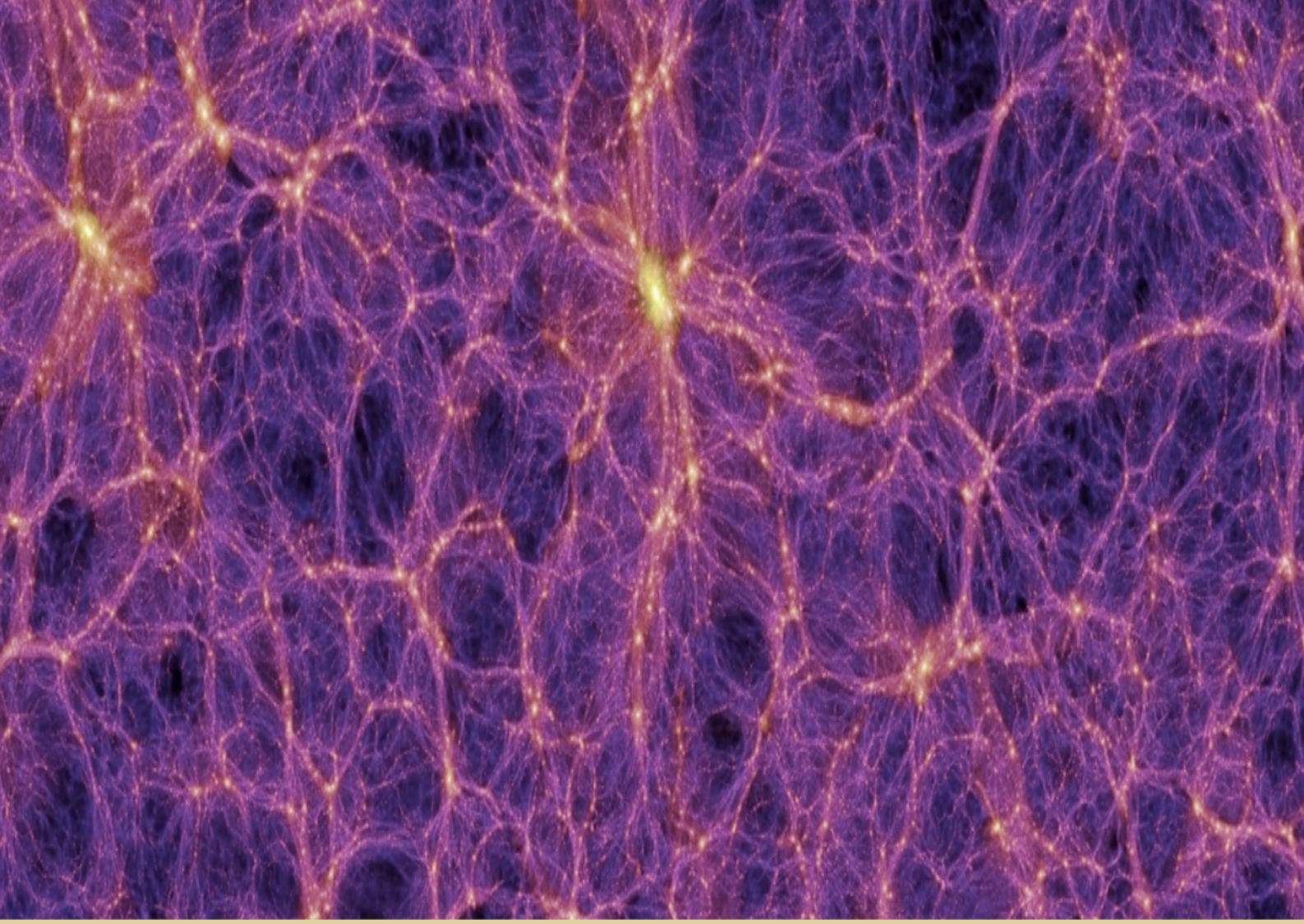


Carlos Manuel Almeida Guedes de Melo

Long Non-coding RNAs at the Service of p53

Tese de Doutoramento em Ciências e Tecnologias da Saúde (Pré-Bolonha), especialização em Biologia Celular e Molecular, orientada pelo Professor Doutor Reuven Agami do Netherlands Cancer Institute (Amsterdam, Holanda) e pela Professora Doutora Maria Celeste Lopes da Faculdade de Farmácia da Universidade de Coimbra e entregue à Faculdade de Farmácia da Universidade de Coimbra.

Julho 2014

· U C ·

Universidade de Coimbra

Image on the cover: 'The redshift zero distribution of dark matter' in Croton DJ et al., MNRAS (2006) 365 (1): 11-28

• U C •

FFUC FACULDADE DE FARMÁCIA
UNIVERSIDADE DE COIMBRA

Carlos Manuel Almeida Guedes de Melo

Long non-coding RNAs at the service of p53

Tese de Doutoramento em Ciências e Tecnologias da Saúde (Pré-Bolonha), especialização em Biologia Celular e Molecular, orientada pelo Professor Doutor Reuven Agami do Netherlands Cancer Institute (Amsterdam, Holanda) e pela Professora Doutora Maria Celeste Lopes da Faculdade de Farmácia da Universidade de Coimbra e entregue à Faculdade de Farmácia da Universidade de Coimbra.

Julho de 2014

· U C ·

UNIVERSIDADE DE COIMBRA

Tese de Doutoramento em Ciências e Tecnologias da Saúde (Pré-Bolonha), especialização em Biologia Celular e Molecular, orientada pelo Professor Doutor Reuven Agami do Netherlands Cancer Institute (Amsterdam, Holanda) e pela Professora Doutora Maria Celeste Lopes da Faculdade de Farmácia da Universidade de Coimbra e entregue à Faculdade de Farmácia da Universidade de Coimbra.

Title (English)	Long non-coding RNAs at the service of p53

Título (Português)	RNAs longos não codificantes ao serviço do p53

Author/Autor	Carlos Manuel Almeida Guedes de Melo
Supervision	Prof. Doctor Reuven Agami Netherlands Cancer Institute, Amsterdam, The Netherlands
Co-Supervision	Prof. Doutora Maria Celeste Fernandes Lopes Faculdade de Farmácia, Universidade de Coimbra, Portugal
Year/Ano	2014

The work presented in this thesis was developed during my PhD research project. The work was performed under the supervision of Doctor Reuven Agami, full professor, the Netherlands Cancer Institute, Amsterdam, and Erasmus Medical Center, Rotterdam, The Netherlands and co-supervised by Doctor Maria Celeste Lopes, full Professor, Faculty of Pharmacy, University of Coimbra, Portugal. This thesis has been supported by the grant SFRH/BD/33472/2008 from Fundação para a Ciência e a Tecnologia, Lisboa, Portugal and by ERC advance grant on enhancer RNA regulation.

The following original articles have been published or are under revision in peer-scientific journals, within the scope of the present PhD Thesis:

Melo CA*, Drost J*, Wijchers PJ, van de Werken H, de Wit E, Oude Vrielink JA, Elkon R, Melo SA, Léveillé N, Kalluri R, de Laat W, Agami R. (2013). eRNAs are required for p53-dependent enhancer activity and gene transcription. *Mol Cell*. 2013 Feb 7;49(3):524-35

Melo CA*, Léveillé N*, Agami R. (2013). eRNAs reach the heart of transcription. *Cell Res*. 2013 Oct;23(10):1151-2

Léveillé N*, **Melo CA***, Rooijers K*, Díaz-Lagares A. et al (2014). Genome wide profiling of p53-responsive enhancer RNAs uncovers an essential lncRNA-dependent epigenetic regulation. (under revision). * These authors contributed equally to this work.

Additional publications (collaboration work):

Léveillé N, Elkon R, Davalos V, Manoharan V, Hollingworth D, Oude Vrielink J, le Sage C, **Melo CA**, Horlings HM, Wesseling J, Ule J, Esteller M, Ramos A, Agami R. (2012). Selective inhibition of microRNA accessibility by RBM38 is required for p53 activity. *Nat Commun*. 2011 Oct 25;2:513-523

“There has to be a lightness; you have to be as light as you can be and not get weighed down and stuck in your emotion, stuck in your body, stuck in your head. You just want to always be trying to elevate somehow.”

Bill Murray

“Genius is 1% inspiration and 99% perspiration.”

Thomas Edison

“Prá frente é que é o caminho!”

Maria da Conceição

Acknowledgments

No one ever told me that pursuing a PhD is not only a professional but also a personal voyage, where you truly mature. For that, I would like to thank those that helped me in this walk and somehow became a part of who I am.

Thinking back in time, I must thank the PDBEB PhD Programme and Fundação para a Ciência e a Tecnologia (FCT). Obrigado pela confiança e por me darem a oportunidade de saciar a minha "sede de conhecimento".

A todos os "PDBEBs", classe 2008, que eu adoro e com quem partilhei imenso (Sofs, Bia, Carol, Pedro, Samira, Silvía, Joana e claro está o Marquito e a minha companheira de farmácia Diana), o meu obrigado.

Reuven was the person that accepted me in his laboratory, without having any idea in what he was getting into. To Reuven I owe my sincere thank you, for giving me this great opportunity, and most importantly for providing me a wealth of problems to solve, and of course the freedom to think and learn.

À Professora Celeste, pelo valioso apoio e pela extrema simpatia com que sempre me recebe! E também por ter aceite co-tutelar a minha tese. Um muito obrigado!

During these four years, far more important than the equipments and reagents I always had available, was the priceless help and support from all the past and present colleagues, whose list of names would be to big to enumerate. Hayley - Harry Potter. Marit and Adam - I hope everything goes well. Arnold - the golden curly boy. Jarno, thank you for the fruitful collaboration. Suzanne, Alejandro, Fabricio, Vittorio, Li, Boris, Joachim, Ruiqi - thank you for always being there to help. Mariette, thank you for everything and for sharing the chocolate. To all the B4 department for the supporting environment. Gözde, thank you for being my partner in craziness. Rita Maria e Rui Manuel, os meus dois Portugas com quem posso treinar o meu Português enferrujado. Mario, mamma mia, thank you for all your help and for the wonderful torta della nona. To one of my companions in crime – Koos, thank you for the patient and valuable help in explaining me all the wonderful world of bioinformatics. Viktoriya, my Piano teacher, thank you for your patience and support. Le Sage, thank you for your friendship, for your support and the precious help!

I would like to thank the members of the jury, for accepting the invitation and being part of this event. Ao Nuno, obrigada por toda a ajuda com as figuras da tese e por estares sempre disponível. Our collaborators for the knowledge and invaluable reagents they shared with us.

To my adoptive family – Nicolas and Rubina. I could not ask for better friends! Thanks for everything! The warm feeling of belonging and for a place that I can also call home. You both are a source of inspiration. Words are not enough to explain the level of complicity. I will miss you both a lot! Also, my apologies for all the hard times I gave you (specially you, Nicolas), I mean the moody Melo.

À minha P. por tudo! Por me aturar (eu sei que é difícil), por me ensinar tanto! E mais importante, por me deixar ser quem sou!

Aos meus Pais, pela grande lição de vida que me proporcionaram! Tudo que sou, devo-o a vocês.

À minha mana! Eu adoro-te, sempre foste a minha inspiração e o meu porto de abrigo quando algo me acontecia! Obrigado pela tua paciência, pela tua racionalidade e pés assentes no chão, quando precisava!

Contents

Abbreviations

3C	3-Chromosomal conformation capture
4C	4-Chromosomal conformation capture
3D-DSL	3-dimensional DNA selection and ligation
ALL	Acute lymphocytic leukemia
AR	Androgen receptor
Bvht	Braveheart
BJ/ET	Primary BJ fibroblast immortalized by the human telomerase reverse transcriptase
CBP	CREB-binding protein
CDKs	Cyclin-dependent kinases
CDKN1A	Cyclin-dependent kinase inhibitor p21
ceRNA	Competing endogenous RNA
CHART-seq	Capture hybridization analysis of RNA targets combined with deep sequencing
ChIP-seq	Chromatin immunoprecipitation combined with deep sequencing
ChIRP	Chromatin isolation by RNA purification
CTD	C-terminal domain
DBD	DNA binding domain
DHT	Dihydrotestosterone

DKC	Dyskeratosis congenital
DMD	Duchenne muscular dystrophy
DNA	Deoxyribonucleic acid
dsRNAs	Double stranded RNAs
DUSP4	Dual specificity phosphatase 4
E2	17β-oestradiol
ENCODE	Encyclopedia of DNA Elements
ER-α	Oestrogen receptor α
eRNAs	Enhancer RNAs
FoxA1	Forkhead box protein A1
FSHD	Facio-scapulo-humoral dystrophy
FDR	False discovery rate
GAS5	Growth arrest specific transcript 5
GR	Glucocorticoid receptor
Gro-seq	Global run-on sequencing
GWAS	Genome-wide association studies
H3K9ac	Histone H3 lysine 9 acetylation
H3K27ac	Histone H3 lysine 27 acetylation
H3K4me1	Histone H3 lysine 4 monomethylation
H3K4me3	Histone 3 lysine 4 trimethylation
H3K27me3	Histone 3 lysine 27 trimethyaltion
H3K36me3	Histone 3 lysine 36 trimethylation
H3K9me3	Histone 3 lysine 9 trimethylation
HATs	Histone acetyl transferases
HDAC	Histone deacetylase
hnRNPI	Heterogeneous Nuclear Ribonucleoprotein I
hnRNPK	Heterogeneous Nuclear Ribonucleoprotein K
HOTTIP	HOXA transcript at the distal tip
hTERT	human Telomerase reverse transcriptase
IER5	Immediate early response 5
IR	Ionizing radiation
K	Lysine
kb	Kilobases
LCR	Locus control region
LED	LncRNA activator of enhancer domains
lincRNAs	Long intergenic non-coding RNAs
lncRNAs	Long non-coding RNAs
LOH	Loss of heterozygosity
MAML1	Mastermind-like-1
MDM2	Murine double minute 2
MEF2C	Myocyte-specific factor 2C

miRNAs	MicroRNAs
MLL	Histone methyltransferase mixed-lineage leukaemia
mRNAs	Messenger RNAs
MSP	Methylation specific PCR
MS2-CP	MS2-coat viral protein
ncRNAs	Non-coding RNAs
ncRNAs-a	Activating non-coding RNAs
nt	Nucleotide
p53BERs	p53-bound enhancer regions
p53RERs	p53-regulated enhancer regions
p53FERs	p53-free enhancer regions
PAPPA	Pregnancy associated plasma protein A
PCAT	Prostate-associated transcripts
PCR2	Polycomb repressive complex 2
PCGEM1	Prostate-specific transcript 1
pRB	Retinoblastoma protein
PR	Progesterone
qRT-PCR	Quantitative reverse transcription polymerase chain reaction
REs	Response elements
RNA	Ribonucleic acid
RNAPII	RNA polymerase II
RNA-seq	RNA sequencing
Ser46	Serine residue 46
siRNA	Small interfering RNA
sncRNAs	Small non-coding RNAs
ssDNA	Single strand DNA
TE	Transposable element
TFs	Transcription factors
TSS	Transcription start sites
T-UCRs	Transcribed ultraconserved regions
UTR	Untranslated region
WDR5	WD repeat-containing protein 5
XIC	X chromosome inactivation center
Xist	X-inactive transcript

Summary

Underlying the multiple steps characteristic of cancer progression is genomic instability, as it provides the requirements necessary for the selective clonal advantage of cancer cells. The acquisition of a mutant genotype allows the outgrowth of these cells and manipulation of their local environment. However, the genome possesses efficient surveillance systems to detect and resolve defects in the DNA.

The tumor suppressor gene TP53, which was found to be mutated in about half of all human cancers, represents a known transcription factor often referred to as the "guardian of the genome" (Lane, 1992). The multifactorial activation of p53 results in a complex transcriptional response that classically culminates in: cell cycle arrest, senescence and/or a programmed cell death (apoptosis), achieved through the regulation of different target genes. Until recently, the p53 transcriptional network mainly revolved around its protein-coding target genes. However, it has been increasingly recognized that the thus-far known protein-coding p53 target genes cannot fully explain its tumor suppression activity (Brady et al., 2011).

The advent of next-generation sequencing techniques unraveled the existence of a large portion of non-coding elements in the genome. These elements can be transcribed into a new class of genomic regulators, the long non-coding RNAs (lncRNAs). The expression of enhancer RNAs (eRNAs), a class of long ncRNAs transcribed from enhancer elements that regulate gene expression, was shown to positively correlate with nearby protein-coding genes (Kim et al., 2010), which led us to hypothesize that part of the p53-regulatory network could be mediated

through its interaction with enhancer-RNA producing domains. For that, we used genome-wide chromatin-binding profiles and discovered p53-bound genome regions located far from any known p53 target gene. These regions essentially revealed, the presence of conserved p53 response elements and presented hallmarks of enhancer domains. We named these new elements as p53-bound enhancer regions (p53BERs), and observed the production of eRNAs in response to p53 induction. Our study further demonstrates that p53BERs enhance p53 transcriptional response accomplished through intra-chromosomal interactions with multiple neighboring target genes. Additionally, we provide evidence concerning the functionality of eRNAs in the enhancement of transcription and demonstrate the importance of these new RNA molecules required for a proper p53 response (Melo et al., 2013). We then set out to globally map all p53-regulated enhancers (p53RERs) and found that while many p53-induced enhancers contained p53-binding sites (p53BERs), most did not (p53-free enhancer regions, p53FERs), suggesting a different mode of activation.

Indeed, we were able to observe that for a subset of p53RERs the presence of a stress-activated lncRNA, termed LED (lncRNA activator of enhancer domains), is required. We show that LED is a direct p53-target that is important for the deposition of histone 3 lysine 9 acetylation (H3K9ac) (a transcriptional active enhancer mark) at a subset of p53RERs. Furthermore, LED's presence is indispensable for a proper p53-induded cell-cycle arrest. Finally, we uncover the promoter-associated hypermehtylation of LED in several cancer cell lines and human tumors.

The findings in this thesis provide evidence that lncRNAs represent functional elements embedded in the non-coding portion of our genome. We show that these newly emerging RNA molecules are important players in cancer biology, and can be a future therapeutic approach in the fight against cancer.

Keywords : LncRNAs, Tumor suppressor p53, Cancer Biology

Resumo

A instabilidade genómica é uma propriedade fundamental subjacente ao processo de carcinogénese, uma vez que confere às células cancerígenas a capacidade para a sua expansão clonal. A aquisição de um genótipo mutante permite não só a expansão destas células, assim como, a manipulação do ambiente em que estão inseridas. O genoma possui, no entanto, sistemas de vigilância que detetam e corrigem defeitos na cadeia de DNA.

O gene supressor tumoral TP53, mutado em cerca de metade de tumores humanos, é um fator de transcrição considerado o 'guardião do genoma' (Lane, 1992). A ativação multifactorial do gene TP53, que atua via indução da transcrição de diferentes genes-alvo, classicamente culmina na paragem do ciclo celular, senescência e/ou apoptose das células cancerígenas. Até recentemente, era considerado que a extensa rede de genes-alvo regulada pelo gene p53 era composta apenas por genes codificadores de proteínas. No entanto, tem sido cada vez mais aceite, que os genes codificadores de proteínas regulados pelo p53 não são causa suficiente para explicar o seu amplo papel como supressor tumoral (Brady et al., 2011).

O desenvolvimento de novas técnicas de sequenciação de última geração veio revelar não só que a maior parte do genoma de eucariotas é composto por DNA não codificante, mas também que estas regiões são transcritas numa nova classe de reguladores genómicos, os long non-coding RNAs (lncRNAs). O termo enhancer RNAs (eRNAs), uma classe de lncRNAs transcritos apartir de enhancers (em inglês) e reguladores da expressão génica, foi correlacionado com a expressão positiva de genes-codificantes de proteínas localizados nas

imediações da região não-codificante (Kim et al., 2010). Sendo assim, formulamos a hipótese de que o p53 exerça parte da sua actividade reguladora via interação com domínios produtores de enhancer-RNAs. Recorremos à técnica de genome-wide chromatin-binding profiles para testar a nossa hipótese e descobrimos regiões-genómicas ligadas a p53, mas que se localizam a grandes distâncias em relação a qualquer um dos genes-alvo de p53. A análise destas regiões permitiu identificar a presença de elementos conservados de ativação de p53 em domínios epigeneticamente caraterizados como enhancers. Nestas regiões, que designámos de p53-bound enhancer regions (p53BERs), constatámos que são produtoras de eRNAs, em resposta à activação de p53. O nosso estudo demonstrou também que os p53BERs aumentam a transcrição regulada pelo p53 através de interacções intra-cromossómicas com múltiplos genes-vizinhos. Adicionalmente, mostrámos não só que uma das funções dos eRNAs consiste no aumento da transcrição mas também na relevância do papel dos eRNAs na resposta regulada pelo p53 (Melo et al., 2013).

No seguimento dos nossos resultados, mapeámos todos os enhancers regulados pelo p53 (p53RERs, em inglês) e descobrimos que, enquanto a maioria dos enhancers induzidos por p53 possuem domínios de ligação ao p53 (p53BERs, em inglês), na sua maioria estes domínios estão ausentes (p53FERs, em inglês) o que sugere um modelo de activação diferente.

De facto, observámos que para um subconjunto de p53RERs é necessária a presença de um lncRNA activado pelo stress, que denominamos por LED (lncRNA activator of enhancer domains). Demonstrámos que o LED é um factor-alvo de p53 e é importante para a deposição do marcador epigenético de acetilação da histona 3 na lisina 9 (H3K9ac, em inglês) (um marcador de enhancer) no subconjunto de p53RERs. A presença de LED é indispensável para a correcta paragem do ciclo celular, induzida pelo p53. Finalmente, detetamos a presença de LED em diversas linhas celulares cancerígenas e tumores humanos.

O trabalho desenvolvido nesta Tese demonstra que os lncRNAs representam elementos funcionais do genoma não-codificante e que desempenham um papel crucial na biologia do cancro, podendo vir a ser utilizados em futuras abordagens terapêuticas.

Palavras chave : LncRNAs, Gene supressor tumoral p53, Biologia do Cancro

Chapter one
Introduction

1.1 Cancer | oncogenes and tumor suppressors

Cancer development is seen as a selective process where cellular fitness is based on the ability of a cell to grow and proliferate in a perturbed environment. This process ultimately leads to cumulative amounts of cellular disruptions that results in a greater clonal advantage and reflect changes in several fundamental properties of a cell. The loss of cellular homeostasis is in part driven by genetic alterations in sets of genes required for the proper stress response of cells.

Two main broad classes of genes can be mutated during the onset of cancer: proto-oncogenes and tumor suppressors. Proto-oncogenes normally promote cell growth and the mutated forms of these genes are called oncogenes, which are permanently active and lead to enhanced cellular growth. The tumor suppressor genes normally restrain growth, and mutations that inactivate them results in cells that no longer show a normal cell growth and division. These are broadly divided into caretaker and gatekeeper genes. Caretaker genes are mainly involved in the maintenance of the integrity of the genome and once mutated lead to a higher genomic instability, while gatekeeper genes are responsible for controlling, or inhibiting tumor cell growth.

Cancer as a sequential process is possible due to a combination of factors that result in a profound cellular deregulation.

1.2 TP53 | tumor suppressor gene

Cancer cells exhibit a highly diverse array of mutation, which often inactivate tumor suppressor genes. TP53 is the most frequently mutated tumor suppressor gene in human cancer with more than 50% occurrence (Beckerman and Prives, 2010). The p53 gene is located on chromosome 17p13 and encodes a (mostly) nuclear phosphoprotein of 53kDa. Since its discovery in 1979, as a target of the SV40 oncogenic virus (Lane and Crawford, 1979; Linzer and Levine, 1979), intensive research has been dedicated to p53 due to its major importance in tumor suppression.

First thought to act as an oncogene cooperating with Ras in the transformation of primary cells (Eliyahu et al., 1984; Parada et al., 1984), several findings questioned the ability of p53 to transform cells. It was only after several groups compared their p53 cDNA sequences, that researchers found that the oncogenic properties of p53 were obtained with mutated forms of p53 that had been isolated from tumor cells (Hinds et al., 1989; Hinds et al., 1990). Subsequent studies revealed that wild-type p53 is in fact a tumor suppressor and one of the most important players in cancer biology (Eliyahu et al., 1989; Finlay et al., 1989). As such, mutant p53 was found in several mouse tumors (Mowat et al., 1985; Wolf and Rotter, 1984), and it is nowadays accepted that TP53 mutations are present in more than 50% of human tumors, being acknowledged as the most common genetic alteration in human cancer (Beckerman and Prives, 2010).

1.2.1 p53 functional domains

TP53 encodes a transcription factor that accumulates upon stress conditions (DNA damage, hypoxia among others), activating or repressing a wide variety of target genes in order to maintain cellular homeostasis. Common to most transcription factors, p53 is composed of distinct domains: an N-terminal transactivation domain (composed of two transactivation subdomains, TAD-I and TAD-II), required for transcriptional activation, a proline-rich region, and an evolutionary conserved sequence-specific core DNA binding domain (DBD), followed by a tetramerization domain and a basic regulatory C-terminal domain (CTD) (Figure 1).

Most of TP53 mutations in human tumors can be found in its DNA binding domain (Brosh and Rotter, 2009), pinpointing p53's main function as a sequence-specific DNA binding protein crucial for tumor suppression. Additionally, p53 can undergo numerous posttranslational modifications such as phosphorylation, acetylation, ubiquitination, among others and recruit different transcriptional cofactors, including histone modifying enzymes and chromatin-remodeling factors (Kruse and Gu, 2008). For example, BRD7, a member of the bromodomain-containing family that recognizes monoacetylated lysine residues, was identified as an important p53 cofactor required for the transcription of a subset of p53 target genes (Drost et al., 2010).

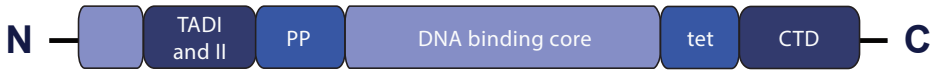


Figure 1. Structural and functional domains of p53.
The major functional domains of p53 are depicted. p53 is composed by an N-terminal transactivation domain a proline-rich region (PP), and a central sequence-specific DNA binding domain (DNA binding core) that is followed by tetramerization (tet) and basic regulatory C-terminal domains (CTD). The transactivation is subdivided into two transactivation subdomains, TAD-I and TAD-II.

1.2.2 p53 regulation and stress response

In unstressed cells, p53 protein levels are maintained at low levels. However, during their lifetime, cells are incessantly exposed to numerous kinds of stress conditions. As a main point of convergence, p53 integrates diverse types of stress-response pathways, and was named the "guardian of the genome" (Figure 2) (Lane, 1992).

Upon stress, p53 protein is stabilized through different posttranslational modifications, leading to its intracellular accumulation. Once activated, the tetrameric p53 acts primarily as a transcription factor, mediating a coordinated and complex transcriptional response, where it positively or negatively regulates hundreds of genes involved in several cellular functions.

The cellular response to p53 activation is highly variable and dependent on several parameters, such as the cellular context, the nature/extent of the damage but also on different p53 posttranslational modifications and interacting proteins. A key regulator of p53 is the murine double minute 2 (MDM2) oncoprotein.

MDM2 is an E3 ubiquitin ligase whose major role serves to keep p53 levels and activity in check (Momand et al., 1992). The interaction between MDM2 and p53 creates an autoregulatory loop where p53 positively regulates the transcription of MDM2, which in turn downregulates p53 (Picksley and Lane, 1993). The oncoprotein MDM2 binds to p53 N-terminal domain and physically blocks its transactivation domain (Lin et al., 1994).

In unstressed cells, MDM2 binds to p53 and catalyzes its ubiquitination and proteasomal degradation (Kubbutat et al., 1997). Additionally, various stress signals (oncogene activation, genotoxic damage, hypoxia, among others) can increase p53 activity resulting in the expression of MDM2 (Barak et al., 1993). However, when a cellular stress signal is sensed, MDM2 can polyubiquitylate itself resulting in its degradation and consequent p53 accumulation and activation. Furthermore, the activation of several oncogenes can disrupt the interaction between p53 and MDM2. This activation triggers the expression of ARF (whose product is a tumor suppressor) that binds and sequesters MDM2 in the nucleolus, decreasing p53 nuclear export, ubiquitination and consequent degradation, therefore allowing p53 accumulation and subsequent activation (Weber et al., 1999; Zhang et al., 1998).

The existence of a regulatory loop between two major cancer players makes it a therapeutically interesting target to pursue. Indeed, small molecules that target the interaction between p53 and MDM2 have been developed. One of such molecule is Nutlin-3a, which binds specifically to the N-terminal region of MDM2, and disrupts the regulatory loop, thereby increasing p53 stability and activation (Vassilev et al., 2004). Additionally p53 stabilization is influenced by a number of other factors that can dictate the amount of p53 available in the cell, adding a profound layer of complexity to the transcriptional response determined by this tumor suppressor. The phosphorylation and acetylation of p53 at its N- and C-terminal domain is one of the most relevant processes. For example, acetylation of p53's lysine residues is considered an important requirement for proper p53 stabilization and activation (Tang et

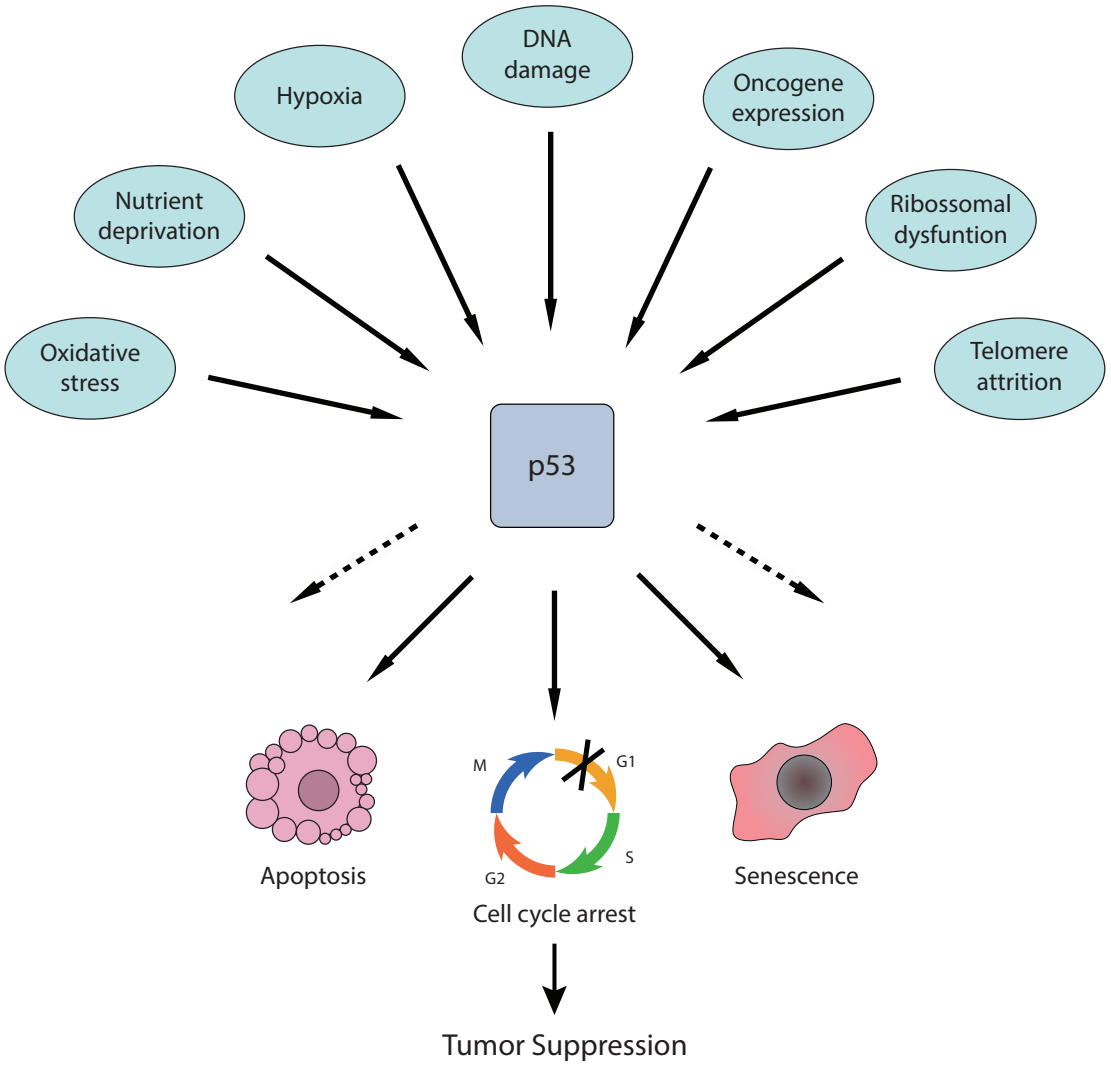


Figure 2. Activation and classical functions of p53

p53 has a key role in integrating various cellular stress signals that culminate in different responses. Ultimately, which response is achieved depends on the differential activation of target-gene expression. The role of p53 in tumor suppression is likely to depend on which cellular response is activated and on the context in which the activation occurs, however, it can classically culminate in three major outcomes: a transient cell cycle arrest, an irreversible growth arrest (senescence) or a programmed cell death (apoptosis).

al., 2008). Ultimately, p53 activation can classically culminate in three major cellular outcomes: a transient cell cycle arrest, an irreversible growth arrest (senescence) or a programmed cell death (apoptosis) (Figure 2).

1.2.3 Transcriptional regulation mediated by p53

1.2.3.1 DNA binding

A crucial step in p53 transcriptional stress response requires binding of p53 to its specific recognition site in the genome, in order to regulate gene expression. As a tetramer p53 will specifically bind to its response elements (REs). These are mainly found within a few thousand nucleotides upstream or downstream from the target transcriptional start site.

The classical p53 RE ligneous two inverted pentameric sequences defined as 5'-PuPuPuCWWGPyPyPy-3' (in which Pu is purine, W can either be an adenine or cytosine and Py represents a pyrimidine) separated by a spacer of 0 to 13 nucleotides (el-Deiry et al., 1992). It was observed, however, that some functional p53 binding sites could include elements that do not match the original p53 RE, adding a new piece to the complex p53 regulatory network (Gohler et al., 2002). One example involves the microsatellite region within PIG3 gene that was shown to have higher affinity for p53, than the canonical binding site in its promoter (Contente et al., 2002). The strong association of p53 with other DNA-binding proteins can explain such observation, especially when p53 is acting in a protein complex. Additionally, different p53 REs will possess different binding potential, which can influence its affinity for the target site, resulting in different levels of activation or inhibition of transcription. Furthermore, it has been shown that both p53 domains, the DBD and the CTD, are able to bind DNA. While the central DBD of p53 is responsible for the sequence specific recognition site in the genome, its CTD binds DNA in a non-sequence specific manner, mainly recognizing structural features of the target DNA (McKinney and Prives, 2002). The observation that p53 contains two distinct binding regions can provide a clue for the observation that on basal conditions a significant amount of p53 is still binding to DNA, despite its low levels in the cell (Espinosa et al., 2003).

Altogether, these findings support the DNA binding potential of p53 at basal and stress conditions, thereby expanding the role and importance of this tumor suppressor to the normal homeostasis of a cell.

1.2.3.2 Modulation of low and high affinity target genes

The p53 stress response acts in three major steps. The first one involves the protein stabilization upon a stress signal, mainly through the release of its negative regulator, MDM2.

The second step can be characterized by an increase in the p53 DNA binding capacity to different subsets of targets. The selective regulation of different clusters of targets will result in a distinct p53 stress response.

As a third step, p53 binding allows the recruitment of specific cofactors (coactivators or corepressors) that result in the transcriptional regulation of the target gene (Figure 3). However, several distinctive stress responses can be mediated by p53 upon different stimuli, through selective binding to different promoters of target genes. This selectivity invokes one of the most fascinating subjects in the p53 field. How is p53 able to choose between the different set of targets in response to different stimuli?

Selectivity can be influenced by several factors such as p53 posttranslational modifications, p53 binding partners and p53RES. These factors are crucial for the target gene choice, from the repertoire of p53-regulated genes. In addition, the cellular context and the differences in the nature/extent of the stress are also important and determinant players. Notably, the cell type variations in the availability of cofactors are also essential aspects to the regulation of specific targets that ultimately drive the cells to a particular response.

Certain posttranslational modifications can also impact the selectivity of p53 for the activation or repression of certain targets. For example, the acetylation of p53, crucial for its proper activation, has been shown to play a role in dictating the preferences for different targets. Previous studies observed that certain p53 lysine residues could be acetylated upon a genotoxic stress. For example, lysine (K) 120 (K120), acetylated by the MYST family of acetyl transferases (TIP60 and hMOF), is required for the transcriptional activation of proapoptotic targets such as PUMA and BAX (Sykes et al., 2006; Tang et al., 2006). Thus, it was suggested that this acetylated form of p53 accumulates more on proapoptotic promoters. One other example involves the K164 within the DBD of p53. This lysine residue is modified by the transcriptional activators p300 and CREB-binding protein (CBP), and together with the acetylation present at the CTD of p53, is an important event for the activation of the majority of p53 target genes (Tang et al., 2008). Additionally, the acetylation of this tumor suppressor is required for the recruitment of cofactors that are necessary for the proper activation of p53 target genes.

Another clear example of p53 posttranslational modification refers to its phosphorylation. Studies showed that upon DNA damage, p53 serine residue 46 (Ser46) is phosphorylated, inducing specifically the transactivation of the proapoptotic gene AIP-1 (Oda et al., 2000). This residue is mainly modified by the kinase HIPK2 at the N-terminal domain of p53 (D'Orazi et al., 2002). These and other posttranslational modifications ensure the appropriate response of p53 to different stimuli. The stress and cell-type specificity pattern of these modifications, regulate the intensity and nature of p53 response. This increases the number of combinatorial possibilities that influence the target gene choice inherent to this tumor suppressor. It has also been suggested that these modifications can impose conformational changes in p53, allowing the recognition of diverse promoters and a selective regulation of target genes (Kim and Deppert, 2003).

Within each target, the panoply of different p53 REs suggests the existence of different

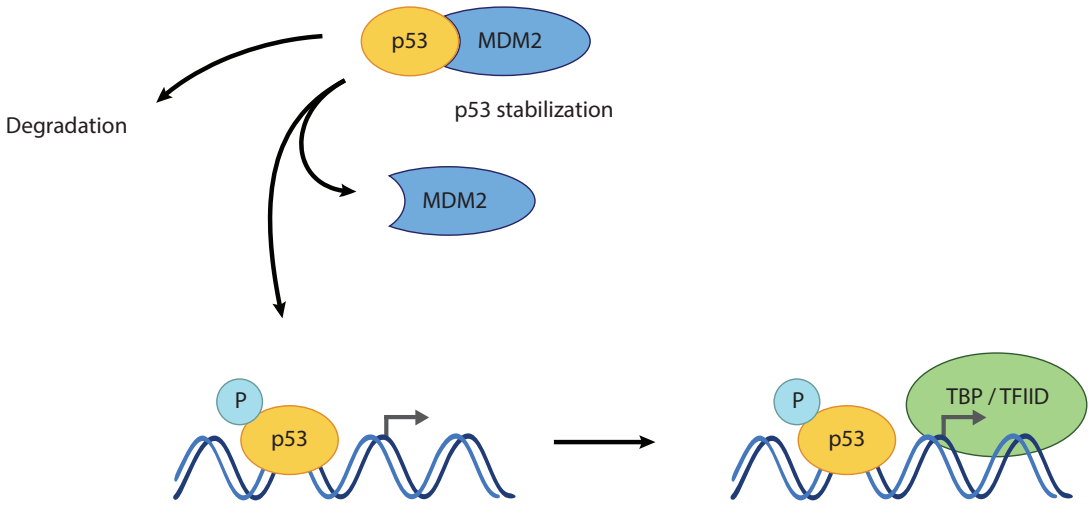


Figure 3. Classical model of p53 activation
The classical model for p53 activation generally consists in three events: (1) stress-induced stabilization mediated by release of MDM2 and phosphporylation (P) of p53; (2) Increase in the p53 DNA binding capacity; (3) Recruitment of cofactors and target gene regulation.

affinities for different consensus sequences. In general, p53 possesses a greater affinity for promoters of cell cycle target genes than promoters of proapoptotic targets (Weinberg et al., 2005). This differential affinity resonates with the amount of p53 in the cell. Depending on the nature/strength of the stress response, p53 will be stabilized and activated to different degrees, affecting the target gene choice and, consequently, the cell fate. Indeed, low levels of p53 preferentially activate high affinity pro cell-cycle arrest genes whereas higher p53 levels will override this default pathway and activate the low affinity promoters of proapoptotic genes (Laptenko and Prives, 2006).

One of the most known high affinity p53 targets is the cyclin-dependent kinase inhibitor p21 (CDKN1A). This was the first p53 target gene to be identified (el-Deiry et al., 1993; Harper et al., 1993), and has been suggested to mediate the p53-induced growth arrest triggered by stress. P21 exerts its cell cycle effects by inhibiting the kinase activities of a class of cyclin-dependent kinases (CDKs), which are required for the phosphorylation of the retinoblastoma protein (pRB). This phosphorylation event is crucial for the dissociation of pRB from the transcription factor E2F1, necessary for the cellular progression from G1 to S phase of the cell cycle (Nakanishi et al., 1999). In addition to the CDKs inhibition, p21 is able to independently inactivate pRB through proteasome-mediated degradation (Broude et al., 2007).

In response to DNA damage, p53 directly activates p21, which induces a transient G1 arrest. This allows the time for the cells to repair their genome before proceeding through the cell cycle (el-Deiry et al., 1993). In case the DNA is not repaired, p53 can trigger apoptosis, through the activation and repression of many other different target genes that include the proapoptotic BCL2 family members (BAX, PUMA, among others) (Fridman and Lowe, 2003). All in all, upon an initial stress signal the cell will classically respond with a reversible cell cycle arrest, while the irreversible apoptotic response requires additional stress stimulations.

The complexity within the p53 transcriptional response creates an intricate network of events that can dictate the cell fate. Which outcome (cell cycle arrest, senescence or apoptosis) prevails is highly influenced by a number of factors, such as the stress input as well as the cellular context.

1.2.4 Non-coding RNAs and p53 response

With the growing number of p53 targets being discovered, it has become clear that p53 is involved in other processes that go beyond the classical responses, previously mentioned. Interestingly, observations suggest that the thus-far known protein coding p53 target genes cannot fully explain the tumor suppression activity of p53 (Brady et al., 2011), raising new possibilities. Until recently the p53 network complexity mainly revolved around its protein coding target genes. The recent discovery of non-coding RNAs (ncRNAs) opened the door for

a new regulatory mechanism between p53 and this new class of targets.

Over the past decade, evidence from numerous next generation sequencing experiments showed that the growing complexity associated to each organism is reflected by a clear increase of the non-coding portion of the genome (Mattick, 2004). Many of these elements are actually transcribed into ncRNAs, establishing their importance as regulators of multiples biological processes. MicroRNAs (miRNAs) were the first class of ncRNAs to be associated with the p53 pathway.

MiRNAs are a class of small (17-25 nucleotide;nt) single stranded RNAs, present in plants and animals (Bartel, 2004). These ncRNAs can posttranslationally silence gene expression by recognizing complementary target sites most often in the 3' untranslated region (UTR) of target messenger RNAs (mRNAs). Generally, nucleotides from the 5'-end of the mature miRNA, named seed sequence, are crucial for binding to its target sequences. This results in translational inhibition and/or transcript decay (Bagga et al., 2005; Giraldez et al., 2006). Several studies have shown a close association between p53 and miRNAs, which can act as tumor suppressors or oncogenes.

For example, miRNA profiling after p53 induction revealed the miRNA-34a, - 34b and -34c as the most responsive miRNAs (Chang et al., 2007; He et al., 2007; Raver-Shapira et al., 2007). Indeed, the miRNA-34 family is highly induced after genotoxic stress in a p53 dependent manner in vitro and in vivo (He et al., 2007; Raver-Shapira et al., 2007). Additionally, ectopic expression of this miRNA family was able to recapitulate the phenotype of p53 activation, promoting cell cycle arrest, apoptosis and senescence (Bommer et al., 2007; Chang et al., 2007; He et al., 2007). These effects were explained by the repression of several direct miRNA-34 targets such as BCL-2 (Bommer et al., 2007). While this miRNA family mainly functions as a tumor suppressor gene, other reports observed that miRNAs could also participate in tumorigenesis.

One example involves the miR-372&373 cluster that favors proliferation and tumorigenesis of primary human cells, by neutralization of downstream signaling of p53 (Voorhoeve et al., 2006). Additionally, it has been observed that TP53 itself can be targeted by miRNAs. MiRNA-504 targets TP53 3'-UTR, leading to reduced protein levels and subsequent impaired p53 function (Hu et al., 2010). Collectively, the existence of a deep complicity between p53 and miRNAs unveils the importance attributed to these ncRNAs.

Recently, another class of ncRNAs, the long non-coding RNAs (lncRNAs), emerged as powerful players of gene regulation. LncRNAs are defined as RNA molecules greater than 200 nucleotides in length that can regulate gene expression trough diverse mechanisms, which are not yet fully understood. Numerous reports revealed their multifaceted roles in many biological processes. For instance, it has been demonstrated that some lncRNAs are bound by p53 in their promoter regions, suggesting a deep relation between this tumor suppressor and these newly emerged RNA molecules.

