國立台灣大學醫學院分子醫學研究所

碩士論文

Graduate Institute of Molecular Medicine

College of Medicine

National Taiwan University

Master Thesis

Daxx 調控非編碼長 RNA 之鑑定及分析

Identification and analysis of Daxx-regulated large intergenic non-coding RNAs

黃天祺

Huang, Tien-Chi

指導教授: 施修明 博士

Advisor: Shih, Hsiu-Ming, Ph.D

中華民國 101 年 7 月

July, 2012

誌謝

人生許多不如意之事，回頭看或許都是上天安排給你的必經之路，念碩士班之前歷經了申請國外學校的各種不順遂，當時相當不能理解，但在施修明老師的實驗室兩年的學習，獲益良多也對自己有了重新的認識，感謝施老師對我的指導以及想法上的包容，讓我了解到自己的不足，也謝謝老師對於我申請學校的支持，讓我順利完成夢想的第一步。

感謝口試委員中央研究院基因體中心阮麗蓉老師，台大分子醫學所徐立中老師對本篇論文提供許多寶貴意見，使本論文更臻完善。

謝謝 R605 實驗室每一個人對我的包容與協助，在這裡真的學到很多。謝謝彥淞學長，在碩士班期間幾乎無私的指導，不管是實驗設計或是技術上的分享或是生活上的建議都給我很大的幫助，感謝 Susan 學姊，在我剛進實驗室對動物操作一無所知的時候細心的指導，謝謝立偉，在繁重的課業中，一起奮戰，無人能敵的冷笑話更成為夏日消暑和實驗空檔的最佳良伴。

謝謝碩士班各位同學的支持與互相的鼓勵，友思、澤翔、小蜂、書維、安洲、宜妙、俞靜，我們一起度過許多難關，希望大家未來都有好的發展。最後，由衷感謝我的家人，在研究和申請留學之路給我相當的包容和支持，如果沒有你們的體諒，我可能沒有辦法繼續堅持我的理想。

謹以此論文獻給曾經幫助過我的人

Contents

List of figures

List of Tables

中文摘要

Daxx 蛋白質最初被證明在細胞質內與 Fas 受體的凋亡區段互相結合，並扮演訊息傳遞重要的角色，然而許多報導指出，Daxx 主要在細胞核內執行基因轉錄調節的功能。為了更進一步瞭解 Daxx 對於基因表現的影響，本實驗利用微陣列晶片系統性的找尋 Daxx 調控的編碼基因(protein-coding gene) 及長非編碼 RNA (large intergenic non-coding RNA)。分析結果指出當抑制 Daxx 表現時，顯著地調控 655 個基因的表達，進一步利用基因功能性分析，發現其中許多基因與細胞型態及癌症生成高度相關。此結果暗示 Daxx 可能參與癌症轉移的調控。另一方面，當抑制 Daxx 表現時，細胞內有 104 個長非編碼 RNA 表現量出現顯著差異。進一步利用生物資訊分析方法，篩選出 Daxx 可能調控的長非編碼 RNA，我們並利用補救實驗及反轉錄定量聚合酶鏈式反應(RT-qPCR) 驗證該分析結果。結果證實 12 個長非編碼 RNA 受到 Daxx 調控，其中包含功能已知的 lincRNA, 如 JPX, NEAT1 及 MIAT。接著，我們利用 RNA 干擾技術發現其中一個長非編碼 RNA，linc4971 的表現對於其上游基因,ZNF703 及 ERLIN2 具有抑制的作用。整體來說，本實驗結果除了提供 Daxx 可能調控的標的基因及長非編碼 RNA，並指出 Daxx 可藉由調控長非編碼 RNA 影響下游基因的表現。

關鍵詞: Daxx、微陣列晶片分析、長非編碼 RNA

English Abstract

The death domain-associated protein (Daxx) participates in various biological processes depending on its sub-cellular localization. In the nucleus, Daxx, as a transcriptional coregulator, interacts with various proteins, including transcription factors, to regulate gene expression. Although many Daxx-regulated genes have been reported, a genome-wide analysis of gene expression profile regulated by Daxx largely remains unclear. Here, we used microarray analysis to identify Daxx-regulated protein-coding genes as well as non-coding RNAs. 655 genes were significantly regulated in Daxx knockdown cells. Gene ontology analysis demonstrated that Daxx-regulated genes showed significant asssociation with cell morphology and cancer, suggesting that Daxx may play a role in tumor formation. Moreover, by combining microarray data and bioinformatic analysis, we identified 45 Daxx-regulated large intergenic non-coding RNAs (lincRNAs). Some identified lincRNAs, such as JPX, NEAT1 and MIAT, are functionally well-known, while most are recently defined transcripts. Notably, knockdown of identified lincRNA linc-4971 resulted in a activation of nearby gene expression, suggesting that this lincRNA may act as regulatory node in Daxx-mediated transcriptional pathways.

Keywords: Daxx, large intergenic non-coding RNA, microarray

Chapter I

Introduction

1. Daxx

Overview of Daxx

The death domain-associated protein (Daxx) is a highly conserved protein whose homologue can be found from zebra fish to human, and ubiquitously expressed in different organs, with particularly high expression in the thymus and testis [1]. Daxx was originally identified as a binding protein to FAS receptor in the cytoplasm to positively regulate FAS-induced apoptosis [2]. Since then, numerous reports have supported the idea of Daxx having a pro-apoptotic role on various stimuli such as FAS, oxidative stress, TGF-β treatment, or UV light [3, 4]. However, several studies also suggested that Daxx might bear anti-apoptotic function. One of the supporting evidence is that Daxx knockout mice showed extensive apoptosis and died in E9.5 [5]. Furthermore, knockdown Daxx with RNAi strategy in cell lines increased apoptosis [6]. Although the debate still remains unresolved, after a decade's research, Daxx has been generally viewed as a multifunctional protein that participates in different biological processes, depending on its sub-cellular localization. Recent identification of the chaperone activity of Daxx has further broadened the function of this controversial protein.

Nuclear localization of Daxx

Daxx proteins predominantly accumulate in the nucleus and co-localize with promyelocytic leukemia nuclear body (PML-NB) and other subnuclear domains [7, 8]. A large portion of Daxx is accumulated in PML-NBs because of its interaction with sumoylated PML protein. This notion was supported by several studies. In PML null cells, Daxx dispersed in the nucleoplasm and associated with heterochromatin. Its ability to induce apoptosis was also reduced [7]. Moreover, Daxx contains a sumo-recognition motif (SIM) resided in the carboxyl terminus, which is responsible for binding to sumoylated PML protein and accumulation in PML-NBs [9]. Since Daxx interacts with numerous proteins in a sumoylation-dependent manner, it was proposed that the accumulation of Daxx in PML-NBs would exclude Daxx binding to other sumoylated factors causing alternation of Daxx function. Indeed, transient overexpression of PML can increase Daxx accumulation and relieve Daxx-mediated gene repression [10, 11], while de-sumoylated PML upon heat shock leads to release Daxx from PML-NBs and enhances transcription repression of Hsp25 [12].

Daxx at heterochromatin

Daxx was reported to associate with heterochromatin in PML null cells. The first evidence showing Daxx association with chromatin is the discovery that Daxx and

HDAC2 together can form a complex with chromatin-associated protein Dek [13]. In addition, Daxx can co-localize and form a complex with α-thalassemia/mental retardation syndrome (ATRX), a member of SWI/SNF, ATP-dependent chromatin remodeler in heterochromatic chromatin-dense regions [14, 15]. Daxx and ATRX can be deposited at PML NBs in G1 and G2 phase, and accumulated in heterochromatin in S phase [15]. However, the exact regulation and function still remain elusive. Interestingly, Daxx can cooperate with ATRX to deposit histone variant H3.3 in telomere and pericentric heterochromatin [16, 17]. Further investigation is needed to link the physiological meaning of Daxx-ATRX localization and their molecular function.

Daxx in the cytoplasm

The sub-cellular localization of Daxx has been a controversial issue since it was discovered as a factor associated with Fas-induced apoptosis [2]. Accumulated evidence, including gel filtration experiments and the immunostaing, indicates that Daxx is a predominant nuclear protein. Immunofluaescence analysis clearly shows a punctate staining with little diffuse pattern in the nucleus using Daxx specific antibody. Moreover, cellular fractionation of different cell lines demonstrated that Daxx largely associates with nuclear fraction [18, 19]. Nevertheless, Daxx can interact with several

proteins, such as apoptosis signal regulating kinase 1 (ASK1), and both proteins are co-localized in the cytoplasm, in correlation with the induction of apoptosis [20]. It should be noted, however, that the bulk of these experiments were based on transient over-expression under various stress conditions [21-23]. How Daxx functions in different sub-cellular localization remains to be determined.

Daxx in transcription regulation

Although Daxx does not show DNA binding ability, it regulates gene transcription via various mechanisms. Firstly, Daxx was shown to associate with multiple proteins involved in transcriptional repression, such as HDAC1 [10], HDAC2 [13], DNA methyltransferase 1 (DNMT1) and its associated protein DNMT1-associated protein (DMAP1)[24, 25]. Secondly, Daxx can regulate gene transcription via interaction with a large number of transcription factors such as Ets1 [26], Pax5 [27], p53 and p53 family members p73 and p63 [28, 29], glucocorticoid receptor [30], androgen receptor [31], Smad4 [32], STAT3[33], RelB [34], RelA (p65) [35], TCF4 [36] (see reviews for more details [37, 38]). With regards to mechanism, a significant number of studies have revealed that Daxx SIM is essential for recognizing sumoylated transcription factors, thereby leading to gene repression. Thirdly, Daxx can affect downstream gene expression via protein stability regulation. Daxx was reported to stabilize Mdm2 and

enhance its intrinsic E3 activity towards p53. The result suggested that disruption of the Mdm2-Daxx interaction might be important for p53 activation in response to DNA damage [39]. Finally, Daxx can facilitate gene activation through its chaperone activity in neuron [40].

Daxx with histone chaperon activity

Recently, Daxx was identified as a novel histone chaperon which could specifically interact with and deposit histone variant H3.3 to certain chromosomal regions [16, 17, 41]. Using biochemical analyses, Drane *et al*. [16] have found that Daxx and ATRX tightly associated with histone H3.3 but not H3.1 and are capable of incorporating H3.3-H4 onto DNA in vitro. This study also indicated that Daxx is associated with H3.3 deposition in pericentromeric heterochromatin in mouse embryonic fibroblasts (MEFs). Goldberg *at el*. [17], using ChIP-Seq technology, revealed that H3.3 was enriched at the direct vicinity of the transcription start sites of genes as well as gene body, and highly correlated with markers with transcription. In addition, H3.3 was enriched in the (TTAGGG)n repeat, which is the conserved telomeric DNA sequence. The former enrichment was HIRA dependent, as H3.3 was absent in protein-coding regions in HIRA-deficient cells, whereas Daxx and ATRX contributed to the latter enrichment. While the majority of H3.3 was deposited by HIRA at protein coding loci, Daxx was

shown to target H3.3 to genic and intergenic regulatory elements and telomere [17, 41]. Combined together, these researches uncovered a novel function of Daxx that could cooperate with ATRX in chromatin assembly and the genome integrity maintenance.

Daxx and diseases

Since Daxx participates in various biological processes, such as apoptosis, transcription regulation and chromatin assembly, deregulation of Daxx function may lead to pathogenesis. Indeed, some evidence suggests a functional implication between Daxx and neurodegenerative diseases, including Alzheimer's disease [42]. In addition, Daxx and its association with PML imply that Daxx may be involved in acute promyelocytic leukemia (APL), which frequently possessingPML/RAR-alpha fusion protein due to the chromosomal translocation. Nevertheless, despite speculations that Daxx may be involved in human diseases, the challenge remains because of lack of loss-off function animal modal to dissect Daxx function *in vivo*.

