Published in final edited form as:

Mol Cell Endocrinol. 2014 January 25; 382(1): . doi:10.1016/j.mce.2013.06.021.

Hormone-Regulated Transcriptomes: Lessons Learned from Estrogen Signaling Pathways in Breast Cancer Cells

Nasun Hah^{1,4} and W. Lee Kraus^{2,3}

¹Gene Expression Laboratory, The Salk Institute for Biological Studies, La Jolla, CA 92037

²The Laboratory of Signaling and Gene regulation, Cecil H. and Ida Green Center for Reproductive Biology Sciences, University of Texas Southwestern Medical Center, Dallas, TX 75390

³Division of Basic Reproductive Biology Research, Department of Obstetrics and Gynecology, University of Texas Southwestern Medical Center, Dallas, TX 75390

Abstract

Recent rapid advances in next generation sequencing technologies have expanded our understanding of steroid hormone signaling to a genome-wide level. In this review, we discuss the use of a novel genomic approach, global nuclear run-on coupled with massively parallel sequencing (GRO-seq), to explore new facets of the steroid hormone-regulated transcriptome, especially estrogen responses in breast cancer cells. GRO-seq is a high throughput sequencing method adapted from conventional nuclear run-on methodologies, which is used to obtain a map of the position and orientation of all transcriptionally engaged RNA polymerases across the genome with extremely high spatial resolution. GRO-seq, which is an excellent tool for examining transcriptional responses to extracellular stimuli, has been used to comprehensively assay the effects of estrogen signaling on the transcriptome of ER⁺-positive MCF-7 human breast cancer cells. These studies have revealed new details about estrogen-dependent transcriptional regulation, including effects on transcription by all three RNA polymerases, complex transcriptional dynamics in response to estrogen signaling, and identification novel, unannotated non-coding RNAs. Collectively, these studies have been useful in discerning the molecular logic of the estrogen-regulated mitogenic response.

Keywords

Antisense RNA; Breast Cancer Cells. Divergent RNA; Enhancer RNA; Estrogen; Estrogen receptor; Estrogen receptor binding site; GRO-seq; Long non-coding RNA; MicroRNA; Mitogenic Growth; RNA polymerase

© 2013 Elsevier Ireland Ltd. All rights reserved.

⁴Address correspondence to: Nasun Hah, Ph.D. Gene Expression Laboratory, The Salk Institute for Biological Studies, 10010 North Torrey Pines Road, La Jolla, CA 92037, Phone: (858) 453-4100, Fax: (858) 535-1979, nhah@salk.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

1. Introduction

1.1. Estrogen signaling and estrogen receptors

Estrogens are a class of endogenous hormones that play critical roles in diverse aspects of human physiology in females and males, including sexual development, reproduction, cardiovascular and neuronal activity, as well as liver, fat, and bone metabolism. Dysregulation of estrogen signaling can lead to a variety of human diseases, such as breast and uterine cancers, osteoporosis, cardiovascular and neurodegenerative diseases, and insulin resistance [1,2]. The actions of estrogens are mediated through estrogen receptor proteins (ER α and ER β), which are members of the nuclear receptor superfamily. They function as ligand-regulated nuclear transcription factors, but also as key components of cytoplasmic membrane-initiated signaling cascades [3–5]. The two ER isoforms are encoded from two separate genes in two different chromosomal locations. They share about 96% homology between their DNA binding domains (DBDs), but only 56% homology between their ligand binding domains (LBDs) and 28% homology between their amino-terminal activation functions 1 (AF-1s) [6]. ER α and ER β can homo- or hetero-dimerize [7,8], indicating that the two isoforms can act together or separately. The two ER isoforms share overlapping functions due, in part, to the significant homology of their DNA binding domains. However, ER α and ER β have different expression patterns in tissues, distinguishable differences in structure, and distinct biological functions [9–12]. The signaling and transcriptional effects of ERs underlie the aforementioned physiological and pathological effects of the estrogen signaling pathways.

1.2. Estrogen signaling and breast cancer

Estrogen and ERs have been shown to play a mitogenic role in breast, uterine, and ovarian cancers [1,11]. The etiology of these hormone-responsive cancers has shown that estrogen stimulates the unregulated cellular proliferation in these target tissues, which can interfere with normal physiological processes and can drive tumor formation or progression [11,13–15]. Many of the effects of estrogen signaling in breast cancers are driven by estrogen-dependent changes in the breast cancer cell transcriptome. Most studies have implicated ER α in these processes, but ER β may play a role as well [11]. Currently, ER α -positive breast cancers are treated with selective ER modulators (SERMs) (e.g., tamoxifen) and aromatase inhibitors (e.g., anastrozole) to block the estrogen signaling pathways in cancer cells [16,17], which ultimately alters ER α -mediate transcription programs and inhibits estrogen-dependent proliferation. Given that these targeted therapies focus on changing the availability of estrogen or the activity of ER α , it is important to understand the set of target genes of the estrogen signaling pathway and how they are regulated in target cells, which will provide insights for developing therapeutics with minimal side effects.

1.3. Transcriptional regulation by estrogen receptors

Many studies over the past three decades have elucidated the molecular mechanisms by which estrogen signaling and nuclear ERs regulate transcription and affect gene expression outcomes. Estrogens, such as the predominant naturally occurring endogenous estrogen 17 β -estradiol (E2), are lipophilic and can diffuse freely into cells, where they initiate cytoplasmic and genomic signaling events that ultimately promote global changes in gene expression in the nucleus [1,18–21]. Cytoplasmic estrogen signaling is mediated by a small pool of cytoplasmic membrane-associated ERs, which stimulate kinase-mediated signaling pathways that lead to changes in the localization and activity of nuclear transcription factors [20,21]. Nuclear estrogen signaling is mediated by nuclear ERs, which function as ligand-regulated transcription factors [1,19]. In the classical or “direct” nuclear pathway, E2 induces the dimerization of ER, which then binds directly to genomic DNA containing estrogen responsive elements (EREs) consisting of two AGGTCA half sites arranged

palindromically around a three-basepair spacer [22,23] (Fig. 1, *top left*). In the non-classical or “indirect” nuclear pathway, ERs bind indirectly to genomic DNA through regulatory elements by “tethering” through other transcription factors, such as AP-1, Sp1, and NF- κ B [24–26] (Fig. 1, *bottom left*). In either case, the binding of ER at enhancers promotes the recruitment of coregulators (e.g., histone modifying and remodeling enzymes, chromatin looping factors) that mediate posttranslational modifications of histones or other transcription factors, as well as chromatin remodeling and looping events [27] (Fig. 1). Liganded ERs and coregulators connect estrogen signaling to promoter-engaged RNA polymerase II (Pol II) and the basal transcriptional machinery through looping mechanisms that induce the changes in occupancy or activity of Pol II [28–31] (Fig. 1). This ultimately leads to profound changes in the transcriptome of cells.

How ERs bind to their cognate binding sites upon estrogen treatment and how the DNA sequences at those binding sites affect ER binding on a genomic scale have been studied extensively [5,18,27,32–34]. However, in order to understand how ERs regulate transcription, it is essential to understand the underlying mechanisms of how Pol II is regulated at estrogen target promoters. In a widely accepted view of estrogen-mediated transcriptional regulation, ERs and coactivators are recruited to target genes upon estrogen stimulation, which in turn modulates the recruitment of the Pol II machinery to their promoters [3,19,27,35]. Under this view of estrogen-dependent transcriptional regulation, the recruitment of the Pol II pre-initiation complex is a rate-limiting step. Interestingly, recent gene specific- and genome-wide studies in flies and mammals have shown an alternative view of gene transcription by Pol II, where Pol II is widely distributed at the promoters of target genes prior to stimulation [36], perhaps as a means of synchronizing transcriptional responses [37]. Estrogen signaling through nuclear and membrane-initiated estrogen signaling may act at both levels, facilitating Pol II loading, as well as release of Pol II from pause sites [34,38–40]. In spite of recent progress characterizing some of the mechanisms of estrogen-regulated transcription, many key questions remain to be answered about which transcripts comprise the set of primary or direct estrogen-regulated target genes and which specific molecular mechanisms control the expression of these genes.

1.4. Genomic studies of estrogen signaling and gene regulation: Approaches used to define primary estrogen target genes

Primary response genes, in the classical sense, are those that are regulated as an immediate response to cellular signaling pathways, without the need for protein synthesis or other secondary regulatory events [41,42]. More recently, the concept of “direct” target genes has emerged, which typically refers to target genes whose promoters come into physical contact with, and are regulated by, a transcription factor bound at a proximal or distal enhancer [28,32–34]. Primary/immediate response genes and direct target genes are two sides of the same coin. Although they are often incorrectly used synonymously, they focus on distinct aspects of the gene regulatory process. For the former, regulatory effects on Pol II are key; for the latter, binding and enhancer-promoter looping events are key. In the section below, we describe some of the approaches that have been used to identify primary/immediate response genes and direct target genes.