1.3 Long non-coding RNAs

The genome is a rapidly evolving entity that constantly reshapes its structure by acquiring, losing or rearranging its DNA. This great plasticity certainly contributed to the evolution and complexity of the human genome.

With the recent explosion of refined genomic and large-scale sequencing technologies, we are now facing the task to redefine certain concepts and definitions no longer in consonance with collected experimental data. For instance, the central dogma of molecular biology stated that the genetic information flows from DNA to RNA to protein (Crick et al., 1961). Although this concept is largely accurate in prokaryotes, where genomes are mainly composed of protein-coding genes, it doesn't reflect the reality of higher eukaryotes, where protein-coding genes occupy less than 3% of the genome. This discrepancy suggests a correlation between the emerging complexities of a genome with the expansion of its non-coding portion. Initially believed to be transcriptionally inactive, the predominant fraction of the genome is in reality pervasively transcribed into thousands of different non-coding RNAs. While this finding challenges the established view of RNA as a modest intermediary in protein synthesis, it also brings a paradigm shift toward the recognition of RNAs as functional molecules that shape the evolution of complex organisms.

Non coding RNAs are transcripts that contain little or no open reading frame and are in general divided into two main categories: small ncRNA (sncRNAs) (<200 bp) and long ncRNAs (lncRNAs) (>200 bp). The initial observation by Andrew Fire and Craig C. Mello in the 1990s, that double stranded RNAs (dsRNAs) were causing gene silencing when exogenously introduced into *C. elegans*, generated a great interest for sncRNAs (Fire et al., 1998; Hammond et al., 2000; Zamore et al., 2000). The family of sncRNA has since expanded into various classes including guide RNAs (small nucleolar RNAs), splicing RNAs (small nuclear RNAs), housekeeping RNAs (ribosomal RNAs and transfer RNAs), regulatory RNAs (microRNAs, PIWI-RNAs) and several transcription-associated RNAs (promoter upstream transcripts, promoter-associated RNAs, transcription initiation RNAs and termini-associated small RNAs).

The curiosity for lncRNAs also spurted in the 1990s with the initial characterization of two important genes: the imprinted H19 (Bartolomei et al., 1991) and the X-inactive transcript (Xist) (Brown et al., 1991). In the first case, the Tilghman's group identified H19 (clone 19 in row H) while screening for genes regulated by α-foetoprotein in the liver. This discovery led to the description of H19 as a 2.3-kb capped, spliced and polyadenylated transcript that had no conserved open reading frames and further established the locus as an archetype of imprinted genes (Bartolomei et al., 1991). In the second case, Xist was identified as a non-coding RNA produced at the X chromosome inactivation center (XIC). Only expressed from the inactive X chromosome, it was then demonstrated that the RNA recruits the polycomb-group proteins to catalyze the deposition of Histone 3 lysine 27 trimethylation (H3K27me3) in cis to epigenetically repress the X chromosome (Brown et al., 1991; Simon et al., 2013).

During the following decades, lncRNAs increased in number and evolved in different classes such as long intergenic non-coding RNAs (lincRNAs), enhancer RNAs (eRNAs) and activating non-coding RNAs (ncRNAs-a). These recent developments led to the discovery of a plethora of functional ncRNAs with surprising and varied functions, ranging from the simple structural module to potent transcriptional, translational and epigenetic regulators.

1.3.1 Origin and epigenetic signatures

Recent technological advances in high-throughput functional genomics resulted in the identification of new regulatory RNA molecules termed long non-coding RNAs. While intense investigation is being directed to identify and understand their functions, few reports have addressed their origin.

Initial analyses of their sequences suggested that lncRNAs are less conserved than protein-coding genes but slightly more than neutrally evolving genomic regions (Guttman et al., 2010; Marques and Ponting, 2009). However, efforts to find distant homologues to human lncRNAs have been less successful when compared to protein coding-genes, suggesting a rapid and species-specific evolution (Cabili et al., 2011; Kutter et al., 2012; Nam and Bartel, 2012; Ulitsky et al., 2011). Interestingly, different hypotheses have been proposed to explain the formation of lncRNAs including transformation of protein-coding genes, duplication of another lncRNA and emergence from transposable element (TE) sequences (Cabili et al., 2011; Derrien et al., 2012; Dinger et al., 2008; Guttman et al., 2009; Jia et al., 2010; Rinn and Chang, 2012; Ulitsky et al., 2011). Although examples have been found for each hypothesis, TEs seem to be an interesting option as they occupy at least half of the human genome (Lander et al., 2001).

TEs are mobile genetic elements divided in two classes: class I (retrotransposons) and class II (DNA transposons). While class I TEs are copied in two steps, involving transcription (DNA → RNA) and reverse transcription (RNA → DNA), class II TEs use a cut-and-paste transposition (no RNA intermediate) that is performed by various transposases. Of note, distinct reports have already highlighted TEs as the source of different genomic regulatory domains such as promoters, enhancers and insulators (Bejerano et al., 2006; Faulkner et al., 2009; Schmidt et al., 2012). It was also demonstrated that the lncRNA Xist is derived from a mixture between a decayed protein coding gene and the progressive accumulation of various TEs (Duret et al., 2006). In addition, Kelley and Rinn (2012) showed that TEs preferentially accumulate at the transcription start sites (TSS) of lincRNAs (a class of lncRNAs). This assumes a special relevance since lincRNAs devoid of TEs are generally expressed at greater levels (Kelley and Rinn, 2012).

A growing body of evidence now suggests that TEs are omnipresent in lncRNAs, suggesting that they might have shaped their evolution and function.

Beside their sequence composition, lncRNAs can also be characterized by epigenetic

and RNA features. LncRNAs are classically capped and polyadenylated transcripts produced by RNA polymerase II. Their genomic location usually harbor epigenetic features of protein-coding genes, such as histone 3 lysine 4 trimethylation (H3K4me3) at the transcription start site and histone 3 lysine 36 trimethylation (H3K36me3) throughout the gene body (Guttman et al., 2009). These features regroup several classes of lncRNAs such as lincRNAs, transcribed ultraconserved regions (T-UCRs), as well as numerous antisense RNAs. However, enhancer RNAs (eRNAs), another class of lncRNAs, present different epigenetic and RNA features, and are unidirectional (1D-eRNA) or bidirectional (2D-eRNA) transcripts produced by RNA polymerase II.

1.3.2 Mechanisms of action

1.3.2.1 LncRNAs as chromatin modifiers

Regulation of gene expression is a highly dynamic process that primarily depends on the chromatin accessibility, which is in turn dictated by the intricate interactions existing between transcription factors, histone modifications, chromatin modifiers and remodelers. Thus, the epigenetic features of a genomic region will largely influence the balance between a more open or closed chromatin state. For instance, active genes are known to have promoter regions decorated with H3K4me3 and histone 3 lysine 27 acetylation (H3K27ac), while repressed or inactive gene promoters are classically marked by H3K9me3 and H3K27me3 (Ernst et al., 2011).

The exact mechanisms that govern the dynamic variation of epigenetic marks remain mostly unknown. However, it is becoming clear that lncRNAs have an essential role in the process. An interesting example is the lncRNA HOTTIP (HOXA transcript at the distal tip) and its impact on the HOXA locus (Wang et al., 2011b). The gene encoding HOTTIP is located upstream of HOXA13 at the 5' tip of the HOXA locus. The transcription of HOTTIP generates a 3.7kb transcript that influences the activation of several 5' HOXA genes. Importantly, it has been shown that HOTTIP gene is physically brought in close proximity with the 5' HOXA gene promoters, by DNA looping. Once produced, HOTTIP RNA can interact with the WD repeat-containing protein 5 (WDR5) protein, which then recruits the histone methyltransferase mixed-lineage leukaemia (MLL). The tethered WDR5-MLL complex finally catalyzes the methylation of local H3K4 and activates the transcription of the HOXA locus, thereby influencing development and cell fate.

In contrast, lncRNAs have also been found to negatively regulate gene expression. Gupta and colleagues recently reported that the lncRNA HOTAIR was able to repress the HOXD cluster genes by a trans-acting mechanism (Figure 4A) (Gupta et al., 2010). HOTAIR operates with the help of two functional domains required to interact with polycomb repressive

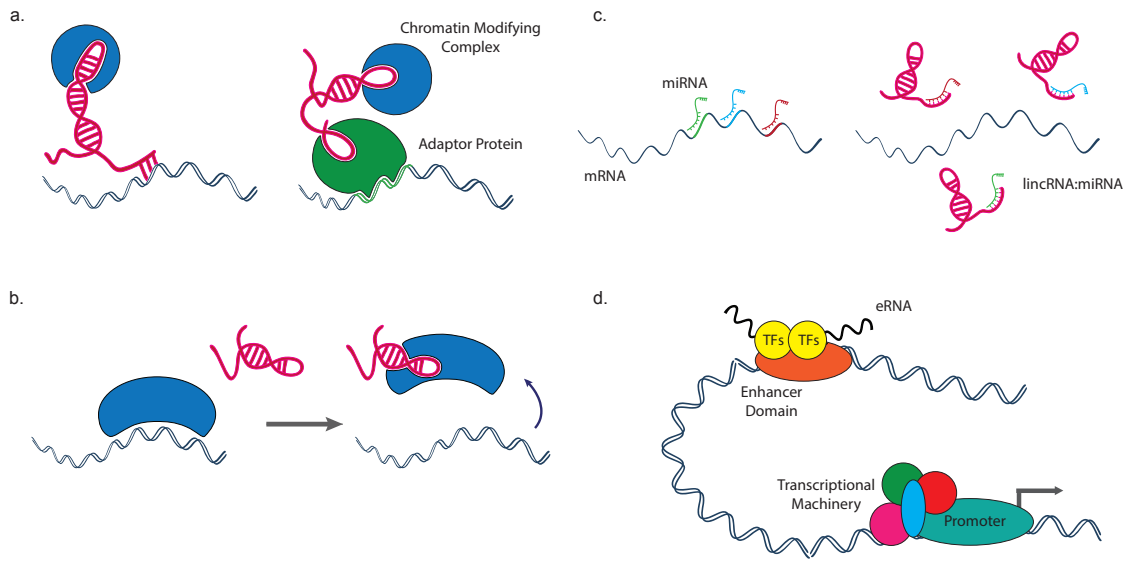


Figure 4. Mechanisms for long noncoding RNA (lncRNA) function
(a) lncRNAs can act as guiders, tethering specific chromatin-modifying complexes (blue) to DNA; this can occur through direct RNA-DNA interaction or mediated by a DNA-binding protein- adapter (green). (b) lncRNAs can function as decoys, binding and sequestering transcription factors from their target DNA sequence. (c) lncRNAs can interact with miRNAs, serving as molecular sponges and preventing them from binding to mRNA. (d) eRNAs, transcribed from enhancer domains, directly affect the expression of neighbouring genes, through chromatin-looping. TFs-transcription factors.

complex 2 (PCR2) and lysine-specific demethylase 1 (LSD1/CoREST1) complexes. By tethering these two repressive chromatin modifiers, HOTAIR organizes their action and facilitates their recruitment to specific promoter regions, where they deposit repressive epigenetic marks such as H3K27me3 (Tsai et al., 2010). Notably, the epigenetic remodeling mediated by HOTAIR has an important role in skin differentiation (Chuong, 2003; Rinn et al., 2007).

A similar repressive mechanism was also employed by the lincRNA-p21. The tumor suppressor p53 was shown to control the expression of lincRNA-p21 in order to repress a subset of genes related to apoptosis (Huarte et al., 2010). Although the mechanism was not entirely clarified, it was suggested that the repressive function of lincRNA-p21 was mediated at promoters through its interaction with the Heterogeneous Nuclear Ribonucleoprotein K (hnRNPK).

An elusive aspect of lncRNAs function resides in their capacity to reach or interact with specific chromatin regions. Although several hypotheses, such as formation of specific RNA-DNA hybrids, interaction with an adapter protein or direct association to a transcription factor have been described, only few reports have so far properly addressed this question. One well-characterized case involves the lncRNA Mistral. Transcribed from a region located between the Hoxa6 and Hoxa7 genes, the lncRNA Mistral was shown to locally recruit and anchor the MLL complex by the formation of an RNA-DNA hybrid (Figure 4A). Interestingly, the 3' region of Mistral can form a stem-loop structure that is specifically recognized by the SET domain of MLL1, gaining the ability to bind single-strand DNA (ssDNA). The local accumulation of the MLL complex favors the deposition of active H3K4me3 that ultimately promotes the transcriptional regulation of Hoxa6 and a7 genes, influencing stem cell differentiation (Bertani et al., 2011).

1.3.2.2 LncRNAs as decoy

The transcriptional activity of a gene depends on the balance between silencing and activating cues. Because of their ability to associate with RNAs and proteins, lncRNAs are interesting candidates to fine-tune or impact such balance.

Apart from their direct and local influence on promoters, lncRNAs can also act as decoys by keeping activators or repressors away from their targets. This is the case for the recently discovered lncRNA braveheart (Bvht). The authors observed that Bvht is necessary for the activation of a gene network involved in cardiac lineage commitment (Klattenhoff et al., 2013). Since Bvht interacts with SUZ12 (a component of PRC2) during cardiomyocyte differentiation, the authors suggested that Bvht mediates epigenetic changes towards cardiac commitment, by sequestering PRC2 from its target genes (Figure 4B).

In addition to protein decoy, lncRNAs can crosstalk with other RNAs by competing for shared miRNAs. The muscle-specific lncRNA MD1 (linc-MD1) participates in the timing

of muscle differentiation by acting as a decoy for miRNAs (Figure 4C) (Cesana et al., 2011). Cesana and colleagues demonstrated that linc-MD1 could function as a competing endogenous RNA (ceRNA) for the miR-133 and miR-135, which negatively regulate mastermind-like-1 (MAML1) and myocyte-specific factor 2C (MEF2C), two prominent factors involved in myogenesis. Supporting the crucial role of linc-MD1 in the control of the myogenic program, it was found that linc-MD1 expression is reduced in myoblasts derived from Duchenne muscular dystrophy (DMD) patients, corroborating their limited ability to undergo terminal differentiation. Importantly, re-introduction of the linc-MD1 in DMD myoblast could partially rescue the timing and the expression of myogenic factors required for the differentiation program.

These observations emphasize the critical function of lncRNAs in many cellular processes, and suggest for their potential influence on the delicate balance between healthy and diseased conditions.

1.3.3 Enhancer RNAs

1.3.3.1 Enhancers

Recent findings described the existence of ncRNAs emerging from regulatory enhancer regions. Enhancers were originally defined as DNA elements capable of modulating gene expression over long distances, independently of their orientation (Bulger and Groudine, 2011).

The distances at which they can act vary greatly (few to hundreds of kilobases, kb). These genomic elements are committed to the regulation of the transcriptome where they contribute to fine-tune and control tissue-specific gene expression. They classically work through a series of events that involve DNA looping formation and stabilization between enhancer-promoter, where regulatory activity is transferred from enhancer onto the transcriptional machinery at the promoter. Moreover, several reports suggest that enhancers can be the subjects of coordinated regulation. This observation is in agreement with the existence of several recognition sites for transcription factors at enhancers (Heinz et al., 2010a; Lupien et al., 2008).

Recently, high throughput sequencing techniques, such as RNA sequencing (RNA-seq), allowed a great advancement in the fields of epigenetics and genomics, contributing to a deep insight into the properties, activity and biology of enhancers. Numerous reports observed that many of these regulatory elements possess an open chromatin state. As a consequence, these enhancers are more sensitive to endonucleases and are thereby pre-marked by the presence of DNase hypersensitivity regions, which allow the binding of transcription factors (He et al., 2010). Additionally, it is known that DNA can wrap around histone proteins forming nucleosomes and that these can exhibit different levels of compaction: less nucleosome compaction is characteristic of an open chromatin configuration (euchromatin), while the

reverse denotes transcriptionally inactive chromatin regions (heterochromatin) (Felsenfeld and Groudine, 2003).

Observations point to the fact that active enhancers have eviction of a central nucleosome, where transcription factors bind, and two well-positioned nucleosomes flanking this "open region" (He et al., 2010). Chromatin signatures were discovered to be a valuable source of information for the prediction of enhancer activity. The various posttranslational histone modifications introduced a new overview concerning enhancer potentiality. Indeed, studies involving chromatin-immunoprecipitation combined with deep sequencing (ChIP-seq), revealed that active enhancers are often characterized by a combination of a higher ratio H3K4me1 over H3K4me3. This ratio, together with CREB-binding protein (CBP) and p300 presence, are commonly used to predict enhancer domains (Heintzman et al., 2007b). Furthermore, the presence of other chemical markers such as acetylation of histone H3 at lysine 27 and lysine 9 (H3K27ac, H3K9ac) provides compelling evidence for the presence of active enhancers (Creyghton et al., 2010b; Ernst et al., 2011).

1.3.3.2 Enhancer transcription and its roles

Transcription at enhancers was firstly reported in the locus control region (LCR) of the beta-globin cluster (Tuan et al., 1992). The authors discovered a ncRNA transcribed from a DNase hypersensitivity region at an enhancer. Since then, advanced sequencing technologies allowed genome wide studies, which demonstrated that transcription at enhancers represents a more widespread phenomenon than originally thought.

Two landmark studies recently demonstrated that RNA polymerase II can be recruited to active enhancers, resulting in enhancer-derived transcription from these genomic elements. Strikingly, enhancer–derived transcripts correlated with the expression of neighboring genes (De Santa et al., 2010; Kim et al., 2010). In one report, the authors observed, through deep sequencing approaches, that a specific cellular stimulation (potassium chloride), in neurons, resulted in the production of bidirectional transcripts. These RNAs extended around 1.2kb from the transcription factor binding site and their expression correlated with neighboring gene activation, thereby originating the term enhancer RNAs (eRNAs) (Kim et al., 2010). These transcripts are most commonly bidirectional (2d-eRNAs) and unpolyadenylated, however, it was observed that eRNAs can also be unidirectional (1d-eRNAs) and polyadenylated (De Santa et al., 2010).

Collectively, enhancer transcription associates with enhancer activity, as assessed by the recruitment of specific stimulus transcriptional activators, histone acetylation and vicinity to stimulus-induced genes through DNA looping (Figure 4D).

Follow-up studies, involving different signaling pathways, grounded the first observations of enhancer transcripts, where several TFs regulate the coordinated expression of these RNA

molecules. However, RNA reverse transcription coupled to massively parallel RNA-seq often fails to detect less abundant or highly unstable RNA species.

New emerging technologies, such as global run-on sequencing (Gro-seq) (Core et al., 2008), allow for a greater perspective over the genome dynamics. This technique is able to measure newly synthesized transcripts, assessing the status of transcriptional activity in any cell, under any condition, establishing its efficacy in the detection and characterization of eRNAs.

In one report (Wang et al., 2011a), the authors pursue the study of cell-lineage-specific factors, such as forkhead box protein A1 (FoxA1), a transcriptional activator that facilitates the DNA binding of other factors, such as androgen receptor (AR). FOXA1 is a key component of the AR transcriptional complex, however, the relationship between AR and FOXA1 is still not resolved. Using ChIP-seq on prostate cancer cells treated with androgen (dihydrotestosterone-DHT), Wang et al (2011) observed a strong enrichment and overlap between FoxA1 and AR binding at enhancer regions. Interestingly, depletion of FoxA1 did not alter some of the AR binding regions, and notably it was found a significant gain in the AR binding program. Using Gro-seq technology in DHT treated cells they were able to observe an increase expression of eRNAs. The expression of these RNA molecules significantly correlated with the AR binding program: eRNA production was high in association with the new AR binding program in the absence of FoxA1, while the lost AR binding regions showed reduced eRNA levels. Collectively, eRNAs can be good indicators of enhancer activity, participating in diversifying the transcriptional responses and with the possibility to give insight into disease progression.

1.3.4 LncRNAs in diseases

Accumulating evidences suggest that lncRNAs are critical regulators of several key biological processes such as transcription, splicing, RNA stability/decay and translation. In respect with their potential and wide range of action, it is not surprising to find an increasing number of dysregulated lncRNAs in a diverse range of diseases including Beckwith-Wiedemann syndrome (h19), diabetes (cdkn2b-as), Huntington's disease (har1a), Alzheimer's disease (51a) and various type of cancers (uca1, hotair, malat1, anril).

In the context of cancer, several lines of evidence suggest that inactivation of tumor suppressive lncRNAs is a driving force either in the initiation or the progression of the disease. For instance, the lncRNA GAS5 (growth arrest specific transcript 5) can interact with the glucocorticoid receptor (GR) and suppress its ability to reach and regulate its target genes (Kino et al., 2010).

Glucocorticoids are known to influence important cellular processes such as cell growth, metabolism and survival. Importantly, while GAS5 overexpression induces growth arrest and apoptosis in breast cancer cell lines, its expression levels are significantly reduced in breast tumors (Mourtada-Maarabouni et al., 2009). Here, GAS5 was found to interact with other

steroid hormone receptors known to be involved in breast cancer, such as androgen (AR) and progesterone (PR) receptors, thus further supporting the implication of GAS5 in tumorigenesis (Brisken, 2013; Chang et al., 2013).

The imprinted lncRNA MEG3 is known to inhibit cell growth by regulating the levels of the tumor suppressor p53 protein (braconi 2011). Similarly to GAS5, MEG3 expression was undetectable in a panel of human tumors (Zhang et al., 2003). In addition, MEG3-null mice had an increased vascularization of the brain, suggesting that MEG3 may prevent tumor growth by inhibiting the angiogenesis (Gordon et al., 2010).

Alternatively, lncRNAs with oncogenic function have been found to be upregulated in cancers. The lncRNA urothelial carcinoma associated 1 (lncRNA-UCA1) was initially identified as being upregulated in bladder cancer and more recently observed to be overexpressed in breast cancer specimens as well (Huang et al., 2014; Wang et al., 2006). The lncRNA-UCA1 promotes the progression of tumor growth by reducing the translation of the tumor suppressor p27 (kip1). It was demonstrated that UCA1 can interact with the Heterogeneous Nuclear Ribonucleoprotein I (hnRNPI) protein and prevent the latter (by decoy) from binding to and enhancing the translation of p27 mRNA. In support of these findings, reducing the expression of UCA1 in MCF-7 cells resulted in a decreased tumor growth when transplanted into mice.

Using large panel of tissue samples in combination with high-throughput sequencing, led to the identification of 121 lncRNAs associated with prostate cancer. Among these prostate-associated transcripts (PCAT), the functional characterization of PCAT-1 pinpointed its implication in cancer cell proliferation. Mechanistically, PCAT-1 is a transcriptional repressor that preferentially regulates several tumor suppressor genes such as BRCA2 (Prensner et al., 2011).

ANRIL is another upregulated lncRNA in prostate cancer tissues that influences the expression of potent tumor suppressors (Kotake et al., 2011; Yap et al., 2010). It was shown that ANRIL can interact with CBX7 (a subunit of PRC1) and SUZ12 (a subunit of PRC2), to mediate in cis the transcriptional repression of INK4a-ARF-INK4b locus. Importantly, genome-wide association studies (GWAS) have also shown that ANRIL locus is associated with increased susceptibility to diseases such as intracranial aneurysm, type 2 diabetes and different types of cancers (Pasmant et al., 2011).

Mutations or sequence variations within lncRNA gene loci have also been associated with heritable diseases. For example, mutations in TERC gene have been associated with dyskeratosis congenita (DKC) (Vulliamy et al., 2006). Some of these mutations destabilize TERC or change its interaction with the telomerase reverse transcriptase (hTERT), which results in a poor telomere maintenance. Affected individuals have an increased risk of developing life-threatening conditions such as cancer and pulmonary fibrosis (Marrone et al., 2007). In the autosomal-dominant facio-scapulo-humeral dystrophy (FSHD) syndrome, a reduction in the copy number of the repressive D4Z4 repeat causes an epigenetic switch that leads to the activation of a lncRNA named DBE-T. The de-repressed DBE-T recruits the Trithorax group

protein Ash1L and coordinates long-range gene regulation that contributes to the progression of the disease (Cabianca et al., 2012).

Collectively, lncRNAs add an important regulatory layer to cellular processes modulating gene expression. Additionally, lncRNA research can represent one important advance towards the understanding of the "missing heritability" concept. This concept hypothesizes that many traits, including diseases, cannot be accounted for by individual genes and seem to have their origins in the non-coding portion of the genome.

1.4 Study outline

Although progress has been made in identifying the genetic and epigenetic causes of cancer and, therefore, uncovering possible therapeutic targets, several challenges remain. Over the past several years, it has become clear that ncRNAs represent essential players in the regulatory mechanisms that govern cellular homeostasis.

The transcription factor p53 represents one major player in the maintenance of this delicate balance. Upon a cellular stress, the p53 pathway plays a determinant role in protecting our cells by integrating several stress signals to activate a specific transcriptional response.

Recently, lncRNAs emerged as powerful regulators of gene expression, with numerous profiling studies showing differential expression signatures associated with cellular transformation (Gupta et al., 2010; Kaneko et al., 2010; Sang et al., 2010). Furthermore, lncRNAs are often associated with various human diseases, such as Alzheimer's disease, heart diseases and cancer (Taft et al., 2010). Indeed, it has currently been demonstrated that lincRNAs (a class of lncRNAs) are an integral component of the p53 transcriptional regulatory network (Huarte et al., 2010).

One important class of lncRNAs is composed of enhancer-derived transcripts called eRNAs. The transcription of eRNAs is commonly correlated to nearby mRNA expression (De Santa et al., 2010; Kim et al., 2010). In addition, enhancer domains may contain binding sites for several transcription factors, orchestrating a cell-type specific transcriptional program. However, how these enhancer RNA producing domains are exploited to quickly respond to cellular insults is unknown. In addition it is unclear to what extent eRNAs carry a transcriptional regulatory function. Therefore the identification and further characterization of activated stress-dependent eRNAs is highly important.

In this Doctoral Thesis we hypothesize that the tumor suppressor p53 regulates enhancer RNA producing domains that are required for a proper and specific p53 transcriptional program. This supports recent observations suggesting that the p53 protein-coding target genes, which have been identified so far, cannot fully account for the tumor suppression activity of p53 (Brady et al., 2011). Moreover, we address the contribution of possible crosstalks between different lncRNA species, which might impact cellular and genetic stability. Thus, this study has

been divided into two related projects.

In the first project, entitled "eRNAs are required for p53-dependent enhancer activity and gene transcription", our main goal was to identify putative enhancer regions that are bound by p53, and may thus require p53 for their activity on target gene promoters. We provide evidence that enhancer domains are bound by p53. Using Nutlin-3a to induce p53 responsiveness, we show that eRNAs are produced from these domains, thereby naming them p53-bound enhancer regions (p53BERs). Further characterization of these regions demonstrated that p53 regulates target genes by producing eRNAs, and affecting the activity of enhancer domains.

In the second project, entitled "Genome wide profiling of p53-responsive enhancer RNAs uncovers an essential lncRNA-dependent epigenetic regulation", we set out to identify all p53-regulated enhancers (bound and unbound). Recent reports suggest an intricate regulatory network between different RNA species (Vance et al., 2014b; Yang et al., 2013). Thus, we further hypothesized on the existence of possible crosstalks between stress-dependent eRNAs and other lncRNAs. Using GRO-seq we profiled the presence of eRNA producing domains directly or indirectly regulated by p53. Furthermore, we identified a stress-activated lncRNA, named LED, which is responsible for the fine-tune regulation of a subset of these enhancers. Interestingly we observed that LED is inactive in human cancer.

Chapter two

eRNAs are required for p53-dependent enhancer activity and gene transcription

Carlos A. Melo[1,2,8], Jarno Drost[1,8], Patrick J. Wijchers[3], Harmen van de Werken[3], Elzo de Wit[3], Joachim A. F. Oude Vrielink[1], Ran Elkon[1], Sonia A. Melo[4], Nicolas Leveille[1], Raghu Kalluri[5,6], Wouter de Laat[3] and Reuven Agami[1,7,#]

[1]Division of Gene Regulation, The Netherlands Cancer Institute, Plesmanlaan 121, 1066 CX, Amsterdam, The Netherlands. [2]Doctoral Programme in Biomedicine and Experimental Biology, Centre for Neuroscience and Cell Biology. [3]Hubrecht Institute-KNAW and University Medical Centre Utrecht, Uppsalalaan 8, 3584CT Utrecht, The Netherlands. [4]Division of Matrix Biology, Department of Medicine, Beth Israel Deaconess Medical Center and Harvard Medical School, Boston, MA. [5]Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, MA. [6]Harvard-MIT Division of Health Sciences and Technology, Boston, MA. [7]Centre for Biomedical Genetics, The Netherlands.
[8]These authors contributed equally to this work
[#]Corresponding author: Reuven Agami, email: r.agami@nki.nl

Running title: p53-induced enhancer RNAs

Mol Cell 2013; 49(3):524-535

2.1 Summary

Binding within or nearby target genes involved in cell proliferation and survival enables the p53 tumor suppressor gene to regulate their transcription and cell cycle progression. Using genome-wide chromatin-binding profiles, we describe binding of p53 also to regions located distantly from any known p53 target gene. Interestingly, many of these regions possess conserved p53-binding sites and all known hallmarks of enhancer regions. We demonstrate that these p53-bound enhancer regions (p53BERs) indeed contain enhancer activity and interact intra-chromosomally with multiple neighboring genes to convey long-distance p53-dependent transcription regulation. Furthermore, p53BERs produce, in a p53-dependent manner, enhancer RNAs (eRNAs) that are required for efficient transcriptional enhancement of interacting target genes and induction of a p53-dependent cell cycle arrest. Thus, our results ascribe transcription enhancement activity to p53 with the capacity to regulate multiple genes from a single genomic binding site. Moreover, eRNA production from p53BERs is required for efficient p53 transcription enhancement.

2.2 Introduction

The p53 gene is the most frequently mutated gene in human cancer as mutations appear in more than 50% of tumors. In cases where tumors harbor the wild-type p53 gene it is believed that alterations in other p53 pathway components account for its inactivation (Vousden and Lu, 2002). Thus, the p53 tumor suppressor pathway plays a central role in tumor suppression (Christophorou et al., 2006). The p53 gene encodes a transcription factor that regulates transcription of target genes through sequence specific interaction with DNA. As most of the inactivating mutations found in cancer are centered in its sequence-specific DNA-binding domain, regulating transcription is thought to be its main function in tumor suppression.

As a transcription factor p53 can regulate transcription of many target genes. Its activation by different types of cellular stresses can elicit several distinct types of cellular responses, including cell cycle arrest and apoptosis. However, the recent observation that the thus-far known protein-coding p53 target genes cannot fully explain the tumor suppression activity of p53 (Brady et al., 2011) raises the possibility that non-coding genomic regions play an important role in the p53 pathway as well. Besides the regulation of protein-coding genes, p53 has already been shown to transcriptionally regulate several non-coding RNA (ncRNA) species. These include miRNAs and, more recently, also p53-regulated lincRNAs were identified (Guttman et al., 2009; Huarte et al., 2010). One of these lincRNAs, termed lincRNA-p21, is specifically required for p53-dependent repression of target genes and induction of apoptosis (Huarte et al., 2010). The observation that many of the mammalian ncRNAs are highly conserved and expressed in a regulated and tissue-specific manner reflects their importance.

Chromatin-binding profiles reveal specific interactions of p53 with promoter regions of nearby target genes, but also with distant locations that harbor characteristics of enhancer domains. Enhancers are non-protein-coding domains that regulate transcription of genes at long distances and are characterized by specific chromatin signatures of histone methylation and acetylation (ENCODE Project Consortium, 2007; Heintzman et al., 2007 and 2009). Recently, RNA polymerase II (RNAPII) was shown to bind to a subset of enhancers and produce bidirectional, non-polyadenylated transcripts defined as enhancer RNAs (eRNAs) (Kim et al., 2010; Wang et al., 2011; De Santa et al., 2010). The transcription of these eRNAs positively correlated with the mRNA levels of the surrounding protein-coding genes (Kim et al., 2010). However, it remains unclear whether eRNAs have any function in gene transcription.

Here we identified several regions that are bound by p53 and harbor all known chromatin features of eRNA-producing enhancer domains. We named these regions p53-bound enhancer regions (p53BERs). We demonstrate that the enhancer activity of p53BERs is dependent on functional p53 and that they interact with multiple surrounding genes. The vast majority of these genes indeed display p53-dependent expression whereas no significant p53 binding could be detected in their promoters. Furthermore, eRNAs are produced from p53BERs in a p53-dependent manner. We show that eRNAs are functional ncRNAs involved in transcription enhancement of genes interacting with the enhancer they are produced from. Thus, our results imply that p53 can regulate transcription of multiple distantly located genes through binding to enhancers and that eRNAs produced from these enhancers are involved in transcription enhancement.

2.3 Results

2.3.1 p53 binds to regions characterized with chromatin signatures of enhancer domains

The recent observation that the currently known p53 target genes do not fully account for p53-dependent tumor suppression (Brady et al., 2011) initiated our search for novel types of p53 target genes. Specifically, we searched for putative enhancer regions that are bound by p53 and thus may require p53 for their activity on target gene promoters. To identify such domains we made use of genome-wide p53-binding datasets that others and we have generated (Wei et al., 2006; Drost et al., 2010) and ranked p53-bound regions based on p53 binding intensity (indicated by its peak height; Table S1). In addition, we searched for regions that harbor histone modifications that have been shown to mark active enhancers (ENCODE Project Consortium, 2011). These include high levels of histone 3 lysine 4 monomethylation (H3K4me1), low levels of H3K4 trimethylation (H3K4me3) and high levels of H3K27 acetylation (H3K27Ac) (Heintzman et al., 2007; Heintzman et al., 2009; Creyghton et al., 2010). Several regions contained evident

markers of enhancer activity combined with p53-binding at levels equivalent or higher when compared to the p21 (CDKN1A) promoter, a prominent high affinity p53 target gene (Figure 1A, S1(A-B) and Table S1). We named our identified enhancers p53-Bound Enhancer Regions (p53BERs). To confirm histone modifications also in our cell model system we performed chromatin immunoprecipitations (ChIPs) for H3K4me1, H3K4me3 and H3K27Ac on one of our identified enhancers, p53BER2, in primary BJ fibroblasts immortalized by hTERT, the human telomerase reverse transcriptase (BJ/ET). Indeed, H3K4me1 was the predominant modification on p53BER2 in combination with high levels of H3K27Ac (marking active genomic regions (Creyghton et al., 2010)), whereas the promoter of the p53 target gene MDM2 harbored, as expected, predominantly H3K4me3 in combination with high levels of H3K27Ac (Figure 1B, C). These results confirm the enhancer signature of our identified p53BERs.

2.3.2 p53BERs contain p53-dependent enhancer activity

Next, we examined the potential transcriptional enhancing activity of the p53BERs using luciferase reporter assays in MCF7 cells (containing wild-type p53). We cloned 1.8–2kb regions, including the p53BER and additional flanking sequences, upstream of a promoter driving luciferase expression. Transfection of all reporters revealed significant transcriptional enhancing activity, ranging from about 4-fold to about 190-fold, whereas two control genomic regions (harboring no enhancer features) did not show any activity in this assay (Figure 1D and S1C). Figure 1D also shows that, as expected from an enhancer region, transcription activation was orientation independent (p53BER2 reverse orientation). Importantly, the enhancer activity of all p53BERs studied here was p53-dependent as co-transfection of a p53 knockdown vector ($p53^{KD}$) significantly inhibited enhancer activity (Figure 1D and S1C). To further confirm p53-dependency of the enhancer activity of p53BERs we introduced point mutations in the p53 consensus sequences within p53BER2 (Figure S1B). Transfection of this reporter revealed that mutating the p53 consensus sequence completely abolished enhancer activity of p53BER2 (Figure 1D). Altogether, we identified p53-bound regions that harbor features of enhancer domains and contain p53-dependent enhancer activity.

2.3.3 p53BERs intra-chromosomally interact with multiple distant genes to convey p53-dependent transcription

Enhancers control expression of one or multiple genes over distance and in an orientation independent manner. This happens through so-called DNA looping, which brings the enhancer in close proximity of its target promoters (Splinter et al., 2011). Thus, p53 could potentially regulate transcription of multiple target genes through binding to a single

enhancer. To identify possible target genes of p53-dependent enhancer domains, we applied chromosome conformation capture (4C) technology in combination with next generation sequencing (Dekker et al., 2002; Splinter et al., 2011) for three p53BERs (1, 2 and 4) in BJ/ET cells. We found that all three p53BERs appeared to interact with multiple neighboring genes on the same chromosome (Figure 2A–C). Similar results were obtained with two different restriction enzyme combinations (Figure S2A). Strikingly, mRNA sequencing of BJ/ET cells stably expressing either a control or p53 knockdown vector revealed that many of these genes require p53 for their proper expression (Figure 2A–C and S2B). To confirm that these genes are indeed regulated by p53 we treated cells with the MDM2-inhibitor Nutlin-3 and analyzed mRNA levels of one prominent neighboring gene for each p53BER. These were DUSP4 (dual specificity phosphatase 4; at ~430 kb), PAPPA (pregnancy-associated plasma protein A; at ~210 kb) and IER5 (immediate early response 5; at ~50 kb), for p53BER1, 2 and 4, respectively. As expected, mRNA levels of all genes were induced upon Nutlin-3 treatment (Figure 2D). Moreover, knocking down p53 decreased basal mRNA levels and interfered with the induction upon Nutlin-3 treatment (Figure 2D). These results imply that DUSP4, PAPPA, and IER5 are transcriptionally regulated by p53 though no significant p53-binding peak could be detected in their promoters by ChIP-seq in BJ/ET cells (Drost et al., 2010). Importantly, we found that the long-range interactions were still present in cells with stable p53 knockdown (Figure 2A–C), suggesting that p53 acts on pre-existing chromatin loops to activate p53BER-target genes.

Collectively, our observations strongly suggest that the identified p53-bound enhancer regions physically contact multiple distant genomic regions to convey p53-dependent transcription regulation.

2.3.4 p53BER2 and p53BER4 produce enhancer RNAs (eRNAs) in a p53-dependent manner

Recently, it was shown that RNA polymerase II (RNAPII) binds to a subset of enhancers where it bi-directionally transcribes so-called enhancer RNAs (eRNAs; Kim et al., 2010). eRNAs are non-coding, non-polyadenylated RNAs with sizes up to 2 kb. In addition to high levels of H3K4me1, low levels of H3K4me3 and high levels of H3K27Ac, eRNA-producing enhancer domains were marked by RNAPII and p300/CBP binding (Kim et al., 2010). We found association of RNAPII and p300 with p53BER2 and p53BER4 in different cell lines (ENCODE Project Consortium, 2011; Figure 3A) and confirmed this for RNAPII in BJ/ET cells by ChIP (Figure 3B). No RNAPII and p300 association was detected in control downstream regions. Furthermore, we found increased p53 binding on p53BER2 and 4 following p53 activation by Nutlin-3 treatment (Figure 3C) and reporter assays using constructs in which p53BER2 was cloned upstream of a promoterless luciferase gene revealed p53-dependent promoter activity (Figure S3A). Next, we asked whether p53 induces transcript production from p53BER2 and p53BER4.

Figure 1. Identification of p53-dependent enhancer domains

(A) UCSC Genome Browser (hg18 assembly) presentation of the p53-binding pattern (in BJ/ET cells expressing either a control or p53KD vector) (Drost et al., 2010) and histone modifications (ENCODE Project Consortium) at the p53BER2 (top) and p53BER4 (bottom) genomic regions. Different cell lines in the histone modification track are associated with a particular color. (B) BJ/ET cells were subjected to ChIP for total H3, H3K4me1 and H3K4me3. The abundance of the histone modifications within the indicated genomic regions was calculated as the ratio between the H3K4me1 and H3K4me3, both corrected to total H3. MDM2 was used as promoter control. Graphs show means and standard deviation (s.d.) for three independent experiments.(C) Cells from (B) were subjected to ChIP with antibodies against total H3 and H3K27Ac. H3K27Ac abundance within the indicated genomic regions was calculated as the ratio between H3K27Ac and total H3. MDM2 was used as promoter control. Region downstream of p53BER2 was used as negative control. Graphs show means and s.d. for three independent experiments. (D) MCF7 cells were co-transfected with the indicated reporter construct and either control or p53 knockdown (p53KD) vector. The relative firefly luciferase/renilla activity was determined and compared to the control promoter vector (Ctrl.). Graphs represent mean and s.d. from three independent experiments. P values were calculated using t-test. Mutant (mut) contains point mutations in the p53-binding motifs of p53BER2.