It was only recently that exome sequencing studies have demonstrated that ATRX and Daxx mutation frequently occurs in pancreatic neuroendocrine tumors (PanNETs) – a class of tumors that arise from islet cells [43]. From the 68 patients examined, 30 carried mutations in MEN1, 17 in Daxx and 12 in ATRX. Jiao *et al.* [43] found framshifts and nonsense mutations that caused loss of ATRX and Daxx protein in a large

percentage of PanNET patients, which suggests inactivation of this complex is required for tumorigenesis – at least in PanNET. The loss of ATRX/Daxx protein was further correlated with alternative lengthening of telomere (ALT) in PanNET and other cancer types [44]. Moreover, mutations in H3.3 and ATRX/Daxx were also found in paediatric glioblastoma, a lethal brain tumor, in a large scale screening, which suggests defects in histone variants, or its deposition, responsible for the telomere integrity, is crucial for carcinogenesis.

2. LincRNAs

Overview of large intergenic non-coding RNAs (lincRNAs)

RNA used to be viewed as a genetic information carrier responsible for being a messenger between DNA and protein. It is now clear that RNA has diverse functions beyond just being an intermediate. High-throughput sequencing data shows the whole genome is pervasively transcribed, while only 1.5% transcripts among genome is protein coding, which suggests such “dark matters” may participate in the regulatory network of various biological processes [45]. Indeed, a subset of RNA, called large intergenic non-coding RNA (lincRNAs) whose length is greater than 200nt has been discovered and characterized for many physiological roles such as gene regulation, development and cancer metastasis [46]. Although lacking open reading frame (ORF), these lincRNAs have mRNA-like characteristics such as splicing and polyadenylation. Moreover, the vast majority of lincRNAs express in a tissue-specific manner, suggesting that those transcripts can be regulated under different cellular contexts.

Identification and annotation of lincRNAs

After the completion of genome project, tilling microarray studies suggested that mammalian cells produced many thousands of long non coding transcripts [47]. These

transcripts have been considered as transcriptional noise due to little or no evolutionary conservation [48, 49]. Recently, by using RNA-sequencing (RNA-seq) technology and computational algorithms, over thousands of lincRNAs were identified and annotated [50]. The abundance of such transcripts suggests that thesemay participate in genome function. Moreover, chromatin map also revealed H3K4me3 and H3K36me3 are common features of over thousand lincRNAs which help us to predict these transcripts systematically [51]. In 2011, John Rinn group integrated RNA–seq data from different resources to comprehensively annotate and characterize human lincRNA [52]. They identified 8195 putative lincRNAs, of which 4662 lincRNAs were defined as a stringent set. Most of lincRNAs are expressed in a highly tissue-specific manner. Moreover, 414 linRNAs reside within intergenic regions associated with specific diseases by genome wide association studies (GWAS). In sum, combined high-throughput sequencing data, chromatin mapping and computational methods in transcriptome reconstruction, numerous human lincRNAs were annotated to be as references to facilitate experimental manipulation and uncover lincRNA function.

Biological function of lincRNAs

The biological significance of linc RNA was first revealed by identification of Xist, a 17 kilobase lincRNA, coating most of inactive X chromosome leading to gene

repression [53]. Regulatory roles of lincRNA were also identified in imprinted clusters. Genomic imprinting is an epigenetic mechanism that regulates a subset of gene expressing allelic specifically in a parent-origin dependent manner. Interestingly, numerous lincRNAs such as H19, Kcnq1ot1, AIR and other lincRNAs were found within imprinting loci responsible for allelic-specific repression. In addition, 231 lincRNAs were characterized in the four human HOX loci which are expressed spatially along developmental axes. Some of these lincRNAs have been shown to play roles both in developmental gene repression and activation [54, 55]. One of which residing in the HOXC locus, termed HOTAIR, can repress transcription *in trans* across 40 kilobases of the HOXD locus [54], while the other one, residing in HOXA locus, called HOTTIP, can activate distally expressed genes [55]. Rapidly growing number of studies further revealed the functions and molecular mechanisms of lincRNAs in regulating pluripotency, p53 pathway as well as cancer metastasis [56-59].

LincRNAs and chromatin modifiers

lincRNAs can cooperate with different chromatin modifiers to exert their functions such as dosage compensation, genomic imprinting and homeotic gene expression. The most well-known associated complex in mammal is Polycomb Repressive Complex 2 (PRC2) which is responsible for catalyzing trimethylation of H3 lysine 27 (H3K27em3).

PRC2 complex contains four core subunits including Eed, Suz12, RbAp48 and catalytic subunit Ezh2. Although PRC2 is important in dictating epigenetic state and gene expression pattern of various mammalian cells, how PRC2 targets to specific loci is still elusive, due to lack of Polycomb-response element (PRE) as identified in Drosophila. The identification of PRC2-associated lincRNAs may provide missing link responsible for PRC2 targeting in the mammalian genome. For instance, Tsix, RepA, and Xist RNAs target Polycomb Repressive Complex 2 (PRC2) *in cis* to achieve X chromosome silencing [60]. LincRNAs also mediate allele-specific repression of imprinting domain via histone mehthyltransferase. Kcnq1ot1, a 90kb lincRNA, expressed from paternal allele, can interact with histone methyltransferase G9a and Ezh2 to silence a cluster of gene in Kcnq1 imprinted locus. Moreover, a transcript from Igfr cluster, called AIR, can recruit G9a to achieve allele-specific repression of Slc22a3, Slc22a2, and Igf2r genes in mouse placenta [61]. In terms of transcripts from HOX loci, HOTAIR not only can bind and target PRC2 to silence HOXD but also alter PRC2 genome-wide occupancy to surpress metastasis suppressor gene expression [54, 58]. Furthermore, HOTAIR is a scaffold PRC2 and LSD1/CoREST/REST complexes which modulate H3 lysine 27 trimethylation and H3 lysine 4 demethylation simultaneously [62].

Inspired by the observation that HOTAIR can bind to PRC2, many lincRNAs was demonstrated physical association with PRC2 using the customized RIP ChIP assay [63].

Moreover, combined native RNA immunoprecipitations (RIP) and RNA-seq, Zhao *et al.* demonstrated that approximately 20% of total lincRNAs can associate with PRC2 complex in mouse embryonic stem cells (mES cells) [64]. Besides PRC2 and G9a, many other chromatin complexes critical for ES cells can associate with lincRNA in a systematic screen [56]. In conclusion, accumulated evidence shows chromatin modifiers can bind to numerous lincRNAs which may alter the chromatin modifier targeting, modulate the enzyme activity or change components of such complex.

Enhancer function of lincRNAs

Recent studies reported that short transcripts derived from enhancer elements, termed eRNAs, that are most likely not polyadenylated [65]. The level of eRNA expression at these enhancers positively correlates with the level of messenger RNA synthesis at nearby genes, suggesting that eRNA synthesis occurs specifically at enhancers promoting mRNA synthesis[65, 66]. Similarly, a group of lincRNA was identified with an enhancer-like function in many human cell lines [67]. Depletion of a number of lincRNAs led to decreased expression of their neighboring protein-coding genes, including several critical regulators in cell differentiation. Of note, the enhancer function of lincRNAs is orientation independent as classical enhancers. Taken together, the results suggest lincRNAs may function as an enhancer to activate nearby genes *in*

cis.

LincRNAs in the nuclear body

The nucleus of eukaryotes is a highly organized structure and is consisted of different nuclear bodies, including nucleoli, nuclear speckles, paraspeckles, PML bodies, and Cajal bodies. The compartmentalization of the nucleus is thought to provide a platform for clustering different components exerting various functions [68]. Several lincRNAs stably, abundantly localized in distinctive nuclear bodies including NEAT1 (paraspeckle), NEAT2 (speckle) and Gomafu/ MIAT (Gomafu-containing nuclear bodies). Previous study showed that NEAT1 is required for paraspeckle formation and the nuclear retention for mRNAs containing Alu repeats [69, 70]. MALAT1 has been reported to affect alternative splicing through its interaction with splicing factors that are accumulated in nuclear speckles [71]. Moreover, Gomafu is also called MIAT (myocardial infarction-association transcript) due to its association with myocardial infarction in a large scale SNP analysis [72]. Nevertheless, the function of Gomafu-containing nuclear domain is unclear.

LincRNAs and human diseases

In the past decade, it has been clear that non-coding RNAs have regulatory roles in many human diseases. Nevertheless, the majority of studies focused on the regulation of microRNAs (miRNAs) and its downstream mRNA targets [73]. The discovery of lincRNAs raise a question: whether these transcripts have essential functions in parthenogenesis. Indeed, recent research indicated that overexpression of HOTAIR is correlated with poor survival of breast cancer and other cancer [74, 75]. HOTAIR was found to alter the PRC2 occupancy and therefore promote cancer metastasis and progression [58]. LincRNA-p21 was identified as a p53 downstream effector associated with hnRNP-K to repress p53-medicated gene expression, indicating that lincRNAs play essential roles for cancer [57] . Global transcriptome of human cancer also revealed that many lincRNAs are aberrantly expressed in lung, breast and brain carcinoma [76]. In conclusion, currrent studies unraveled the physiological roles of lincRNAs and deregulation of lincRNAs may result in many other disease phenotypes.

Specific aims

Although Daxx-mediated transcription has been reported, how Daxx regulates transcriptional network is yet unclear. Here, we used microarray technology to identify Daxx-regulated transcripts, including protein-coding genes and lincRNAs. This is the first time genome-wide analysis of Daxx-mediated transcription. Since Daxx has been known to interact with more than 40 proteins including numerous transcription factors, our results further provides a roadmap to elucidate Daxx-mediated transcriptional profile and biological consequences. Moreover, our previous studies indicated that Daxx can regulate HOTAIR stability and therefore affect global hisotne modification. We hypothesize Daxx can regulate other functional lincRNAs. Hence, we will combine microarray data and bioinformatic analysis to identify Daxx-regulated lincRNAs, potentially acting as important Daxx downstream effector(s) in various biological processes.

Chapter II

Materials and Methods

Materials and methods

Cell culture and transfection

HEK293T, HeLa cells were obtained from the American Type Culture Collection (ATCC). Cells were cultured in DMEM (high glucose, Gibco) supplemented with 10 % fetal bovine serum (Gibco), 10 mM NaHCO3 (Sigma) and penicillin/streptomycin (Gibco). Transfections were performed by using Lipofectamine 2000 (Invitrogen) according to the manufacturer's instruction.

Plasmids and RNA interference

The mammalian constructs expressing HA-tagged Daxx, shDaxx#1, shDaxx#2, were described previously [36, 77]. $Daxx_{RFL}$ plasmids contain four sense mutations in the middle of the shDaxx#2 binding site. Specific siRNA oligonucleotides against lincRNAs were synthesized by Life Technologies^{TM} .

RT-quantitative PCR (RT-qPCR)

Total RNA from cells was extracted using TRIzol (Life Technologies) and NucleoSpin RNA II (Macherey Nagel). For RNA decay rates, cells were incubated for the appropriate time after the addition of acinomycin D (Sigma) in final concentration

5ug/ml. Five microgram RNA of each sample was then reverse transcribed using ThermoScript RT-PCR system (Life Technologies). Quantitative Real-time PCR was performed using SYBR Green PCR master mix (Life Technologies) and ABI 7500 sequence detection system (Life Technologies). Relative expression values were calculated (ΔΔCT method) was normalized with GAPDH level. Primers information is listed as follows.

