

# Bioinformatic Evaluation of Features on Cis-regulatory Elements at 6q25.1

N Sreekar and Smeeta Shrestha

Department of Genetics, School of Basic and Applied Sciences, Dayananda Sagar University, Bengaluru, India.

Bioinformatics and Biology Insights  
Volume 17: 1–10  
© The Author(s) 2023  
Article reuse guidelines:  
sagepub.com/journals-permissions  
DOI: 10.1177/11779322231167971



**ABSTRACT:** Eukaryotic non-coding regulatory features contribute significantly to cellular plasticity which on aberration leads to cellular malignancy. Enhancers are cis-regulatory elements that contribute to the development of resistance to endocrine therapy in estrogen receptor (ER)-positive breast cancer leading to poor clinical outcome. ER is vital for therapeutic targets in ER-positive breast cancer. Here, we review and report the different regulatory features present on ER with the objective to delineate potential mechanisms which may contribute to development of resistance. The UCSC Genome Browser, data mining, and bioinformatics tools were used to review enhancers, transcription factors (TFs), histone marks, long non-coding RNAs (lncRNAs), and variants residing in the non-coding region of the ER gene. We report 7 enhancers, 3 of which were rich in TF-binding sites and histone marks in a cell line-specific manner. Furthermore, some enhancers contain estrogen resistance variants and sites for lncRNA. Our review speculates putative models suggesting potential aberrations in gene regulation and expression if these regulatory landscapes and assemblies are altered. This review gives an interesting perspective in designing integrated in vitro studies including non-coding elements to study development of endocrine resistance in ER-positive breast cancer.

**KEYWORDS:** Estrogen receptor, non-coding DNA, gene regulation, enhancer, estrogen resistance, transcription factors, RNA, chromatin state

**RECEIVED:** December 18, 2022. **ACCEPTED:** March 17, 2023.

**TYPE:** Review

**FUNDING:** The author(s) received no financial support for the research, authorship, and/or publication of this article.

**DECLARATION OF CONFLICTING INTERESTS:** The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

**CORRESPONDING AUTHOR:** Smeeta Shrestha, Lee Kong Chian School of Medicine, Nanyang Technological University (NTU), 636921, Singapore. Email: smeetashrestha@gmail.com

## Introduction

Promoters, enhancers, and insulators are some regulatory features located on eukaryotic non-coding genomic regions that contribute to the spatiotemporal development of cells.<sup>1</sup> One such important feature is the enhancer element which when bound by specific proteins expedite gene expression regardless of their distance, orientation, and position from the associated target gene.<sup>2,3</sup> Interaction of enhancer with regulatory proteins form a complex gene regulatory network that determines cell identity during development.<sup>4,5</sup> Chromatin immunoprecipitation, microarrays, and high-throughput sequencing have helped to elucidate chromatin architecture and associated enhancer function.<sup>6</sup>

Chromatin signature of enhancers helps to understand basic characterization of enhancer organization and cell-type specific activities.<sup>7</sup> Methylation marks on histones near enhancers correlate to its function in different cell types.<sup>8,9</sup> Genes that are actively transcribed show histone H3 monomethylated at Lys 1 (H3K4me1) and acetylated at Lys 27 (H3K27ac),<sup>10-12</sup> but developmental genes that are poised for future activation contain repressive trimethylated at Lys 27 (H3K27me3) near enhancers.<sup>10</sup> DNA methylation is important in determining enhancer landscape and eviction of transcription factors (TFs) contributing to cell-type specific enhancer activity<sup>13</sup> and gene expression.<sup>14</sup> The TF identity, binding site, and number reflect the complexity of the enhancer regulatory network.<sup>15</sup> Active enhancers also transcribe a class of non-coding RNA, enhancer RNA (eRNA) which assist in enhancer-promoter looping, chromatin modification, and regulation of gene expression.

These eRNAs exhibit tissue and lineage specificity and are markers of cell state and function.<sup>2</sup> In some cases, an aggregation of multiple enhancers, termed as super enhancer, can contribute to gene expression patterns leading to a robust phenotype.<sup>16,17</sup> Super enhancer<sup>18</sup> is located near genes that have high H3K4Me1, H3K27Ac marks, and p300 protein occupancy.<sup>19</sup> They are known to produce higher levels of eRNA which is responsible for cell identity, cell fate, transcription in cancer cells, and tumor heterogeneity.<sup>2</sup> The presence of disease associated non-coding variants on enhancers is estimated to explain a large proportion of heritability in comparison to coding variants.<sup>20</sup> Mutations in enhancer lead to loss of potential motif for TF-DNA interaction and aberrations in epigenetic modifications leading to enhancer-associated gene dysregulation.<sup>17,19</sup>

Estrogen receptor (ER) signaling is key driver of ER+ breast carcinogenesis and inhibition of ER signaling is the main focus of ER+ breast cancer therapy.<sup>21</sup> Intricate mutations in the ER pathway lead to development of endocrine resistance.<sup>21</sup> Those changes may be associated with ER $\alpha$  modifications either at the genetic, regulatory, or protein level. Evidence shows interaction of genomic features at pre- or post-transcription to contribute to the development of resistance. Specially enhancer reprogramming driven by assembly of TFs contribute to endocrine resistance.<sup>22-24</sup> Studies have shown estrogen mediated epigenetic remodeling of enhancers at ER $\alpha$  binding in endocrine resistance breast cancer.<sup>25</sup> Also, reprogramming of enhancer by TF assemblies supports endocrine-resistant breast cancer<sup>26</sup> and breast cancer-specific epigenomic



landscape contribute to dysregulated oncogene expression in breast tumorigenesis.<sup>27</sup> These studies report role of regulatory features contributing to development of endocrine resistance. However, there is limited information on the dynamics of DNA regulatory elements on estrogen receptor 1 (ESR1), estrogen receptor estrogen response element (ER-ERE) signaling, and how variation in enhancer landscape connects to chromatin reconfiguration leading to altered gene expression. Organization of the gene regulation is multi-tiered which contribute as target sites to address ER modulation.<sup>28</sup> This study uses ENCODE data base and data mining to review genomic regulatory features present on *ER1* gene to illustrate the non-coding regulatory features, enhancers, variants, TFs, and methylation signatures on ER which can contribute to its modulation. In addition, it provides insights to cohesively integrate this information to understand development of endocrine resistance.