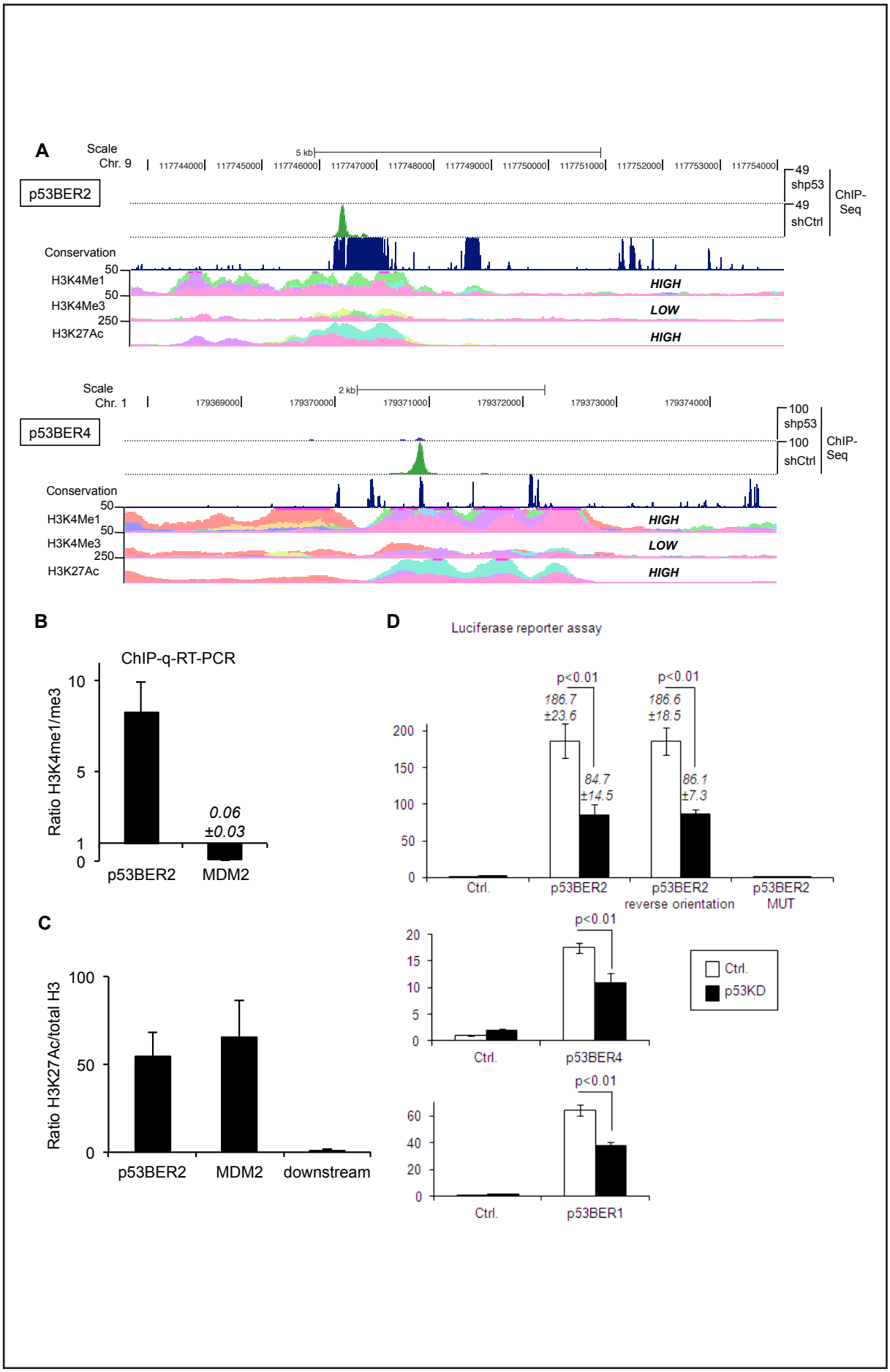
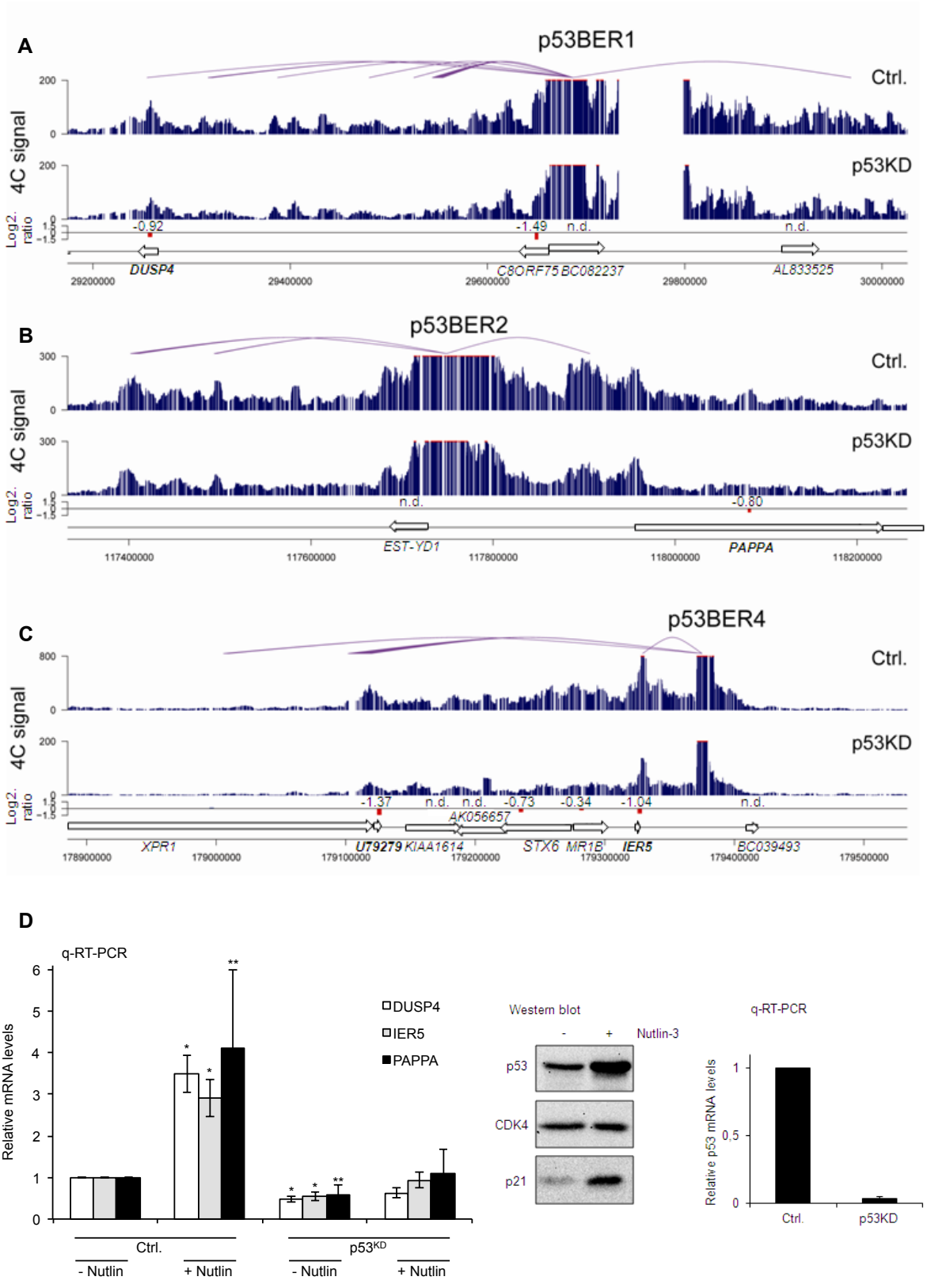

A
Scale
Chr. 9
5 kb
117744000
117745000
117746000
117747000
117748000
117749000
117750000
117751000
117752000
117753000
117754000
p53BER2
49
shp53
49
shCtrl
ChIP-Seq
Conservation
50
H3K4Me1
HIGH
50
H3K4Me3
LOW
250
H3K27Ac
HIGH
Scale
Chr. 1
2 kb
179369000
179370000
179371000
179372000
179373000
179374000
p53BER4
100
shp53
100
shCtrl
ChIP-Seq
Conservation
50
H3K4Me1
HIGH
50
H3K4Me3
LOW
250
H3K27Ac
HIGH
B
ChIP-q-RT-PCR
Ratio H3K4me1/me3
10
5
1
0
0.06
±0.03
p53BER2
MDM2
C
Ratio H3K27Ac/total H3
100
50
0
p53BER2
MDM2
downstream
D
Luciferase reporter assay
p<0.01
186.7
±23.6
84.7
±14.5
p<0.01
186.6
±18.5
86.1
±7.3
200
150
100
50
0
Ctrl.
p53BER2
p53BER2
reverse orientation
p53BER2
MUT
p<0.01
20
15
10
5
0
Ctrl.
p53KD
Ctrl.
p53BER4
p<0.01
60
40
20
0
Ctrl.
p53BER1

Figure 2. p53BERs show long-range interactions with multiple distantly located genes
(A) – (C) Long-range interactions of p53BER1, 2 and 4 with surrounding regions as determined by 4C in wild type and p53KD BJ/ET cells. Plots represent running mean analysis in which 4C sequencing read counts were averaged over windows of 31 restriction ends. To account for the fact that the majority of the data is very close to the viewpoint the data range of the vertical axis we set to the 98% quantile value for the analyzed region (fragments with values above the 98% quantile value are marked red on top). Arcs designate significant interactions (P<10-5, see also Figure S4) of the p53BERs with surrounding loci. Relative changes in mRNA expression upon p53KD were measured by 3'-end mRNA sequencing in BJ/ET cells and is represented at the bottom (red bars and numbers). N.d., not detectable. Note that significant interactions are largely unaffected in p53KD BJ/ET cells. (D) BJ/ET cells expressing either a control or p53 knockdown vector were treated with Nutlin-3 for 24 h or left untreated. Left and right panel: relative DUSP4, IER5, PAPPA and p53 mRNA levels were measured by q-RT-PCR and are represented as fold induction from values measured in untreated control cells. Graphs show means and s.d. for three independent experiments. P values were calculated using t-test. * p value < 0.01, ** p value < 0.05. Middle panel: western blot analysis was performed to detect p53 stabilization and p21 induction upon Nutlin-3 treatment. CDK4 was used as loading control.


A
p53BER1
Ctrl.
p53KD
4C signal
Log2. ratio
-0.92
-1.49
n.d.
n.d.
DUSP4
C8ORF75 BC082237
AL833525
29200000
29400000
29600000
29800000
30000000
B
p53BER2
Ctrl.
p53KD
4C signal
Log2. ratio
n.d.
-0.80
EST-YD1
PAPPA
117400000
117600000
117800000
118000000
118200000
C
p53BER4
Ctrl.
p53KD
4C signal
Log2. ratio
-1.37
n.d.
n.d.
-0.73
-0.34
-1.04
n.d.
AK056657
XPR1
U79279 KIAA1614
STX6
MR1B
IER5
BC039493
178900000
179000000
179100000
179200000
179300000
179400000
179500000
D
q-RT-PCR
Relative mRNA levels
DUSP4
IER5
PAPPA
Ctrl.
p53KD
- Nutlin
+ Nutlin
- Nutlin
+ Nutlin
Western blot
- + Nutlin-3
p53
CDK4
p21
q-RT-PCR
Relative p53 mRNA levels
Ctrl.
p53KD

We activated p53 in BJ/ET cells with Nutlin-3 and performed quantitative reverse transcription polymerase chain reaction (q)-RT-PCR analysis, which revealed transcript induction at both loci upon Nutlin-3 treatment (Figure 3D). Similarly, in MCF7 cells too, transcript production could be induced from p53BER2 and 4 upon ionizing radiation or Nutlin-3 treatment (Figure S3(B-C)). In contrast, no transcript was detected from these enhancer regions either in cells knocked down for p53, or in Nutlin-3-treated MDA-MB-436 cells, which contain mutant p53 (Figure S3(D-F)), further confirming p53-dependency. Northern blot analysis detected Nutlin-3-induced production of a transcript at p53BER2 with a size of about 600 nucleotides (Figure 3E). In addition, we were able to identify the 5' start site of the p53BER2-produced transcript (Figure S3G). Last, transcripts produced from p53BER2 and p53BER4 were reproducibly detected upon p53 activation by q-RT-PCR on random-primed cDNA, whereas q-RT-PCRs on oligo(dT)-primed cDNA failed to do so, implying that the detected transcripts are likely to be non-polyadenylated. This is in accordance with the recent observation that eRNAs lack poly(A) tails (Kim et al., 2010). In conclusion, our observations indicate that p53BER2 and p53BER4 produce eRNAs in a p53-dependent manner.

2.3.5 Enhancer RNAs (eRNAs) are functional non-coding RNAs that enhance transcription

The production of eRNAs at enhancers has been shown to correlate with transcription of neighboring genes, but their functional involvement is unclear (Kim et al., 2010). Therefore, we set out to investigate if eRNAs are involved in enhancement of transcription. We cloned three regions surrounding the p53 binding site in p53BER2 into a vector containing 24 copies of the MS2 hairpin RNA structure to generate a chimera RNA with the 24 copies of MS2 (Figure 4A). The MS2 hairpin binds tightly to the MS2-coat viral protein (MS2-CP; Haim-Vilmovsky et al., 2011). We attached the Gal4 DNA-binding domain to MS2-CP and designed a luciferase reporter vector containing 5x Gal4 binding sites upstream of a promoter. If the cloned region produces RNA that is functionally involved in transcription enhancement, co-transfection of all components (eRNA-MS2, Gal4-MS2-CP and the luciferase reporter) would localize the RNA nearby the promoter and is expected to enhance reporter activity (Figure 4A). Intriguingly, co-transfection of region 1, but not 2 or 3, significantly enhanced reporter activity (Figure 4B). We performed MS2-CP-pull-down assays to verify transcript production from region 1 and to exclude plasmid contamination in the (semi-) q-RT-PCR (Figure 4C). Since p53BERs contain p53-dependent promoter activity (Figure S3A), we tested whether p53 binding is required for enhancing activity in this assay. Figure 4C shows that mutating all p53 sites in p53BER2 markedly reduced transcript levels, as determined by MS2-CP pull-down assays, supporting p53-dependent eRNA production from this region. Moreover, transfection of this reporter did not result in any transcription enhancement activity by region 1 (Figure 4D), showing that the

enhancing effect is dependent on p53 binding and production of the transcript. Furthermore, removing the MS2-CP tethering component significantly reduced the enhancing effect of region 1 on the reporter, indicating that transcript localization is involved in the observed effect (Figure 4E). The remaining MS2-CP-independent transcriptional activation can be attributed to the over-expression of the eRNA-MS2 in the nucleus.

Altogether, these results demonstrate for the first time a role for eRNAs in transcription enhancement.

2.3.6 eRNAs produced from p53BERs are required for transcription enhancement of neighboring genes and for an efficient p53-dependent cell cycle arrest

As the identified p53BERs interact with and regulate transcription of neighboring genes at distant locations, we set out to investigate if the eRNAs produced from p53BER2 and p53BER4 are involved in p53-dependent transcriptional regulation of genes encoded in their vicinity. We designed siRNAs to target the eRNAs that are produced from these regions. Transfection of MCF7 cells with si-p53BER2 and si-p53BER4, but not control siRNAs, reduced corresponding eRNA transcript levels following p53 activation by Nutlin-3 (Figure S4). The genes nearest to p53BER2 and p53BER4 are PAPPA and IER5, respectively. We showed in Figure 2 and S2B that the PAPPA and IER5 genes are transcribed in a p53-dependent manner and that they physically interact with the nearest p53BER domains. Moreover, no p53-binding peak was detected in the promoters of PAPPA and IER5 by ChIP-seq in BJ/ET cells (Drost et al., 2010) that can explain such p53-dependent regulation. Therefore, we examined the effect of si-p53BER2 and si-p53BER4 on the transcription induction of PAPPA and IER5, respectively, following p53 activation. We transfected cells with a negative control siRNA, two independent siRNAs targeting p53BER transcripts, or a siRNA targeting p53, and determined the relative induction of PAPPA and IER5 mRNA levels upon Nutlin-3 treatment. As expected, transcription of PAPPA and IER5 was activated upon Nutlin-3 treatment (Figure 5A, B). This induction was significantly decreased when p53 was knocked down, again confirming that transcription is p53-dependent. Intriguingly, transfection of either of two independent siRNAs targeting the eRNAs produced from p53BER2 and p53BER4 significantly inhibited the induction of PAPPA and IER5, respectively, upon Nutlin-3 treatment (Figure 5A, B). This effect was further confirmed by RNAPII ChIP, showing that significantly less RNAPII is present at the PAPPA and IER5 genes upon eRNA knockdown (Figure S5A). Thus, we conclude that eRNAs produced from p53BER2 and p53BER4 are required for the efficient induction of PAPPA and IER5 transcription, respectively, following p53 activation.

To determine if other p53BER-interacting genes also show eRNA-dependent transcription, we investigated the levels of the p53BER4 interacting gene U79279. RNA-seq showed marked

Figure 3. p53BER2 and p53BER4 carry features of eRNA-producing enhancer domains and produce transcripts in a p53-dependent manner

(A) UCSC Genome Browser (hg18 assembly) presentation showing RNAPII and p300 binding at p53BER2 (top) and p53BER4 (bottom) (ENCODE Project Consortium). (B) BJ/ET cells were cultured in the presence of Nutlin-3 for 24 h and subsequently subjected to ChIP for RNAPII. Rabbit IgG was used as negative control. Protein binding to the indicated genomic regions was quantified by calculating the percentage of input that is chromatin-bound. Graphs show means and s.d. for three independent experiments. (C) BJ/ET cells were cultured in the absence or presence of Nutlin-3 for 24 h and subsequently subjected to ChIP for p53. Protein binding to the indicated genomic regions was quantified by calculating the percentage of input that is chromatin-bound and is represented as fold induction from values measured in untreated cells. A non p53-bound region was used as negative control. Graphs show means and s.d. for three independent experiments. (D) BJ/ET cells were treated with Nutlin-3 for 24 h or left untreated. The relative transcript levels produced at p53BER2 and p53BER4 were determined by q-RT-PCR in relation to GAPDH. Graphs show mean and s.d. for three independent experiments. (E) MCF7 cells were treated with Nutlin-3 for 24 h or left untreated. Northern blot analysis was performed to detect transcripts produced at p53BER2. tRNA-Met was used as loading control.

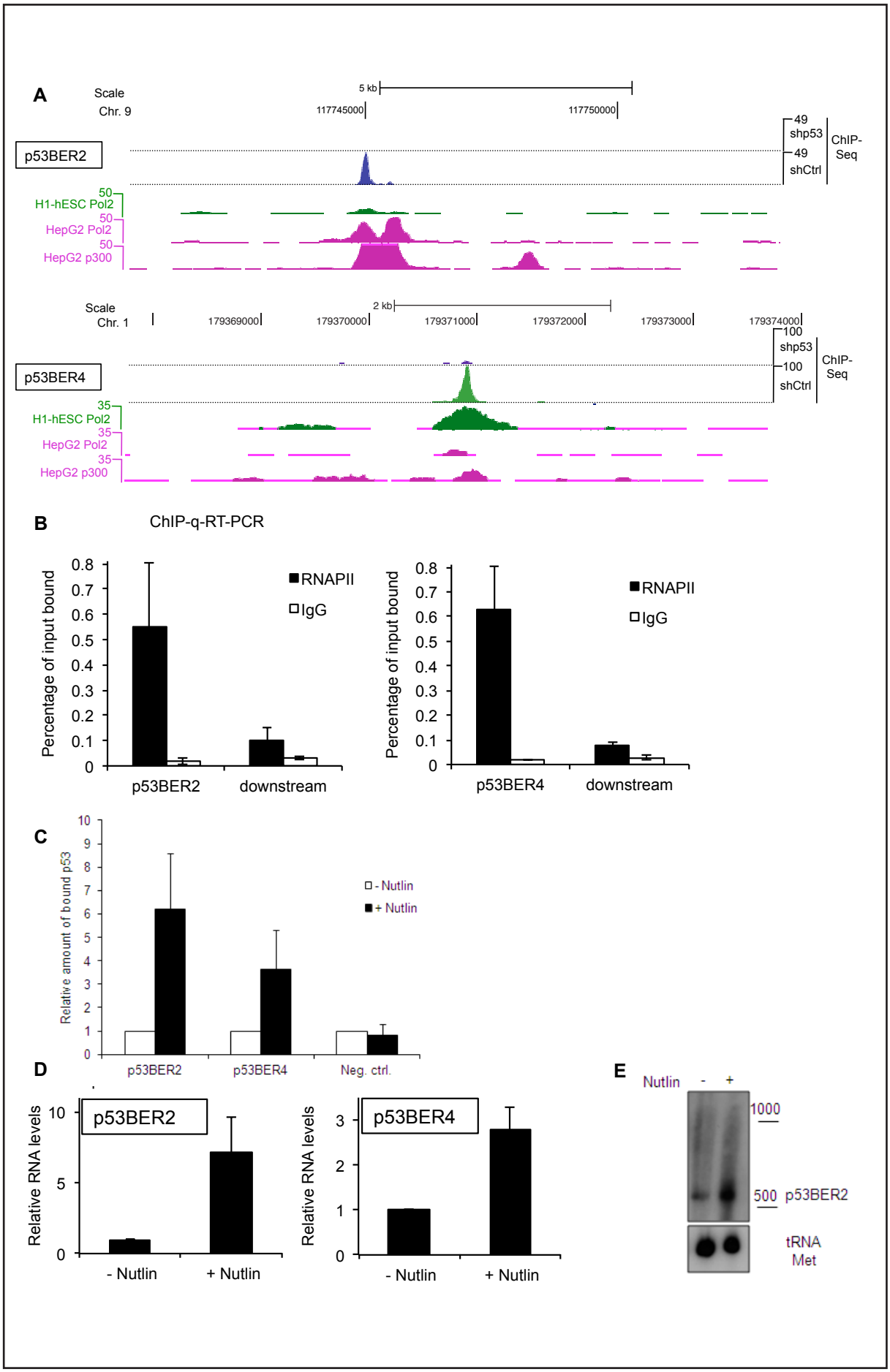
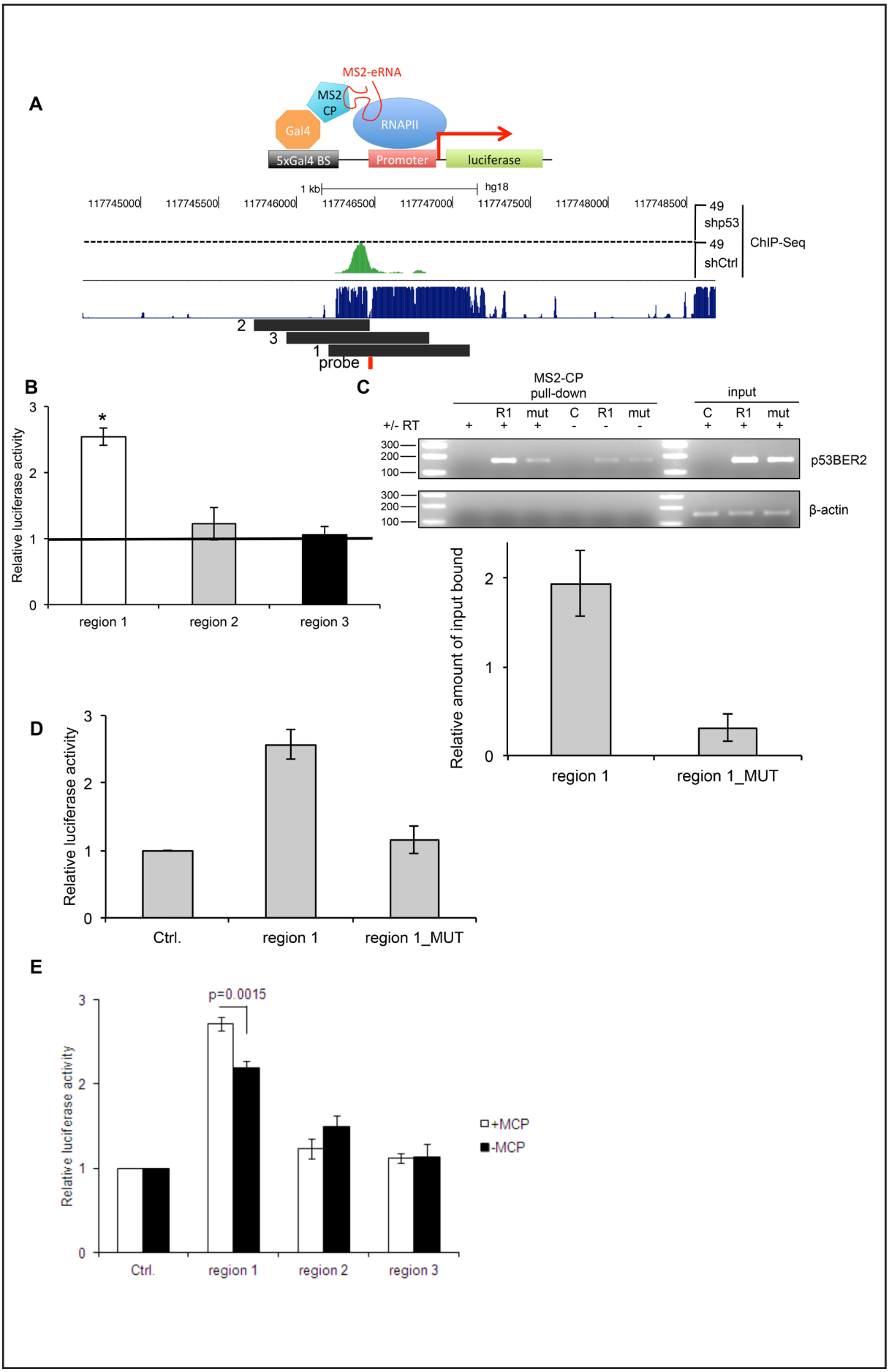

A
Scale
Chr. 9
5 kb
117745000
117750000
p53BER2
49 shp53
49 shCtrl
ChIP-Seq
50 H1-hESC Pol2
50 HepG2 Pol2
50 HepG2 p300
Scale
Chr. 1
2 kb
179369000
179370000
179371000
179372000
179373000
179374000
p53BER4
100 shp53
100 shCtrl
ChIP-Seq
35 H1-hESC Pol2
35 HepG2 Pol2
35 HepG2 p300
B
ChIP-q-RT-PCR
Percentage of input bound
RNAPII
IgG
p53BER2
downstream
p53BER4
downstream
C
Relative amount of bound p53
- Nutlin
+ Nutlin
p53BER2
p53BER4
Neg. ctrl.
D
p53BER2
p53BER4
Relative RNA levels
- Nutlin
+ Nutlin
E
Nutlin - +
1000
500
p53BER2
tRNA Met

Figure 4. eRNAs are involved in enhancement of transcription

(A) Top: schematic representation of the tethering system. Bottom: location of the tested regions and the Northern probe (red bar) that was used in Figure 3D. (B) MCF7 cells were co-transfected with a pSuper-24MS2 vector containing the indicated p53BER2 regions, and the MS2-CP:Gal4 and reporter constructs (as described in panel A). The relative firefly luciferase/renilla activity was determined compared to the control promoter vector (set to 1). Graphs represent mean and s.d. from three independent experiments. P values were calculated using t-test. * P value < 0.01. (C) MCF7 cells were transfected with empty pSuper-24MS2 vector (C), pSuper-region 1-24MS2 (R1), or pSuper-region 1_MUT-24MS2. MS2-CP-pull-down assays followed by RT-PCR with p53BER2 primers were performed to verify transcript production. Expected PCR product size 166 bp. Bottom graph: transcript levels produced from region 1 and region 1_MUT were determined using MS2-CP-pull-down assays followed by RT-PCR. Graphs show means and s.d. for three independent experiments. (D) MCF7 cells were co-transfected as in (B) with either wild-type region 1 or region 1 containing point mutations in the p53-binding motifs (MUT). The relative firefly luciferase/renilla activity was determined and compared to the control promoter vector (Ctrl.). Graphs represent mean and s.d. from three independent experiments. (E) MCF7 cells were co-transfected with the indicated p53BER2 regions, the reporter construct and either MS2-CP:Gal4 or empty vector. Error bars represent standard-error of the means (SEM; n=7). P values were calculated using t-test.


A
MS2-eRNA
MS2 CP
Gal4
RNAPII
5xGal4 BS
Promoter
luciferase
1 kb
hg18
117745000
117745500
117746000
117746500
117747000
117747500
117748000
117748500
49 shp53
49 shCtrl
ChIP-Seq
2
3
1
probe
B
Relative luciferase activity
region 1
region 2
region 3
C
MS2-CP pull-down
input
+/- RT
R1
mut
C
p53BER2
β-actin
Relative amount of input bound
region 1
region 1_MUT
D
Relative luciferase activity
Ctrl.
region 1
region 1_MUT
E
p=0.0015
Relative luciferase activity
+MCP
-MCP
Ctrl.
region 1
region 2
region 3

p53-dependent transcription of this gene (Figure 2C). Once again, transfection of siRNAs targeting p53BER4-produced eRNAs resulted in a significantly decrease in U79279 transcript induction, further supporting a role for eRNAs in enhancement of transcription (Figure S5B).

Next, we set out to investigate if eRNA production is required for p53 function following stress induction. We transfected MCF7 cells with control siRNAs, siRNAs targeting the p53BER2 eRNA, or p53. After treatment with ionizing radiation (IR), cell cycle profiles were analyzed by flow cytometry. As expected, while control transfected MCF7 cells entered a G1 arrest upon IR, knocking down either p53 or p21 did not (Figure S6A), confirming that the arrest is indeed dependent on a functional p53 pathway. Interestingly, targeting p53BER2-produced transcripts (two distinct siRNAs) resulted in a significant decrease in the amount of cells arrested in G1 (Figure S6A, B).

Furthermore, similar to the knockdown of p53BER2-produced eRNAs, knocking down PAPPA also significantly inhibited induction of a p53-dependent cell cycle arrest (Figure 5C and S6C). Neither knockdown of the p53BER2-produced transcript nor knockdown of PAPPA significantly affected p21 mRNA levels (Figure S6D). Altogether, these results indicate that the eRNAs produced from p53BER2 are required for an efficient p53-dependent cell cycle arrest, and that this effect could be mediated through inhibition of PAPPA induction by p53.

2.4 Discussion

We describe here a mechanism whereby p53 induces target gene transcription through the production of eRNAs from distantly located enhancer domains. We identified several such enhancer regions (p53BERs) that are bound by p53 with high affinity, require p53 for their activity and produce eRNAs in a p53-dependent manner. Recently, enhancer-like features were also ascribed to a different class of ncRNAs (Ørom et al., 2010). Knockdown of these ncRNAs was demonstrated to result in decreased expression of neighboring genes. In contrast to eRNAs (Kim et al., 2010), these ncRNAs are polyadenylated and the regions encoding them harbor the chromatin signature of protein-coding genes (high H3K4me3 at 5´-end and downstream H3K36me3 (Ørom et al., 2010)). We report that the suppression of p53-induced eRNAs results in altered p53-dependent transcription of genes encoded in the vicinity of the p53BER and that eRNA expression enhances transcriptional activity in an RNA tethering reporter assay. Therefore, our results demonstrate for the first time the involvement of eRNAs in enhancement of target gene transcription, implying they are not merely by-products of gene transcription. This classifies eRNAs as a novel family of regulatory RNAs that enhance gene transcription.

In addition, we show that our identified p53BERs interact with multiple distant protein-coding genes and we link them to the p53 network present in our cellular model system. As

enhancers can interact and regulate multiple target genes (recently reviewed in Splinter et al., 2011 and Ong et al., 2011), enhancer-bound p53 would be able to regulate transcription of multiple genes via one single genomic binding site. We propose that enhancer-bound p53 can enhance transcription of distant target genes by getting in close proximity of target gene promoters via intra-chromosomal interactions. Transcription of target genes is enhanced upon p53 activation, which requires p53-dependent eRNA production (Figure 6). As the long-range interactions are not dependent on p53, the eRNAs seem to act on pre-existing chromatin conformations, which may allow for rapid induction of p53 target genes in response to diverse stimuli.

One of the genes regulated by p53BER2 is the PAPPA gene. PAPPA is an IGFBP-protease that cleaves IGFBP4 and IGFBP5 (Lawrence et al., 1999; Conover et al., 2004), which have both shown to be involved in cell cycle regulation and senescence (Kojima et al., 2012). Therefore, the involvement of PAPPA in cellular proliferation might be explained by its ability to regulate IGF levels. A recent publication reported p53-binding in intron 1 of the PAPPA gene by overexpressing p53 in the p53-negative cell line MDA-MB157 cells, which results in PAPPA repression (Chander et al., 2011). In contrast to this result, we did not detect any p53 binding within the first intron of PAPPA using p53 ChIP-seq in BJs and other cell lines (Figure S6E for BJ/ET cells). Furthermore, we report positive regulation of PAPPA transcription by p53 upon RAS^{V12} expression, Nutlin-3 treatment and ionizing radiation (Table S2). These results imply cell type specific regulation of PAPPA by p53. Enhancers regulate transcription in a cell type specific manner (Ong and Corces, 2012), which could thus explain the difference in PAPPA regulation in different cell types.

Li et al. (2012) recently showed that in mouse ES cells p53 represses transcription of ES cell-specific genes through binding and interfering with distal enhancers (Li et al., 2012). It has become clear in recent years that enhancers are marked by different histone modifications in different cell types, reflecting their cell type specificity (Ong et al., 2012). This means that p53 can, by binding to enhancers and regulating their activity, affect target gene transcription in a cell type-dependent manner. Further research is needed to determine what other factors determine whether p53 is required for, or interferes with enhancer activity.

Moreover, an enhancer can interact with and regulate multiple genes. We show that eRNA knockdown interferes with transcription enhancement. Inhibiting eRNAs produced from a single enhancer could thus affect transcription of multiple genes. Therefore, eRNA inhibition could be an efficient strategy to interfere with enhancer activity. In addition, deletion of a single enhancer locus could result in the aberrant regulation of a set of genes. The genomic region including p53BER2 is located at 9q33.1. This region is associated with loss of heterozygosity (LOH) in human cancer (Schultz et al., 1995; Callahan et al., 2003). Based on the data we present in this report, it would be interesting to determine whether loss of p53BERs contributes to human cancer.

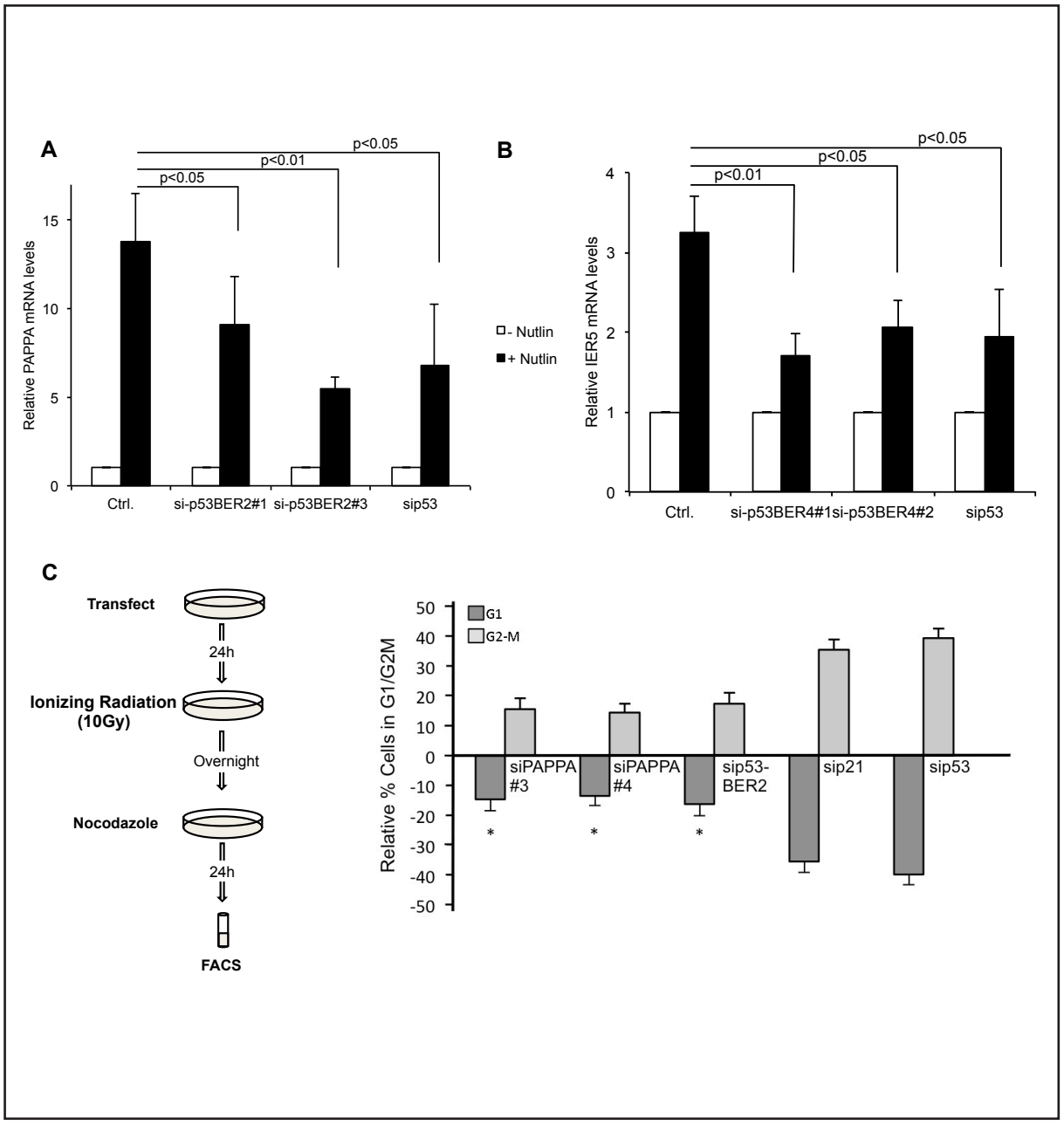


Figure 5. p53-dependent eRNAs are required for transcription of genes surrounding p53BER2 and p53BER4

(A) BJ/ET and (B) MCF7 cells were transfected with the indicated siRNAs and either treated with Nutlin-3 for 24 h or left untreated. Relative PAPPA (A) and IER5 (B) mRNA levels were measured by q-RT-PCR and is represented as fold induction from values measured in untreated cells. Graphs represent mean and s.d. from at least three independent experiments. P values were calculated using t-test. (C) Left is a schematic representation of the time line used for cell cycle analysis. At the right side the cell-cycle-distribution changes of MCF7 cells transfected with control siRNA or siRNAs against p53BER2 (sip53BER2) and PAPPA (siPAPPA#3, #4) treated with ionizing radiation (10 Gy). To capture cycling cells in G2/M, cells were incubated 24 h with nocodazole before flow cytometric analysis. For each condition, the difference between the control transfected and indicated siRNAs was calculated and plotted as the relative variation of cells present in both G1 and G2/M phase. Graphs represent mean and s.d. from three independent experiments. P values were calculated using t-test. * p value < 0.01.

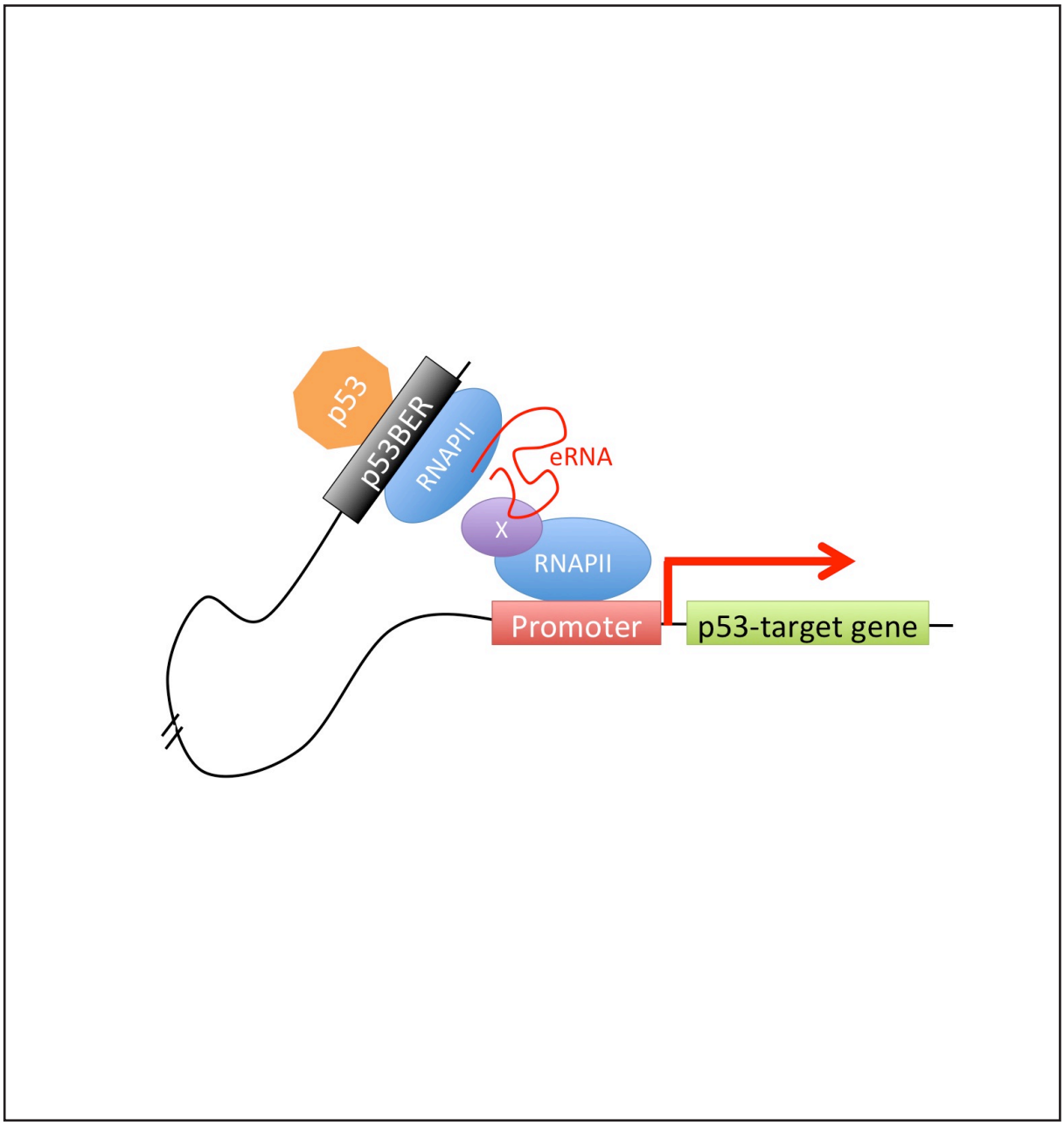


Figure 6. Schematic model depicting how p53BERs may enhance target gene transcription. p53-independent intra-chromosomal interactions bring p53-bound p53BERs in close proximity of target genes, through yet unidentified protein complex (marked by ?). It seems likely that p53 is bound at p53BERs in a poised state and that upon p53 activation eRNAs are produced and transcription enhancement takes place. eRNAs are produced in a p53-dependent manner and affect transcription enhancement by a currently unknown mechanism. Possibly, eRNAs interact with proteins (X) that activate transcription.

2.5 Experimental Procedures

Cell culture, transfection and retroviral transduction

Cells were cultured in DMEM medium with 10% FCS and antibiotics. Transfection of MCF7 cells for reporter assays was performed with polyethylenimine. SiRNA transfections of BJ/ET and MCF7 cells were performed using Dharmafect 1 and 3 (Dharmacon), respectively. SiRNA sequences were as follows: p53BER2#1 GCACAGATTCCGTGTAAAT; p53BER2#3 GCTGGACACTGGGTAAATC; p53BER4#1 TTAGGGGAAGCTGCTATGT; p53BER4#2 CAGCCTTGTGGTTTCACAG; sip21 GACCATGTGGACCTGTCAC; sip53 GACTCCAGTGGTAATCTAC. SiRNAs targeting PAPPA were purchased from Dharmacon. Retroviruses were made by calcium phosphate transfection of Ecopack 2 cells (Clontech) and harvesting 40 and 64 h later.

Constructs and antibodies

pRetrosuper (pRS), pRS-blast, pRS-hygro, pBabe-puro-Ras^{V12}ER^{TAM}, pMSCV-blast-Ras^{V12}ER^{TAM} and pBabe-H2B-GFP-hTERT were described previously (Voorhoeve et al., 2003; Brummelkamp et al., 2002; Kolfschoten et al., 2005). pRS-p53 was previously described (Voorhoeve et al., 2003). To monitor enhancer activity the sequences of interest were cloned in pGL3-promoter luciferase reporter vectors (Promega). Primer sequences are shown in Supplementary Information, Table S3. Mutations in the p53-binding sequences of p53BER2 were made in two steps using Quickchange Site-Directed mutagenesis kit (Agilent) according to manufacturer's instructions (primer 1, 5'-CTTCTGAGAAACTCATGGAGATTTCTGTGCATGCCTGAAC-3'; primer 2, 5'-ATGTCTGTGCATGCCTGAAATTCTCTGACAAAGCCAAGCA-3').

For ChIP antibodies against the following proteins were used: H3 (Cell Signaling Technology), H3K4me1 (Abcam, ab8895), H3K4me3 (Upstate, clone MC315), H3K27Ac (Abcam, ab4729) and RNAPII (Upstate, clone CTD4H8).

For Western blotting antibodies against p53 (DO1), p21 (F5) and CDK4 (C22) were purchased from Santa Cruz Biotechnology. Anti-beta-actin antibody was from Abcam (ab8227).

Encode Consortium histone modification and protein binding data

For histone modifications we used the ENCODE Promoter-Associated Histone Mark dataset (H3K4me1, H3K4me3, H3K27Ac) on 9 Cell Lines (Gm12878, H1 ES, HepG2, HMEC, HSMM, HUVEC, K562, NHEK, NHLF). For RNAPII and p300, we used the Histone Modifications by ChIP-seq from ENCODE/Broad Institute dataset from HepG2 and H1 ES cells. Presence of histone modifications and protein binding were verified by ChIP-Q-RT-PCR in BJ/ET cells.

Quantitative RT–PCR, 3'-end mRNA sequencing and Northern blot

Total RNA was extracted with Trizol (Life Technologies) according to the manufacturer's instructions. cDNA was transcribed with Superscript III (Life Technologies) using random hexamers (for eRNAs) or oligo(dT) (for mRNAs) in accordance with the manufacturer's protocol. To determine RNA expression levels, primers were designed with Primer Express v. 3.0 Software. Real-time PCR was performed with SYBR Green PCR master mix (Applied Biosystems) and Chromo 4 Real Time PCR Detector (Bio-Rad Laboratories). Primer sequences are shown in Supplementary Information, Table S4.

3'-mRNA sequencing was performed according to the manufacturer's instructions (Illumina mRNA-seq). To sequence from the 3'-end, first strand cDNA synthesis was performed using oligo(dT) instead of random primers.

For Northern blot, 30 µg of total RNA (DNase treated) was run on a gel composed of 15 % polyacrylamide (19:1), 48% urea and 1x TBE. RNA was transferred to BrightStar-Plus Positively Charged Nylon Membranes (Ambion) at 200 mA for 1 h in 0.5x TBE. Subsequently, RNA was cross-linked to the membrane (1200 J for 1 min). Pre-hybridization was performed in ULTRAhyb®-Oligo solution (Ambion) for 1 h at 42°C. Hybridization was performed in ULTRAhyb®-Oligo solution (Ambion) containing 150 ng of biotin-tagged oligo (BioTeg; Sigma) by overnight incubation at 42°C. Next, membranes were washed in 2x SSPE/0.5% SDS (2x 15 min), 0.2x SSPE/0.5% SDS (2x 30 min) and 2x SSPE (1x 5 min). Signal detection was performed using BrightStar BioDetect Kit according to the manufacturer's instructions (Ambion). p53BER2 probe sequence: 5'-CCCCACTTTCCACTGGGTCC-3'.

p53BER2 5'-end identification

Total RNA from Nutlin-3 treated BJ/ET cells was extracted with Trizol (Life Technologies) according to the manufacturer's instructions. Total RNA was used for first strand cDNA was synthesis (SuperScript III reverse transcriptase; Life Technologies) using a p53BER2-specific RT primer (5'- CCCCACTTTCCACTGGGTCC-3'). Single-strand cDNA was converted to double-strand cDNA (Life Technologies) and end repaired for 30 min at 20°C (NEBNext End Repair Enzym Module; New England Biolabs; supplemented with 1 µl Klenow DNA polymerase; Life Technologies). After purification (QIAquick PCR Purification Kit; QIAGEN), 1 µl 45 µM annealed P5-splinkerette (5'-AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTT CCGATCT-3') was ligated overnight at 16°C to the cDNA. The P5-splinkerette-ligated cDNA was purified (QIAquick PCR Purification Kit; QIAGEN) and used as template for PCR with P5 primer (5'-AATGATACGGCGACCACCGAGATCT-3') and p53BER2-specific nested primers (nested_1 5'-CCCAGTGGAAAGTGGGGGAG-3'; nested_2 5'-GTGGAAAGTGGGGGAGTGGTG-3') using Phusion Hot Start DNA polymerase (Finnzymes). After gel purification (QIAgen Gel Extraction Kit; QIAGEN) the DNA was cloned into pCR-

Blunt (Zero Blunt PCR Cloning kit; Life Technologies) and transformed into competent bacteria. Plasmid DNA was purified and sequenced.

Luciferase reporter assays

MCF7 cells were co-transfected with 400 ng of pRS or pRS-p53^{KD} in combination with 100 ng luciferase reporter (pGL3-promoter; Promega) and 5 ng renilla plasmid. Three days after transfection, luciferase assays were performed in accordance with the manufacturer's instructions (Dual Luciferase system; Promega). For RNA tethering, MCF7 cells were co-transfected with chimeras of different regions around p53BER2 and 24 copies of MS2 in pSUPER (Brummelkamp et al., 2002), pCS2-empty or pCS2-MCP:GAL4, and pGL3 containing 5 copies of the GAL4 binding site in front of the SV40 promoter. Three days after transfection, luciferase assays were performed in accordance with the manufacturer's instructions (Dual Luciferase system; Promega).