RT-qPCR Primers for protein-coding genes

Gene	5’ Forward	5’ Reverse
Daxx	TGCAGACACCCCCGAAGCCT	TGCCATTCCACTAGGGCCCTCA
ERLIN2	ACCAGGGCGCTTGACCCTCT	GAGCCTTTATCCGTGGCGGC
WHSC1L1	GGATTACAGAAAGGGTCGGCGG	TGCGGAGACGGAGCTGTCACT
ZNF703	CCATTGAGCTGGACGCCAAGAA	TTGAGTTTGGAGGAGGGCGGC
UNC5D	GCTGCCGAGGTGGAATGGCTG	GAGAGCCGTGCCTGCCTGATG
PROSC	TTCCAGCATGCGGTTGAAGTAGGA	TTCACGTCTGCTGCGCACTT
DUSP26	ACGACTCGCCAGCCTTTGACA	ATCTTCCCTCCTGGCTGGCTCA
GAPDH	TCTTTTGCGTCGCCAGCCGAG	TGACCAGGCGCCCAATACGAC

RT-qPCR Primers for lincRNAs (working primers)

Name	Accession Number	Chromosome position (hg19)	*	Forward	Reverse
linc-7720	TCONS_00017720	chr10:125115957-125277984	+	AGAGAGGGCACCTGGAGAATGGTT	TTGCTACGGTCCTTGGGCTCTTCT
linc-3393	TCONS_l2_00023393	chr5:53691181-53710899	+	AGCACATGACCTCAAGTCTTTGGA	CCAGCCTTCCATCTTTCCGGC
linc-4971	TCONS_00014971	chr8:37184475-37189507	-	GAGCGGCTGGAGTTCACGGC	GGAGGGGTGGTCTCGCACAC
linc-5145	TCONS_00015145	chr8:118652591-118695917	-	TGGAGTGTACAGCTGAAGGATCAGG	AGTCACCTTTGTGCTTATGGCAA
linc-5292	TCONS_00015292	chr8:67383684-6738720	+	TGGATACCCATCATTCTGGCCCTT	TACAGGCTTAGGGGGCCACAGA
linc-3782	TCONS_00013782	chr7:39659893-39662092	-	GGATTGAGGATGGTGGGTTGCACA	GCTGTGAGCCACCAAAAAGCAGC
linc-2761	TCONS_00002761	chr2:105421977-105467916	+	ATGCTCCATGTCCCCTCTGCCA	AGAGGCAGGGTCCTGACAACTCT
Pri-mir34	NR_029610	chr1:9211727-9211836	+	CCTCCAAGCCAGCTCAGTTG	TGACTTTGGTCCAATTCCTGTTG
JPX	NR_024582	chrX:73164159-73290217	+	GCCTGGGAGTCCACCACCAC	TCAGGCTGTCTTCCGCCCTCA
NEAT1	NR_028272.1	chr11:65190269-65194003	+	TGCCATCTCACAGGCAGGGGAA	GCGTATGCAAGTCTGACGCCCA
MIAT	NR_033320	chr22:27053446-27072440	-	GGCAGCATTTCATGGCCTCCGT	AGCACGGGAAAGGGCCCAAA
HOTAIR	NR_003716.3	chr12:54356092-54362540	-	GTCCCTAATATCCCGGAGGT	GCAGGCTTCTAAATCCGTTC

*strand

RT-qPCR primers for Non-specific or low expression lincRNAs

Name	Accession Number	Chromosome position (hg19)		Forward	Reverse
linc-0646	TCONS_l2_00020646	chr4:90602647-90641106	+	GGCTGCTACCTCCCCATGAGTC	TGCCTTCTGGACACTGTACGACTA
linc-0645	TCONS_l2_00020645	chr4:90602647-90615273	+	AGCCCTCAGAGCTCCTCTGCT	ACCAGTGGCTGGCACTCATGC
linc-6893	TCONS_l2_00026893	chr7:130565099-130582733	-	CCAGTTGCACCCATAGTCCGGC	CACTTTCCTTGGGGCCTGCGG
linc-4143	TCONS_00024143	chr15:67276460-67280258	+	AAGCCAGCTCCCTCCGAGAA	TGCCTGAGGAGGACCAATGC
linc-8290	TCONS_l2_00028290	chr8:90598098-90618591	-	TACACTGGGGTAGTGTGGGCCT	TCCCAGATGCAGCTCACGCA
linc-0472	LINC00472	chr6:72124149-72130448	-	GCACCTCGGGCTGCGGATTT	CCCTGGCCACATGGCCCAAC
linc-6817	AK056817.1	chr5:42985503-42993435		GGACGCCCAGAGCAAGCAGG	AGCGCCACTCCTCCTTCGGA
linc-2946	TCONS_00012946	chr7:7294785-7317410	-	TGTCACCGGGAGCAGCAGTT	TGGTCTTGCTTTCTGTCTCATTCAC
linc-5474	LOC100505474	chr18:53750587-53804767		TGGTGCTCAGCCCTTGCACT	TGAGGTGGTGGTTGCATGCGT
linc-1655	LOC100131655	chr18:74506688-74534251		ACCCACAACAGCTCCGACTCCA	TGCCTCAGTCCTTGGTGCAGA
linc-1471	LOC541471	chr2:112124591-112252692	-	CAGCAACATGGCAGGCCAGA	AATGTGCTGTGGAGCAGGCACG
LOC100506305	NR_038904	chr11:13001081-13011095		TGGTCCTCCCTGCACTTCACAGT	GCTTGCATTCACCTCTGCCGGT

Microarray analysis

RNA extraction

Total RNA was extracted by the commercial kit (NucleoSpin, Macherey-Nagel®). The concentration of RNA was quantified by Nano drop 2000 spectrophotometer (Thermo) and the quality was monitored by the Agilent 2100 Bioanalyzer (Agilent).

Knockdown validation using RT-qPCR

To determine Daxx and lincRNA knock down efficiency, we used real-time qPCR to check the expression level of interested genes. The qPCR was performed using self-designed primers and 2 X SYBR green mix (invitrogen) and run on ABI 7500. Expression of interested genes was normalized with Gapdh level. The fold decrease of Daxx or HOTAIR expression compared to control represented knockdown efficiency. To avoid experimental variations, 3 pairs of independent knockdown experiments were selected based on comparable knockdown efficiency.

Label and Agilent microarray hybridization

The SurePrint G3 Human GE 8x60k Microarray (G4858A-028004) used for this study has probes to 27,958 Entrez Gene RNAs and 7,419 lincRNAs. Using invitrogen

SuperScript plus Indirect cDNA labeling System (L1014-04), we labeled total RNA for hybridization according to manufacturer's instructions. Alexa fluor® 555 labeled cDNA was prepared from 15ug total RNA and purified by purification module in the kit. Dye incorporation and cDNA yield were checked by gel image. Total Alexa fluor® 555 labeled cDNA was hybridized following the protocol as described in Gene expression Hybridization Kit (Agilent, cat: 5188-5342) and Gene Expression Wash Buffer Kit (Agilent, cat: 51885327)

Scan protocol and data processing

Slides were scanned immediately after washing on the Agilent DNA Microarray Scanner (G2565) using one color scan setting for 8x60k array slides (Scan Area 61x21.6 mm, Scan resolution 3μm, Dye channel is set to Green and Green PMT is set to 100%). The scanned images were analyzed with Feature Extraction Software 10.7.3.1 (Agilent) using default parameters (protocol GE1_105_Dec08 and Grid: 028004_D_F_20100430) to obtain background subtracted and spatially detrended Processed Signal intensities.

Data analysis

The signal intensity obtained from microarray was analyzed using GeneSpring GX 11.5.1(Agilent Technologies Inc., Santa Clara, CA). The signals underwent background

subtraction and normalized to 75 percentile using median of control samples as the baseline transformation. The genes and lincRNAs with statistical significance were selected based on p-value (p-value<0.05). In addition, hierarchical clustering was performed to create heat map. The distance metric and linkage method are Pearson centered and centroid, respectively. The gene ontology functional analysis was generated through the use of Ingenuity Pathways Analysis (Ingenuity System®, www.ingenuity.com).

Accession number

The microarray data is deposited at the Gene Expression Omnibus (GEO) under accession number GSE35792.

Chapter III

Results

Results

Transcriptional profile of knockdown of Daxx

In order to identify target genes and the transcription network controlled by Daxx, we knocked down Daxx using short hairpin RNA (shRNA) and performed microarray analysis. Three biological pairs of experiments were performed (Figure 1A). After statistical filtering, 655 genes were differentially expressed, of which 236 genes were up-regulated and 419 genes were down-regulated (p-value<0.05, fold change> 1.5) (Figure 1B and Table 1). Gene ontology (GO) analysis indicated most significant biological functions of Daxx-mediated genes are cell morphology and cell death (figure 2A). In addition, GO analysis also indicated that Daxx target genes might be involved in several physiological processes such as inflammatory response, developmental disorder and cancer (figure2B). Pathway analysis showed that Daxx might regulate key genes in several unexpected pathways such as B cell signal pathway and paxillin pathway. Furthermore, according to gene expression pattern, we predicted what transcription factor Daxx could regulate and how it could regulate this transcription factors activity. The result indicates Daxx may repress the transcription activity of various essential proteins such as NF-κB, p53 and SMAD4 [37]. Some of these lincRNAs have been

reported before, supporting the reliability of this prediction, while others have not been identified previously such as ELK-1, RARG and SMAD3, which provides novel research candidates (Table 2). In sum, revealing Daxx-mediated transcription landscape in a genome-wide scale not only facilitates us to identify Daxx target genes but also elucidates the role of Daxx in various biological functions.

The identification of Daxx-regulated lincRNAs

Many lincRNAs have been identified as functional regulators in cells but how these non-coding transcripts are regulated is still largely unknown. Since Daxx functions as a transcription coregulator, it is reasonable to speculate some lincRNAs can be modulated by Daxx and participate in Daxx-mediated functions. The Agilent microarray chip we used has probes to 27,958 Entrez Gene RNAs and 7,419 lincRNAs so we were able to determine expression of lincRNAs in the same analysis. Our result showed that 104 lincRNAs were significantly affected in Daxx knockdown cells compared to control (p-value<0.05, FC>1.5). Because the design of probes for lincRNAs in the Agilent microarray chip was based on outdated data, we verified those signals with latest annotation information (Figure 3). First, we compared the defined regions of lincRNA probes in the microarray with latest announced stringent 4662

human lincRNAs [52]. Only 93 lincRNAs were overlapped. Second, we further filtered out if these lincRNAs are not in the UCSC online data base (http://genome.ucsc.edu/cgi-bin/hgGateway) and 45 lincRNAs were left (Figure 4). Then, we designed primers based on predicted lincRNAs sequence on UCSC website and used quantitative real-time PCR to measure these transcripts. We initially focused on up-regulated lincRNAs in microarray data. Reconstitution experiments were performed to avoid siRNA off-target effect. Excluding lincRNAs with low expression and non-specific regulation, 12 lincRNAs were validated (Figure 5). Among these, some are well-characterized non-coding transcripts such as JPX, NEAT1 and pri-mir34. Combined with microarray data and bioinformatic analysis, we showed that Daxx is capable of regulating lincRNA expression.

Evolutionary conservation of Daxx-mediated lincRNAs

To determine whether Daxx-regulated lincRNAs are evolutionarily conserved, we compared human loci of these lincRNAs with the mutual loci among different species. The conservation analysis revealed that only a few lincRNAs are evolutionarily constrained. Notably, lincRNA-4971 displays a higher degree of conservation than other lincRNAs (Figure 6). Taken together, although Daxx is a conserved protein, downstream lincRNAs may vary among different species. Identification of secondary

structure and associated proteins of these lincRNAs may facilitate our search for functional orthologs.

