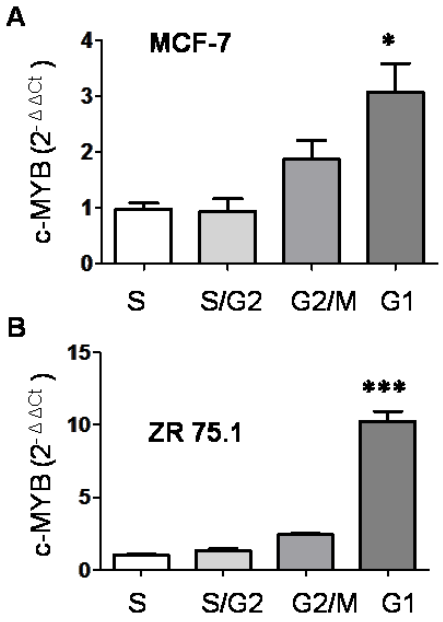

Fig. S7 c-MYB levels are low during the S-phase and peaks in G1-phase of the cell cycle in breast cancer cell lines. Real time PCR analysis of c-MYB mRNA was carried out on the samples of Fig. 3 and *SI Appendix* Fig. S4. In thymidine synchronized MCF-7 (A) as well as in ZR 75.1 (B) cell lines the c-MYB level of expression is low during the S-phase reaching the highest level during the G1-phase of the cell cycle.

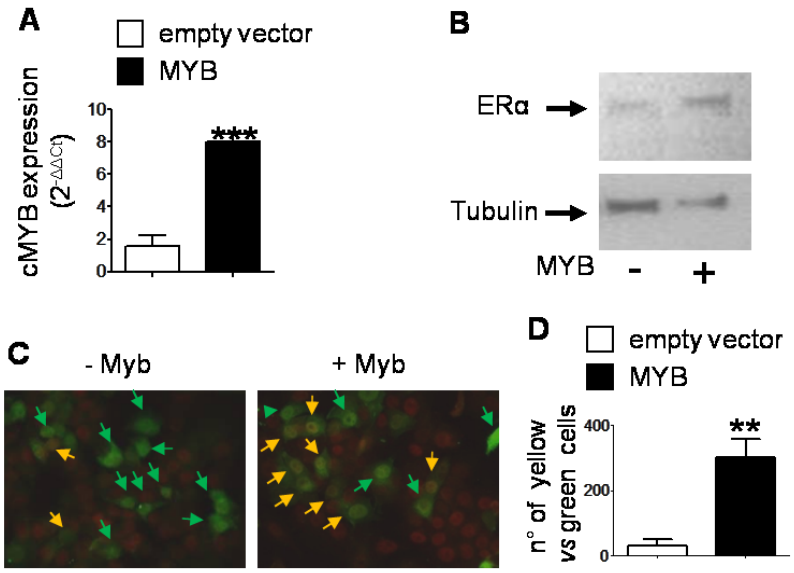

Fig. S8 Ectopical expression of c-MYB increases ERα expression during the S-phase of the cell cycle. A) Semiquantitative real time PCR analysis to measure *c-MYB* mRNA levels during the S-phase of the cell cycle in thymidine synchronized MCF-7 cells transfected with pCDNA3HA-MYB (black bar) or with the empty vector (white bar). *** P <0.001 MYB *versus* empty vector transfected cells: P value was calculated with one way ANOVA followed by Bonferroni's post test. B) Western blot analysis of the ERα levels in protein extracts from the same cells as in (A). C) Immunocytochemistry experiments aimed at assessing whether c-MYB overexpression induces endogenous ERα expression during the S-phase of the cell cycle. To this purpose MCF-7 cells were transfected with an EGFP overexpressing vector (pEGFP-N1) together with a c-MYB overexpressing vector (PCDNA3-HA-MYB) or the empty vector (PCDNA3) as control; the plasmid ratio in the transfection was 1 (pEGFP-N1) : 10 (pCDNA3HA-MYB or pCDNA3), thus each EGFP expressing cell had a high likelihood to be transfected also with the c-MYB (or empty vector) expression vector. Transfected cells were thymidine synchronized and analyzed by immunocytochemistry in the S-phase of the cell cycle. Immunocytochemistry was carried out with an anti-EGFP (green) and anti-ERα (red) antibodies; co-localization (yellow staining) were mainly present in c-MYB overexpressing cells. D) Graph values represent the ratio between yellow cells (green positive + red positive) *versus* green cells obtained from immunocytochemistry experiment indicated in (C); bars in the graph are average values of the cell counts obtained from n. 3 fields (20X); ** P <0.01 yellow *versus* green cells. P values were calculated with one way ANOVA followed by multiple Bonferroni's test.