Chromatin immunoprecipitation assay

ChIP assays were performed as described (Dahl et al., 2008) with some minor modifications. In brief, cells were cross-linked for 8 min with 1% formaldehyde and neutralized with 125 mM glycine. After centrifugation, the pellet was resuspended in lysis buffer (50 mM Tris pH 8.0, 10 mM EDTA, 1% SDS) and sonicated to achieve an average fragment size of 300 – 800 bp. Chromatin was diluted in RIPA buffer (10 mM Tris pH 7.5, 140 mM NaCl, 1 mM EDTA, 0.5 mM EGTA, 1% Triton X-100, 0.1 % SDS, 0.1 % sodium deoxycholate) and incubated overnight with the indicated antibody. A negative control for immunoprecipitation was performed in the presence of normal rabbit IgG (Santa Cruz). The next day beads were added and incubated for 2 – 3 h. After washing, DNA was eluted overnight in elution buffer (20 mM Tris pH 7.5, 5 mM EDTA, 50 mM NaCl, 1% SDS, 50 µg/ml proteinase K) followed by DNA purification using PCR purification kit (Roche). Real-time PCR was performed with SYBR Green PCR master mix (Applied Biosystems) and Chromo 4 Real Time PCR Detector (Bio-Rad Laboratories). Primers sequences are shown in Supplementary Information, Table S5.

MS2-CP pull-down assay

MS2-CP pull-down assays were performed as described (Slobodin and Gerst, 2011). Briefly, MCF7 cells were transfected with pSUPER containing region 1 or region 1_MUT (p53 binding site mutant) attached 24 copies of MS2. After 72 h, cells were cross-linked for 10 min with 0.01% formaldehyde and neutralized with 125 mM glycine. Cells were snap-frozen in liquid nitrogen and thawed in lysis buffer (50 mM Tris pH 7.5, 150 mM NaCl, 1% Triton X-100, 1.8 mM MgCl2, protease inhibitor cocktail (Roche)) with occasional vortexing and

subsequently sonicated. After centrifugation, supernatant was blocked using avidin solution (Sigma; 10 µg per 1 mg of protein extract). Streptavidin-conjugated beads (GE Healthcare) were pre-coupled to MS2-CP-SBP protein and blocked in lysis buffer containing 2% BSA and yeast tRNA (0.1 mg per 100 µl of beads). The blocked beads were added to the blocked protein extract and incubated overnight. After the pull-down, beads were washed twice with lysis buffer, twice with washing buffer (20 mM Tris pH 7.5, 300 mM Nacl, 0.5% NP40) and twice with PBS. Beads were eluted using biotin elution solution (6 mM biotin (Sigma) in PBS) and equal volume of reverse cross-link buffer was added (100 mM Tris pH 7.0, 10 mM EDTA, 20 mM DTT, 2% SDS). Reverse cross-linking was performed at 70°C for 2 h. RNA was extracted from eluted and input fractions with Trizol (Life Technologies) according to the manufacturer's instructions. cDNA was transcribed with Superscript III (Life Technologies) using random hexamers in accordance with the manufacturer's protocol. RT-PCR was performed using p53BER2-specific primers (FW 5'-CAGTCTACATTCCCTGGCCTTTG-3', RV 5'-GACATGAGCCATTTTACCCTTAATCC'-3').

Chromosome conformation capture on chip (4C)

4C templates were prepared as described previously (Splinter et al., 2011). In brief, at least 10^7 BJ/ET cells (with or without p53 knockdown) were harvested using trypsin-EDTA, and trypsin was quenched with FCS-containing medium. Cells were cross-linked with 2% formaldehyde for 10 min at room temperature, followed by quenching with glycine (125 mM final) and centrifugation for 8 min at 600 g (4°C). Cells were lysed in 150 mM NaCl, 50 mM Tris-HCl (pH 7.5), 5 mM EDTA, 0.5% NP-40, 1.0% Triton X-100, and nuclei were pelleted for 8 min at 600 x g (4°C). Nuclei were digested overnight with 400 U DpnII (NEB) or Csp6I (Fermentas), followed by proximity ligation in 7 ml using 100 U T4 DNA ligase (Roche) overnight at 16°C. DNA circles were phenol-chloroform extracted and ethanol precipitated using glycogen as a carrier (20 µg/ml). DNA circles were further digested with 50 U Csp6I (for DpnII circles) or DpnII (for Csp6I circles) overnight at 37°C, followed by heat-inactivation and subsequently ligated in 14 ml using 200 U T4 DNA ligase. Trimmed circles, the 4C template, were ethanol precipitated using glycogen as a carrier (20 µg/ml). 16 identical 50 µl PCR reactions were performed per viewpoint using the Expand Long Template PCR system (Roche) with 200 ng 4C template per PCR reaction. PCR conditions were as follows: 95°C for 2 min; 36 cycles of 95°C for 15 s, 55°C for 1 min and 68°C for 3 min; followed by a final step of 68°C for 7 min. All 16 PCR reactions were pooled and purified for next-generation sequencing using the Roche High Pure PCR Product Purification Kit. Each DpnII experiment was validated in a biological replicate experiment with Csp6I, and vice versa.

4C primers carried additional 5' overhangs composed of adapter sequences for Illumina single read sequencing, and were sequenced on a GA-II or Hi-seq 2000 machine (Illumina). Mapping of 4C data was performed essentially as in Splinter et al., 2011. To visualize and identify regions with significant enrichment of 4C signal we generate local 4C domainograms.

To generate these graphs genomic windows of a given size are compared to their directly flanking genomic windows using a Wilcoxon rank-sum test. Formally, the window $W_{i..i+w-1}$ is compared to the up- and downstream windows $W_{i-w..i-1}$ and $W_{i+w..i+2w-1}$, respectively, where W is a genomic window, i is the index of a fragment in the genome and w is the size of the window. To calculate the enrichment along the region of interest a sliding window approach is employed. By performing a statistical test over a range of window sizes (w) and plotting the resulting matrix of p-values along the chromosome a multiscale representation of the 4C data is obtained. Windows that have a p-value $<10^{-5}$ in the Wilcoxon rank-sum test are represented by arcs in Figure 2A-C.

2.6 Supplemental Information

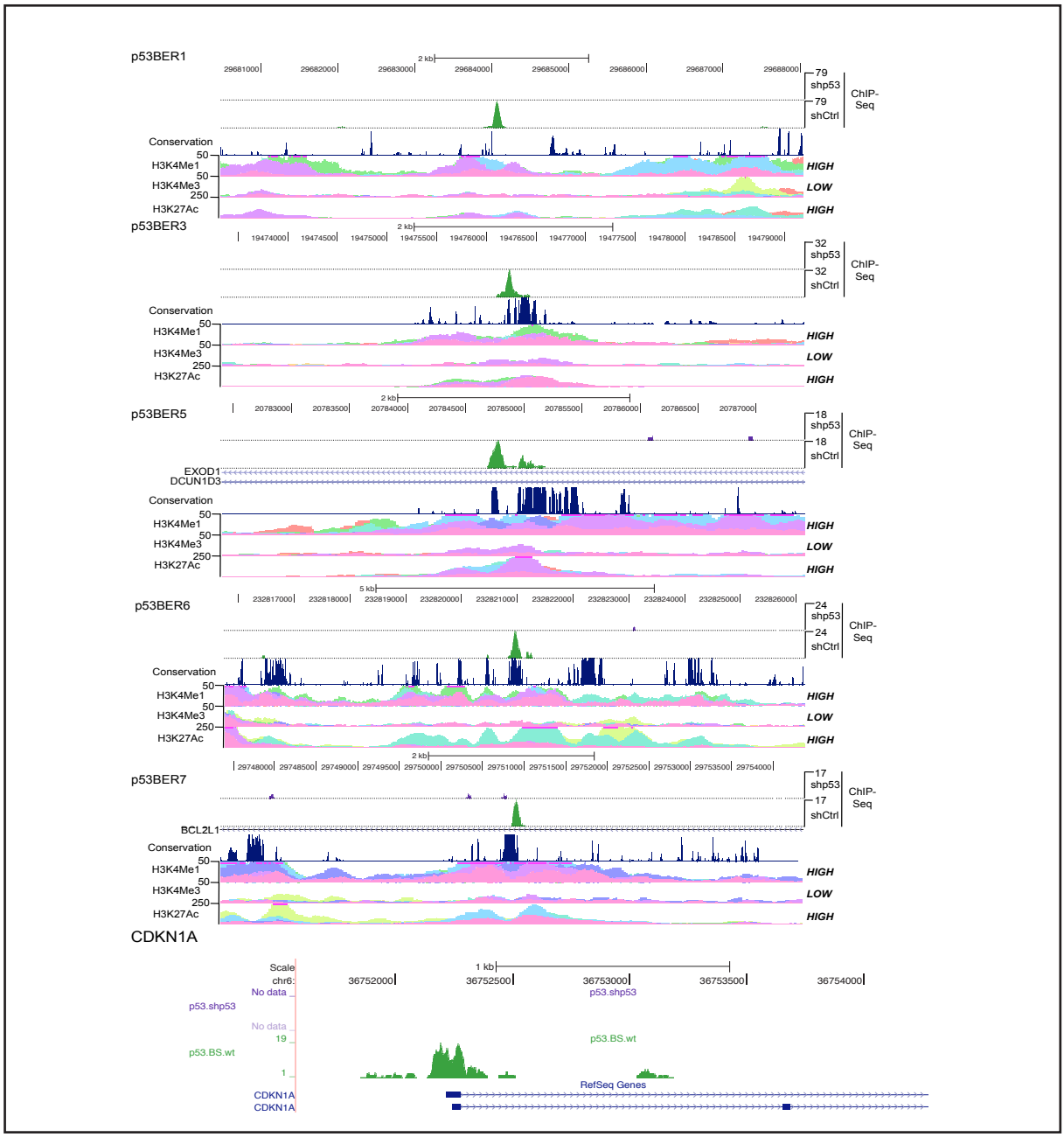
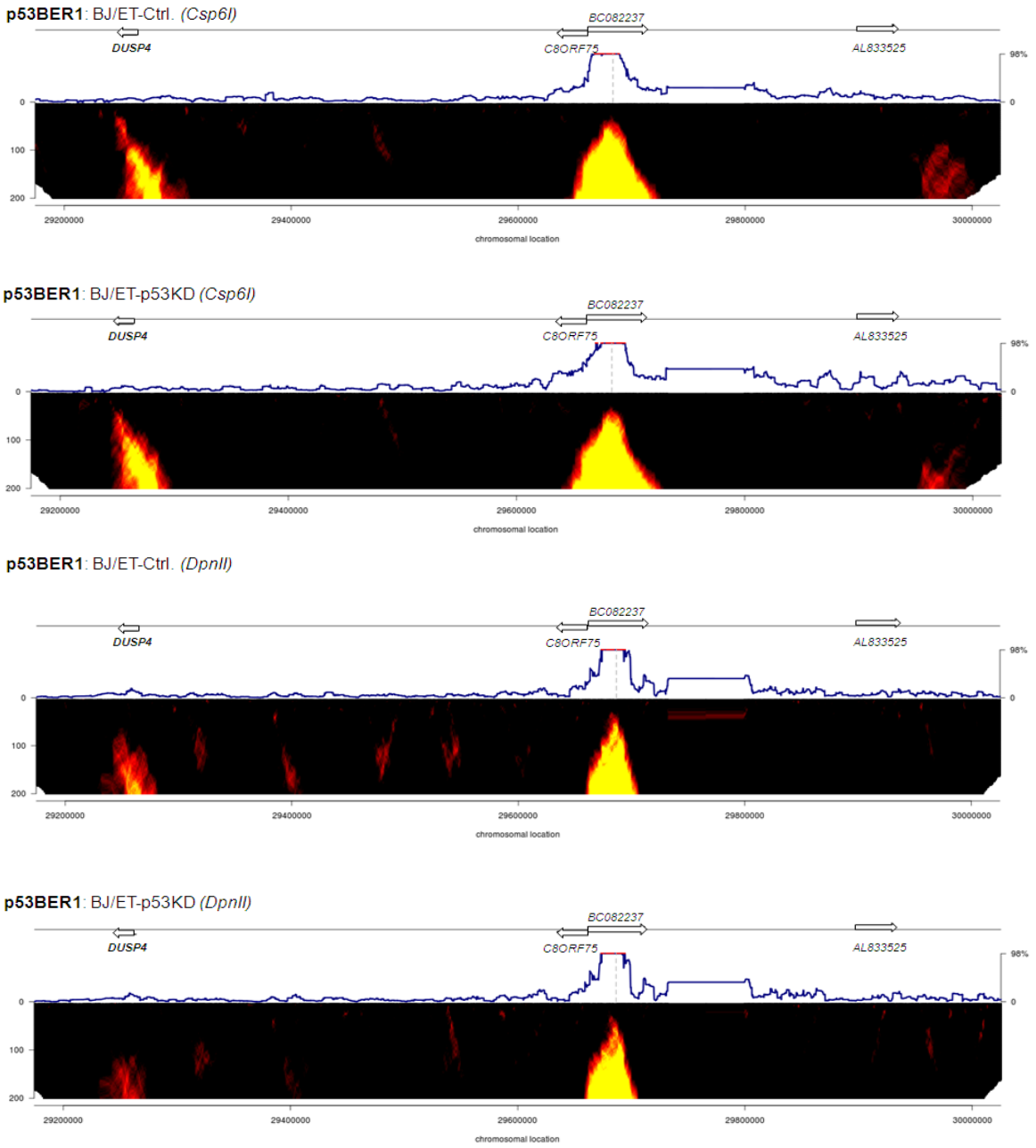
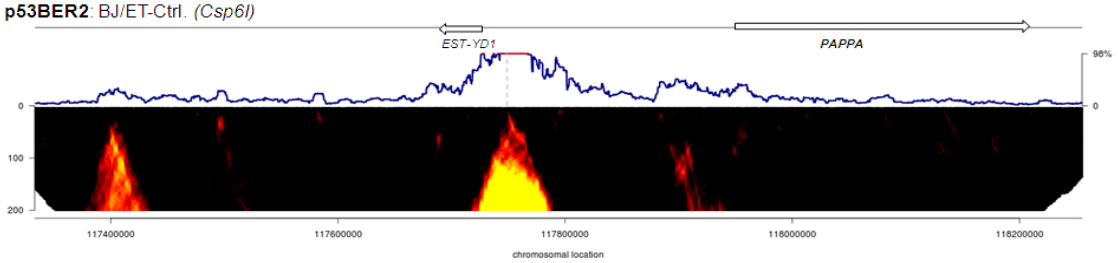


Figure S1 (related to Figure 1). Identification of p53-dependent enhancer domains
(A) UCSC Genome Browser (hg18 assembly) presentation of the p53-binding pattern (Drost et al., 2010) and histone modifications around the indicated p53BERs and the high affinity p53 target gene p21 (CDKN1A; ENCODE Project Consortium). (B) Location of p53 consensus sequences within p53BERs. (C) MCF7 cells were co-transfected with the indicated reporter construct and either control or p53 knockdown (p53KD) vector. The relative firefly luciferase/renilla activity was determined and compared to the control promoter vector (Ctrl.). Graphs represent mean and s.d. from three independent experiments. *P* values were calculated using *t*-test. * p value < 0.01, ** p value < 0.05.

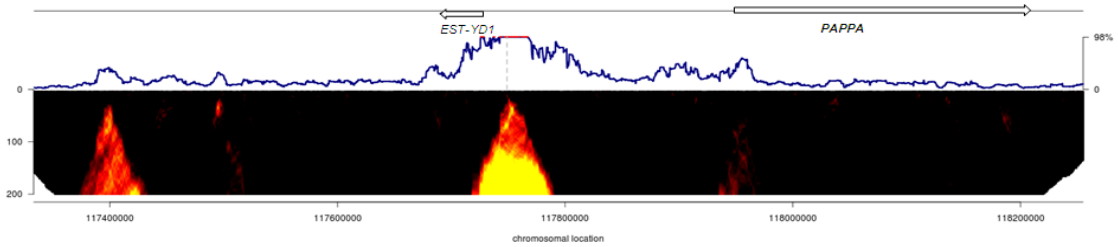
A


p53BER1: BJ/ET-Ctrl. (Csp6I)
BC082237
DUSP4
C8ORF75
AL833525
98%
0
0
100
200
29200000
29400000
29600000
29800000
30000000
chromosomal location
p53BER1: BJ/ET-p53KD (Csp6I)
BC082237
DUSP4
C8ORF75
AL833525
98%
0
0
100
200
29200000
29400000
29600000
29800000
30000000
chromosomal location
p53BER1: BJ/ET-Ctrl. (DpnII)
BC082237
DUSP4
C8ORF75
AL833525
98%
0
0
100
200
29200000
29400000
29600000
29800000
30000000
chromosomal location
p53BER1: BJ/ET-p53KD (DpnII)
BC082237
DUSP4
C8ORF75
AL833525
98%
0
0
100
200
29200000
29400000
29600000
29800000
30000000
chromosomal location

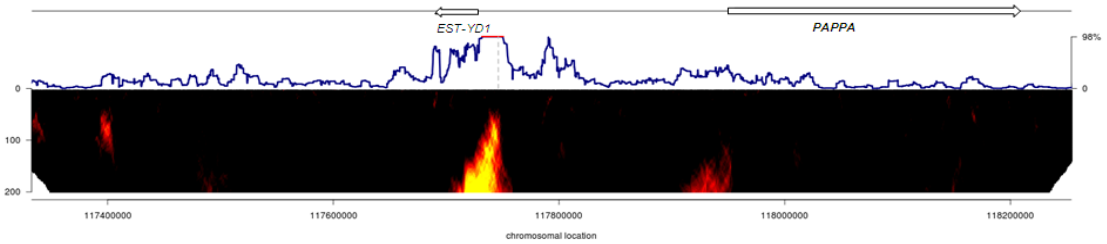
p53BER2: BJ/ET-Ctrl. *(Csp6I)*



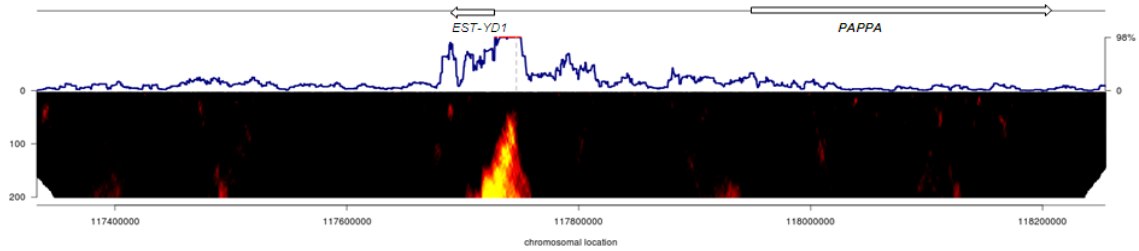
p53BER2: BJ/ET-p53KD *(Csp6I)*

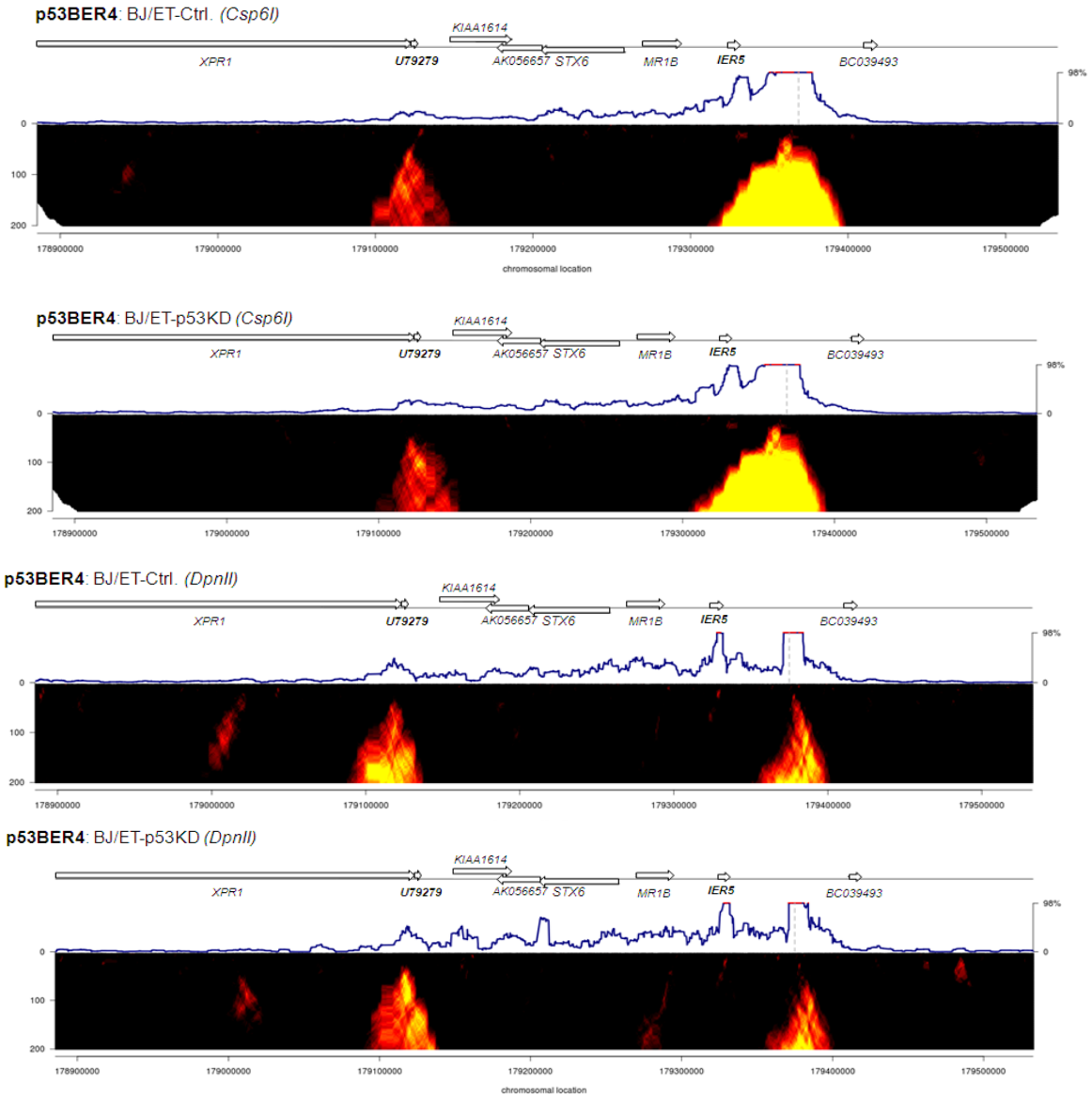


p53BER2: BJ/ET-Ctrl. *(DpnII)*



p53BER2: BJ/ET-p53KD *(DpnII)*




p53BER4: BJ/ET-Ctrl. (Csp6I)
p53BER4: BJ/ET-p53KD (Csp6I)
p53BER4: BJ/ET-Ctrl. (DpnII)
p53BER4: BJ/ET-p53KD (DpnII)
KIAA1614
XPR1
U79279
AK056657 STX6
MR1B
IER5
BC039493
98%
0
100
200
178900000
179000000
179100000
179200000
179300000
179400000
179500000
chromosomal location

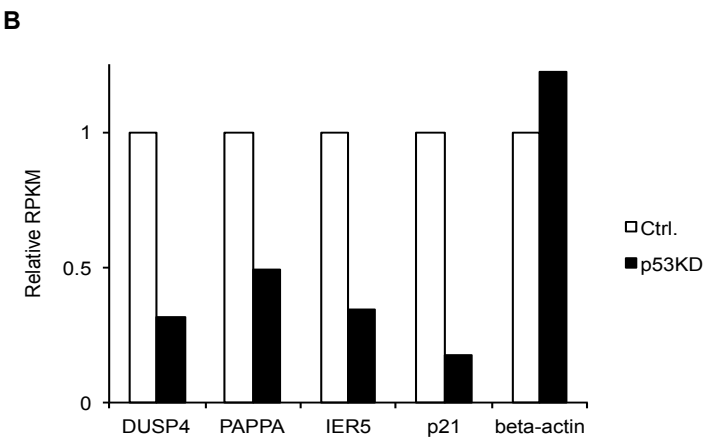

B
Relative RPKM
1
0.5
0
DUSP4
PAPPA
IER5
p21
beta-actin
Ctrl.
p53KD

Figure S2 (related to Figure 2). p53BERs show long-range interactions with multiple distantly located genes

(A) Domainograms (de Wit et al., 2008, see Methods) visualizing significance of interactions at different window sizes of each p53BERs with surrounding chromosomal regions in BJ/ET cells. Color ranges (see scale bar) reflect different levels of significance, from black (low significance, $P=1x10^{-2}$) to yellow (high significance, $P=1x10^{-8}$). For each locus, two independent biological replicates are shown with different restriction enzymes (DpnII and Csp6I). (B) BJ/ET cells expressing either a control or p53 knockdown vector were subjected to 3'-end mRNA sequencing. Depicted are RPKM values as fold difference from control cells.

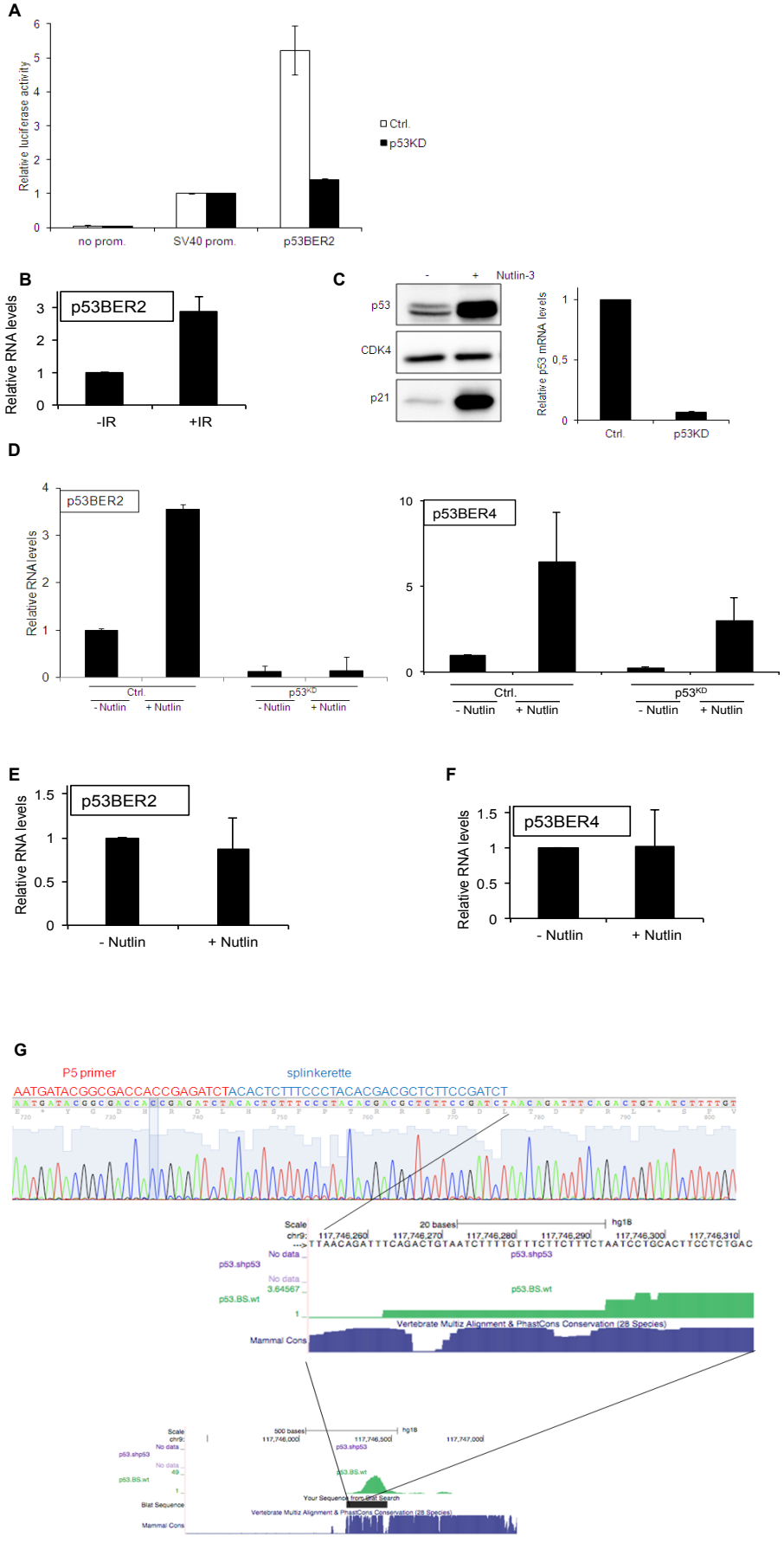

A
Relative luciferase activity
Ctrl.
p53KD
no prom.
SV40 prom.
p53BER2
B
p53BER2
Relative RNA levels
-IR
+IR
C
Nutlin-3
p53
CDK4
p21
Relative p53 mRNA levels
Ctrl.
p53KD
D
p53BER2
p53BER4
Relative RNA levels
Ctrl.
p53KD
- Nutlin
+ Nutlin
E
p53BER2
Relative RNA levels
- Nutlin
+ Nutlin
F
p53BER4
Relative RNA levels
- Nutlin
+ Nutlin
G
P5 primer
splinkerette
AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCT
Scale
20 bases
hg18
chr9:
117,746,260
117,746,270
117,746,280
117,746,290
117,746,300
117,746,310
TTAACAGATTTCAGACTGTAATCTTTTGTTTCTTCTTTCTAATCCTGCACTTCCTCTGAC
No data
p53.shp53
p53.BS.wt
3.64567
Vertebrate Multiz Alignment & PhastCons Conservation (28 Species)
Mammal Cons
500 bases
117,746,000
117,746,500
117,747,000
Blat Sequence
Your Sequence from BLAT Search

Figure S3 (related to Figure 3). p53BER2 and p53BER4 produce enhancer RNAs (eRNAs) in a p53-dependent manner.

(A) MCF7 cells were co-transfected with reporter constructs containing a luciferase gene not driven by a promoter (no prom.), SV40 promoter (SV40 prom.), or p53BER2 and either control or p53 knockdown (p53KD) vector. The relative firefly luciferase/renilla activity was determined and compared to the SV40 promoter vector. Graphs represent mean and s.d. from three independent experiments. (B) MCF7 cells were treated with ionizing radiation (10 Gy) or left untreated. After 8 h, the relative p53BER2 transcript levels were determined by q-RT-PCR. Graphs show mean and s.d. for three independent experiments. (C) MCF7 cells expressing either a control or p53 knockdown vector were treated with Nutlin-3 for 24 h or left untreated and harvested. Western blot analysis was performed to detect p53, p21 and CDK4 (loading control). (D) Cells form (C) were treated with Nutlin-3 for 24 h or left untreated. The relative p53BER2 and p53BER4 transcript levels were determined by q-RT-PCR. Graphs show mean and s.d. for three independent experiments. (E-F) MDA-MB-436 (mutant p53) were treated with Nutlin-3 for 24 h or left untreated. The relative p53BER2 (E) and p53BER4 (F) transcript levels were determined by q- RT-PCR. Graphs show mean and s.d. for three independent experiments. (G) 5'-end of p53BER2 was identified using 5' P5 linker ligation method (see Methods).

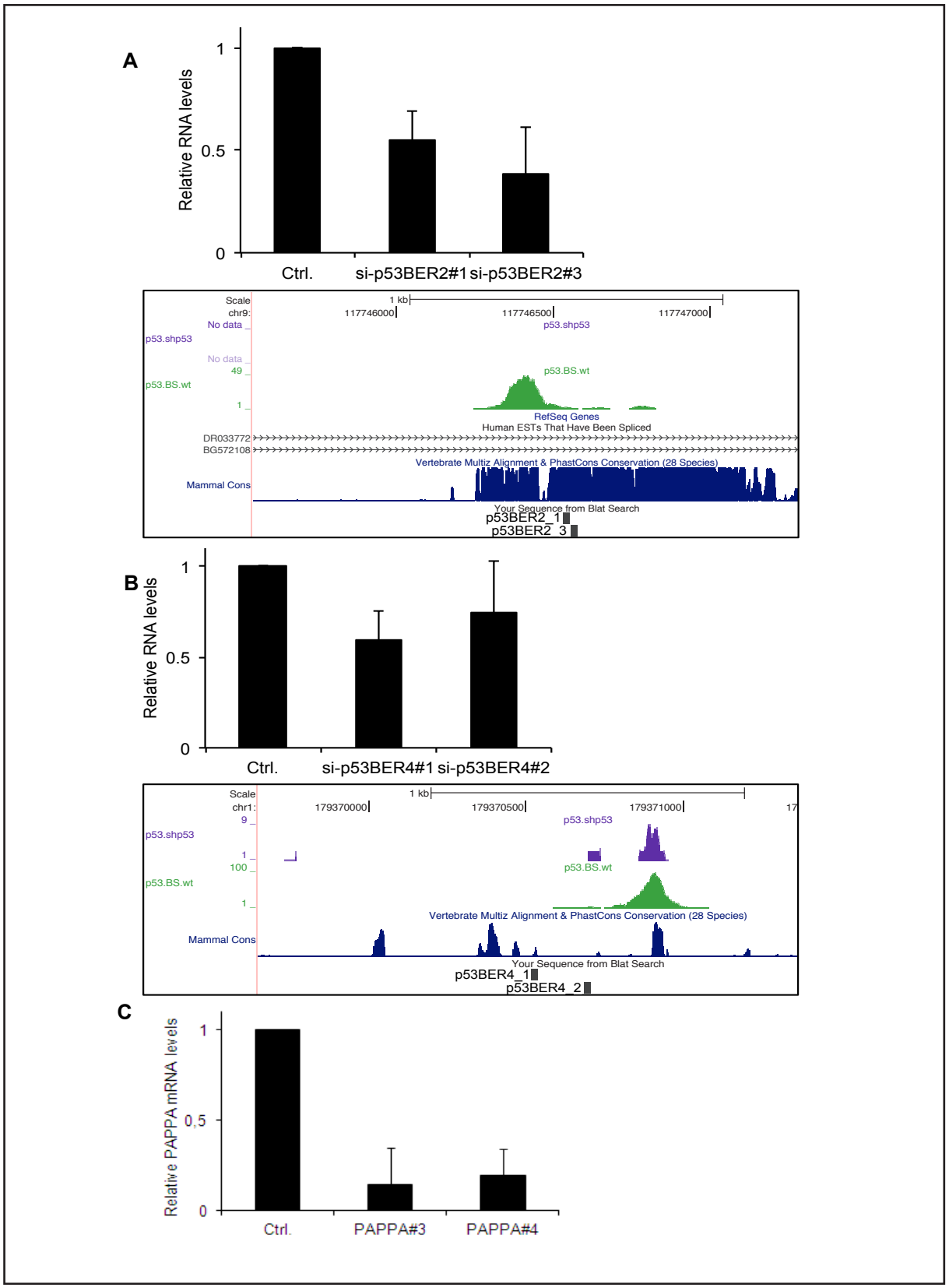


Figure S4 (related to Figure 5). Schematic representation of the siRNA locations and efficiencies at p53BER2

(A) and p53BER4 (B). Efficiency of PAPPA siRNAs (C). Graphs shows means and s.d. for three independent experiments.

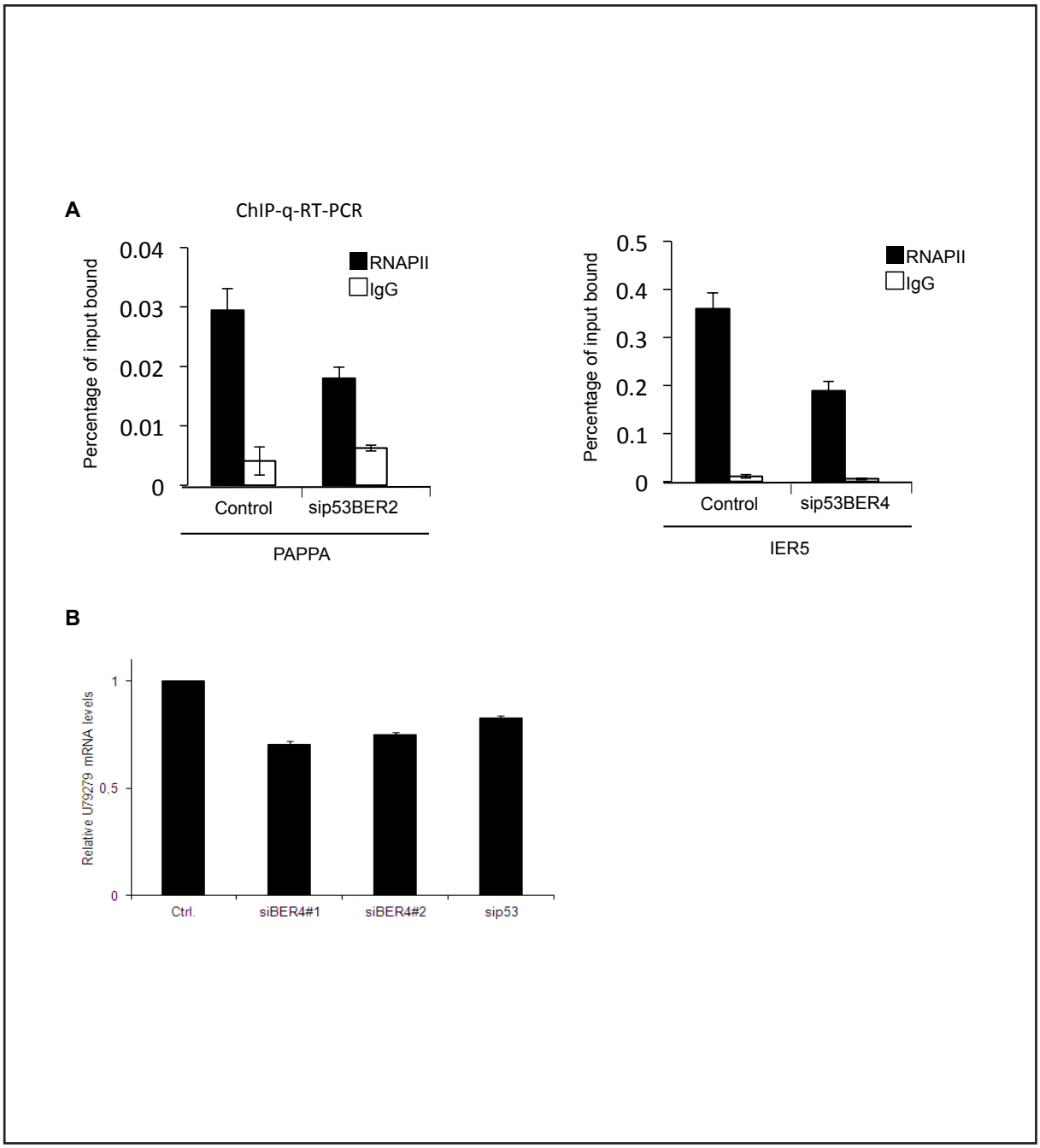
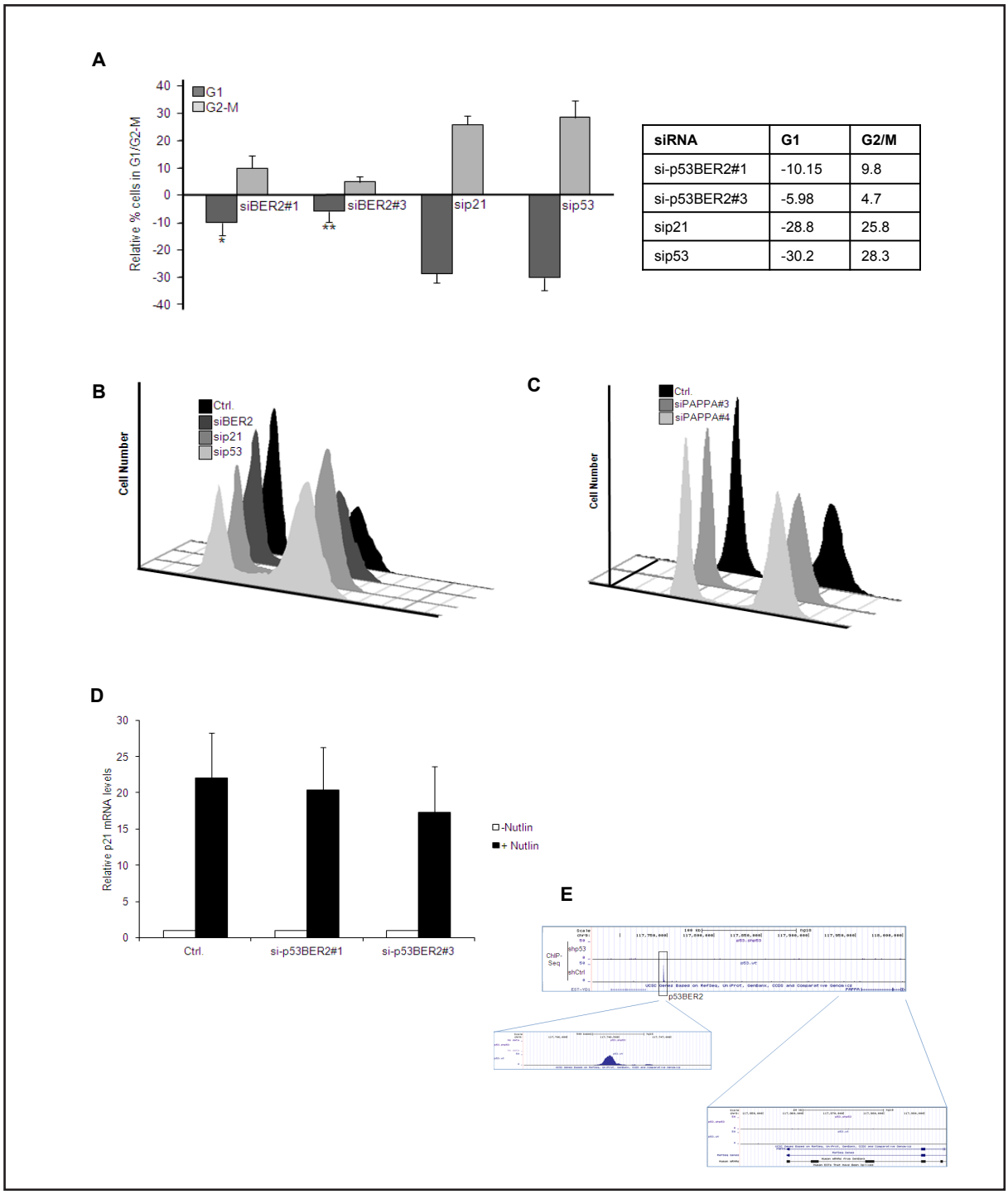


Figure S5 (related to Figure 5). eRNAs produced from p53BERs are required for transcription enhancement of neighboring genes

(A) BJ/ET cells (PAPPA) and MCF7 cells (IER5) were cultured in the presence of Nutlin-3 for 24 h and subsequently subjected to ChIP for RNAPII. Rabbit IgG was used as negative control. Protein binding to the indicated genomic regions was quantified by calculating the percentage of input that is chromatin-bound. ANOVA was used on the percent input values to determine the p-values for the conditions used (sip53BER2 and sip53BER4) vs. scramble, with variance of biological replicates as error term, using the aov() function of R (version 2.15.1). BER2: BER2kd vs scramble: pValue=0.0016; BER4: BER4kd vs scramble: pValue=0.0123. (B) MCF7 cells were transfected with the indicated siRNAs and treated with Nutlin-3 for 24 h. Relative U79279 mRNA levels were measured by q-RT-PCR and is represented as fold induction from values measured in control transfected cells. Graphs represent mean and s.d. from at least three independent experiments.



siRNA	G1	G2/M
si-p53BER2#1	-10.15	9.8
si-p53BER2#3	-5.98	4.7
sip21	-28.8	25.8
sip53	-30.2	28.3

Figure S6 (related to Figure 5). eRNAs produced from p53BER2 are required for p53-dependent cell cycle arrest

(A) Graphs represent mean and s.d. from three independent experiments. *P* values were calculated using *t*-test. * p value < 0.01, ** p value < 0.05. (B) FACS plots of the experiment described in Figure 5C. (C) FACS plots of the PAPPA siRNAs used in the experiment descrided in Figure 5C. (D) MCF7 cells were transfected with the indicated siRNAs and either treated with Nutlin-3 for 24 h or left untreated. Relative p21 mRNA levels were measured by q-RT-PCR and is represented as fold induction from values measured in untreated cells. Graphs represent mean and s.d. from at least three independent experiments. (E) Genomic overview of p53 binding at p53BER2 and PAPPA gene loci as determined by p53 ChIP-seq (Drost et al., 2010).