## Methods

### *Regulatory elements present on Estrogen receptor 1 and its features*

The University of California Santa Cruz (UCSC) Human Genome Browsers of builds GRCh37/hg19 at chr6:151,921,777-152,102,547 and GRCh38/hg38 at chr6:151,599,479-151,779,568 was evaluated. The Encyclopedia of DNA Elements (ENCODE) Project<sup>29,30</sup> tracks were used to identify the gene regulatory elements on the *ESR1* gene. The UCSC GeneHancer<sup>31</sup> track features were overlaid on the above coordinates to identify DNA regulatory elements corresponding to *ESR1* gene. Further DNA regulatory tracks such as histone marks (H3K4Me1, H3K4Me3, H3K9Me3, H3K27Ac), chromatin state segmentation, and transcription factor-binding sites (TFBSs) were mapped on the GeneHancer identified regulatory sequence on ESR1. Tracks specific for cell lines were used to integrate and compare the chromatin states and TF binding at the regulatory region. The human genome build, GRCh38/hg38 was used to evaluate potential candidate cis-regulatory elements (cCREs) at ESR1. Flowchart illustrating the various databases used to annotate features on ESR1 in Supplementary Figure 1.

### *Single nucleotide polymorphisms on Estrogen receptor 1 regulatory elements*

The single nucleotide polymorphisms (SNPs) located on the non-coding DNA regulatory elements on ESR1 was evaluated using literature search, Human Gene Mutation Database (HGMD),<sup>32</sup> and 1000 Genome Phase 1 and Phase 3 available from the UCSC hg19 genome build. SNPs present across the 7 enhancers were tabulated, evaluated, and characterized based on their association with estrogen resistance, clinical relevance, and allele frequency. Regulome Database<sup>33</sup> was used to investigate the chromatin accessibility of these SNPs.

### *Enhancer RNA signatures on Estrogen receptor 1 regulatory element*

Enhancer RNA in Cancer (eRIC) database<sup>34</sup> was used to identify the eRNAs expressed in  $\leq 1$  mb of *ESR1* gene in *The Cancer Genome Atlas* (TCGA) Breast invasive carcinoma (BRCA) cell type. The eRIC database was used to confirm the eRNA target genes in BRCA cells.

## Result

### *Estrogen receptor 1 gene contains Enhancer regulatory elements*

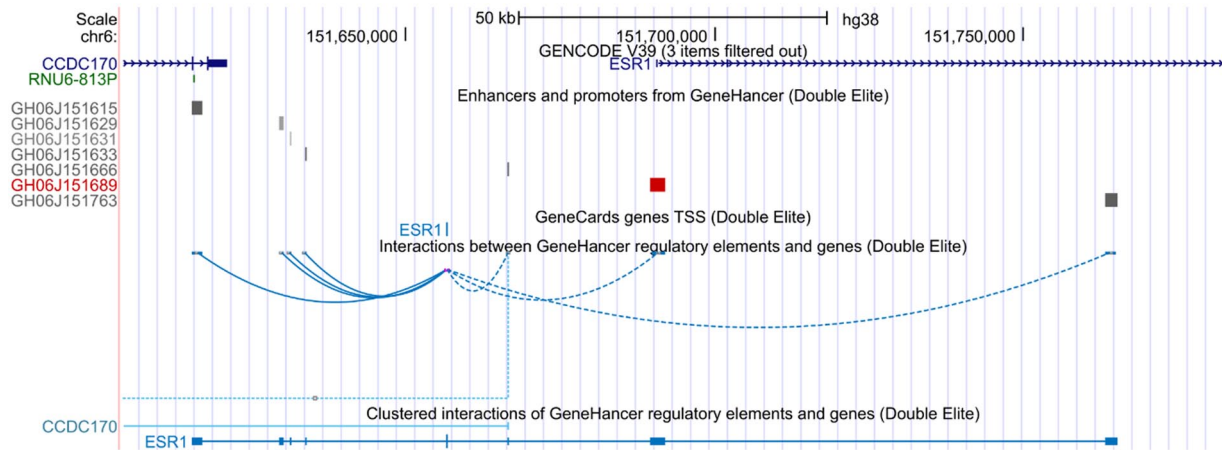
We report 7 regulatory marks present at chr 6:151,921,777-152,102,547 genomic region. Six were annotated to be double elite enhancers and 1 was a promoter. Elite enhancers are DNA regulatory elements that are evidenced from more than 1 source, whereas double elite enhancers are association of multiple elite enhancers (Figure 1). The first enhancer (GH06J151615) was located on *CCDC170* gene, 4 enhancers (GH06J151629, GH06J151631, GH06J151633, GH06J151666) were present at the intergenic region, and the last enhancer (GH06J151763) was present in the second intron of *ESR1* gene. Interaction of enhancers GH06J151615, GH06J151629, GH06J151631, and GH06J151633 was in the forward direction (solid line), whereas enhancers GH06J151666, GH06J151689, and GH06J151763 showed reverse directed interaction (dotted line) shown in Figure 1. ENCODE scored strengths of the genomic regulatory elements based on the literature, number of unique TFs, and overlaps with ultra-conserved non-coding genomic elements. GH06J151633 showed the highest score of 11.03, whereas GH06J151631 showed the lowest confidence score of 4.5 (Table 1).

### *Regulatory elements on Estrogen receptor 1 interact with different genes*

Six of the DNA regulatory elements on ESR1 interacted with both *CCDC170* and *ESR1* genes except GH06J151763, which interacted only with ESR1. In addition, enhancers GH06J151666, GH06J151689, and GH06J151763 interacted with Spectrin repeat containing nuclear envelope 1 (*SYNE1*), GH06J151689 with Zinc Finger and BTB domain containing 2 (*ZBTB2*) genes, and GH06J151615 with an lncRNA such as lnc-RMND1-2 (Table 1).

### *Histone marks on regulatory elements of Estrogen receptor 1*

Enhancers associated with histone marks, H3K4Me1, H3K4Me3, H3K9Me3, and H3K27ac help initiate active transcription.<sup>35</sup> Enhancers GH06J151615, GH06J151633, GH06J151631, and GH06J151763 reported histone H3K4Me1 peaks in human hepatic cancer cells (HepG2). Enhancers GH06J151631 and GH06J151763 showed active H3K4Me1



**Figure 1.** Double elite features and regulatory regions on *ESR1* gene loci 6q25.1 region of human genome using ENCODE on UCSC Browser. A total of 7 double elite regulatory elements were observed using GeneHancer database available in UCSC ENCODE Browser using Human Hg38 version near the vicinity of *ESR1* gene. GH06J151615 enhancer was seen in *CCDC170* gene, whereas GH06J151629, GH06J151631, GH06J151633, and GH06J151666 were observed in the intergenic region. GH06J151689 was the promoter and GH06J151763 was present in the second intron of the transcript variant 4 of the *ESR1* gene.

ENCODE indicates Encyclopedia of DNA Elements; UCSU, University of California Santa Cruz.

and H3K4Me3 marks in human mammary epithelial cells (HMECs). Moderate H3K4Me1 and H3K9Me3 peaks were observed in GH06J151615, GH06J151629, and GH06J151763 in human embryonic stem cell (H1-hESC) cells (Figure 2).