1.4.1. Expression microarrays and RNA-seq—Numerous gene expression microarray studies have been used to identify and define estrogen-regulated target genes in breast cancer cells, including MCF-7 cells, a widely used human breast adenocarcinoma cell line [18,27]. These studies have been useful in characterizing the global effects of estrogen signaling and defining key concepts in estrogen-mediated gene regulation. However, the results from analyses using the same cell lines have shown variable results from experiment to experiment due to a variety of factors, including cell growth and estrogen treatment

conditions, as well as the array platforms used [18,27,32,43]. Estimates of the number of target genes that are regulated upon estrogen stimulation based on expression microarrays varies from 100 to 1,500 [5,18,27]. It is likely that these numbers include indirect or secondary (or tertiary) effects of long treatments with estrogen (i.e., >4 hours to days). To define the primary estrogen-regulated transcriptome, cycloheximide, an mRNA translation inhibitor, was used to prevent secondary effects of estrogen signaling. Interestingly, only 20 to 30% of the total estrogen-regulated gene set was regulated in the presence of cycloheximide, indicating that only a small set of the genes identified in expression microarray experiments are genuine primary or immediate targets of estrogen signaling [43]. This approach helped to alter our thinking about the timescale of estrogen effects and what constitutes an estrogen target gene. However, this approach presents a number of problems, including the toxicity of cycloheximide, as well as a failure to inhibit the actions of non-coding regulatory RNAs, which may impact gene regulation by ligand-activated ERs. Furthermore, technical issues with expression microarrays, such as high levels of noise, relatively low sensitivity, and poor coverage of non-coding RNAs in traditional array platforms, suggest that they are not the best approach for defining direct target genes. RNA expression analyses using next generation sequencing approaches (e.g., RNA-seq) have circumvented many of the issues with microarrays, providing a sensitive, robust and unbiased approach for determining the effects of estrogen signaling on the steady-state levels of mRNAs [44–46].

1.4.2. ChIP-based genomic localization analyses—Genomic localization analyses, in which genomic ER binding sites are mapped using ChIP-based technologies, have been used in an attempt to define direct ER-regulated target genes [5,18,27]. These studies have provided a wealth of information about the location, nature, molecular mechanisms, and functions of ER enhancers. Studies using ChIP-chip or ChIP-seq have found that ER binds to between 5,000 and 10,000 locations across the genome [32,34], perhaps as many as 30,000 [47], depending on the sensitivity of the assay and how the data are analyzed. The available ER genomic localization studies have provided a surprisingly consistent view of the general patterns of ER binding across the genome. Unexpectedly, only <5% of ER binding sites defined by these studies are located within 1 kb of the annotated transcription start sites, with the majority of sites located distally to gene promoters, many by as much as 100 kb [32,34,48] (Fig. 1). Although these intergenic ER binding sites are functional and can mediate estrogen-dependent gene expression, their large number and distal locations make it difficult to use them by themselves as a means to define direct estrogen-regulated target promoters. ChIP-based analyses of direct ER target genes have been greatly facilitated by genomic looping analyses (e.g., ER-based ChIA-PET), which can connect specific ER binding events with target gene promoters [33].

Studies using ChIP-chip and ChIP-seq to monitor the estrogen-dependent changes in Pol II localization as a surrogate of defining direct estrogen-regulated target promoters have also provided new insights [34]. However, the technical aspects of ChIP do not distinguish productive Pol II from preinitiation complex or pre-loaded Pol II at estrogen-regulated gene promoters, and much of the Pol II localization were detected beyond annotated transcription start sites. ChIP with antibodies to specific phosphorylated forms of Pol II (e.g., Ser5 or Ser2 phosphorylation of the heptapeptide repeat in the carboxyl-terminal domain of the largest subunit of Pol II) can help to distinguish newly loaded and elongating forms of Pol II, respectively, which can provide additional information about the impact of estrogen signaling on the transcription process [40,49]. In addition, Pol II-based ChIA-PET has also aided in the identification of enhancer events that lead to functional gene regulatory outcomes [50].

2. Global Analyses of the Estrogen-Regulated Transcriptome Using GRO-seq

Previous genomic analyses of steroid hormone regulated transcription have provided a wealth of information about the global actions of the hormones and their cognate receptors [18]. Although these studies have revealed new information about the mode of receptor binding and the factors that cooperate with the receptors as they access the chromatin template, they have fallen short of defining the primary/immediate response transcriptomes and the specific molecular mechanisms of hormone-dependent transcription by Pol II. New genomic approaches are providing opportunities to (1) identify primary response genes, (2) establish the hierarchy of network interactions that characterize secondary changes in gene expression, and (3) identify novel, unannotated, non-coding RNAs. In the space below, we discuss the use of global nuclear run-on coupled with massively parallel sequencing (GRO-seq) to explore new facets of the steroid hormone regulated transcriptome, focusing primarily on results obtained from studying the estrogen signaling pathway in breast cancer cells.

2.1. Characterizing transcriptomes using GRO-seq

GRO-seq is a direct, high throughput sequencing method adapted from conventional nuclear run-on methodologies, which is used to obtain a map of the position and orientation of all transcriptionally engaged RNA polymerases across the genome with extremely high spatial resolution [51]. Given its high sensitivity and temporal resolution, GRO-seq is an useful approach to examine transcriptional responses to extracellular stimuli. In the protocol for GRO-seq, nuclei are isolated from cells or tissues after stimulation or under certain physiological, disease, or designed experimental conditions. Isolated nuclei are subjected to a short run-on in the presence of a nucleotide analog, Bromo-UTP (brUTP), allowing RNA polymerases to transcribe approximately 100 bases of nascent, brUTP-labeled RNA (Fig. 2). The run-on reactions are performed in the presence of sarkosyl or high salt to prevent transcription initiation by unengaged RNA polymerases and to remove impediments to elongation for actively engaged and transcribing RNA polymerases. RNA is collected from the nuclei, base-hydrolyzed, and subjected to immunoprecipitated with brUTP antibody-conjugated beads to enrich for the labeled nascent RNAs. The nascent RNAs are ligated to RNA adaptors, reverse transcribed, amplified, and subjected to deep sequencing [51] (Fig. 2). The sequenced short reads are mapped to a reference genome, which provides information about regions of the annotated and unannotated genome that are actively transcribed (Fig. 3). Novel transcription units can be defined as discrete, contiguous regions of transcription using bioinformatic approaches, including a two-state hidden Markov model (HMM) [52].

GRO-seq has advantages that make it uniquely suited to serve as a general method for assessing changes in transcription caused by estrogen stimulation [52,53]. First, GRO-seq is able to resolve instantaneous changes in the recruitment of RNA polymerases I, II and III. This rapid temporal resolution can be used to identify the most immediate transcriptional effects of estrogen signaling. Second, GRO-seq is a direct measure of transcription. In contrast, methods that use RNA abundance as a measure of transcription, based either on microarray or deep sequencing technologies, measure a complex function of both RNA stability and processing. Third, unlike RNA polymerase ChIP-seq, GRO-seq identifies the orientation of transcription, allowing the detection of antisense and divergent transcripts, which are clearly a significant fraction of the transcriptome and may have important regulatory roles. Fourth, GRO-seq allows the identification and characterization of transcripts lacking existing annotations in an unbiased way, which is also significant for

detecting novel transcripts. Its high sensitivity allows the detection of transcribed regions that are not readily detected using steady-state RNA analyses.

2.2. Effects of estrogen on the breast cancer transcriptome

GRO-seq has been used to comprehensively assay the effects of estrogen signaling on the transcriptome of the ER⁺-positive MCF-7 human breast cancer cell line, revealing new facets of estrogen-dependent transcriptional regulation [52,54,55]. Unlike other approaches, GRO-seq measures the position and direction of transcriptionally competent polymerases (Pols I, II, and III) in the genome at extremely high resolution (Fig. 3). This map is indicative of instantaneous transcriptional activity at a given locus. Thus, it allows the detection of any transcripts that are generated and regulated as primary endpoints of estrogen signaling in unbiased way. Remarkably, about one quarter of the >30,000 primary transcripts detected by GRO-seq in MCF-7 cells are regulated by estrogen signaling, and the effects on the transcriptome are widespread across transcript types [52]. Estrogen rapidly and robustly regulates not only annotated transcripts, such as protein coding mRNAs and non-coding regulatory RNAs, but also unannotated transcripts, such as antisense, divergent, enhancer, and intergenic RNAs (Fig. 4) [52]. GRO-seq provides a comprehensive picture of the effects of rapid estrogen signaling on the transcriptome. In the space below, we discuss the effects of estrogen signaling on different classes of estrogen-regulated transcripts and their possible effects on the biology of estrogen signaling.