Table S1

Rank #	chromosome	start	end	Neighbouring gene	p53BER	Peak height
1	chr8	29683912	29684200	DUSP4	1	79
2	chr9	117746224	117746551	PAPPA	2	49
3	chr4	77922127	77922401			
4	chr4	76030785	76031048			
5	chr3	154815119	154815393			
6	chr5	52822299	52822583			
7	chr12	19476074	19476349		3	32
8	chr1	179370714	179371060	IER5	4	100*
9	chr3	196389658	196389937			
10	chr6	36752063	36752408	CDKN1A		19
11	chr16	20784627	20784882		5	18
12	chr4	123938821	123939097			
13	chr8	126270765	126270982			
14	chr17	52149897	52150142			
15	chr1	232820805	232821097		6	24*
16	chr10	89592873	89593111			
17	chr13	113573373	113573719			
18	chr4	157912082	157912296			
19	chr20	29750782	29750998		7	17
20	chr1	142722044	142722314			

Table S2

Cells	Treatment	Fold-change	Technique
BJ	Nutlin 6 hrs	↑ 7.7	RNA-seq*
	Nutlin 19 hrs	↑ 11.07	
BJ	+ RAS	↑ 3.8	RNA-seq*
	+ RAS; p53kd	↓ 2.4	
BJ	p53.kd	↓ 2.5	Microarray**
BJ	IR (5 Gy; 6 hrs)	↑ 2.1	Microarray**

Table S3

Region	Forward (5' – 3')	Reverse (5' – 3')
p53BER1	GTCTAAAAAGTACCTTTACAGC	GATTAATCCATTGATGAGGGAAGAG
p53BER2	CTAGTATATGAACTGTTATTGTCC	TTTTTTAGGTAGAATTTTCCAGC
p53BER3	GAAAGGAGGGCTGGTTGCTTTATTTTGGCTGAAGGCC	CCTTCTGGGAGTCCCAGAAGGGCAAGGGTCTCTCAC
p53BER4	GCTCACGCCTGTAATCTCAGCACTTTG	CTTGTGGCCAACTATGACACTCCAG
p53BER5	CCAAAAATAAATAAATAAAAATGC	TTGCTGATTGATCATCTGTTG
p53BER6	CGTGTCTGCTTGGTTTGGGCTC	CTCTGAGCAATTTCTCAATAAAAC
p53BER7	CGAGTAGCTGGGAATACAGGTGCGAAAC	GAACTCTTACTTTAATGGGAGAGACAG

Table S4

Gene	Forward (5' – 3')	Reverse (5' – 3')
p53BER2	CCAGTGGAAAGTGGGGGAGT	CCTGAAATCTCTGCTTGGCTTTG
p53BER4	TGGCACTGGGCTTAGGTCTTTT	CCCACAAGGGCTCTCAAGTTC
p53	CTCCTCTCCCCAGCCAAAGA	GGAACATCTCGAAGCGCTCA
p21	TACCCTTGTGCCTCGCTCAG	GAGAAGATCAGCCGGCGTTT
PAPPA	TGCCGAGAGAATAAGCACAAGG	GGTGGAGGTGGGTCACAGG
U79279	GACTGAGGCAGGAGAATTGC	TGGTTAGGTTGCTGCTCCTT

Table S5

Region	Forward (5'-3')	Reverse (5'-3')
p53BER2	GTTACAAATACGGGAGCTGGAC	AAAGGCCAGGGAATGTAGACTG
Downstream p53BER2	CTAGAACTTTGTCTGATGTGGTCAA	GGACAGCTCTTTATTCCATTATGTC
p53BER4	ACTCTTAAACAGCCTTGTGGTTTC	CAGAAACACCCCACTACAGTCTC
Downstream p53BER4	AAGTAGAGTCCTGGGAGAAACAGA	GGCCCTAGACAAACATCTAAAGAAC
IER5	ATCTCTCTGGGCAAGATCTACAAC	GGTCGCTCAGGTAGACTTGG
MDM2	GGGCTATTTAAACCATGCATTTTC	GTCCGTGCCCACAGGTCTA

Chapter three

Genome wide profiling of p53-responsive enhancer RNAs uncovers an essential lncRNA-dependent epigenetic regulation

Nicolas Léveillé[1,#], Carlos A. Melo[1,2,#], Koos Rooijers[1,#], Angel Díaz-Lagares[3], Sonia A. Melo[4], Gözde Kormaz[1], Rui Lopes[1], Farhad Akbari Moqadam[5], Ana R. Maia[6], Patrick J. Wijchers[7], Geert Geeven[7], Monique L. den Boer[5], Raghu Kalluri[4], Wouter de Laat[7], Manel Esteller[3,8,9] and Reuven Agami[1,10]

[1]Division of Gene Regulation, The Netherlands Cancer Institute, Plesmanlaan 121, 1066 CX, Amsterdam, The Netherlands. [2]Doctoral Programme in Biomedicine and Experimental Biology, Centre for Neuroscience and Cell Biology, Coimbra, Portugal. [3]Cancer Epigenetics and Biology Program (PEBC), Bellvitge Biomedical Research Institute (IDIBELL), L'Hospitalet de Llobregat, 08908 Barcelona, Catalonia, Spain. [4]Department of Cancer Biology, University of Texas MD Anderson Cancer Center, Houston, 77054, USA. [5]Department of Pediatric Oncology and Hematology, Erasmus University Medical Center, Erasmus MC, Rotterdam, the Netherlands. [6]Department of Cell Biology and Cancer Genomics Center, The Netherlands Cancer Institute, Plesmanlaan 121, 1066 CX, Amsterdam, Netherlands. [7]Hubrecht Institute-KNAW and University Medical Centre Utrecht, Uppsalalaan 8, 3584CT Utrecht, The Netherlands. [8]Department of Physiological Sciences II, School of Medicine, University of Barcelona, 08036 Barcelona, Catalonia, Spain. [9]Institució Catalana de Recerca i Estudis Avançats (ICREA), 08010 Barcelona, Catalonia, Spain. [10]Erasmus MC, Rotterdam University, The Netherlands.
[#]These authors contributed equally to this work.
Correspondence should be addressed to R.A. (r.agami@nki.nl)

Under revision

3.1 Summary

Recently, p53 was shown to bind enhancers to regulate key target genes. Here, we used enhancer RNA (eRNA) production to globally map p53-regulated enhancers. Intriguingly, though many p53-induced enhancers contained p53-binding site, most did not. As long-noncoding RNAs (lncRNAs) are prominent regulators of chromatin function, we hypothesized that p53-induced lncRNAs contribute to activation of enhancers by p53. Among p53-induced lncRNAs, we identified LED whose suppression attenuated p53 function. Chromatin binding and eRNA expression analyses showed that LED associates and activates strong enhancers. Interestingly, LED-bound enhancers were enriched among p53-induced enhancers.

One prominent target of LED was located at an enhancer within CDKN1A, a potent p53-responsive cell-cycle inhibitory gene. LED suppression attenuated CDKN1A enhancer activity, CDKN1A induction, and cell-cycle arrest following p53 activation. Finally, promoter-associated hypermethylation studies indicated silencing of LED in human cancer cell lines and tumors. Preference to p53 wild-type tumors connects LED-silencing to p53-mediated tumor suppression.

3.2 Introduction

For several decades the foundations of molecular biology leaned against the dogma that genetic information is stored in protein-coding genes (Crick et al., 1961). Although this concept was, and is still, largely true in prokaryotes, where genomes are mainly composed of protein-coding genes, it does not hold true for higher eukaryotes, where protein-coding sequences occupy less than 3% of the genome. Once considered transcriptionally inactive or simply referred to as "junk DNA", the predominant fraction of the genome is in fact pervasively transcribed into thousands of different non-coding RNAs (ncRNAs), which can further be divided into two groups: small non-coding RNAs (sncRNAs) and long non-coding RNAs (lncRNAs). In addition, lncRNA genes have been classified based on the epigenetic state of their chromatin. For instance, the long intergenic non-coding RNAs (lincRNAs) are known for the presence of high histone 3 lysine 4 trimethylation (H3K4me3) at their promoters and high H3K36me3 along their transcribed regions, also referred as the K4K36 signature (Guttman et al., 2009).

Alternatively, enhancer RNAs (eRNAs) are produced from transcriptionally active enhancer regions which are epigenetically defined by high level of H3K4me1, low level of H3K4me3 (Heintzman et al., 2007a; Visel et al., 2009), and high level of histone 3 lysine 27 acetylation (H3K27Ac) and H3K9Ac (Creyghton et al., 2010a).

Importantly, lincRNAs have recently emerged as potent regulators of gene expression.

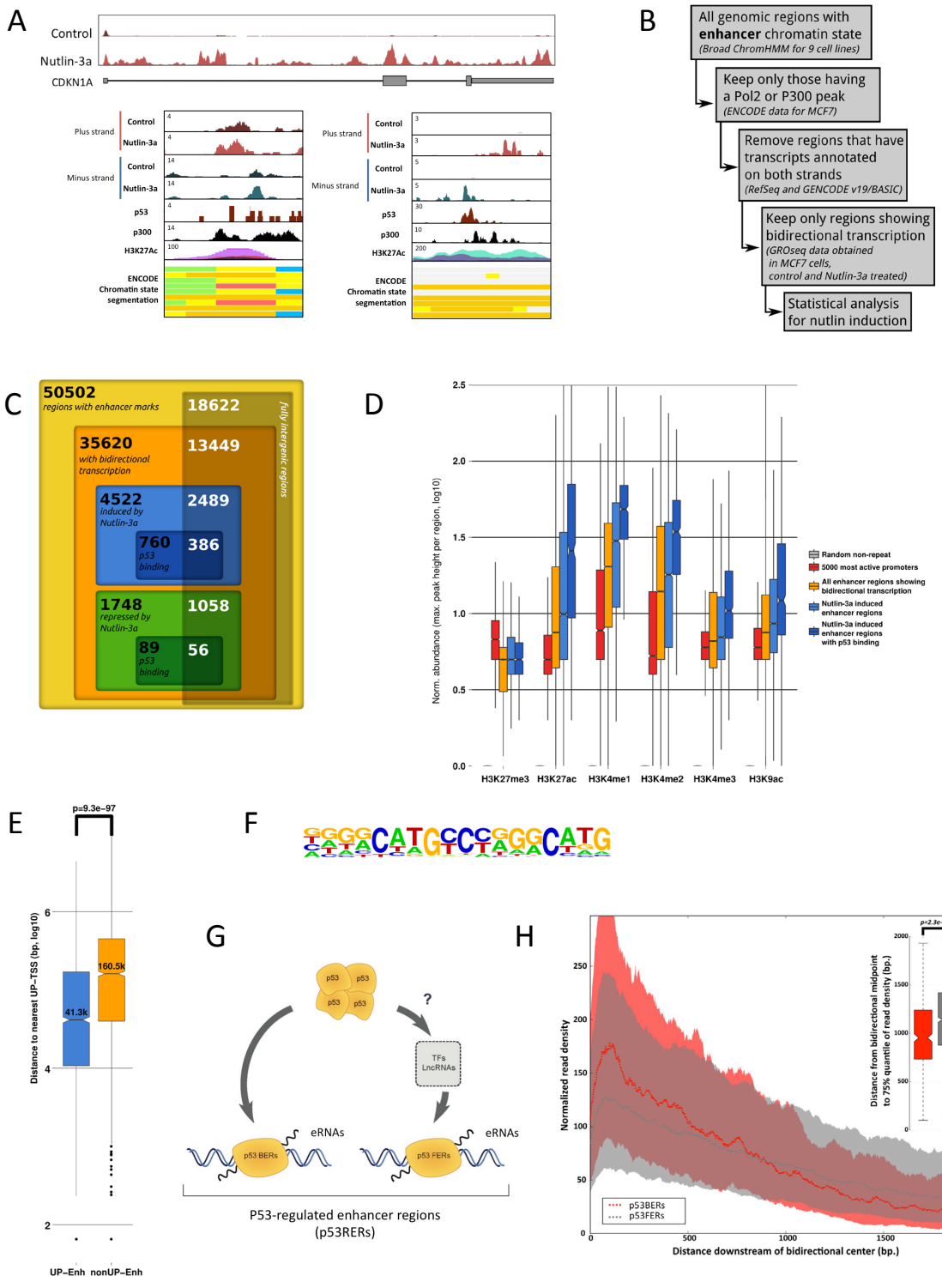
Recent publications have shown that lincRNAs are able to form complexes with various chromatin modifiers and to specifically direct them to different genomic regions. For example, the lincRNA-p21 was shown to interact with and guide the heterogeneous nuclear ribonucleoprotein K (hnRNP-K) to repress a subset of p53 target genes (Huarte et al., 2010). However, while lincRNAs can mediate their effect in cis and in trans, eRNAs have been so far mainly characterized for their function in cis. Classically expressed as bidirectional transcripts from enhancer regions, eRNAs can alter the expression of their neighboring genes through the formation of DNA loops, which help to bridge the interaction between enhancers and nearby promoters. Several transcription factors were found to be important coordinators of eRNA expression (Lam et al., 2013; Li et al., 2013; Melo et al., 2013b).

An interesting case revealed that the tumor suppressor p53 directly regulates the expression of eRNAs upon cellular stresses (Melo et al., 2013b). *P53* function is frequently compromised in tumors, in part as a consequence of somatic mutations which occur in more than 50% of all human cancers (Beckerman and Prives, 2010). Moreover, it was also shown that p53 is inactivated in various cancers by dysregulation of its regulatory pathway, such as the amplification and over-expression of its negative regulators MDM2 and MDM4 (Gembarska et al., 2012; Momand et al., 1998). Upon cellular stresses, p53 is activated and acts primarily as a transcription factor to mediate and coordinate a complex transcriptional response that regulates hundreds of target genes.

Until recently, the p53 network was mainly characterized by its impact on protein-coding target genes (Wei et al., 2006). However, we now begin to discover and appreciate the great potential of ncRNAs in the intricate regulatory network of p53. The recent discoveries that p53 can mediate its function in collaboration with diverse lncRNAs, suggest a potential role for this novel regulatory layer in disease such as cancer, and therefore urge the importance of an in-depth reassessment of the p53 transcriptional stress-response. Here, by using Global Run-On sequencing (GRO-seq) we mapped p53-responsive enhancers bound by p53. Surprisingly, we also found a large group of p53-activated enhancers that were not associated with p53. While motif-search analysis identified the p53 signature in the enhancers bound by p53, no transcription factor signature was uncovered in the p53-unbound enhancer group.

We hypothesized that p53-responsive lncRNAs may play a role in p53-mediated enhancer activation. Indeed, we show that a prominent p53-induced lncRNA termed LED (LncRNA activator of Enhancer Domains) is required for p53-induced cell cycle arrest and is involved in the activation of a subset of p53-bound and unbound enhancers by inducing an epigenetic modification. Strikingly, promoter-associated hypermethylation of LED was uncovered in several cancer cell lines and human tumors with preference to p53 wild type status, suggesting its implication in tumorigenesis. Altogether, we propose that LED is an important epigenetic regulator and a potential tumor suppressor of the p53 pathway.

Figure 1. Identification of p53 regulated enhancer RNAs (p53RERs)
(A) GRO-Seq snapshot showing the induction of CDKN1A/p21 transcription upon nutlin-3a treatment (upper scheme). Display of nutlin-3a-induced bidirectional transcription at p53-bound and unbound enhancers (lower schemes). Binding of p300, presence of H3K27 acetylation and chromatin states in 9 cell lines are also presented. (B) Diagram showing the outline of the algorithm to identify enhancers using chromatin segmentation data, and ChIP-Seq and GRO-Seq data. (C) Venn diagram showing the number of retrieved regions at the steps of the enhancer identification algorithm. (D) Boxplot showing the abundance of several enhancer marks at different regions (grey: 5000 random non-repeat regions, red: promoters of the 5000 most abundant genes as identified by GRO-seq in the nutlin-3a-treated condition, orange: all putative enhancer regions showing bidirectional transcription, blue: subset of the putative enhancers showing significant induction upon nutlin-3a treatment (induced p53RERs), dark blue: subset of the induced p53RERs having a p53 peak within 1kb (p53BERs). (E) Boxplot showing the distances between enhancer region and the nearest annotated gene induced upon nutlin-3a treatment, for induced p53RERs (UP Enhancers) and nutlin-3a unresponsive enhancers (nonUP enhancers). (F) Motif identified in induced p53BERs using HOMER. (G) Schematic representation of p53 regulated-enhancers (p53RERs). P53 can directly bind to enhancers (p53BERs) or regulated intermediate factors (e.g. lncRNAs or TFs) to indirectly influence another subset of enhancers (p53FERs). (H) Normalized density of transcription downstream of the point of bidirectional transcription for p53BERs (red) and p53FERs (grey). The dotted line indicates the median across all regions and the shaded area indicates the middle 50% quantile. The boxplot in the inset shows the distance from the point of bidirectional transcription to the 75% quantile of read density.


A
Control
Nutlin-3a
CDKN1A
Plus strand
Minus strand
p53
p300
H3K27Ac
ENCODE Chromatin state segmentation
B
All genomic regions with enhancer chromatin state
(Broad ChromHMM for 9 cell lines)
Keep only those having a Pol2 or P300 peak
(ENCODE data for MCF7)
Remove regions that have transcripts annotated on both strands
(RefSeq and GENCODE v19/BASIC)
Keep only regions showing bidirectional transcription
(GROseq data obtained in MCF7 cells, control and Nutlin-3a treated)
Statistical analysis for nutlin induction
C
50502
regions with enhancer marks
18622
fully intergenic regions
35620
with bidirectional transcription
13449
4522
induced by Nutlin-3a
2489
760
p53 binding
386
1748
repressed by Nutlin-3a
1058
89
p53 binding
56
D
Norm. abundance (max. peak height per region, log10)
H3K27me3
H3K27ac
H3K4me1
H3K4me2
H3K4me3
H3K9ac
Random non-repeat
5000 most active promoters
All enhancer regions showing bidirectional transcription
Nutlin-3a induced enhancer regions
Nutlin-3a induced enhancer regions with p53 binding
E
p=9.3e-97
Distance to nearest UP-TSS (bp, log10)
41.3k
160.5k
UP-Enh
nonUP-Enh
F
G
p53
TFs LncRNAs
eRNAs
p53 BERs
p53 FERs
P53-regulated enhancer regions (p53RERs)
H
Normalized read density
p53BERs
p53FERs
Distance downstream of bidirectional center (bp.)
Distance from bidirectional midpoint to 75% quantile of read density (bp.)
p=2.3e-13

3.3 Results

3.3.1 Genome-wide identification of p53-regulated enhancer RNAs

To detect active enhancers we relied on the observation that eRNA production marks enhancer activity. Using Global Run-On sequencing (GRO-seq) of MCF-7 cells treated with nutlin-3a, a specific activator of p53, we obtained a genome-wide quantitative snapshot of transcriptional activity. As expected, the activation of CDKN1A/p21, and many other known target genes of p53, was readily apparent (Figure 1A). Moreover, also several previously described p53-induced eRNAs could be confirmed (Figure 1A) (Melo et al., 2013b). We proceeded to generate a global view of the putative p53-regulated enhancers in MCF-7 cells. By taking the union of the enhancer domains defined by the Broad chromatin segmentation (Ernst et al., 2011), we selected only regions showing RNA polymerase II (RNAPII) and p300 binding in MCF-7 cells (using publicly available ChIP-seq data (Nikulenkov et al., 2012)) and excluded those having annotated transcripts on both strands, as well as those having transcription start sites. The remaining regions were extended by 1kb for the purpose of read counting and used in conjunction with the aforementioned GRO-seq data (Figure 1B). This analysis resulted in the detection of 50,502 putative enhancers of which 6,270 were regulated (at least one direction) by nutlin-3a treatment, and referred here as p53-regulated enhancer regions or p53RERs. Since p53 mainly functions as an activator of transcription (Wang et al., 2009), the vast majority (72%) of the differentially expressed eRNAs showed induction upon nutlin-3a treatment, and the number of activated enhancers were more often bound by p53 than the repressed enhancers (Figure 1C). Using ENCODE ChIP-seq data obtained from MCF-7 cells, the presence of enhancer-specific histone modifications was confirmed (Figure 1D). Moreover, on average, nutlin-3a induced enhancers are positioned closer to p53-regulated canonical genes (median 41kb), than to non-regulated genes (median 161kb) (Figure 1E). This observation is in agreement with the notion that eRNAs are potent regulators of neighboring target genes (Lam et al., 2013; Li et al., 2013; Melo et al., 2013b). In support of a role for nutlin-3a-regulated enhancer regions within the p53 pathway, the gene ontology analysis on neighboring genes revealed enrichment for genes involved in DNA damage response/signal transduction by p53 (GO term 0030330, p=1.6e-3). Moreover, a de novo motif analysis using HOMER (Heinz et al., 2010b) confirmed the presence of a p53 response element at p53-bound enhancer regions (p53BERs) (Figure 1F).

Next we reanalyzed published p53 ChIP-seq data (Nikulenkov et al., 2012) to identify which of the p53RERs were direct targets of p53 (hereafter referred to as p53 bound-enhancer regions or p53BERs). The enhancers in the remaining subset of p53RERs were considered p53-free enhancer regions, or p53FERs, as no enrichment for any known transcription factor (TF) signature was found. One possible reason for this observation is that activation of these enhancers could be mediated by a combination of several different factors (Figure 1G). Indeed,

the analysis of TF binding sites (using the ENCODE Uniform TFBS data) showed that while several TFBSs were enriched in induced p53RERs, with respect to all enhancers, no TFBS was specifically enriched in the p53-free enhancer group (Figure S1A). An alternative explanation is that the transcriptional activation of enhancers is mediated by p53-dependent lncRNAs, as lncRNAs were recently shown to be able to associate with and modulate regulatory elements (Vance et al., 2014a; Yang et al., 2013). In support of the hypothesis that the activation of p53BERs and p53FERs is fundamentally different, we observed a faster transcription drop-off for the first group compared to the second (Figure 1H and Figure S1B).

3.3.2 LED, a p53-induced lncRNA required for the p53-stress response

We therefore set to identify relevant lncRNAs by profiling the transcriptome of nutlin-3a treated MCF-7 cells. Using RNA-sequencing (RNA-seq) in combination with an annotation catalog comprised of Ensembl, Refseq and the Broad Linc Catalog (Cabili et al., 2011), we identified 194 nutlin-3a-responsive lncRNA genes (Figure 2A). We then reasoned that the most up-regulated transcripts might have a greater biological importance, and consequently selected the top 3 most activated lncRNAs (i.e. RP3-510D11.2, loc643401 and linc00086 (*hereinafter* referred to as LED) for further characterization (Figure 2B). Validation confirmed that the selected lncRNAs were induced upon both nutlin-3a and ionizing radiation (IR) treatment (Figure 2C). Next, we determined whether these lncRNAs were regulated by p53. As expected, we observed that p53 depletion decreased both the basal and nutlin-3a-induced levels of all tested lncRNAs (Figure 1D). Moreover, we demonstrated the direct binding of p53 at each lncRNA locus by using publicly available p53 chromatin immunoprecipitation sequencing (ChIP-seq) data and ChIP-qPCR (Figures 2E and S2A). Altogether, these results demonstrate that our selected lncRNAs are *bona fide* p53 targets. Next, we assessed whether the depletion of our selected p53-induced lncRNAs phenotypically influenced the p53-stress response using short interfering RNA (siRNA). Among the investigated candidates, only LED significantly influenced the G1 checkpoint arrest following nutlin-3a treatment, as shown by flow cytometry (Figures 2F and 2G). To corroborate this finding we first evaluated cellular entry into mitosis using phospho-H3(ser10) staining. Cells treated with siRNAs targeting LED showed a significant increase of phospho-H3(ser10) compared to cells transfected with a non-targeting siRNA (Figures S2B and S2C). Furthermore, cell proliferation assays confirmed this observation, as LED-suppressed cells, but not control cells, proliferated more following nutlin-3a treatment in comparison with control-transfected cells (Figures 2H-I and S2D). Thus, the induction of LED lncRNA is required for efficient sustenance of p53 stress-response. To investigate the mechanism by which LED impacts the p53-stress response, we performed gene expression analysis by RNA-seq following knockdown of LED. 1983 genes were responsive to LED depletion (FDR less than 1 %), of which 1340 were up-regulated and 643 down-regulated

Figure 2. Novel stress-regulated lncRNA LED

(A) Outline of pipeline for identification of p53 regulated lncRNAs. MCF-7 cells were treated with 8 μM nutlin-3a for 16 h and subjected to RNA sequencing (RNA-seq). (B) Display of genomic location and RNA-seq data showing the nutlin-3a induction of selected lncRNAs in MCF-7 cells. Values are represented by RPM (reads per million). (C) Stress-dependent regulation of selected lncRNAs upon nutlin-3a (8 μM) and ionizing radiation (10 Gy) treatment in MCF-7 cells measured by qRT-PCR. Values are represented by fold induction (n=3; p-values were calculated with a two-tailed student's t-test. ***p<0.005, **p<0.01, *p<0.05). (D) P53-dependent regulation of validated lncRNAs in MCF-7 cells transfected with a control or p53 siRNA in the presence or absence of nutlin-3a. Values are represented by fold induction (n=3; p-values were calculated with a two-tailed student's t-test. **p<0.01, *p<0.05). (E) Schematic representation of p53 response element (p53 RE) in LED gene body. Chromatin immunoprecipitation performed in nutlin-3a treated MCF-7 cells using IgG or p53 antibodies followed by qPCR in the p53 RE region. Values represent the percentage of input. (n=3; p-values were calculated with a two-tailed student's t-test, *p<0.05). (F) Relative LED RNA levels measured by qRT-PCR from MCF-7 cells transfected with a control or two independent LED siRNAs (LED-kd) (n=3; p-values were calculated with a two-tailed student's t-test. ***p<0.005, *p<0.05). (G) Relative cell cycle variation (LED-kd minus control-kd) of MCF-7 cells transfected with a control or two independent LED siRNAs, treated with nutlin-3a for 16 h. To capture cycling cells in G2/M, cells were treated with nocodazole for 24 h, prior to flow cytometric analysis (n=3; p-values were calculated with a two-tailed student's t-test. **p<0.01, *p<0.05). (H) Schematic representation of cell proliferation assay in MCF-7 treated with 2 pulses of nutlin-3a and transfected with a control or LED siRNA. (n=3; p-values were calculated with a two-tailed student's t-test. ***p<0.005). (I) Colony formation assay using the same condition as in (H).

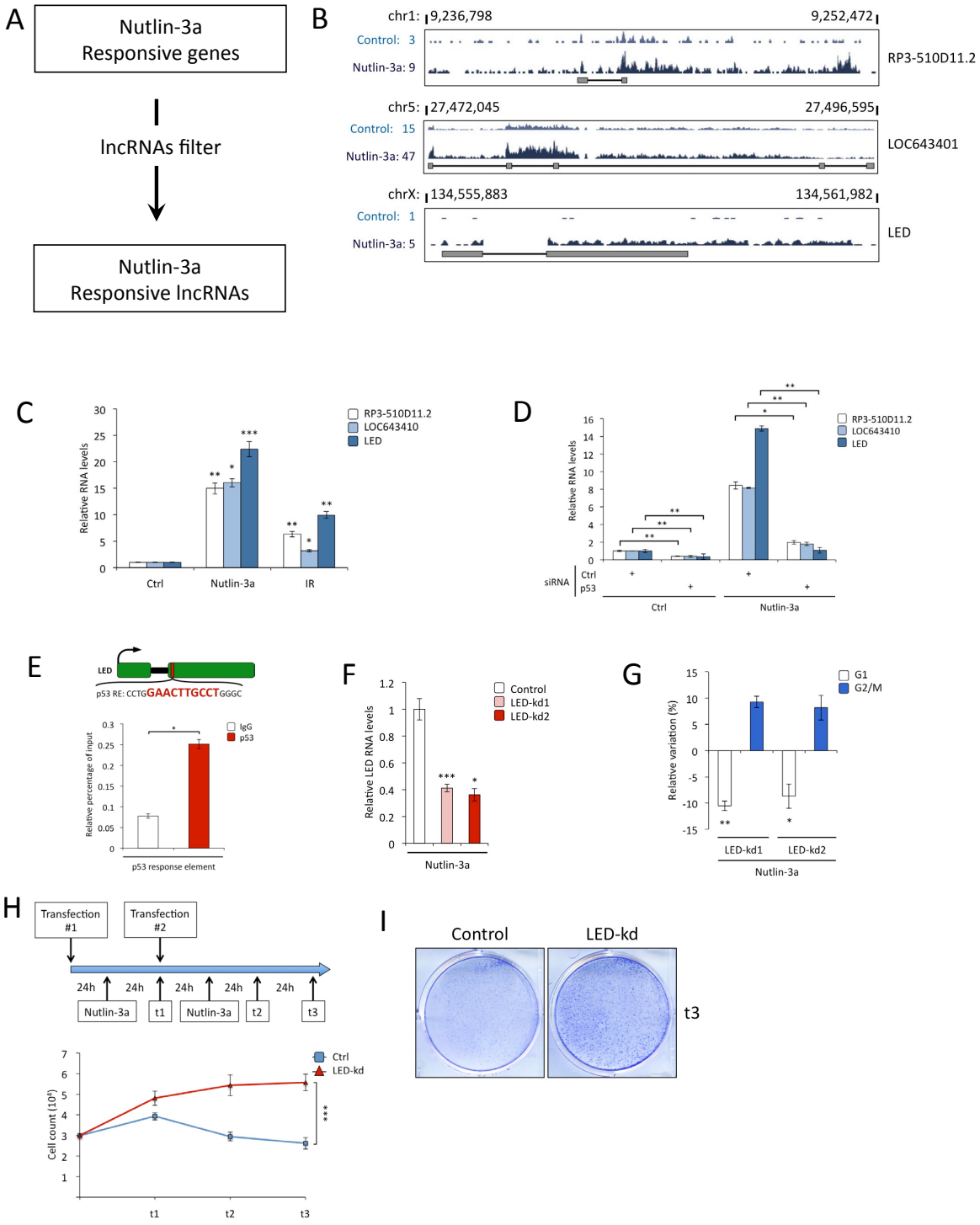
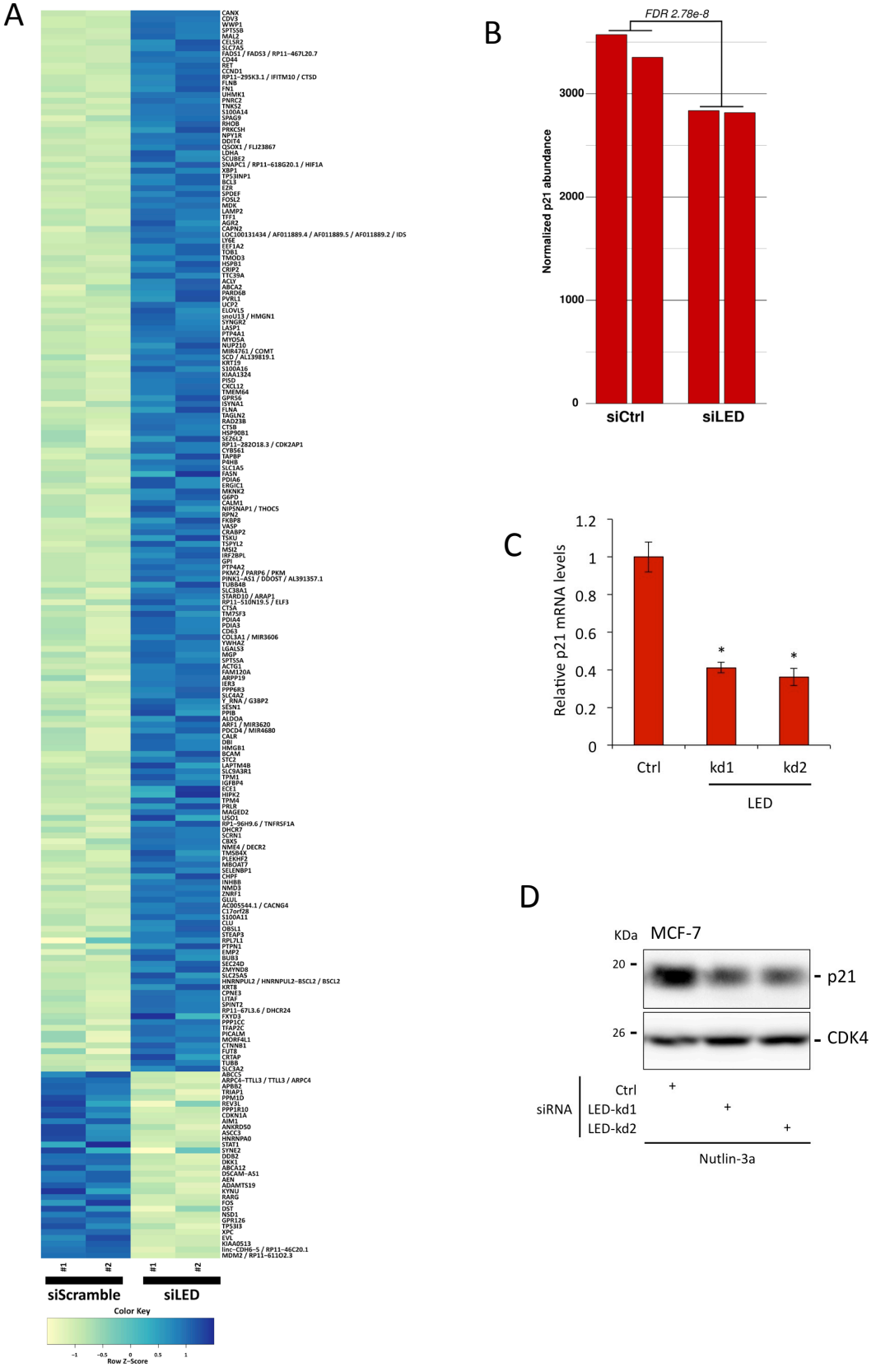

A
Nutlin-3a
Responsive genes
lncRNAs filter
Nutlin-3a
Responsive lncRNAs
B
chr1: 9,236,798
9,252,472
Control: 3
Nutlin-3a: 9
RP3-510D11.2
chr5: 27,472,045
27,496,595
Control: 15
Nutlin-3a: 47
LOC643401
chrX: 134,555,883
134,561,982
Control: 1
Nutlin-3a: 5
LED
C
Relative RNA levels
RP3-510D11.2
LOC643410
LED
Ctrl
Nutlin-3a
IR
D
Relative RNA levels
siRNA
Ctrl
p53
Ctrl
Nutlin-3a
E
LED
p53 RE: CCTGGAACTTGCCTGGGC
Relative percentage of input
IgG
p53
p53 response element
F
Relative LED RNA levels
Control
LED-kd1
LED-kd2
Nutlin-3a
G
Relative variation (%)
G1
G2/M
LED-kd1
LED-kd2
Nutlin-3a
H
Transfection #1
Transfection #2
24h
Nutlin-3a
t1
t2
t3
Ctrl
LED-kd
Cell count (10⁴)
I
Control
LED-kd
t3

Figure 3. LED is required for the proper p53 stress response

(A) RNA sequencing heatmap showing a subset of genes differentially expressed upon LED knockdown (only the subset with absolute fold-change > 30% and FDR < 5% is shown). (B) Barplot derived from the RNA-sequencing showing the normalized p21 mRNA levels in control and LED knockdown conditions. (C) Relative mRNA levels of p21 upon transfection of a control or two independent LED siRNAs (LED-kd), measured by qRT-PCR in MCF-7 cells treated with nutlin-3a (n=3; p-values were calculated with a two-tailed student's t-test. *p<0.05, ***p<0.005). (D) Western blot showing p21 protein levels from MCF-7 cells transfected with a control or two independent LED siRNA (LED-kd) treated with nutlin-3a for 16 h. Detection of CDK4 protein levels is also displayed.


A
#1
#2
#1
#2
siScramble
siLED
Color Key
Row Z-Score
B
FDR 2.78e-8
Normalized p21 abundance
3000
2000
1000
0
siCtrl
siLED
C
Relative p21 mRNA levels
1.2
1
0.8
0.6
0.4
0.2
0
Ctrl
kd1
kd2
LED
D
KDa
MCF-7
20
26
p21
CDK4
siRNA
Ctrl
LED-kd1
LED-kd2
Nutlin-3a

(Figures 3A). Interestingly, LED knockdowns significantly reduced the levels of the cell-cycle regulator p21 (FDR of 2.7e-8) (Figure 3B), without influencing p53 levels (FDR of 0.12). We further validated this observation by showing the LED-dependent regulation of p21 at both the mRNA (Figure 3C) and protein levels (Figure 3D). Altogether, our results indicate that LED is required for an efficient p53-dependent checkpoint by maintaining high levels of p21.

3.3.3 LED fine-tunes gene expression by associating with enhancer domains

Next, to investigate whether LED, a *bona fide* lncRNA of ~4kb, exerts its function in the nucleus or in cytoplasm, we examined its subcellular localization. (Figures 4A-C and S4A). Since LED is partially located in the nucleus we set out to assess its putative interaction with chromatin. We performed chromatin isolation by RNA purification technique (ChIRP) (Chu et al., 2011) using two pools of anti-sense oligos to LED (odd and even) and confirmed the specific enrichment for LED, but not GAPDH RNA (Figure S4B). Then, we sequenced the DNA fragments co-purified in the two pools, aligned reads to the genome, and processed them using peak calling software in a pipeline developed for ChIRP-seq data. Overlap of the peaks from the odd and even purifications indicated LED binding in 1698 putative sites.

To investigate the nature of the genomic features present at LED-associated domains, we made use of chromatin state annotations previously defined by Ernst and colleagues (Ernst et al., 2011). Intriguingly, although LED associated-sites were present in all chromatin states, significant enrichment was observed in strong enhancer regions (Figures 4D and S4C). Moreover, intersection with the GRO-Seq data revealed that a subgroup of LED-bound enhancers was sensitive to nutlin-3a (Figure 4E). This observation suggests that LED might participate in the nutlin-3a-dependent regulation of enhancers.

To assess the regulatory potential of LED on enhancers, we first selected a subset of LED-associated enhancer domains (Figure 4F). Then, we reasserted that bound enhancers harbor hallmarks of active enhancers (Birney et al., 2007; Heintzman et al., 2009; Heintzman et al., 2007a) (Figure S4D). Furthermore, we performed ChIP for H3K4me1 and H3K4me3, to confirm the relative deposition of these histone modifications in our cell system (Figure 4G). Next, we tested a selected group of LED-associated enhancers for eRNA production by qRT-PCR (after DNase-treatment of isolated RNA) on MCF-7 cells incubated with or without nutlin-3a. As previously observed with the GRO-Seq, this analysis confirmed the nutlin-3a-dependent transcriptional induction of eRNAs at all tested LED-associated enhancers (Figure 4H). This nutlin-3a induction of eRNAs was specific, as the abundance of a control, LED-unbound, FOXC1 enhancer (FOXC1e) RNAs remained unaffected. Strikingly, RNAi-mediated LED knockdown reduced the level of activation of these putative eRNAs (Figure 4I), indicating direct regulation of eRNA production by LED. Intriguingly, we noticed among the LED-

associated enhancers a prominent peak located within the first intron of p21.

We further validated the association of LED to p21 enhancer (p21e) domain using ChIRP-qPCR (Figure S4E). To evaluate the enhancing potential of p21e, we cloned a 1.2kb fragment into a pGL3-promoter luciferase reporter vector. As expected from an enhancer domain, p21e activated the luciferase gene in an orientation-independent manner (Figure 4J). Using northern-blotting and RNAPII ChIP experiments, we further supported the presence of an antisense eRNA at p21e locus and its regulation by LED (Figures 4K and 4L). Last, we examined whether p21e could interact with distant promoters by DNA looping, using circular chromosome conformation capture (4C) experiments.

This analysis failed to reveal long-distance enhancer-promoter interactions, suggesting that p21e acts within its functional domain on the p21 promoter (Figure S4F). Collectively, these results demonstrate that LED associates with chromatin regions marked as enhancers and regulates the production of eRNAs.

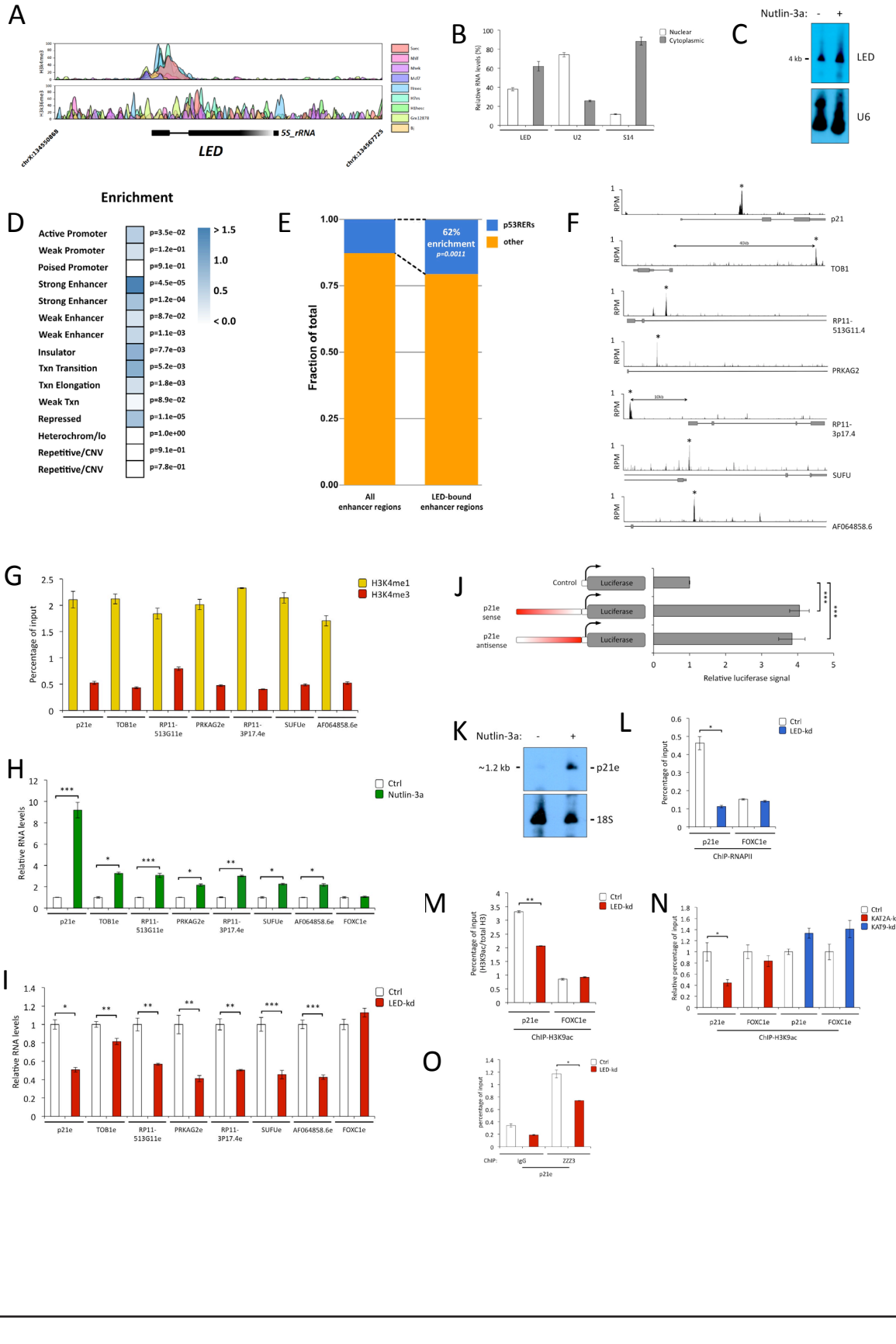
3.3.4 LED influences the epigenetic landscape of bound enhancers

To further delineate the mode of action by which LED regulates enhancers, we hypothesized that LED controls enhancer activity by remodeling the epigenetic state of enhancer domains. To investigate this possibility, we first assessed whether LED influences the deposition of active enhancer histone marks, such as H3K27ac and H3K9ac. ChIP analyses revealed that the levels of H3K9ac, but not H3K27ac, were decreased at p21 enhancer domain upon LED knockdown (Figures 4M and S4G). Similar results were also obtained with another LED-associated enhancer (Figure S4H).

Interestingly, only two histone acetyl transferases (HATs), KAT2A and KAT9, are so far known to mediate H3K9 acetylation (Agalioti et al., 2002; Karam et al., 2010). However, by knocking down each histone acetyl transferase, we observed that only KAT2A-depleted cells had reduced levels of H3K9ac at LED-associated enhancers (Figure 4N and S4I-J). In addition, a recent report demonstrated that KAT2A is a subunit of the histone acetyl transferase ATAC complex, which is found at both promoters and enhancers (Krebs et al., 2011). Thus, we examined if the ATAC complex was present at LED-bound enhancers and whether LED knockdown interferes with its binding to enhancers. By performing a ChIP experiment, using an antibody recognizing ZZZ3 (a subunit of ATAC), we observed the presence of ATAC at LED-bound enhancers and demonstrated that LED knockdown hampers its association (Figure 4O and S4K). These results indicate that LED is required for an efficient deposition of H3K9ac at specific enhancer loci, and suggests that this action is mediated in collaboration with the histone acetyl transferase KAT2A and the complex ATAC.