Chromatin signature of Daxx-mediated lincRNAs

By identification of the chromatin structure with H3K4me4 and H3K36me3, called K4-K36 domain that reside outside known protein-coding genes, Guttman *et al* discovered over a thousand lincRNAs [51]. Using the public database of histone modifications, we searched for the chromatin signature of loci containing Daxx mediated-lincRNAs and K4-K36 distribution was found suggesting Daxx-mediated lincRNAs share common feature with other lincRNAs (Figure 7).

Tissue specificity of Daxx-mediated lincRNAs

As the majority of lincRNAs shows tissue specific expression pattern, functional lincRNAs may have distinctive roles depending on the cellular context. Using UCSC online resource with expression information across 22 human tissues and cell lines, we demonstrated lincRNA expression profile in different tissues (Figure 8). JPX and NEAT1 are highly abundant in all cell types because of their known essential function in X inactivation and paraspeckle formation, respectively. Daxx is highly expressed in testes and thyroid organs. Nevertheless, the result indicated that specific expression of

these lincRNAs is not completely correlated with Daxx expression pattern.

Daxx-regulated lincRNAs affect gene expression in cis

A few reports indicated that lincRNAs can act *in cis* [67]; in an enhancer-like function or repress neighboring genes via recruiting chromatin modifiers [78]. To examine whether Daxx-regulated lincRNAs have similar mechanisms in gene expression regulation, we checked the expression profile of genes around lincRNAs within a window of 10 genes on either side from microarray data. The result demonstrated that up-regulation of linc-4971 and lin-2761 may result in down-regulation of nearby genes while linc-5292 and linc-3782 may have enhancer-like function (Figure 9). The study suggested the notion that lincRNAs in *cis* function is independent of orientation [67]. In order to validate this phenomenon, we used siRNA technique to knockdown linc-4971. We examined the expression of neighboring genes and found that knockdown of linc-4971 led to a derepression of nearby genes such as ERLIN2 and ZNF703. The result suggests that expression of linc-4971 may repress nearby genes *in cis*.

Chapter IV

Discussion

Discussion

Daxx-mediated transcription network

Although Daxx has been shown to act as a transcriptional coregulator, a genome-wide screening of gene expression profile regulated by Daxx has not been demonstrated. Here, we used microarray analysis to identify Daxx target genes systematically and constructed a Daxx-regulated transcription network. GO analysis showed that gene response by Daxx knockdown significantly correlated with cell death and inflammatory response, in accordance with previous studies [33, 38]. Recently, mutations in Daxx were reported in PanNETs and other tumors [43, 79]. GO analysis also showed a significant enrichment in genes involved in cell morphology and cancer, suggesting the role of Daxx in tumor formation (Figure 2A). Moreover, putative Daxx-regulated transcription factors were listed based on gene expression profile. Many of them were reported in previous studies [38], while some are the first time to appear on the list of Daxx-related proteins such as ELK1 and RARG. Further study is required to validate this prediction. Taken together, microarray analysis not only provides an informative database which will help us to identify genes controlled by Daxx but also reveals the most significant biological function of Daxx in a large-scale analysis.

Identification of Daxx-regulated lincRNAs

Although the whole human genome has been sequenced, to determine the transcription of a specific lincRNA remains a challenge. One reason is that transcriptome is more complex than we thought. For example, lincRNAs have several features such as alternative splicing, antisense transcription, tissue specificity and little evidence of evolutionary conservation. More works will be needed to understand the biology of lincRNAs, such as precise annotation of lincRNA's exon and intron, determination of secondary structure, identification of lincRNA-interacting proteins and elucidation of how these lincRNAs be regulated in various cellular contexts. In addition, although phenotypic consequences of depletion of numerous lincRNAs were presented, the molecular mechanisms are still elusive. Biochemical and genetic studies will be required to elucidate the role of lincRNAs in biological processes.

Here, we identified lincRNAs with differential expression upon Daxx depletion. By comparing selected lincRNAs in the Agilent microarray chip with current annotation resources, we revealed that not all identified lincRNAs have been annotated as defined lincRNAs. This may result from the probe design on microarray may be based on the chromatin state and other evidence rather than the latest RNA-seq data or transcript assembly. In comparison with current integrative database, we effectively identified Daxx-regulated lincRNAs. However, microarray still has several limitations in detection of lincRNAs. For example, probes may not cover whole transcription unit so the result

might be false negative. Moreover, microarray cannot differentiate strand-specific transcripts efficiently. Recent developed high throughput technology (RNA-seq) may be required to circumvent such shortcomings. .

Of note, many identified lincRNAs show low expression level so we failed to detect their expression using RT-qPCR. One report indicated the maximal expression levels of lincRNAs are lower than protein-coding genes, with a ~10-fold lower median maximal expression level [52]. Even though lincRNAs with basal expression level may have function, choosing abundant transcripts as the research target seems more biologically logical and feasible.

Twelve Daxx-regulated lincRNAs were validated in our study. Some are well-known non-coding RNAs. Pri-miR34 is a precursor transcript of miR34 which is regulate by p53. The increase of pri-miR34 may result from p53 stabilization while Daxx was depleted [39]. Ectopic miR-34 expression induces apoptosis, cell-cycle arrest or senescence which are in line with the phenotype of cells with Daxx depletion [80]. Interestingly, two known nuclear long non-coding RNAs, NEAT1 and MIAT, are regulated by Daxx. NEAT1 is required for biogenesis of nuclear paraspeckle while MIAT is resided in a novel nuclear body with unclear function [69, 81]. Daxx can form numerous foci in the nucleus but the current evidence only supports its colocalization with PML nuclear body. It is intriguing to know whether Daxx can colocalize with

NEAT1 and MIAT and regulate the levels of these nuclear non-coding RNAs via interaction or via a transcriptional control.

Evolutionary conservation of Daxx-regulated lincRNAs

The evolutionary conservation of lincRNAs in different species is still in a debate. With little evolutionary evidence, it has been suggested that lincRNAs may just be the outcome of transcriptional noise [48]. However, a group of lincRNA discovered by the chromatin-state map has highly evolutionary conservation in mammals [51]. Among Daxx-regulated lincRNAs, linc-4971 locus shows more conservative regions than linc-3782 locus. The latter case may be due to expansion of retrotransposon or evolution of non-coding transcripts from a coding one.

Chromatin status of genomic loci containing Daxx-regulated lincRNAs

Chromatin signature with H3K4me3 and H3K36me3 is the hallmark of active transcription [82]. The chromatin states of loci containing Daxx-mediated lincRNAs also show H3K4me3-H3K36me3 distribution (Figure 7). Notably, there are no protein-coding genes resided in all tested loci, indicating that identified lincRNAs are *bona fide* non-coding transcription units.

Tissue specificity of Daxx-regulated lincRNAs

LincRNAs was shown to have much more tissue specificity than protein-coding

genes [52]. Compared with expression pattern derived from RNA-seq data, Daxx-regulated lincRNAs show little correlation with Daxx expression level. This may be due to Daxx is expressed abundantly among tissues as an essential protein controlling cell survival [1]. Of note, Daxx is highly expressed in testes and thyroid. According to a recent research, approximate one third of lincRNAs are specific to testes suggesting a biological function of Daxx in regulation of these testes-specific lincRNAs [52]. Conditional Daxx knockout mice in testes may reveal the biological relevancy of these lincRNAs regulated by Daxx.

Daxx-regulated lincRNAs affect gene expression *in cis*

Numerous lincRNAs can affect regional genes *in cis* [60, 67, 78]. Therefore, we examined whether Daxx-regulated lincRNAs have similar function. Using microarray data, we found upregulation of linc-4971 and linc-2761 may result in local repression of nearby genes while upregulation of linc-5292 and linc-3782 may function as enhancers, supporting by a correlation of nearby gene activation. Of note, linc-4971 resided in chromosome 8p11-12 whose amplification has been found in ~15% of human breast cancers and was associated with poor prognosis [83, 84]. Moreover, 23 genes from the 8p11-12 region were correlated with cancer progression [85]. Several genes have been validated for their oncogenic function including WHSC1L1, DDHD2, ERLIN2 and

ZNF703 [85, 86]. We found that knockdown of linc-4971 can activate its upstream genes ERLIN2 and ZNF703 while PROSC is not affected (Figure 10). As the result, we identified a lincRNA resided in 8p11-12 which can modulate regional gene expression, suggesting lincRNAs may act as regulatory nodes in Daxx-mediated transcriptional pathways. In addition, in the microarray data, two lincRNAs showed enhancer-like function as described [67]. However, we cannot exclude the possibility that this effect was resulted from shared upstream regulation or neighboring transcriptional effect.

Chapter V

Figures

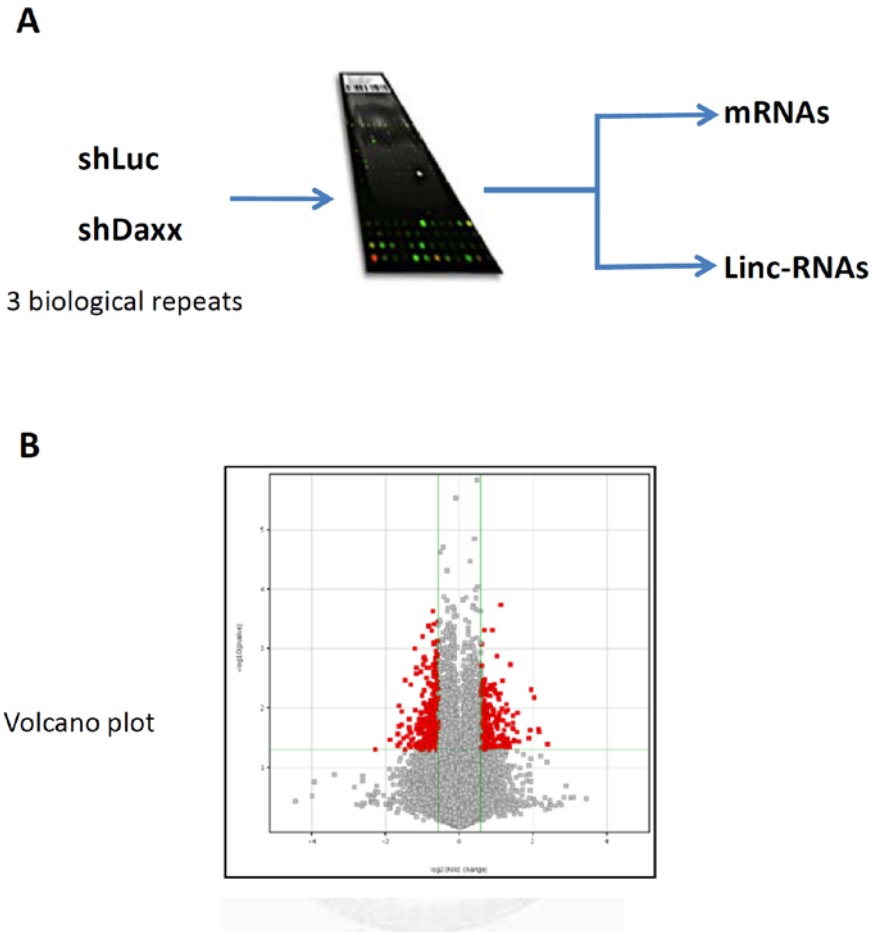

Figure 1. Microarray analysis of Daxx-mediated transcription. (A) The experimental design of microarray analysis. 293T cells were transfected by indicated constructs with shot hairpin RNA targeting to Daxx or luciferase, respectively. RNA was extracted and subjected to microarray analysis. Tree biological batches were performed. **(B)** Differentially expressed genes with statistical significance were shown by dark dots in the volcano plot (student t-test p-value<0.05, fold change>1.5).