2.2.1. Protein-coding transcripts (messenger RNAs)—Many studies have examined changes in the steady-state mRNA levels upon steroid hormone stimulation using expression microarrays. In fact, time courses of estrogen treatment in MCF-7 cells are one of the most commonly performed experiments in the published nuclear receptor literature [18,27]. These studies have typically examined treatment times on the order of hours to days, with the shortest treatments in the range of 2 to 6 hours. The results from expression microarray analyses performed by different groups have set the number of estrogen-regulated genes anywhere from ~100 to ~1500 [27,32,43]. Interestingly, expression microarray experiments identified only a small number of genes (< 100) that were directly regulated by estrogen within an hour of treatment [43,56]. Even though the number of regulated genes that were identified varies from laboratory to laboratory, expression microarray studies have indicated that the number of genes called up or down-regulated genes steadily increases as over the time course of estrogen treatment. Interestingly, microarray studies have shown that estrogen-dependent up-regulation of gene expression occurs within 1 to 8 hours of estrogen treatment, whereas down-regulation occurs with sustained estrogen treatment [32,48].

In contrast to expression microarray experiments, GRO-seq has allowed the detection of immediate effects of estrogen signaling on Pol II transcription, on the order of minutes to hours [52]. Approximately 3,000 protein coding transcripts are regulated by relatively short treatments with E2 in MCF-7 cells, representing ~15% of all annotated RefSeq genes and ~33% of expressed RefSeq genes in the MCF-7 transcriptome. This is a significantly larger fraction than what has been determined previously using expression microarrays with longer treatment times, reflecting differences in the sensitivity of both approaches, as well transcripts with rapid and transient inductions that are not detected as E2-regulated at later treatment time points.

Hierarchical clustering has defined classes of estrogen-regulated transcripts that share similar patterns of regulation based on GRO-seq data, including (1) a class of rapidly down-regulated transcripts (within 10 minutes), which comprises ~50% of the E2-regulated protein coding transcripts, (2) two classes of transcripts with maximal transcription at relatively short E2 treatment time points (10 or 40 min), and (3) one class of transcripts with

continued increases in transcription at a longer E2 treatment time point (160 minutes) (Fig. 5A) [52]. Gene ontology analyses on each class of genes revealed some unique aspects of estrogen signaling in breast cancer cells (Fig. 5B). Notably, both rapidly down-regulated genes and genes that reach maximal transcription at 40 min share the similar gene ontologies, such as transcription, nucleic acid metabolism, cell surface receptors, and G protein-coupled signaling pathways [52]. Furthermore, the genes that steadily increase their expression over the time course of estrogen treatment are enriched in ontologies related to translation, ribosome biogenesis, and protein synthesis. These results highlight the essence of a mitogenic transcriptional response, where estrogen signaling prepares the cells for continued signaling and transcriptional responses in the short term, as well as the translation of protein effectors from newly expressed mRNAs in the long term. Interestingly, gene ontologies related to cell cycle regulation are not significantly enriched in the primary estrogen target gene sets [52], perhaps suggesting an indirect regulatory mechanism via GPCR signaling pathways.

Studies mapping ER binding sites defined by ChIP-seq relative to transcripts defined by GRO-seq give clues to the underlying mechanisms of regulation. Previous studies indicated that most ER binding sites are located distally from the promoters of estrogen-regulated target genes, in some cases more than 100 kb away [28,32,34,48]. Defining the primary targets of estrogen signaling using GRO-seq, however, has shown that about half of all estrogen target genes immediately up-regulated by short (i.e., 10 to 40 min) treatments with E2 have ER binding sites within 10 kb of the transcription start site [52]. In contrast, down-regulated genes are not enriched for promoter-proximal ER binding despite the rapid down-regulation of transcription [52]. The latter result suggests that (1) down-regulation may not be a direct effect of the estrogen signaling pathway and (2) the rapid down-regulation of the genes in response to E2 may be due to a passive mechanism where the transcription machinery is withdrawn from the down-regulated loci to facilitate activated responses elsewhere. FoxA1, a pioneer factor for ER [28,32], is strongly enriched at functional, but not non-functional, ER binding sites [55]. Whether FoxA1 binding distinguishes immediately up- and down-regulated estrogen target genes has not yet been determined, but may provide additional insights about the mechanisms of ER-dependent gene regulation.

Collectively, GRO-seq-based studies of estrogen-regulated protein-coding gene expression indicate that signaling-dependent transcriptional responses are coordinated in time-dependent manner. Recent studies have applied GRO-seq to the study of other signal-regulated transcriptional responses, including Toll-like receptor 4 (TLR4)-dependent gene expression in macrophages revealing temporally distinct patterns of TLR4-dependent gene activation required for homeostasis and immune responses [57], again illustrating the utility of GRO-seq for these types of analyses.

2.2.2. Primary microRNA transcripts—In addition to mRNAs, GRO-seq also provides information about the transcriptional regulation of primary microRNA transcripts, which are processed to yield microRNAs [52]. MicroRNAs are ~22 nt noncoding regulatory RNAs that act to inhibit the translation or promote the degradation of target mRNAs, thereby mediating the posttranscriptional regulation of gene expression [58]. MicroRNA transcripts are synthesized by Pol II (primarily) or Pol III (in some cases), either as part of a “host” gene in which they are embedded or from an intergenic region using their own promoter [58] (Fig. 4). Transcription of the latter can be unambiguously detected by GRO-seq, while transcription of the former cannot usually be distinguished from transcription of mRNA genes containing microRNAs due to overlapping signals in the assay [52]. The pattern of estrogen regulation of primary microRNA transcripts mirrors that observed for protein-coding transcripts (i.e., approximately half up-regulated and half down-regulated), with the

regulation occurring on a similar time scale. Interestingly, GRO-seq analyses have revealed a coordinated response between estrogen-regulated primary microRNA transcripts and the protein-coding genes that they ultimately regulate [52]. This mode of regulation allows the estrogen signaling pathway to reinforce patterns of regulation by targeting mRNA transcripts in two ways – through transcriptional effects on mRNA genes and on the microRNA genes that target those mRNAs.

2.2.3. Enhancer transcripts at ER α binding sites (eRNAs)—Since the initial discovery and characterization of ER binding sites (ERBS), great efforts have been made to understand the features and functions of these ER enhancers. High throughput genomic technologies have shown that a majority of ERBSs are distally located from target genes and likely function as distal *cis*-regulatory elements to regulate target genes [32,59]. Studies with other sequence-specific DNA binding transcription factors have provided useful parallels for understanding the function of ER enhancers. Enhancers are thought to promote communication with promoters of target gene through (1) chromatin loops or (2) tracking of enhancer-bound transcription factors through intervening chromatin to the promoters [60–62]. Recent studies have shown that many of these enhancer regions overlap with sites of active transcription by RNA pol II, which produces short non-coding transcripts called enhancer RNAs (“eRNAs”) [52,55,63–66]. Recent studies have shown that eRNAs may play a key role in the activity of enhancers and the expression of their cognate target genes [65,67–69].

GRO-seq, which has been a useful tool for identifying and characterizing enhancer transcripts [52,55,66,69], has shown that the genomic binding sites for ER and other steroid receptors overlap with sites of hormone-induced transcription [52,55,66,69] (Figs. 3B and 4). Some transcripts produced from ERBSs appear to be long non-coding RNAs, while others are relatively short (~3 to 5 kb) and are bidirectionally transcribed [52,55]. The expression of the latter, which are more typical of the original definition of eRNAs (i.e., short, bidirectional; [65]), are predominately up-regulated upon estrogen stimulation with kinetics that precede or match ER binding and the induction of the target genes [55,69]. The production of eRNAs at ERBSs is strongly correlated with the enrichment of a number of genomic features that have been shown to be associated with enhancers (e.g., H3K4me1, H3K27ac, p300/CBP, RNA pol II, and an open chromatin architecture), as well as enhancer looping to target gene promoters [55,69]. Androgen receptor (AR) enhancer transcription has also been explored using GRO-seq, revealing that cell-lineage-specific factors (e.g., the “pioneer” factor FoxA1) can both facilitate and restrict AR actions on structurally and functionally distinct classes of enhancers, including a class of enhancers that do not require nucleosome remodeling to induce specific enhancer-promoter looping and gene activation [66]. GRO-seq-identified enhancer transcription also provides a unique unbiased tool for de novo identification of functional enhancers in any cell type [55].

In spite of our ability to identify active/dynamic and steady-state eRNA production, we still know little about the molecular mechanisms of enhancer function. In addition, we lack a comprehensive and integrated view of enhancer chromatin features, eRNA production, and gene looping events in a signal-regulated gene regulatory system. This information is needed to understand the molecular mechanisms of action of signal-regulated enhancers and their cell type-specific functions. Nonetheless, recent studies have begun to provide some clues. The act of transcription may help to create an open chromatin environment that promotes enhancer function [70]. Alternatively, the stable accumulation of eRNAs may play a functional, perhaps even structural, role and may facilitate gene looping [65,67,70–73]. Recent studies of ER enhancers in MCF-7 cells and RevErb enhancers in macrophages suggest a functional role for stable eRNAs in promoter looping and target gene activation [68,69]. Further studies in this area are required.