Figure 4. LED binds preferentially to enhancers and regulates enhancer RNA production from p53RERs

(A) H3K4 trimethylation and H3K36 trimethylation (K4K36 signature) in the region of the LED gene locus for several cell lines, as determined by the ENCODE consortium. (B) LED subcellular localization. Percentage of total RNA found in nuclear and cytoplasmic fractions of nutlin-3a treated MCF-7 cells. U2 and S14 genes were used as positive controls for nuclear and cytoplasmic fraction, respectively. (C) Northern blot analysis showing a ~4 kb transcript (LED) in MCF-7 cells incubated with or without nutlin-3a. U6 was used as a loading control. (D) Enrichment of LED ChIRP peaks in genomic features defined by ENCODE. The different colors represent different levels of enrichment. (E) Bar graph showing the fraction of induced p53RERs of all found putative enhancer regions (left) and of all LED-bound enhancer regions (right). The enrichment of induced p53RERs in the LED-bound fraction is significant with p=0.0011 (hypergeometric distribution). (F) ChIRP sequencing data, showing LED association to a subset of enhancer domains. Values are represented by RPM (reads per million). Star (*) indicates the significant binding site. (G) ChIP-qPCR for total H3, H3K4me1 and H3K4me3 on LED-associated p53RERs in MCF-7 cells. Values were corrected to total H3. Mean ± SD are shown. (H) Stress-dependent regulation of LED-associated p53RERs following nutlin-3a (8 μM) treatment in MCF-7 cells was measured by qRT-PCR. FOXC1e (not associated with LED) was used as a control. Values are represented by fold induction (n=3; p-values were calculated with a two-tailed student's t-test. ***p<0.005, **p<0.01, *p<0.05). (I) LED-dependent regulation of p53RERs (p53BERs and p53FERs) upon nutlin-3a (8 μM) treatment in MCF-7 cells transfected with a control or LED siRNA (LED-kd). FOXC1e was used as a control. Values are represented by fold induction (n=3; p-values were calculated with a two-tailed student's t-test. **p<0.01, *p<0.05). (J) MCF-7 cells were transfected with the indicated reporter constructs, and the relative firefly/renilla luciferase activity was assessed and normalized to the empty promoter vector (Control) (n=3; p-values were calculated with a two-tailed student's t-test. ***p<0.005) (K) Northern blot analysis showing a ~1.2 Kb antisense transcript (p21e) on control and nutlin-3a treated MCF-7 cells. 18S was used as a loading control. (L) ChIP-qPCR for RNA polymerase II (RNAPII) on p21e, and FOXC1e regions in nutlin-3a treated MCF-7 cells transfected with a control or LED siRNA (LED-kd). Values represent the percentage of input (n=3; p-values were calculated with a two-tailed student's t-test. *p<0.05). (M) ChIP-qPCR for total H3 and H3K9Ac on p21e, and FOXC1e regions in nutlin-3a treated MCF-7 cells transfected with a control or LED siRNA (LED-kd). Values represent the percentage of input and were calculated as the ratio between H3K9Ac over total H3 (n=3; p-values were calculated with a two-tailed student's t-test. **p<0.01). (N) ChIP-qPCR for total H3 and H3K9Ac on p21e, and FOXC1e regions in nutlin-3a treated MCF-7 cells transfected with a control or KAT2A and KAT9 siRNA (KAT2A-kd, KAT9-kd). Values represent the percentage of input and were calculated as the ratio between H3K9Ac over total H3 (n=3; p-values were calculated with a two-tailed student's t-test. *p<0.05). (O) ChIP-qPCR for ZZZ3 on p21 enhancer region in nutlin-3a treated MCF-7 cells transfected with a control or LED siRNA. Values represent the percentage of input (n=3; p-values were calculated with a two-tailed student's t-test. *p<0.05).


A
5S_rRNA
LED
B
Nuclear
Cytoplasmic
Relative RNA levels (%)
LED
U2
S14
C
Nutlin-3a: - +
4 kb
LED
U6
D
Enrichment
Active Promoter p=3.5e-02
Weak Promoter p=1.2e-01
Poised Promoter p=9.1e-01
Strong Enhancer p=4.5e-05
Strong Enhancer p=1.2e-04
Weak Enhancer p=8.7e-02
Weak Enhancer p=1.1e-03
Insulator p=7.7e-03
Txn Transition p=5.2e-03
Txn Elongation p=1.8e-03
Weak Txn p=8.9e-02
Repressed p=1.1e-05
Heterochrom/lo p=1.0e+00
Repetitive/CNV p=9.1e-01
Repetitive/CNV p=7.8e-01
> 1.5
1.0
0.5
< 0.0
E
62% enrichment p=0.0011
p53RERs
other
Fraction of total
All enhancer regions
LED-bound enhancer regions
F
p21
TOB1
RP11-513G11.4
PRKAG2
RP11-3p17.4
SUFU
AF064858.6
RPM
G
H3K4me1
H3K4me3
Percentage of input
p21e
TOB1e
RP11-513G11e
PRKAG2e
RP11-3P17.4e
SUFUe
AF064858.6e
H
Ctrl
Nutlin-3a
Relative RNA levels
FOXC1e
I
Ctrl
LED-kd
J
Control
p21e sense
p21e antisense
Luciferase
Relative luciferase signal
K
Nutlin-3a: - +
~1.2 kb
p21e
18S
L
Ctrl
LED-kd
Percentage of input
ChIP-RNAPII
M
Percentage of input (H3K9ac/total H3)
ChIP-H3K9ac
N
KAT2A-kd
KAT9-kd
Relative percentage of input
O
ChIP:
IgG
ZZZ3

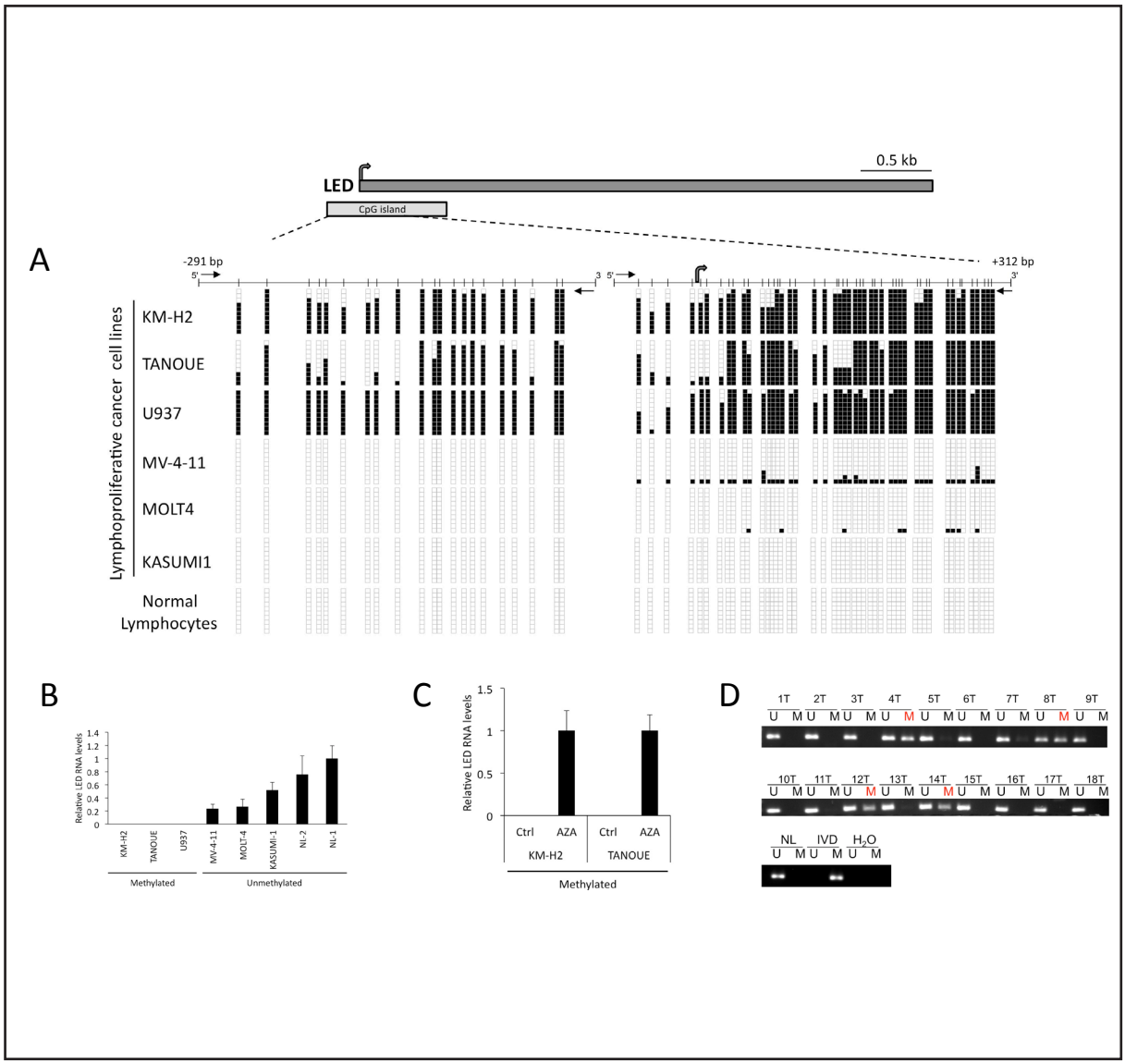


Figure 5. DNA methylation-associated silencing of LED in lymphoproliferative tumors.
(A) Schematic representation of LED genomic loci and CpG island. Bisulfite genomic sequencing analysis of LED CpG island in human lymphoproliferative cancer cell lines and normal lymphocytes as tissue control. Location of bisulfite genomic sequencing PCR primers (black arrows), CpG dinucleotides (vertical lines) and the transcriptional start site (grey arrow) are shown. Ten single clones are represented for each sample. Presence of unmethylated or methylated CpGs is indicated by white or black squares, respectively. (B) LED expression levels in methylated or unmethylated human lymphoproliferative cell lines and in normal lymphocytes (NL) as control. Values were determined by qRT-PCR in triplicates and are expressed as mean ± sem (n=2-4). (C) Restored LED expression after treatment with DNA demethylating agent 5-aza-2′-deoxycytidine (AZA) in LED CpG island methylated cell lines. Values were determined in triplicates by qRT-PCR and are expressed as mean ± s.e.m (n=3). (D) Methylation-specific PCR analyses for LED methylation in primary leukemias. The presence of a band under the U lane indicates unmethylated alleles, while the presence of a band under the M lane indicates methylated alleles. Normal lymphocytes and in vitro methylated DNA (IVD) are shown as negative and positive controls for methylated alleles, respectively.

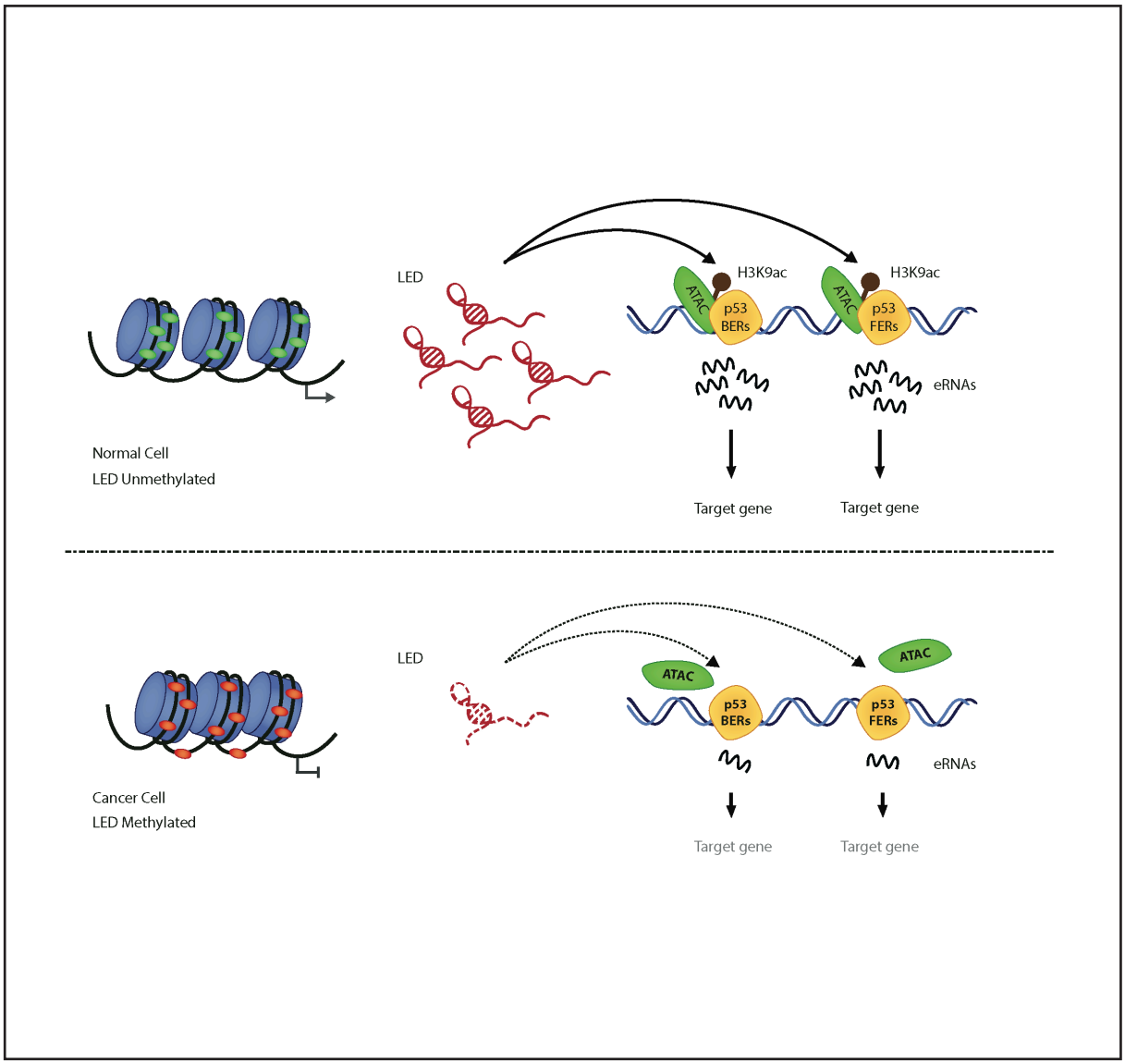


Figure 6. Schematic representation of LED function in normal and cancer cell.
P53-bound enhancer regions (p53BERs), p53-free enhancer regions (p53FERs), ATAC complex and H3K9 acetylation are displayed.

3.3.5 LED is inactivated by promoter hypermethylation in cancer

Gene expression comparison analysis suggests not only that LED is activated by p53, but also that its function is intimately linked to the transcriptional response of p53. We therefore examined whether LED is inactivated in cancer. Inspection of the LED promoter sequence identified a large CpG island region (Figure 5A). As CpG islands are often subject to hypermethylation and silencing, we asked whether LED promoter hypermethylation leads to a reduced LED expression in cancers. We first measured the methylation status of LED CpG islands in 135 cancer cell lines covering a wide range of cancers. Notably, we find LED promoter methylation in ~44 % (59/135) of all tested cell lines, with a large proportion in Leukemia (Figures S5A-C). Moreover, we observed a strong preference for methylation in p53 wild-type (wt) cell lines (60 %; 29/48) as compared to p53 mutants (34 %; 30/87, p = 0.004 [Chi-square]) (Figure 5A, Table 1A). Most importantly, we then assessed and validated the transcriptional silencing of LED by its promoter-associated hypermethylation on several cancer cell lines. As expected, there was a significant anti-correlation between LED expression and its methylation status (Figures 5B and S5D-E). Also, treatment of LED-promoter-hypermethylated cell lines with the DNA-demethylating agent 5-AZAcytidine resulted in LED re-expression (Figure 5C). Finally, we evaluated the prevalence of LED-promoter hypermethylation in various human tumors. Using methylation-specific PCR (MSP), we observed LED-promoter hypermethylation in various tumor-types, most prominently reaching 22 % of all samples in Acute Lymphocytic Leukemia (Figure 5D and Table 1B).

3.4 Discussion

Coordination of gene expression components within response programs is a delicate task crucial for the maintenance of cellular homeostasis. One key player for such coordination is the tumor suppressor p53, which organizes the implementation of an appropriate cellular response to stress cues such as DNA damage and emerging oncogenes. With the discovery that the genome is pervasively transcribed (Core et al., 2008), it is likely that novel p53-sensitive transcripts and regulatory networks will be uncovered. While many target promoters of p53 are well-established, little is known about the role of this master tumor suppressor as enhancer factor.

3.4.1 p53 role as an enhancer factor

Enhancer RNAs (eRNAs) were recently suggested as transcriptional regulators (Hah et al., 2013a; Melo et al., 2013b). Moreover, eRNA level emerges as robust readout for determining

A

Association of LED CpG island hypermethylation with TP53 mutational status in human cancer cell lines.

Methylation	TP53 wt	TP53 mut
Methylated	29 (60%)	30 (34%)
Unmethylated	19 (40%)	57 (66%)
Total (135)	48	87

B

Frequency of LED promoter associated-hypermethylation (M) in cancer patients

Tumor type	Methylation	Total	%
Acute Lymphocytic Leukemia	21	95	22
Cutaneous Lymphomas	1	6	16
Folicular Lymphomas	1	10	10
Melanomas	4	46	9
Acute Myeloid Leukemia	1	13	8
Chronic Lymphocytic Leukemia	1	33	3

Table 1A. Association of LED CpG island hypermethylation with TP53 mutational status in human cancer cell lines.

Table 1B. Frequency of LED promoter associated-hypermethylation (M) in cancer patients.

enhancer activity, as it correlates with the levels of neighboring target genes. The GRO-seq is a very powerful technique to globally measure newly synthesized eRNAs and to infer enhancer activity in a genome-wide manner. Here we used GRO-seq to map and quantify eRNAs induced by the p53 inducer nutlin-3a, and identified hundreds of regulated enhancer domains. While many enhancers are direct targets of p53, most nutlin-3a regulated enhancer domains were not bound by this transcription factor. Thus, it is likely that those enhancers are indirectly induced by regulators controlled by p53. However, bioinformatics tools failed to identify prominent binding sites of such regulators within p53-free enhancer regions. Interestingly, this group of enhancers shows a reduced "sharpness" of transcription compared to p53-bound enhancers, pointing to a different mode of activation. Indeed, at least for a subset of these enhancers we demonstrate direct activation by a long non-coding RNA induced by p53. Whether a more "diffused" enhancer transcription depends on the presence of the lncRNA alone remains to be pursued. However, in support of our finding, two recent studies reported that not only transcription factors but also *trans* acting lncRNAs are present at transcriptional regulatory regions (Vance et al., 2014a; Yang et al., 2013). For instance, the lncRNAs PCGEM1 and PRNCR1 were found to interact with the androgen receptor (AR) at enhancer domains and to mediate enhancer-promoter interaction.

Thus, for the first time, we demonstrate the contribution of a p53-induced lncRNA, termed LED, in the regulation of enhancer-derived transcripts.

3.4.2 Role of LED lncRNA as epigenetic modifier and activator of enhancers

LED is a direct transcriptional target of p53. Suppression of LED expression attenuated the activation of target enhancer domains, as demonstrated by reduced eRNA production and by a lower H3K9 acetylation. We found that LED was associated with different genomic loci and especially enriched at enhancer domains producing eRNAs. Notably, a subgroup of these enhancers is regulated by p53. Moreover, some, but not all, LED-bound enhancers were concomitantly bound by p53. Despite this observation, all tested LED-bound p53-induced eRNAs responded to siRNA-mediated LED depletion. This suggests that LED is a p53-induced factor able to convey p53-dependency in the absence of p53 binding. How exactly does LED trigger enhancer activation? Modulation of the chromatin epigenetic state plays an important role in the regulation of gene expression. Thus far, several studies have put forward the idea that lncRNAs are important epigenetic regulators. For example, the lncRNA HOTAIR represses gene expression by interacting with and guiding the polycomb repressive complex 2 (PRC2) to target promoters, where it contributes to chromatin compaction by catalyzing the methylation of histone H3 at lysine 27 (Gupta et al., 2010). Alternatively, ribonucleoprotein complexes such as HOTTIP:MLL/WDR5 activate gene expression by promoting the deposition of an active mark (H3K4me3) on promoters (Wang et al., 2011b). Here we complement these

observations by showing that LED is essential for the acetylation of H3K9 at bound enhancers, a modification associated with active gene transcription. This finding is consistent with the fact that LED is required for RNAPII loading and eRNA transcription at bound enhancers. Moreover, genome-wide deposition of H3K9 acetylation was previously reported to be enriched at regulatory elements such as promoters, enhancers and repetitive sequences (Ernst et al., 2011). Consequently, active transcriptional programs may primarily be epigenetically governed by the action of a subset of activating lncRNAs. Importantly, siRNA-mediated depletion of the histone acetyl transferase KAT2A reduced the levels of H3K9 acetylation at LED-bound enhancer regions. In addition, KAT2A was, similarly to LED, essential for an efficient RNAPII loading at enhancers. Interestingly, the interplay between KAT2A and the p53 pathway has already been established by previous work (Gamper and Roeder, 2008). In metazoan, KAT2A operates via two distinct histone acetyltransferases (HAT) complexes: SAGA and ATAC (Krebs et al., 2011). While SAGA mainly functions at promoters, ATAC acts at both promoters and enhancers. However, only its association to enhancer regions appears to be cell-type specific. This characteristic is reminiscent of a process driven by non-coding RNAs, as their expression is renowned to be cell-type specific. Indeed, our results demonstrate that LED is essential for the association of ZZZ3 protein, a subunit of ATAC, to LED-bound enhancers. These data suggest that LED guides or helps the association of KAT2A/ATAC to certain enhancers following p53 activation. However, further experiments will be required to clarify the underlying molecular mechanism.

3.4.3 Dysregulation of lncRNAs in cancer

P53 is one of the most commonly inactivated gene in human cancer, with somatic mutations occurring in approximately half of all human cancers (Hollstein et al., 1991). In addition, alterations in the p53 pathway often represent an alternative route to attenuate the function of wild-type p53 in tumor (Leveille et al., 2011; Voorhoeve et al., 2006). Here we demonstrate that LED lncRNA is largely silenced in p53 wild-type primary human acute lymphocytic leukemia (ALL). This important observation pinpoints the dysregulation of lncRNAs as a potent mechanism in tumorigenesis. In support of this concept, other lncRNAs have been linked with cancer. For example, the oncogenic lncRNA *HOTAIR* is highly expressed in breast tumors and promotes cancer metastasis by guiding PRC2 to specific genomic loci (Gupta et al., 2010). The lncRNA ANRIL and SChLAP1 are overexpressed in prostate cancers and antagonize the tumor-suppressive activity of INK4a/b and SWI/SNF complex, respectively (Prensner et al., 2013; Yap et al., 2010). Finally, tumor-suppressive lncRNAs such as GAS5 have been shown to be down-regulated in cancer (Mourtada-Maarabouni et al., 2009). Collectively, our results highlight a novel tumor suppressive mechanism involving a p53-induced lncRNA acting on enhancers (Figure 6). The existence of a crosstalk between different lncRNA species uncovers

an emerging regulatory network with potential considerable impacts in cancer development.

3.5 Experimental Procedures

Analysis of GRO-seq data and determination of enhancer regions

GRO-seq protocol was performed as previously described (Wang et al., 2011a). GRO-seq data was aligned to hg19 (including unassembled contigs) using bowtie2 (v. 2.0.6) with parameters "--seed 42 --sensitive". Alignments with mapping quality lower than 10 and non-primary alignments were not considered in further analyses. Broad ChromHMM data for 9 cell lines (Ernst et al., 2011) was used to screen putative enhancer regions. Along each chromosome, positions that were marked as enhancer regions (feature IDs 4, 5, 6 and 7) in at least one cell line were merged into regions. Transcription start sites, as annotated by RefSeq (obtained from the UCSC database server 9th of august 2012) and GENCODE v19/BASIC were extended by 1000 bases and used to blacklist positions (i.e. those positions were excluded as putative enhancer regions). Each merged region was tested for the presence of p300 and Pol2 as determined by ENCODE in MCF-7 cells (accessions GSM822295 and GSM1010800 (Johnson et al., 2007; Lee et al., 2012)) and each region without p300 and Pol2 peak was removed. The remaining regions were considered putative enhancer regions with enhancer marks. GRO-seq counts were obtained for each region, after extending each region by 1kb. Regions having detectable transcription on both strands were considered putative enhancer regions with bidirectional transcription and used for downstream analyses. edgeR (Robinson et al. 2010) was used to determine statistical significance of differential expression of the enhancer regions (separately for each strand).

Generation of omnibus annotation

Ensembl annotations (v37.65), RefSeq gene annotations (obtained from UCSC database server 9th of August 2012) and the Broad Linc RNA catalog(Cabili et al., 2011) were merged in a single GTF (annotation file) using the gffread utility supplied with the Cufflinks package (v. 1.3.0), using parameters "-M -K -F -G". This essentially collapses overlapping exons/transcripts so as to create an omnibus with low degree of redundancy yet high coverage of known and novel transcripts.

Analysis of RNA-seq data

RNA-seq samples were processed according to Illumina's protocol. Raw RNA-seq data was aligned using TopHat2 (v. 2.0.3) (Trapnell et al., 2009), using parameters "-m1 -p4

-F0.0 --segment-length 21 --segment-mismatches 1" and an exon annotation GTF file that was generated as described before. Reads with mapping quality less than 10 and non-primary alignments were discarded. Remaining reads were counted using HTSeq-count (http://www-huber.embl.de/users/anders/HTSeq/doc/overview.html), per gene ID. Statistical analysis of the differential expression of genes was performed using edgeR (Robinson et al., 2010). Genes with FDR for differential expression lower than 0.01 were considered significant.

Analysis of p53 ChIP data

p53 ChIP-seq data obtained from MCF-7 cells upon untreated and p53-stimulated conditions was obtained from SRA project SRP007261. Alignment was done using bowtie2 (v. 2.0.6) with parameters "--seed 42 --sensitive" to hg19 (including unassembled contigs). Only primary alignments with quality of at least 10 were kept. Peaks were called by MACS (v. 2.0.10.20130501) using default parameters. Peaks with a fold-change (w.r.t. input) lower than 3.0 or a -log10(q-value) lower than 2.0 were discarded.

Sequence motif enrichment analysis

For enhancer regions the midpoint of bidirectional transcription was established, after pooling the GRO-seq data of the nutlin-3a-treated replicates. 200 bases of DNA around the midpoint of bidirectional transcription were extracted. For the analysis of sequence enrichment of BERs and FERs, the backgrounds were enhancers with bidirectional transcription that had an FDR for differential expression of 0.75 or greater (for both strands). Additionally, for FERs the background contained only enhancers that had no p53 peak within 1kb. HOMER (Heinz et al., 2010b) was used to search for overrepresented sequence motifs, using parameters "-nogo -nlen 4 -len 18 -S 5 -mis 2".

Constructs

p21e domain sense and antisense were PCR amplified using gDNA derived from MCF-7 cell lines and subsequently cloned (Nhe1/Xho1) into pGL3-promoter luciferase reporter vector. All primers used are listed in Supplementary table 1.

Cell Culture and Transfection

MCF-7 cells were cultured in DMEM containing 10 % FBS, penicillin and streptomycin at 37°C and 5 % CO2. Identification and validation of lncRNA regulated by nutlin-3a was carried out by treating MCF-7 cells with 2 μM to 8 μM of nutlin-3a for a period of 16 h. To induce a p53 stress-response cells were also treated with 10 Gy of ionizing radiation for 16 h. RNAi

experiments were performed using Dharmafect transfection reagent-1 and between 10 nM to 100 nM of small interfering RNA (siRNA). For epigenetic study, cells were treated with 2 μM 5-aza-2'-deoxycytidine (A3656, Sigma) for 72 hours.

Protein analysis

Whole-cell lysates were prepared as previously described (Leveille et al., 2011). Protein detection was performed using primary antibodies detecting p53 (DO1, Santa Cruz, 1:1000), p21 (Sc-397, Santa Cruz, 1:1000), CDK4 (Sc-260, Santa Cruz, 1:1000), phospho-histone H3 (ser 10) (9701, cell signaling, 1:100). Proteins were visualized using adequate secondary antibody (Dako) and ECL reagents (GE Healthcare).

Quantitative Real-Time PCR

RNA isolation and cDNA preparation were carried out as previously described (Melo et al., 2013b). Real-time quantitative (q)PCR analysis was performed using the LightCycler® 480 SYBR Green I Master mix (Roche). The GAPDH was used as an internal control.

Flow cytometry

Control or nutlin-3a treated cells were arrested in mitosis using 250 ng/ml nocodazole for 24h to 36h. Cells were then trypsinized, washed and resuspended in PBS containing 0.6 % NP-40, 50 mg/ml RNaseA and 50 mg/ml propidium iodide for 10min. Cell cycle profiles were captured using FACSCalibur (BD Biosciences) and analyzed with the Flowjo software.

Immunofluorescence

Cells were first plated at a density of 3 x 10^5 cells/well and concomitantly reverse transfected with a control siRNA or siRNAs against LED or p21. After 24 h, cells were trypsinized and seeded on microscope cover slips coated with polylysine. Next, cells were fixed with 3 % formaldehyde and subsequently permeabilized with PBS-Triton X-100 (0.3 %) solution. After blocking 1h with 2% PBS-Milk, cells were successively incubated with the primary antibody phospho-histone H3 (ser 10) and the Alexa Fluor 488 Dye-conjugated secondary antibody. Images were captured using an AxioCam MRc CCD camera (Carl Zeiss Microimaging).

RNA fluorescence In situ hybridization

MCF-7 cells treated with nutlin-3a (8 μM) and non-treated controls were grown on coverslips in 6 well plates overnight. The media was aspirated and cells washed 3x in cold PBS.

Fixation solution (5 mL 10x PBS, 5 mL 37 % formaldehyde (100% formalin) and 40 mL DEPC H2O) was added and cells were incubated for 20 minutes at 4°C. Cells were washed 3x in cold PBS and 70 % cold ethanol was used to permeabilize cells at 4°C for 24 h. Cells were washed with cold PBS and left in hybridization buffer (1 g dextran sulfate, 7 ml DEPC water, 1 ml formamide and 1 ml 20x SSC buffer) for 1 h. 50 ng of stellaris probe was used in hybridization buffer and cells were kept incubating for 48 h at 37°C. After washing with wash SSC buffer coverslips were covered with Draq5 during 20 minutes for nuclear staining, washed with cold PBS and mounted using antifade buffer (850 µl DEPC H2O, 100 µl 20x SSC, 40 µl 10 % glucose, 10 µl Tris). Images were captured in a Zeiss confocal microscope.

ChIRP

ChIRP was performed as previously described (Chu et al., 2012). ChIRP probes (48 x 40-mer) targeting LED were designed at http://www.singlemoleculefish.com/designer.html. Probes were divided into 2 sets (odd and even). The input, and odd and even probe samples were sequenced individually. After clipping of adapters from the obtained reads, data was aligned to hg19 using bowtie2 (v. 2.0.6) using parameters "--seed 42 -N 1 -p 2". Reads with mapping quality less than 10 and non-primary alignments were excluded from further analysis. Peak detection was run using MACS (v. 2.0.10.20130501) (Zhang et al., 2008) using parameters " -g hs -B -p 0.1". Peaks with a -log10(q-value) >= 5 and an enrichment >= 4 with respect to the input were kept, and peaks found in the odd and even samples were intersected. Overlapping peaks in both samples that had a position-wise Pearson correlation of abundance >= 0.2 and at least 25 reads in both samples were merged. From the resulting set of peaks, plasmid contaminants were discarded (Supplementary table 2).

ChIP

MCF-7 (5×10^6) cells were first transfected with a control or specific siRNAs. Next, cells were fixed with 1% Formaldehyde for 8 minutes (min) at room temperature and subsequently quenched with 125 mM glycine for 5 min on ice. The cells were pelleted 10 min at 470g and re-suspended in 300 µl of cold lysis buffer (50 mM tris-HCl pH 8.0, 10 mM EDTA and 1% SDS) supplemented with protease inhibitor cocktail (Roche). The suspension was sonicated 20 min (30 sec on/off at maximum power) and further diluted with 800 ul of dilution buffer (10 mM tris-HCl pH 7.5, 140 mM NaCl, 1 mM EDTA, 0,5 mM EGTA and 1 % triton-X100) supplemented with protease inhibitor cocktail. The lysate was centrifuged 10 min at maximum speed and the soluble fraction (chromatin) was transfer to a new tube. For each ChIP reaction, 100 µl of chromatin preparation was diluted with 300 µl of dilution buffer and incubated on an end-to-end rotator with 2 to 10 µg of antibody at 4°C overnight. Then, 30 µl of protein A/G beads, previously blocked 1h with PBS/BSA (0.1 %) solution, was add to each ChIP reaction

and incubated 2h to 3h at 4°C. The immune-purified chromatin was washed 2 x 5 min with the dilution buffer and 1 x with TE (50 mM tris-HCl pH 8.0 and 10 mM EDTA) and finally eluted in 300 µl elution buffer (20 mM tris-HCl pH 7.5, 5 mM EDTA, 50 mM NaCl and 1 % SDS) at 65°C overnight. Eluted chromatin was purified using QIAquick PCR purification kit (Qiagen) and subjected to real-time PCR analysis. Antibodies and amounts used in this study: pol2 (8 µg, CTD4H8, Upstate), H3K9ac (3 µl, ab4441, abcam), H3K27ac (3 µl, ab4729, abcam), H3K4me1 (6 µl, ab8895, abcam), H3K4me3 (3 µl, MC315, Upstate), histone H3 (5 µl, 2650, Cell Signaling), ZZZ3 (15 µl, kindly provided by Dr. Làszlò Tora).

Chromosome Conformation Capture on Chip (4C)

4C templates preparation and analysis were performed as previously described (Splinter et al., 2011). Briefly, 10^7 of MCF-7 cells were harvested and crosslinked with formaldehyde 2% for 10 min at room temperature, and neutralized with 125 M glycine. After washing with PBS, cells were lysed in 150 mM NaCl, 50 mM Tris-HCl (ph 7.5), 5 mM EDTA, 0.5 % NP40, 1 % Triton X-100 and nuclei were recovered by spinning 8 min at 600g. Nuclei were digested overnight with 400 U DpnII (NEB) and Csp6I (NEB) and re-ligated in 7ml with 100 U T4 DNA ligase (Roche) overnight at 16°C. Purified DNA circles were digested with 50 U DpnII (Csp6I circles) and Csp6I (DpnII circles) overnight at 37°C, followed by heat inactivation and ligation in 14 ml with 200 U T4 DNA ligase. The 4C template was then purified and used for PCR amplification.

DNA methylation analysis

The Methyl Primer Express v1.0 software was used to identify the CpG islands and design specific primers for the methylation analysis (Supplementary Table). DNA methylation status was established by bisulfite genomic sequencing of multiple clones or methylation-specific PCRs in DNA samples previously treated with sodium bisulfite (EZ DNA methylation Gold kit, Zymo Research). The Illumina 450K methylation array was used to anlyze the methylation status in multiple human cancer cell lines. For epigenetic drug treatments, cells were treated with 1 µM 5-aza-2′-deoxycytidine (Sigma, St Louis, MO, USA) for 72 h.

3.6 Supplemental Information

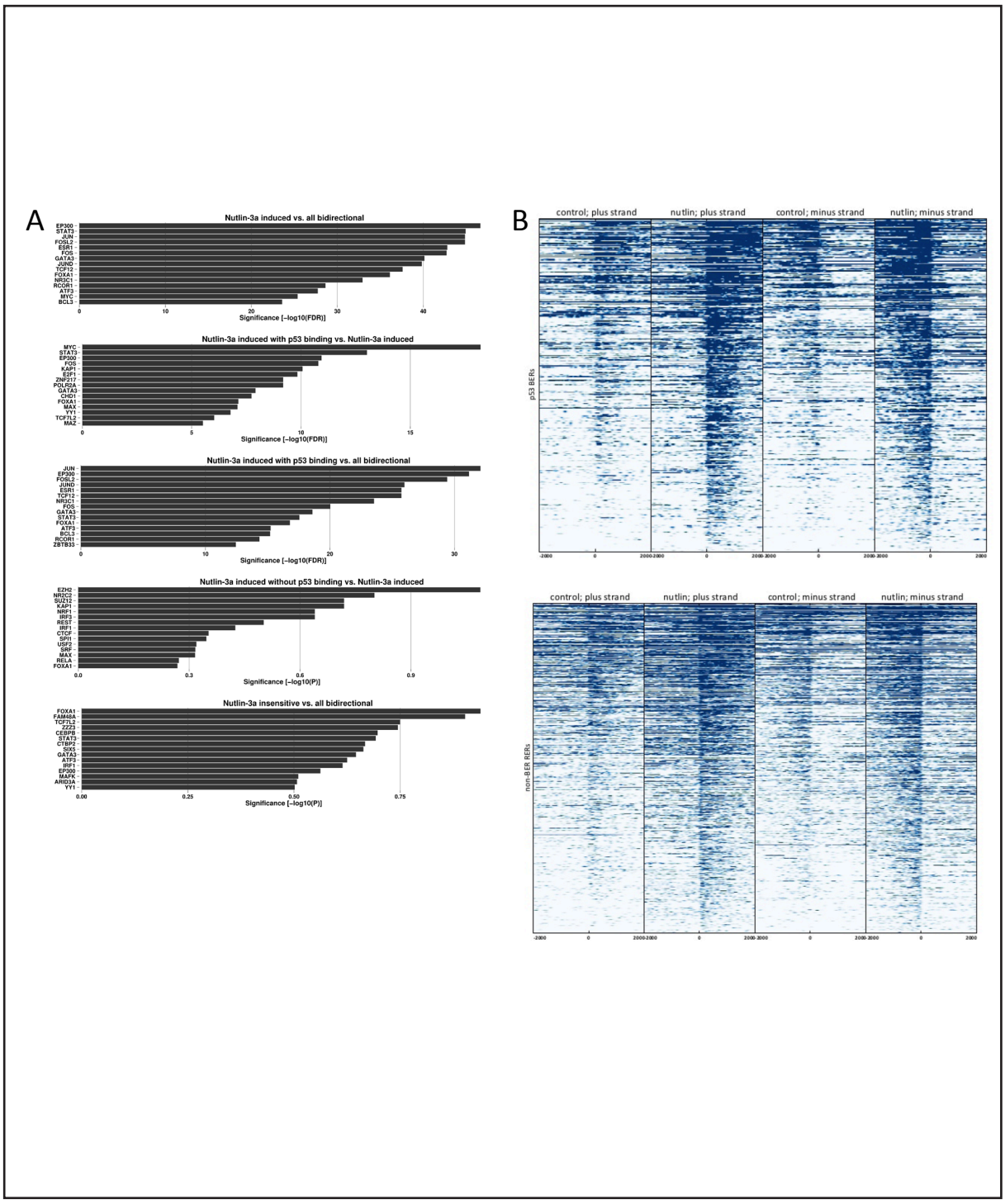


Figure S1 (related to Figure 1). Identification of p53 regulated enhancer RNAs (p53RERs) (A) Top 15 transcription factor enrichments. Top to bottom: p53RERs vs all bidirectional enhancers, p53BERs vs. p53RERs, p53BERs vs. all bidirectional, p53FERs vs. p53RERs, nutlin-3a-insensitive vs. all bidirectional (as negative control). Note that in the last two panels the significance is indicated by uncorrected p-value, while in the other panels the FDR is used. (B) GRO-seq read densities centered at the point of bidirectional transcription for p53BERs (upper panel) and p53FERs (lower panel).

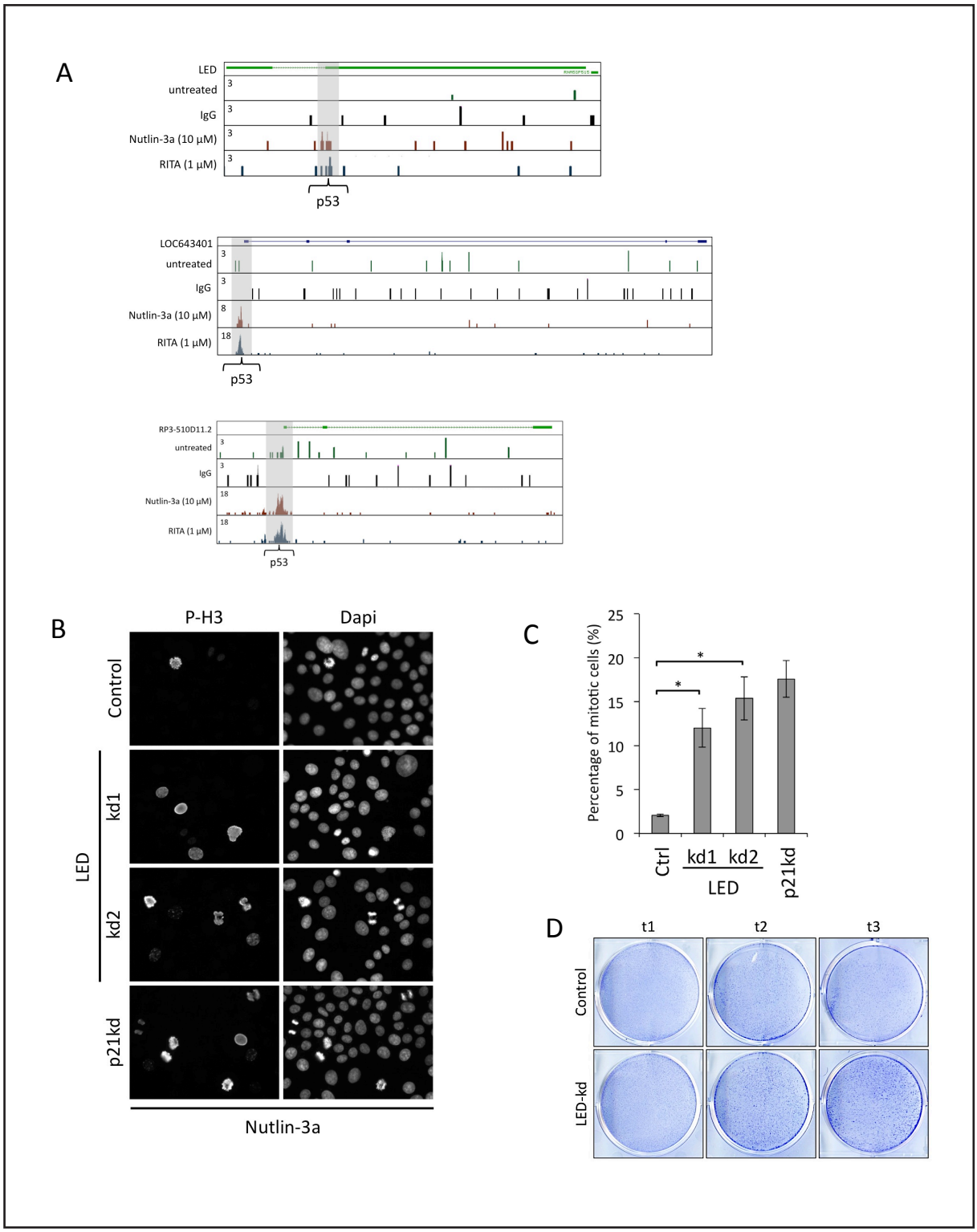
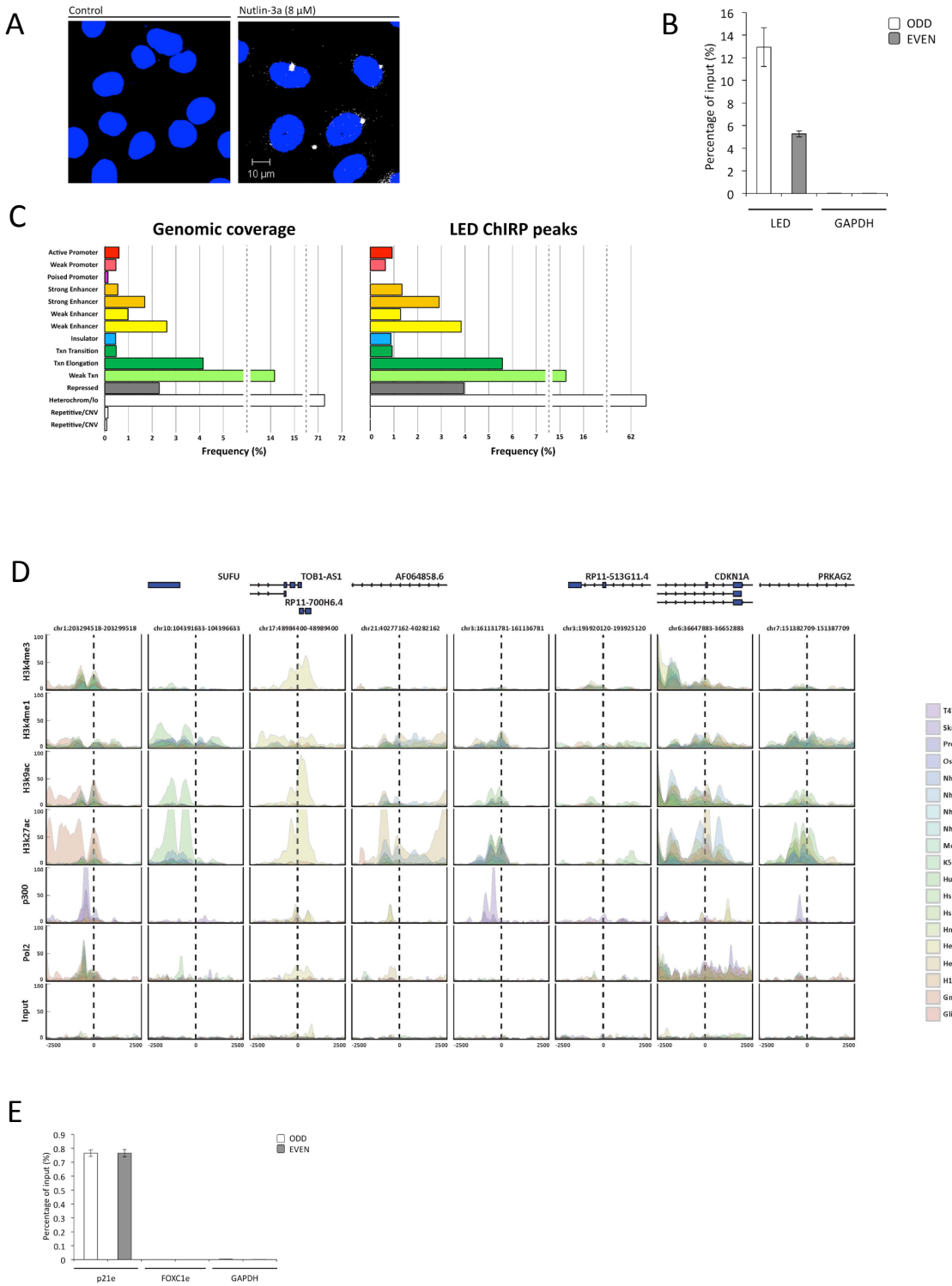


Figure S2 (related to Figure 2). Novel stress-regulated lncRNA LED

(A) P53 or IgG ChIP in MCF-7 cells treated with or without nutlin-3a (8 μM) and RITA (1 μM) treatments. The binding of p53 within LED gene body is displayed. (B) Immunostaining detection of the mitotic marker phospho-histone H3 ser10 (P-H3) and DAPI staining in MCF-7 cells treated as in (Figure 2G) and using a siRNA targeting p21 (p21kd) as a positive control. (C) Quantification of the marker P-H3 from (A) indicating the percentage of mitotic cells. (n=3; p-values were calculated with a two-tailed student's t-test. *p<0.05). (D) Colony formation assay using the same condition as in (Figure 2H).