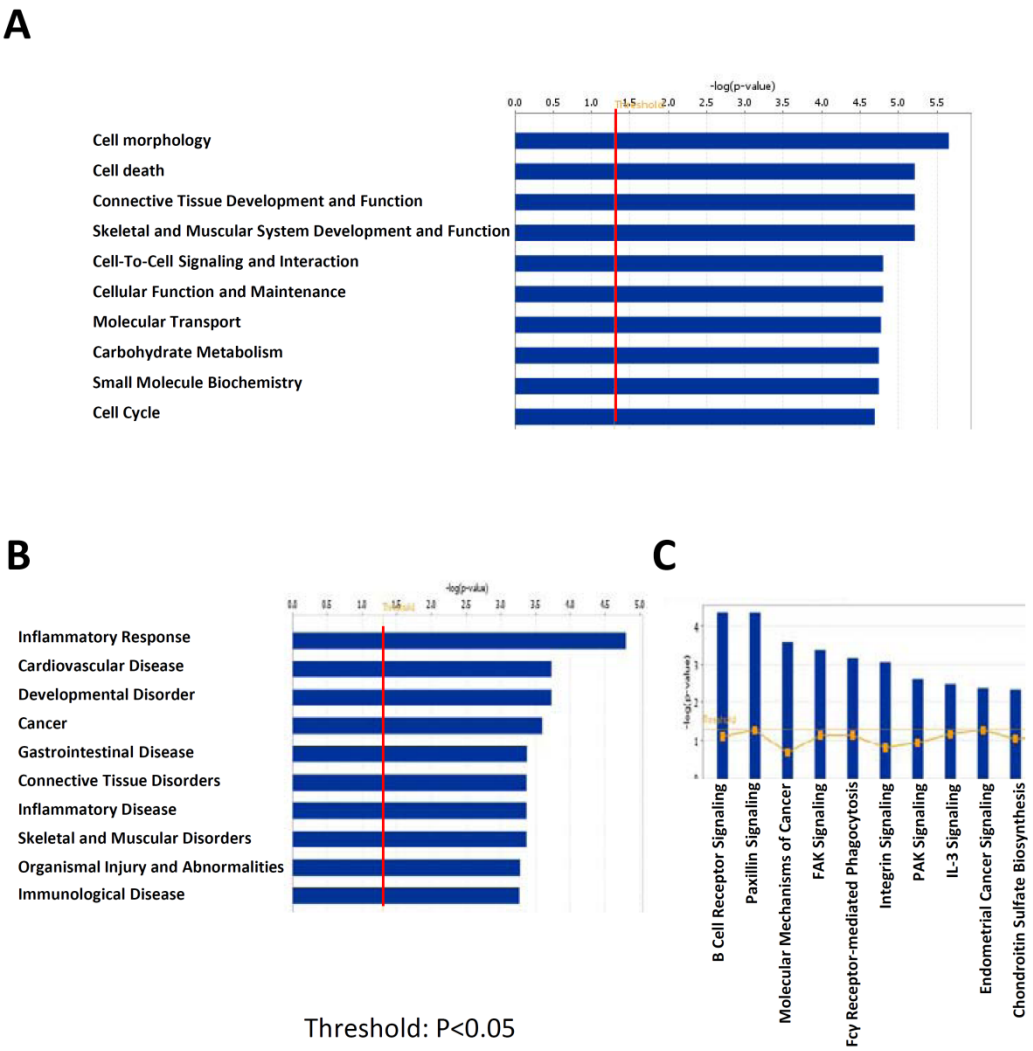

Figure 2. GO analysis of Daxx-regulated genes.

(A) Top 10 molecular and cellular functions of Daxx-regulated genes. **(B)** Top significant physiological functions of Daxx-regulated genes. **(C)** Most significant canonical pathways of genes regulated by Daxx. Analyses were performed with Ingenuity Pathways Analysis (IPA). Red lines indicate p-value= 0.05. Fisher exact t test.

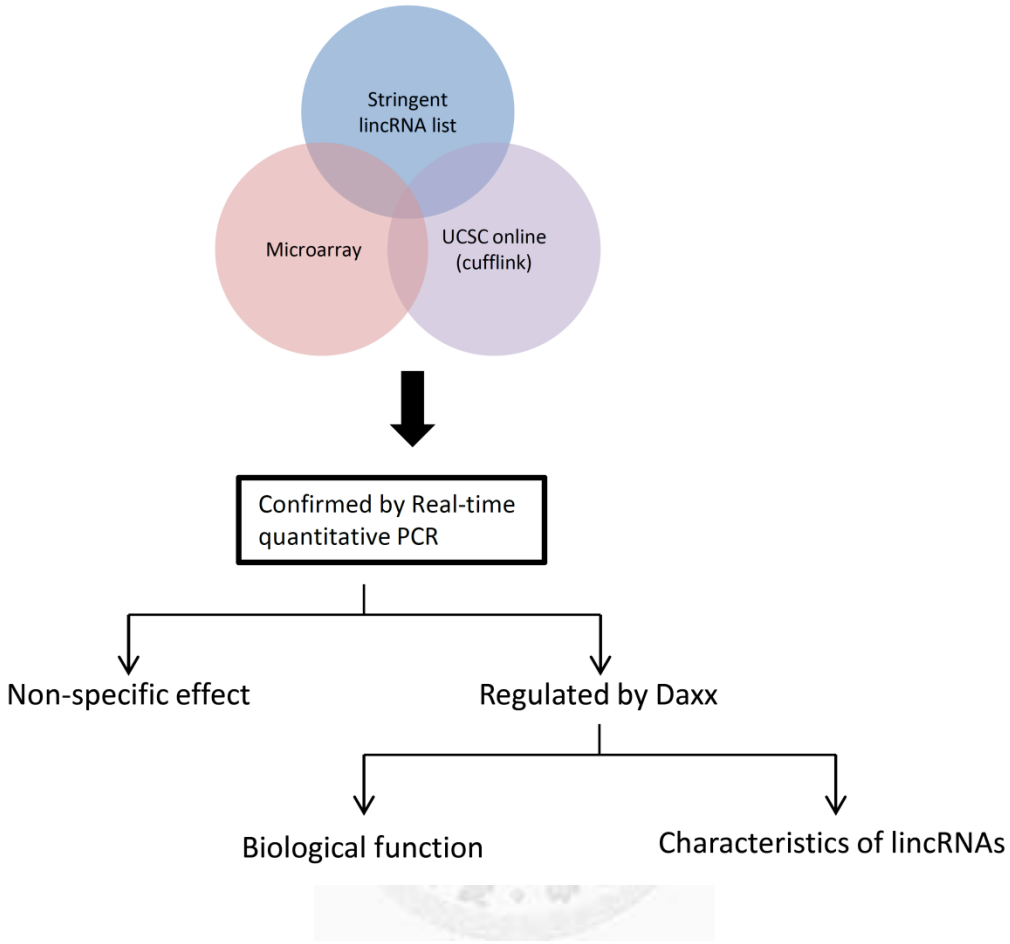

Figure 3. Strategy for the identification of Daxx-regulated lincRNAs.

RNA samples were derived from knockdown Daxx 293T cells and subjected to microarray analysis. By comparison with current annotation and online database, Daxx-mediated lincRNAs were identified as the research candidate. Reconstitution experiments were performed to eliminate the outcome resulted from the off-target effect. The function of Daxx-mediated lincRNAs was further tested by siRNA knockdown experiments.

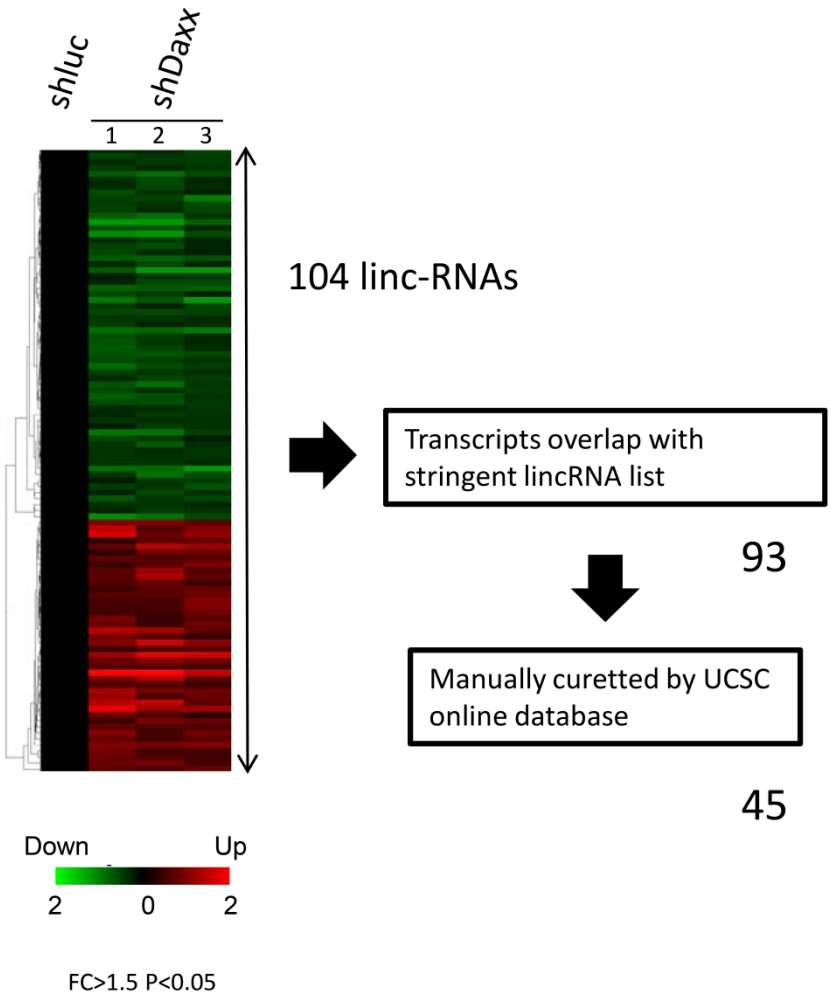

Figure 4. **Scheme of identification of differentially expressed lincRNAs.**

The expression profile of 104 lincRNAs in tree biological repeats of experiments were shown by hierarchical clustering heat map. Probes in the Agilent microarray chip for lincRNAs do not overlap with the 4462 lincRNAs were filtered out. Subsequently, lincRNAs were manually excluded by comparing with UCSC online database. 45 lincRNAs were identified in this study.