2.2.4. Long non-coding transcripts (lncRNAs)—GRO-seq has also been a useful technique for identifying long non-coding RNAs (lncRNAs), a class of transcripts that are transcribed from intergenic regions in genome and are, by definition, >200 nucleotides in length [74,75]. Many of these Pol II transcripts are 5' capped, polyadenylated, and spliced, like mRNAs, but have little or no coding potential. Some lncRNAs are primarily cytoplasmic, while others are primarily nuclear or chromatin-associated. They have been implicated in the regulation of histone modifications, chromatin structure and looping, gene expression, RNA processing and stability, and microRNA function. Biologically, lncRNAs are involved in developmental and physiological processes, as well as disease states, such as cancer [75,76].

The sensitivity of GRO-seq has allowed the detection of lncRNAs that were not readily detected by analyses of steady-state RNA analyses (e.g., RNA-seq) (Fig. 4) [52,77,78]. These studies have revealed that while lncRNA genes are transcribed at about the same level, on average, as protein-coding genes, lncRNAs are less stable as a class than mRNAs [52,77–79]. The combination with GRO-seq and RNA-seq has allowed for the identification, annotation, and characterization of over two thousand lncRNAs in MCF-7 cells, many of which are estrogen-regulated [77,78]. Most estrogen-upregulated lncRNAs have ERBSs in the proximal promoter regions of the genes that code for them and they are regulated by estrogen with similar kinetics as the estrogen-regulated mRNAs described above. The expression of many of these lncRNAs correlates with the expression of mRNAs that code for proteins involved in cell proliferation and cancer cell signaling, perhaps suggesting related functions for the lncRNAs through so-called “guilt-by-association” analyses [77,78]. Interestingly, knockdown of some of these lncRNAs robustly alters the estrogen-dependent proliferation of breast cancer cells in culture, suggesting potential roles in breast cancer cell proliferation, as well as potential value as biomarkers of breast cancer [77,78]. Further studies are needed to assess the role of estrogen-regulated lncRNAs in physiological processes. Recent studies have shown, however, that other steroid hormone-regulated lncRNAs can play key roles in physiological processes [80–85]. This should be a rich area for further exploration into the detailed mechanisms controlling the biological response to estrogen signaling.

2.2.5. Other unannotated non-coding transcripts: Divergent and antisense transcripts—An advantage of GRO-seq is the ability to identify and characterize novel transcripts, which has allowed determining the types of unannotated non-coding transcripts in cells, including divergent and antisense RNAs. Divergent transcripts, which are short Pol II transcripts that initiate from the promoter regions of Pol II genes, but run in the opposite direction, were described in a series of recent studies using GRO-seq and other genomic methods [51,79,86] (Figs. 3A and 4). Although divergent transcription was previously observed in yeast and mammalian cells using gene-specific approaches, these studies using high-throughput sequencing approaches have facilitated the identification and characterization of divergent transcripts. Studies analyzing short RNAs from mouse embryonic stem cells and nascent RNAs from human fibroblasts indicated higher RNA densities at the promoters of actively expressed genes, some of which was due to antisense RNAs that map upstream of the promoter and originate on the opposite strand of the transcribed canonical promoter [51,79,86]. The prevalence of divergent transcription at active promoters has led to the suggestion that it may play a regulatory role in gene expression. Another study that used RNA-seq and GRO-seq to identify lncRNAs de novo has shown that many lncRNAs originate from divergently transcribed protein-coding gene promoters, which may facilitate coordinate regulation of gene expression for the lncRNAs and their corresponding protein coding genes [79]. Interestingly, short divergent transcripts are also observed at enhancers, where transcripts are generated bidirectionally, as described above [52,79]. Divergent transcripts at promoters and enhancers are regulated by estrogen

stimulation in MCF-7 cells, generally showing similar magnitudes and dynamics of expression as the mRNAs to which they are divergent [52]. Although the functions of divergent transcripts are largely unknown, they may play a role in (1) generating open chromatin structures at active gene promoters (or enhancers) by creating nucleosome-free regions, (2) introducing negative superhelical tension in promoter DNA, or (3) promoting transcription initiation by additional RNA polymerases.

Antisense transcripts are non-coding (or possibly even coding) RNAs that are transcribed (typically by Pol II) from the opposite strand of annotated protein-coding genes [87,88] (Fig. 4). Like divergent transcripts, the functions of antisense transcripts are not well characterized. Most appear to be non-coding and some may simply be lncRNAs that happen to be transcribed antisense to mRNAs. Antisense transcription may help to generate open chromatin architecture at actively transcribed loci and antisense transcripts may hybridize to sense transcripts to regulate their stability and function. Estrogen regulates the expression of antisense transcripts with similar kinetics and magnitudes as mRNAs and lncRNAs. Further studies are required, however, to elucidate the functions of these transcripts within the estrogen signaling pathway.

2.3. Dynamics of estrogen-dependent Pol II transcription

The high sensitivity of GRO-seq provides outstanding temporal resolution of hormone-dependent transcriptional responses during time course experiments, allowing detection of nearly instantaneous changes in the recruitment or activity of Pol II (or Pols I or III). Studies in MCF-7 cells have shown that estrogen signaling propagates very rapidly to the genome [52]. The earliest transcriptional responses can be detected by GRO-seq within minutes (Fig. 5A). A striking feature of these results is that estrogen-regulated transcription is rapid and transient. This is illustrated by protein-coding transcripts, which are immediately up- or down-regulated upon estrogen stimulation (within 10 minutes), as well as transiently up-regulated transcripts that reach maximal expression within an hour and return to the basal levels within 3 hours [52]. The extent of the rapid and transient dynamics of gene expression regulation was not previously detected using expression microarrays, even though the changes in transcription detected by GRO-seq are accompanied by similar, but delayed, changes in the steady-state levels of the mRNAs [52]. These results suggest that much of the early actions of E2 are mediated at the level of transcription, rather than on RNA stability or degradation. Genes whose transcription is rapidly regulated are most likely primary (or direct) targets of the estrogen signaling pathway.

GRO-seq also provides a dynamic view of the initial “wave” of Pol II following up- or down-regulation of a gene in response to estrogen stimulation. The leading or lagging edge of the Pol II wave is clearly resolved as Pol II transcribes through the gene body over time (Fig. 6). The distance traveled over time for the Pol II wave has allowed accurate determination of Pol II transcription rates across hundreds of genes in MCF-7 cells [39,52]. These studies have shown that transcription rates can vary as much as 4-fold for different genes. Interestingly, the rates are slowest near the promoter and increase during the first ~15 kb transcribed by Pol II. In addition, the rates correlate with Pol II density, resulting in higher transcription rates for genes with higher Pol II density. Estrogen signaling stimulates gene expression by increasing Pol II loading and or initiation, whereas another signaling pathway (i.e., TNF signaling) reduces Pol II residence time at promoter-proximal pause sites [54]. These results have helped to identify unexpected variation in the rate of transcription in response to estrogen signaling and provide new insights about the molecular mechanisms of estrogen signaling. Together, the observation made using GRO-seq have provided an unprecedented picture of kinetics hormone-regulated transcription by Pol II.

2.4. Estrogen-regulated Pol I and Pol III transcription

As noted above, GRO-seq allows the detection of transcription by Pools I and III, in addition to Pool II. Studies in MCF-7 cells have shown that estrogen rapidly and robustly up-regulates the transcription of ribosomal RNAs (rRNAs; Pool I) and transfer RNAs (tRNAs; Pool III) [52]. In fact, tRNA production from about one third of the approximately 500 tRNA genes in humans is upregulated in response to short treatments with estrogen and generally continues to increase over time. At least one tRNA for each of the 20 amino acids is represented in the upregulated set. Furthermore, estrogen up-regulates the expression of a large fraction of mRNAs coding for proteins with gene ontologies suggesting a role in protein biosynthesis (e.g., ribosome biogenesis, ribosome cellular compartment, ribosome biogenesis, rRNA metabolic process, rRNA processing, tRNA aminoacylation, and tRNA processing), with continued up regulation over hours of treatment. Together, these results demonstrate a potent effect of estrogen signaling on the protein biosynthetic machinery, including strong, immediate, and likely direct effects on transcription by all three RNA polymerases.

2.5. The Logic of an Estrogen-Regulated Mitogenic Transcriptional Response

The detailed analyses of the estrogen-regulated transcriptome described above have revealed the logic of the estrogen-dependent mitogenic transcriptional response (Fig. 7). The most immediate effects of estrogen signaling on the genome results in the regulation of mRNAs encoding proteins involved in transcription, nucleic acid metabolism, and G-protein-coupled and cell surface signaling [52]. In this way, estrogen signaling propagates the hormone-dependent transcriptional response, leading to secondary and sustained effects. Over the longer term, estrogen signaling up-regulates the protein biosynthetic machinery as described above. This is likely a means by which the estrogen signaling pathway prepares the cell for translation of the mRNAs that are newly synthesized in response to estrogen signaling. The immediate and sustained effects of estrogen signaling on the expression of mRNAs, rRNAs, and tRNAs described here underlie the mitogenic effects of estrogen signaling in breast cancers. The extent to which the global effects of estrogen on the expression of other non-coding RNAs, such as lncRNAs and microRNAs, contribute to the mitogenic response have not been investigated in detail, but gene-specific studies have suggested that they are likely to play a role as well [89–91].