#### *Transcription factor-binding sites on Estrogen receptor 1 regulatory elements*

Enhancers consist of dense clusters of TFBSs which are cell-type specific and involve in modulation of gene expression.<sup>36</sup> Five of the *ESR1* regulatory elements (GH06J151615, GH06J151633, GH06J151666, GH06J151689, and GH06J151763) report cell-type TFBS (Figure 3). A total of 19 TFs show binding to enhancer GH06J151615 of which CEBPB and RXRA were reported in H1-hESC cells, whereas HNF4A and HNF4G were reported in hepatic cells. The GH06J151633 enhancer was contained binding site for 13 TFs. A TF SMC3, which is a DNA repair protein binds to GH06J151633 enhancer in H1-hESC cells, whereas FOXA1 and TEAD4 TFs are bound to it in HepG2 cells. The type of TF and enhancer TFBS is cell-type dependent. ZNF263, a zinc finger TF repressor, is bound to GH06J151666 enhancer in HMECs. The *ESR1* regulatory element GH06J151689 is a promoter, with 39 TFs. Interestingly, the IKZF1 a chromatin remodeling factor bound to GH06J151689 specifically in HI-hESC cells, but bound to KAP1, a nucleosome remodeling protein in HMEC cells. Finally, GH06J151763, an *ESR1* intronic enhancer, bounds to 23 TFs. It specifically bound to SP4 in H1-hES cells, but in HMEC cells, it bound to KAP1 and to SMC3, a DNA repair protein in HepG2 cells (Figure 3).

#### *Variants on Estrogen receptor 1 Enhancer elements*

A total of 4 databases—1000 genome Phase 1 and Phase 3,<sup>37</sup> ClinVar,<sup>38</sup> HGMD,<sup>32</sup> and literature search with terms, “Breast

cancer AND estrogen receptor AND Genetic”—were used to find and annotate the variants across the 7 enhancers on *ESR1* (Supplementary Table 1). Overall, 245 variants were identified of which 12 were responsible for estrogen resistance and sum of 179 variants were located across the 7 regulatory elements of *ESR1* gene. However, only rs6929137 and rs9397435 were present on GH06J151615 and GH06J151629 enhancers and were linked to estrogen resistance (Table 2). The potential function of rs6929137, rs9397435, and rs140068132 located on enhancers were analyzed using Regulome DB and SNP functionality scored based on its chromatin state in H1-hESC, HMEC, and HepG2 cell lines (Supplementary Table 2). The rs6929137 present on GH06J151615 showed strong transcription activity in H1-hESC cells, low activity in mammary epithelium cells, and enhancer activity in HepG2 cells. The SNP, r9397435 on GH06J151629 reported enhancer activity in H1-hESC and HMEC cells but showed strong transcription activity in HepG2 cells. Similiary, rs140068132 present on GH06J151633 reported low activity in H1-hESC cells, repressed state in HMEC cells and enhancer activity in HepG2 cell (Supplementary Table 2). Distinct variants across different enhancer show different regulatory potential across different cell lines.

#### *Candidate cis-regulatory elements on estrogen receptor 1 enhancer elements*

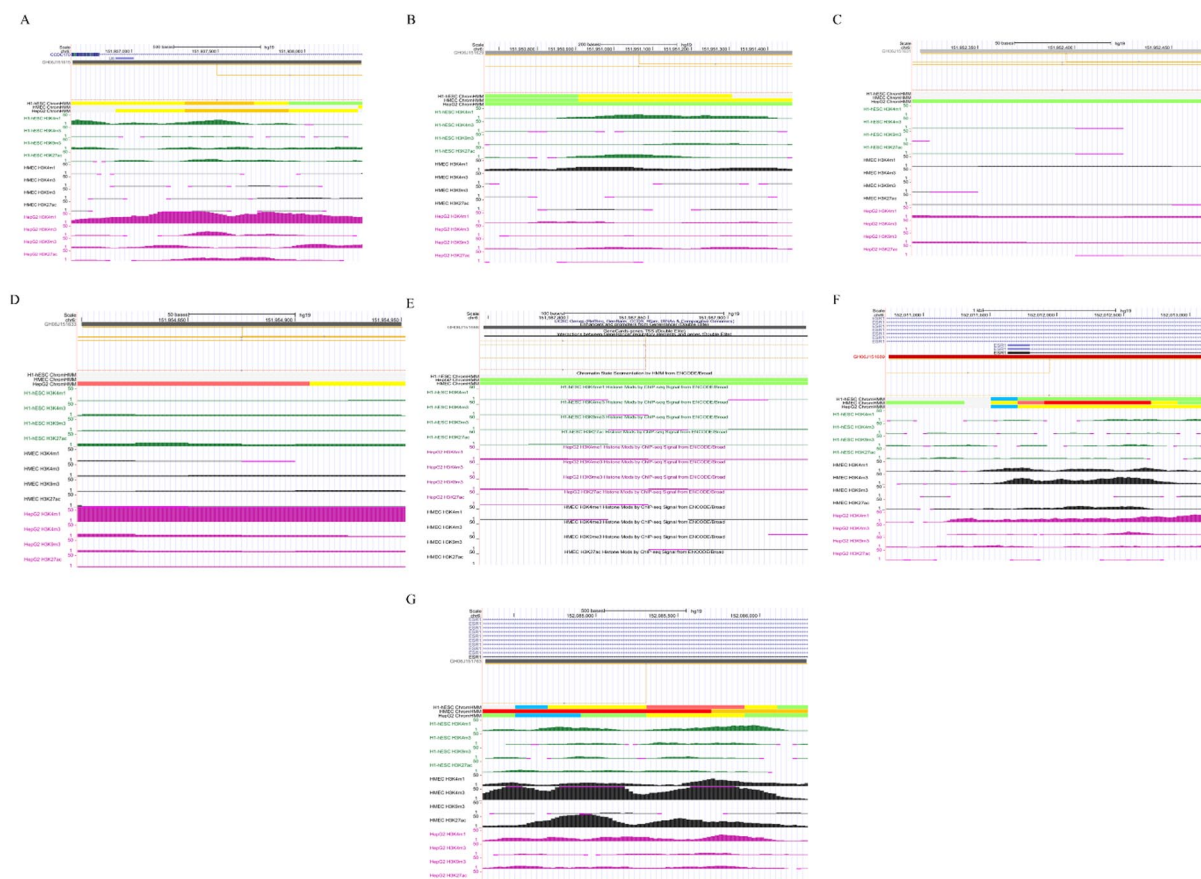
The cCREs are sequences on non-coding DNA that regulate transcription of neighboring genes. We identified 156 potential cCREs on *ESR1* at chr6:151616119-152130808 (hg38) (Supplementary Table 3). Majority (88/156) of cCREs were classified as distal enhancer-like sites (dELS). A total of 6 of 14 SNPs present on cCREs had clinical significance. The variants rs9397435, rs121913044, and rs397509428 were linked to estrogen resistance and rs121913043 produced mutant ER

**Table 1.** Elite enhancer coordinates, enhancer interactions, GWAS catalog, transcription factor, total number of SNPs.