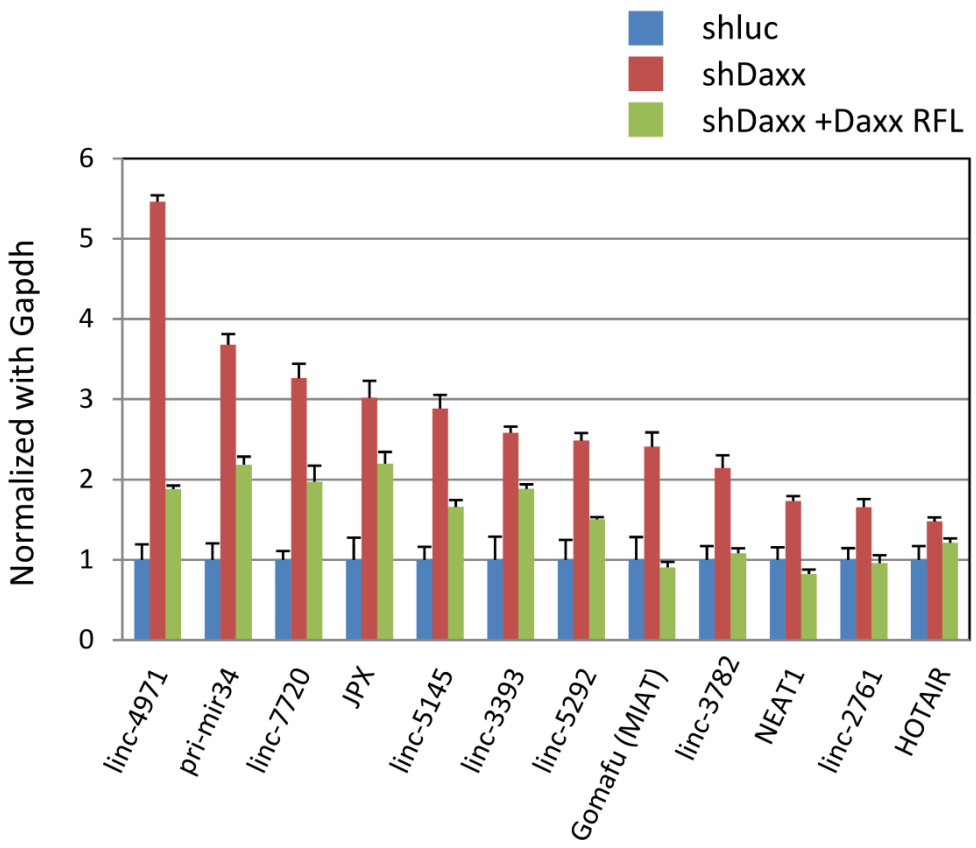

Figure 5. Validation of differential expressed lincRNAs upon depletion of Daxx.

Daxx knockdown in 293T cells resulted in up-regulation of 12 lincRNAs, while reconstitution of Daxx leads to rescue such phenotype. Relative expression level were determined by RT-qPCR and normalized with GAPDH. Error bars represent S.D. from three technical replicates.

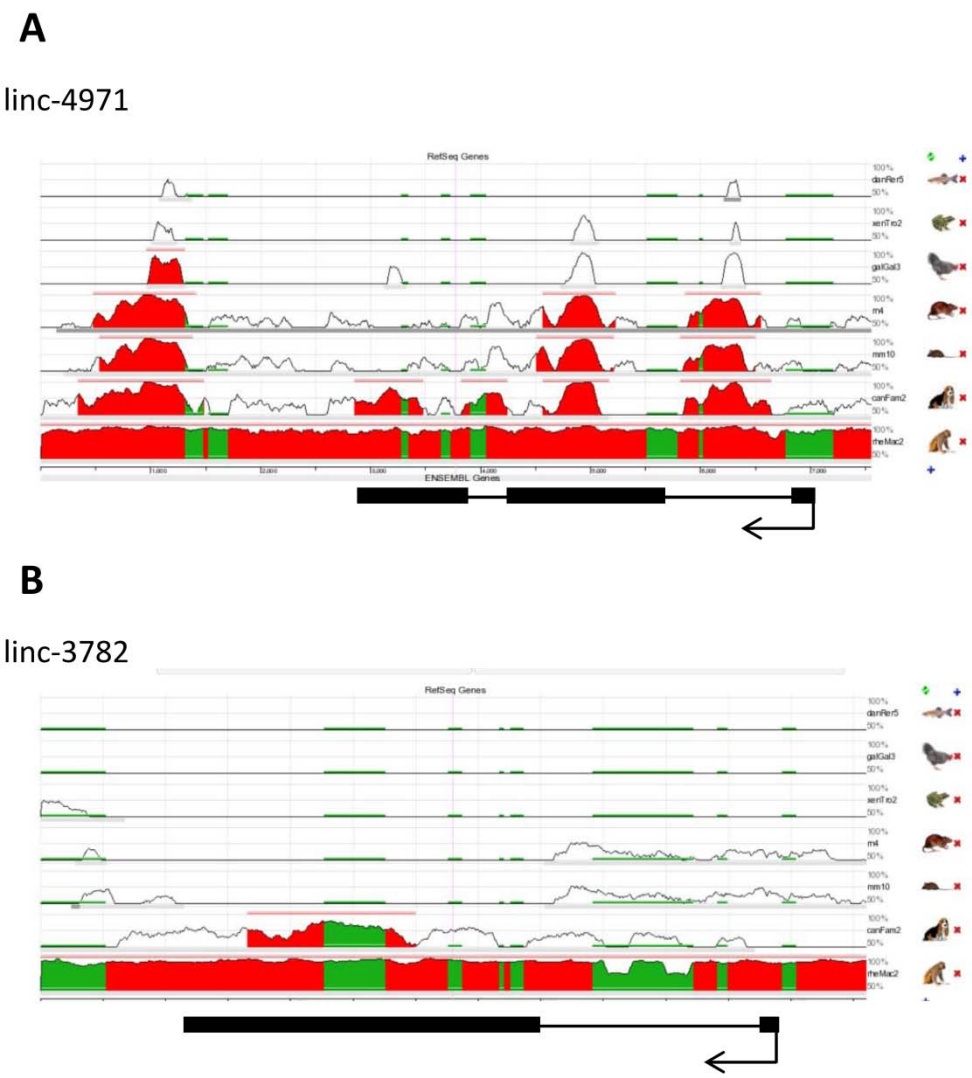

Figure 6. Evolutionary conservation of Daxx-mediated lincRNAs.

(A-B) Evolutionary conservation (as % of sequence identity, Y-axis) of human loci containing lincRNAs is shown using the Evolutionary Conserved Regions (ECR, http://ecrbrowser.dcode.org/) browser for a number of vertebrates: chicken (galGal), frog (xenTro), zebrafish (danRer), rat (ratNor), mouse (musMus), dog (canFam), and macaque (resMac). Red, and green colors correspond to intergenic, intronic and

repetitive genomic regions, respectively. Two loci were presented here. Linc-4971 locus shows more conservative regions than linc-3782 locus.

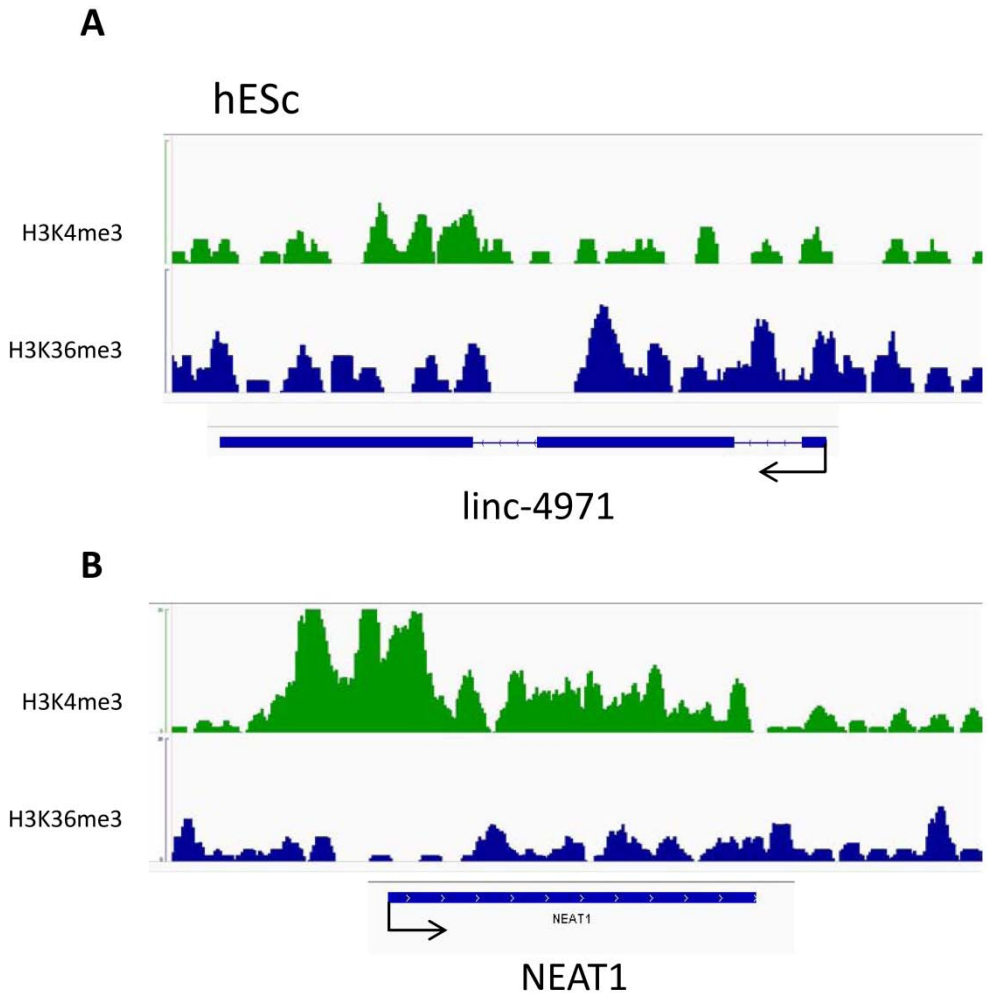

Figure 7. Intergenic K4-K36 domain in identified lincRNAs loci.

(A-B) Trimethylation of lysine 4 of histone H3 (H3K4me3) and trimethylation of lysine 36 of histone H3 (H3K36me3) distribute on the genomic loci containing Daxx-mediated lincRNAs. Each histone modification is plotted as the number of DNA fragment obtained by ChIP-seq data from ENCODE database. Blue boxes indicate non-coding transcripts.

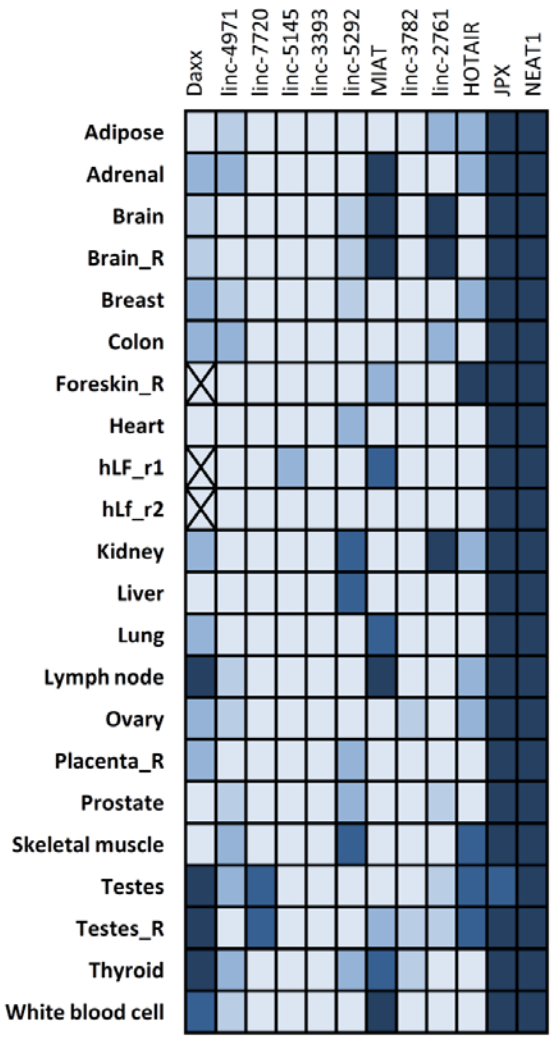

Figure 8. Tissue specificity of Daxx-regulated lincRNAs.

The expression patterns of Daxx-regulated lincRNAs in various tissues were derived from UCSC online resource with expression information across 22 human tissues and cell lines. Notably, many lincRNAs show basal expression in different tissues. Expression abundance scores range is displayed from light blue to dark blue respectively.