3. Conclusions and Future Perspectives

GRO-seq in combination with new bioinformatic approaches has allowed the detection of estrogen-regulated transcription by RNA Pools I, II, and III at both annotated and unannotated genomic regions in MCF-7 breast cancer cells. By focusing on short treatment times with estrogen, prior to the activation of secondary targets, has allowed the identification of the primary target genes of estrogen signaling. This unique approach has revealed a picture of estrogen-regulated gene expression that differs considerably from previous views based on expression microarray studies and genomic ChIP-based studies. A striking finding is that estrogen can regulate a surprisingly large fraction (~26%) of the breast cancer transcriptome in a rapid, robust, and transient manner. In addition to protein coding transcripts, nearly every other class of transcript that has been described to date is also regulated by estrogen, including annotated and unannotated non-coding transcripts. GRO-seq analyses are useful for studying the mechanisms of estrogen-regulated transcription as well, allowing determination of the rates of Pool II transcription in cells. Collectively, the results from GRO-seq analyses in MCF-7 cells has revealed new levels of estrogen-dependent transcriptional regulation, providing the most comprehensive measurement of the primary and immediate effects of estrogen signaling to date.

Acknowledgments

The authors would like to thank members of the Kraus lab for providing critical comments on the review. The work in the lab of W.L.K. is supported by grants from the NIH and the Texas Cancer Prevention and Research Institute of Texas.

Abbreviations

ChIP	chromatin immunoprecipitation
E2	17 β -estradiol
ER	estrogen receptor alpha
ERBS	estrogen receptor binding site
eRNA	enhancer RNA
GRO-seq	global nuclear run-on and sequencing
HMM	hidden Markov model lncRNA, long non-coding RNAs
mRNA	messenger RNA
Pol (I/II/III)	RNA polymerase (I, II, III)
rRNA	ribosomal RNA
SERM	selective estrogen-receptor modulator
tRNA	transfer RNAs

References

1. Deroo BJ, Korach KS. Estrogen receptors and human disease. *J Clin Invest.* 2006; 116:561–70. [PubMed: 16511588]
2. Cadenas C, Bolt HM. Estrogen receptors in human disease. *Arch Toxicol.* 2012; 86:1489–90. [PubMed: 22941009]
3. Bjornstrom L, Sjoberg M. Mechanisms of estrogen receptor signaling: convergence of genomic and nongenomic actions on target genes. *Mol Endocrinol.* 2005; 19:833–42. [PubMed: 15695368]
4. Levin ER. Membrane oestrogen receptor alpha signalling to cell functions. *J Physiol.* 2009; 587:5019–23. [PubMed: 19687123]
5. Welboren WJ, Sweep FC, Span PN, Stunnenberg HG. Genomic actions of estrogen receptor alpha: what are the targets and how are they regulated? *Endocr Relat Cancer.* 2009; 16:1073–89. [PubMed: 19628648]
6. Warner M, Nilsson S, Gustafsson JA. The estrogen receptor family. *Curr Opin Obstet Gynecol.* 1999; 11:249–54. [PubMed: 10369199]
7. Bai Z, Gust R. Breast cancer, estrogen receptor and ligands. *Arch Pharm (Weinheim).* 2009; 342:133–49. [PubMed: 19274700]
8. Cheung E, Schwabish MA, Kraus WL. Chromatin exposes intrinsic differences in the transcriptional activities of estrogen receptors alpha and beta. *EMBO J.* 2003; 22:600–11. [PubMed: 12554660]
9. Harrington WR, Sheng S, Barnett DH, Petz LN, Katzenellenbogen JA, Katzenellenbogen BS. Activities of estrogen receptor alpha- and beta-selective ligands at diverse estrogen responsive gene sites mediating transactivation or transrepression. *Mol Cell Endocrinol.* 2003; 206:13–22. [PubMed: 12943986]
10. Katzenellenbogen BS, Katzenellenbogen JA. Estrogen receptor transcription and transactivation: Estrogen receptor alpha and estrogen receptor beta: regulation by selective estrogen receptor modulators and importance in breast cancer. *Breast Cancer Res.* 2000; 2:335–44. [PubMed: 11250726]

11. Pearce ST, Jordan VC. The biological role of estrogen receptors alpha and beta in cancer. *Crit Rev Oncol Hematol*. 2004; 50:3–22. [PubMed: 15094156]
12. Pfaffl MW, Lange IG, Daxenberger A, Meyer HH. Tissue-specific expression pattern of estrogen receptors (ER): quantification of ER alpha and ER beta mRNA with real-time RT-PCR. *APMIS*. 2001; 109:345–55. [PubMed: 11478682]
13. Platet N, Cathiard AM, Gleizes M, Garcia M. Estrogens and their receptors in breast cancer progression: a dual role in cancer proliferation and invasion. *Crit Rev Oncol Hematol*. 2004; 51:55–67. [PubMed: 15207254]
14. Prall OW, Rogan EM, Sutherland RL. Estrogen regulation of cell cycle progression in breast cancer cells. *J Steroid Biochem Mol Biol*. 1998; 65:169–74. [PubMed: 9699870]
15. Rochefort H, Platet N, Hayashido Y, Derocq D, Lucas A, Cunat S, Garcia M. Estrogen receptor mediated inhibition of cancer cell invasion and motility: an overview. *J Steroid Biochem Mol Biol*. 1998; 65:163–8. [PubMed: 9699869]
16. Osborne CK, Schiff R. Estrogen-receptor biology: continuing progress and therapeutic implications. *J Clin Oncol*. 2005; 23:1616–22. [PubMed: 15755967]
17. Cazzaniga M, Bonanni B. Breast cancer chemoprevention: old and new approaches. *J Biomed Biotechnol*. 2012; 2012:985620. [PubMed: 22851887]
18. Cheung E, Kraus WL. Genomic analyses of hormone signaling and gene regulation. *Annu Rev Physiol*. 2010; 72:191–218. [PubMed: 20148673]
19. Hall JM, Couse JF, Korach KS. The multifaceted mechanisms of estradiol and estrogen receptor signaling. *J Biol Chem*. 2001; 276:36869–72. [PubMed: 11459850]
20. Hammes SR, Levin ER. Extranuclear steroid receptors: nature and actions. *Endocr Rev*. 2007; 28:726–41. [PubMed: 17916740]
21. Hammes SR, Levin ER. Minireview: Recent advances in extranuclear steroid receptor actions. *Endocrinology*. 2011; 152:4489–95. [PubMed: 22028449]
22. Berry M, Nunez AM, Chambon P. Estrogen-responsive element of the human pS2 gene is an imperfectly palindromic sequence. *Proc Natl Acad Sci U S A*. 1989; 86:1218–22. [PubMed: 2919170]
23. Kumar V, Chambon P. The estrogen receptor binds tightly to its responsive element as a ligand-induced homodimer. *Cell*. 1988; 55:145–56. [PubMed: 3167974]
24. Heldring N, Isaacs GD, Diehl AG, Sun M, Cheung E, Ranish JA, Kraus WL. Multiple sequence-specific DNA-binding proteins mediate estrogen receptor signaling through a tethering pathway. *Mol Endocrinol*. 2011; 25:564–74. [PubMed: 21330404]
25. Kushner PJ, Agard DA, Greene GL, Scanlan TS, Shiau AK, Uht RM, Webb P. Estrogen receptor pathways to AP-1. *J Steroid Biochem Mol Biol*. 2000; 74:311–7. [PubMed: 11162939]
26. Stender JD, Kim K, Charn TH, Komm B, Chang KC, Kraus WL, Benner C, Glass CK, Katzenellenbogen BS. Genome-wide analysis of estrogen receptor alpha DNA binding and tethering mechanisms identifies Runx1 as a novel tethering factor in receptor-mediated transcriptional activation. *Mol Cell Biol*. 2010; 30:3943–55. [PubMed: 20547749]
27. Kininis M, Kraus WL. A global view of transcriptional regulation by nuclear receptors: gene expression, factor localization, and DNA sequence analysis. *Nucl Recept Signal*. 2008; 6:e005. [PubMed: 18301785]
28. Carroll JS, Liu XS, Brodsky AS, Li W, Meyer CA, Szary AJ, Eeckhoutte J, Shao W, Hestermann EV, Geistlinger TR, Fox EA, Silver PA, Brown M. Chromosome-wide mapping of estrogen receptor binding reveals long-range regulation requiring the forkhead protein FoxA1. *Cell*. 2005; 122:33–43. [PubMed: 16009131]
29. Hsu PY, Hsu HK, Singer GA, Yan PS, Rodriguez BA, Liu JC, Weng YI, Deatherage DE, Chen Z, Pereira JS, Lopez R, Russo J, Wang Q, Lamartiniere CA, Nephew KP, Huang TH. Estrogen-mediated epigenetic repression of large chromosomal regions through DNA looping. *Genome Res*. 2010; 20:733–44. [PubMed: 20442245]
30. Pan YF, Wansa KD, Liu MH, Zhao B, Hong SZ, Tan PY, Lim KS, Bourque G, Liu ET, Cheung E. Regulation of estrogen receptor-mediated long range transcription via evolutionarily conserved distal response elements. *J Biol Chem*. 2008; 283:32977–88. [PubMed: 18728018]