A
Control
Nutlin-3a (8 μM)
10 μm
B
Percentage of input (%)
ODD
EVEN
LED
GAPDH
C
Genomic coverage
LED ChIRP peaks
Active Promoter
Weak Promoter
Poised Promoter
Strong Enhancer
Strong Enhancer
Weak Enhancer
Weak Enhancer
Insulator
Txn Transition
Txn Elongation
Weak Txn
Repressed
Heterochrom/lo
Repetitive/CNV
Repetitive/CNV
Frequency (%)
Frequency (%)
D
SUFU
TOB1-AS1
AF064858.6
RP11-700H6.4
RP11-513G11.4
CDKN1A
PRKAG2
H3k4me3
H3k4me1
H3k9ac
H3k27ac
p300
Pol2
Input
T47d
Sknshra
Progfib
Osteobl
Nhlf
Nhek
Nhdfad
Nha
Mcf7
K562
Huvec
Hsmmt
Hsmm
Hmec
Hepg2
Helas3
H1hesc
Gm12878
Gliobla
E
Percentage of input (%)
ODD
EVEN
p21e
FOXC1e
GAPDH

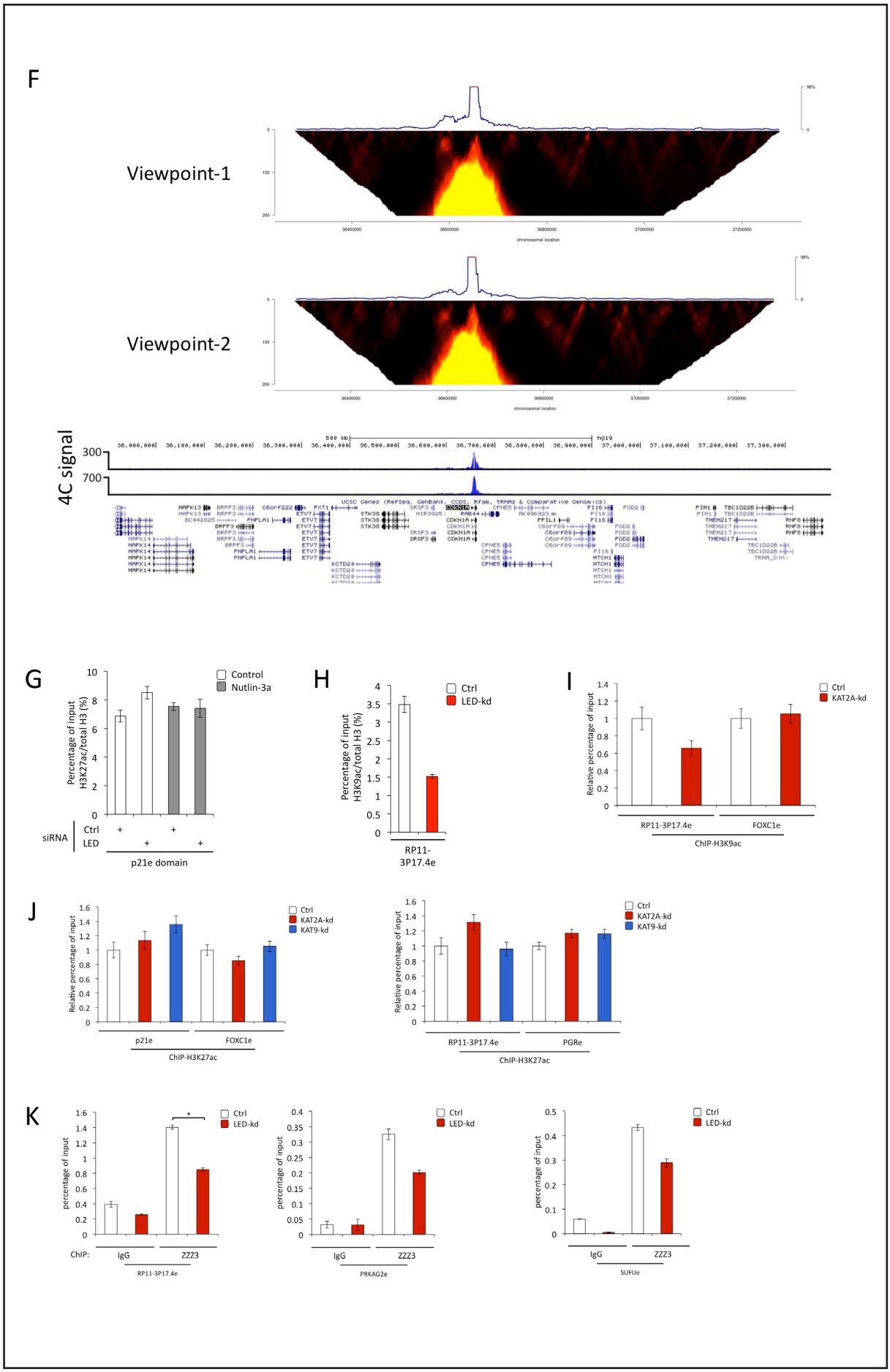
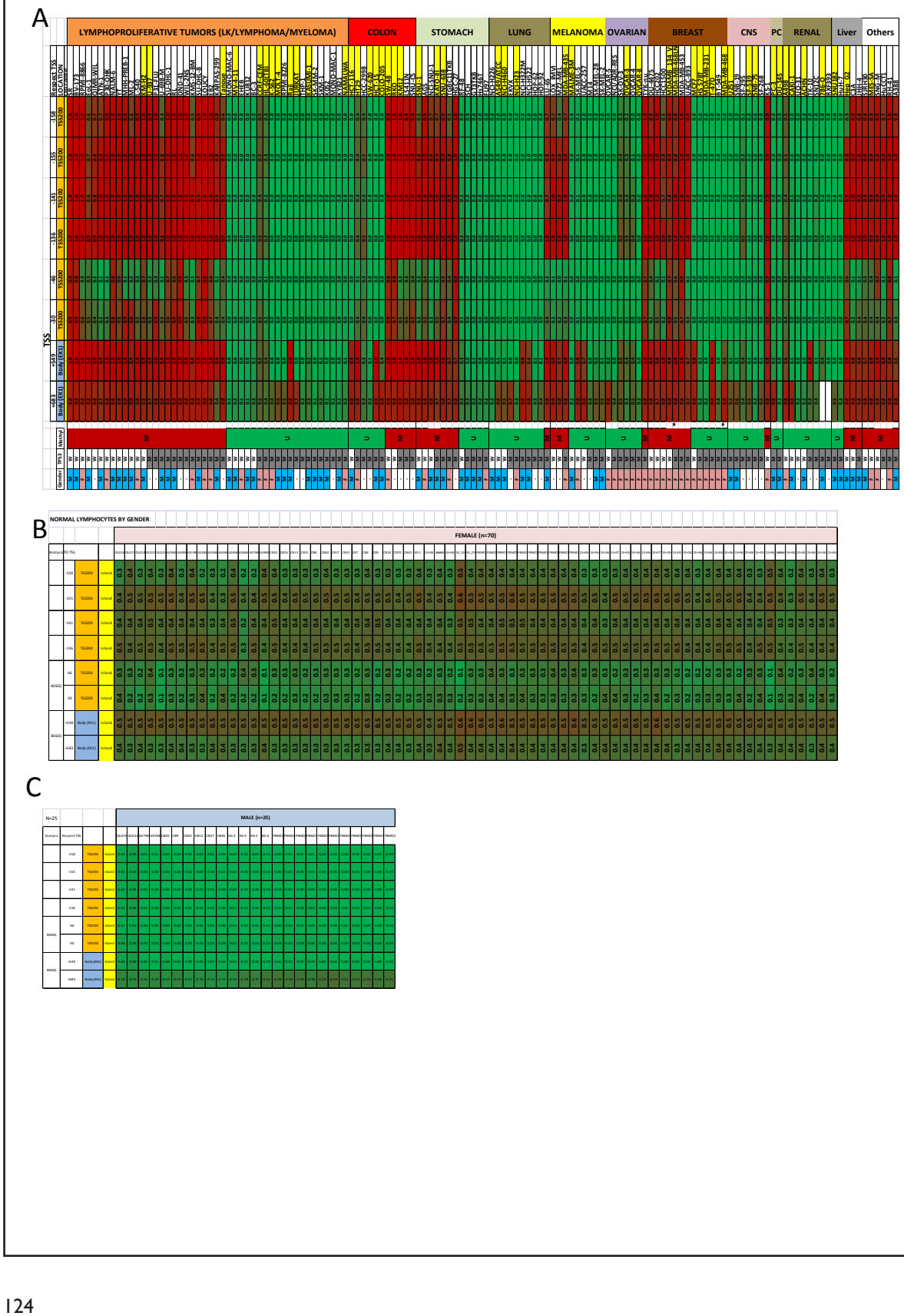

F
Viewpoint-1
Viewpoint-2
4C signal
300
700
G
Control
Nutlin-3a
Percentage of input H3K27ac/total H3 (%)
siRNA
Ctrl
LED
p21e domain
H
Ctrl
LED-kd
Percentage of input H3K9ac/total H3 (%)
RP11-3P17.4e
I
Ctrl
KAT2A-kd
Relative percentage of input
RP11-3P17.4e
FOXC1e
ChIP-H3K9ac
J
Ctrl
KAT2A-kd
KAT9-kd
Relative percentage of input
p21e
FOXC1e
ChIP-H3K27ac
RP11-3P17.4e
PGRe
ChIP-H3K27ac
K
Ctrl
LED-kd
percentage of input
ChIP:
IgG
ZZZ3
RP11-3P17.4e
PRKAG2e
SUFUe

Figure S4 (related to Figure 4). LED binds preferentially to enhancers and regulates enhancer RNA production from p53RERs.

(A) RNA Fluorescence *in situ* hybridization (FISH), with probes against LED in MCF-7 cells incubated with or without nutlin-3a. (B) ChIRP enrichment for LED with ODD and EVEN probes tiling LED transcript. GAPDH was used as a control (mean ± SD). (C) Distribution of the ENCODE features over the genome (left panel) and LED ChIRP peaks in the different ENCODE features (right panel). (D) Levels of several histone marks (H3K27Ac, H3K4me1, H3K4me3 and H3K9Ac), p300 and RNAPII in different cell lines in the loci (+/- 5 kb) of selected LED-induced enhancers, as determined by the ENCODE consortium. (E) DNA ChIRP-qPCR enrichment for p21e, but not for FOXC1e or GAPDH, using ODD and EVEN LED-tiling probes in MCF-7 cells. Values are represented as percentage of input. Mean ± SD are shown. (F) Domainograms (de Wit et al., 2008) visualizing significance of interactions at different window sizes (viewpoints) for p21 enhancer domain with surrounding chromosomal regions in MCF-7 cells. Color ranges (see scale bar) reflect different levels of significance, from black (low significance, *P*=1x10-2) to yellow (high significance, P=1x10-8). To account for the fact that the majority of the data are very close to the viewpoint, we set the data range of the vertical axis to the 98% quantile value for the analyzed region. Values are represented by intensity of 4C signal. Schematic representation of p21 gene with the location of the bait used for the 4C experiments is also shown. (G) ChIP-qPCR for total H3 and H3K27Ac on p21 enhancer domain (p21e) in control and nutlin-3a treated MCF-7 cells transfected with a control or LED siRNAs (LED-kd). Values represent the percentage of input and were calculated as the ratio between H3K27Ac over total H3. Mean ± SD are shown. (H) ChIP-qPCR for total H3 and H3K9Ac on RP11-3P17.4e in nutlin-3a treated MCF-7 cells transfected with a control or LED siRNA (LED-kd). Values represent the percentage of input and were calculated as the ratio between H3K9Ac over total H3. Mean ± SD are shown. (I) ChIP-qPCR performed as in (I) but with a control or KAT2A siRNA (KAT2A-kd). (J) ChIP-qPCR for total H3, H3K9Ac and H3K27ac on p21e and RP11-3P17.4e in nutlin-3a treated MCF-7 cells transfected with a control, KAT2A or KAT9 siRNA. Values represent the percentage of input and were calculated as the ratio between H3K9Ac or H3K27ac over total H3. FOXC1e and PGRe enhancer RNAs were used as a control. Mean ± SD are shown. (K) ChIP-qPCR for ZZZ3 on RP11-3P17.4, PRKAG2 and SUFU enhancer regions in nutlin-3a treated MCF-7 cells transfected with a control LED siRNA. Values represent the percentage of input. Mean ± SD are shown. (n=3; p-values were calculated with a two-tailed student's t-test. *p<0.05).


A
LYMPHOPROLIFERATIVE TUMORS (LK/LYMPHOMA/MYELOMA)
COLON
STOMACH
LUNG
MELANOMA
OVARIAN
BREAST
CNS
PC
RENAL
Liver
Others
TSS
Normal
B
NORMAL LYMPHOCYTES BY GENDER
FEMALE (n=70)
C
MALE (n=25)

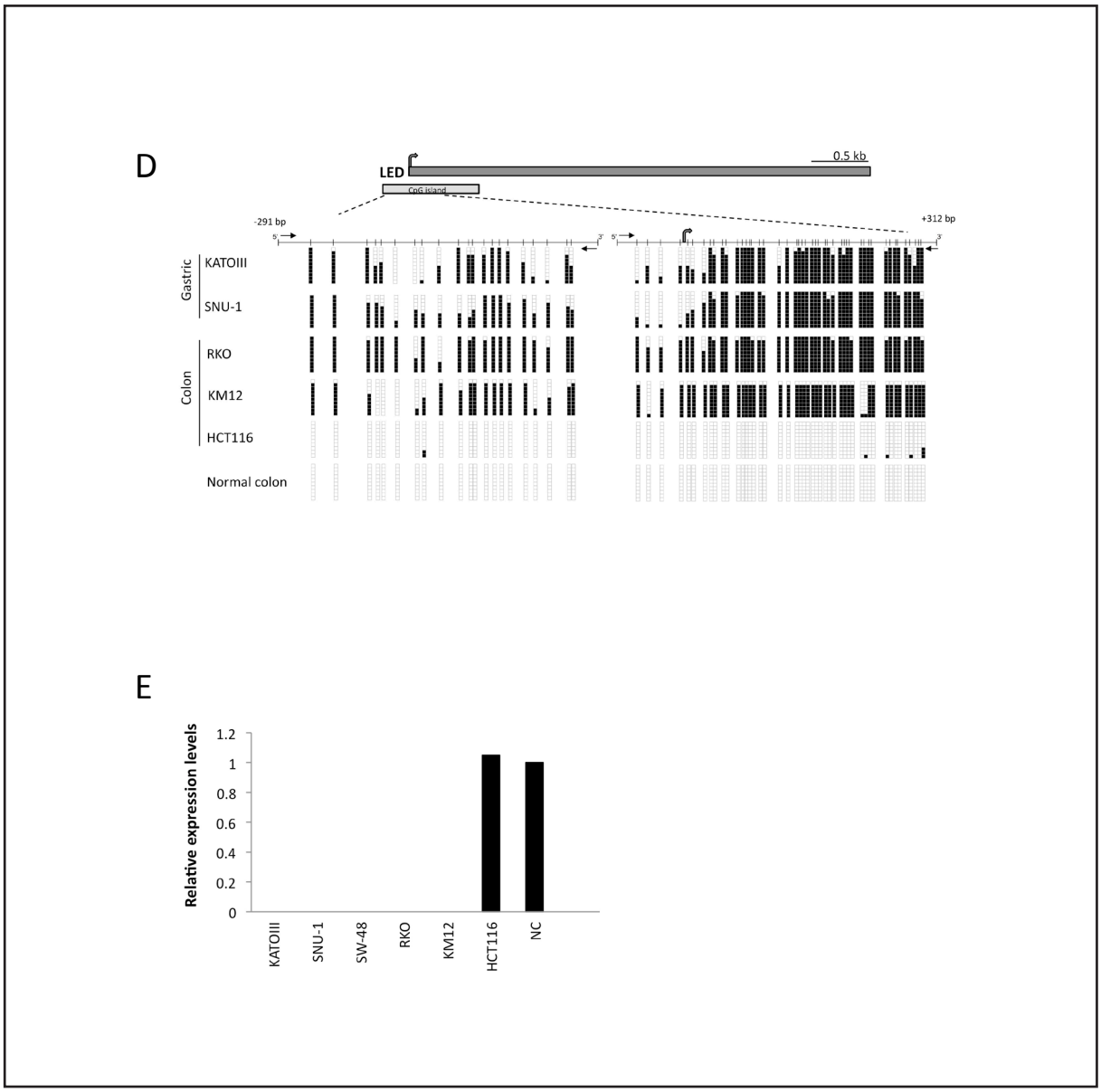


Figure S5 (related to figure 5.). DNA methylation-associated silencing of LED in colon and gastric cancers

(A) DNA methylation-associated silencing of LED in several cancer cell lines. Methylation was measured in different LED regions represented by TSS (transcription start site: -158, -155, -141, -136, -46, -30, +549, +683). Presence of unmethylated or methylated CpGs is indicated by green or red, respectively. Methylation is considered when both CpGs at (-155, -141, -136) and at (+549,+683) are methylated or when all the CpGs are methylated. Indication of TP53 status (W-wild type ; M-mutant) and gender (M-male ; F-female) are represented. (B-C) LED methylation in normal lymphocytes by gender- female (B) and male (C) and analysed as in (A). (D) Schematic representation of LED gemomic loci and CpG island. Bisulfite genomic sequencing analysis of LED CpG island in human solid cancer cell lines and normal colon as as tissue control. Location of bisulfite genomic sequencing PCR primers (black arrows), CpG dinucleotides (vertical lines) and the transcriptional start site (grey arrow) are shown. Ten single clones are represented for each sample. Presence of unmethylated or methylated CpGs is indicated by white or black squares, respectively. (E) LED expression levels in methylated or unmethylated human solid cancer cell lines and in normal colon (NC) as control. Values were determined by qRT-PCR in triplicates.

Chapter four
Discussion

4.1 Redefining "Junk" DNA

Presently we are experiencing a "genomic revolution". It has been demonstrated that our genome is highly dynamic, responding to numerous environmental cues achieved by the concerted action of transcriptional networks, which orchestrate cellular events.

Eukaryotic organisms have evolved complex gene regulatory circuits that constantly redefine and shape our genome. Intimately related to this complex circuitry, the concept of C-value enigma discusses the fact that genome size does not correlate with organismal complexity (Gregory, 2001). For example, the cells of some salamanders may contain 40 times more DNA than those of humans. It was not until the development of the genome programs and the subsequent breakthrough in sequencing technologies, that we could fully appreciate the widespread importance of transcriptional regulation as a fundamental concept of the genome.

Large-scale transcriptome studies revealed that our genome is composed of only 2-3% protein-coding genes, whereas 98% accounts for intronic and intergenic regions, ncRNAs and short or long interspersed elements (Lander et al., 2001). In addition, the characterization and annotation of all functional DNA sequences, carried out by projects, such as the ENCODE (Encyclopedia of DNA Elements), led us to reconsider the non-coding portion of the genome, often called the genomic dark matter. Enhancers, for example, are part of the 98% of the human genome that is non-coding. Researchers are now turning their focus of attention into the once

called "Junk DNA" elements, and slowly uncovering the players of this incredible genomic choreography. "Let there be light" into the working elements of our genome and what those elements do.

4.2 eRNAs reach the heart of transcription

Enhancers are short genomic sequences that regulate transcription of remotely located target genes. Specific epigenetic signatures, such as histone methylation and acetylation, are good predictors of enhancer activity. Primarily, active enhancers are marked by high level of histone 3 lysine 4 monomethylation (H3K4me1), low level of H3K4me3 (Heintzman et al., 2007a;Visel et al., 2009) and high level of histone 3 lysine 27 acetylation (H3K27Ac) (Creyghton et al., 2010a). Most importantly, it was recently demonstrated that these genomic elements are actively transcribed. A landmark paper by Kim et al (2010) described enhancer RNAs, a new class of non-coding RNAs (ncRNAs) that are produced from polymerase II bound enhancers. This study unraveled a complex connection between eRNA-producing domains and their target genes. Most significantly, they observed that the level of eRNA expression positively correlated with the expression of nearby coding genes (Kim et al., 2010). However, several questions were posed concerning the functionality of these RNA molecules. Are eRNAs just a byproduct of pervasive transcription?

Follow up studies revealed the potential and functional importance of eRNA-producing domains by studying key transcription factors (Melo et al., 2013; Wang et al., 2011a). For instance, using ChIP-seq, recent reports have demonstrated the importance of eRNAs for enhancer activity. For example, Wang et al, (2011) demonstrated that androgen receptor (AR) and FOXA1 are able to bind enhancer regions. Using Gro-seq they observed specific stimulus-dependent expression of eRNAs from these enhancer domains. Most importantly, the expression of these enhancer-derived transcripts significantly correlated with the new AR binding program in the absence of FOXA1, thereby demonstrating the importance of these RNA molecules in the regulation of gene expression (Wang et al., 2011a). However, their functional role remained unclear.

Following this line of thought, in this Doctoral Thesis we were able to demonstrate that the tumor suppressor p53 binds with high affinity to specific enhancer regions (termed p53-bound enhancer regions or p53BERs). Detailed analyses unveiled the existence of eRNAs that were stimulated upon cellular treatment with Nutlin-3a, a p53 specific activator. Most importantly, eRNA depletion led to a specific repression of neighboring protein-coding genes.

Two recent follow-up studies (Li et al, 2013; Lam et al, 2013) showed that eRNAs can, in fact, either positively or negatively impact the transcriptional regulation of nearby coding genes. Li and colleagues (2013) set out to investigate the DNA binding profile of oestrogen receptor α (ER-α) in human breast cancer cells. Using GRO-Seq on 17β-oestradiol (E2) treated cells,

they observed a correlation between ER-α-bound enhancers, eRNA induction, and oestrogen-upregulated expression of neighboring genes. In a similar way, Lam *et al* (2013) found that the nuclear receptors Rev-Erbs directly regulate macrophage gene expression by negatively influencing the transcriptional activity of enhancers.

To further understand the importance of eRNA transcriptional activity, we used a tethering system. We cloned the eRNAs to an expression vector, where they were expressed as a chimera with an RNA guiding moiety. These chimeric RNAs bind to an adaptor protein (MS2-coat viral protein: MS2-CP) that simultaneously binds to specific sequences on a reporter plasmid. We observed that bringing the eRNA, produced from p53BERs, into physical proximity to a promoter, was sufficient for the expression of the reporter gene. Our results provided one of the first examples towards the functional characterization of eRNAs in enhancer activity (Melo et al., 2013). In a similar experimental setup, Li *et al.* (2013) demonstrated that FOXC1 eRNA was able to increase a FOXC1 native-promoter activity. Furthermore, the authors observed that depletion of the sense eRNA producing enhancer cassette resulted in the abrogation of the promoter signal. This activity was fully restored upon tethering the sense FOXC1 eRNA to this half-enhancer site (Li et al., 2013). Using a different but compelling strategy, Lam *et al.* (2013) extended these observations to Rev-Erbs induced eRNAs. Using various reporter constructs, they demonstrated that the production of sequence-specific eRNAs appears to be a prerequisite for the emergence of enhancing activity.

Altogether, we were able to demonstrate and confirm that p53 can bind enhancer regions producer of eRNAs. Our own and follow-up studies showed the importance of these RNA molecules for transcription enhancement. However, how do these RNA molecules reach the heart of transcription?

4.3 eRNAs mechanism of action

We postulated that eRNA tethering to promoters could be accomplished through DNA looping. This event is most commonly involved in transcriptional regulation, allowing distal binding sites to act in a coordinated manner. Using circularized chromosome conformation capture combined with deep sequencing (4C technology), it is possible to produce genome-wide maps that reflect the frequency of interactions between a specific locus with other regions in the genome. Indeed, we observed that p53BERs interact intrachromosomally with neighboring genes, thereby conveying a long-distance p53-dependent transcriptional activation. Most importantly we observed that the long-range interactions were still present in the absence of p53 suggesting that this tumor suppressor acts on pre-existing chromatin loops to activate p53BER-target genes. Interestingly other reports showed the importance of eRNA production influencing target gene transcription, in part through the stabilization of enhancer-promoter looping, as suggested by Li *et al.* (2013).

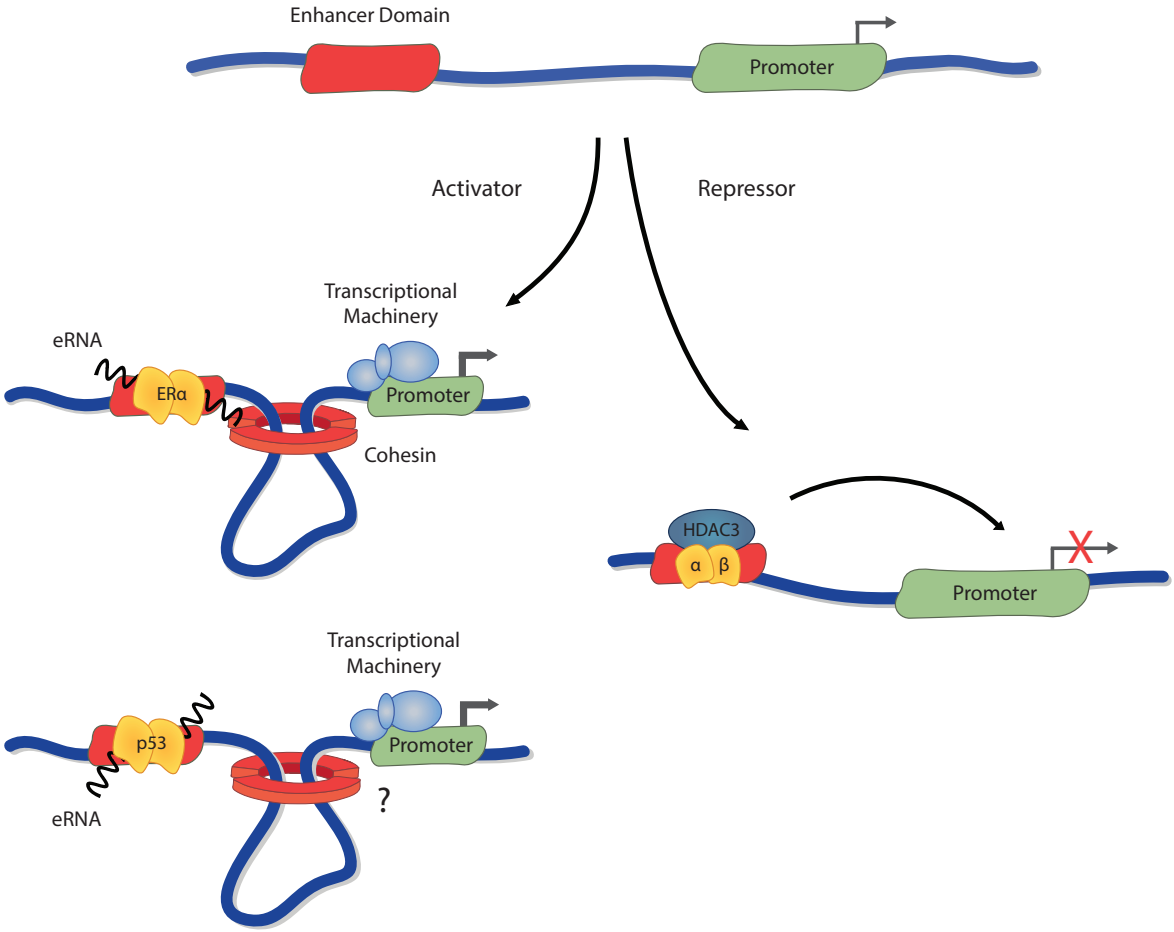


Figure 1. Representative scheme showing the influence of eRNAs on the expression of neighboring gene.
Key transcription factors can activate (ER-α and p53) or repress (Rev-Erb proteins) enhancer domains by inducing or silencing eRNA production. While induced eRNAs lead to a greater activation of neighboring genes, transcriptional repressors (Rev-Erb α/β) can achieve their function through epigenetic silencing (HDAC deacetylation) of enhancer domains and the repression of eRNA production. Both scenarios have opposite impact on the transcriptional regulation of neighboring genes.

To investigate whether eRNA production from ER-α-bound enhancer impact chromatin structure, the authors applied a three-dimensional DNA selection and ligation (3D-DSL) procedure. They found that E2 (17β-estradiol) treatment intensified the frequency of interactions between enhancers and promoters. Supporting this notion, targeting eRNAs with antisense oligos reduced the number of enhancer-promoter interactions. This finding raise the important question of whether, in this working system, the eRNA production could contribute to the stabilization of enhancer-promoter looping. As several reports have already indicated cohesin as a key component for the formation and stabilization of chromatin structures (Merkenschlager and Odom, 2013), the authors performed immunoprecipitation with antibodies against Rad21 (one subunit of the cohesin complex), followed by deep sequencing (ChIP-seq). Indeed, a subset of Rad21 binding sites was clearly overlapping with E2-induced eRNA-producing enhancer domains. This suggested an interaction between eRNAs and the cohesin complex, which was then confirmed by RNA immunoprecipitation experiments. However, the quantitative effect of eRNA production on cohesin recruitment remains unclear. Intriguingly, we observed that p53BERs act on pre-determined chromatin loops, suggesting a different activation mechanism concerning eRNA function. However, it is possible that the cohesin complex, among other factors that are not dependent on eRNA production, might be present at p53BERs.

One other possibility, involves the fact that upon stress, an immediate cellular response is needed. Acting on pre-determined chromatin loops, p53 can rapidly activate a myriad of target genes, through the production of enhancer-derived transcripts. Additionally, Hah and colleagues (2013) also observed that reducing the amount of eRNA had no significant effect on estradiol-induced enhancer-promoter looping, opposing to what has been reported by Li and colleagues (2013) (Hah et al., 2013b). However, this difference may be explained by the different experimental techniques used, chromosome conformation capture (3C) versus 3D-DSL procedure.

In sum, these findings reflect different mechanistic underlying eRNA function at specific loci or enhancers (Figure 1), with numerous intriguing questions to be answered. For example, the properties or rules governing the specificity of action of the eRNAs are still undefined. One enticing idea concerns the potential role of eRNA structure in mediating protein recruitment to specific promoter targets to activate transcription. This type of activation was proposed for long intergenic non-coding RNAs (lincRNAs) (Huarte et al., 2010) and might as well occur with eRNAs. We were able to demonstrate that eRNA knockdown significantly reduced RNAPII occupancy at the promoter of the target gene, implying that eRNA synthesis facilitates RNAPII recruitment. In accordance to our results, a recent report by Mousavi *et al.* (2013) demonstrated that the eRNA produced from the core enhancer (CE) element of the MYOD gene was necessary for an efficient RNAPII recruitment and a permissive chromatin environment at the MYOD promoter, through chromatin remodeling complexes (Mousavi et al., 2013). Collectively, these findings propose a model whereby the eRNA, through the recruitment of master transcription factors and/or chromatin remodeling complexes, favors

an open chromatin state and consequently regulates gene expression. Alternatively, sequence homology between eRNAs and promoters could also dictate the formation of a RNA-DNA hybrid complex at specific targets leading to the subsequent alteration in chromatin structure and transcription activity at target genes.

4.4 Transcriptional regulatory modules associated to lncRNAs

Recent reports support eRNA regulation by master transcription factors. However, ncRNAs are highly dynamic molecules involved in various cellular processes. The existence of complex regulatory networks governing the genome uncovers numerous possibilities yet to be discovered. For example, it has been shown that miR-103 and miR-107, implicated in cancer, binds to the 3' untranslated region (UTR) of DICER mRNA. Interestingly, Dicer is an enzyme involved in the miRNA processing machinery, providing an explanation for the fact that the expression of miR-103 and miR-107 globally affects miRNA production (Martello et al., 2010). These and other findings suggest that ncRNAs are integral members of a complex system, in which multiple players can have their role in the "genomic script".

Our study further demonstrates the intrinsic complexity of ncRNAs, more specifically eRNAs, unraveling new regulatory mechanisms. Using Global Run-ON sequencing (GRO-seq) we mapped the presence of p53- (bound and unbound) regulated enhancers (p53RERs). Furthermore, we were able to uncover the existence of additional players, besides transcription factors, aiding the regulation of enhancer RNA producing domains. The differences related to the presence or absence of p53 and the transcriptional "sharpness" between the different groups of p53RERS: p53-bound enhancer regions (p53BERs) and p53-free enhancer regions (p53FERs), led us to suggest different modes of activation. Indeed, we provide evidence that a subset of p53BERs and p53FERs are regulated by a lncRNA induced by p53. Using chromatin isolation by RNA purification, we were able to observe that the lncRNA LED (LncRNA activator of Enhancer Domains) binds and epigenetically regulates the activity of eRNA producing domains.

Recent reports tackle this type of interactions between different genomic elements. Yang *et al.* (2013) reports the existence of a lncRNA termed PCGEM1 (prostate-specific transcript 1) that binds to AR and associates with genomic locations characteristic of AR-bound enhancers. Furthermore, they observed that PCGEM1 association with AR-bound enhancers increased the AR-mediated gene activation without affecting AR levels (Yang et al., 2013). Indeed, in our study we observed a significant enrichment of induced p53RERs in the LED-bound fraction of enhancers. Moreover, we show that Nutlin-3a induced enhancers are positioned closer to p53-regulated canonical genes, than to non-regulated genes. A more recent example involves a lncRNA transcribed from a conserved enhancer element upstream of the PAX6 gene, called Paupar (Vance et al., 2014b). Through a different experimental approach, called CHART-seq (capture hybridization analysis of RNA targets), the authors were able to map the different

genomic binding sites of this lncRNA. They provide evidence that the functional Paupar binding sites are located within the regulatory elements of genes that are affected by Paupar deletion. They also observed an overlap between a subset of Paupar binding sites and PAX6 occupancy. In fact, they demonstrate that Paupar, together with the transcription factor PAX6, is able to modulate the activity of transcriptionally regulatory elements of neurodevelopmental genes. This report provides evidence that transcriptional active enhancers (Paupar) can act both locally (affecting PAX6 gene) and distally, sharing a common mechanistic basis for the majority of lncRNAs.

We further deepen this type of regulatory process by showing that LED alone or together with p53 modulates the activity of multiple p53RERs. Also, our study uncovers a new lncRNA-dependent epigenetic regulation of eRNA producing domains. Finally, we disclosed LED inactivation in human cancer, thereby highlighting its important role as a potential tumor suppressor of the p53 pathway. Collectively our findings represent an important contribution towards the elucidation of the heterogeneous molecular mechanisms associated with lncRNAs.

Transcription has therefore become an integral part of enhancer domains. Understanding the eRNA dynamics during several cellular events, such as cancer, is a matter of immediate importance. Hence, while key findings start to shed light on the mechanisms and biological relevance of these RNA molecules, a long walk lies ahead shadowed by what we still do not know. Only the continuing study will shed light on its events and with it create new therapeutic approaches to disease.

Chapter five
Concluding Remarks and Future Perspectives

Dark matter, that's how science described 98% of the human genome that lies between our 21,000 genes, since our DNA was first sequenced. However, as Lavoisier once said: "In nature nothing is wasted, nothing is created, everything is transformed", once thought to be an apparent biological wasteland, the dark matter of our genome revealed to be the leading character of crucial gene regulatory activity. Projects such as the ENCODE, revealed that around 80% of our genome is biochemically active. We can now predict, with high confidence, functional regulatory elements lying in the "black holes" of our genome. Recent works uncovered some of the many features associated to one class of regulatory genomic elements, the enhancers. Furthermore, it became clear that transcription associated to enhancers represent an important event intimately connected to enhancer activity. Since the discovery of enhancer-RNAs (eRNAs), numerous studies revealed the importance of these lncRNA molecules in the tissue-specific regulation of gene expression. However, the biological relevance of eRNAs remains largely unknown. Recent reports have started to shed some light on this subject. For example, we were first to show that suppression of the eRNA, expressed from the p53-bound enhancer region 2 (p53BER2), as well its protein-coding target gene (PAPPA), had a significant inhibitory effect on the p53-induced cell cycle arrest (Melo et al., 2013). These findings prompt important scientific questions concerning the use of eRNA producing domains as therapeutic targets, allowing for a locus- and tissue-specific gene regulation. One other important issue involves numerous genome-wide association studies (GWAS). GWAS

are normally performed by genotyping naturally occurring DNA sequences, known as single nucleotide polymorphism (SNPs) across the genome in the case of study (disease) and control (without disease) populations (Hirschhorn and Daly, 2005). Numerous studies were able to identify many disease-associated SNPs. Moreover, more than 90% of these SNPs were occurring not in the genes themselves, but in the non-coding portion of the genome, pointing to regulatory elements, such as enhancers as culprits (Manolio et al., 2009). In accordance with this hypothesis, ENCODE consortium and other groups successfully mapped thousands of disease-associated SNPs to enhancer elements that were identified through epigenomic profiling (Consortium et al., 2012). The observation that enhancers and enhancer-derived transcripts are often cell-type specific may provide possible explanations for the tissue- and disease-specific nature of many susceptibility alleles. Collectively, we now possess the right tools to establish important connections between several diseases and the genome control circuitry behind it. Understanding how transcriptional regulation may contribute to the pathogenesis of diseases is a matter of utmost importance. Finally, regulating the levels of gene expression may prove a more reliable therapeutic strategy than replacing a defective protein or turning off a gain-of-function allele. Therefore, an endeavour into the role of these cis-regulatory elements in disease carries great promise for therapeutical interventions.

Our work advances science towards a better understanding of lncRNAs as important elements involved in the control of gene expression. The determination of their mode of action as well as their biological relevance is instrumental towards a deep understanding of human biology and the diseases associated with it.

Chapter six
References

Agalioti, T., Chen, G., and Thanos, D. (2002). Deciphering the transcriptional histone acetylation code for a human gene. Cell *111*, 381-392.

Bagga, S., Bracht, J., Hunter, S., Massirer, K., Holtz, J., Eachus, R., and Pasquinelli, A.E. (2005). Regulation by let-7 and lin-4 miRNAs results in target mRNA degradation. Cell *122*, 553-563.

Barak, Y., Juven, T., Haffner, R., and Oren, M. (1993). mdm2 expression is induced by wild type p53 activity. The EMBO journal *12*, 461-468.

Bartel, D.P. (2004). MicroRNAs: genomics, biogenesis, mechanism, and function. Cell *116*, 281-297.

Bartolomei, M.S., Zemel, S., and Tilghman, S.M. (1991). Parental imprinting of the mouse H19 gene. Nature *351*, 153-155.

Beckerman, R., and Prives, C. (2010). Transcriptional regulation by p53. Cold Spring Harbor perspectives in biology *2*, a000935.

Bejerano, G., Lowe, C.B., Ahituv, N., King, B., Siepel, A., Salama, S.R., Rubin, E.M., Kent, W.J., and Haussler, D. (2006). A distal enhancer and an ultraconserved exon are derived from a novel retroposon. Nature *441*, 87-90.

Bertani, S., Sauer, S., Bolotin, E., and Sauer, F. (2011). The noncoding RNA Mistral activates Hoxa6 and Hoxa7 expression and stem cell differentiation by recruiting MLL1 to chromatin. Molecular cell *43*, 1040-1046.

Birney, E., Stamatoyannopoulos, J.A., Dutta, A., Guigo, R., Gingeras, T.R., Margulies, E.H., Weng, Z., Snyder, M., Dermitzakis, E.T., Thurman, R.E., *et al.* (2007). Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. Nature *447*, 799-816.

Bommer, G.T., Gerin, I., Feng, Y., Kaczorowski, A.J., Kuick, R., Love, R.E., Zhai, Y., Giordano, T.J., Qin, Z.S., Moore, B.B., *et al.* (2007). p53-mediated activation of miRNA34 candidate tumor-suppressor genes. Current biology : CB *17*, 1298-1307.

Brady, C.A., Jiang, D., Mello, S.S., Johnson, T.M., Jarvis, L.A., Kozak, M.M., Kenzelmann Broz, D., Basak, S., Park, E.J., McLaughlin, M.E., *et al.* (2011). Distinct p53 transcriptional programs dictate acute DNA-damage responses and tumor suppression. Cell *145*, 571-583.

Brisken, C. (2013). Progesterone signalling in breast cancer: a neglected hormone coming into the limelight. Nature reviews. Cancer *13*, 385-396.

Brosh, R., and Rotter, V. (2009). When mutants gain new powers: news from the mutant p53 field. Nature reviews. Cancer *9*, 701-713.

Broude, E.V., Swift, M.E., Vivo, C., Chang, B.D., Davis, B.M., Kalurupalle, S., Blagosklonny, M.V., and Roninson, I.B. (2007). p21(Waf1/Cip1/Sdi1) mediates retinoblastoma protein degradation. Oncogene *26*, 6954-6958.

Brown, C.J., Ballabio, A., Rupert, J.L., Lafreniere, R.G., Grompe, M., Tonlorenzi, R., and Willard, H.F. (1991). A gene from the region of the human X inactivation centre is expressed exclusively from the inactive X chromosome. Nature *349*, 38-44.

Bulger, M., and Groudine, M. (2011). Functional and mechanistic diversity of distal transcription enhancers. Cell *144*, 327-339.

Cabianca, D.S., Casa, V., Bodega, B., Xynos, A., Ginelli, E., Tanaka, Y., and Gabellini, D. (2012). A long ncRNA links copy number variation to a polycomb/trithorax epigenetic switch in FSHD muscular dystrophy. Cell *149*, 819-831.

Cabili, M.N., Trapnell, C., Goff, L., Koziol, M., Tazon-Vega, B., Regev, A., and Rinn, J.L. (2011). Integrative annotation of human large intergenic noncoding RNAs reveals global properties and specific subclasses. Genes & development *25*, 1915-1927.

Cesana, M., Cacchiarelli, D., Legnini, I., Santini, T., Sthandier, O., Chinappi, M., Tramontano, A., and Bozzoni, I. (2011). A long noncoding RNA controls muscle differentiation by functioning as a competing endogenous RNA. Cell *147*, 358-369.

Chang, C., Lee, S.O., Yeh, S., and Chang, T.M. (2013). Androgen receptor (AR) differential roles in hormone-related tumors including prostate, bladder, kidney, lung, breast and liver. Oncogene.

Chang, T.C., Wentzel, E.A., Kent, O.A., Ramachandran, K., Mullendore, M., Lee, K.H., Feldmann, G., Yamakuchi, M., Ferlito, M., Lowenstein, C.J., *et al.* (2007). Transactivation of miR-34a by p53 broadly influences gene expression and promotes apoptosis. Molecular cell *26*, 745-752.

Christophorou, M.A., Ringshausen, I., Finch, A.J., Swigart, L.B., and Evan, G.I. (2006). The pathological response to DNA damage does not contribute to p53-mediated tumour suppression. Nature *443*, 214-217.

Chu, C., Qu, K., Zhong, F.L., Artandi, S.E., and Chang, H.Y. (2011). Genomic maps of long noncoding RNA occupancy reveal principles of RNA-chromatin interactions. Mol Cell *44*, 667-678.
Chu, C., Quinn, J., and Chang, H.Y. (2012). Chromatin isolation by RNA purification (ChIRP). J Vis Exp.

Chuong, C.M. (2003). Homeobox genes, fetal wound healing, and skin regional specificity. J Invest Dermatol *120*, 9-11.

Consortium, E.P., Bernstein, B.E., Birney, E., Dunham, I., Green, E.D., Gunter, C., and Snyder, M. (2012). An integrated encyclopedia of DNA elements in the human genome. Nature *489*, 57-74.

Contente, A., Dittmer, A., Koch, M.C., Roth, J., and Dobbelstein, M. (2002). A polymorphic microsatellite that mediates induction of PIG3 by p53. Nature genetics *30*, 315-320.

Core, L.J., Waterfall, J.J., and Lis, J.T. (2008). Nascent RNA sequencing reveals widespread pausing and divergent initiation at human promoters. Science *322*, 1845-1848.