ENHANCER NAME	NUMBER OF BPS	LOCATION—HG19	ELITE CONFIDENCE VALUE	INTERACTIONS	GWAS CATALOG	TOTAL NUMBER OF SNPS	TFBS
GH06J151615	1666	chr6:151936661-151938326	9.06	HSALNG0054371-001, HSALNG0054371-002, LOC107986528, RNU6-813P, RF00026-946, ESR1, CCDC170	Bone density	23	PoIR2A, GATA2, CEBPB, GATA3, ZNF217, E2F1, TCF7L2, PoIR2A, EP300, MYC, FOXA1, NFIC, ESR1, HNF4A, SP1, HNF4G, RXRA, PoIR2A, MYC.
GH06J151629	800	chr6:151950665-151951464	6.49	ESR1, HSALNG0054372, Inc-RMND1-2, CCDC170	NA	32	NA
GH06J151631	150	chr6:151952321-151952470	4.5	ESR1, HSALNG0054372, Inc-RMND1-2, CCDC170	Estrogen-receptor negative breast cancer, heel bone mineral density, breast carcinoma, cancer, breast size	6	NA
GH06J151633	150	chr6:151954801-151954950	11.03	ESR1, HSALNG0054372, Inc-RMND1-2, CCDC170	Breast carcinoma	6	TEAD4, EP300, GATA3, FOXA1, E2F1, PoIR2A, ESR1, MYC, FOXA2, CTCF, RAD21, ZNF217, SMC3.
GH06J151666	202	chr6:151987748-151987949	9.31	CCDC170, HSPA8P15, ESR1, HSALNG0054375, piR-43325-084, Inc-ESR1-3, SYNE1	NA	3	ZNF263.
GH06J151689	2412	chr6:152010741-152013152	1485.08	ESR1, ENSG00000233823, ZBTB2, CCDC170, HSALNG0054376, Inc-ESR1-3, HSALNG0054377, SYNE1	Bone fracture, hip bone mineral density	59	E2F1, FOXA1, ESR1, KAP1, CTCF, CTCF, RAD21, SMC3, PoIR2A, IKZF1, PoIR2A, MTA3, FOXM1, ATF2, RELA, RAD21, EP300, CHD2, SP1, TCF3, POU2F2, YY1, STAT3, TBL1XR1, SP1, MXI1, BHLHE40, BATE.
GH06J151763	1962	chr6:152084325-152086286	5.45	Inc-CCDC170-2, ENSG00000233823, ESR1, LOC107986529, HSALNG0054378, Inc-ESR1-3, SYNE1	Heel bone mineral density	49	CTCF, SMC3, STAT3, PoIR2A, MYC, SP4, NFYB, SP2, CHD2, FOXA1, PBX3, SP1, Fos, CTCF, GATA2, YY1, PoIR2A, ZNF143, MYC, RAD21, SMC3, KAP1, CTCFL.

Abbreviations: GWAS, Genome-wide association studies; NA, not available; SNP, single nucleotide polymorphism; TFBS, transcription factor-binding site; UCSC, University of California Santa Cruz.

Table provides a cumulative overview of the identified enhancers and its associated factors such as Transcription factors, GWAS catalog, and number of SNPs found in the enhancer. A total of 7 elite enhancers were reported using hg19 version of GeneHancer available in UCSC Genome Browser. The lowest confidence was seen for enhancer GH06J151631, and the highest confidence score was for the enhancer GH06J151689. The enhancers showed interactions with various non-coding RNAs, CCDC170 and ESR1. Also, SNPs rs6929137 and rs9397435 were found in enhancers—GH06J151615 and GH06J151629, respectively.



**Figure 2.** Chromatin and histone state across 7 enhancers. Each cell line illustrates peaks for histone marks, H3K4Me1, H3K4Me3, H3K9Me3, and H3K27ac. (A) GH06J151615, (B) GH06J151629, (C) GH06J151631, (D) GH06J151633, (E) GH06J151666, (F) GH06J151689, and (G) GH06J151763. The first panel of tracks shows the horizontal bar of chromatin states retrieved from the ChromHMM database across the 3 cell lines. The second panel of vertical peaks colored in H1-hESC (green), HMEC (black), and HepG2 (pink). H1-hESC indicates human embryonic stem cell; HepG2, human hepatic cancer cell; HMEC, H3K4Me3 marks in mammary epithelial cell.

alpha protein.<sup>39,40</sup> The rs121913044 was located on DNase and H3K4Me3 signatures, whereas rs397509428 resided on CTCF and H3K4Me3 sites (Table 2) (Supplementary Figure 2.)