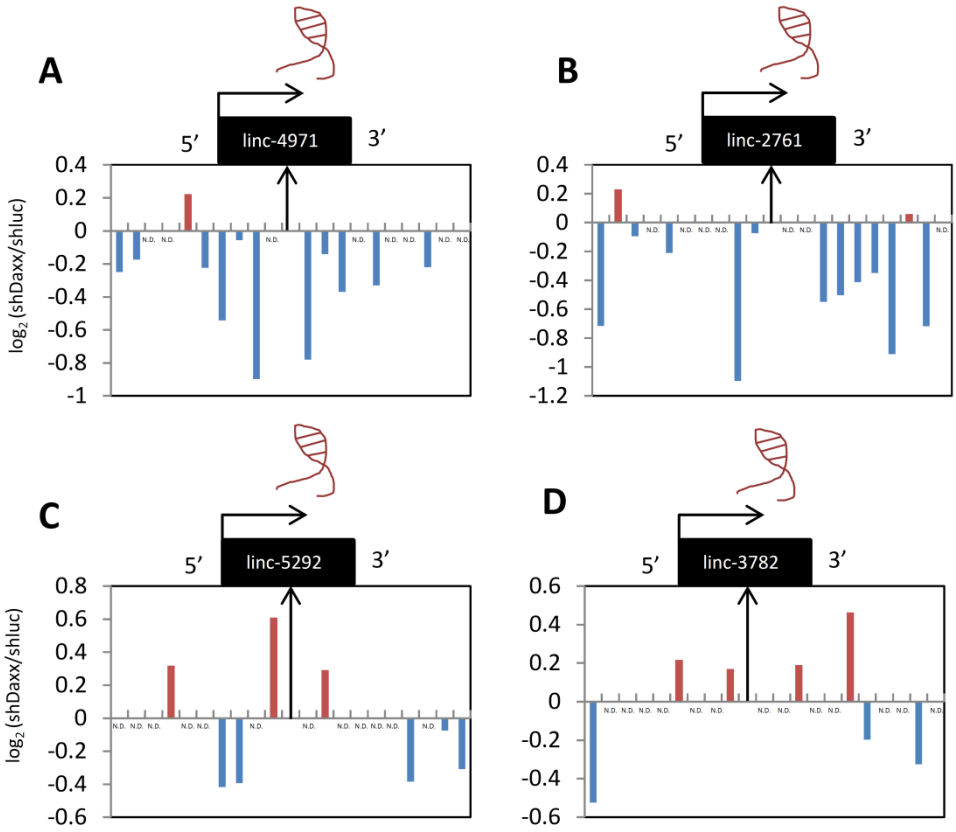

Figure 9. Daxx-mediated lincRNAs act *in cis*.

The expression profile of genes around identified lincRNAs within a window of 10 genes on either side. (blue-downregulation, red-upregulation). Data were subtracted from microarray analysis. (A-B) Up-regulation of linc-4971 and lin-2761 correlates with down-regulation of nearby genes and (C-D) linc-5292 and linc-3782 may have enhancer-like function. N.D. not determined.

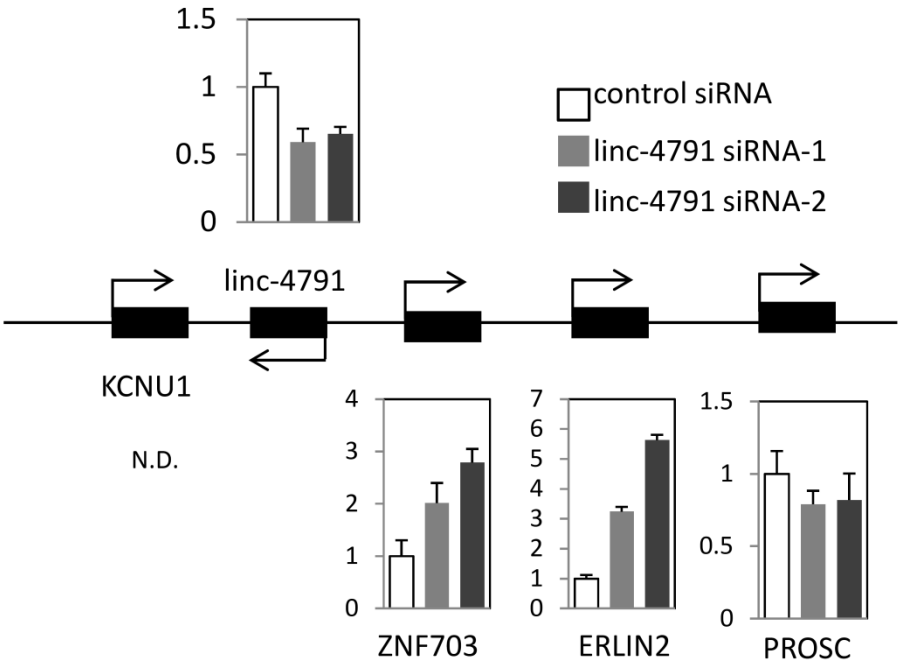

Figure 10. Knockdown of linc-4971 results in activation of upstream genes.

Two specific siRNAs targeting to linc-4971 were treated on 293T cells. Black boxes indicate transcription units. Transcription direction was shown by the arrow. Relative expression level were determined by RT-qPCR and normalized with GAPDH. Error bars represent S.D. from three biological replicates.

Chapter VI

Tables

Table 1. Numbers of differentially expressed genes upon depletion of Daxx

	P all	P<0.05	P<0.02	P<0.01	P<0.005	P<0.001
FC all	42545	3552	1566	757	396	89
FC>1.1	25495	3330	1485	714	374	83
FC>1.5	3199	655	299	150	80	15
FC>2	663	119	49	20	12	3
FC>3	122	15	4	3	1	0

FC: fold change

Table 2 Putative Daxx-regulated transcription factors

Transcription regulator	predication	Regulation of z score	P value of overlap
NFKB1*	Repress	2.826	4.39E-04
NFkB (complex)*	Repress	2.669	5.88E-02
ELK1	Repress	2.406	3.57E-03
RARG	Repress	2.317	4.19E-05
TP53*	Repress	2.272	3.08E-04
SMAD3	Repress	2.230	7.22E-03
CREBBP*	Repress	2.227	1.95E-05
RELA*	Repress	2.074	1.18E-02
CTNNB1*	Repress	2.017	7.97E-03
MTPN		1.952	2.34E-02
TP63*		1.799	3.46E-02
IRF8		1.751	9.49E-02
SRF		1.635	2.07E-03
NOTCH1		1.607	9.18E-03
SMAD4*		1.513	5.97E-04

*** The interaction or regulation has been reported**

References

1. Salomoni, P. and A.F. Khelifi, *Daxx: death or survival protein?* Trends Cell Biol, 2006. **16**(2): p. 97-104.
2. Yang, X., et al., *Daxx, a novel Fas-binding protein that activates JNK and apoptosis.* Cell, 1997. **89**(7): p. 1067-76.
3. Wu, S., H.N. Loke, and A. Rehemtulla, *Ultraviolet radiation-induced apoptosis is mediated by Daxx.* Neoplasia, 2002. **4**(6): p. 486-92.
4. Perlman, R., et al., *TGF-beta-induced apoptosis is mediated by the adapter protein Daxx that facilitates JNK activation.* Nat Cell Biol, 2001. **3**(8): p. 708-14.
5. Michaelson, J.S., et al., *Loss of Daxx, a promiscuously interacting protein, results in extensive apoptosis in early mouse development.* Genes Dev, 1999. **13**(15): p. 1918-23.
6. Michaelson, J.S. and P. Leder, *RNAi reveals anti-apoptotic and transcriptionally repressive activities of DAXX.* J Cell Sci, 2003. **116**(Pt 2): p. 345-52.
7. Zhong, S., et al., *Promyelocytic leukemia protein (PML) and Daxx participate in a novel nuclear pathway for apoptosis.* J Exp Med, 2000. **191**(4): p. 631-40.
8. Ishov, A.M., et al., *PML is critical for ND10 formation and recruits the PML-interacting protein daxx to this nuclear structure when modified by SUMO-1.* J Cell Biol, 1999. **147**(2): p. 221-34.
9. Lin, D.Y., et al., *Role of SUMO-interacting motif in Daxx SUMO modification, subnuclear localization, and repression of sumoylated transcription factors.* Mol Cell, 2006. **24**(3): p. 341-54.
10. Li, H., et al., *Sequestration and inhibition of Daxx-mediated transcriptional repression by PML.* Mol Cell Biol, 2000. **20**(5): p. 1784-96.
11. Lehembre, F., et al., *Regulation of Pax3 transcriptional activity by SUMO-1-modified PML.* Oncogene, 2001. **20**(1): p. 1-9.
12. Nefkens, I., et al., *Heat shock and Cd2+ exposure regulate PML and Daxx release from ND10 by independent mechanisms that modify the induction of heat-shock proteins 70 and 25 differently.* J Cell Sci, 2003. **116**(Pt 3): p. 513-24.
13. Hollenbach, A.D., et al., *Daxx and histone deacetylase II associate with chromatin through an interaction with core histones and the chromatin-associated protein Dek.* J Cell Sci, 2002. **115**(Pt 16): p. 3319-30.
14. Xue, Y., et al., *The ATRX syndrome protein forms a chromatin-remodeling complex with Daxx and localizes in promyelocytic leukemia nuclear bodies.* Proc Natl Acad Sci U S A, 2003. **100**(19): p. 10635-40.
15. Ishov, A.M., O.V. Vladimirova, and G.G. Maul, *Heterochromatin and ND10 are*

cell-cycle regulated and phosphorylation-dependent alternate nuclear sites of the transcription repressor Daxx and SWI/SNF protein ATRX. J Cell Sci, 2004. **117**(Pt 17): p. 3807-20.

16. Drane, P., et al., *The death-associated protein DAXX is a novel histone chaperone involved in the replication-independent deposition of H3.3.* Genes Dev, 2010. **24**(12): p. 1253-65.
17. Goldberg, A.D., et al., *Distinct factors control histone variant H3.3 localization at specific genomic regions.* Cell, 2010. **140**(5): p. 678-91.
18. Pluta, A.F., W.C. Earnshaw, and I.G. Goldberg, *Interphase-specific association of intrinsic centromere protein CENP-C with HDaxx, a death domain-binding protein implicated in Fas-mediated cell death.* J Cell Sci, 1998. **111 (Pt 14)**: p. 2029-41.
19. Lalioti, V.S., et al., *The insulin-sensitive glucose transporter, GLUT4, interacts physically with Daxx. Two proteins with capacity to bind Ubc9 and conjugated to SUMO1.* J Biol Chem, 2002. **277**(22): p. 19783-91.
20. Ko, Y.G., et al., *Apoptosis signal-regulating kinase 1 controls the proapoptotic function of death-associated protein (Daxx) in the cytoplasm.* J Biol Chem, 2001. **276**(42): p. 39103-6.
21. Charette, S.J., et al., *Inhibition of Daxx-mediated apoptosis by heat shock protein 27.* Mol Cell Biol, 2000. **20**(20): p. 7602-12.
22. Jung, Y.S., et al., *Physical interactions and functional coupling between Daxx and sodium hydrogen exchanger 1 in ischemic cell death.* J Biol Chem, 2008. **283**(2): p. 1018-25.
23. Junn, E., et al., *Interaction of DJ-1 with Daxx inhibits apoptosis signal-regulating kinase 1 activity and cell death.* Proc Natl Acad Sci U S A, 2005. **102**(27): p. 9691-6.
24. Puto, L.A. and J.C. Reed, *Daxx represses RelB target promoters via DNA methyltransferase recruitment and DNA hypermethylation.* Genes Dev, 2008. **22**(8): p. 998-1010.
25. Muromoto, R., et al., *Physical and functional interactions between Daxx and DNA methyltransferase 1-associated protein, DMAP1.* J Immunol, 2004. **172**(5): p. 2985-93.
26. Li, R., et al., *EAP1/Daxx interacts with ETS1 and represses transcriptional activation of ETS1 target genes.* Oncogene, 2000. **19**(6): p. 745-53.
27. Emelyanov, A.V., et al., *The interaction of Pax5 (BSAP) with Daxx can result in transcriptional activation in B cells.* J Biol Chem, 2002. **277**(13): p. 11156-64.
28. Kim, E.J., J.S. Park, and S.J. Um, *Identification of Daxx interacting with p73, one of the p53 family, and its regulation of p53 activity by competitive interaction*