31. Tan SK, Lin ZH, Chang CW, Varang V, Chng KR, Pan YF, Yong EL, Sung WK, Cheung E. AP-2gamma regulates oestrogen receptor-mediated long-range chromatin interaction and gene transcription. *EMBO J.* 2011; 30:2569–81. [PubMed: 21572391]
32. Carroll JS, Meyer CA, Song J, Li W, Geistlinger TR, Eeckhoute J, Brodsky AS, Keeton EK, Fertuck KC, Hall GF, Wang Q, Bekiranov S, Sementchenko V, Fox EA, Silver PA, Gingeras TR, Liu XS, Brown M. Genome-wide analysis of estrogen receptor binding sites. *Nat Genet.* 2006; 38:1289–97. [PubMed: 17013392]
33. Fullwood MJ, Liu MH, Pan YF, Liu J, Xu H, Mohamed YB, Orlov YL, Velkov S, Ho A, Mei PH, Chew EG, Huang PY, Welboren WJ, Han Y, Ooi HS, Ariyaratne PN, Vega VB, Luo Y, Tan PY, Choy PY, Wansa KD, Zhao B, Lim KS, Leow SC, Yow JS, Joseph R, Li H, Desai KV, Thomsen JS, Lee YK, Karuturi RK, Herve T, Bourque G, Stunnenberg HG, Ruan X, Cacheux-Rataboul V, Sung WK, Liu ET, Wei CL, Cheung E, Ruan Y. An oestrogen-receptor-alpha-bound human chromatin interactome. *Nature.* 2009; 462:58–64. [PubMed: 19890323]
34. Welboren WJ, van Driel MA, Janssen-Megens EM, van Heeringen SJ, Sweep FC, Span PN, Stunnenberg HG. ChIP-Seq of ERalpha and RNA polymerase II defines genes differentially responding to ligands. *EMBO J.* 2009; 28:1418–28. [PubMed: 19339991]
35. Carroll JS, Brown M. Estrogen receptor target gene: an evolving concept. *Mol Endocrinol.* 2006; 20:1707–14. [PubMed: 16396959]
36. Adelman K, Lis JT. Promoter-proximal pausing of RNA polymerase II: emerging roles in metazoans. *Nat Rev Genet.* 2012; 13:720–31. [PubMed: 22986266]
37. Boettiger AN, Levine M. Synchronous and stochastic patterns of gene activation in the *Drosophila* embryo. *Science.* 2009; 325:471–3. [PubMed: 19628867]
38. Aiyar SE, Sun JL, Blair AL, Moskaluk CA, Lu YZ, Ye QN, Yamaguchi Y, Mukherjee A, Ren DM, Handa H, Li R. Attenuation of estrogen receptor alpha-mediated transcription through estrogen-stimulated recruitment of a negative elongation factor. *Genes Dev.* 2004; 18:2134–46. [PubMed: 15342491]
39. Danko, CG.; Hah, N.; Luo, X.; Martins, AL.; Core, L.; Lis, JT.; Siepel, A.; Kraus, WL. Signaling Pathways Differentially Affect RNA Polymerase II Initiation, Pausing, and Elongation Rate in Cells. *Molecular Cell.* 2013. <http://dx.doi.org/10.1016/j.molcel.2013.02.015>
40. Kininis M, Isaacs GD, Core LJ, Hah N, Kraus WL. Postrecruitment regulation of RNA polymerase II directs rapid signaling responses at the promoters of estrogen target genes. *Mol Cell Biol.* 2009; 29:1123–33. [PubMed: 19103744]
41. Herschman HR. Primary response genes induced by growth factors and tumor promoters. *Annu Rev Biochem.* 1991; 60:281–319. [PubMed: 1883198]
42. Winkles JA. Serum- and polypeptide growth factor-inducible gene expression in mouse fibroblasts. *Prog Nucleic Acid Res Mol Biol.* 1998; 58:41–78. [PubMed: 9308363]
43. Lin CY, Strom A, Vega VB, Kong SL, Yeo AL, Thomsen JS, Chan WC, Doray B, Bangarusamy DK, Ramasamy A, Vergara LA, Tang S, Chong A, Bajic VB, Miller LD, Gustafsson JA, Liu ET. Discovery of estrogen receptor alpha target genes and response elements in breast tumor cells. *Genome Biol.* 2004; 5:R66. [PubMed: 15345050]
44. Sinicropi D, Qu K, Collin F, Crager M, Liu ML, Pelham RJ, Pho M, Rossi AD, Jeong J, Scott A, Ambannavar R, Zheng C, Mena R, Esteban J, Stephans J, Morlan J, Baker J. Whole transcriptome RNA-Seq analysis of breast cancer recurrence risk using formalin-fixed paraffin-embedded tumor tissue. *PLoS One.* 2012; 7:e40092. [PubMed: 22808097]
45. Yamaga R, Ikeda K, Horie-Inoue K, Ouchi Y, Suzuki Y, Inoue S. RNA sequencing of MCF-7 breast cancer cells identifies novel estrogen-responsive genes with functional estrogen receptor-binding sites in the vicinity of their transcription start sites. *Horm Cancer.* 2013; 1007/s12672-013-0140-3
46. Wang Z, Gerstein M, Snyder M. RNA-Seq: a revolutionary tool for transcriptomics. *Nat Rev Genet.* 2009; 10:57–63. [PubMed: 19015660]
47. Hurtado A, Holmes KA, Ross-Innes CS, Schmidt D, Carroll JS. FOXA1 is a key determinant of estrogen receptor function and endocrine response. *Nat Genet.* 2011; 43:27–33. [PubMed: 21151129]