Creyghton, M.P., Cheng, A.W., Welstead, G.G., Kooistra, T., Carey, B.W., Steine, E.J., Hanna, J., Lodato, M.A., Frampton, G.M., Sharp, P.A., *et al.* (2010a). Histone H3K27ac separates active from poised enhancers and predicts developmental state. Proc Natl Acad Sci U S A *107*, 21931-21936.

Creyghton, M.P., Cheng, A.W., Welstead, G.G., Kooistra, T., Carey, B.W., Steine, E.J., Hanna, J., Lodato, M.A., Frampton, G.M., Sharp, P.A., *et al.* (2010b). Histone H3K27ac separates active from poised enhancers and predicts developmental state. Proceedings of the National Academy of Sciences of the United States of America *107*, 21931-21936.

Crick, F.H., Barnett, L., Brenner, S., and Watts-Tobin, R.J. (1961). General nature of the genetic code for proteins. Nature *192*, 1227-1232.

D'Orazi, G., Cecchinelli, B., Bruno, T., Manni, I., Higashimoto, Y., Saito, S., Gostissa, M., Coen, S., Marchetti, A., Del Sal, G., *et al.* (2002). Homeodomain-interacting protein kinase-2 phosphorylates p53 at Ser 46 and mediates apoptosis. Nature cell biology *4*, 11-19.

De Santa, F., Barozzi, I., Mietton, F., Ghisletti, S., Polletti, S., Tusi, B.K., Muller, H., Ragoussis, J., Wei, C.L., and Natoli, G. (2010). A large fraction of extragenic RNA pol II transcription sites overlap enhancers. PLoS biology *8*, e1000384.

de Wit, E., Braunschweig, U., Greil, F., Bussemaker, H.J., and van Steensel, B. (2008). Global chromatin domain organization of the Drosophila genome. PLoS Genet *4*, e1000045.

Derrien, T., Johnson, R., Bussotti, G., Tanzer, A., Djebali, S., Tilgner, H., Guernec, G., Martin, D., Merkel, A., Knowles, D.G., *et al.* (2012). The GENCODE v7 catalog of human long noncoding RNAs: analysis of their gene structure, evolution, and expression. Genome Res *22*, 1775-1789.

Dinger, M.E., Amaral, P.P., Mercer, T.R., Pang, K.C., Bruce, S.J., Gardiner, B.B., Askarian-Amiri, M.E., Ru, K., Solda, G., Simons, C., *et al.* (2008). Long noncoding RNAs in mouse embryonic stem cell pluripotency and differentiation. Genome Res *18*, 1433-1445.

Drost, J., Mantovani, F., Tocco, F., Elkon, R., Comel, A., Holstege, H., Kerkhoven, R., Jonkers, J., Voorhoeve, P.M., Agami, R., *et al.* (2010). BRD7 is a candidate tumour suppressor gene required for p53 function. Nature cell biology *12*, 380-389.

Duret, L., Chureau, C., Samain, S., Weissenbach, J., and Avner, P. (2006). The Xist RNA gene evolved in eutherians by pseudogenization of a protein-coding gene. Science *312*, 1653-1655.

el-Deiry, W.S., Kern, S.E., Pietenpol, J.A., Kinzler, K.W., and Vogelstein, B. (1992). Definition of a consensus binding site for p53. Nature genetics *1*, 45-49.

el-Deiry, W.S., Tokino, T., Velculescu, V.E., Levy, D.B., Parsons, R., Trent, J.M., Lin, D., Mercer, W.E., Kinzler, K.W., and Vogelstein, B. (1993). WAF1, a potential mediator of p53 tumor suppression. Cell *75*, 817-825.

Eliyahu, D., Michalovitz, D., Eliyahu, S., Pinhasi-Kimhi, O., and Oren, M. (1989). Wild-type p53 can inhibit oncogene-mediated focus formation. Proceedings of the National Academy of Sciences of the United States of America *86*, 8763-8767.

Eliyahu, D., Raz, A., Gruss, P., Givol, D., and Oren, M. (1984). Participation of p53 cellular tumour antigen in transformation of normal embryonic cells. Nature *312*, 646-649.

Ernst, J., Kheradpour, P., Mikkelsen, T.S., Shoresh, N., Ward, L.D., Epstein, C.B., Zhang, X., Wang, L., Issner, R., Coyne, M., *et al.* (2011). Mapping and analysis of chromatin state dynamics in nine human cell types. Nature *473*, 43-49.

Espinosa, J.M., Verdun, R.E., and Emerson, B.M. (2003). p53 functions through stress- and promoter-specific recruitment of transcription initiation components before and after DNA damage. Molecular cell *12*, 1015-1027.

Faulkner, G.J., Kimura, Y., Daub, C.O., Wani, S., Plessy, C., Irvine, K.M., Schroder, K., Cloonan, N., Steptoe, A.L., Lassmann, T., *et al.* (2009). The regulated retrotransposon transcriptome of mammalian cells. Nature genetics *41*, 563-571.

Felsenfeld, G., and Groudine, M. (2003). Controlling the double helix. Nature *421*, 448-453.

Finlay, C.A., Hinds, P.W., and Levine, A.J. (1989). The p53 proto-oncogene can act as a suppressor of transformation. Cell *57*, 1083-1093.

Fire, A., Xu, S., Montgomery, M.K., Kostas, S.A., Driver, S.E., and Mello, C.C. (1998). Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans. Nature *391*, 806-811.

Fridman, J.S., and Lowe, S.W. (2003). Control of apoptosis by p53. Oncogene *22*, 9030-9040.

Gamper, A.M., and Roeder, R.G. (2008). Multivalent binding of p53 to the STAGA complex mediates coactivator recruitment after UV damage. Mol Cell Biol *28*, 2517-2527.

Gembarska, A., Luciani, F., Fedele, C., Russell, E.A., Dewaele, M., Villar, S., Zwolinska, A., Haupt, S., de Lange, J., Yip, D., *et al.* (2012). MDM4 is a key therapeutic target in cutaneous melanoma. Nat Med *18*, 1239-1247.

Giraldez, A.J., Mishima, Y., Rihel, J., Grocock, R.J., Van Dongen, S., Inoue, K., Enright, A.J., and Schier, A.F. (2006). Zebrafish MiR-430 promotes deadenylation and clearance of maternal mRNAs. Science *312*, 75-79.

Gohler, T., Reimann, M., Cherny, D., Walter, K., Warnecke, G., Kim, E., and Deppert, W. (2002). Specific interaction of p53 with target binding sites is determined by DNA conformation and is regulated by the C-terminal domain. The Journal of biological chemistry *277*, 41192-41203.

Gordon, F.E., Nutt, C.L., Cheunsuchon, P., Nakayama, Y., Provencher, K.A., Rice, K.A., Zhou, Y., Zhang, X., and Klibanski, A. (2010). Increased expression of angiogenic genes in the brains of mouse meg3-null embryos. Endocrinology *151*, 2443-2452.

Gregory, T.R. (2001). Coincidence, coevolution, or causation? DNA content, cell size, and the C-value enigma. Biological reviews of the Cambridge Philosophical Society *76*, 65-101.

Gupta, R.A., Shah, N., Wang, K.C., Kim, J., Horlings, H.M., Wong, D.J., Tsai, M.C., Hung, T., Argani, P., Rinn, J.L., *et al.* (2010). Long non-coding RNA HOTAIR reprograms chromatin state to promote cancer metastasis. Nature *464*, 1071-1076.

Guttman, M., Amit, I., Garber, M., French, C., Lin, M.F., Feldser, D., Huarte, M., Zuk, O., Carey, B.W., Cassady, J.P., *et al.* (2009). Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals. Nature *458*, 223-227.

Guttman, M., Garber, M., Levin, J.Z., Donaghey, J., Robinson, J., Adiconis, X., Fan, L., Koziol, M.J., Gnirke, A., Nusbaum, C., *et al.* (2010). Ab initio reconstruction of cell type-specific transcriptomes in mouse reveals the conserved multi-exonic structure of lincRNAs. Nat Biotechnol *28*, 503-510.

Hah, N., Murakami, S., Nagari, A., Danko, C.G., and Kraus, W.L. (2013a). Enhancer transcripts mark active estrogen receptor binding sites. Genome Res *23*, 1210-1223.

Hah, N., Murakami, S., Nagari, A., Danko, C.G., and Kraus, W.L. (2013b). Enhancer transcripts mark active estrogen receptor binding sites. Genome Res *23*, 1210-1223.

Hammond, S.M., Bernstein, E., Beach, D., and Hannon, G.J. (2000). An RNA-directed nuclease mediates post-transcriptional gene silencing in Drosophila cells. Nature *404*, 293-296.

Harper, J.W., Adami, G.R., Wei, N., Keyomarsi, K., and Elledge, S.J. (1993). The p21 Cdk-interacting protein Cip1 is a potent inhibitor of G1 cyclin-dependent kinases. Cell *75*, 805-816.

He, H.H., Meyer, C.A., Shin, H., Bailey, S.T., Wei, G., Wang, Q., Zhang, Y., Xu, K., Ni, M., Lupien, M., *et al.* (2010). Nucleosome dynamics define transcriptional enhancers. Nature genetics *42*, 343-347.

He, L., He, X., Lim, L.P., de Stanchina, E., Xuan, Z., Liang, Y., Xue, W., Zender, L., Magnus, J., Ridzon, D., *et al.* (2007). A microRNA component of the p53 tumour suppressor network. Nature *447*, 1130-1134.

Heintzman, N.D., Hon, G.C., Hawkins, R.D., Kheradpour, P., Stark, A., Harp, L.F., Ye, Z., Lee, L.K., Stuart, R.K., Ching, C.W., *et al.* (2009). Histone modifications at human enhancers reflect global cell-type-specific gene expression. Nature *459*, 108-112.

Heintzman, N.D., Stuart, R.K., Hon, G., Fu, Y., Ching, C.W., Hawkins, R.D., Barrera, L.O., Van Calcar, S., Qu, C., Ching, K.A., *et al.* (2007a). Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. Nat Genet *39*, 311-318.

Heintzman, N.D., Stuart, R.K., Hon, G., Fu, Y., Ching, C.W., Hawkins, R.D., Barrera, L.O., Van Calcar, S., Qu, C., Ching, K.A., *et al.* (2007b). Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. Nature genetics *39*, 311-318.

Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y.C., Laslo, P., Cheng, J.X., Murre, C., Singh, H., and Glass, C.K. (2010a). Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Molecular cell *38*, 576-589.

Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y.C., Laslo, P., Cheng, J.X., Murre, C., Singh, H., and Glass, C.K. (2010b). Simple combinations of lineage-determining transcription factors

prime cis-regulatory elements required for macrophage and B cell identities. Mol Cell *38*, 576-589.

Hinds, P., Finlay, C., and Levine, A.J. (1989). Mutation is required to activate the p53 gene for cooperation with the ras oncogene and transformation. Journal of virology *63*, 739-746.

Hinds, P.W., Finlay, C.A., Quartin, R.S., Baker, S.J., Fearon, E.R., Vogelstein, B., and Levine, A.J. (1990). Mutant p53 DNA clones from human colon carcinomas cooperate with ras in transforming primary rat cells: a comparison of the "hot spot" mutant phenotypes. Cell growth & differentiation : the molecular biology journal of the American Association for Cancer Research *1*, 571-580.

Hirschhorn, J.N., and Daly, M.J. (2005). Genome-wide association studies for common diseases and complex traits. Nature reviews. Genetics *6*, 95-108.

Hollstein, M., Sidransky, D., Vogelstein, B., and Harris, C.C. (1991). p53 mutations in human cancers. Science *253*, 49-53.

Hu, W., Chan, C.S., Wu, R., Zhang, C., Sun, Y., Song, J.S., Tang, L.H., Levine, A.J., and Feng, Z. (2010). Negative regulation of tumor suppressor p53 by microRNA miR-504. Molecular cell *38*, 689-699.

Huang, J., Zhou, N., Watabe, K., Lu, Z., Wu, F., Xu, M., and Mo, Y.Y. (2014). Long non-coding RNA UCA1 promotes breast tumor growth by suppression of p27 (Kip1). Cell Death Dis *5*, e1008.

Huarte, M., Guttman, M., Feldser, D., Garber, M., Koziol, M.J., Kenzelmann-Broz, D., Khalil, A.M., Zuk, O., Amit, I., Rabani, M., *et al.* (2010). A large intergenic noncoding RNA induced by p53 mediates global gene repression in the p53 response. Cell *142*, 409-419.

Jia, H., Osak, M., Bogu, G.K., Stanton, L.W., Johnson, R., and Lipovich, L. (2010). Genome-wide computational identification and manual annotation of human long noncoding RNA genes. Rna *16*, 1478-1487.

Johnson, D.S., Mortazavi, A., Myers, R.M., and Wold, B. (2007). Genome-wide mapping of in vivo protein-DNA interactions. Science *316*, 1497-1502.

Kaneko, S., Li, G., Son, J., Xu, C.F., Margueron, R., Neubert, T.A., and Reinberg, D. (2010). Phosphorylation of the PRC2 component Ezh2 is cell cycle-regulated and up-regulates its binding to ncRNA. Genes & development *24*, 2615-2620.

Karam, C.S., Kellner, W.A., Takenaka, N., Clemmons, A.W., and Corces, V.G. (2010). 14-3-3 mediates histone cross-talk during transcription elongation in Drosophila. PLoS Genet *6*, e1000975.

Kelley, D., and Rinn, J. (2012). Transposable elements reveal a stem cell-specific class of long noncoding RNAs. Genome Biol *13*, R107.

Kim, E., and Deppert, W. (2003). The complex interactions of p53 with target DNA: we learn as we go. Biochemistry and cell biology = Biochimie et biologie cellulaire *81*, 141-150.

Kim, T.K., Hemberg, M., Gray, J.M., Costa, A.M., Bear, D.M., Wu, J., Harmin, D.A., Laptewicz, M., Barbara-Haley, K., Kuersten, S., *et al.* (2010). Widespread transcription at neuronal activity-regulated enhancers. Nature *465*, 182-187.

Kino, T., Hurt, D.E., Ichijo, T., Nader, N., and Chrousos, G.P. (2010). Noncoding RNA gas5 is a growth arrest- and starvation-associated repressor of the glucocorticoid receptor. Sci Signal *3*, ra8.

Klattenhoff, C.A., Scheuermann, J.C., Surface, L.E., Bradley, R.K., Fields, P.A., Steinhauser, M.L., Ding, H., Butty, V.L., Torrey, L., Haas, S., *et al.* (2013). Braveheart, a long noncoding RNA required for cardiovascular lineage commitment. Cell *152*, 570-583.

Kotake, Y., Nakagawa, T., Kitagawa, K., Suzuki, S., Liu, N., Kitagawa, M., and Xiong, Y. (2011). Long non-coding RNA ANRIL is required for the PRC2 recruitment to and silencing of p15(INK4B) tumor suppressor gene. Oncogene *30*, 1956-1962.

Krebs, A.R., Karmodiya, K., Lindahl-Allen, M., Struhl, K., and Tora, L. (2011). SAGA and ATAC histone acetyl transferase complexes regulate distinct sets of genes and ATAC defines a class of p300-independent enhancers. Mol Cell *44*, 410-423.

Kruse, J.P., and Gu, W. (2008). SnapShot: p53 posttranslational modifications. Cell *133*, 930-930 e931.

Kubbutat, M.H., Jones, S.N., and Vousden, K.H. (1997). Regulation of p53 stability by Mdm2. Nature *387*, 299-303.

Kutter, C., Watt, S., Stefflova, K., Wilson, M.D., Goncalves, A., Ponting, C.P., Odom, D.T., and Marques, A.C. (2012). Rapid turnover of long noncoding RNAs and the evolution of gene expression. PLoS Genet *8*, e1002841.

Lam, M.T., Cho, H., Lesch, H.P., Gosselin, D., Heinz, S., Tanaka-Oishi, Y., Benner, C., Kaikkonen, M.U., Kim, A.S., Kosaka, M., *et al.* (2013). Rev-Erbs repress macrophage gene expression by inhibiting enhancer-directed transcription. Nature *498*, 511-515.

Lander, E.S., Linton, L.M., Birren, B., Nusbaum, C., Zody, M.C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., FitzHugh, W., *et al.* (2001). Initial sequencing and analysis of the human genome. Nature *409*, 860-921.

Lane, D.P. (1992). Cancer. p53, guardian of the genome. Nature *358*, 15-16.

Lane, D.P., and Crawford, L.V. (1979). T antigen is bound to a host protein in SV40-transformed cells. Nature *278*, 261-263.

Laptenko, O., and Prives, C. (2006). Transcriptional regulation by p53: one protein, many possibilities. Cell death and differentiation *13*, 951-961.

Lee, B.K., Bhinge, A.A., Battenhouse, A., McDaniell, R.M., Liu, Z., Song, L., Ni, Y., Birney, E., Lieb, J.D., Furey, T.S., *et al.* (2012). Cell-type specific and combinatorial usage of diverse transcription factors revealed by genome-wide binding studies in multiple human cells. Genome Res *22*, 9-24.

Leveille, N., Elkon, R., Davalos, V., Manoharan, V., Hollingworth, D., Oude Vrielink, J., le Sage, C., Melo, C.A., Horlings, H.M., Wesseling, J., *et al.* (2011). Selective inhibition of microRNA accessibility by RBM38 is required for p53 activity. Nat Commun *2*, 513.

Li, W., Notani, D., Ma, Q., Tanasa, B., Nunez, E., Chen, A.Y., Merkurjev, D., Zhang, J., Ohgi, K., Song, X., *et al.* (2013). Functional roles of enhancer RNAs for oestrogen-dependent transcriptional activation. Nature *498*, 516-520.

Lin, J., Chen, J., Elenbaas, B., and Levine, A.J. (1994). Several hydrophobic amino acids in the p53 amino-terminal domain are required for transcriptional activation, binding to mdm-2 and the adenovirus 5 E1B 55-kD protein. Genes & development *8*, 1235-1246.

Linzer, D.I., and Levine, A.J. (1979). Characterization of a 54K dalton cellular SV40 tumor antigen present in SV40-transformed cells and uninfected embryonal carcinoma cells. Cell *17*, 43-52.

Lupien, M., Eeckhoute, J., Meyer, C.A., Wang, Q., Zhang, Y., Li, W., Carroll, J.S., Liu, X.S., and Brown, M. (2008). FoxA1 translates epigenetic signatures into enhancer-driven lineage-specific transcription. Cell *132*, 958-970.

Manolio, T.A., Collins, F.S., Cox, N.J., Goldstein, D.B., Hindorff, L.A., Hunter, D.J., McCarthy, M.I., Ramos, E.M., Cardon, L.R., Chakravarti, A., *et al.* (2009). Finding the missing heritability of complex diseases. Nature *461*, 747-753.

Marques, A.C., and Ponting, C.P. (2009). Catalogues of mammalian long noncoding RNAs: modest conservation and incompleteness. Genome Biol *10*, R124.

Marrone, A., Sokhal, P., Walne, A., Beswick, R., Kirwan, M., Killick, S., Williams, M., Marsh, J., Vulliamy, T., and Dokal, I. (2007). Functional characterization of novel telomerase RNA (TERC) mutations in patients with diverse clinical and pathological presentations. Haematologica *92*, 1013-1020.

Martello, G., Rosato, A., Ferrari, F., Manfrin, A., Cordenonsi, M., Dupont, S., Enzo, E., Guzzardo, V.,

Rondina, M., Spruce, T., *et al.* (2010). A MicroRNA targeting dicer for metastasis control. Cell *141*, 1195-1207.

Mattick, J.S. (2004). RNA regulation: a new genetics? Nature reviews. Genetics *5*, 316-323.

McKinney, K., and Prives, C. (2002). Efficient specific DNA binding by p53 requires both its central and C-terminal domains as revealed by studies with high-mobility group 1 protein. Molecular and cellular biology *22*, 6797-6808.

Melo, C.A., Drost, J., Wijchers, P.J., van de Werken, H., de Wit, E., Oude Vrielink, J.A., Elkon, R., Melo, S.A., Leveille, N., Kalluri, R., *et al.* (2013). eRNAs are required for p53-dependent enhancer activity and gene transcription. Molecular cell *49*, 524-535.

Merkenschlager, M., and Odom, D.T. (2013). CTCF and cohesin: linking gene regulatory elements with their targets. Cell *152*, 1285-1297.

Momand, J., Jung, D., Wilczynski, S., and Niland, J. (1998). The MDM2 gene amplification database. Nucleic Acids Res *26*, 3453-3459.

Momand, J., Zambetti, G.P., Olson, D.C., George, D., and Levine, A.J. (1992). The mdm-2 oncogene product forms a complex with the p53 protein and inhibits p53-mediated transactivation. Cell *69*, 1237-1245.

Mourtada-Maarabouni, M., Pickard, M.R., Hedge, V.L., Farzaneh, F., and Williams, G.T. (2009). GAS5, a non-protein-coding RNA, controls apoptosis and is downregulated in breast cancer. Oncogene *28*, 195-208.

Mousavi, K., Zare, H., Dell'orso, S., Grontved, L., Gutierrez-Cruz, G., Derfoul, A., Hager, G.L., and Sartorelli, V. (2013). eRNAs promote transcription by establishing chromatin accessibility at defined genomic loci. Molecular cell *51*, 606-617.

Mowat, M., Cheng, A., Kimura, N., Bernstein, A., and Benchimol, S. (1985). Rearrangements of the cellular p53 gene in erythroleukaemic cells transformed by Friend virus. Nature *314*, 633-636.

Nakanishi, M., Kaneko, Y., Matsushime, H., and Ikeda, K. (1999). Direct interaction of p21 cyclin-dependent kinase inhibitor with the retinoblastoma tumor suppressor protein. Biochemical and biophysical research communications *263*, 35-40.

Nam, J.W., and Bartel, D.P. (2012). Long noncoding RNAs in C. elegans. Genome Res *22*, 2529-2540.

Nikulenkov, F., Spinnler, C., Li, H., Tonelli, C., Shi, Y., Turunen, M., Kivioja, T., Ignatiev, I., Kel, A., Taipale, J., *et al.* (2012). Insights into p53 transcriptional function via genome-wide chromatin occupancy and gene expression analysis. Cell Death Differ *19*, 1992-2002.

Oda, K., Arakawa, H., Tanaka, T., Matsuda, K., Tanikawa, C., Mori, T., Nishimori, H., Tamai, K., Tokino, T., Nakamura, Y., *et al.* (2000). p53AIP1, a potential mediator of p53-dependent apoptosis, and its regulation by Ser-46-phosphorylated p53. Cell *102*, 849-862.

Parada, L.F., Land, H., Weinberg, R.A., Wolf, D., and Rotter, V. (1984). Cooperation between gene encoding p53 tumour antigen and ras in cellular transformation. Nature *312*, 649-651.

Pasmant, E., Sabbagh, A., Vidaud, M., and Bieche, I. (2011). ANRIL, a long, noncoding RNA, is an unexpected major hotspot in GWAS. Faseb J *25*, 444-448.

Picksley, S.M., and Lane, D.P. (1993). The p53-mdm2 autoregulatory feedback loop: a paradigm for the regulation of growth control by p53? BioEssays : news and reviews in molecular, cellular and developmental biology *15*, 689-690.

Prensner, J.R., Iyer, M.K., Balbin, O.A., Dhanasekaran, S.M., Cao, Q., Brenner, J.C., Laxman, B., Asangani, I.A., Grasso, C.S., Kominsky, H.D., *et al.* (2011). Transcriptome sequencing across a prostate cancer cohort identifies PCAT-1, an unannotated lincRNA implicated in disease progression. Nat Biotechnol *29*, 742-749.

Prensner, J.R., Iyer, M.K., Sahu, A., Asangani, I.A., Cao, Q., Patel, L., Vergara, I.A., Davicioni, E., Erho, N., Ghadessi, M., *et al.* (2013). The long noncoding RNA SChLAP1 promotes aggressive prostate cancer and antagonizes the SWI/SNF complex. Nat Genet *45*, 1392-1398.

Raver-Shapira, N., Marciano, E., Meiri, E., Spector, Y., Rosenfeld, N., Moskovits, N., Bentwich, Z., and Oren, M. (2007). Transcriptional activation of miR-34a contributes to p53-mediated apoptosis. Molecular cell *26*, 731-743.

Rinn, J.L., and Chang, H.Y. (2012). Genome regulation by long noncoding RNAs. Annu Rev Biochem *81*, 145-166.

Rinn, J.L., Kertesz, M., Wang, J.K., Squazzo, S.L., Xu, X., Brugmann, S.A., Goodnough, L.H., Helms, J.A., Farnham, P.J., Segal, E., *et al.* (2007). Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs. Cell *129*, 1311-1323.

Robinson, M.D., McCarthy, D.J., and Smyth, G.K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics *26*, 139-140.

Sang, X., Zhao, H., Lu, X., Mao, Y., Miao, R., Yang, H., Yang, Y., Huang, J., and Zhong, S. (2010). Prediction and identification of tumor-specific noncoding RNAs from human UniGene. Medical oncology *27*, 894-898.

Schmidt, D., Schwalie, P.C., Wilson, M.D., Ballester, B., Goncalves, A., Kutter, C., Brown, G.D., Marshall, A., Flicek, P., and Odom, D.T. (2012). Waves of retrotransposon expansion remodel genome organization and CTCF binding in multiple mammalian lineages. Cell *148*, 335-348.

Simon, M.D., Pinter, S.F., Fang, R., Sarma, K., Rutenberg-Schoenberg, M., Bowman, S.K., Kesner, B.A., Maier, V.K., Kingston, R.E., and Lee, J.T. (2013). High-resolution Xist binding maps reveal two-step spreading during X-chromosome inactivation. Nature *504*, 465-469.

Splinter, E., de Wit, E., Nora, E.P., Klous, P., van de Werken, H.J., Zhu, Y., Kaaij, L.J., van Ijcken, W., Gribnau, J., Heard, E., *et al.* (2011). The inactive X chromosome adopts a unique three-dimensional conformation that is dependent on Xist RNA. Genes Dev *25*, 1371-1383.

Sykes, S.M., Mellert, H.S., Holbert, M.A., Li, K., Marmorstein, R., Lane, W.S., and McMahon, S.B. (2006). Acetylation of the p53 DNA-binding domain regulates apoptosis induction. Molecular cell *24*, 841-851.

Taft, R.J., Pang, K.C., Mercer, T.R., Dinger, M., and Mattick, J.S. (2010). Non-coding RNAs: regulators of disease. The Journal of pathology *220*, 126-139.

Tang, Y., Luo, J., Zhang, W., and Gu, W. (2006). Tip60-dependent acetylation of p53 modulates the decision between cell-cycle arrest and apoptosis. Molecular cell *24*, 827-839.

Tang, Y., Zhao, W., Chen, Y., Zhao, Y., and Gu, W. (2008). Acetylation is indispensable for p53 activation. Cell *133*, 612-626.

Trapnell, C., Pachter, L., and Salzberg, S.L. (2009). TopHat: discovering splice junctions with RNA-Seq. Bioinformatics *25*, 1105-1111.

Tsai, M.C., Manor, O., Wan, Y., Mosammaparast, N., Wang, J.K., Lan, F., Shi, Y., Segal, E., and Chang, H.Y. (2010). Long noncoding RNA as modular scaffold of histone modification complexes. Science *329*, 689-693.

Tuan, D., Kong, S., and Hu, K. (1992). Transcription of the hypersensitive site HS2 enhancer in erythroid cells. Proceedings of the National Academy of Sciences of the United States of America *89*, 11219-11223.

Ulitsky, I., Shkumatava, A., Jan, C.H., Sive, H., and Bartel, D.P. (2011). Conserved function of lincRNAs in vertebrate embryonic development despite rapid sequence evolution. Cell *147*, 1537-1550.

Vance, K.W., Sansom, S.N., Lee, S., Chalei, V., Kong, L., Cooper, S.E., Oliver, P.L., and Ponting, C.P. (2014a). The long non-coding RNA Paupar regulates the expression of both local and distal genes. Embo J *33*, 296-311.

Vance, K.W., Sansom, S.N., Lee, S., Chalei, V., Kong, L., Cooper, S.E., Oliver, P.L., and Ponting, C.P. (2014b). The long non-coding RNA Paupar regulates the expression of both local and distal genes. The EMBO journal *33*, 296-311.

Vassilev, L.T., Vu, B.T., Graves, B., Carvajal, D., Podlaski, F., Filipovic, Z., Kong, N., Kammlott, U., Lukacs, C., Klein, C., *et al.* (2004). In vivo activation of the p53 pathway by small-molecule antagonists of MDM2. Science *303*, 844-848.

Visel, A., Blow, M.J., Li, Z., Zhang, T., Akiyama, J.A., Holt, A., Plajzer-Frick, I., Shoukry, M., Wright, C., Chen, F., *et al.* (2009). ChIP-seq accurately predicts tissue-specific activity of enhancers. Nature *457*, 854-858.

Voorhoeve, P.M., le Sage, C., Schrier, M., Gillis, A.J., Stoop, H., Nagel, R., Liu, Y.P., van Duijse, J., Drost, J., Griekspoor, A., *et al.* (2006). A genetic screen implicates miRNA-372 and miRNA-373 as oncogenes in testicular germ cell tumors. Cell *124*, 1169-1181.

Vousden, K.H., and Lu, X. (2002). Live or let die: the cell's response to p53. Nat Rev Cancer *2*, 594-604.

Vulliamy, T.J., Marrone, A., Knight, S.W., Walne, A., Mason, P.J., and Dokal, I. (2006). Mutations in dyskeratosis congenita: their impact on telomere length and the diversity of clinical presentation. Blood *107*, 2680-2685.

Wang, B., Xiao, Z., and Ren, E.C. (2009). Redefining the p53 response element. Proc Natl Acad Sci U S A *106*, 14373-14378.

Wang, D., Garcia-Bassets, I., Benner, C., Li, W., Su, X., Zhou, Y., Qiu, J., Liu, W., Kaikkonen, M.U., Ohgi, K.A., *et al.* (2011a). Reprogramming transcription by distinct classes of enhancers functionally defined by eRNA. Nature *474*, 390-394.

Wang, K.C., Yang, Y.W., Liu, B., Sanyal, A., Corces-Zimmerman, R., Chen, Y., Lajoie, B.R., Protacio, A., Flynn, R.A., Gupta, R.A., *et al.* (2011b). A long noncoding RNA maintains active chromatin to coordinate homeotic gene expression. Nature *472*, 120-124.

Wang, X.S., Zhang, Z., Wang, H.C., Cai, J.L., Xu, Q.W., Li, M.Q., Chen, Y.C., Qian, X.P., Lu, T.J., Yu, L.Z., *et al.* (2006). Rapid identification of UCA1 as a very sensitive and specific unique marker for human bladder carcinoma. Clin Cancer Res *12*, 4851-4858.

Weber, J.D., Taylor, L.J., Roussel, M.F., Sherr, C.J., and Bar-Sagi, D. (1999). Nucleolar Arf sequesters Mdm2 and activates p53. Nature cell biology *1*, 20-26.

Wei, C.L., Wu, Q., Vega, V.B., Chiu, K.P., Ng, P., Zhang, T., Shahab, A., Yong, H.C., Fu, Y., Weng, Z., *et al.* (2006). A global map of p53 transcription-factor binding sites in the human genome. Cell *124*, 207-219.

Weinberg, R.L., Veprintsev, D.B., Bycroft, M., and Fersht, A.R. (2005). Comparative binding of p53 to its promoter and DNA recognition elements. Journal of molecular biology *348*, 589-596.

Wolf, D., and Rotter, V. (1984). Inactivation of p53 gene expression by an insertion of Moloney murine leukemia virus-like DNA sequences. Molecular and cellular biology *4*, 1402-1410.

Yang, L., Lin, C., Jin, C., Yang, J.C., Tanasa, B., Li, W., Merkurjev, D., Ohgi, K.A., Meng, D., Zhang, J., *et al.* (2013). lncRNA-dependent mechanisms of androgen-receptor-regulated gene activation programs. Nature *500*, 598-602.

Yap, K.L., Li, S., Munoz-Cabello, A.M., Raguz, S., Zeng, L., Mujtaba, S., Gil, J., Walsh, M.J., and Zhou, M.M. (2010). Molecular interplay of the noncoding RNA ANRIL and methylated histone H3 lysine 27 by polycomb CBX7 in transcriptional silencing of INK4a. Molecular cell *38*, 662-674.

Zamore, P.D., Tuschl, T., Sharp, P.A., and Bartel, D.P. (2000). RNAi: double-stranded RNA directs the ATP-dependent cleavage of mRNA at 21 to 23 nucleotide intervals. Cell *101*, 25-33.

Zhang, X., Zhou, Y., Mehta, K.R., Danila, D.C., Scolavino, S., Johnson, S.R., and Klibanski, A. (2003). A pituitary-derived MEG3 isoform functions as a growth suppressor in tumor cells. J Clin Endocrinol Metab *88*, 5119-5126.

Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nusbaum, C., Myers, R.M., Brown, M., Li, W., *et al.* (2008). Model-based analysis of ChIP-Seq (MACS). Genome Biol *9*, R137.

Zhang, Y., Xiong, Y., and Yarbrough, W.G. (1998). ARF promotes MDM2 degradation and stabilizes p53: ARF-INK4a locus deletion impairs both the Rb and p53 tumor suppression pathways. Cell *92*, 725-734.

Annexes

Table S1
Primers and probes used for chapter two

Primers	**5'--> 3'**
Metylation	
LED-Bisulfite seq-Fw 1	ATTTTGGGAAATGGTTTTAAAG
LED-Bisulfite seq-Rev 1	CCCAAACCAACATAAATCC
LED-Bisulfite seq-Fw 2	ATTTATGTTGGTTTGGG
LED-Bisulfite seq-Rev 2	ATTCCCTCCTTCTAATTTC
LED-MSP-M-Fw	GGTACGTCGAGTAGGCGGGGTC
LED-MSP-M-Rev	CTACGAAAAACCACCCCCCCG
LED-MSP-U-Fw	GGGGTATGTTGAGTAGGTGGGGTT
LED-MSP-U-Rev	ACTACAAAAAACCACCCCCCCAA
QPCR	
QPCR-LED-Fw	TTCTGCTGCGGCGACTTGAC
QPCR-LED-Rev	CTTCTTCTCCTCTTGTTCAG
QPCR-p21-Fw	TACCCTTGTGCCTCGCTCAG
QPCR-p21-Rev	GAGAAGATCAGCCGGCGTTT
QPCR-p53-Fw	ATGGAGGAGCCGCAGTCAGAT
QPCR-p53-Rev	GCAGCGCCTCACAACCTCCGTC
QPCR-p21e-Fw	ATGAGACTTGCCCCTGAGTCATAC
QPCR-p21e-Rev	TAACCTCTATCCACCCTCTGCTAG
QPCR-MAP2e-Fw	CCACAGCAACACCCAATTCTTATT
QPCR-MAP2e-Rev	TCTGTGTTTGTGCCTATTTTCCAC
QPCR-TOB1e-Fw	GAAGTTTGTTTTGGACAGTGTTTG
QPCR-TOB1e-Rev	AGTCTCTCCATAAAGCCTGATGTT
QPCR-KYNUe-Fw	GTTATGACACACCTATGGGACAAA
QPCR-KYNUe-Rev	TCAGGTCTACAAAACTCACTTTGC
QPCR-KYNUe-Fw	GCTCAGATTTTGATCACCAGATAA
QPCR-KYNUe-Rev	GGGTTCATACTGTGTGAGAGAAGA
QPCR-STAT4e-Fw	CTTTATGGGATTTCATGTCTGGA
QPCR-STAT4e-Rev	ATCTTTGTCCATTGTCAATCATTT
QPCR-RP11-3P17.4e-Fw	TTCTTGTGGTAGGAAGGAGGTAAC
QPCR-RP11-3P17.4e-Rev	CCAGAAGGAAATAGAGCAGAAAAC
QPCR-POU5F1Be-Fw	ATTACCTGTATGAAGGGACCGTAG
QPCR-POU5F1Be-Rev	TAGATTCCAGGATAGTGCTTCTCC
QPCR-FOXC1e-Fw	CATGAAAGGTGAAGCGGAAATAC
QPCR-FOXC1e-Rev	TGAAGGAGCAGGTGAAACG
QPCR-GAPDH-Fw	TGCACCACCAACTGCTTAGC
QPCR-GAPDH-Rev	GGCATGGACTGTGGTCATGAG
QPCR-linc000475-Fw	CTGCAGGAGGCTACACAGTC
QPCR-linc000475-Rev	CAATTCAAATGACCACATGC
QPCR-LOC643401-Fw	CCAACAATCTCTGGCTGCA

Table S1 (continued)

Primers	5'--> 3'
QPCR	
QPCR-LOC643401-Rev	CACATTACCACAGGAGCTGA
QPCR-RP3-510D11.2-Fw	CCAGACCGACGGGACAGCG
QPCR-RP3-510D11.2-Rev	GCTTCCCTGTCCTCCTCCTA
QPCR-S14-Fw	GGCAGACCGAGATGAATCCTC
QPCR-S14-rev	CAGGTCCAGGGGTCTTGGTCC
QPCR-U2-Fw	CATCGCTTCTCGGCCTTTTG
QPCR-U2-rev	TGGAGGTACTGCAATACCAGG
ChIP	
p21-prom-proximal-Fw	GAGGAAGAAGACTGGGCATGTCT
p21-prom-proximal-Rev	GCTTGGAGCAGCTACAATTACTGAC
MDM2 prom-Fw	GGGCTATTTAAACCATGCATTTTC
MDM2 prom-Rev	GTCCGTGCCCACAGGTCTA
BBC3 prom-Fw	GCGAGACTGTGGCCTTGTGT
BBC3 prom-Rev	CGTTCCAGGGTCCACAAAGT
FAS prom-Fw	ACAGGAATTGAAGCGGAAGTCT
FAS prom-Rev	GAGTTCCGCTCCTCTCTCCAA
GADD45 prom-Fw	AGCGGAAGAGATCCCTGTGA
GADD45 prom-Rev	CGGGAGGCAGTGCAGATG
CCNG1 prom-Fw	ACCTCCTGCCTTCCTGTCTTTA
CCNG1 prom-Rev	GCACAAGCTAAGCCCTGGAT
RRM2B prom-Fw	AAGAGAAGCCAGCTGAGATTAAAA
RRM2B prom-Rev	TTTAGCCAAATCACATCATTCCTA
ZMAT3 prom-Fw	CTCTGGATAAAGGAAATCAGCAAT
ZMAT3 prom-Rev	GGATCTTTATTTTGAGGGAAACCT
siRNAs	
LED-kd1	CCUCGGUGAUGCUUAACGUC
LED-kd2	CUUCCCCUACUUCUACAUCG
p21e-kd1	UGGAGCCCACAGUUUCACGU
p21e-kd2	ACCCUCUGCUAGGCAUGAGU
RP3-510D11.2-KD2	CUCCACCUGCGCAAAACGAA
RP3-510D11.2-KD3	GGGCGGAUUCUCCUUGGACU
RP3-510D11.2-KD4	GCGGCCAUCAGUAACAGCAA
LOC643401-KD1	CAGCUCCUGUGGUAAUGUGG
LOC643401-KD2	UCACCAGGGUGAGUGAACAA
LOC643401-KD3	GUGUGAGGUGGUGCUAUGUA
LOC643401-KD4	AGGAGCACAAUUCUGCAACA
ChIRP	modified with C3(biotin)
LED-1	GTAGTTTCTGGAGCCCGATG

Table S1 (continued)

Primers	**5'--> 3'**
ChIRP	modified with C3(biotin)
LED-2	CCCTCTTTGGCCAATATGAG
LED-3	CAAATGCCAATGCCTTTCTG
LED-4	GGGAACTCAGACCTTCCTTC
LED-5	CAAGGAAAGCGACATTCCTA
LED-6	GATCTTCCCTGTGACTTTCC
LED-7	AGCATTGTACACCAGTTAGC
LED-8	GTTTTCAGTCCCAGGAACTG
LED-9	CAGCTTCATTCAGGACCTGA
LED-10	TGTCACTCTTGTCCCATAGA
LED-11	ATCTGTGCATCAGACACAAG
LED-12	TGACACCAGGGTTTGAAAGG
LED-13	CTCAGCAAAGAGCTTTCCTC
LED-14	CGCAACTTTATCCCAGTAGG
LED-15	GACTTCCGTGACCCTCTTAT
LED-16	ACCGCCTAGAAAACAGACTG
LED-17	GTATGTCATTCCGGTCAGTG
LED-18	ACACAGTGAATCACACACAG
LED-19	AAACAAGAACAGCTCTCTGG
LED-20	ACAAGAAGACGTGACAACAC
LED-21	GCCTCCAGTTGGGATTTTAG
LED-22	CTCTGAACAGGAAACACGGT
LED-23	CAGGTCAAGTCGGTTCTCAG
LED-24	CGCCAGACTCTAAAAAGGAA
LED-25	CCCTCAGGATTCTACTGACT
LED-26	CTGTGTCGAGTCTGTTTCAG
LED-27	GAGAGGAGTCTCAGTGCTAC
LED-28	GATTCTGCAGTAACCTTGCT
LED-29	GCCAGGTCCTTTATCTTCTC
LED-30	TCTGATGGGACAGAATTTGC
LED-31	CTCCTTGTCAGGATGCTCTG
LED-32	CATTAGCACCAGTGCGTTTG
LED-33	CAAGGACTGGGGTCCATTTC
LED-34	GGGGATAAAGAGGACCATCT
LED-35	ATCTTGCTTTGCCAAGTCAT
LED-36	TTTTCCTCTGTAGCAGGGAT
LED-37	TAAAAACTCCAGGTTGTGCA
LED-38	CGTATTCACTGCTTACAAGC
LED-39	TTCTTCTTGGGGCTCTGTTT

Table S1 (continued)

Primers	5'--> 3'
ChIRP	modified with C3(biotin)
LED-40	CACTGGTAGCTCAGGAAAAG
LED-41	GGTCATCTTCCATCTTTGCT
LED-42	GGATTTCTTCTGGACCTCAC
LED-43	CATTTGGAGAAGGCTCTCAG
LED-44	AAGACAGGTAGACCCATCTC
LED-45	GAAACTTCCCACTTAGCTGC
LED-46	ATGTATTTGCTCTAGTGCCC
LED-47	GGTGCACACAAAGTGAAAAT
LED-48	GCGGAATCAGATTTCAAAGT
RNA FISH	modified with Quasar 570
LED-1	GTAGTTTCTGGAGCCCGATG
LED-2	CCCTCTTTGGCCAATATGAG
LED-3	CAAATGCCAATGCCTTTCTG
LED-4	GGGAACTCAGACCTTCCTTC
LED-5	CAAGGAAAGCGACATTCCTA
LED-6	GATCTTCCCTGTGACTTTCC
LED-7	AGCATTGTACACCAGTTAGC
LED-8	GTTTTCAGTCCCAGGAACTG
LED-9	CAGCTTCATTCAGGACCTGA
LED-10	TGTCACTCTTGTCCCATAGA
LED-11	ATCTGTGCATCAGACACAAG
LED-12	TGACACCAGGGTTTGAAAGG
LED-13	CTCAGCAAAGAGCTTTCCTC
LED-14	CGCAACTTTATCCCAGTAGG
LED-15	GACTTCCGTGACCCTCTTAT
LED-16	ACCGCCTAGAAAACAGACTG
LED-17	GTATGTCATTCCGGTCAGTG
LED-18	ACACAGTGAATCACACACAG
LED-19	AAACAAGAACAGCTCTCTGG
LED-20	ACAAGAAGACGTGACAACAC
LED-21	GCCTCCAGTTGGGATTTTAG
LED-22	CTCTGAACAGGAAACACGGT
LED-23	CAGGTCAAGTCGGTTCTCAG
LED-24	CGCCAGACTCTAAAAAGGAA
LED-25	CCCTCAGGATTCTACTGACT
LED-26	CTGTGTCGAGTCTGTTTCAG
LED-27	GAGAGGAGTCTCAGTGCTAC
LED-28	GATTCTGCAGTAACCTTGCT

Table S1 (continued)

Primers	5'--> 3'
RNA FISH	modified with Quasar 570
LED-29	GCCAGGTCCTTTATCTTCTC
LED-30	TCTGATGGGACAGAATTTGC
LED-31	CTCCTTGTCAGGATGCTCTG
LED-32	CATTAGCACCAGTGCGTTTG
LED-33	CAAGGACTGGGGTCCATTTC
LED-34	GGGGATAAAGAGGACCATCT
LED-35	ATCTTGCTTTGCCAAGTCAT
LED-36	TTTTCCTCTGTAGCAGGGAT
LED-37	TAAAAACTCCAGGTTGTGCA
LED-38	CGTATTCACTGCTTACAAGC
LED-39	TTCTTCTTGGGGCTCTGTTT
LED-40	CACTGGTAGCTCAGGAAAAG
LED-41	GGTCATCTTCCATCTTTGCT
LED-42	GGATTTCTTCTGGACCTCAC
LED-43	CATTTGGAGAAGGCTCTCAG
LED-44	AAGACAGGTAGACCCATCTC
LED-45	GAAACTTCCCACTTAGCTGC
LED-46	ATGTATTTGCTCTAGTGCCC
LED-47	GGTGCACACAAAGTGAAAAT
LED-48	GCGGAATCAGATTTCAAAGT
Northern probes	modified with TEG-biotin
LED	CTTTATTTATTTGCAGTCTGTTTTCTAGGC
p21e	CCTTTGGCCTTTTTAGCTGTATGACTCAGG