#### Enhancer RNA on estrogen receptor 1 regulatory element

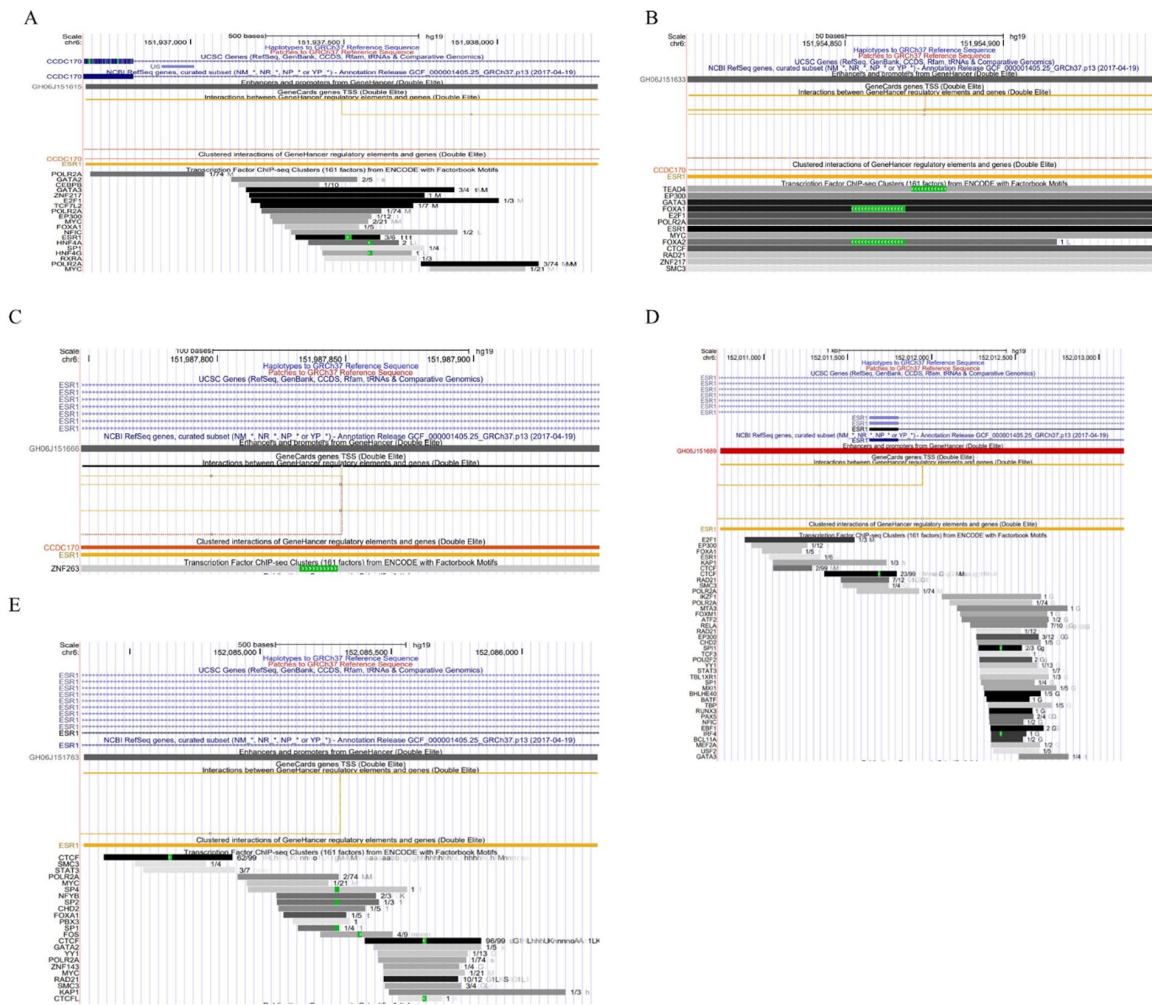
Enhancer RNAs are lncRNA transcribed from DNA sequence of regulatory regions. The eRIC database reported 7 eRNA regulating ESR1 from the TCGA BRCA cell-type expression database. The eRNA ENSR00000205343 was reported to be transcribed from the position, chr6: 151763273-151765000 containing GH06J151763 and cCREs with the mean expression of 3.14 units (Figure 4A). The ENSR00000205343, ENSR00000205344, and ENSR00000323738 were seen to be transcribed from cCREs containing dELS and CTCF bound signatory marks. All the 7 eRNA identified in BRCA cell type was known to act on ESR1 and its neighboring genes such as Required for Meiotic Nuclear Division 1 (*RMND1*) and *ZBTB2* (Figure 4B).

#### Discussion

The non-coding DNA vastly outnumber coding DNA supporting their role as designers of eukaryotic complexity.

Enhancers are DNA regulatory elements which regulate transcription of one or more genes regardless of its orientation or relative distance from the target promoter.<sup>41</sup> They contribute significantly to cell growth, differentiation, and disease. Estrogen plays an important role in breast cancer development and endocrine therapy prevents the effect of estrogen on breast cancer cells. Endocrine therapy significantly reduced relapse rate by half in breast cancer patients; however, 1 of 3 patients still develop hormone resistance and relapse.<sup>42</sup> Hormone resistance in breast cancer is an acquired phenotype in which the cells stop responding to hormone treatment. The molecular mechanism driving this resistance in ER-positive breast cancer patients is poorly understood. Recent studies report global reprogramming of estrogen responsive regions in the DNA could contribute to this insensitivity. Here, we use the UCSC Genome Browser, ENCODE, and associated databases to review and report DNA regulatory elements on ESR1 genomic region to postulate the role of genomic regulatory marks contributing to hormone resistance.

The genetic regulatory elements that define endocrine response in ESR1-positive breast cancer patients has remained poorly understood. The spatial organization of the genome inside the cell nucleus is pivotal to cell function.<sup>43</sup>



**Figure 3.** Transcription factor binding to the ESR1 DNA regulatory elements. Using UCSC we identified several tissue specific and common transcription factors binding to ESR1 enhancers. 19 TFs were bound onto GH06J151615 (A), 13 TFs on GH06J151633 (B), 39 TFs on the promoter (D), and 23 TFs on the intronic enhancer, GH06J151763 (E). Single as seen binding to GH06J151666 enhancer (C). Enhancers GH06J151629 and GH06J151631 did not harbor any transcription factor. DNA indicates deoxyribonucleic acid; TF, transcription factor; UCSC, University of California Santa Cruz.

Non-coding genomic regions harbor specific nucleotide sequence favoring certain interaction and bindings in its 3-dimensional (3D) conformations that may facilitate certain repertoire of gene expression leading to development of resistance.

Using data mining approaches, we review and report 7 double elite regulatory elements spanning across *ESR1* and *CCDC170* genes at intergenic, promoter, exonic, and non-coding intronic regions. These enhancers formed clusters and enact as a super enhancer<sup>44</sup> (Supplementary Table 4). Furthermore, they are sites for lncRNA binding and facilitate gene expression.<sup>45</sup> The UCSC database tracks, ChromHMM, histone marks, TFs, SNPs, cCREs, and eRNA illustrate the genomic complexity of ESR1 regulation.

The 7 ESR1 regulatory sequences hosted histone marks and transcriptional factors which varied across cell type example some enhancers active in normal mammary epithelial cells but not in hepatoma and vice versa. Similarly, enhancers at

intergenic and on introns of ESR1 contained active histone marks, whereas those on exon showed no activity.

There was no reported TF binding on 2/7 enhancers. Intergenic variants rs6929137 and rs9397435 are responsible for estrogen resistance. The rs140068132 variant found at TF binding site was present on GH06J151633 enhancer and is known to be associated with substantially lower risk of breast cancer, particularly ER-negative breast cancer.<sup>46,47</sup> ESR1 enhancers contained cCREs at sites of transcription initiation in breast invasive carcinoma (BRCA) cell type. The variants rs121913044 and rs397509428 which were associated with estrogen resistance were located at CREs harboring DNase, histone, and CTCF marks. We speculate that the development of estrogen resistance could be due to loss/alteration of above chromatin signatures.