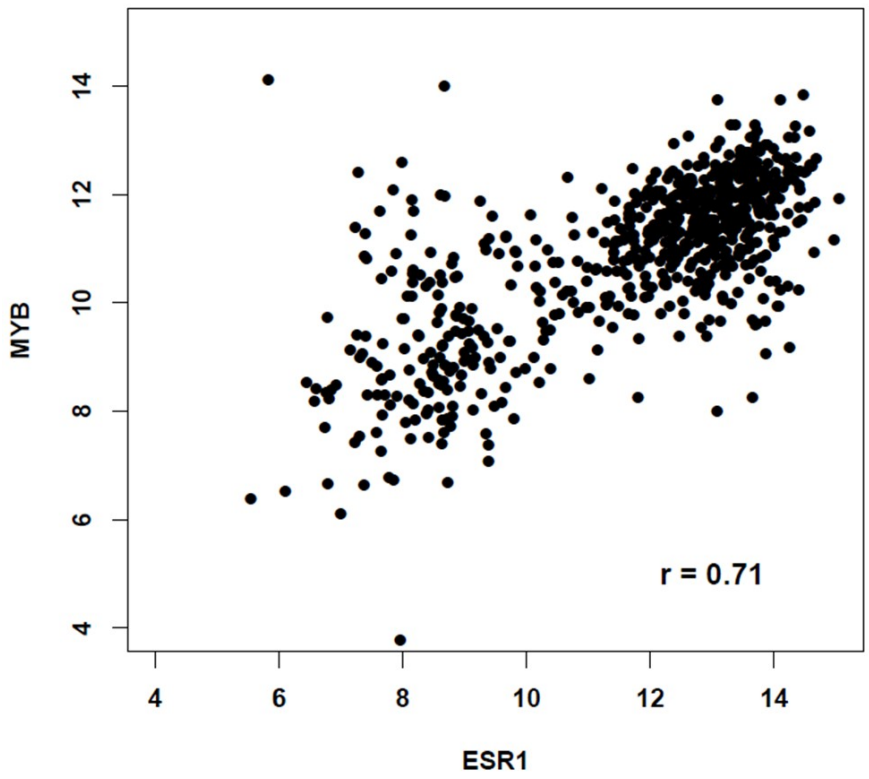

Fig. S9 Correlation between ERα and c-MYB expression in breast cancer patients. The "205225_at" probe set was considered as reporters for ERα gene expression, while "204798_at" was used for c-MYB (number of patients = 684).