with PML. Nucleic Acids Res, 2003. **31**(18): p. 5356-67.
29. Zhao, L.Y., et al., *Negative regulation of p53 functions by Daxx and the involvement of MDM2.* J Biol Chem, 2004. **279**(48): p. 50566-79.
30. Lin, D.Y., et al., *Promyelocytic leukemia protein (PML) functions as a glucocorticoid receptor co-activator by sequestering Daxx to the PML oncogenic domains (PODs) to enhance its transactivation potential.* J Biol Chem, 2003. **278**(18): p. 15958-65.
31. Lin, D.Y., et al., *Negative modulation of androgen receptor transcriptional activity by Daxx.* Mol Cell Biol, 2004. **24**(24): p. 10529-41.
32. Chang, C.C., et al., *Daxx mediates the small ubiquitin-like modifier-dependent transcriptional repression of Smad4.* J Biol Chem, 2005. **280**(11): p. 10164-73.
33. Muromoto, R., et al., *Physical and functional interactions between Daxx and STAT3.* Oncogene, 2006. **25**(14): p. 2131-6.
34. Croxton, R., et al., *Daxx represses expression of a subset of antiapoptotic genes regulated by nuclear factor-kappaB.* Cancer Res, 2006. **66**(18): p. 9026-35.
35. Park, J., et al., *Inhibition of NF-kappaB acetylation and its transcriptional activity by Daxx.* J Mol Biol, 2007. **368**(2): p. 388-97.
36. Huang, Y.S. and H.M. Shih, *Daxx positively modulates beta-catenin/TCF4-mediated transcriptional potential.* Biochem Biophys Res Commun, 2009. **386**(4): p. 762-8.
37. Shih, H.M., et al., *Daxx mediates SUMO-dependent transcriptional control and subnuclear compartmentalization.* Biochem Soc Trans, 2007. **35**(Pt 6): p. 1397-400.
38. Lindsay, C.R., V.M. Morozov, and A.M. Ishov, *PML NBs (ND10) and Daxx: from nuclear structure to protein function.* Front Biosci, 2008. **13**: p. 7132-42.
39. Tang, J., et al., *Critical role for Daxx in regulating Mdm2.* Nat Cell Biol, 2006. **8**(8): p. 855-62.
40. Michod, D., et al., *Calcium-dependent dephosphorylation of the histone chaperone DAXX regulates H3.3 loading and transcription upon neuronal activation.* Neuron, 2012. **74**(1): p. 122-35.
41. Lewis, P.W., et al., *Daxx is an H3.3-specific histone chaperone and cooperates with ATRX in replication-independent chromatin assembly at telomeres.* Proc Natl Acad Sci U S A, 2010. **107**(32): p. 14075-80.
42. Akterin, S., et al., *Involvement of glutaredoxin-1 and thioredoxin-1 in beta-amyloid toxicity and Alzheimer's disease.* Cell Death Differ, 2006. **13**(9): p. 1454-65.
43. Jiao, Y., et al., *DAXX/ATRX, MEN1, and mTOR pathway genes are frequently altered in pancreatic neuroendocrine tumors.* Science, 2011. **331**(6021): p.

1199-203.
44. Heaphy, C.M., et al., *Altered telomeres in tumors with ATRX and DAXX mutations.* Science, 2011. **333**(6041): p. 425.
45. Carninci, P., et al., *The transcriptional landscape of the mammalian genome.* Science, 2005. **309**(5740): p. 1559-63.
46. Wang, K.C. and H.Y. Chang, *Molecular mechanisms of long noncoding RNAs.* Mol Cell, 2011. **43**(6): p. 904-14.
47. Bertone, P., et al., *Global identification of human transcribed sequences with genome tiling arrays.* Science, 2004. **306**(5705): p. 2242-6.
48. Ponjavic, J., C.P. Ponting, and G. Lunter, *Functionality or transcriptional noise? Evidence for selection within long noncoding RNAs.* Genome Res, 2007. **17**(5): p. 556-65.
49. Struhl, K., *Transcriptional noise and the fidelity of initiation by RNA polymerase II.* Nat Struct Mol Biol, 2007. **14**(2): p. 103-5.
50. Guttman, M., et al., *Ab initio reconstruction of cell type-specific transcriptomes in mouse reveals the conserved multi-exonic structure of lincRNAs.* Nat Biotechnol, 2010. **28**(5): p. 503-10.
51. Guttman, M., et al., *Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals.* Nature, 2009. **458**(7235): p. 223-7.
52. Cabili, M.N., et al., *Integrative annotation of human large intergenic noncoding RNAs reveals global properties and specific subclasses.* Genes Dev, 2011. **25**(18): p. 1915-27.
53. Lee, J.T., *Lessons from X-chromosome inactivation: long ncRNA as guides and tethers to the epigenome.* Genes Dev, 2009. **23**(16): p. 1831-42.
54. Rinn, J.L., et al., *Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs.* Cell, 2007. **129**(7): p. 1311-23.
55. Wang, K.C., et al., *A long noncoding RNA maintains active chromatin to coordinate homeotic gene expression.* Nature, 2011. **472**(7341): p. 120-4.
56. Guttman, M., et al., *lincRNAs act in the circuitry controlling pluripotency and differentiation.* Nature, 2011. **477**(7364): p. 295-300.
57. Huarte, M., et al., *A large intergenic noncoding RNA induced by p53 mediates global gene repression in the p53 response.* Cell, 2010. **142**(3): p. 409-19.
58. Gupta, R.A., et al., *Long non-coding RNA HOTAIR reprograms chromatin state to promote cancer metastasis.* Nature, 2010. **464**(7291): p. 1071-6.
59. Loewer, S., et al., *Large intergenic non-coding RNA-RoR modulates reprogramming of human induced pluripotent stem cells.* Nat Genet, 2010. **42**(12): p. 1113-7.

60. Zhao, J., et al., *Polycomb proteins targeted by a short repeat RNA to the mouse X chromosome.* Science, 2008. **322**(5902): p. 750-6.
61. Nagano, T., et al., *The Air noncoding RNA epigenetically silences transcription by targeting G9a to chromatin.* Science, 2008. **322**(5908): p. 1717-20.
62. Tsai, M.C., et al., *Long noncoding RNA as modular scaffold of histone modification complexes.* Science, 2010. **329**(5992): p. 689-93.
63. Khalil, A.M., et al., *Many human large intergenic noncoding RNAs associate with chromatin-modifying complexes and affect gene expression.* Proc Natl Acad Sci U S A, 2009. **106**(28): p. 11667-72.
64. Zhao, J., et al., *Genome-wide identification of polycomb-associated RNAs by RIP-seq.* Mol Cell, 2010. **40**(6): p. 939-53.
65. Kim, T.K., et al., *Widespread transcription at neuronal activity-regulated enhancers.* Nature, 2010. **465**(7295): p. 182-7.
66. Wang, D., et al., *Reprogramming transcription by distinct classes of enhancers functionally defined by eRNA.* Nature, 2011. **474**(7351): p. 390-4.
67. Orom, U.A., et al., *Long noncoding RNAs with enhancer-like function in human cells.* Cell, 2010. **143**(1): p. 46-58.
68. Mao, Y.S., B. Zhang, and D.L. Spector, *Biogenesis and function of nuclear bodies.* Trends Genet, 2011. **27**(8): p. 295-306.
69. Clemson, C.M., et al., *An architectural role for a nuclear noncoding RNA: NEAT1 RNA is essential for the structure of paraspeckles.* Mol Cell, 2009. **33**(6): p. 717-26.
70. Chen, L.L. and G.G. Carmichael, *Altered nuclear retention of mRNAs containing inverted repeats in human embryonic stem cells: functional role of a nuclear noncoding RNA.* Mol Cell, 2009. **35**(4): p. 467-78.
71. Tripathi, V., et al., *The nuclear-retained noncoding RNA MALAT1 regulates alternative splicing by modulating SR splicing factor phosphorylation.* Mol Cell, 2010. **39**(6): p. 925-38.
72. Ishii, N., et al., *Identification of a novel non-coding RNA, MIAT, that confers risk of myocardial infarction.* J Hum Genet, 2006. **51**(12): p. 1087-99.
73. Mendell, J.T. and E.N. Olson, *MicroRNAs in stress signaling and human disease.* Cell, 2012. **148**(6): p. 1172-87.
74. Geng, Y.J., et al., *Large intervening non-coding RNA HOTAIR is associated with hepatocellular carcinoma progression.* J Int Med Res, 2011. **39**(6): p. 2119-28.
75. Kogo, R., et al., *Long noncoding RNA HOTAIR regulates polycomb-dependent chromatin modification and is associated with poor prognosis in colorectal cancers.* Cancer Res, 2011. **71**(20): p. 6320-6.
76. Gibb, E.A., et al., *Human cancer long non-coding RNA transcriptomes.* PLoS One,

2011. **6**(10): p. e25915.
77. Chang, C.C., et al., *Structural and Functional Roles of Daxx SIM Phosphorylation in SUMO Paralog-Selective Binding and Apoptosis Modulation.* Mol Cell, 2011. **42**(1): p. 62-74.
78. Yap, K.L., et al., *Molecular interplay of the noncoding RNA ANRIL and methylated histone H3 lysine 27 by polycomb CBX7 in transcriptional silencing of INK4a.* Mol Cell, 2010. **38**(5): p. 662-74.
79. Schwartzentruber, J., et al., *Driver mutations in histone H3.3 and chromatin remodelling genes in paediatric glioblastoma.* Nature, 2012. **482**(7384): p. 226-31.
80. Hermeking, H., *The miR-34 family in cancer and apoptosis.* Cell Death Differ, 2010. **17**(2): p. 193-9.
81. Ip, J.Y. and S. Nakagawa, *Long non-coding RNAs in nuclear bodies.* Dev Growth Differ, 2011.
82. Mikkelsen, T.S., et al., *Genome-wide maps of chromatin state in pluripotent and lineage-committed cells.* Nature, 2007. **448**(7153): p. 553-60.
83. Garcia, M.J., et al., *A 1 Mb minimal amplicon at 8p11-12 in breast cancer identifies new candidate oncogenes.* Oncogene, 2005. **24**(33): p. 5235-45.
84. Yang, Z.Q., et al., *Multiple interacting oncogenes on the 8p11-p12 amplicon in human breast cancer.* Cancer Res, 2006. **66**(24): p. 11632-43.
85. Yang, Z.Q., et al., *Transforming properties of 8p11-12 amplified genes in human breast cancer.* Cancer Res, 2010. **70**(21): p. 8487-97.
86. Holland, D.G., et al., *ZNF703 is a common Luminal B breast cancer oncogene that differentially regulates luminal and basal progenitors in human mammary epithelium.* EMBO Mol Med, 2011. **3**(3): p. 167-80.