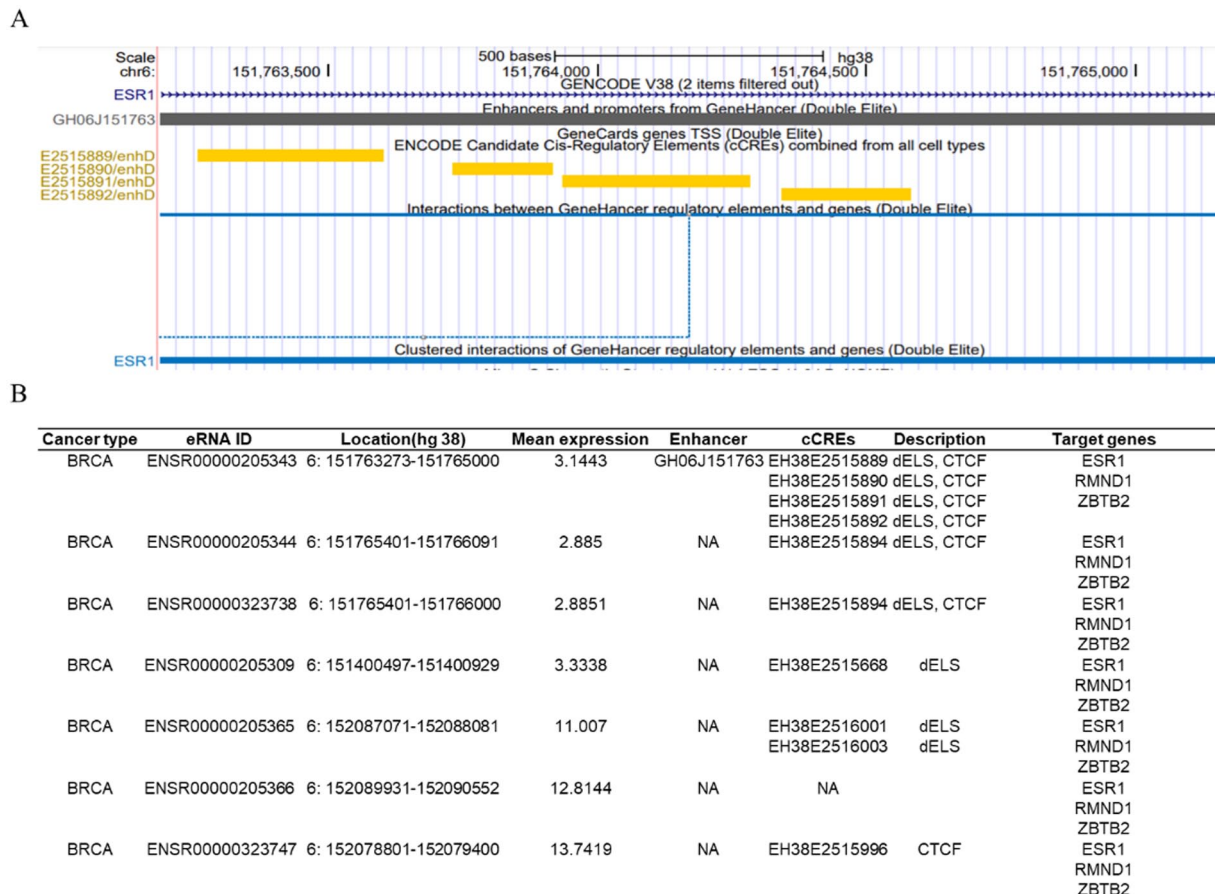
The human phenotype, normal or disease is a complex interaction of genes and regulatory elements in a certain genomic conformation. Most genes in the human genome

**Table 2.** Curative table of SNPs responsible for estrogen resistance along population and enhancer status.

SNP ID	ANCESTRAL ALLELE	VARIANT ALLELE	MAF	ALLELE WITH MAF	EUR	SAS	ENHANCER	CCRES	CLASSIFICATION
rs6929137	G	A	0.35	A	G=70% A=30%	G=70% A=30%	GH06J151615	NA	NA
rs9397435	A	G	0.12	G	A=93% G=7%	A=94% G=6%	GH06J151629	EH38E2515791	dELS
rs3734805	A	A	0.11	C	A=92% C=8%	A=95% C=5%	NA	NA	NA
rs12665607	T	A	0.1	A	T=92% A=8%	T=94% A=6%	NA	NA	NA
rs2046210	A	T	0.41	A	G=68% A=32%	G=68% A=32%	NA	NA	NA
rs12662670	T	C, G	0.11	G	T=92% A=8%	T=92% A=8%	NA	NA	NA
rs9383589	A	G	0.1	G	A=92% G=8%	A=92% G=8%	NA	NA	NA
rs104893956	C	T	NA	NA	NA	NA	NA	NA	NA
rs121913044	T	A	NA	NA	NA	NA	NA	EH38E2515955	DNase, H3K4Me3
rs397509428	G	T	NA	NA	NA	NA	NA	EH38E2515973	DNase, H3K4Me3, CTCF
rs1131692059	G	A	NA	NA	NA	NA	NA	EH38E2515973	DNase, H3K4Me3, CTCF
rs1562787582	T	A	NA	NA	NA	NA	NA	EH38E2516014	dELS
rs140068132	A	G	0.02	G	A=100%	A=100%	GH06J151633	EH38E2515796	dELS
rs2747652	C	T	0.46	T	T=45% C=55%	T=48% C=52%	NA	NA	NA

Abbreviations: MAF, minor allele frequency; EUR, Europe; SAS, South Asian; cCRE, candidate cis-regulatory element; NA, not applicable; SNP, single nucleotide polymorphism; TA, transcription factors.

Table also provides the overview of SNP bound with cCRE and its signature. Comprehensive list of SNPs responsible for estrogen resistance along with MAF and population status in European and South Asian population. We also looked if the SNP was bound to any of cCREs present in the region and noted its signature. SNP rs9397435 was present on GH06J151629 enhancer also contained dELS signature. SNP rs397509428 and rs1131692059 accompanied on DNase, H3K4Me3, and CTCF signatory marks on chromatin, whereas rs121913044 was present on DNase and H3K4Me3 site. SNP rs1562787582 was present on dELS region. Here, we also added 2 interesting SNPs rs140068132 which altered TF binding and rs2747652 responsible for estrogen negative breast cancer.



**Figure 4.** cCRE on enhancer GH06J151763. (A) Using UCSC, we identified 4 cCREs found on the enhancer GH06J151763. All the 4 cCREs had the distal enhancer-like sites and CTCF-binding signatures. (B) Seven eRNAs were identified using eRIC database. All the 7 eRNAs were found to act on *ESR1*, *RMND1*, and *ZBTB2* genes. Interestingly, eRNA ENSR00000205343 was transcribed by GH06J151763 enhancer. We also looked if any cCREs were present on the eRNA transcribed region and found that except ENSR00000205366 eRNA all the other eRNAs contained cCREs in them. eRNA ENSR00000323747 showed the highest mean expression of 13.74, and ENSR00000205344 and ENSR00000323738 showed the lowest mean expression of 2.885 in BRCA cell lines retrieved from TCGA database which was available in eRIC database.