SI Appendix Tables

Table S1 – low and high grade breast cancers

ID	Grade	ERα	PgR	histotype	haplotype	5'-*SNP* n. of transcripts (1)	*end* n. of transcripts (2)	Pausing rate	px *vs* PX	TLI (3)
1	Low	Positive	negative	IDC	PX/px	610265	61407	9.94	340	0.1
2	Low	Positive	positive	IDC	PX/Px	20328	1359	14.96	794	0.3
3	Low	Positive	positive	IDC+ LCIS + ILC	px/px	49359	57502	0.86	n.d.	4.6
4	Low	Positive	positive	IDC	PX/Px	49359	15008	3.29	0.6	n.a
5	Low	Positive	positive	IDC + ILC	px/px	5275	2546	2.07	n.d.	0.2
6	Low	Positive	positive	IDC	PX/PX	2787	8808	0.32	n.d.	6.6
7	Low	Positive	positive	IDC	PX/px	55503	88458	0.63	2897	4.2
8	Low	Positive	positive	IDC + ILC	PX/px	11816	1245	9.49	5	0.8
9	Low	Positive	positive	IDC+DCIS	px/px	51579	36304	1.42	n.d.	7.3
10	Low	Positive	positive	IDC	px/px	895	166	5.39	n.d.	4.3
11	Low	Positive	positive	IDC	PX/px	55503	100146	0.55	n.d.	2.3
12	High	positive	positive	IDC	Px/px	6525	30	216.88	n.d.	3.6
13	High	positive	positive	IDC	PX/px	10130	10649	0.95	4.0	n.a
14	High	positive	positive	IDC	px/px	2534	699	3.62	n.d.	n.a
15	High	positive	positive	IDC	PX/PX	970	18	55.33	n.d.	n.a
16	High	positive	negative	IDC	PX/Px	19170	3193	6.00	506	n.a
17	High	positive	negative	IDC	PX/px	3793	426	8.91	0.7	n.a
18	High	positive	positive	IDC	px/px	19030	5208	3.65	n.d.	n.a
19	High	positive	positive	IDC	px/px	25515	1561	16.34	n.d.	n.a
20	High	positive	negative	IDC	PX/PX	76633	17368	4.41	n.d.	1.7
21	High	positive	positive	IDC+DCIS	PX/PX	204685	75886	2.70	n.d.	3.5
22	High	positive	negative	IDC + DCIS	px/px	472145	5521	85.52	n.d.	n.a

(1) and (2): absolute number of template molecules in 0.1 µg total RNA measured with the standard curve procedure; amplicons were mapping (1) before (5'-*SNP* amplicon), and (2) after (*end* amplicon) the two intronic SNPs (see map of Fig. 2A). (3) TLI = Thymidine Labeling Index, % of cells incorporating ^{3}H-thymidine (n. 2000 analysed in total). The results reported in Fig. 2D refer to all the specimens of *SI Appendix* Table S1. n.d.: not detectable. n.a.: not available

Table S2 – primary breast cancers and corresponding relapses

Patient	Type	ER	PgR	5'-*SNP* n. of transcripts (1)	*end* n. of transcripts (2)	Pausing rate
1	Primary	Positive	positive	27255	6435	4.2
1	Relapse	negative	positive	6289	152	41.3
2	Primary	positive	positive	260151	40504	6.4
2	Relapse	negative	negative	15612	177	88.0
3	Primary	positive	negative	55502	117593	0.4
3	Relapse	negative	positive	2188	373	5.9
4	Primary	positive	positive	204685	42318	4.8
4	Relapse	negative	positive	2849	n.d.	>100
5	Primary	positive	positive	81860	13451	6.1
5	Relapse	negative	positive	60607	929	65.2
6	Primary	positive	positive	1573	475	3.3
6	Relapse	negative	positive	54,4	n.d.	>50
7	Primary	positive	positive	58425	52679	1.1
7	Relapse	negative	positive	203	n.d.	>100
8	Primary	positive	positive	2844	2865	1.0
8	Relapse	positive	negative	92223	20707	4.5
9	Primary	positive	positive	11207	8680	1.3
9	Relapse	positive	negative	2961	444	6.7

(1) and (2): absolute number of template molecules in 0.1 µg of total RNA measured with the standard curve procedure; amplicons were mapping (1) before (5'-*SNP* amplicon) and (2) after (*end* amplicon) the two intronic SNPs (see map of Fig. 2A). The results reported in Fig. 2E refer to all the specimens of *SI Appendix* Table S2. n.d.= not detectable

Table S3 - normal and neoplastic endometrium

Patient	Type	px *vs* PX	genotype
1	normal endometrium	0.93	PX/px
2	normal endometrium	1.02	PX/px
3	normal endometrium	4.35	PX/px
4	normal endometrium	1.69	PX/Px
5	normal endometrium	1.43	PX/px
6	endometrial tumor	>1000 (1)	PX/px
7	endometrial tumor	7.14	PX/px
8	endometrial tumor	74.1	PX/px

(1) the PX allele was undetectable and the number of the PX template molecules was lower than 1000 times as compared to the px template; thus the ratio could not be calculated because of the limit of the detection system. The results shown in Fig. 2C were obtained from the biopsies of patient 5 and patient 6.