48. Lin CY, Vega VB, Thomsen JS, Zhang T, Kong SL, Xie M, Chiu KP, Lipovich L, Barnett DH, Stossi F, Yeo A, George J, Kuznetsov VA, Lee YK, Charn TH, Palanisamy N, Miller LD, Cheung E, Katzenellenbogen BS, Ruan Y, Bourque G, Wei CL, Liu ET. Whole-genome cartography of estrogen receptor alpha binding sites. *PLoS Genet.* 2007; 3:e87. [PubMed: 17542648]
49. Heidemann M, Hintermair C, Voss K, Eick D. Dynamic phosphorylation patterns of RNA polymerase II CTD during transcription. *Biochim Biophys Acta.* 2013; 1829:55–62. [PubMed: 22982363]
50. Li G, Ruan X, Auerbach RK, Sandhu KS, Zheng M, Wang P, Poh HM, Goh Y, Lim J, Zhang J, Sim HS, Peh SQ, Mulawadi FH, Ong CT, Orlov YL, Hong S, Zhang Z, Landt S, Raha D, Euskirchen G, Wei CL, Ge W, Wang H, Davis C, Fisher-Aylor KI, Mortazavi A, Gerstein M, Gingeras T, Wold B, Sun Y, Fullwood MJ, Cheung E, Liu E, Sung WK, Snyder M, Ruan Y. Extensive promoter-centered chromatin interactions provide a topological basis for transcription regulation. *Cell.* 2012; 148:84–98. [PubMed: 22265404]
51. Core LJ, Waterfall JJ, Lis JT. Nascent RNA sequencing reveals widespread pausing and divergent initiation at human promoters. *Science.* 2008; 322:1845–8. [PubMed: 19056941]
52. Hah N, Danko CG, Core L, Waterfall JJ, Siepel A, Lis JT, Kraus WL. A rapid, extensive, and transient transcriptional response to estrogen signaling in breast cancer cells. *Cell.* 2011; 145:622–34. [PubMed: 21549415]
53. Core LJ, Lis JT. Transcription regulation through promoter-proximal pausing of RNA polymerase II. *Science.* 2008; 319:1791–2. [PubMed: 18369138]
54. Danko CG, Hah N, Luo X, Martins AL, Core L, Lis JT, Siepel A, Kraus WL. Signaling pathways differentially affect RNA polymerase II initiation, pausing, and elongation rate in cells. *Mol Cell.* 2013; 50:212–22. [PubMed: 23523369]
55. Hah N, Murakami S, Nagari A, Danko C, Kraus WL. Enhancer transcripts mark active estrogen receptor binding sites. *Genome Res.* 2013; 23. [PubMed: 23034409]
56. Creighton CJ, Cordero KE, Larios JM, Miller RS, Johnson MD, Chinnaiyan AM, Lippman ME, Rae JM. Genes regulated by estrogen in breast tumor cells in vitro are similarly regulated in vivo in tumor xenografts and human breast tumors. *Genome Biol.* 2006; 7:R28. [PubMed: 16606439]
57. Escoubet-Lozach L, Benner C, Kaikkonen MU, Lozach J, Heinz S, Spann NJ, Crotti A, Stender J, Ghisletti S, Reichart D, Cheng CS, Luna R, Ludka C, Sasik R, Garcia-Bassets I, Hoffmann A, Subramaniam S, Hardiman G, Rosenfeld MG, Glass CK. Mechanisms establishing TLR4-responsive activation states of inflammatory response genes. *PLoS Genet.* 2011; 7:e1002401. [PubMed: 22174696]
58. Krol J, Loedige I, Filipowicz W. The widespread regulation of microRNA biogenesis, function and decay. *Nat Rev Genet.* 2010; 11:597–610. [PubMed: 20661255]
59. Carroll RS, Brown M, Zhang J, DiRenzo J, Font De Mora J, Black PM. Expression of a subset of steroid receptor cofactors is associated with progesterone receptor expression in meningiomas. *Clin Cancer Res.* 2000; 6:3570–5. [PubMed: 10999746]
60. Bulger M, Groudine M. Functional and mechanistic diversity of distal transcription enhancers. *Cell.* 2011; 144:327–39. [PubMed: 21295696]
61. Kolovos P, Knoch TA, Grosveld FG, Cook PR, Papantonis A. Enhancers and silencers: an integrated and simple model for their function. *Epigenetics Chromatin.* 2012; 5:1. [PubMed: 22230046]
62. Ong CT, Corces VG. Enhancer function: new insights into the regulation of tissue-specific gene expression. *Nat Rev Genet.* 2011; 12:283–93. [PubMed: 21358745]
63. De Santa F, Barozzi I, Mietton F, Ghisletti S, Polletti S, Tusi BK, Muller H, Ragoussis J, Wei CL, Natoli G. A large fraction of extragenic RNA pol II transcription sites overlap enhancers. *PLoS Biol.* 2010; 8:e1000384. [PubMed: 20485488]
64. Djebali S, Davis CA, Merkel A, Dobin A, Lassmann T, Mortazavi A, Tanzer A, Lagarde J, Lin W, Schlesinger F, Xue C, Marinov GK, Khatun J, Williams BA, Zaleski C, Rozowsky J, Roder M, Kokocinski F, Abdelhamid RF, Alioto T, Antoshechkin I, Baer MT, Bar NS, Batut P, Bell K, Bell I, Chakraborty S, Chen X, Chrast J, Curado J, Derrien T, Drenkow J, Dumais E, Dumais J, Duttgupta R, Falconnet E, Fastuca M, Fejes-Toth K, Ferreira P, Foissac S, Fullwood MJ, Gao H, Gonzalez D, Gordon A, Gunawardena H, Howald C, Jha S, Johnson R, Kapranov P, King B,

- Kingswood C, Luo OJ, Park E, Persaud K, Preall JB, Ribeca P, Risk B, Robyr D, Sammeth M, Schaffer L, See LH, Shahab A, Skancke J, Suzuki AM, Takahashi H, Tilgner H, Trout D, Walters N, Wang H, Wrobel J, Yu Y, Ruan X, Hayashizaki Y, Harrow J, Gerstein M, Hubbard T, Reymond A, Antonarakis SE, Hannon G, Giddings MC, Ruan Y, Wold B, Carninci P, Guigo R, Gingeras TR. Landscape of transcription in human cells. *Nature*. 2012; 489:101–8. [PubMed: 22955620]
65. Kim TK, Hemberg M, Gray JM, Costa AM, Bear DM, Wu J, Harmin DA, Laptewicz M, Barbara-Haley K, Kuersten S, Markenscoff-Papadimitriou E, Kuhl D, Bito H, Worley PF, Kreiman G, Greenberg ME. Widespread transcription at neuronal activity-regulated enhancers. *Nature*. 2010; 465:182–7. [PubMed: 20393465]
 66. Wang D, Garcia-Bassets I, Benner C, Li W, Su X, Zhou Y, Qiu J, Liu W, Kaikkonen MU, Ohgi KA, Glass CK, Rosenfeld MG, Fu XD. Reprogramming transcription by distinct classes of enhancers functionally defined by eRNA. *Nature*. 2011; 474:390–4. [PubMed: 21572438]
 67. Melo CA, Drost J, Wijchers PJ, van de Werken H, de Wit E, Oude Vrielink JA, Elkon R, Melo SA, Leveille N, Kalluri R, de Laat W, Agami R. eRNAs Are Required for p53-Dependent Enhancer Activity and Gene Transcription. *Mol Cell*. 2013; 49:524–35. [PubMed: 23273978]
 68. Lam MT, Cho H, Lesch HP, Gosselin D, Heinz S, Tanaka-Oishi Y, Benner C, Kaikkonen MU, Kim AS, Kosaka M, Lee CY, Watt A, Grossman TR, Rosenfeld MG, Evans RM, Glass CK. Rev-Erbs repress macrophage gene expression by inhibiting enhancer-directed transcription. *Nature*. 2013
 69. Li W, Notani D, Ma Q, Tanasa B, Nunez E, Chen AY, Merkurjev D, Zhang J, Ohgi K, Song X, Oh S, Kim HS, Glass CK, Rosenfeld MG. Functional roles of enhancer RNAs for oestrogen-dependent transcriptional activation. *Nature*. 2013
 70. Natoli G, Andrau JC. Noncoding transcription at enhancers: general principles and functional models. *Annu Rev Genet*. 2012; 46:1–19. [PubMed: 22905871]
 71. Lai F, Orom UA, Cesaroni M, Beringer M, Taatjes DJ, Blobel GA, Shiekhattar R. Activating RNAs associate with Mediator to enhance chromatin architecture and transcription. *Nature*. 2013; 494:497–501. [PubMed: 23417068]
 72. Orom UA, Shiekhattar R. Noncoding RNAs and enhancers: complications of a long-distance relationship. *Trends Genet*. 2011; 27:433–9. [PubMed: 21831473]
 73. Orom UA, Shiekhattar R. Long non-coding RNAs and enhancers. *Curr Opin Genet Dev*. 2011; 21:194–8. [PubMed: 21330130]
 74. Mercer TR, Mattick JS. Structure and function of long noncoding RNAs in epigenetic regulation. *Nat Struct Mol Biol*. 2013; 20:300–7. [PubMed: 23463315]
 75. Rinn JL, Chang HY. Genome regulation by long noncoding RNAs. *Annu Rev Biochem*. 2012; 81:145–66. [PubMed: 22663078]
 76. Wapinski O, Chang HY. Long noncoding RNAs and human disease. *Trends Cell Biol*. 2011; 21:354–61. [PubMed: 21550244]
 77. Gadad SS, Sun M, Kraus WL. Identification and functional characterization of estrogen-regulated long noncoding RNAs in breast cancer cells: From genomics to molecular and cellular biology. *Endocr Rev*. 2012; 33:SUN-565.
 78. Sun M, Danko GG, Gadad SS, Hah N, Kraus WL. Characterization of estrogen-regulated non-coding RNAs. *Endocr Rev*. 2011; 32:OR14–2.
 79. Sigova AA, Mullen AC, Molinie B, Gupta S, Orlando DA, Guenther MG, Almada AE, Lin C, Sharp PA, Giallourakis CC, Young RA. Divergent transcription of long noncoding RNA/mRNA gene pairs in embryonic stem cells. *Proc Natl Acad Sci U S A*. 2013; 110:2876–81. [PubMed: 23382218]
 80. Askarian-Amiri ME, Crawford J, French JD, Smart CE, Smith MA, Clark MB, Ru K, Mercer TR, Thompson ER, Lakhani SR, Vargas AC, Campbell IG, Brown MA, Dinger ME, Mattick JS. SNORD-host RNA Zfas1 is a regulator of mammary development and a potential marker for breast cancer. *RNA*. 2011; 17:878–91. [PubMed: 21460236]
 81. Ginger MR, Shore AN, Contreras A, Rijnkels M, Miller J, Gonzalez-Rimbau MF, Rosen JM. A noncoding RNA is a potential marker of cell fate during mammary gland development. *Proc Natl Acad Sci U S A*. 2006; 103:5781–6. [PubMed: 16574773]