BRCA indicates breast invasive carcinoma; cCRE, candidate cis-regulatory element; eRIC, enhancer RNA in cancer; eRNA, enhancer RNA; ESR1, estrogen receptor 1; RMND1, Required for Meiotic Nuclear Division 1; TCGA, The Cancer Genome Atlas; UCSC, University of California Santa Cruz; ZBTB2, Zinc Finger and BTB domain containing 2.

are likely regulated by multiple enhancers. Transcriptional enhancers are required for the differential expression of the human ER in breast cancers.<sup>48</sup> Development of estrogen resistance could be contributed by features present within the ESR1 genomic landscape such as SNPs, CREs, and histone marks, which contribute to ESR1 regulation/expression. ER $\alpha$  is well-known TF involved in maintenance of female reproductive system, bone mass and plays an important role in protecting cardiovascular and central nervous systems. Genomic studies comparing normal, MCF-7, tamoxifen-resistant, and fulvestrant-resistant cells showed loss/alterd CTCF and ER $\alpha$ -binding sites with development of estrogen resistance.<sup>49</sup>

However, till date there are limited studies on DNA regulatory marks of ESR1 to understand its role in development of estrogen resistance. This study reviews the regulatory regions and its associated features speculating its potential role in ESR1 activity. Driver mutations directly impact the protein product; however, variations in the enhancers might affect the TF signature and gene expression profile.<sup>50</sup> We believe the

ESR1 DNA regulatory signatures and mutations could contribute to the development of estrogen resistance during long-term hormone therapy.

Using our observations, we put forth 3 assumptions to better understand the development of estrogen resistance attributed by non-coding DNA regulatory features. In the first assumption, the mutation in enhancer results in loss of TF binding consequently led to reduced transcription. In the second assumption, the loss of TF binding site results in a modified mRNA production, and in the third assumption, the entire enhancer landscape is perturbed due to loss of histone modification resulting in modified/alterd enhancer landscape leading to production of mutated mRNA, no mRNA, and/or loss of enhancer-promoter interaction. These assumptions mentioned could be the repercussion of long-term exposure to tamoxifen which altered cytosol metabolite meliu leading to change in chromatin conformation and enhancer landscape. At present, the study is limited only to in silico analysis and no in vitro studies have been conducted to validate these assumptions.



## Conclusions

We review the potential role of non-coding DNA regulatory elements present on ESR1 in regulating estrogen resistance in breast cancer. This study can be used for development of better diagnostic test in identification of estrogen resistance. We postulate the role of intergenic enhancer and its features in contributing to estrogen resistance. In this study, we speculate that change in enhancer landscape and spatio-temporal gene expression results in the development of estrogen resistance over the course of hormone therapy.

## Author Contribution

SN – Literature and database search; SS – designed study and write manuscript.

## Supplemental Material

Supplemental material for this article is available online.