Table S4 – factors tested for their ability to increase or decrease Pol II elongation through the px site in transient transfection experiments

Stimulus	px fold increase*	P value
Dexametasone (10^{-7} M)	1.2	0.63
Progesterone (10^{-7} M)	1.1	0.91
Retinoic acid (10^{-7} M)	0.6	0.09
Genistein (25 uM)	0.7	0.28
ERalpha	1.5	0.11
P53	1.1	0.83
c-MYB	1.9	<0.0001
c-MYC	1.1	0.32
NFYA	0.71	0.009
CBP	1.3	0.92
SRC1	1.5	0.08
SRC3	1.2	0.06
CDK9/CyclinT1	1.2	0.25
FUS (siRNA)	0.7	0.15
PTMA	1.1	0.15
H89	1.2	0.29
Trichostatin A	0.85	0.09
Phorbol ester (PMA)	0,24	0.0006
Tyrphostin	1.5	0.18
Staurosporine	1.3	0.39
LeptomycinB	0.30	0.004
Adenine 100uM	0.6	0.24
ATP 100uM	0.7	0.22
Lysine 100uM	0.80	0.62
LiCl	0.46	0.005
LY294002	1.11	0.93

* Expressed as RLU/LacZ of the vector carrying the 78bp fragment *versus* RLU/LacZ of the empty vector (the assay is shown in *SI Appendix* Fig. S5B)

Table S5 - association between proliferation marker TLI (Thymidine Labeling Index) and pausing

Patient	5'-*SNP* n. of transcripts	*end* n. of transcripts	pausing ratio	TLI
2	20328	1359	15.0	0.3
1	610265	61407	9.9	0.1
8	11816	1245	9.5	0.8
21	204685	75886	2.7	1.7
5	5275	2546	2.1	0.2
9	51579	36304	1.4	7.3
3	49359	57502	0.9	4.6
7	55503	88458	0.6	4.2
11	55503	100146	0.6	2.3
6	2787	8808	0.3	6.6

Data are taken from *SI Appendix* Table S2; patients are ordered by decreasing pausing ratio. Inverse correlation between pausing ratio and TLI was observed for all patients, but three exception (patients 10, 12, 22) that were omitted from the list.

Supporting Methods

Primers, probes and plasmid used in real-time PCR for amplification and quantitation of the primary ERα transcripts. The following primers and probes specific for ERα intron 1 and the 36B4 coding sequence were used: *start* amplicon: forward 5'-GCGCTGCGTTCAGAGTCAA-3', reverse 5'-GGTAAGTGGGTGGAGAGTACGTTT-3', probe 5'-TTCTCTCGCCGGGCAGCTGAA-3'; *5'-SNP* amplicon: forward 5'-GGCTCAAACTACAGGGCTTAAACA-3', reverse 5'-TGGAATCCCAGCACTTTGG-3', probe 5'-TCTCCTGCTTTGGCC-3'; *end* amplicon: 5'-forward 5'-AATAAACAATTTCCCCTCAAGGTTAA-3', reverse 5'-CCTCTTGCCGTCTGTTGCA-3', probe 5'-TAATAGGCAACACCTTTTG-3' 36B4 amplicon: forward 5'-GGCGACCTGGAAGTCCAACT-3', reverse 5'-CCATCAGCACCACAGCCTTC-3', probe 5'-ATCTGCTGCATCTGCTTGGAGCCCA-3'. Primers were designed with Primer express 2.0 software (Applied Biosystems) and preset to an annealing temperature of 60°C. The thermal profile of the 40 amplification cycles for each amplification reaction was: denaturation step at 95° C for 15 s and an annealing/extension step at 60° C for 60 s. The primer/probe efficiency were determined in amplification experiments on genomic DNAs (obtained from MCF-7 and from endometrial samples) setting the *end* amplicon to 1: these values were found to be constant in all reactions and used as normalizing factors. All amplicons generated a single amplification band, thus reflecting the specificity of the primers.