82. Kelly VR, Xu B, Kuick R, Koenig RJ, Hammer GD. Dax1 up-regulates Oct4 expression in mouse embryonic stem cells via LRH-1 and SRA. *Mol Endocrinol*. 2010; 24:2281–91. [PubMed: 20943815]
83. Kino T, Hurt DE, Ichijo T, Nader N, Chrousos GP. Noncoding RNA gas5 is a growth arrest- and starvation-associated repressor of the glucocorticoid receptor. *Sci Signal*. 2010; 3:ra8. [PubMed: 20124551]
84. Sun L, Goff LA, Trapnell C, Alexander R, Lo KA, Hacisuleyman E, Sauvageau M, Tazon-Vega B, Kelley DR, Hendrickson DG, Yuan B, Kellis M, Lodish HF, Rinn JL. Long noncoding RNAs regulate adipogenesis. *Proc Natl Acad Sci U S A*. 2013; 110:3387–92. [PubMed: 23401553]
85. Xu B, Gerin I, Miao H, Vu-Phan D, Johnson CN, Xu R, Chen XW, Cawthorn WP, MacDougald OA, Koenig RJ. Multiple roles for the non-coding RNA SRA in regulation of adipogenesis and insulin sensitivity. *PLoS One*. 2010; 5:e14199. [PubMed: 21152033]
86. Seila AC, Core LJ, Lis JT, Sharp PA. Divergent transcription: a new feature of active promoters. *Cell Cycle*. 2009; 8:2557–64. [PubMed: 19597342]
87. Bickel KS, Morris DR. Silencing the transcriptome's dark matter: mechanisms for suppressing translation of intergenic transcripts. *Mol Cell*. 2006; 22:309–16. [PubMed: 16678103]
88. Lee JT. Epigenetic regulation by long noncoding RNAs. *Science*. 2012; 338:1435–9. [PubMed: 23239728]
89. Spizzo R, Almeida MI, Colombatti A, Calin GA. Long non-coding RNAs and cancer: a new frontier of translational research? *Oncogene*. 2012; 31:4577–87. [PubMed: 22266873]
90. Shore AN, Herschkowitz JI, Rosen JM. Noncoding RNAs involved in mammary gland development and tumorigenesis: there's a long way to go. *J Mammary Gland Biol Neoplasia*. 2012; 17:43–58. [PubMed: 22402938]
91. Yan B, Wang Z. Long noncoding RNA: its physiological and pathological roles. *DNA Cell Biol*. 2012; 31(Suppl 1):S34–41. [PubMed: 22612272]

Highlights

- Global nuclear run-on sequencing (GRO-seq) can be used to map transcriptomes
- GRO-seq tracks the position and orientation of actively transcribing RNA polymerases
- GRO-seq has been used to explore estrogen responses in breast cancer cells
- GRO-seq has identified novel, unannotated, estrogen-regulated non-coding RNAs
- These studies have elucidated the molecular logic of mitogenic responses

Figure 1. Assembly and function of enhancer complexes at estrogen receptor binding sites
 Estrogen receptors (ERs) can bind to genomic DNA directly through estrogen response elements (EREs; green boxes) (*top left*) or indirectly by tethering through other transcription factors (e.g., AP-1, SP1, NF- κ B) (*bottom left*). The binding of ER may be facilitated by the prior binding of pioneer factors, such as FOXA1 and AP2 (*top left*). Upon binding to chromatin in response to estrogen, ER recruits co-regulators (e.g., Mediator, SRC, p300/CBP) to enhancers to promote transcription of target gene expression through short-range interactions (promoter proximal enhancers; – looping) or long-range interactions (promoter distal enhancers; + DNA looping) with target gene promoters (*right*). The outcome is the regulation of RNA polymerase II (Pol II) recruitment or activity, leading to increased target gene expression.

Figure 2. Protocol for GRO-seq

Global run-on coupled with massively parallel sequencing (GRO-seq) is a direct, high throughput sequencing method used to obtain a map of the position and orientation of all transcriptionally engaged RNA polymerases (Pols I, II, and III) across the genome [51]. **(A)** RNAs are transcribed throughout the genome by actively engaged RNA polymerases (Pol) in cells. **(B)** In order to map the location and density of RNA polymerases using GRO-seq, nuclei isolated from cells are subjected to a short run-on reaction in the presence of an NTP analog (Bromo-UTP) to label nascent RNAs. The reaction is performed in the presence of Sarkosyl, which blocks unengaged RNA polymerases from initiating transcription and removes impediments to elongating RNA polymerases during the run-on. Only newly synthesized portions of the RNA molecules are labeled with Bromo-UTP. **(C)** RNAs are isolated from the run-on reactions and base hydrolyzed to generate short fragments for short-read sequencing. **(D)** Bromo-UTP incorporated, newly synthesized RNAs are enriched by

Bromo-UTP specific antibody conjugated-beads. Isolated nascent RNAs are ligated with RNA-adaptors, reverse-transcribed, and amplified for high throughput sequencing.

Figure 3. Genome browser views of GRO-seq data

Genome browser views of GRO-seq reads mapped to the human reference genome. The samples were prepared from ER⁺-positive MCF-7 human breast cancer cells treated with or without estradiol (E2) for 40 min [39,52]. Red indicates plus (+) strand reads and blue indicates minus (-) strand reads. Gene annotations with introns and exons are shown as schematics below the browser tracks. The transcription start sites (TSSs; *bent arrows*) and direction of transcription (*repeated arrowheads*) are shown. **(A)** Transcription of a protein coding gene. Upon stimulation with E2, transcription of *LHX4* is induced, whereas transcription of *ACBD6* is unregulated. The location of divergent promoter transcription is shown. **(B)** Transcription of the distal enhancer of the *SMAD7* gene. Upon stimulation with E2, transcription of the *SMAD7* distal ER⁺ enhancer is up-regulated, generating enhancer RNAs (eRNAs). Transcription of the distal enhancer corresponds to transcription of the *SMAD7* gene.

Figure 4. GRO-seq detects different types of Pol II-transcribed RNAs

A two state hidden Markov model (HMM)-based algorithm can be used with GRO-seq data to map transcription units that produce different types of RNAs. Each schematic illustrates a different type of transcription unit being transcribed by RNA polymerase II (Pol II). The transcription start sites (TSS; *bent arrows*), plus and minus DNA strands, actively transcribed RNAs, and orientation of the RNA polymerases are shown. Messenger RNAs (mRNAs) are 5' capped, spliced, and polyadenylated protein coding RNAs, which may contain microRNAs embedded in their introns. MicroRNAs may also be transcribed as primary transcripts from an intergenic region using their own promoter. Enhancer RNAs (eRNAs) are short, typically divergently/bidirectionally transcribed RNAs that originate from enhancers, including ER binding sites. Long non-coding RNAs (lncRNAs) are mRNA-like transcripts that are longer than 200 nucleotides, but lack coding potential. They are typically transcribed from intergenic regions, but their transcription units may overlap protein-coding genes. Antisense RNAs (asRNAs) originate from the opposite strand of an annotated protein coding or lncRNA gene with significant overlap of the transcribed regions. Lastly, divergent RNAs are short non-coding RNAs that are transcribed from the opposite strand of a RefSeq gene promoter, other Pol II promoter, or enhancer. Non-coding RNAs transcribed by RNA Pol I and III are not shown, but they are detected by GRO-seq.

Figure 5. Dynamics and biology of an estrogen-dependent transcriptional response

GRO-seq analysis of an estrogen-dependent transcriptional response in ER⁺-positive MCF-7 human breast cancer cells subjected to a time course of estrogen treatment. **(A)** Estrogen-dependent transcriptional responses for protein coding genes can be classified into four distinct classes based on the dynamics of the response over a time course of estradiol (E2) treatment. The graphs are based on data from Hah et al. (2011) [52], but are stylized and are not factually quantitative. GRO-seq detects instantaneous changes in transcription that reflect the rapid actions of estrogen signaling on the transcriptome. Steady-state mRNA levels parallel the transcriptional response, but are clearly temporally delayed. **(B)** Gene ontology analyses of each class of genes from (A) reflect distinct biological effects the estrogen-dependent transcriptional response [52].

Figure 6. Determination of Pol II transcription rates and dynamics using GRO-seq

GRO-seq allows the monitoring of instantaneous changes in RNA polymerase location, orientation, and density, as illustrated in these stylized browser tracks based on actual data. When applied to a time course of induction (e.g., up-regulated gene, A; *left*) or inhibition (e.g., down-regulated gene, C; *right*), GRO-seq can be used to determine the leading edge (*left*) or lagging edge (*right*) of the “wave” of RNA polymerase transcribing along the body of a gene. Using the edge of the RNA polymerase wave, the distance traveled over time can be measured, allowing calculation of the rate of transcription, as described [39].

Figure 7. The logic of a mitogenic transcriptional response

Estrogen signaling has immediate and lasting effects on the transcriptome in breast cancer cells, as indicated. The immediate responses prepare the cell for secondary signaling and transcriptional responses. The later responses support ribosome biogenesis, protein synthesis, and protein processing, preparing newly transcribed mRNAs for translation.