## REFERENCES

- Jin F, Li Y, Ren B, Natarajan R. Enhancers: multi-dimensional signal integrators. *Transcription*. 2011;2:226-230. doi:10.4161/TRNS.2.5.17712
- Arnold PR, Wells AD, Li XC. Diversity and emerging roles of enhancer RNA in regulation of gene expression and cell fate. *Front Cell Dev Biol*. 2020;7:377. doi:10.3389/FCCELL.2019.00377/BIBTEX
- Banerji J, Rusconi S, Schaffner W. Expression of a  $\beta$ -globin gene is enhanced by remote SV40 DNA sequences. *Cell*. 1981;27:299-308. doi:10.1016/0092-8674(81)90413-X
- Ong CT, Corces VG. Enhancers: emerging roles in cell fate specification. *EMBO Rep*. 2012;13:423. doi:10.1038/EMBOR.2012.52
- Narlikar L, Ovcharenko I. Identifying regulatory elements in eukaryotic genomes. *Brief Funct Genomic Proteomic*. 2009;8:215-230. doi:10.1093/BFGP/ELP014
- Kidder BL, Hu G, Zhao K. ChIP-Seq: technical considerations for obtaining high-quality data. *Nat Immunol*. 2011;12:918-922. doi:10.1038/NI.2117
- Calo E, Wysocka J. Modification of enhancer chromatin: what, how and why? *Mol Cell*. 2013;49:825-837. doi:10.1016/J.MOLCEL.2013.01.038
- Heinz S, Romanoski CE, Benner C, Glass CK. The selection and function of cell type-specific enhancers. *Nat Rev Mol Cell Biol*. 2015;16:144-154. doi:10.1038/NRM3949
- Gates LA, Foulds CE, O'Malley BW. Histone marks in the "drivers seat": functional roles in steering the transcription cycle. *Trends Biochem Sci*. 2017;42:977. doi:10.1016/J.TIBS.2017.10.004
- Rada-Iglesias A, Bajpai R, Swigut T, Bruggmann SA, Flynn RA, Wysocka J. A unique chromatin signature uncovers early developmental enhancers in humans. *Nature*. 2011;470:279-285. doi:10.1038/NATURE09692
- Creyghton MP, Cheng AW, Welstead GG, et al. Histone H3K27ac separates active from poised enhancers and predicts developmental state. *Proc Natl Acad Sci U S A*. 2010;107:21931-21936. doi:10.1073/PNAS.1016071107/-/DCSUPPLEMENTAL/PNAS.201016071SI.PDF
- Zentner GE, Tesar PJ, Scacheri PC. Epigenetic signatures distinguish multiple classes of enhancers with distinct cellular functions. *Genome Res*. 2011;21:1273-1283. doi:10.1101/GR.122382.111
- Wiench M, John S, Baek S, et al. DNA methylation status predicts cell type-specific enhancer activity. *EMBOJ*. 2011;30:3028. doi:10.1038/EMBOJ.2011.210
- Cedar H, Bergman Y. Programming of DNA methylation patterns. *Annu Rev Biochem*. 2012;81:97-117. doi:10.1146/ANNUREV-BIOCHEM-052610-091920
- Li L, Wunderlich Z. An enhancer's length and composition are shaped by its regulatory task. *Front Genet*. 2017;8:63. doi:10.3389/FGENE.2017.00063/BIBTEX
- Song W, Sharan R, Ovcharenko I. The first enhancer in an enhancer chain safeguards subsequent enhancer-promoter contacts from a distance. *Genome Biol*. 2019;20:1-14. doi:10.1186/S13059-019-1808-Y/FIGURES/4
- Fuxman Bass JI, Sahni N, Shrestha S, et al. Human gene-centered transcription factor networks for enhancers and disease variants. *Cell*. 2015;161:661. doi:10.1016/J.CELL.2015.03.003
- Tang F, Yang Z, Tan Y, Li Y. Super-enhancer function and its application in cancer targeted therapy. *NPJ Precis Oncol*. 2020;4:2-7. doi:10.1038/s41698-020-0108-z
- Carullo NVN, Day JJ. Genomic enhancers in brain health and disease. *Genes (Basel)*. 2019;10:43. doi:10.3390/GENES10010043
- Corradin O, Scacheri PC. Enhancer variants: evaluating functions in common disease. *Genome Med*. 2014;6:85. doi:10.1186/S13073-014-0085-3
- Rasha F, Sharma M, Pruitt K. Mechanisms of endocrine therapy resistance in breast cancer. *Mol Cell Endocrinol*. 2021;532:111322. doi:10.1016/J.MCE.2021.111322
- Bi M, Zhang Z, Jiang YZ, et al. Enhancer reprogramming driven by high-order assemblies of transcription factors promotes phenotypic plasticity and breast cancer endocrine resistance. *Nat Cell Biol*. 2020;22:701-715. doi:10.1038/S41556-020-0514-Z
- Yun MR, Lim SM, Kim SK, et al. Enhancer remodeling and MicroRNA alterations are associated with acquired resistance to ALK inhibitors. *Cancer Res*. 2018;78:3350-3362. doi:10.1158/0008-5472.CAN-17-3146
- Hamdan FH, Johnsen SA. Perturbing enhancer activity in cancer therapy. *Cancers (Basel)*. 2019;11:634. doi:10.3390/CANCERS11050634
- Sklias A, Halaburkova A, Vanzan L, et al. Epigenetic remodelling of enhancers in response to estrogen deprivation and re-stimulation. *Nucleic Acids Res*. 2021;49:9738-9754. doi:10.1093/NAR/GKAB697
- Bi M, Zhang Z, Jiang YZ, et al. Enhancer reprogramming driven by high-order assemblies of transcription factors promotes phenotypic plasticity and breast cancer endocrine resistance. *Nat Cell Biol*. 2020;22:701-715. doi:10.1038/s41556-020-0514-z
- Huang H, Hu J, Maryam A, et al. Defining super-enhancer landscape in triple-negative breast cancer by multiomic profiling. *Nat Commun*. 2021;12:1-16. doi:10.1038/s41467-021-22445-0
- Deng S, Feng Y, Pauklin S. 3D chromatin architecture and transcription regulation in cancer. *J Hematol Oncol*. 2022;15:1-23. doi:10.1186/S13045-022-01271-X
- Lu B, Klingbeil O, Tarumoto Y, et al. A transcription factor addition in leukemia imposed by the MLL promoter sequence. *Cancer Cell*. 2018;34:970-981.e8. doi:10.1016/J.CCELL.2018.10.015
- Abascal F, Acosta R, Addleman NJ, et al. Expanded encyclopaedias of DNA elements in the human and mouse genomes. *Nature*. 2020;583:699-710. doi:10.1038/s41586-020-2493-4
- Fishilevich S, Nudel R, Rappaport N, et al. GeneHancer: genome-wide integration of enhancers and target genes in GeneCards. *Database J Biol Databases Curation*. 2017;2017:1-17. doi:10.1093/DATABASE/BAX028
- Stenson PD, Mort M, Ball EV, et al. The human gene mutation database: towards a comprehensive repository of inherited mutation data for medical research, genetic diagnosis and next-generation sequencing studies. *Hum Genet*. 2017;136(6):665. doi:10.1007/S00439-017-1779-6
- Boyle AP, Hong EL, Hariharan M, et al. Annotation of functional variation in personal genomes using RegulomeDB. *Genome Res*. 2012;22:1790-1797. doi:10.1101/GR.137323.112
- Zhang Z, Lee JH, Ruan H, et al. Transcriptional landscape and clinical utility of enhancer RNAs for eRNA-targeted therapy in cancer. *Nat Commun*. 2019;10:1-12. doi:10.1038/s41467-019-12543-5
- Coppola CJ, Ramaker RC, Mendenhall EM. Identification and function of enhancers in the human genome. *Hum Mol Genet*. 2016;25:R190-R197. doi:10.1093/HMG/DDW216
- Panigrahi A, O'Malley BW. Mechanisms of enhancer action: the known and the unknown. *Genome Biol*. 2021;22:1-30. doi:10.1186/S13059-021-02322-1
- Auton A, Abecasis GR, Altshuler DM, et al. A global reference for human genetic variation. *Nature*. 2015;526:68-74. doi:10.1038/nature15393
- Landrum MJ, Lee JM, Riley GR, et al. ClinVar: public archive of relationships among sequence variation and human phenotype. *Nucleic Acids Res*. 2014;42:D980. doi:10.1093/NAR/GKT1113
- Reese JC, Katzenellenbogens BS. Mutagenesis of cysteines in the hormone binding domain of the human estrogen receptor alterations in binding and transcriptional activation by covalently and reversibly attaching ligands. *J Biol Chem*. 1991;266:10880-10887. doi:10.1016/S0021-9258(18)99101-5
- Reese JC, Katzenellenbogens BS. Characterization of a temperature-sensitive mutation in the hormone binding domain of the human estrogen receptor studies in cell extracts and intact cells and their implications for hormone-dependent transcriptional activation. *J Biol Chem*. 1992;267:986-9873. doi:10.1016/S0021-9258(19)50174-0
- Benton ML, Talipini SC, Kostka D, Capra JA. Genome-wide enhancer annotations differ significantly in genomic distribution, evolution, and function. *BMC Genomics*. 2019;20:1-22. doi:10.1186/S12864-019-5779-X/FIGURES/7
- Abe O, Abe R, Enomoto K, et al. Relevance of breast cancer hormone receptors and other factors to the efficacy of adjuvant tamoxifen: Patient-level meta-analysis of randomised trials. *Lancet*. 2011;378:771-784. doi:10.1016/S0140-6736(11)60993-8/ATTACHMENT/F68DD577-772B-4631-BFEF-E259042F9B C5/MMC1.PDF
- Zhou T, Zhang R, Ma J. The 3D genome structure of single cells. *Annu Rev Biomed Data Sci*. 2021;4:21-41. doi:10.1146/annurev-biodatasci-020121-084709
- Jiang Y, Qian F, Bai X, et al. SEDb: a comprehensive human super-enhancer database. *Nucleic Acids Res*. 2019;47:D235-D243. doi:10.1093/NAR/GKY1025

45. Hou Y, Zhang R, Sun X. Enhancer lncrnas influence chromatin interactions in different ways. *Front Genet.* 2019;10:936. doi:10.3389/FGENE.2019.00936/BIBTEX
46. Fejerman L, Ahmadiyah N, Hu D, et al. Genome-wide association study of breast cancer in Latinas identifies novel protective variants on 6q25. *Nat Commun.* 2014;5. doi:10.1038/NCOMMS6260
47. Hoffman J, Fejerman L, Hu D, et al. Identification of novel common breast cancer risk variants at the 6q25 locus among Latinas. *Breast Cancer Res.* 2019;21. doi:10.1186/S13058-018-1085-9
48. Tang Z, Treilleux I, Brown M. A transcriptional enhancer required for the differential expression of the human estrogen receptor in breast cancers. *Mol Cell Biol.* 1997;17:1274-1280. doi:10.1128/MCB.17.3.1274
49. Achinger-Kawecka J, Valdes-Mora F, Luu PL, et al. Epigenetic reprogramming at estrogen-receptor binding sites alters 3D chromatin landscape in endocrine-resistant breast cancer. *Nat Commun.* 2020;11:1-17. doi:10.1038/s41467-019-14098-x
50. Walavalkar K, Notani D. Beyond the coding genome: non-coding mutations and cancer. *Front Biosci (Landmark Ed).* 2020;25:1828-1838. doi:10.2741/4879