The standard curve was generated by real time PCR amplification using as template serial plasmid dilutions (from 1 million to one copy) of a plasmid containing the 532 bp region spanning the two amplicons (5'*SNP* and *end*). The threshold Ct was plotted with the logarithm of the copy number and the resulting curve showed a linear trend down to 10 (5'-*SNP* amplicon) or 100 (*end* amplicon) molecules of template.

ChIP primers. Primers used for real time PCR amplifications of Fig. 5F are: (5')-amplicon (upstream the attenuation site) forward 5'-GCCAAGTCCTGGGCCAATAT-3', reverse 5'- TCTCGCCAGACCAGACATAGG-3'; (At)-amplicon (on the attenuation site) forward 5'-AGGGTTATGTGGCAATGACG-3', reverse 5'- CCTGCACCAGAATATGTTACCT-3'; (3')-amplicon (downstream the attenuation site) forward 5'- CTGACAGCAGCTTAGGCTAACCA -3', reverse 5'-GCCCCACACTTAAGATTAAAAAATGA -3'. The primer sequences of the human *ERα* intron 1 used in Fig. 6A, Fig. 6E and *SI Appendix* Fig. S5B were hESR1-Forward 5'-AGGGTTATGTGGCAATGACG-3' hESR1-Reverse 5'- CCTGCACCAGAATATGTTACCT-3' (close to the SNPs) and hESR1-Forward 5'-CAGAGTCAAGTTCTCTCGCC-3' hESR1-Reverse 5'-GAGACCTTCCCAAGTGACTC-3' (-35 Kb from the SNPs); for the *CYCB2* promoter the PCR conditions were previously reported (9).

Plasmids. All ERα fragments used in the run-on analysis were subcloned into a pBluescript II KS vector (Stratagene). Exons 1, 2, 4 fragments were obtained by PCR amplification from the cDNA vector pSWMTERα (kindly provided by G. Greene) using the following primers: exon 1, forward 5'-ATGACCCTCCAC-3', reverse 5'-GTAGTAGGGCACCTGCTGG-3'; exon 2, forward 5'-GCCAAATTCAG-3', reverse 5'-AGATTCCATAGCCATAGTT, exon 4, forward GACCAGAGAGGAA-3', reverse 5'-TCTTCGCCCAGTTG-3'. The exon 8 fragment Bsrg1/Cla1 from pSWMTERα was subcloned into pBluescript II KS, while the intron 1 Pst1/BssHII fragment was subcloned from the λGHER1 (kindly provided by P. Chambon, IGBMC/CNRS UMR7104/Inserm U964/Collège de France, Université de Strasbourg, Illkirch-Cedex, Strasbourg, F). Single strand probes were generated according to the manufacturer's instruction using a helper phage M13K07 (New England Biolabs). For the functional study, the *76inter1* fragment containing the two RFLPs in the four different haplotype configurations were chemically synthesized

(Primm); *Nco*I sites were introduced in the DNA termini to allow subcloning into the Nco1 site of the pGL3 control vector (Promega). The plasmid used for the absolute quantification of the real time PCR templates with the standard curve procedure was obtained by cloning the PCR genomic fragment of 532 bp (generated with the more external primers used in real time PCR for the *5'-SNP* and *end* amplicons) into pCR 2.1 vector (Invitrogen). pHA-Myb expression vector was kindly gift by Giuseppe Raschellà (4).

Northern blot, western blot. Northern and western blots were performed as previously described (8). For northern blot 20 μg total RNA was used in the assay. The entire ERα cDNA was labeled with ^{32}P and used as a probe; membranes were exposed to autoradiography film for 2 days.
Western blots were carried out on 20 μg total proteins; immunodetection of the western blot was done with the anti-ERα antibody H222 (1:2000) and an anti-rat IgG conjugated with horseradish peroxidase as the secondary antibody (1: 5000); detection was performed using an enhanced chemiluminescence (ECL) system (Amersham). The primary antibody was monoclonal H222 (kindly provided by G. Greene, University of Chicago). Detection was performed using an ABC kit (Vector Laboratories).

Thymidine synchronization and cytofluorimetry.

For the synchronization experiments, cells were treated at least for a length of time necessary for duplication with 200 μM thymidine. The block was released by washing the cells twice with PBS and re-supplementing the culture with fresh pre-warmed medium. Synchronously cycling cells were then collected at different time points for cytofluorimetry, run on, northern blot and western blot analysis.

For the cytofluorimetry experiments, cells were trypsinized and suspended in cold PBS; aggregates were dissolved by gentle pipetting before drop-by-drop instillation of cold ethanol (final concentration 62%). Before detection with a cytofluorimeter (Beckton Dickinson, FACS Calubur), cells were pelleted and resuspended in a staining solution (10 μg/ml RNAseA, 1% NP40, 5 μg/ml propidium iodide).

Immunocytochemistry. H222 anti-ERα antibody was diluted 1:2000 in phosphate-buffered saline (PBS) with 10% rabbit serum and 0.3% Tween 20, while the secondary antibody was a biotinylated anti rat IgG (Vector Laboratories) diluted 1:2000 in PBS with 1% goat serum and 0.1% Tween 20. Antibody/antigen detection was performed using an ABC kit (Vector Laboratories) according to the manufacturer's instructions. Immunostaining was visualized by exposure to 3,3'-diaminobenzidine.

Cell culture and transfection. MCF-7, T47D and ZR 75.1 cell lines were routinely grown at 37° C in RPMI 1640 (Sigma) supplemented with 10 % FBS. Transfections and reporter assays were carried out with the calcium phosphate precipitate procedures as described earlier (8). Transfections of c-MYB were carried out with LIPOFECTAMINE LTX (Invitrogen). Briefly, $1.8x10^5$ MCF7cells were plated in triplicate in 24-mm plate and transfected with following the manufacturer's instructions. Five hours later, the liposome complexes were removed and cells were grown in DMEM plus 1% FBS medium for 24 h.

SI Appendix reference list

1. Ponglikitmongkol M, Green S, Chambon P (1988) Genomic organization of the human oestrogen receptor gene. *EMBO J* 7:3385-8.
2. Ma ZQ, et al. (1993) Activated estrogen receptor mediates growth arrest and differentiation of a neuroblastoma cell line. *Proc Natl Acad Sci USA* 90: 3740-3744
3. Krek W, DeCaprio JA (1995) Cell synchronization. *Methods Enzymol* 254:114-24.
4. Tanno B, et al. (2010) Expression of Slug is regulated by c-Myb and is required for invasion and bone marrow homing of cancer cells of different origin. *J Biol Chem* 285:29434-45.
5. Farré D, et al. (2003) Identification of patterns in biological sequences at the ALGGEN server: PROMO and MALGEN. *Nucleic Acids Res* 31:3651-3.
6. Herrington DM, et al. (2002) Common estrogen receptor polymorphism augments effects of hormone replacement therapy on E-selectin but not C-reactive protein. *Circulation* 105:1879-82.
7. Drabsch Y, et al. (2007) Mechanism of and requirement for estrogen-regulated MYB expression in estrogen-receptor-positive breast cancer cells. *Proc Natl Acad Sci U S A* 104(34):13762-13767.
8. Gurtner A, et al. NF-Y (2008) dependent epigenetic modifications discriminate between proliferating and postmitotic tissue. *PLoS ONE* 3:e2047.
9. Ciana P, et al. (2003) Estrogen receptor α, a molecular switch converting transforming growth factor-α-mediated proliferation into differentiation in neuroblastoma cells. *J Biol Chem* 278:31737-44.